



International Conference on Transportation and Development 2024

Transportation Safety and Emerging Technologies

Selected Papers from the International Conference
on Transportation and Development 2024

Atlanta, Georgia
June 15–18, 2024



EDITED BY
Heng Wei, Ph.D., P.E.



TRANSPORTATION
& DEVELOPMENT
INSTITUTE
ASCE

Downloaded from ascelibrary.org by Texas State University on 06/14/24. Copyright ASCE. For personal use only; all rights reserved.

INTERNATIONAL CONFERENCE ON TRANSPORTATION AND DEVELOPMENT 2024

TRANSPORTATION SAFETY AND EMERGING TECHNOLOGIES

SELECTED PAPERS FROM THE INTERNATIONAL CONFERENCE
ON TRANSPORTATION AND DEVELOPMENT 2024

June 15–18, 2024
Atlanta, Georgia

SPONSORED BY
Georgia Department of Transportation
The Transportation & Development Institute
of the American Society of Civil Engineers

EDITED BY
Heng Wei, Ph.D., P.E.



TRANSPORTATION
& DEVELOPMENT
INSTITUTE

Published by the American Society of Civil Engineers

Published by American Society of Civil Engineers
1801 Alexander Bell Drive
Reston, Virginia, 20191-4382
www.asce.org/publications | ascelibrary.org

Any statements expressed in these materials are those of the individual authors and do not necessarily represent the views of ASCE, which takes no responsibility for any statement made herein. No reference made in this publication to any specific method, product, process, or service constitutes or implies an endorsement, recommendation, or warranty thereof by ASCE. The materials are for general information only and do not represent a standard of ASCE, nor are they intended as a reference in purchase specifications, contracts, regulations, statutes, or any other legal document. ASCE makes no representation or warranty of any kind, whether express or implied, concerning the accuracy, completeness, suitability, or utility of any information, apparatus, product, or process discussed in this publication, and assumes no liability therefor. The information contained in these materials should not be used without first securing competent advice with respect to its suitability for any general or specific application. Anyone utilizing such information assumes all liability arising from such use, including but not limited to infringement of any patent or patents.

ASCE and American Society of Civil Engineers—Registered in U.S. Patent and Trademark Office.

Photocopies and permissions. Permission to photocopy or reproduce material from ASCE publications can be requested by sending an e-mail to permissions@asce.org or by locating a title in ASCE's Civil Engineering Database (<http://cedb.asce.org>) or ASCE Library (<http://ascelibrary.org>) and using the "Permissions" link.

Errata: Errata, if any, can be found at <https://doi.org/10.1061/9780784485514>

Copyright © 2024 by the American Society of Civil Engineers.
All Rights Reserved.
ISBN 978-0-7844-8551-4 (PDF)
Manufactured in the United States of America.

Preface

It is my great pleasure to present to you the proceedings from the ASCE International Conference on Transportation and Development (ICTD 2024), organized by the Transportation and Development Institute (T&DI) of ASCE. ICTD is ASCE's flagship conference in transportation and development. A special feature of this event was the partnership with Georgia Department of Transportation (GDOT) that helped deliver sessions showcasing their projects as part of the ICTD program. In addition, the 19th International Conference on Automated People Movers and Automated Transit Systems (APM-ATS 2024) was co-located with ICTD 2024 to cater to the technical content needs of the APM-ATS community. The event was held from June 15 to 18, 2024, at the Hilton Atlanta in the heart of the beautiful downtown Atlanta – the capital city of the state of Georgia.

ASCE ICTD 2024's three days of technical program featured three super-plenary sessions, one plenary session, and 42 technical breakout sessions that included leaders from ASCE, government agencies, private industry, and the academic sector, covering key areas of the entire spectrum of transportation and development. Technical sessions were organized by Georgia DOT, and ASCE-T&DI's CAV Impacts Committee, Uncrewed Aerial Systems Committee, MODaaS Committee, Data Sensing & Analytics Committee, AI in Transportation Committee, Transportation Safety Committee, Street and Highway Operations Committee, Active Transportation Committee, Highway Pavements Committee, Highway Construction Committee, Sustainable Transportation Committee, Freight and Logistics Committee, Infrastructure Systems Committee, Economics and Finance Committee, Aviation Planning & Operations Committee, Public Transport Committee, and Rail Transport Committee.

Technical posters divided into 25 different topics areas were on display throughout the conference. The poster program concluded with a dedicated poster reception, providing attendees with additional content to learn about cutting edge research and practice, and have the opportunity to speak directly with the authors. The conference program also included a variety of special events including offsite socials for students and younger members, a welcome reception, a seated awards luncheon, and several networking events in the exhibit hall.

The conference was preceded by two (2) workshops on the following topics as preconference events:

- AI and Emerging Technologies for Integrated Transportation Cybersecurity
- Government Relations and Public Relations University

In addition, the conference featured seven (7) tours on June 18, 2024, including the following:

- Hartsfield-Jackson Atlanta International Airport (ATL) - Behind the scenes
- Automated People Mover (APM) at Hartsfield-Jackson Atlanta International Airport (ATL)
- MARTA Control Center: Behind the Scenes
- Delta Technical Operations

- GDOT Traffic Management Center and Highway Emergency Response Operators (TMC/HEROs)
- Curiosity Labs
- Atlanta Beltline

The co-located APM-ATS 2024 Conference shared three super-plenaries and Awards Banquet with ICTD 2024 and included two separate plenary sessions and 18 technical breakout sessions. Technical sessions covered topics ranging from APM application at airports to APM engineering, design, operations, and maintenance, and from ATS planning to Personal Rapid Transit and Urban Systems and Major Activity Centers. In addition, the conference shared the welcome reception and other social and networking events with ICTD 2024. Last but not least, APM-ATS preceded with an Airport Operations and Maintenance Managers' Workshop and concluded with a technical tour of the Hartsfield-Jackson Atlanta International Airport Automated People Mover Systems.

ASCE ICTD 2024 has followed the great success of past ICTDs and attracted significant interest indicated by the rich technical program. A large number of papers were accepted for publication in the proceedings. Each paper went through a rigorous peer review by technical experts before becoming a publication of ASCE – the world's largest publisher of Civil Engineering content.

The proceedings for this conference have been organized into three (3) volumes based on the topical distribution as follows:

Volume I: Transportation Safety and Emerging Technologies

- Transportation Safety
- AI in Transportation
- CAV Impacts
- Data Sensing and Analytics
- Intelligent Transportation Systems
- Uncrewed Aerial Systems

Volume II: Transportation Planning, Operations, and Transit

- Active Transportation
- Freight & Logistics
- Social Equity, Justice, and Welfare
- Sustainable Transportation & Urban Development
- Transportation and Public Health
- Transportation Economics & Finance
- Workforce Development, Diversity and Inclusion
- APM-ATS
- Public Transport
- Rail Transport
- Street & Highway Operations

- Volume III: Pavements and Infrastructure Systems
 - Airport Pavements

- Highway Construction
- Highway Pavements
- Infrastructure Systems

All these accomplishments were due to the incredible efforts of ASCE ICTD 2024 Conference Co-Chairs, Mr. Russell R. McMurry, and Ms. Marsha Anderson Bomar, the Conference Steering Committee, and the terrific support from ASCE staff. I would also like to express my sincere gratitude to all the authors and conference participants for their significant contributions. I am also grateful to all paper reviewers for their outstanding volunteer efforts. I would extend my special thanks to the T&DI technical committee volunteers, conference sponsors, exhibitors, moderators, and speakers for their help in making ASCE ICTD 2024 a great success!

Our unique integration of private, government, and academic leaders makes ASCE ICTD event series an excellent platform for information exchange, experience sharing, and professional networking. I hope you found ASCE ICTD 2024 to be a wonderful and rewarding experience, and I sincerely hope that ASCE ICTD becomes a ‘can’t miss’ conference in each year that follows. On behalf of the conference leadership and committee, and ASCE T&DI, I wish you all the best in your professional endeavors and hope to see you at the ICTD 2025 and the co-located Pavements 2025, scheduled to be held at the Renaissance Resort in Glendale, Arizona from June 8 to 11, 2025.

ASCE ICTD 2024 Proceedings Editor



A handwritten signature in black ink, appearing to read "Heng Wei".

Heng Wei, Ph.D., P.E., F.ASCE
The University of Cincinnati

Acknowledgments

ICTD Steering Committee

- Marsha Anderson Bomar, AICP, F. ASCE, ENV SP, F. ITE, Commissioner Emeritus, Atlanta DOT (Co-Chair)
- Russell R. McMurry, P.E., Commissioner, Georgia DOT (Co-Chair)
- Tom Klin, GHD (Program Chair)
- Robert Bryson, P.E., M.ASCE, City of Milwaukee, Retired (Roadways Technical Chair & Poster Chair)
- Walt Kulyk, P.E., M.ASCE, Federal Transit Administration, Retired (Rail & Transit Technical Chair)
- Samuel Labi, Ph.D., M.ASCE, Purdue University (Planning & Development Technical Chair)
- Amiy Varma, Ph.D., P.E., M.ASCE, AAAJ, LL (Aviation Technical Chair)
- Cole Kopca, S.M.ASCE, PacTrans, University of Washington (Younger Member Activities)
- Hampton Worthey, EIT, Michael Baker International (Local Area Chair)
- Heng Wei, Ph.D., P.E., F.ASCE, University of Cincinnati (Proceedings Editor)

APM-ATS Steering Committee

- Rich Czarnomski, A.M.ASCE, Lea+Elliott (Co-Chair)
- Matt Lesh, A.M.ASCE, Spectrum Mobility (Co-Chair)
- William J. Sproule, Ph.D., P.E., F.ASCE, Michigan Technological University (Program Chair)
- Dr. Murthy V. A. Bondada Ph.D., P.E., F.ASCE, F.ITE, Atlantic Transportation Engineers, LLC

ASCE Staff

Muhammad Amer, ENV SP, M.ASCE, Managing Director, Transportation & Development Institute (T&DI) and Future World Vision

Sara Jessica Palmer, CMP, Manager, Transportation & Development Institute (T&DI) Conferences

Melissa Willis, Senior Marketing Coordinator, Transportation & Development Institute (T&DI)

Corinne Addison, Manager, Book Production, ASCE

Tom Sours, Senior Manager, Event and Media Sales, ASCE

Contents

Transportation Safety

Assessing the Impact of Training Programs for Highway Work Zone Employees	1
Apurva Pamidimukkala, Sai Sneha Channamallu, and Sharareh Kermanshachi	
A Review on the Mental Health Stressors of Construction Workers	10
Deema Almaskati, Sharareh Kermanshachi, and Apurva Pamidimukkala	
Assessing Driver Behavior during Bicyclist Overtaking: An Integrated Simulator Study	22
Parisa Masoumi, Anam Ardeshtiri, Eazaz Sadeghvaziri, and Mansoureh Jeihani	
Assessment of Pedestrian and Bicycle Safety at Commercial Driveways along Major Corridors.....	34
Cong Chen, Elzbieta Bialkowska-Jelinska, Kristine Williams, Pei-Sung Lin, Shubhankar Chintamani Shindgikar, Tia Boyd, and Gina Bonyani	
Unraveling Crash Causation: A Deep Dive into Non-Motorists on Personal Conveyance	47
Subasish Das, Rohit Chakraborty, and Mahmuda Sultana Mimi	
Understanding How We Can Make Commercial Driveways Safer: Exploratory Analysis through Case Studies.....	59
Tia Boyd, Kristine Williams, Pei-Sung Lin, Cong Chen, Elzbieta Bialkowska-Jelinska, and Gina Bonyani	
Effect of Crash Barriers on Driver Behavior	70
Hasara Senevirathna, Nipuna Seneviratne, Vasantha Wickramasinghe, and Sunanda Dissanayake	
Assessment of Stopping Sight Distance Restrictions on Vertical Alignments: A Case Study	82
Samantha Islam and Benjamin Burton	
COVID-19 Pandemic and Distraction-Related Motorcycle Crashes in Kentucky	92
Bharat Kumar Pathivada, Noelle Buhay, Dylan Justice, Arunabha Banerjee, and Kirolos Haleem	
Bridge Health-Informed Route Planning: Challenges and Promises	104
Yonas Kassa, Will Heller, Brandon Lacy, Brian Ricks, and Robin Gandhi	
Time Series Analysis of Highway Work Zone Crashes in North Carolina	117
Kazi Tahsin Huda, Yuting Chen, Don Chen, and Srinivas S. Pulugurtha	

Project-Based Crash Analysis for Crash Risk Reduction during Pavement Preservation.....	128
Sirisha Gangireddy, Ruijie Bian, Tara Tolford, and Hany Hassan	
Using the Estimated Available Friction at the Posted Speed as a Pavement Safety Performance Measure on Freeways	141
Behrokh Bazmara, Gerardo W. Flintsch, Edgar de León Izeppi, and Samer W. Katicha	
Exploring Statistical Methods in Developing Safety Performance Functions	152
Xingjing Xu and Ilir Bejleri	
Motorcycle Crashes in California: Analysis of Crash Severity and Contributing Factors	166
Amal Mamlouk, Ashkan Teymour, and Masoud Ghodrat Abadi	
Assessing the Impact of Cellular Coverage Areas on Distracted Driving, Crashes, and Injuries.....	178
S. Hernandez and Md. M. Rahman	
Estimating Driver Behaviour at Unsignalized Crosswalks.....	192
H. M. N. S. Kumari, I. P. W. N. Thathsara, W. M. V. S. K. Wickramasinghe, and Sunanda Dissanayake	
Comprehensive Analysis of Red Light Running in the United States: Strategies, Successes, and Future Directions.....	204
Saquib M. Haroon, Cipriano De Luna Gutierrez, Sushmita Bhandari, Alyssa Ryan, and Henrick Haule	
Assessing the Influence of Various No-Right-Turn-on-Red Signs: A Driving Simulator Study	215
Parisa Masoumi, Abolfazl Taherpour, Mansoureh Jeihani, Anam Ardeshiri, and Samira Ahangari	
Safety Benefits of Parcel Delivery Modes Using Geographically Weighted Negative Binomial Regression	226
Zhenyu Wang, Pei-Sung Lin, and Yaye Mallon Keita	
Optimizing Speed Control Guidance at Urban Signalized Intersections: A Driving Simulator Study on Driver Behavior and Sociodemographic Factors	238
Parisa Masoumi, Eazaz Sadeghvaziri, and Mansoureh Jeihani	
Exploring the Role of Human Behavior in Road Crash Occurrence through Analysis of Crash Variability in Highly Similar Road Environments.....	249
Nitesh Acharya and Michael Henry	

Investigating the Influence of Weather on Weigh-in-Motion Measurements Using In-Pavement Strain Sensors	261
Xinyi Yang, Xingyu Wang, Ying Huang, Pan Lu, Yihao Ren, and Joseph Podolsky	
Navigating Road Safety and Equity: A GIS Analysis of Crash Data in Atlanta, GA.....	271
Eazaz Sadeghvaziri, Ramina Javid, and Lila Turbiville	
Comparison of the Effect of In-Crosswalk Traffic Signs on Pedestrian Safety	281
Mulugeta D. Amare and Daba S. Gedafa	
Noteworthy Practices for Cross-Agency Traffic Safety Coordination	294
Taylor Dinehart, Tia Boyd, Jeff Kramer, Lama Alfaseeh, and Sarah Caper	
Predicting Bicycle-Involved Crashes in the SCAG Region: A Machine Learning Analysis Using HSIS Data from California State.....	305
Ramina Javid, Eazaz Sadeghvaziri, Seyedmirsajad Mokhtarimousavi, and Hananeh Omidi	
Evaluating Factors of Interstate Crashes in New Jersey	315
Fahmida Rahman	
Evaluation of Sinusoidal Rumble Strips on Noise Mitigation and Bicycle Accommodation	324
Pei-Sung Lin, Cong Chen, Yaye Keita, Shubhankar Shindgikar, and Ela Białkowska-Jelińska	
Effects of Law Enforcement Presence, Frequency, and Duration on Speeding.....	336
Pei-Sung Lin, Yaye Keita, Shubhankar Shindgikar, Emmeth Duran, Ela Białkowska-Jelińska, and Rakesh Rangaswamy	
Causal Insights into Speeding Crashes	348
Subasish Das, Ahmed Hossain, Swastika Barua, Shahrbanoo Kavianpour, and Abbas Sheykhfard	
Extended Runtime at Signalized Intersection Using Backup Power after the Loss of Utility Power.....	360
Pei-Sung Lin, Achilleas Kourtellis, and Zhenyu Wang	
Before-After Analysis of Highway Crashes on the Dhaka-Chattogram National Highway of Bangladesh	372
Afsara Tasquia, Shadman Sakib, Md. Moshiur Rahman, Javed Bari, Kazi Md. Shifun Newaz, and Faria Tabassum	

AI in Transportation

Machine Learning Based Multi-Modal Transportation Network Planner	380
Neeraj Menon Manghat, Vaishak Gopalakrishna, Sai Bonthu, Victor Hunt, Arthur Helmicki, and Doug McClintock	
Privacy-Preserved Federated Reinforcement Learning for Autonomy in Signalized Intersections	390
Negar Asadi, Seyed Hamid Hosseini, Mahdi Imani, Daniel P. Aldrich, and Seyedeh Fatemeh Ghoreishi	
Field Study on Unconfined Compressive Strength and Drilling Data of DSM Columns: A Machine Learning Approach	404
Seyed Meisam Alavi, Sajjad Shakeri Talarposhti, Ahmad Ali Khodaei Ardabili, and Milad Aghamolaei	
Fast Machine Learning-Based High Fidelity Mesoscopic Modeling Tool for Traffic Simulation.....	417
Neeta A. Eapen, Robert B. Heckendorf, and Ahmed Abdel-Rahim	
Improved Prediction of Soil Thermal Properties Using Recurrent Neural Networks	431
Scott Michael Slone, Zachary Zody, Robert Ibey, and Wade A. Lein	
Exploring the Determinants of Pedestrian Crash Severity Using an AutoML Approach.....	442
Amir Rafe and Patrick A. Singleton	
Automated Traffic Safety Assessment Tool Utilizing Monocular 3D Convolutional Neural Network Based Detection Algorithm at Signalized Intersections	456
Ahmed Mohamed, Lizhe Li, and Mohamed M. Ahmed	
A Deep Learning Approach for Multiple Weather Conditions Detection Based on In-Vehicle Dash Camera Video.....	468
Lizhe Li and Mohamed M. Ahmed	
Vehicle-to-Vehicle Conflicts Detection at Signalized Intersections: Leveraging Multi-Cameras Utilizing Convolutional Neural Network-Based Pose Estimation Algorithm.....	479
Ahmed Mohamed, Lizhe Li, and Mohamed M. Ahmed	
Pedestrian Tracking at Signalized Intersections Leveraging Multi-Camera Field of Views Using Covolutional Neural Network-Based Pose Estimation Algorithm	490
Ahmed Mohamed and Mohamed M. Ahmed	

Clarifying Rear-End Crash Propensity with Spatial Spillover Effect: Artificial Bayes-GLM Method.....	502
Wei Lin, Heng Wei, Zhixia Li, and Dong Nian	
Highway Abandoned Objects Recognition Based on Open Vocabulary Object Detection Approach.....	515
Siwen Liu, Yifan Zhuang, Nanfang Zheng, Yufei Ji, Pei Liu, and Ziyuan Pu	
Is ChatGPT a Reliable Source of Transportation Equity Information for Scientific Writing?	525
Boniphace Kutela, Shoujia Li, Subasish Das, and Jinli Liu	
 <i>CAV Impacts</i>	
Enhancing AV Safety: A Bagging Classifier Approach for Predicting Crash Outcomes	538
Sai Sneha Channamallu, Deema Almaskati, Sharareh Kermanshachi, and Apurva Pamidimukkala	
Exploring Riders' Preferences of Using Shared Autonomous Vehicles	550
Apurva Pamidimukkala, Sharareh Kermanshachi, Jay Michael Rosenberger, and Ann Foss	
Automated Vehicles vs. Human Drivers: Modeling Driving Behavior Using Data from Field Experiments.....	560
Pei Li, Steven T. Parker, and David A. Noyce	
Evaluating Autonomous Vehicles' Impact on Emergency Evacuation Clearance Time: A Comparative Study.....	573
Asad Ali, Mingwei Guo, Ying Huang, Talha Ahmed, and Pan Lu	
Optimizing Lane Configuration for Efficient Platoon Control during Lane Closures	583
Vamsi Krishna Maddineni and Mohamed M. Ahmed	
Safety and Operational Impacts of Varying Market Penetration and Level of Autonomy of Connected and Automated Vehicles on I-80 Freeway Segment	592
Mandip Sigdel and Mohamed M. Ahmed	
Safety Assessment of Cooperative Adaptive Cruise Control Truck Platooning on Freeway Work Zones	605
Vamsi Krishna Maddineni and Mohamed M. Ahmed	
Dynamically Expanding Capacity of Autonomous Driving with Near-Miss Focused Training Framework	616
Ziyuan Yang, Zhaoyang Li, Jianming Hu, and Yi Zhang	

Delivering Tomorrow: Analyzing Automated Delivery Vehicle Narratives through Media Mining	627
Subasish Das and Boniphace Kutela	
Impact of Automated Driving under Virtual Traffic Signals on Traffic Resilience of Urban Road Networks	640
Yankai Zhang, Jiajing Chen, Bo Liang, Zihuan Xu, Lei Mo, Zhao Zhang, and Ruijie ‘Rebecca’ Bian	
Assessing Cybersecurity Risks and Traffic Impact in Connected Autonomous Vehicles.....	652
Saurav Silwal, Lu Gao, Yunpeng Zhang, Ahmed Senouci, and Yi-Lung Mo	
Prediction of Daily Disengagements of Automated Vehicles Using Explainable Machine Learning Approach	663
Boniphace Kutela, Sunday Okafor, Norris Novat, Tumlumbe Juliana Chengula, and John Kodi	
<i>Data Sensing and Analytics</i>	
Machine Learning Strategies for Optimizing Urban Parking: A Comparative Evaluation.....	678
Sai Sneha Channamallu, Sharareh Kermanshachi, Jay Michael Rosenberger, and Apurva Pamidimukkala	
Ego-Centric Pedestrian Trajectory Prediction Considering Camera Motion Parameters	690
Yufei Ji, Pei Liu, Nanfang Zheng, Siwen Liu, and Ziyuan Pu	
Prototype Data Dashboard for Multi-Source Transportation and Community Health Data Analytics	699
Chengyue Wang, Talha Azfar, and Ruimin Ke	
Leveraging Computer Vision for Transportation Infrastructure Health Monitoring	710
Claudia Marin-Artieda, Ogheneruona Akpareva, and Stephen Arhin	
Harnessing Synthetic Image Datasets for Enhanced Scene Understanding in Construction Work Zones	722
Sultan Al Shafian and Da Hu	
Deep Learning Enhanced Crack Detection for Tunnel Inspection	732
Sultan Al Shafian, Da Hu, and Wenbing Yu	
Performance Evaluation of Bridge Approach Slabs and Joints through Sub-mm 3D Laser Imaging System.....	742
Guolong Wang, Kelvin C. P. Wang, Guangwei Yang, and Joshua Qiang	

Intelligent Transportation Systems

A Review of Factors Affecting Severity of Autonomous Vehicle Crashes.....	755
Deema Almaskati, Sharareh Kermanshachi, and Apurva Pamidimukkala	
Autonomous Vehicle Safety: A Comprehensive Analysis of Crash Injury Determinants	767
Sai Sneha Channamallu, Deema Almaskati, Sharareh Kermanshachi, and Apurva Pamidimukkala	
Computer Vision Based Detector Integrated with V2X Framework for Vulnerable Road User Safety.....	780
Sai Bonthu, Vaishak Gopalakrishna, William Martin, Nick Hegemier, Victor Hunt, and Arthur Helmicki	
Examining Drivers of Electric Vehicle Adoption.....	788
Apurva Pamidimukkala, Sharareh Kermanshachi, Jay Michael Rosenberger, and Greg Hladik	
Real-Time Traffic Control and Safety Measure Analysis Using LiDAR Sensor during Traffic Signal Failures	798
A. Ansariyar and M. Jeihani	
A Comparative Analysis of Probe Data Sources for Provision of Work Zone Congestion Information to the Public	811
Dorcus Okaidjah, Skylar Knickerbocker, and Christoher Day	
Investigating the Role of Dimension Reduction in Counteracting Machine Learning Performance Bias in TSMO Strategies	821
Mustafa Attallah and Jalil Kianfar	
Hierarchical Classification—Regression (HiClassR) to Improve Incident Clearance Time Prediction.....	833
Mustafa Attallah and Jalil Kianfar	
An Exploratory Analysis of Factors Affecting Adoption of Electric Vehicles	845
Apurva Pamidimukkala, Sharareh Kermanshachi, Jay Michael Rosenberger, and Greg Hladik	
Smartwatches in Transportation: Unleashing Innovations and Advancements—A Comprehensive Systematic Review	855
Eazaz Sadeghvaziri, Ramina Javid, and Mohammadsoroush Tafazzoli	

Uncrewed Aerial Systems

Use of Synthetic Aperture Radar and Uncrewed Aerial Vehicle for Assessing Pavement Condition	866
Amit Gajurel, Nripojyoti Biswas, Hiramani Chimauriya, and Anand J. Puppala	
Using AI and Change Detection in Geospatial UAS Airfield Pavement Inspection for Pavement Management	876
Michael T. McNerney, Grant Bishop, and Valerie Saur	
Integration of Small Unmanned Aircraft Systems and Deep Learning for Efficient Airfield Pavement Crack Detection and Assessment.....	884
Md. Abdullah All Sourav, Halil Ceylan, Sunghwan Kim, and Matthew Brynick	

Assessing the Impact of Training Programs for Highway Work Zone Employees

Apurva Pamidimukkala, Ph.D.¹; Sai Sneha Channamallu²;
and Sharareh Kermanshachi, Ph.D., P.E., F.ASCE³

¹Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

²Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: sxc5928@mavs.uta.edu

³Associate Vice Chancellor and Associate Dean of Research, Penn State Univ., Middletown, PA (corresponding author). Email: svk5464@psu.edu

ABSTRACT

Work zone safety, especially amid increasing highway construction and repair, is paramount in ensuring the well-being of construction workers. Between 2003 and 2019, road construction sites saw a total of 2,103 worker fatalities, averaging 131 deaths annually. During this period, Texas had the highest number of road construction-related worker deaths. With traffic congestion intensifying and a rise in on-site hazards, there is a pressing need for workers to be adequately educated on the multitude of risks they face, from physical injuries to exposure to toxic substances. However, the efficacy of current training programs in inculcating proper safety practices remains underexplored. This study seeks to develop comprehensive training materials highlighting the myriad of hazards in work zones and the preventive measures to be adopted. An experimental highway construction worker group in Texas underwent this tailored training. Their knowledge was gauged before and after the session through a standardized questionnaire comprising 10 questions aligned with the training content. The data was rigorously analyzed, unveiling a marked elevation in the participants' safety awareness post-training. These findings can guide construction firms and regulatory bodies in instituting training measures that heighten the safety quotient in work zones.

Keywords: Hazards, occupational safety, training, employees, work zone

INTRODUCTION

The heavy construction industry in the United States, particularly highway and street construction, is a significant economic contributor, valued at about \$237 billion in 2017 and employing approximately 1 million individuals (US Department of Commerce, 2019). Despite representing a mere 0.8% of the working population, it accounted for 3% of the fatalities within the industry (BLS, 2018; US Census Bureau, 2019). The Federal Highway Administration reports that there is a work zone injury every 13 minutes and a fatality every 11 hours, with 2,103 workers losing their lives at road construction sites from 2003 to 2019, averaging 124 fatalities per year (BLS, 2022). Fatal accidents are not confined to unfavorable weather or road conditions. In 2008, it was found that 29% of fatal crashes took place on dry roads, while 35% happened on wet roads (Traffic Safety Facts, 2008). These figures highlight the persistent risks in road construction, which pose dangers not only to workers but also to motorists navigating through these zones (CDC, 2022; Nipa & Kermanshachi, 2022).

The intricacy and variety of dangers present in roadwork zones have notably risen because of the growing number of highway projects and the swift emergence of large-scale highway endeavors globally (Li et al., 2018). A work zone, as defined by the Federal Highway Administration, is an area of trafficway with construction, maintenance, or utility work activities. These zones can vary in duration and type of activity, ranging from long-term stationary construction to short-term mobile maintenance tasks (FHWA, 2018; Kermanshachi et al., 2020). The review of existing research indicates that the existence of a work zone is probable to lead to a higher rate of accidents (Yang et al., 2015). The hazards in these zones are multifaceted, including the risk of being struck by vehicles or construction equipment, falls, electrocutions, and caught-in-between accidents, collectively known as the "Fatal Fours" in construction safety which accounted for 58.6% of the construction deaths in 2018 (OSHA Education Center, 2023; Subramanya et al., 2020).

Falls consistently rank as one of the primary causes of fatalities and severe injuries in the construction industry. To combat this, the implementation of effective fall protection systems is paramount. These systems, which include both passive (such as guardrails and safety nets) and active (like harnesses and lifelines) measures, play a crucial role in ensuring worker safety. It is imperative that these systems are not only chosen wisely but also installed correctly. Moreover, workers must receive thorough training to understand and utilize these systems effectively. This comprehensive approach to fall safety is essential to prevent accidents and injuries (Highway worker safety program, 2022; Safapour et al., 2021). Electrocution hazards represent another significant concern in construction settings. These dangers can emerge from either direct contact with live electrical components or indirect contact through tools or other conductive materials. To mitigate these risks, it is crucial to provide robust safety training focused on electrical hazards. Additionally, the deployment of appropriate personal protective equipment (PPE), such as insulated gloves and dielectric footwear, is vital to safeguard workers against electrical injuries (Long et al., 2014; Rad & Kermanshachi, 2018). Struck-by-object incidents are also a major contributor to construction-related deaths. Heavy equipment like trucks, cranes, and excavators often play a role in these accidents. To protect workers from these hazards, it is essential to enforce the use of protective gear, including hard hats, safety glasses, and high-visibility clothing. Such PPE helps in minimizing injuries from flying or falling objects and increases worker visibility around heavy machinery (Rashad, 2022; Pamidimukkala & Kermanshachi, 2023).

Caught-in-between hazards, often involving cranes, heavy equipment, and masonry walls, pose significant risks to construction workers. These dangers necessitate strict compliance with safety protocols and the implementation of effective engineering controls. Safety measures may include proper training, ensuring machinery is adequately guarded, and maintaining a safe distance from moving equipment to prevent these types of accidents (Namian et al., 2020; Safapour et al., 2023). Additionally, working conditions during nighttime and in low-visibility environments present distinct challenges in highway construction. These conditions require heightened safety measures, such as enhanced lighting, reflective clothing, and increased awareness of surrounding hazards. This is particularly important for workers operating at night, who must be highly visible and vigilant to avoid accidents (Pamidimukkala & Kermanshachi, 2022). Lastly, overexertion, though often underestimated, is a significant risk factor in the construction industry. It can lead to a multitude of nonfatal and fatal accidents. Factors contributing to overexertion include heavy lifting, repetitive motions, and prolonged periods of strenuous activity. Addressing this issue requires a focus on ergonomic practices, adequate rest

breaks, and training in proper lifting techniques to reduce the occurrence of strain-related injuries (Rathnasiri et al., 2023; Kermanshachi et al., 2018).

Safeguarding against hazards in the workplace typically follows a tiered approach, beginning with the most effective strategies and concluding with the least (Ammar & Dadi, 2023). At the top of this hierarchy are engineering controls, which either eliminate the hazard entirely or create a physical barrier between the worker and the hazard. Following this are administrative controls, which entail alterations in work practices aimed at minimizing the time, frequency, and intensity of exposure to occupational hazards (Highway worker safety program, 2022; Kermanshachi & Safapour, 2019). The last line of defense in this hierarchy is the utilization of PPE, which serves as a crucial measure in hazard control. It is imperative that workers receive thorough training to understand and utilize these effectively (Huebschman et al., 2003; Subramanya et al., 2022).

The Occupational Safety and Health Act (OSHA) of 1971 has been pivotal in shaping safety protocols in high-risk zones. Prior to 1970, there were no comprehensive and uniform provisions governing the protection of workers against occupational safety and hazards. The number of work-related accidents was alarmingly high, with as many as 14,000 deaths and 2 million injuries annually (OSHA Education Center, 2023; Nipa et al., 2022). The introduction of OSHA marked a significant turning point, mandating employers to maintain a hazard-free workplace and ensure their workforce is trained in hazard recognition and control (Safapour et al., 2017). Despite these regulations, construction sites, particularly work zones on highways and streets, continue to confront various safety challenges, highlighting the ongoing need for vigilance and improvement in occupational safety measures. The over 25,000 work zone accidents annually in the state of Texas proves there is a need for increased safety measures regarding roadway construction areas (U.S. Department of Labor, 2004; Pamidimukkala et al., 2020).

Despite the significant number of fatalities, recent reports anticipate a consistent rise in infrastructure investment in the upcoming years (FMI, 2018). As the number of highway projects increases, employees will face greater safety hazards, leading to a higher likelihood of accidents, injuries, and fatalities. Consequently, it is crucial to implement appropriate measures to effectively reduce the heightened safety risks. Therefore, the objectives of this study are to (1) develop training materials for employers and employees, (2) deliver training sessions and educate them about the recognition and prevention of hazardous situations associated with work zones and roadway construction that result in injuries and fatalities, and (3) evaluate the training sessions. This initiative is crucial for enhancing the overall safety of construction sites, particularly in high-risk areas like roadways and highways. This study aims to create a safer work environment in construction zones by equipping workers and employers with the knowledge necessary to identify and mitigate risks effectively.

METHODOLOGY

As illustrated in Figure 1, a three-step methodology was designed to fulfill the objectives of the research. The first step was to review related published documents and collect materials on safety and health challenges of the workforce active in work zones and roadway construction. In addition, strategies and best practices which were adopted to prevent and/or overcome hazardous situations during roadway, highway, and bridge construction were identified and collected. In this regard apart from scholarly publications, two of the major and primary sources used were OSHA and National Cooperation Highway Research Program (NCHRP) published documents. The review and collection of the existing documents were divided into (1) guidance for

workers/labours/engineers active in roadway construction fields and sites, and (2) guidance for construction employers of roadway workforce. In the second step, the training modules and supplementary materials were developed and in the last step the training materials were pilot tested on several workers and employers working in roadway construction.

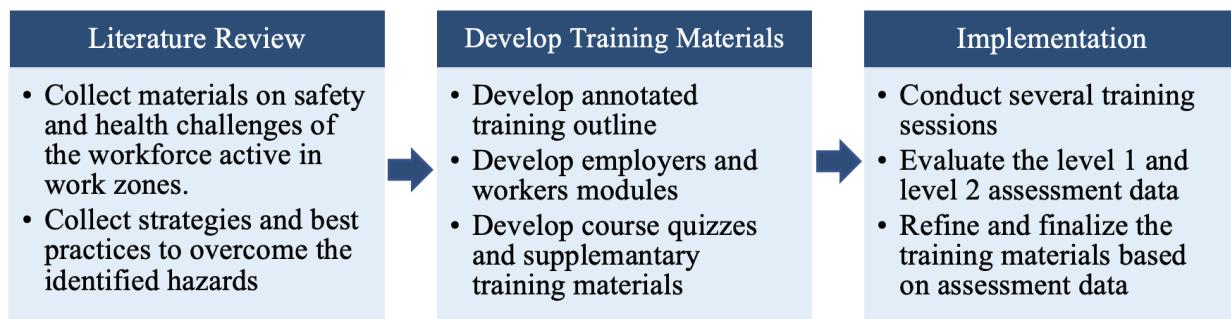


Figure 1. Methodology

RESULTS

Develop Training Materials

The team developed new modules for the following two sets of courses: 4-hour training for workforce in roadway and highway construction (e.g., labours, inspectors, heavy machine operators, etc.) and 3-hour training for small contractor employers active in roadway and highway construction industry. The organized outline for the training materials instructing construction employees and employers is presented in Table 1.

As presented in Table 1, the research team has developed modules on hazards associated with work zones and roadway construction in the training materials to instruct workers on the exposures to such hazards and their preventive methods along with hierarchical controls. The hierarchy of controls is a way of determining which actions will best control exposures; therefore, the team included controls to prevent such hazards at work zones in roadway construction. As it is very important that all workers receive thorough training on how to work next to motor vehicle traffic at nights in a manner that reduces their risk of experiencing a safety hazard, a module on night work hazards was developed. The research team has also developed a module on workers' rights, and employer responsibilities to instruct workers and their employers about their rights and responsibilities at construction work zones. All the modules were developed in the PowerPoint and each of the slides have a footnote consisting of key message, background, interactivity, notes, and references. The research team has also developed and included class handouts for all training modules. Class handouts contain multiple quiz questions based on the topic of the module. These questions were also included in the PowerPoint presentations of the modules. Figure 2 presents the layout of the modules developed in the PowerPoint. Both PDF and PowerPoint versions of the modules were developed and stored. Next, the team has developed supplementary training materials including instructor and participant guides and training assessments.

Table 1. Outline of Occupational Safety and Health Training Materials for Roadway Construction Employees and Employers

Type	Name	Description
Modules	Module 0	Introduction to the Program
	Module 1	General Information
	Module 2	Accidents in Construction
	Module 3	Personal Safety
	Module 4	Falls
	Module 5	Struck-by
	Module 6	Electrocutions
	Module 7	Caught-in/between
	Module 8	Soft Tissue Injuries
	Module 9	Night Work Hazards
	Module 10	Fire Protection and Prevention
	Module 11	Signs, Signals and Barricades
	Module 12	Workers' Rights/ Employers' Responsibilities
	Module 13	OSHA's Construction Safety Programs
	Module 14	Summary
Supplementary	Pre-training Assessment	Assessment of the participants at the beginning of training session
	Post-training Assessment	Assessment of the participants at the end of training session
	Training Evaluation Form	Evaluation of the training event

In order to assess the efficacy of the training program and improve the contents, two assessment methodologies were devised. Level 1 training session evaluation assesses trainees' responses to the training, including their perceptions of the training environment, instructors, and the overall quality and usefulness of the training. In level 2 learning assessment, the team measures the skills, knowledge, and safety attitude the trainees acquire and retain. The team implemented the pre- and post-test assessment by distributing the 10 case examples before and after the workshop. This method of assessment yields a more accurate results regarding the effectiveness of the training materials. Figure 3 presents an example of Level 1 and Level 2 assessments.

Implementation

After finalizing the training materials, the research team pilot tested the materials to the roadway workers and employers. The training sessions for employers and workers is of three hours and four hours respectively. After the end of each session, the trainees were requested to fill the level 1 and level 2 training assessments. The attendees were instructed to provide their comments on how to enhance and improve the training materials on level 1 assessment. Next, the team has implemented the level 2 learning assessment by distributing ten questions before and after the workshop. After each training session is completed, the responses collected for both level 1 and 2 training assessments were analysed qualitatively and quantitatively. In level 1 assessment, the team evaluated the criteria on quality of the training modules and topics, skills of the instructors, length of the training, effectiveness of supplementary materials including handouts/quizzes and case studies, etc. It was revealed that, the trainees expressed satisfaction

with the design and contents of the training materials and no significant feedback was received from the trainees regarding any necessary modifications to the modules.

INTRODUCTION

- Soft tissue injuries affect muscles, nerves, tendons, ligaments, joints and spinal discs.
- They are the most prevalent type of injury in the construction industry!
- Although not fatal, these types of injuries can be disabling, cause constant pain, and prevent an individual from leading a normal lifestyle.

US Department of Labor
Occupational Safety and Health Administration



Susan Harwood Training Program

QUESTIONS

3. Which class of high-visibility clothing should workers who work where traffic speeds may exceed 50 miles per hour?

A. 1
B. 2
C. 3
D. 4

US Department of Labor
Occupational Safety and Health Administration

Susan Harwood Training Program

Key Message
Introduction
Background Information

- Many activities can lead to soft tissue injuries. These injuries are not fun! Just ask anyone who has sprained an ankle or been laid up with a serious back injury.
- STIs are injuries to the musculoskeletal structure – the joint tissues, ligaments, tendons, and muscles – and are better known as strains and sprains.
- Although most soft tissue injuries are not fatal, they can cause years of pain and suffering for workers and their families.

Interactivity
 None
Notes
 None
References
 Islam, M. Rashad, "Construction safety: health, practices and OSHA" (2022)
 Construction Safety and Health, "Caught-in between hazards." Retrieved from <https://www.osha.gov/sites/default/files/2018-12/fy07_sh-16586-07_3_caughtinbetween_hazards_trainer_guide.pdf>, Retrieved 26 November 2022

Provide participants with handouts of questions to which they may respond. Allocate ten minutes for them to answer the questions. After that, provide the correct answers and discuss them with the participants

Answers
 3. C

Figure 2. Example PowerPoint Slides

Training Evaluation Form

	Course Evaluation						
	Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree
1. Learning objectives of the course were clearly explained	<input type="radio"/>						
2. Goals and objectives of this course were useful for my learning needs	<input type="radio"/>						
3. Course content was relevant to my daily work	<input type="radio"/>						
4. Materials were well organized and easy to understand	<input type="radio"/>						
5. The time distribution was adequate for every module	<input type="radio"/>						
6. This educational experience will help me to maintain safety at workplace	<input type="radio"/>						
7. I would recommend this training to others in my field	<input type="radio"/>						
Instructor Evaluation	Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree
8. Instructor presented information in clear, understandable manner	<input type="radio"/>						
9. Instructor demonstrated thorough knowledge of subject	<input type="radio"/>						
10. Instructor answered questions satisfactorily	<input type="radio"/>						
11. Instructor facilitated the discussions effectively	<input type="radio"/>						
12. What aspects of the course were most beneficial to you? Why?	<hr/> <hr/>						
13. What changes (if any) would you suggest to improve this course?	<hr/> <hr/>						

Pre-and Post-Training Assessment

1. What is the most effective method in the order of hierarchical controls?
 - Engineering Controls
 - Administrative controls
 - Personal protective equipment
 - None of the above
2. State whether the following statement is True/False. Personal protective equipment (PPE) is the last line of defense against electrical hazards in the workplace.
 - True
 - False
3. What are the construction fatal flaws which contribute for more than 50% of accidents in construction work zone?
 - Falls
 - Struck-by
 - Caught-in-between
 - Electrocution
 - All the above
4. State whether the following statement is True/False. A signal person should be stationed behind vehicles that have obstructed the way in construction work zone.
 - True
 - False
5. Which is not the primary causes of Fall-Related accidents?
 - Improper connected walking/working surfaces.
 - Improper use of fall protection equipment
 - Unprotected sides, edges and holes
 - Safety Net systems
6. State whether the following statement is True/False. Soft tissue injuries are injuries that affect your muscles, nerves, tendons, bones, joints, and spinal discs.
 - True
 - False
7. Which are the types of caught-in-between hazards on construction sites?
 - Pinned-in-between
 - Pinched-between parts
 - Buried-in-between
 - All the above
8. What requirements should be followed when performing utility work near energized electrical hazards?
 - NFPA 70A
 - NFPA 70B
 - NFPA 70C
 - NFPA 70E
9. In which class of fire does the electrical wiring and fuse box fall under?
 - Class A
 - Class B
 - Class C
 - Class D
10. Which of the following are worker's rights?
 - A safe workplace
 - Access to accurate information that employers have collected on hazards in the workplace
 - File a complaint with OSHA if they believe hazardous safety or health conditions exist in the workplace
 - All the above

Figure 3. Level 1 Assessment (left), and Level 2 Assessment (Right)

In level 2 assessment, the team analysed the obtained skills and knowledge of the trainees using pre- and post-training evaluation technique. In this regard, the responses of the attendees on pre- and post-training case studies were comparatively analysed to measure their level of improvements in their knowledge and skills in recognition and prevention of hazards associated with work zones and roadway construction. The results revealed that the average percentage of questions correctly answered by all the participants' during pre-training was 61%, and the average percentage correct responses on the post-training assessment questions were raised to 89%, after the training sessions were delivered. Figure 4. presents the comparison scores of pre- and post-training assessments of each of ten questions.

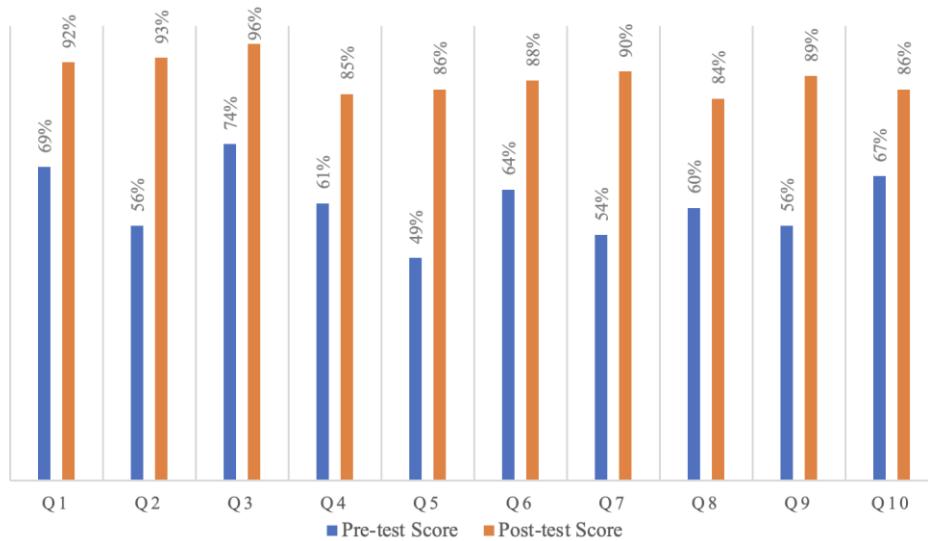


Figure 4. Pre- and Post- Test Comparison Scores

CONCLUSION

The objective of this study was to develop and implement occupational hazard training for workers and employers in roadway construction. An extensive literature review was performed and a training outline was designed based on the reviewed materials. Next, training modules and supplementary materials were developed, and several pilot training workshops were conducted for workers and employers in Texas. After conducting the training sessions, the participants' knowledge was assessed before and after the training sessions through pre- and post-training tests. The questionnaire comprised ten questions regarding the subject matter addressed in the training session. The collected assessments were subjected to quantitative analysis, which demonstrated a statistically significant disparity in the participants' knowledge levels before and after the training. The materials developed in this study may aid highway/roadway construction companies improve the safety of their workforce.

REFERENCES

- Ammar, A., and Dadi, G. 2023. Evaluation of Near-Miss Reporting Program Perceived by Employees' Challenges and Opportunities. In *International Conference on Transportation and Development 2023* (pp. 695-705).

- BLS (Bureau of Labor Statistics). 2018. Bureau of Labor Statistics (BLS), Census of Fatal Occupational Injuries (CFOI)—Current and Revised Data, Retrieved from <https://www.bls.gov/iif/oshcfoi1.html>. Accessed 30 October 2023.
- BLS (Bureau of Labor Statistics). 2022. Bureau of Labor Statistics, National census of Fatal Occupational Injuries. Retrieved from <https://www.bls.gov/news.release/pdf/cfoi.pdf>. Accessed 5 June 2023.
- CDC (Centers for Disease Control and Prevention). 2022. Centers for Disease Control and Prevention. Highway work zones and signs, signals, and barricades. Retrieved from <https://www.cdc.gov/niosh/topics/highwayworkzones/default.html>. Accessed 12 July 2023.
- Subramanya, K., Kermanshachi, S., and Ramaji, I. 2022. Health and safety of construction field workforce active in extreme weather conditions. In *Construction Research Congress 2022* (pp. 737-747).
- FHWA. 2018. Work Zone Facts and Statistics. Available at: https://ops.fhwa.dot.gov/wz/resources/facts_stats.html. Accessed 5 March 2022.
- FMI. 2018. FMI 2018 FMI Overview: Featuring FMI's Fourth Quarter 2017 Construction Outlook Retrieved from https://www.fminet.com/wp-content/uploads/2018/01/Overview2018_FINAL.pdf. Accessed 23 April 2022.
- Highway Worker Safety Program. 2022. The Associate General Contractors of America. Retrieved from <https://www.agc.org/highway-worker-safety>. Accessed 12 July 2023.
- Huebschman, C. R., Garcia, C., Bullock, D. M., and Abraham, D. M. 2003. *Construction work zone safety*.
- Kermanshachi, S., Dao, B., Rouhanizadeh, B., Shane, J., and Anderson, S. 2020. Development of the project complexity assessment and management framework for heavy industrial projects. *International Journal of Construction Education and Research*, 16(1), 24-42.
- Kermanshachi, S., and Safapour, E. 2019. Identification and quantification of project complexity from perspective of primary stakeholders in US construction projects. *Journal of Civil Engineering and Management*, 25(4), 380-398.
- Kermanshachi, S., Thakur, R., and Govan, P. 2018. Discovering the impact of late change orders and rework on labor productivity: a water treatment case study analysis using system dynamics modeling. In *Construction Research Congress 2018* (pp. 691-701).
- Li, Y., Hu, Y., Xia, B., Skitmore, M., and Li, H. 2018. "Proactive behavior-based system for controlling safety risks in urban highway construction megaprojects." *Autom. Constr.* 95 (Nov): 118–128.
- Long, S., Smith, B. K., Ng, E. H., and Sun, C. 2014. Work zone safety: physical and behavioral barriers in accident prevention (No. cmr 14-013). Missouri. Dept. of Transportation. Division of Construction and Materials.
- Nipa, T. J., and Kermanshachi, S. 2022. Resilience measurement in highway and roadway infrastructures: Experts' perspectives. *Progress in disaster science*, 14, 100230.
- Nipa, T. J., Kermanshachi, S., and Patel, R. K. 2022. Analysis of the resilience management dimensions based on project complexity level. In *Construction Research Congress 2022* (pp. 80-89).
- Rad, S. K., and Kermanshachi, S. 2018. Development of project life cycle communication ladder framework using factor Analysis method. In *Construction Research Congress 2018* (pp. 543-552).
- Rashad, I. M. 2022. *Construction Safety Health, Practices, and OSHA*. McGraw Hill.

- Rathnasiri, N., De Silva, N., and Wijesundara, J. 2023. State of the art in work zone safety: A systematic review. *International Journal of Transportation Science and Technology*.
- OSHA Education Center. 2023. OSHA 30-Hour Construction Study Guide. Retrieved from <https://www.oshaeducationcenter.com/30-hour-study-guides/>. Accessed 19 August 2023.
- Pamidimukkala, A., and Kermanshachi, S. 2022. Assessment of Effectiveness of Occupational Hazards Training for Women in the Construction Industry. In *International Conference on Transportation and Development 2022* (pp. 270-279).
- Pamidimukkala, A., and Kermanshachi, S. 2023. Occupational Challenges of Women in Construction Industry: Development of Overcoming Strategies Using Delphi Technique. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 15(1), 04522028.
- Pamidimukkala, A., Kermanshachi, S., and Safapour, E. 2020. Challenges in Post-Disaster Housing Reconstruction: Analysis of Urban vs. Rural Communities. In *Creative Construction e-Conference 2020* (pp. 103-110). Budapest University of Technology and Economics.
- Namian, M., Kermanshachi, S., Khalid, M., and Al-Bayati, A. J. 2020. Construction safety training: Exploring different perspectives of construction managers and workers. In *2020 ASEE Virtual Annual Conference Content Access*.
- Safapour, E., Kermanshachi, S., Shane, J., and Anderson, S. 2017. Exploring and assessing the utilization of best practices for achieving excellence in construction projects. In *proceedings of the 6th CSCE International Construction Specialty Conference* (pp. 1-9).
- Safapour, E., Kermanshachi, S., and Pamidimukkala, A. 2021. Post-disaster recovery in urban and rural communities: Challenges and strategies. *International Journal of Disaster Risk Reduction*, 64, 102535.
- Safapour, E., Kermanshachi, S., and Ramaji, I. 2023. Selection of best practices that enhance phase-based cost and schedule performances in complex construction projects. *Engineering Management Journal*, 35(1), 84-99.
- Subramanya, K., Kermanshachi, S., and Rouhanizadeh, B. 2020. Modular construction vs. traditional construction: Advantages and limitations: A comparative study. In *Creative Construction e-Conference 2020* (pp. 11-19). Budapest University of Technology and Economics.
- Traffic Safety Facts. 2008. National Highway Traffic Safety Administration. Report No. DOT HS 810 631, US Department of Transportation, Washington, D.C., 2009.
- US Census Bureau. 2019. Construction Expenditures Branch-Construction Spending Survey. Retrieved from http://www.census.gov/construction/c30/historical_data.html, (Accessed 11 April 2019).
- US Department of Commerce. 2019. Gross-Domestic-Product-(GDP)-by-Industry Data, Retrieved from <https://apps.bea.gov/iTable/iTable.cfm?ReqID=51&step=1>. Accessed 11 April 2023.
- US Department of Labor, Occupational Safety and Health Administration. 2004. OSH Act of 1970. Retrieved from <https://www.osha.gov/laws-regulations/completeoshaact>. Accessed 25 November 2023.
- Yang, H., Ozbay, K., Ozturk, O., and Xie, K. 2015. Work zone safety analysis and modeling: a state-of-the-art review. *Traffic injury prevention*, 16(4), 387-396.

A Review on the Mental Health Stressors of Construction Workers

Deema Almaskati¹; Sharareh Kermanshachi, Ph.D., P.E.²; and Apurva Pamidimukkala, Ph.D.³

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: deemanabeel.almaskati@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Penn State Univ.
Email: svk5464@psu.edu

³Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

ABSTRACT

The mental health of construction workers is often neglected despite their overrepresentation in deaths associated with suicide and other mental health conditions. The nature of the construction sector is inherently stressful, and failure to manage individual stress levels can manifest in depression, which can eventually lead to suicide ideation and suicide. This study identifies prevalent mental health conditions within the industry, defines mental health stressors on construction worksites and determines their frequency of citation, and discusses best practices for the promotion of a healthy work environment. Through the identification of 13 factors negatively affecting their mental health, the findings highlight the significance of protecting not only the physical but also the mental health of construction personnel. The most frequently cited mental health stressors were found to be excessive work demand, work life imbalance, and poor working environments. Additionally, the study showed that the effective management of workplace mental health involves the implementation of a strategy considering the integration of primary, secondary, and tertiary interventions. The promotion of a help-seeking culture may also help reduce the stigma around mental health and diminish internal shame, specifically for male construction workers.

Keywords: construction industry; mental health; occupational stress; suicide; opioid; construction safety

INTRODUCTION

There has been an extensive focus within the industry on safety hazards and how their presence can be reduced to safeguard the physical wellbeing of construction employees. However, this attention is often not reciprocated for the protection of workers' mental health. In comparison to other industries, mental health issues are most common in the construction industry, which experiences high suicide rates in multiple countries globally including the United States, Australia, and the United Kingdom (Samaraweera et al. 2022). The construction industry has the second-highest suicide rate in the US among all occupational categories (Eyllon et al. 2020). The suicide rate for men in the construction industry was 56 fatalities per 100,000 people, which is roughly double the national rate for male suicide (Sussell 2023).

Construction projects are characterized by strict timelines and constrained budgets, which, alongside a slew of other factors, result in a stressful work environment (Nipa et al. 2022). Failure to manage stress levels can have negative implications on an individual's mental health.

According to the World Health Organization, mental health refers to a person's ability to fulfill their potential, successfully handle day-to-day stressors, work productively, and make a positive contribution to society (Sun et al. 2022). Poor mental health is a significant health and safety issue that negatively impacts productivity, resulting in both personal and professional losses and posing a threat to the safety and performance of construction projects (Nwaogu et al. 2020). Anxiety, depression, and substance use disorders are common mental health conditions observed in construction personnel (Frimpong et al. 2023). It is estimated that over 55% of construction workers in the United Kingdom have experienced mental health issues at some point in their lives and 42% of them experience poor mental health due to their current place of employment (Kotera et al. 2019). There has also been increased evidence that younger construction workers are at a higher risk of developing poor mental health conditions in comparison to their older colleagues, suffering from high levels of depression, anxiety, and sleep disorders (Frimpong et al. 2023). Additionally, they have a much higher propensity for substance use disorders and are three times more likely to develop opioid or alcohol dependency (Dong et al. 2019).

Due to the alarming statistics on poor mental health in construction alongside high suicide rates, researchers worldwide have increasingly attempted to understand the factors that contribute to these issues (Nwaogu et al. 2020). For instance, Frimpong et al. (2022) sought to determine the frequency of physical and mental health conditions affecting Ghana's youth in the construction industry, whereas Dong et al. (2019) analyzed data on the use of opioids and associated overdose fatalities in the construction industry. Other studies such as those conducted by Sun et al. (2022) and Golzad et al. (2023) have identified mental health causation factors. However, these studies failed to determine the factors' frequency of citation. Additionally, there was a lack of discussion on the best practices and strategies that can be put in place to reduce the influence of these factors on workers' mental health.

As construction projects are highly dependent on the labor of their workforce, protecting not only the physical but also the mental health of employees is paramount to the success of the industry as a whole. Thus, the following objectives were developed to fulfill the aims of this study: (1) recognize the mental health issues prevalent in the industry, (2) identify common mental health stressors, (3) rank the mental health stressors based on their frequency of citation within the literature, and (4) discuss best practices to support employee mental health and promote mental wellness. Through this review, construction professionals and supervisors will better understand how the workplace can be designed to better support the wellbeing of its workers and how resources can be best allocated to manage the mental health of construction workers.

METHODOLOGY

In order to document the key findings of earlier pertinent studies and synthesize the current body of knowledge on mental health in the construction sector, this study performed a comprehensive literature review. An initial search was carried out utilizing online research databases such as Google Scholar, Elsevier, and American Society of Civil Engineers. Relevant studies were compiled through the inputting the following keywords in various combinations: "construction", "mental health", "psychosocial hazards", "opioid", "substance use", and "suicide". The articles gathered then underwent a screening and exclusion process to determine their suitability for this study, as shown in Figure 1. Following the keyword search, inappropriate document types, duplicates, and studies in non-English languages were excluded. Articles were

then filtered out based on their title and abstract. Next, their full texts were reviewed to determine whether the paper was relevant to the objectives of this study. Additional studies were identified from the reference lists of the eligible literature selected and underwent the same screening and exclusion process. Through the evaluation of 98 studies, the literature review resulted in the identification of factors contributing to increased mental health problems in construction personnel and the discussion on best practices.

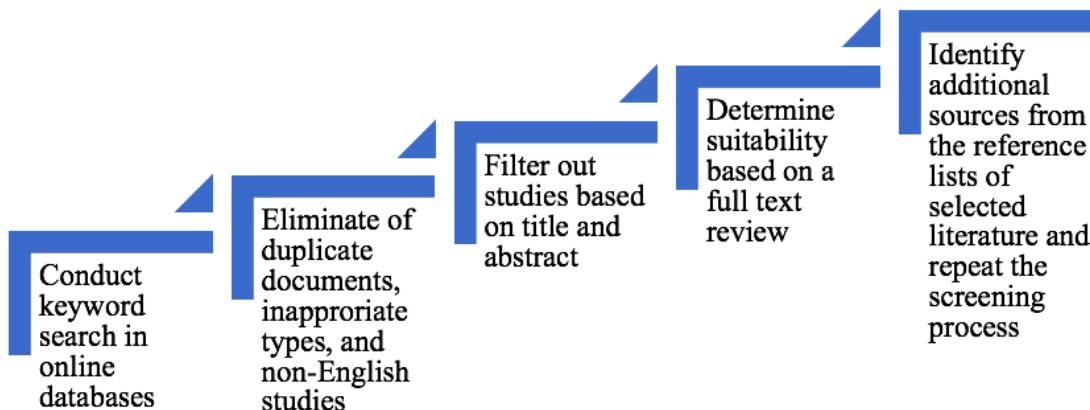


Figure 1. Screening and exclusion process for selection of relevant studies

RESULTS & DISCUSSIONS

Mental Health Conditions in the Construction Industry

On a global scale, suicide is the most common cause of death of construction workers with high mortality rates in countries including as the United States, Canada, Denmark, United Kingdom, and Australia (King and Lamontagne 2021). The effects of suicide are unquantifiable, having negative impacts on the individual's family, friends, and the economy as a whole (Tijani et al. 2023). Suicide and suicide ideation are not the only mental health conditions prevalent in the construction sector. Construction workers also have high rates of opioid use disorders, which are often used for the management of pain associated with work-related musculoskeletal disorders (Shaw et al. 2020). Long-term use of these drugs for pain management can have significant risks, leading to overdose, suicide, and the manifestation of other mental health issues such as depression (Dale et al. 2021). Those employed in the construction industry have been found to be disproportionately impacted by opioid use disorders, having the highest mortality rate from opioids amongst all other occupations in the United States (Dale et al. 2021; Gomes et al. 2022). Other common mental health conditions prevalent within the industry include depression, mania, anxiety disorders, alcohol use, sleep disorders, and post-traumatic stress disorder (Chan et al. 2020a; Frimpong et al. 2023).

Mental Health Stressors in Construction

The prevalence of mental health conditions in the construction industry can be better understood through the recognition of the factors causing occupational stress. The nature of the industry places a lot of pressure on construction workers, who must complete projects on time

within a certain budget while meeting a specific level of standards (Sun et al. 2022). While the contribution of their work to society cannot be overstated, the mental health of construction workers can often be overlooked to prioritize project completion. Due to the demanding nature of the industry, construction workers face increased levels of work-related stress, a high-risk factor for the development of depression, which, consequently, is a high-risk factor for suicide (Nwaogu et al. 2020). To better support their employees and foster a workplace that prioritizes their mental wellbeing, it is essential that construction employers understand the factors that negatively impact their workers' mental health. Thus, a list of mental health stressors in the construction sector has been compiled and is outlined in Table 1.

Excessive Job Demand

Excessive job demand was the most frequently cited mental health stressor and is likely to result in burnout and other mental health problems (Sun et al. 2020). Factors such as work overload, long work hours, overlapping deadlines, and excessive workloads should be mitigated through job reallocation and other interventions to reduce fatigue associated with high job demand (Kermanshachi and Safapour 2019; Golzad et al. 2023).

Work Life Imbalance

Work life imbalance was the second most frequently cited stressor in and it occurs when an individual is unable to integrate professional and personal demands effectively due to the demanding nature of their job (Grant et al. 2019). Previous studies have found poor work life balance to have detrimental impacts on the perceived health and safety of construction workers (Chan et al. 2020b).

Poor Working Environment

Construction personnel may have concerns regarding unsafe work environments, lack of safety equipment, or lack of safety support because of the hazardous nature of construction sites (Milner et al. 2017; Safapour et al. 2017; Pamidimukkala et al. 2022). In addition to emphasizing the importance of hazard awareness and management (Pamidimukkala and Kermanshachi 2022), site supervisors must ensure that productivity is not prioritized at the expense of the safety of their workers (Safapour and Kermanshachi 2020; Karthick et al. 2022).

Alcohol and Drug Use

Physical and mental stresses of the job may drive construction workers to increasingly depend on the use of alcohol and other drugs (Mushi and Manege 2018). Alcohol consumption in particular is associated with psychomotor dysfunction and reduced cognitive ability, compromising work ability and productivity (Sawick and Szóstak 2020).

Low Job Support

An employee may experience a lack of adequate emotional and practical support needed to effectively perform their job. Because construction projects are often short-term in nature, it can be challenging for companies to offer their employees long-term support (Wu et al. 2019).

Table 1. Identification of factors impacting mental health of construction personnel

No.	Mental Health Stressor	Description	Previous Study	Frequency	Rank
1	Excessive job demand	Burnout caused by long work hours, tight deadlines, and other factors	Leung et al. 2016; Kermanshachi et al. 2018	35	1
2	Work life imbalance	Dedicates a disproportionate amount of time and energy to work obligations in lieu of personal commitments and relationships	Milner et al. 2017	32	2
3	Poor working environment and safety concerns	Presence of hazards in the workplace (noise, poor lighting, faulty equipment and machines, etc.)	Bowen et al. 2018; Karthick et al. 2023	28	3
4	Alcohol and drug use	Substance dependency or use caused by physical ailments or stressful nature of the job	Dale et al. 2021	25	4
5	Low job support	Lack of sufficient support from supervisors and colleagues	Boschman et al. 2013	21	5
6	Low job control	Lack of task autonomy or minimal opportunity to be involved in decision-making	Chan et al. 2020a	19	6
7	Job insecurity	Fear of losing one's existing job or of not having job stability	Langdon and Sawang 2018; Pamidimukkala et al. 2021	18	7
8	Role conflict	Incompatible tasks and expectations from different managerial persons the employee interacts with or reports to	De Silva et al. 2017	16	8
9	Workplace injustice	Discrimination based on age, gender, race, or religious status	Chan et al. 2020a	15	9
10	Interpersonal conflict	Tensions and disagreements between employees in the workplace	Golzad et al. 2023	14	10
11	Low recognition and reward opportunities	Inadequate reward or recognition	Sun et al. 2022	11	11
12	Financial concern	Fears regarding financial difficulties and instability or insufficient income	Bowers et al. 2018	9	12
13	Role ambiguity or confusion	Uncertainty or lack of clarity on job duties	Lingard and Turner 2017	7	13

Low Job Control

Low job control is characterized by a lack of autonomy or ability to participate in decision-making, the enforcement of rigid work schedules and strict rules, and authoritarian work cultures (Chan et al. 2020a). This mental health stressor can have detrimental effects on work stress and can be mitigated through increasing worker autonomy through the fostering of a supportive work environment (Golzad et al. 2023).

Job Insecurity

The temporary nature of construction projects can cause concern for on-site construction personnel regarding the certainty of their future employment (Sun et al. 2022). The construction sector is characterized by its project-based nature, and thus many lower-level construction workers have been known to suffer from job insecurity (Golzad et al. 2023). Insecurity surrounding job certainty and future career prospects has been linked to anxiety and work stress as well as poor work performance (Chan et al. 2020b).

Role Conflict

In certain instances, an employee may be faced with two or more expectations from two or more different parties. Role conflict occurs when the individual is unable to meet all of these expectations simultaneously due to the incompatibility and inconsistency in the demands of these expectations (Dodanwala et al. 2021). This type of confusion may prevent the individual from fulfilling their job expectations and may hinder their productivity.

Workplace Injustice

Workplace injustice is a broad term that encompasses all forms of disrespect and inequality within the workplace including all forms of harassment, bullying, and discrimination (Nwaogu et al. 2023). Gender discrimination is frequently reported in construction workplaces with female professionals experiencing a lack of respect from subordinates in addition to suffering verbal, physical, and sexual harassment (Pamidimukkala and Kermanshachi 2023). The gender pay gap is another form of workplace injustice faced by female construction professionals which occurs when they are being paid lower salaries than their male counterparts (Chan et al. 2020a).

Interpersonal Conflict

Interpersonal conflicts are more likely to occur in construction projects due to their inherent complexity and uncertainty in addition to the involvement of various stakeholders (Kermanshachi et al. 2020; Sun et al. 2022). Ineffective conflict management can have significant negative effects on project quality and the mental health of those involved (Safapour et al. 2019; Wang and Liu 2021).

Low Recognition and Reward Opportunities

Employee demotivation and the belief that their efforts are insufficient are likely outcomes of inadequate rewards and recognition (Aung et al. 2023). Reward and recognition opportunities

include verbal affirmations, increased job responsibilities, greater job autonomy, and monetary rewards (Azeez et al. 2019).

Financial Concern

Financial concern is a mental health stressor that causes anxiety on job security in addition to how an individual will be able to fulfill their financial obligations (Pamidimukkala and Kermanshachi 2021). Financial strain has been associated with reduced cognition and productivity at work (Bowers et al. 2018).

Role Ambiguity

Role ambiguity and confusion occurs when the employee is unsure of their role within the workplace and is unable to obtain a clear understanding of job expectations (Wu et al. 2019). This stressor can occur when the individual is unclear on their responsibilities due to project complexity, limited project information, and tight time frames (Leung and Chan 2012; Kermanshachi et al. 2023). Role ambiguity can be reduced through enhanced information clarity and effective job allocation (Dodanwala et al. 2021).

Best Practices

In order to maintain a healthy workplace, employers are obligated to support their employees and provide a work environment conducive to the promotion of positive health and wellbeing (Campbell and Gunning 2022). The best practices for effective workplace intervention to manage occupational mental health should consider a combination of primary, secondary, and tertiary interventions. Primary interventions aim to minimize mental health stressors, secondary interventions are targeted towards helping employees manage those stressors, and tertiary interventions involve the provision of counseling or other tools in response to the employee's mental health condition (Nwaogu et al. 2022). These interventions can be applied for alcohol and drug use, a mental health stressor identified in Table 1. As previously discussed, the construction industry is characterized by high rates of opioid use with the majority of these workers using prescription opioids for the management of pain due to musculoskeletal disorders or other occupational injuries (Dale et al. 2021). A primary intervention for opioid dependency would involve reducing the employee's risk of developing work injuries and musculoskeletal disorders, whereas a secondary intervention would be providing the employee with alternative pain management tools, and a tertiary intervention would be supporting the employee through the provision of counseling or other mental health services aimed at reducing opioid dependency (Ompad et al. 2019; Gomes et al. 2022).

The prevalence of mental health conditions in the construction industry can also be mitigated through prevention strategies focused on promoting a help-seeking culture (Newaz et al. 2022). Many construction workers avoid seeking out help due to internal shame and the stigma surrounding mental health (Kotera et al. 2019). Elevated suicide rates have been observed for male construction workers (Samaraweera et al. 2022), a statistic that Campbell and Gunning (2020) theorized was directly related to failure to seek support due to mental health shame associated with masculinity. To mitigate this, the introduction of self-compassion training in order to reduce shame associated with mental health conditions is suggested (Kotera et al. 2019).

Such training tools may be utilized in combination with enhanced access to mental health services. Providing employees with easily accessible mental health tools, such as online and telehealth services can greatly reduce the stigma surrounding mental health and help individuals manage any work-related stress (King and Lamontagne 2021).

CONCLUSION

In order to highlight the significance of the issue of mental health in construction, this study aims to present a thorough assessment of the existing body of knowledge. Because the profitability of the construction industry as a whole is highly dependent on labor of its workforce, it is critical to safeguard the mental and physical well-being of its workers. Anxiety, alcohol use, depression, obsessive compulsive disorder, sleep disorders, mania, opioid and other drug use, and suicide were all found to be common mental health conditions among construction personnel. The industry was also found to suffer from disproportionate rates of both suicides and opioid use disorders.

While the nature of the industry is inherently stressful, construction employers can better understand the prevalence of mental health issues by identifying the factors that lead to increased occupational stress. Thus, this study compiled a list of 13 mental health stressors in construction including role conflicts, job ambiguity, safety concerns, and workplace injustice. It was revealed that excessive work demand, work life imbalance, poor working environments, substance use, and low job control were the top 5 most frequently cited factors contributing to poor mental health conditions. These stressors can be effectively managed through the utilization of a strategy that incorporates primary, secondary, and tertiary interventions. Additionally, the encouragement of a help-seeking culture may also assist employers in lessening the internal shame and stigma associated with mental health, particularly for male construction workers.

REFERENCES

- Aung, Z. M., San Santoso, D., and Dodanwala, T. C. (2023). "Effects of demotivational managerial practices on job satisfaction and job performance: Empirical evidence from Myanmar's construction industry." *Journal of Engineering and Technology Management*, 67.
- Azeez, M., Gambatese, J., and Hernandez, S. (2019). "What do construction workers really want? A study about representation, importance, and perception of US construction occupational rewards." *Journal of Construction Engineering and Management*, 145(7).
- Boschman, J. S., Van der Molen, H. F., Sluiter, J. K., and Frings-Dresen, M. H. W. (2013). "Psychosocial work environment and mental health among construction workers." *Applied ergonomics*, 44(5), 748-755.
- Bowen, P., Govender, R., Edwards, P., and Cattell, K. (2018). "Work-related contact, work-life conflict, psychological distress and sleep problems experienced by construction professionals: an integrated explanatory model", *Construction Management and Economics*, 36, 153-174.
- Bowers, J., Lo, J., Miller, P., Mawren, D., and Jones, B. (2018). "Psychological distress in remote mining and construction workers in Australia." *Medical Journal of Australia*, 208(9), 391-397.

- Campbell, M. A., and Gunning, J. G. (2020). "Strategies to improve mental health and well-being within the UK construction industry." *Proceedings of the Institution of Civil Engineers-Management, Procurement and Law*, 173(2), 64-74.
- Chan, A. P. C., Chiang, Y. H., Wong, F. K. W., Liang, S., and Abidoye, F. A. (2020b). "Work-life balance for construction manual workers." *Journal of Construction Engineering and Management*, 146(5), p.04020031.
- Chan, A. P., Nwaogu, J. M., and Naslund, J. A. (2020a). "Mental ill-health risk factors in the construction industry: Systematic review." *Journal of construction engineering and management*, 146(3), 04020004.
- Dale, A. M., Buckner-Petty, S., Evanoff, B. A., and Gage, B. F. (2021). "Predictors of long-term opioid use and opioid use disorder among construction workers: Analysis of claims data." *American journal of industrial medicine*, 64(1), 48-57.
- De Silva, N., Samanmali, R., and De Silva, H. L. (2017). "Managing occupational stress of professionals in large construction projects." *Journal of Engineering, Design and Technology*, 15(4), 488-504.
- Dodanwala, T. C., Shrestha, P., and Santoso, D. S. (2021). "Role conflict related job stress among construction professionals: The moderating role of age and organization tenure." *Construction Economics and Building*, 21(4), 21-37.
- Dong, X. S., Brooks, R. D., and Cain, C. T. (2019). "Overdose fatalities at worksites and opioid use in the construction industry."
- Eyllon, M., Vallas, S. P., Dennerlein, J. T., Garverich, S., Weinstein, D., Owens, K., and Lincoln, A. K. (2020). "Mental health stigma and wellbeing among commercial construction workers: a mixed methods study." *Journal of occupational and environmental medicine*, 62(8), 423-430.
- Frimpong, S., Antwi, A. B., Sunindijo, R. Y., Wang, C. C., Ampratwum, G., Dansoh, A., Boateng, E. S., Hagan, J. A., and Mensah, P. A. (2022). "Health status of young construction workers in the Global South: The case of Ghana." *Safety science*, 148.
- Frimpong, S., Sunindijo, R. Y., Wang, C. C., Boadu, E. F., Dansoh, A., and Fagbenro, R. K. (2023). "A scoping review of research on mental health conditions among young construction workers." *Construction Innovation*.
- Golzad, H., Teimoory, A., Mousavi, S. J., Bayramova, A., and Edwards, D. J. (2023). "Mental Health Causation in the Construction Industry: A Systematic Review Employing a Psychological Safety Climate Model." *Buildings*, 13(10), 2442.
- Gomes, T., et al. (2022). "Lives lost to opioid toxicity among Ontarians who worked in the construction industry." On behalf of: The Ontario Drug Policy Research Network, The Office of the Chief Coroner for Ontario, Ontario Forensic Pathology Service Public Health Ontario.
- Grant, C. A., Wallace, L. M., Spurgeon, P. C., Tramontano, C., and Charalampous, M. (2019). "Construction and initial validation of the E-Work Life Scale to measure remote e-working." *Employee Relations*, 41(1), 16-33.
- Karthick, S., Kermanshachi, S., and Namian, M. (2022). "Physical, mental, and emotional health of construction field labors working in extreme weather conditions: Challenges and overcoming strategies." In *Construction Research Congress 2022*, 726-736.
- Karthick, S., Kermanshachi, S., and Pamidimukkala, A. (2023). "Analysis of the Health and Safety Challenges Faced by Construction Workers in Extreme Hot Weather Conditions." *Journal of Legal Affairs and Dispute Resolution in Engineering*, 15(1).

- Kermanshachi, S., Thakur, R., and Govan, P. (2018). "Discovering the impact of late change orders and rework on labor productivity: a water treatment case study analysis using system dynamics modeling." In *Construction Research Congress 2018*, 691-701.
- Kermanshachi, S., and Safapour, E. (2019). "Identification and quantification of project complexity from perspective of primary stakeholders in US construction projects." *Journal of Civil Engineering and Management*, 25(4), 380-398.
- Kermanshachi, S., Dao, B., Rouhanizadeh, B., Shane, J., and Anderson, S. (2020). "Development of the project complexity assessment and management framework for heavy industrial projects". *International Journal of Construction Education and Research*, 16(1), 24-42.
- Kermanshachi, S., Nipa, T. J., and Dao, B. (2023). "Development of complexity management strategies for construction projects." *Journal of Engineering, Design and Technology*, 21(6), 1633-1657.
- King, T. L., and Lamontagne, A. D. (2021). COVID-19 and suicide risk in the construction sector: preparing for a perfect storm. *Scandinavian journal of public health*, 49(7), 774-778.
- Kotera, Y., Green, P., and Sheffield, D. (2019). "Mental health shame of UK construction workers: Relationship with masculinity, work motivation, and self-compassion." *Revista de Psicología del Trabajo y de las Organizaciones*, 35(2), 135-143.
- Langdon, R. R., and Sawang, S. (2018). "Construction workers' well-being: What leads to depression, anxiety, and stress?." *Journal of construction engineering and management*, 144(2).
- Leung, M. Y., and Chan, I. Y. S. (2012). "Exploring stressors of Hong Kong expatriate construction professionals in Mainland China: focus group study". *Journal of Construction Engineering and Management*, 138 (1), 78-88.
- Leung, M. Y., Liang, Q., and Olomolaiye, P. (2016). "Impact of job stressors and stress on the safety behavior and accidents of construction workers." *Journal of Management in Engineering*, 32(1).
- Lingard, H., and Turner, M. (2017). "Promoting construction workers' health: a multi-level system perspective." *Construction management and economics*, 35(5), 239-253.
- Milner, A., Maheen, H., Currier, D., and LaMontagne, A. D. (2017). "Male suicide among construction workers in Australia: A qualitative analysis of the major stressors precipitating death." *BMC Public Health*, 17, 1-9.
- Mushi, F. V., and Manege, S. L. (2018). "Alcohol abuse and illicit drug use at construction sites: Perception of workers at construction sites." *International Journal of Construction Engineering and Management*, 7(2), 65-72.
- Newaz, M. T., Giggins, H., and Ranasinghe, U. (2022). "A Critical Analysis of Risk Factors and Strategies to Improve Mental Health Issues of Construction Workers." *Sustainability*, 14(20).
- Nipa, T. J., Kermanshachi, S., and Patel, R. K. (2022). "Analysis of the resilience management dimensions based on project complexity level". In: *Construction Research Congress 2022*, 80-89.
- Nwaogu, J. M., Chan, A. P., Hon, C. K., and Darko, A. (2020). "Review of global mental health research in the construction industry: A science mapping approach." *Engineering, construction and architectural management*, 27(2), 385-410.
- Nwaogu, J. M., Chan, A. P., and Naslund, J. A. (2022). "Measures to improve the mental health of construction personnel based on expert opinions." *Journal of management in engineering*, 38(4).

- Nwaogu, J. M., Chan, A. P., and Akinyemi, T. A. (2023). "Conceptualizing the dynamics of mental health among construction supervisors." *International Journal of Construction Management*, 23(15), 2593-2613.
- Ompad, D. C., Gershon, R. R., Sandh, S., Acosta, P., and Palamar, J. J. (2019). "Construction trade and extraction workers: A population at high risk for drug use in the United States, 2005–2014." *Drug and alcohol dependence*, 205.
- Pamidimukkala, A., and Kermanshachi, S. (2021). "Impact of Covid-19 on field and office workforce in construction industry." *Project Leadership and Society*, 2, 100018.
- Pamidimukkala, A., Kermanshachi, S., and Jahan Nipa, T. (2021). "Impacts of COVID-19 on health and safety of workforce in construction industry." In *International Conference on Transportation and Development 2021*, 418-430.
- Pamidimukkala, A., and Kermanshachi, S. (2022). "Assessment of Effectiveness of Occupational Hazards Training for Women in the Construction Industry." In *International Conference on Transportation and Development 2022*, 270–279.
- Pamidimukkala, A., Kermanshachi, S., and Jahan Nipa, T. (2022). "Safety Risks of Reconstruction Workers in Clean-Up and Recovery Phase due to Natural Hazards." In *Construction Research Congress 2022*, 520-530.
- Pamidimukkala, A., and Kermanshachi, S. (2023). "Occupational Challenges of Women in Construction Industry: Development of Overcoming Strategies Using Delphi Technique." *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 15(1).
- Safapour, E., Kermanshachi, S., and Kamalirad, S. (2019). "Development of effective communication network in construction projects using structural equation modeling technique." In *ASCE International Conference on Computing in Civil Engineering 2019*, 513-521.
- Safapour, E., Kermanshachi, S., Shane, J., and Anderson, S. (2017). "Exploring and assessing the utilization of best practices for achieving excellence in construction projects." In *proceedings of the 6th CSCE International Construction Specialty Conference*, 1-9.
- Safapour, E., and Kermanshachi, S. (2020). "Identification and categorization of factors affecting duration of post-disaster reconstruction of interdependent transportation systems." In *Construction Research Congress 2020*, 1290-1299. Reston, VA: American Society of Civil Engineers.
- Samaraweera, A., Jayasanka, T. A. D. K., Gamage, V., Rameezdeen, R., and Alankarage, S. (2022). *Conceptual Framework for Suicide Prevention Process in Construction*.
- Sawicki, M., and Szóstak, M. (2020). "Impact of alcohol on occupational health and safety in the construction industry at workplaces with scaffoldings.: *Applied Sciences*, 10(19), 6690.
- Shaw, W. S., Roelofs, C., and Punnett, L. (2020). "Work environment factors and prevention of opioid-related deaths." *American journal of public health*, 110(8), 1235-1241.
- Sun, C., Hon, C. K., Way, K. A., Jimmieson, N. L., and Xia, B. (2022). "The relationship between psychosocial hazards and mental health in the construction industry: A meta-analysis." *Safety science*, 145, 105485.
- Sussell, A. (2023). "Suicide Rates by Industry and Occupation—National Vital Statistics System, United States, 2021." *MMWR. Morbidity and Mortality Weekly Report*, 72.
- Tijani, B., Falana, J. N., Jin, X., and Osei-Kyei, R. (2023). "Suicide in the construction industry: Literature review." *International Journal of Construction Management*, 23(10), 1684-1693.

Wang, D., and Liu, Y. (2021). "The effect of political skill on relationship quality in construction projects: The mediating effect of cooperative conflict management styles." *Project Management Journal*, 52(6), 563-576.

Wu, G., Hu, Z., and Zheng, J. (2019). "Role stress, job burnout, and job performance in construction project managers: the moderating role of career calling." *International journal of environmental research and public health*, 16(13).

Assessing Driver Behavior during Bicyclist Overtaking: An Integrated Simulator Study

Parisa Masoumi¹; Anam Ardestiri, Ph.D.²; Eazaz Sadeghvaziri, Ph.D.³;
and Mansoureh Jeihani, Ph.D.⁴

¹Ph.D. Student, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: pamas2@morgan.edu

²Traffic and ITS Engineer, Mead & Hunt, Inc., Columbia, MD.
Email: anam.ardeshiri@meadhunt.com

³Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Macon, GA. Email: sadeghvaziri_e@mercer.edu

⁴Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD.
Email: mansoureh.jeihani@morgan.edu

ABSTRACT

This study used a driving simulator to investigate driver behavior near bike lanes. The interaction between cars and bicyclists during passing maneuvers is a key factor in bike-related accidents. The study tested two scenarios: one with a road containing a bike lane and one without. Twenty-nine participants had two driving sessions, each matching one of the scenarios. The study used car and bike simulators allowing two participants to drive simultaneously on the same road network and interact live. Driver performance was evaluated by analyzing their interactions during 30-s overtaking maneuvers. Statistical analysis assessed the bike lane's impact on driver behavior. The presence of a bicyclist led to significant changes in driver lateral movement and speed. When more space was available from the bike lane, drivers felt safer and increased speed to overtake bicycles. In conclusion, this study revealed that most drivers gradually approached bicyclists and accelerated to pass them quickly.

Keywords: Bicycle Lane, Bicyclist Safety, Overtaking, Lateral Movement, Speed, Driving Simulator

INTRODUCTION

The Federal Highway Administration (FHWA) states that “Complete Street (CS) is safe, and feels safe, for all users. Complete Streets benefit everyone who uses them, including drivers, pedestrians, bicyclists, public transit riders, children, the elderly, people with disabilities, and freight vehicles” (*Complete Streets / FHWA*, n.d.). There are many factors determining the practicality and feasibility of the CS approach in an urban setting, such as travel demand and pattern, land-use, topology, availability of shared modes/vehicles, transit system reliability, roadway and sidewalk conditions, and how streets are designed and constructed. CS is more about inclusive multimodal planning rather than merely adding bicycle lanes, bus lanes, or exclusive lanes for one mode of transportation. CSs aim to ensure that both motorized and non-motorized road users have a variety of modes of transportation accessible to use in a safe and effective manner (Anderson et al., 2015). Many advocates place a strong emphasis on the implementation of equitable transportation policies to encourage the use of various green modes of transportation, such as walking, cycling, and other micro-mobility options, in accordance with

public health promotion programs. Providing a safe road environment can encourage more people to walk and ride bicycles (Brown et al., 2015; Reynolds et al., 2009).

In the last two decades, cycling has become very popular in urban environments, especially among younger travelers. Bicycle commute in the U.S. raised by 51% from 2000 to 2016, (Mcleod, 2014). As the need for mobility and the number of non-motorized users continues to grow, concerns about the frequency and severity of crashes are becoming increasingly significant. Despite all the long- and short-term benefits of this mode, bicyclists remain vulnerable road users (VRUs) since there aren't enough crushable materials to shield them in the event of a crash. According to the World Health Organization (WHO), VRUs are involved in 50% of all traffic fatalities each year (World Health Organization, 2009). Indeed, interacting with motorized vehicles remains the most hazardous aspect of bicycle travel (Chaurand & Delhomme, 2013; Frings et al., 2014; Hamann & Peek-Asa, 2013; Schepers et al., 2014). According to a 2013 research report by the WHO, bicyclists and pedestrians are involved in 27% of all fatal crashes, and 4,500 pedestrians and more than 700 bicyclists are killed in crashes involving motor vehicles every year in the U.S. (Organization, 2013). National Highway Traffic Safety Administration's 2017 report revealed that urban areas account for 70% of pedal cyclist fatalities while rural regions account for 30% (Administration, 2017). Additionally, 61% of these fatalities happened at non-intersections. According to these crash statistics, interactions between drivers and bicyclists in metropolitan areas without intersections have resulted in a significant number of bicycle crashes.

Numerous studies have used naturalistic experiments to assess how various elements affect the passing behavior of motor vehicles and bicyclists during overtaking maneuvers (Fournier et al., 2020). Most of these studies used naturalistic experiments to explain overtaking behavior. An example of such findings is from Shackel and Parkin's study (Shackel & Parkin, 2014), where the authors found that platoon formations of oncoming vehicles resulted in shorter passing distances. Multiple studies have highlighted the correlation between motor vehicle type and lateral position as another significant factor. Walker's research (Walker, 2007) revealed that professional drivers operating large vehicles typically allow shorter lateral movement distances. Additionally, De Ceunynck et al. (De Ceunynck et al., 2017) noted that close overtaking of bicycles by buses is more frequently observed in bus lanes.

The choice of passing distance and time has direct implications on the safety and comfort of bicyclists. Insufficient passing distances can increase the risk of crashes, while excessively long passing times may lead to frustration and impatience among both drivers and bicyclists. A few studies have been conducted on examining the interaction between vehicle and bicyclist during passing maneuvers. The study conducted by Mehta et al. found that drivers on arterial routes will often change their lanes in order to keep a safe distance from bicyclists (Mehta et al., 2015). Love et al. conducted a study to evaluate compliance with the three-foot law, measuring the distance between overtaking motor vehicles and bicyclists. The authors found that 3 ft or less of passing clearance was common in standard lanes and lanes with a shared lane marking, but not in streets with 5 ft wide bicycle lanes (Love et al., 2012). Chuang et al. incorporated road-related factors, such as clear line separation and avoidance of road hazards, which led to greater lateral distance between motorists and bicyclists. In addition to vehicle speed, the authors emphasized the significance of passing time as a crucial factor. By utilizing an instrumented bicycle, the authors discovered that the mean lateral distance between vehicles and bicycles was notably smaller when the passing time ranged from 0.1 to 0.4 seconds compared to passing times longer than 0.4 seconds. These findings suggest that passing time, along with relative speed, plays a significant role in shaping overtaking behavior (Chuang et al., 2013).

In this study, drivers' performance will be as they ride alongside a bicycle by examining their interactions prior to, during, and after the overtaking moment. By focusing on passing distance and time, this study aims to provide valuable insights into the factors influencing drivers' decisions and behaviors during overtaking maneuvers, contributing to the development of effective strategies and guidelines for promoting safer interactions between motorists and bicyclists on the road.

METHODS

Several studies have suggested that utilizing a driving simulator is a helpful tool to analyze driver behavior under various traffic circumstances and various roadway layouts (Bella & Silvestri, 2015, 2016). A car and a bicycle simulator at Morgan State University's Safety and Behavioral Analysis Center (SABA) center were used in this study to evaluate driver's behavior while interacting with bicyclists. The driving simulator was equipped with an open cabin containing a steering wheel, gas pedal, brake pedal, rear and side mirrors, simulated engine and traffic noise, and three 40-inch LCD screens positioned in front of the cabin to display simulated roadway network that could mimic real-world conditions (Figure 1 (a)). The bike simulator (Figure 1 (b)) had one LCD screen positioned in front of the bicycle. These simulators were integrated, allowing any two participants to ride and drive simultaneously on the same virtual road network. The participants were unaware that they were traveling along the same route. Second-by-second lateral position and speed data of the subject vehicles were recorded to analyze the driver behavior during the overtaking period between the driver and the bicyclist.

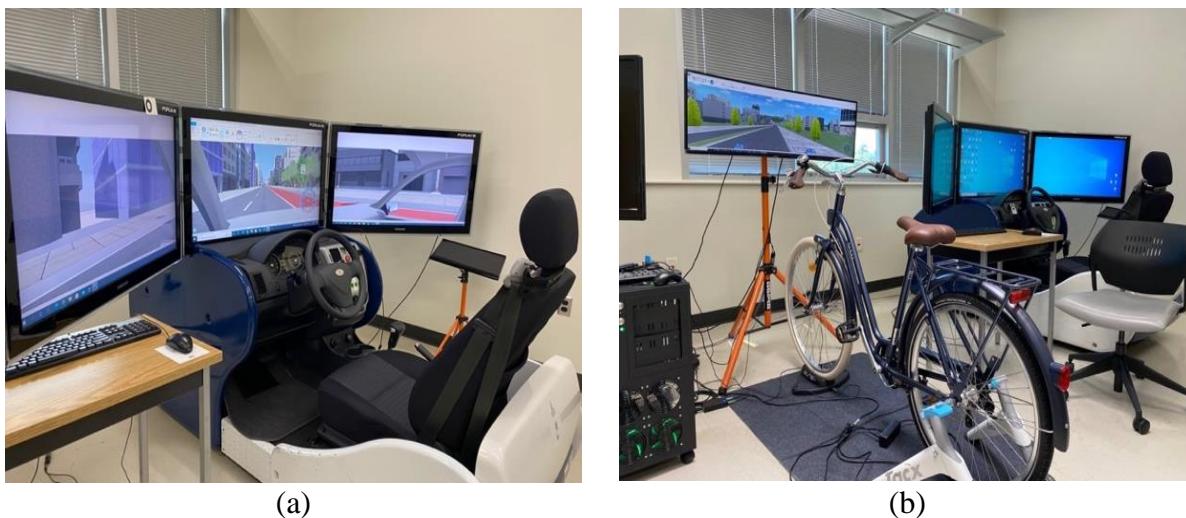


Figure 1. (a) Views from car simulator and (b) bike simulator at MSU

The SABA Lab used local advertising and emails to recruit people in the Baltimore metro area to take part in a paid study. A total of 29 participants (15 female and 14 male) of various age groups and diverse socio-demographic backgrounds participated in this study. All participants were required to provide proof of a valid U.S. driver's license or learner's permit. The Driving Simulator's virtual environment and the study's research objectives were explained to participants before the experiments started. They were instructed to drive the car as they would in the real world. Two different networks were designed in the simulators to investigate the

behavioral responses of drivers in the presence and absence of bicycle lane configurations. This approach enabled the analysis and comparison of driver's behavior under different conditions, allowing for the examination of potential variations and patterns associated with the presence or absence of bicycle lanes. The road network consisted of three intersections. This study consisted of two different scenarios, including a scenario without bicycle lane (Scenario 1) and a scenario with a bicycle lane on the right side of the road (Scenario 2). The overall traffic flow is conserved 50 kmph. The road's length was the same in both scenarios—8,530-foot one-way road. Road lane configurations and descriptions of scenarios are shown as initial snapshots in Figure 2.



Figure 2. (a) Scenario without bicycle lane (Scenario 1) (b) Scenario with bicycle lane design (Scenario 2)

TABLE 1. Description of Scenario 1 and Scenario 2

Scenario	Number of Lanes	Car Lane	Bike Lane	Driver's Speed Limit (km/h)
Scenario 1 (Without bike lane-Mild Traffic)	3 lanes	3 lanes (11 feet each)	-	50
Scenario 2 (With bike lane-Mild Traffic)	3 lanes	2 lanes (11 feet each)	1 lane (6 feet)	50

ANALYSIS

In an aggregate analysis, the average lateral position and speed profiles of all participants were plotted where the “integrated” bicycle appeared in the network. For scenario 1 these variables were recorded at the same location along the road as scenario 2. The area of interaction between the driver and the bicyclist was defined as the time period starting 15 seconds prior to the driver passing the bicyclist and continuing for 15 seconds after the passing point. The rationale behind selecting a 15-second timeframe was to capture a comprehensive picture of the behaviors and dynamics leading up to the passing maneuver and its immediate aftermath. Therefore, by starting the area of interaction 15 seconds before the passing point, the study aims to include the relevant pre-passing behaviors and factors that may influence the driver's decision

to initiate the passing maneuver. Furthermore, extending the region of interaction to 15 seconds post-passing allows for the examination of any potential consequences or reactions that occur immediately after the passing point. This includes observing how the drivers adjust their positions, speed, and behaviors following the completion of the overtaking maneuver.

By considering both lateral movement and speed data during 30 seconds of the passing point, we can gain a comprehensive understanding of the dynamics of overtaking maneuvers. This detailed analysis can provide valuable insight into the relationship between lateral movement, speed, and safe overtaking practices, contributing to a deeper understanding of the factors influencing safe passing behavior.

Descriptive Analysis

The study collected descriptive statistics based on the data from pre-survey questionnaires, which provided information about the socio-demographic characteristics of the participants (Table 2). Only 33% of the participants were familiar with the concept of CS.

TABLE 2. Participant Demographic

Variables		Percent
Gender	Male	48.3
	Female	51.7
Age	18-25	44.8
	26-35	13.8
	36-45	20.7
	46-55	10.3
	56-65	6.9
	65+	3.5

In a post-survey questionnaire, when asked about their experience with unprotected bike lane, nearly half of participants found it distracting. Ninety percent of participants found CS a safe approach for drivers and bicyclists after the experiment.

Aggregate Data Analysis

To study the driver's performance during the whole overtaking maneuver of the bicycle in scenario 2, the average lateral position of all participants was evaluated 15 seconds prior and 15 seconds past the passing point. The average lateral position profiles of all participants in each scenario are presented in Figure 3. The black dashes show where the user vehicle passed the bicycle. The lateral movements for scenario 1 were recorded at the same spot as scenario 2. As mentioned in the methodology section, Scenario 1, has three car lanes, while scenario 2 has two car lanes and one bicycle lane. To have consistency between the two graphs, the graph coordinates from scenario 1 originate from the right curbside of the road, whereas in scenario 2 they originate from the left side of the bicycle lane. Based on the average lateral locations of all participants, it was found that drivers exhibited a tendency to drift away from the bicycle while overtaking it. This behavior was likely a means to ensure a safe distance between the vehicle and the bicyclist.

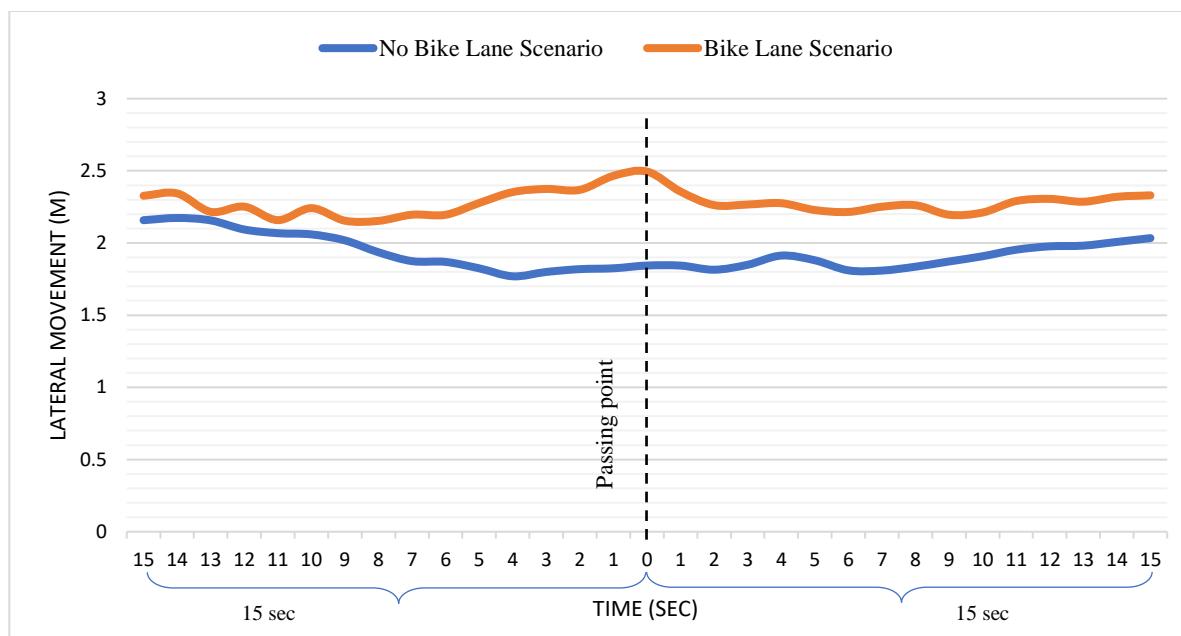


Figure 3. Average lateral movement 15-sec. prior & 15-sec. post passing point

The average speed of all participants in **No Bike Lane and Bike Lane Scenarios** were calculated separately and shown in Figure 4. In **No Bike Lane Scenario**, the average speed of the drivers 15 seconds before overtaking the bicyclist was 36km/hr, which was higher than the average speed at the same point in the Bike Lane Scenario. The average speed in **No Bike Lane Scenario** then increased to a maximum average speed of 55km/h just after the passing point before dropping down to 37 km/hr. This was significantly different from the results in **Bike Lane Scenario**, which possessed a similar trend line but a higher average speed at the passing point (62 km/h). With the presence of a dedicated bicycle lane, drivers' average speed 15 seconds before overtaking the bicyclist was 30 km/hr, increasing up to 62km/hr at the passing point and finally dropping to 42km/hr roughly 15 seconds after overtaking the bicyclist. Moreover, at about 9 seconds prior to the passing point, the average speed in both scenarios was almost the same. Similarly, the drivers' average speed for was nearly the same for the last 7 seconds of both scenarios.

As a result, the drivers were inclined to accelerate when passing the bicyclist. The drivers increased their speed to pass the bicycle as quickly as possible, most likely to minimize the interaction with the bicyclist. The drivers reached their speed normal speed within 10 sec after passing the bicycle. The speed and the lateral movement graphs together demonstrated that when drivers were laterally far from the bicycle lane, they felt secure enough to accelerate and pass the bicycle.

Statistical Analysis

t-test

Many studies use statistical analysis to develop policies to improve traffic safety, investigate and forecast travel behavior, and pinpoint deficiencies in transportation policy (Javid &

Sadeghvaziri, 2023b, 2023a; Mehryaar & Bandelt, 2022; Sadeghvaziri et al., n.d.). Statistical analysis was used in this study to assess how different road designs influence driving behavior. The analysis revealed a noteworthy distinction in lateral movements between scenario 2 (Mean = 2.28, SD = 0.09) and scenario 1 (Mean = 1.93, SD = 0.12); $t(30) = 12.368$, $p = 0.000$. To compare the average speed of drivers in two different scenarios, a paired samples t-test was conducted. The results revealed a statistically significant difference in speed between scenario 2 (mean = 50.64, SD = 9.049) and scenario 1 (Mean = 48.68, SD = 5.682). The t-test yielded a $t(30) = 2.649$ and a p-value of 0.0127, indicating that the observed difference is statistically significant due to the presence of a bicycle lane.

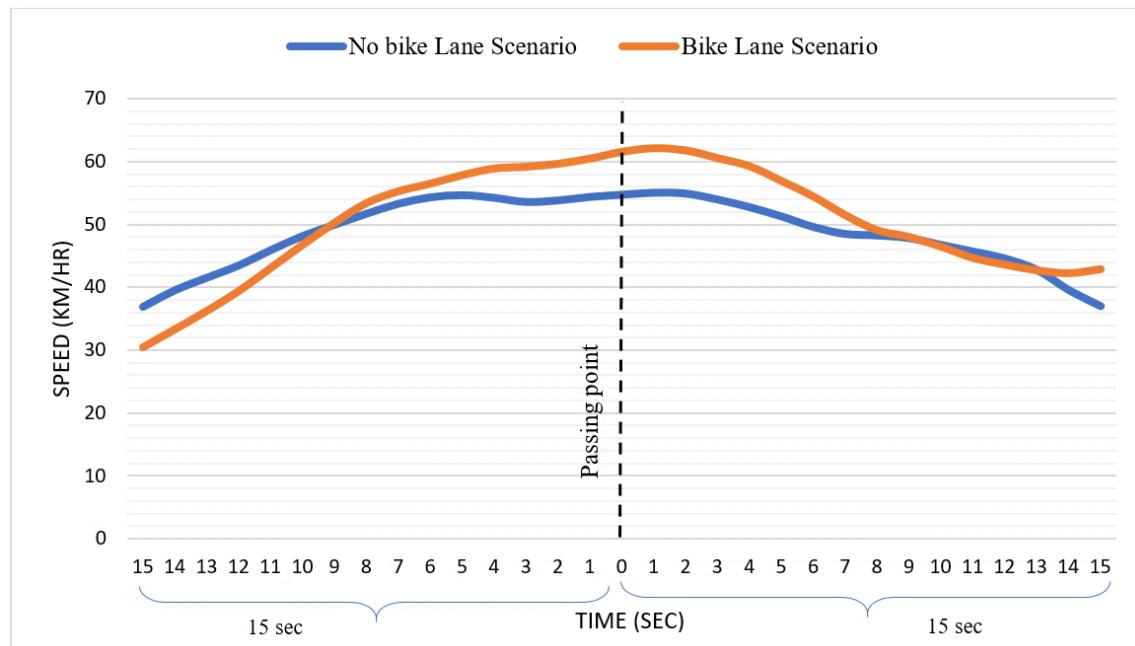


Figure 4: Average participant's speed 15-sec prior & 15-sec post passing-point

Spearman's Rank Correlation (speed and lateral movement)

The Spearman's rank correlation is a statistical measure used to assess the monotonic relationship between two scenarios. Unlike the Pearson correlation coefficient, which measures linear association, the Spearman's rank correlation focuses on the ordinal relationship between the variables. It is particularly useful when dealing with non-linear relationships or when the data is in the form of ranks or categories. In this case, the goal of using Spearman's rank correlation was to examine the association between the average lateral movement of participants in scenario 1 and scenario 2. By calculating the correlation coefficient, the authors aimed to understand whether there was a consistent monotonic relationship between the lateral movement in the two scenarios, without assuming a specific functional form or linearity. The Spearman's rank correlation coefficient of -0.225 suggests that the average lateral movement of participants in scenarios 1 and 2 does not have a strong linear relationship. A correlation coefficient close to zero or with a small magnitude, such as -0.225, indicates that there is little to no linear association between the variables. This means that changes in one variable (e.g., average lateral movement in scenario 2) are not consistently and directly related to changes in the other variable

(e.g., average lateral movement in scenario 1) in a linear fashion. While the Spearman's rank correlation coefficient does capture the presence of any monotonic relationship, it does not provide evidence for a specific linear relationship. Therefore, based on the coefficient of -0.225, it can be concluded that the average lateral movement in scenarios 1 and 2 are not strongly linearly related (Figure 5 (a)).

The Spearman's rank correlation was performed on the average speed of drivers for both scenarios, and the results show a coefficient of 0.9589, which indicates a strong positive monotonic relationship between the scenarios. This means that as one variable increases, the other also increases, and vice versa. It is important to note that the high Spearman's correlation coefficient of 0.9589 suggests a strong association between the variables, but it does not necessarily imply a strictly linear relationship. While a Pearson's correlation coefficient close to 1 does indicate a strong linear association, it is worth emphasizing that the Spearman's rank correlation coefficient focuses on the monotonic relationship rather than linearity (Figure 5 (b)).

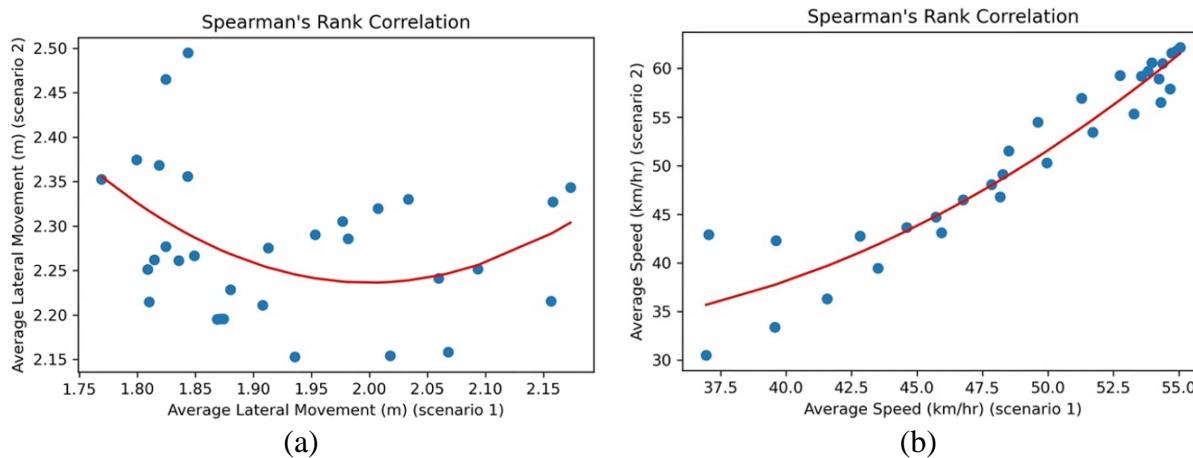


Figure 5: Spearman's Rank Correlation: relationship between (a) average lateral movement and (b) average speed for two scenarios

According to the t-test results, there is a significant difference in the average speeds between the two scenarios during the 30-second period in which drivers passed the bicycle. However, based on the Spearman's rank correlation analysis, the variables exhibit a monotonic relationship. This implies that while the speeds in the two scenarios differ, they still follow a similar pattern of increase and decrease, most likely due to similar roadway geometry, despite the bike lane in scenario 2.

Pearson Correlation (speed and lateral movement)

The Pearson correlation coefficient is a measure of the linear relationship between two variables. It quantifies the strength and direction of the linear association between the variables. The coefficient ranges from -1 to 1, where -1 represents a perfect negative linear relationship, 1 represents a perfect positive linear relationship, and 0 indicates no linear relationship between the variables. The Pearson correlation was performed on the average lateral movement of drivers for both scenarios (Figure 6 (a)). The results showed a coefficient of -0.2030 which indicated a weak negative linear relationship between the scenarios. This means that as the lateral movement

increases for one scenario, the lateral movement tends to slightly decrease for the other scenario, and vice versa. However, it is important to note that the coefficient is close to 0, suggesting a weak association. Therefore, the correlation between lateral movement in the two scenarios is not significant, indicating that there is no strong linear relationship between them.

The average speed in the two scenarios was found to be significantly different according to the t-test conducted during the 30-second period of passing the bicycle. However, the Pearson correlation coefficient of 0.9459 reveals a strong positive linear relationship between the average speed in the two scenarios. Therefore, despite the initial difference in speeds, the two scenarios shared a pattern in their average speed changes over time (Figure 6 (b)).

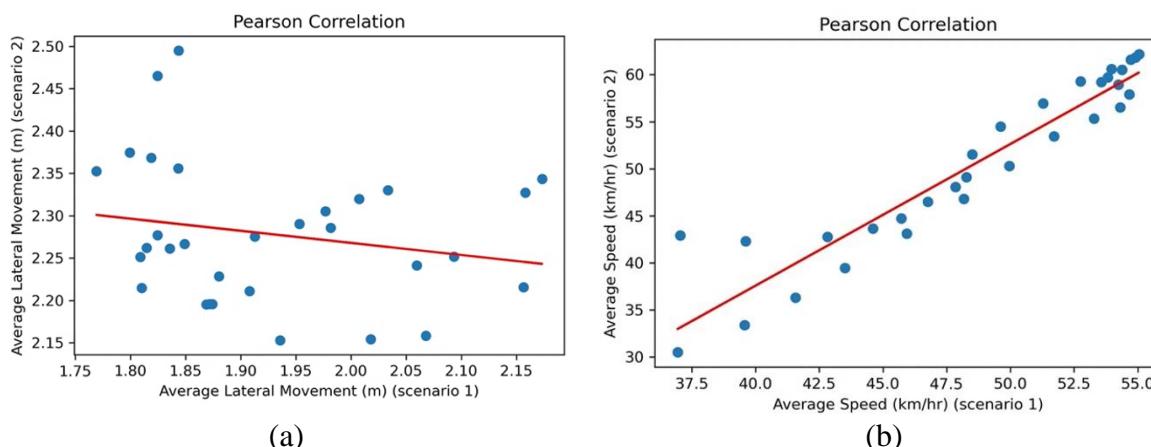


Figure 6. Pearson Correlation: relationship between (a) average lateral movement and (b) average speed for two scenarios

CONCLUSIONS

This paper presents the findings of a study on driver's behavior when car drivers pass bicyclists using a car and a bike simulator. The study's objective was to assess how a bicycle lane affected the lateral movement and speed of drivers as they passed bicyclists. The results of this study provide insights to better understand drivers' decision-making process regarding their lateral movement and speeds when overtaking bicyclists. The authors designed the study based on the assumption that drivers' speeds and lateral positions are influenced by the road's design and the presence of bicyclists. This study therefore considered two different scenarios to analyze driver's speed and lateral movement: a scenario without a bicycle lane (scenario 1), and a scenario with a bicycle lane (scenario 2). Overall, scenario 2 demonstrated a higher average speed in the vicinity of bicyclist compared to scenario 1. However, the increase in speed, when compared to scenario 1, shows that when drivers have more space between them and the bicycle lane, they feel safe enough to overtake the bicycle with minor acceleration. It was found that most drivers tend to gradually approach bicyclists and accelerate to pass them swiftly. Statistical analyses revealed the significant differences between both scenarios in the speed and the lateral movement of drivers.

Possible future research could focus on evaluating the efficacy of painted bicycle lanes, in combination with various types of barriers such as curb separators, bollards, planters or greenery, plastic or rubber barriers, median or island separation, delineator posts, and raised cycle tracks to

alert drivers. Additionally, investigating how drivers' speed and lateral movements are influenced by a more consistent density of bicyclists can provide valuable insights for transportation planning and safety. These insights promote a deeper understanding of driver behavior, identification of risk factors, safer road infrastructure design, driver education and training, and active transportation. By gaining knowledge about driver responses to bicyclist density, informed policies, infrastructure improvements, and educational initiatives can be developed to create safer and more sustainable transportation systems. Furthermore, expanding the scope of this study by analyzing larger datasets and considering different scenarios would significantly enhance our understanding of the topic. By including a larger dataset, we can capture a broader range of variables and factors that may influence the phenomenon being studied. This would allow for more robust statistical analyses and provide a more representative picture of the relationships and trends involved. Additionally, considering different scenarios, such as varying traffic conditions or road configurations, would enable us to identify patterns and variations in driver behavior across different contexts. This comprehensive understanding would not only deepen our knowledge of the specific phenomenon under investigation but also provide valuable insights into the generalizability and applicability of our findings. Ultimately, a more comprehensive understanding of the topic would support evidence-based decision-making, inform policy development, and contribute to the development of effective interventions and strategies in the field of transportation and road safety. Examining the impact of pedestrian movements on speed and lateral position may also yield unique findings for novel research.

ACKNOWLEDGMENTS

This study was supported by the Sustainable Mobility & Accessibility Regional Transportation Equity Research (SMARTER) Center, a Regional University Transportation Center of the U.S. Department of Transportation University Transportation Centers Program at Morgan State University.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Parisa Masoumi, Anam Ardestiri, Eazaz Sadeghvaziri, Mansoureh Jeihani; data collection: Parisa Masoumi, Anam Ardestiri, Eazaz Sadeghvaziri, Mansoureh Jeihani; analysis and interpretation of results: Parisa Masoumi, Anam Ardestiri, Eazaz Sadeghvaziri, Mansoureh Jeihani; draft manuscript preparation: Parisa Masoumi, Anam Ardestiri, Eazaz Sadeghvaziri, Mansoureh Jeihani. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- National Highway Traffic Safety Administration. (2017). *Traffic Safety Facts-2015 Data-Bicyclists and Other Cyclists*.
- Anderson, G., Searfoss, L., Cox, A., Schilling, E., Seskin, S., and Zimmerman, C. (2015). Safer streets, stronger economies: Complete streets project outcomes from across the United States. *Institute of Transportation Engineers. ITE Journal*, 85(6), 29.
- Bella, F., and Silvestri, M. (2015). Effects of safety measures on driver's speed behavior at pedestrian crossings. *Accident Analysis & Prevention*, 83, 111–124.

- Bella, F., and Silvestri, M. (2016). Driver's braking behavior approaching pedestrian crossings: A parametric duration model of the speed reduction times. *Journal of Advanced Transportation*, 50(4), 630–646.
- Brown, B. B., Werner, C. M., Tribby, C. P., Miller, H. J., and Smith, K. R. (2015). Transit use, physical activity, and body mass index changes: Objective measures associated with complete street light-rail construction. *American Journal of Public Health*, 105(7), 1468–1474.
- Chaurand, N., and Delhomme, P. (2013). Cyclists and drivers in road interactions: A comparison of perceived crash risk. *Accident Analysis & Prevention*, 50, 1176–1184.
- Chuang, K.-H., Hsu, C.-C., Lai, C.-H., Doong, J.-L., and Jeng, M.-C. (2013). The use of a quasi-naturalistic riding method to investigate bicyclists' behaviors when motorists pass. *Accident Analysis & Prevention*, 56, 32–41.
- Complete Streets | FHWA. (n.d.). Retrieved June 18, 2023, from <https://highways.dot.gov/complete-streets>.
- De Ceunynck, T., Dorleman, B., Daniels, S., Laureshyn, A., Brijs, T., Hermans, E., and Wets, G. (2017). Sharing is (s)caring? Interactions between buses and bicyclists on bus lanes shared with bicyclists. *Transportation Research Part F: Traffic Psychology and Behaviour*, 46, 301–315.
- Fournier, N., Bakhtiari, S., Valluru, K. D., Campbell, N., Christofa, E., Roberts, S., and Knodler, M., Jr. (2020). Accounting for drivers' bicycling frequency and familiarity with bicycle infrastructure treatments when evaluating safety. *Accident Analysis & Prevention*, 137, 105410.
- Frings, D., Parkin, J., and Ridley, A. M. (2014). The effects of cycle lanes, vehicle to kerb distance and vehicle type on cyclists' attention allocation during junction negotiation. *Accident Analysis & Prevention*, 72, 411–421.
- Hamann, C., and Peek-Asa, C. (2013). On-road bicycle facilities and bicycle crashes in Iowa, 2007–2010. *Accident Analysis & Prevention*, 56, 103–109.
- Javid, R., and Sadeghvaziri, E. (2023a). Assessing Gender Gaps in Citi Bike Ridership in New York City (SSRN Scholarly Paper 4358660).
- Javid, R., and Sadeghvaziri, E. (2023b). Equity Analysis of Bikeshare Access: A Case Study of New York City. *Findings*.
- Love, D. C., Bread, A., Burns, S., Margulies, J., Romano, M., and Lawrence, R. (2012). Is the three-foot bicycle passing law working in Baltimore, Maryland? *Accident Analysis & Prevention*, 48, 451–456.
- Mcleod, K. (2014). Where We Ride: Analysis of Bicycle Commuting in American Cities. <https://www.semanticscholar.org/paper/Where-We-Ride%3A-Analysis-of-Bicycle-Commuting-in-Mcleod/d9c3f7cffb1dbc3267725c6692d31c4c1187b62b>.
- Mehryaar, E., and Bandelt, M. J. (2022). Geological Modeling Along Tunnel Projects Using Machine Learning Techniques. *ARMA/DGS/SEG International Geomechanics Symposium*.
- Mehta, K., Mehran, B., and Hellinga, B. (2015). Evaluation of the passing behavior of motorized vehicles when overtaking bicycles on urban arterial roadways. *Transportation Research Record*, 2520(1), 8–17.
- World Health Organization. (2013). *Global status report on road safety 2013: Supporting a decade of action: Summary*. World Health Organization.
- Reynolds, C. C., Harris, M. A., Teschke, K., Cripton, P. A., and Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: A review of the literature. *Environmental Health*, 8, 1–19.

- Sadeghvaziri, E., Javid, R., and Jeihani, M. (n.d.). *Investigating Walking and Biking Activities Among Low-Income African Americans*.
- Schepers, P., Hagenzieker, M., Methorst, R., van Wee, B., and Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis & Prevention*, 62, 331–340.
- Shackel, S. C., and Parkin, J. (2014). Influence of road markings, lane widths and driver behaviour on proximity and speed of vehicles overtaking cyclists. *Accident Analysis & Prevention*, 73, 100–108.
- Walker, I. (2007). Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. *Accident Analysis & Prevention*, 39(2), 417–425.
- World Health Organization. (2009). Global status report on road safety: Time for action. Rapport de Situation Sur La Sécurité Routière Dans Le Monde : Il Est Temps d'agir, 287.

Assessment of Pedestrian and Bicycle Safety at Commercial Driveways along Major Corridors

Cong Chen, Ph.D., P.E., M.ASCE¹; Elzbieta Balkowska-Jelinska²; Kristine Williams³;
Pei-Sung Lin, Ph.D., P.E.⁴; Shubhankar Chintamani Shindgikar⁵; Tia Boyd, AICP⁶;
and Gina Bonyani⁷

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: congchen1@usf.edu

²Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: elzbieta@usf.edu

³Retired; formerly, Univ. of South Florida; Principal, KMW Associates LLC, Garden, MI.
Email: KMWAssociates@outlook.com

⁴Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: lin@usf.edu

⁵Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: shindgikar@usf.edu

⁶Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: tiaboyd@usf.edu

⁷Florida Dept. of Transportation, Tallahassee, FL. Email: Gina.Bonyani@dot.state.fl.us

ABSTRACT

This research evaluated the impacts of commercial driveway and connecting street features on pedestrian and bicycle safety through crash frequency and severity analyses. Data for driveway-related crashes involving pedestrians or bicycles in Florida over a five-year period were analyzed. The study also reviewed the characteristics of commercial driveways along selected major corridors with high pedestrian and bicycle crash frequencies, which were included in the analysis. The results revealed several significant factors affecting pedestrian and bicycle crash occurrence at commercial driveways, including the number of lanes on connecting streets, driveway design features, median opening type, traffic control device type, and painted bike lanes. Significant factors affecting crash severity were also identified, including shoulder type, alcohol or drug involvement, driveway number of lanes, bike lane type, driveway throat length, and connecting street annual average daily traffic (AADT) in the crash year. These findings provide important insights on addressing pedestrian and bicycle safety issues at commercial driveways, which could be used as a reference for future access management guideline improvement.

INTRODUCTION

Driveways are crucial in road networks linking the road and adjacent properties. Different road users, such as motorists, pedestrians, bicyclists, and others, engage in interactions as they enter, exit, or pass by driveways. Due to the intricate interactions among various road users, driveways primarily contribute to traffic crashes and congestion (Williamson and Zhou, 2014). The American Association of State Highway and Transportation Officials (AASHTO) considers driveways as low-volume intersections (AASHTO, 2011) where many road users interact and

create conflict points. The conflict points are prevalent for commercial driveways, with higher road user flow than residential driveways. Hence, commercial driveways can significantly impact the flow and safety of traffic.

The majority of accidents occurring at commercial driveways involve pedestrians and bicyclists. This is largely due to the urban setting of these driveways, where pedestrian and bicyclist traffic is high, especially around commercial hubs such as shopping malls, dining establishments, and recreational areas. There is little data on pedestrian and bicyclist crashes and fatalities related to commercial driveways. However, according to the National Highway Traffic Safety Administration (NHTSA, 2023a), 85% of fatal bicyclist crashes occurred in urban areas in 2021. This represents a substantial increase of 16% compared to 2011.

In 2021, there were 7,388 pedestrian fatalities in the United States, including 817 that occurred in Florida. This marks a significant rise from the figures of 2011, which recorded 4,457 pedestrian deaths nationwide, including 490 in Florida. In 2021, the fatality rate per 100,000 population was 2.23 nationally and 3.75 in Florida, and in 2011, the rate was 1.43 nationally and 2.57 in Florida (NHTSA 2023b). In 2021, there were 966 bicyclist fatalities in the United States, including 197 that occurred in Florida compared to 2011, 677 fatalities nationally and 125 in Florida. In 2021, the fatality rate per 100,000 population was 0.29 nationally and 0.90 in Florida, and in 2011, the rate was 0.22 nationally and 0.66 in Florida (NHTSA, 2023a). For both pedestrians and bicyclists, the fatality rates in Florida consistently surpass national averages as shown in Figure 1. The red line in Figure 1 indicates the percentage of bicyclist fatalities in urban areas in the United States.

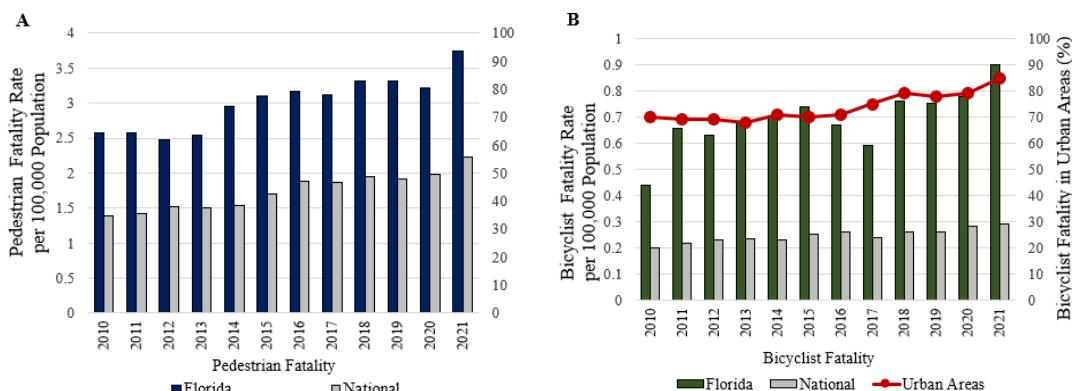


Figure 1. Fatality rates per 100,000 population in Florida and the U.S., for (A): pedestrian, and (B): bicyclist.

The safety of bicyclists and pedestrians where bike lanes and sidewalks intersect with driveways relies on visibility, driver anticipation, and, for bicyclists, their level of skill. (Williams et al., 2014). A common safety problem at driveways is the tendency for drivers turning right out of a driveway to monitor oncoming traffic on the left without looking for bicyclists and pedestrians to their right. Crashes involving bicyclists and pedestrians become less common when bicyclist and pedestrian travel patterns are consistent with motorists' expectations (NASEM, 2021). For example, separated bike lanes with a one-way configuration can reduce these conflicts.

Visibility issues at driveways (Figure 2) and vehicles encroaching on sidewalks upon exit (Figure 3) pose significant hazards to pedestrians and cyclists. A landscaped buffer between the

curb and sidewalk enhances visibility for pedestrians entering or leaving vehicles (Dixon et al., 2009). This buffer also provides space for vehicles turning into driveways to stop before reaching the sidewalk, safely allowing the vehicle to wait for pedestrian activity without protruding into the active travel lane (Gattis et al., 2010). Integrating ADA-compliant sidewalk cross-slopes with the driveway's vertical slope is vital at driveway-sidewalk intersections, defining pedestrian space and enhancing accessibility.



Figure 2. Limited pedestrian sight distance at driveway Source: Gattis et al., (2010).



Figure 3. Exiting vehicle encroaching on active sidewalk Source: Williams et al., (2014).

There is very little research on the commercial driveway safety evaluation of bicyclists and pedestrians. The authors feel this paper addresses this research gap in the literature and provides insight into the pedestrian and bicycle crash frequency and their causes and contributing factors in this research.

METHODOLOGY

In this study, an assessment of pedestrian and bicyclist safety at commercial driveways was conducted through the following methodology: (1) Available geospatial data collection and identification of high crash corridors; (2) Driveway characteristics review and data collection; (3) Development of two regression models.

Available geospatial data collection and identification of high crash corridors

According to the FDOT Crash Analysis Reporting User Manual (FDOT, 2009), a crash is considered “driveway access related” if the crash is in a driveway access or is influenced by traffic entering or exiting a driveway access. In this research, the 2015-2019 driveway-related crash data, roadway characteristics and traffic volume data, and land use data were collected from a variety of sources (FDEP, 2017; FDOT, 2021a; FDOT, 2021b). These data were imported, integrated, and processed in an ArcGIS environment assuring the same geographic coordinate system, which provided the initial foundation for the site screening process of commercial driveways located along corridors. The objective of this step was to identify the corridors and driveway sites with the highest numbers of pedestrian and bicycle commercial

driveway-related crashes. Candidate commercial driveway sites for further data collection and analysis were selected through a screening process by (1) identifying commercial area based on land use data; (2) ranking roadway segments based on total number of pedestrian and bicycle crashes; and (3) selecting representative corridors and driveways based on roadway/driveway characteristics and spatial distribution.

Driveway Characteristics Review and Data Collection

The driveway characteristics and non-motorist facility data (i.e., presence of sidewalk, crosswalk, pedestrian/bicycle signal, etc.) of each driveway site, as listed in Table 1, were manually reviewed and recorded. The review was conducted using satellite layer and street view images in Google Maps. The characteristics of driveways and adjacent roads were documented by reviewing the street view images as someone leaving the driveway and moving along the adjacent roads. A total of 9,889 commercial driveways along corridors were reviewed and included in the analysis.

Table 1. Reviewed Attributes for Driveways, Adjacent Road, and Non-Motorist Facility

Attributes	
Roadway	Non-motorist facility
Functional Class (Defined by FDOT)	Bike Lane Type (No Bike Lane, Conventional, Other Bike Lane Types (e.g., buffered bike, keyhole lane, etc.))
Access Class (Defined by FDOT)	Painted Bike Lane (N/A, No, Yes)
Context Class (Defined by FDOT)	Sidewalk (Present, Absent)
Number of Lanes on Connecting Street (1, 2, 3, 4 or more)	Marked Crosswalk (Present, Absent)
Posted Speed (<35mph, 40-45 mph, ≥45 mph)	Pedestrian Crossing Signal Availability (Present, Absent)
Annual Average Daily Traffic (AADT)	Pedestrian Refuge Island (Present, Absent)
Median Type (Undivided or Painted Median, Non-traversable Median, Two-way Left Turn Lane)	
Median Opening Type (No Physical Median, No Opening, Directional Opening, Full Median Opening)	
Driveway	
Driveway Location (Latitude and Longitude)	
Driveway Design (Flush Radial, Curb Radial, Curb Flare, Wide-open frontage or other types)	
Driveway Number of Lanes (1, 2, 3, 4 or more)	
Driveway Traffic Operations (One-way Entry/Exit, Right-in/Right-out (No Opening or Channelizing Island), Right-in/Left-in/Right-out (Directional Opening), Full Traffic Movement, Left-in/Left-out)	
Driveway Channelization (None, Painted/Island)	
Traffic Control Device at Driveway (No Control, Sign Control, Traffic Signal)	
Right-Turn Lane Type (Exclusive (serves one site), Shared/Continuous, No Right-turn Lane)	
Driveway Throat Length (Adequate, Short (less than the length of two cars), None)	
Temporary Closure for Work Zone (Yes/No)	

After the data review was completed, the reviewed data were integrated with the roadway characteristics data and driveway-related crash data based on the FDOT roadway ID, road section number information, and geospatial location proximity. Specifically, each driveway-related crash record was examined to determine if it fell into the influence areas of driveways before it was linked with the closest driveways, as shown in Figure 4. The driveway influence area defined by the Florida Design Manual is 76 m (250 ft) upstream and downstream from the driveway location. Due to the large number of driveways along urban corridors, the influence areas of multiple driveways may overlap. Therefore, the influence areas under such scenarios

were divided into segments based on the distance between adjacent driveways, using the method applied by Williamson and Zhou (2014).



Figure 4. Example of crash clusters and crash assignment to corridor driveways.

Regression Models

Utilizing statistical analysis methods can help unveil how driveway-related crashes are distributed across various values of potential contributing factors, highlighting which factors demand specific attention. In current literature, regression modeling techniques are commonly employed to delve deeper into the individual effects of contributing factors on both the frequency and severity of driveway-related crashes (Avelar et al., 2013; Chen et al., 2018; Gorthy, 2017; Kwigizile et al., 2014; Le et al., 2018). Different models were therefore used in our study, including a negative binomial model for crash frequency analysis, and a multinomial logit model for crash severity analysis. The details of these models are as follows:

Negative binomial model for crash frequency analysis

In this study, a non-negative binomial regression model was employed to investigate the influence of roadway, driveway, and traffic characteristics on driveway-related crash distributions. A general structure of the negative binomial model is presented and explained below (Gorthy, 2017):

$$\ln \lambda_i = \beta X_i + \varepsilon_i$$

where,

λ_i is the expected number of crashes or crash rate for driveway i ;

X_i is a vector of explanatory variables to be examined;

β is a vector of coefficients to be estimated; and,

e^{ε_i} is a gamma-distributed error term with the mean equal to 1.

The explanatory variables for the negative binomial regression model included the crash, traffic, roadway, and driveway characteristics found in Table 1. Nominal variables were converted to variables with a limited number of categorical values for modeling purposes. A Type 3 analysis through Chi-square test was conducted to examine the significance of each explanatory variable, considering all other variables are present in the model (SAS Institute Inc,

2022). All the significant variables are identified at the level of significance $p=0.05$ (95% confidence level).

Multinomial logit model for crash severity analysis

The multinomial logit model (MNL) was used to examine crash severity outcomes at commercial driveways and the contributing factors relative to roadway features, traffic conditions, driveway characteristics, and other crash-related variables. Florida adopted the “KABCO” injury scale defined by the National Highway Traffic Safety Administration (NHTSA, 2017) to document crash and road user injury severities, where “K” indicates fatality, “A” represents incapacitating injury, “B” denotes non-incapacitating injury, “C” is possible injury, and “O” is no injury (property damage only). Fatal and incapacitating injury crashes are of the highest interest in traffic safety, but often account for a very small proportion of all crashes. Therefore, three injury severity levels were defined for this study, including no injury (**NI**, including all “O” crashes), minor injury (**MI**, including “B” and “C” crashes) and severe injury (**SI**, including “K” and “A” crashes).

As noted by Kim et al. (2007), the positivity or negativity of a coefficient estimated from a logit model with three or more crash severity levels in the response variable cannot be intuitively interpreted as the increase or decrease in the probability of that crash severity. To properly evaluate the influence of contributing factors on crash severity outcomes, a direct average pseudo-elasticity analysis is necessary, which works by altering the values of each contributing factor and examining the probability change. For this study, the variables were all converted to 0-1 indicator variables for logit modeling. The average pseudo-elasticity is defined by the percentage change in probability when an indicator variable is changed from 0 to 1 (and 1 to 0), and is calculated as follows:

$$E_{x_{nk}}^{P_{ni}} = \frac{P_{ni}[x_{nk} = 1] - P_{ni}[x_{nk} = 0]}{P_{ni}[x_{nk} = 0]}$$

where, $E_{x_{nk}}^{P_{ni}}$ is the direct average pseudo-elasticity of the k th variable from the vector x_n . P_{ni} is the probability of crash n resulting in injury severity level i and is defined as the following according to the basic structure of multinomial logit model:

$$P_{ni} = \frac{e^{\beta_i x_n}}{\sum_{i'} e^{\beta_{i'} x_n}}$$

where β_i is the vector of coefficients estimated specific to crash severity level i and x_n is a vector of exogenous variables for crash n . This average pseudo-elasticity method has been used in several authentic traffic safety studies (Shankar and Mannering, 1996; Ulfarsson and Mannering, 2004), and therefore is also used in this study to evaluate the marginal effects of the contributing factors.

For the crash frequency analysis datasets, each data entry is a driveway record with all linked features and crash frequency at the driveway site. For the crash severity analysis datasets, each data entry is a crash record linked with all roadways, traffic, and driveway characteristics, and additional relevant variables representing crash characteristics, such as alcohol involvement,

lighting condition, weather condition, road surface, shoulder type, etc. were also included in the modeling process.

RESULTS AND DISCUSSION

This section discusses in detail the results of the negative binomial model for crash frequency analysis and the multinomial logit model for crash severity analysis.

Crash Frequency Analysis

Negative binomial modeling and significance tests were used to analyze pedestrian and bicycle crash frequency at commercial driveways along major corridors, identify the significant variables, and estimate their influence. Modeling analysis results are presented in Tables 3-6 accordingly., A total of five variables was found to be significant in determining the number of pedestrian and bicycle crashes at commercial driveways along corridors at a 5% significance level, including number of lanes on connecting street, driveway design feature, median opening type, traffic control device, and painted bike lane, as shown in Table 2.

Table 2. Significant Variables for Pedestrian and Bicycle Crash Frequency Prediction at Commercial Driveways along Corridors

Source	Degree of Freedom	Chi-Square	Pr > ChiSq
Number of Lanes on Connecting Street	3	8.28	0.0406
Driveway Design Feature	3	42.66	<.0001
Median Opening Type	3	11.65	0.0087
Traffic Control Device	2	33.33	<.0001
Painted Bike Lane	2	15.84	0.0004

Looking into more detail using the negative binomial model, Table 3 presents the estimated influence on pedestrian and bicycle crash frequency at commercial driveways of these significant variables for each categorical value.

Crash Severity Analysis

The findings from the multinomial logit analysis are summarized below.

- **Type of Shoulder:** Type of shoulder was found to be a significant factor affecting the pedestrian and bicycle crash severity at commercial driveways along corridors. Specifically, the presence of a paved shoulder significantly reduced the risk of both minor injuries and severe injuries by 2.0% and 37.6%, respectively. These results suggest a significant statistical dependence between paved shoulders and pedestrian and bicycle crash severity reduction. Therefore, paved shoulders should be considered near commercial driveways in areas with high pedestrian and bicycle activity.
- **Alcohol or Drug Involved:** As expected, alcohol or drug involvement either by the motorist or the non-motorist or both, significantly increased the risk of severe injuries in

commercial driveway crashes. Compared to crashes without alcohol or drug involvement, alcohol involvement increased the potential of severe injury crashes by 208.95%. Although not specific to access management per se, this result demonstrates the extremely negative impact of alcohol or drug involvement on traffic safety and verifies the necessity of law enforcement and zero tolerance on driving under the influence (DUI).

Table 3. Negative Binomial Modeling Results for Pedestrian and Bicycle Crash Frequency at Commercial Driveways along Corridors

Parameter	Estimate	Std Error	Odds Ratio	Wald Chi-Square	Pr > ChiSq
Intercept	-2.6027	0.3043	0.0741	73.14	<.0001
Number of Lanes on Connecting Street (One-way, including all lanes)					
One Lane*					
Two Lanes	0.5657	0.3071	1.7607	3.39	0.0655
Three Lanes	0.39	0.3061	1.4770	1.62	0.2025
Four Lanes or More	0.2801	0.3107	1.3233	0.81	0.3673
Driveway Design Feature					
Curb Flare*					
Flush Radial	-0.733	0.1744	0.4805	17.67	<.0001
Curb Radial	0.3014	0.0895	1.3517	11.34	0.0008
Wide-open Access and Other	-0.1407	0.1195	0.8687	1.39	0.2389
Median Opening Type					
No Physical Median*					
No Opening	-0.2392	0.1118	0.7873	4.58	0.0323
Directional Opening	-0.4607	0.1721	0.6308	7.17	0.0074
Full Median Opening	0.0464	0.135	1.0475	0.12	0.7313
Traffic Control Device					
No Control*					
Sign Control	0.4202	0.0805	1.5223	27.28	<.0001
Signal Control	0.8666	0.2318	2.3788	13.98	0.0002
Painted Bike Lane					
Not Applicable (if no bike lane)*					
Not Painted	0.3341	0.0833	1.3967	16.11	<.0001
Painted	0.1275	0.1043	1.1360	1.49	0.2216
Dispersion	1.2514	0.1978			

* Indicates the base condition used for analysis

- **Driveway Number of Lanes:** Two-lane driveways were found to significantly increase the risk of severe injury in pedestrian and bicycle crashes by 162.2%. Driveways with four or more lanes or wide-open access increased the risk of severe injury in pedestrian and bicycle crashes by 231.5%. A possible reason for these findings is that higher number

of lanes on commercial driveways is generally associated with more daily driveway trips and higher non-motorist activity, therefore leading to higher risk for more severe pedestrian and bicycle crashes. These results show a significant dependence between pedestrian and bicycle crash severity and driveway number of lane configuration and are worth further in-depth investigation.

- **Bike Lane Type:** Having no bike lane was found to be a significant factor in reducing the probability of minor injury pedestrian/bicycle crashes by 8.8%; however, the pseudo-elasticity analysis showed that absence of a bike lane increased the risk of severe injury pedestrian/bicycle crashes by 27.1%. These results are reasonable, because bicyclists typically travel on roadway shoulders or sidewalks if there is no bike lane. When a bicycle crash occurs, it is more likely that the bicyclist is traveling next to the travel lane with higher exposure. Therefore, a dedicated bike lane is recommended to reduce pedestrian/bicycle crash severity. These results also verify the need to conduct both MNL modeling and pseudo-elasticity analysis since the MNL results in Table 4 do not reveal this influence on severe injuries. If there is no bike lane available, it is very likely that bicyclists are traveling on the sidewalk, keeping a distance from mainstream traffic. If bicyclists choose to travel next to the travel lane when there is no bike lane available, crashes are more likely to occur and those that occur are more likely to result in severe injuries or fatalities.

Table 4. Multinomial Logit Modeling Results for Pedestrian and Bicycle Crash Severity at Commercial Driveways along Corridors

Variable	Specific to Injury Severity	Estimated Coefficient	Standard Error	P-value
Intercept	MI	1.47098***	0.2216	<0.0001
Intercept	SI	-1.36825***	0.4343	0.0016
Type of Shoulder				
Paved Shoulder	MI	-0.45372**	0.1935	0.019
Paved Shoulder	SI	-0.91696***	0.2864	0.0014
Alcohol or Drug Involvement				
Alcohol or Drug Involved	SI	1.54170***	0.596	0.0097
Driveway Number of Lanes				
Two Lanes	MI	0.45875**	0.2218	0.0386
Two Lanes	SI	1.47117***	0.441	0.0008
Four or More OR Wide-open Access	SI	1.68231***	0.6403	0.0086
Bike Lane Type				
No Bike Lane	MI	-0.33891**	0.1485	0.0225
Driveway Throat Length				
Short Throat Length	SI	0.45598**	0.2117	0.0312
Connecting Street AADT at Crash Year				
50,000<AADT≤60,000	MI	0.63122**	0.2514	0.012
Note: *** , ** , * ==> Significance at 1%, 5%, 10% level.				

- **Driveway Throat Length:** Insufficient driveway throat length also significantly increased the risk of severe injury pedestrian and bicycle crashes by 46.4%, as shown in Table 5. Similar to the conclusions drawn from vehicular safety findings, these results confirm the importance of sufficient driveway throat length for commercial driveways for pedestrian and bicycle safety along commercial corridors.
- **Connecting Street AADT at Crash Year:** For pedestrian and bicycle crashes at commercial driveways along corridors, the AADT range of 50,000 and 60,000 was the only category found to be significant and was estimated to increase the potential for minor injury crashes by 15.5%.

Table 5. Average Pseudo-elasticity of Significant Variables for Pedestrian and Bicycle Crash Severity at Commercial Driveways along Corridors

Variable	Injury Severity Level		
	NI	MI	SI
Type of Shoulder			
Paved Shoulder	53.3%	-2.0%	-37.6%
Alcohol or Drug Involved			
Alcohol or Drug Involved Crash	-28.56%	-29.54%	208.95%
Driveway Number of Lanes			
Two Lanes	-37.2%	-1.4%	162.2%
Four or More OR Wide-open Access	-31.2%	-32.6%	231.5%
Bike Lane Type			
No Bike Lane	27.5%	-8.8%	27.1%
Driveway Throat Length			
Short Throat Length	-5.7%	-5.9%	46.4%
Connecting Street AADT at Crash Year			
50,000<AADT≤60,000	-38.2%	15.5%	-37.8%

CONCLUSIONS

Driveway accidents and fatalities represent one of the primary types of crashes both across the nation and in Florida. Significant efforts are being made to apply treatments and develop policies to reduce these crashes. This research evaluated the crash frequency and severity of vehicle related pedestrian and bicycle crashes occurring on commercial driveways using five-years of data from 2015-2019. It also identified the significant factors that contribute to the occurrence and severities of these crashes.

The crash frequency analysis showed four variables to be significant in improving or worsening the crash frequency at driveways. The two driveway design variables were found to be significant: flush radial and curb radial. The flush radial design reduced the crash frequency by 52%, while the curb radial design tended to increase the crash frequency by 35.2%. Radial return designs are often used on high-volume driveways, which have higher crash potential. On flush shoulder roadways, FDOT prefers sidewalk placement outside the clear zone or five feet beyond the shoulder pavement to provide adequate protection for pedestrians or bicyclists.

The median opening type variable reduced crash frequency. The no opening type and directional opening type median were able to reduce crash frequency by 21.3% and 36.9%, respectively. Physical medians (both no opening and directional opening) provide buffer space for pedestrians and bicyclists to wait to cross, reducing collision risk with through traffic. No median opening or a directional median opening limits vehicular turning movements, thereby also reducing driveway conflicts. The presence of traffic control devices, however, increases crash frequency. If a sign control is present, the crash frequency increases by 52.2%, while the traffic signal control increases it by 137.9%.

Driveways with sign or traffic signal controls tend to have higher traffic volume and more complex traffic than locations with no traffic controls and, therefore, experience higher crash frequencies. If there is no painted bike lane, the crash frequency increases by 39.7%. Conventional bike lanes without paint do not necessarily provide protection. Motor vehicles must cross bike lanes to enter or exit driveways, leading to conflicts with bicyclists in the bike lane.

Three crash severity levels were used in the analysis, including no injury (NI, base category), minor injury (MI), and severe injury (SI). The sum of the probabilities for all three crash severity levels is equal to 1, so increasing the probability of one severity level will decrease the probability of another, or the other two, crash severity levels. The multinomial analysis showed that six variables were significant in determining their influence on crash severity levels. The paved shoulder variable showed a 2% decrease in the probability of minor injury (MI) and a 37.6% decrease in the probability of severe injury (SI). Paved shoulders should be considered near commercial driveways in areas with high pedestrian and bicycle activity.

Alcohol and drug involvement increased the probability of severe injury by 208.95%. Although not specific to access management, this confirms the severe adverse impact of substance use on traffic safety. If there are two lanes in the driveway, the probability of minor injury reduces by 1.4% but increases the probability of severe injury by 162.2%. Also, if there are four lanes or more or wide-open access, the probability of severe injury increases by 231.5%. Multiple driveway lanes suggest more complex traffic conditions, relatively higher vehicle speeds, and more pedestrian and bicycle exposure, therefore inducing severe injury.

If there is no bike lane, the probability of minor injury reduces by 8.85%. If no bike lane is available, many bicyclists travel on the sidewalk to avoid mainstream traffic. If they travel next to the travel lane, severe injuries or fatalities are likely when a crash occurs. Short driveway throat length increases the probability of severe injury by 46.4%. Sufficient driveway throat length at commercial driveways is important to pedestrian and bicycle safety along corridors. Lastly, if the AADT is between 50,000 and 60,000, the probability of minor injury increases by 15.5%. Only this AADT categorical value was statistically significant in explaining crash severity outcomes.

There are some limitations in this research. This research relies on accurate geolocation of these crash data included in the analysis. It was found during the data integration process that some crashes were improperly geolocated, and some crash reports did not identify which commercial driveway location(s) were involved in the crash. Therefore, manual review and correction were included in the data process based on detailed information on the crash reports. Considering the large number of crashes included in the analysis, these inaccuracies should impose only a minor or negligible influence on our research findings.

Another limitation relates to the influence of relationships among the studied variables. Although each variable in the modeling procedure represents a different roadway or driveway

characteristic, some variables are slightly or moderately interrelated. For example, the number of lanes on connecting street and connecting street categories' 5-year average AADT are highly interrelated, with both generally associated with heavier traffic volumes. Nonetheless, the researchers did find different influences from different crash groups. Despite an overall upward trend on all crash frequencies for both variables, different effects were observed regarding pedestrian and bicycle crashes at commercial driveways. The number of lanes on connecting street category was significant overall for predicting pedestrian and bicycle crashes, but none of the categorical values were significant, and AADT was not significant for these crashes in the negative binomial modeling process.

Overall, these findings provide important insights on addressing pedestrian and bicycle safety issues at commercial driveways, which could be used as a reference for future access management guideline improvement.

REFERENCES

- AASHTO. (2011). *A Policy On Geometric Design*. Washington, DC.
- Avelar, R. E., Dixon, K. K., Brown, L. S., Mecham, M. E., and Van Schalkwyk, I. (2013). "Influence of Land Use and Driveway Placement on Safety Performance of Arterial Highways." *Transportation Research Record*, SAGE Publications Inc, 2398(1), 101–109.
- Chen, C., Wu, Q., Zhang, G., Liu, X. C., and Prevedouros, P. D. (2018). "Extracting Arterial Access Density Impacts on Safety Performance Based on Clustering and Computational Analysis." *Journal of Transportation Engineering, Part A: Systems*, American Society of Civil Engineers, 144(4), 4018008.
- Dixon, K., van Schalkwyk, I., and Layton, R. (2009). "Balancing Urban driveway design demands based on stopping sight distance." *Transportation Research Record*, (2120), 18–27.
- FDEP. (2017). *Statewide Land Use Land Cover*. Geospatial Open Data.
- FDOT (Florida DOT). (2009). *Crash data reports user manual for the crash analysis reporting system (CARS)*. Tallahassee, Florida.
- FDOT (Florida DOT). (2021a). *SSOGis – Web Application*.
- FDOT (Florida DOT). (2021b). *Transportation Data Portal*.
- Gattis, J. L., Gluck, J. S., Barlow, J. M., Eck, R. W., Hecker, W. F., and Levinson, H. S. (2010). NCHRP Report 659: Guide for the Geometric Design of Driveways. Washington, D.C.
- Gorthy, S. C. (2017). *Analysis of Right-In, Right-Out Commercial Driveway Safety, Operations and Use of Channelization as Compliance Countermeasure*. Clemson University.
- Kim, J.-K., Kim, S., Ulfarsson, G. F., and Porrello, L. A. (2007). "Bicyclist injury severities in bicycle-motor vehicle accidents." *Accident; analysis and prevention*, 39(2), 238–51.
- Kwigizile, V., Mulokozi, E., Xu, X., Teng, H., and Ma, C. (2014). *Investigation of the Impact of Corner Clearance on Urban Intersection Crash Occurance*.
- Le, T. Q., Gross, F., and Harmon, T. (2018). *Safety Evaluation of Corner Clearance at Signalized Intersections*. McLean, VA.
- NASEM (National Academies od Sciences Engineering and Medicine). (2021). *Guide for pedestrian and bicyclist safety at alternative and other intersections and interchanges*. Washington, DC.
- NHTSA (National Highway Traffic Safety Administration). (2017). *Model Minimum Uniform Crash Criteria (MMUCC) Guideline*.

- NHTSA (National Highway Traffic Safety Administration). (2023a). *Bicyclist and Other Cyclist Traffic Safety Facts*.
- NHTSA (National Highway Traffic Safety Administration). (2023b). *Ranking of State Pedestrian Fatality Rates*.
- SAS Institute Inc. (2022). *SAS*.
- Shankar, V., and Mannering, F. (1996). "An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity." *Journal of Safety Research*, 27(3), 183–194.
- Ulfarsson, G. F., and Mannering, F. L. (2004). "Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents." *Accident Analysis & Prevention*, 36(2), 135–147.
- Williams, M. K., Stover, G. V., Dixon, K. K., and Demosthenes, P. (2014). *Access Management Manual*. 543p.
- Williamson, M., and Zhou, H. (2014). "A Study of Safety Impacts of Different Types of Driveways and their Density." *Procedia - Social and Behavioral Sciences*, 138, 576–583.

Unraveling Crash Causation: A Deep Dive into Non-Motorists on Personal Conveyance

Subasish Das, Ph.D.¹; Rohit Chakraborty²; and Mahmuda Sultana Mimi³

¹Assistant Professor, Dept. of Civil Engineering, Ingram School of Engineering, Texas State Univ., San Marcos, TX. ORCID: <https://orcid.org/0000-0002-1671-2753>.

Email: subasish@txstate.edu

²M.S. Student, Dept. of Civil Engineering, Ingram School of Engineering, Texas State Univ., San Marcos, TX. ORCID: <https://orcid.org/0000-0002-1660-9764>.

Email: rohitchakraborty@txstate.edu, qnb9@txstate.edu

³Dept. of Civil Engineering, Ingram School of Engineering, Texas State Univ., San Marcos, TX

ABSTRACT

Non-motorists using personal conveyances, like skateboards, face increased safety challenges due to reduced visibility, particularly at intersections and in low light conditions. Safety studies on non-motorists using personal conveyances rely on limited hospital datasets or controlled/naturalistic riding conditions, lacking comprehensive identification of e-scooter riding risk factors. To address this research gap, this study collected real-world traffic crash data and quantified safety risks related to key contributing factors. Utilizing the 2020–2021 Fatality Analysis Reporting System (FARS) fatal crash data, this study examined the patterns of crashes associated with non-motorists using personal conveyances at segments and intersections. The findings provide valuable insights into key risk factors that can guide stakeholders, municipalities, and campus administrators in developing effective mitigation strategies to reduce safety risks associated with non-motorists using personal conveyances. By addressing these safety concerns, personal conveyances devices can be integrated as a safe and sustainable shared mobility option in urban and campus environments.

Keywords: non-motorists, personal conveyance, traffic crashes, crash severity, Bayesian network.

INTRODUCTION

Personal conveyances encompass a wide array of devices used by non-motorists and pedestrians for mobility and leisure, which may be motorized or human-powered, excluding those propelled through pedaling (NHTSA, 2023). This category includes rideable toys, motorized rideable toys, and personal mobility assistance devices, such as self-balancing devices, wheelchairs, and mobility scooters. These options are popular in urban areas for their environmental benefits and convenience in first- and last-mile travel. However, users face challenges like risky behaviors and conflicts with other road users, leading to concerns about non-motorist injuries. Non-motorists, often vulnerable road users, lack comprehensive regulations, raising safety concerns. The absence of crash data hinders understanding non-motorist safety, and this study addresses the research gap by analyzing 2020-2021 crash data in the United States using Bayesian Network analysis.

This study will investigate factors contributing to injury crashes involving non-motorist safety and personal conveyances, offering insights into the complex dynamics between non-motorists and other road users in urban areas. Through a thorough analysis, the research seeks to

recommend evidence-based policy interventions and safety measures, designed to mitigate risks associated with personal conveyances and enhance the well-being of sustainable transportation users. Ultimately, the study strives to contribute to creating safer and more accessible environments for non-motorists, promoting the adoption of sustainable transportation options and ensuring the safety of those using personal conveyances for mobility and leisure.

LITERATURE REVIEW

The rising popularity of eco-friendly non-motorized travel modes has spurred research on usage patterns and safety. This literature review consolidates findings from various studies to comprehensively understand challenges and opportunities in non-motorist crashes.

Safety Studies on Crashes

Several studies have examined the factors related to the non-motorist crashes and the non-motorized modes. Salas-Niño (2022) studied the legal framework for micro-mobility devices in a major U.S. city, finding a higher number of non-motorist injuries compared to traditional vehicles. This highlights the global need for effective safety regulations. Brunner et al. (2022) examined threat response and crash avoidance in non-motorists, emphasizing the importance of response latency and capabilities. Their study, involving 36 participants, simulated safety impacts of automated driving and traffic measures. Haworth et al. (2021) compared safety behaviors of shared and private non-motorists in Brisbane, Australia. They found a decline in shared e-scooter usage linked to a reduction in illegal behaviors, coinciding with an increase in private e-scooter usage.

Shah et al. (2021) and Pobudzei et al. (2023) investigated e-scooter and bicycle crashes, identifying distinct crash types. E-scooter crashes exhibited specific patterns. Approximately 10% of these crashes resulted in injuries or fatalities, and most crashes were concentrated in the city center. E-scooter casualties were more frequent on weekends and during evening and nighttime hours. A higher percentage of these casualties were intoxicated, and hit-and-run cases surpassed those involving cyclists (Pobudzei et al., 2023). These underscore the importance of tailoring safety interventions and campaigns for e-scooters (Shah et al., 2021).

Mayer et al. (2020) investigated the road safety implications of e-scooters in Austria and Germany, where they derived measures to enhance road safety concerning e-scooters. Huemer et al. (2022) observed 4,514 e-bike and e-scooter riders in Braunschweig, Germany, finding 13.4% engaged in distracting tasks like using headphones, leading to more violations and reduced safety gear use. Cluster analysis identified high-risk groups of young and middle-aged male riders of electric bikes and e-scooters, suggesting targeted campaigns for these demographics to address safety concerns. Sandt et al. (2022) tackled emerging e-scooter safety issues, outlining community efforts for injury prevention. The digest covered usage patterns, safety trends, contributing factors to crashes, injuries, fatalities, and city initiatives for e-scooter management and regulation. Ma et al. (2021) explored e-scooter safety risks through naturalistic riding experiments, emphasizing significant impacts like severe vibrations, especially on concrete pavements, and proximity to objects in constrained riding environments. The findings underscore heightened safety challenges for e-scooter riders.

Clewlow et al. (2022) described the U.S. Department of Transportation's Safety Data Initiative (SDI), which utilized GPS trip trace data to improve road networks and safety for

vulnerable road users. Their work showcased the potential for data-driven insights in enhancing safety for a wide range of road users. Azimian and Jiao (2022) addressed e-scooter injury crashes in Austin, focusing on variables like demographics, income, land use, and education. Their results highlighted the need for infrastructure development, educational campaigns, and stricter enforcement to enhance e-scooter safety.

Safety Studies on Non-Motorists using other Personal Conveyances

Several types of personal conveyance devices such as roller skates, skateboards, baby carriages, toy skates have been used by non-motorists. In California, users of skateboards, roller skates, and push scooters cover up to 48 million miles per year (Fang and Handy, 2019). In 2012, 14 skateboarders died in automobile collisions in California. The estimated fatality rate suggests 20.9 to 23.0 deaths per 100 million miles traveled (Fang and Handy, 2017). A study compared the injury severity of skateboarders and long boarders. Longboard riders, due to increased dimensions allowing higher speeds, experienced significantly more fatal or severe injuries like skull fractures, traumatic brain injuries, and intracranial hemorrhage compared to traditional skateboarders (Fabian et al., 2014). Valdez (2016) offered safety guidelines for self-balancing motorized skateboard riders, recommending an age restriction (children under 13) and advising against using the device in or near moving traffic.

Previous studies in traffic safety analysis lack sufficient research on non-motorist safety analysis. Yasmin et al., (2021) tackled limitations in non-motorized crash prediction models arising from a lack of true exposure data. They addressed these issues by creating an integrated framework that combines non-motorized demand and crash prediction for comprehensive mobility and safety analysis. The proposal includes developing aggregate-level models for non-motorist generation and attraction, trip exposure matrices for safety evaluation, and crash frequency and severity proportion models at a zonal level.

Safety Studies on Bayesian Networks

Bayesian networks aid in motorcycle crash severity analysis by revealing relationships between potential factors. The Bayes' rule can predict future events based on preceding ones. (Das et al., 2023). Several studies used Bayesian networks for traffic crash analysis and pedestrian safety for determining the influence of the contributing factors on predicting crash outcomes (Kitali et al., 2021). The models can bring out the most probable factor or even combination that leads to an accident (Davis, 2003; Ma et al., 2018). In case of complex interrelationships that include multiple crash attributes and crash outcomes, the Bayesian networks were proved to be effective (De Oña et al., 2013).

The Bayesian Network approach is suitable for handling uncertain data types, like uncertainty in information provided by respondents, making it effective for acceptability estimation. Unlike traditional regression models, Bayesian Approaches offer flexibility in understanding interdependence between multiple factors and assessing counterfactual scenarios (Garces et al., 2016). Moreover, Bayesian Networks can also be used to investigate the impact of influential factors on crash severity for large vehicles such as large-trucks (Wu et al., 2023).

The discussed studies emphasize key insights into non-motorist safety, identifying factors like rider behaviors, environmental conditions, and infrastructure design. However, additional research is needed to uncover more variables and develop effective policies and countermeasures to mitigate non-motorist crashes.

METHODOLOGY

Data Collection

This study acquired fatal crash data from FARS for 2020-2021. According to the 2021 FARS/CRSS Coding and Validation Manual, the non-motorist on personal conveyance data includes crash data regarding roller skates, in-line skates, skateboards, skates, scooters, motorized skateboard and others (2023). Table 1 shows a 43% increase in no intersection crashes and a 26% increase in four-way intersection crashes from 2020 to 2021. Conversely, T-intersection and 'others' crashes decreased by 12% and 50%, respectively, while the overall number of crashes rose by 25%.

Table 1. Non-motorists Fatal Crashes by Intersection Type

Year	Not an Intersection	Four-Way Intersection	T-Intersection	Others	Grand Total
2020	42	27	17	2	88
2021	60	34	15	1	110
Grand Total	102	61	32	3	198

Bayesian Network

A Bayesian Network (BN) determines a joint probability distribution over a set of random variables U , which is an annotated directed acyclic graph (DAG). Consider, $U = \{A_1, \dots, A_n, C\}$ where n stands for the number of RIFs, the variables A_1, \dots, A_n are the RIFs, and C is the class variable (for example, frontage road related injury types). Consider a graph structure where the class variable is the root, that is, $\prod C = \emptyset$ ($\prod C$ denotes the set of parents of C in U) and each RIF has the class variable as its unique parent, i.e., $\prod A_i = \{C\}$ for $1 \leq i \leq n$. A BN defines a unique joint probability distribution over U given by

$$P(A_1, \dots, A_n, C) = P(C) \cdot \prod_{i=1}^n P(A_i|C) \quad (1)$$

The DAG on $\{A_1, \dots, A_n\}$ is a tree if $\prod A_i$ contains only one parent for all A_i , except for one variable without parents (referred as the root). If, function π can define a tree over A_1, \dots, A_n if there is exactly one i such that $\pi(i) = 0$ (i.e. the root of the tree), and there is no sequence i_1, \dots, i_k such that $\pi(ij) = i_{j+1}$ for $i \leq j < k$ and $\pi(i_k) = i_1$ (i.e., no cycles). Such a function defines a tree network where $\prod A_i = \{C, \dots, A_{\pi(i)}\}$ if $\pi(i) > 0$ and $\prod A_i = \{C\}$ if $\pi(i) > 0$, and $\prod A_i = \{C\}$ if $\pi(i) = 0$.

A Bayesian belief net or a decision diagram is a graphical representation of factors contributing to a conclusion or uncertainty (Howard and Matheson, 1984; Pearl, 1988). This diagram calculates the likelihood of the outcome given all the model's factors. Conditional probabilities establish a relationship between the inputs and the result. Decisions can be incorporated into the influence diagram so that the decision-maker can understand how each alternative impacts the probability of a result (Shachter, 2007).

RESULTS AND DISCUSSIONS

A comprehensive BN model was developed in this study by using Netica's advanced Expectation Maximization (EM) technique. EM was employed to fit the network to the input-output combinations obtained from the Monte Carlo runs (Lauritzen, 1995). The BN model parameters, also called conditional probability distributions, were estimated through the learning process. These distributions offer information about the likelihood of different states within each node of the BN. Figure 1 illustrates the final BN network layout, depicting the relationships between different nodes. Each node's conditional probability tables (CPTs, represented as belief bars) demonstrate the probabilities associated with the node's various states. These CPTs capture the relationships and dependencies between the variables within the BN, permitting probabilistic inference and analysis. The model's structure and conditional probability distributions obtained through the learning process are powerful tools for investigating the relationships and dependencies between variables related to non-motorist crashes.

Figure 1 also shows the initial data of non-motorist collisions modeled through a BN, focused on different intersection types. Most crashes occurred not at an intersection (51.5%) at a four-way intersection (30.8%). Most crashes occurred at an urban location (93.9%). The most common season for crashes was Autumn (36.4%) followed by Summer (32.3%), and most crashes occurred in the daylight (46.0%). Weekdays were the most common days for crashes to occur (79.8%). The most common sex was male (78.8%), and the most common age group was 45 to 65 years old (36.4%), followed by older than 65 years old (33.8%). In cases where drug use and drinking were known, most crashes did not involve drugs (37.4%) or drinking (42.9%). Most crashes were single-vehicle crashes (94.4%) and were often not a collision with a motor vehicle in transit (98.5%). These crashes at the intersections may occur due to collisions with properties or pedestrians based on different intersection types.

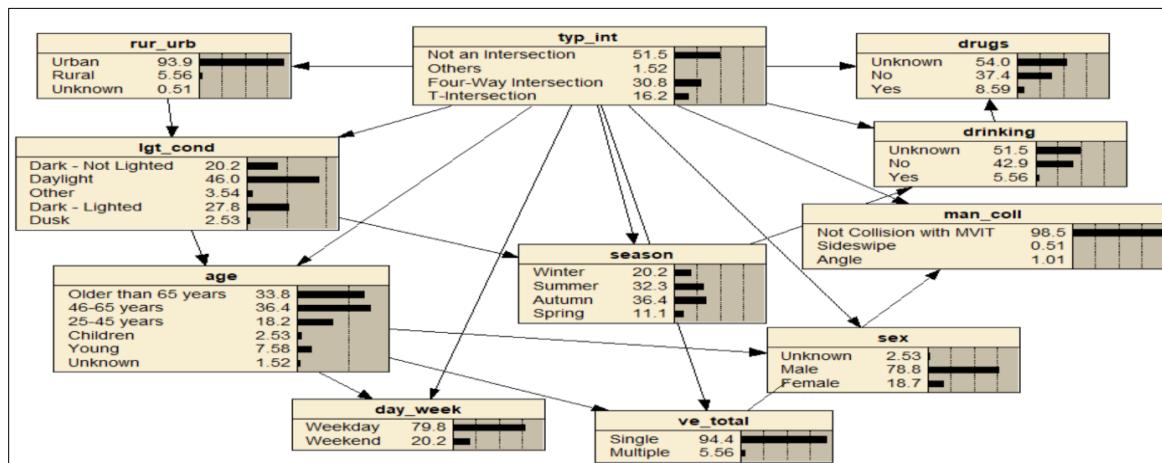


Figure 1. BN of full data

Table 2 shows the conditional probability chart for intersection types by lighting condition and age. Higher scores are displayed in dark green, and lower in lighter green to almost yellow. One of the highest probability scores occurs in the age group of 46 to 65 years old, at dusk, and not at an intersection (1.000). Probability scores of 1.000 also occur with an age group of older than 65 years old, with dusk lighting, and at a four-way intersection, with an age group young, at

dusk, and at a T-intersection, with an age group of 25 to 45 years old, other lighting conditions, and a T-intersection, and with an age group of 25-45 years old, in the dark but with lighting, and at other intersection types. Overall, the highest probability scores occur between 25 to 45 years old and 46 to 65 years old.

Table 2. Conditional Probability Table for Intersection Type by Lighting Condition and Age

Intersection Type	Lighting	Children	Young	25-45 years	46-65 years	Older than 65 years	Unknown
Not an Intersection	Daylight	0.022	0.065	0.065	0.348	0.478	0.022
Not an Intersection	Dark - Not Lighted	0.037	0.148	0.259	0.444	0.111	0.000
Not an Intersection	Dark - Lighted	0.000	0.080	0.280	0.520	0.120	0.000
Not an Intersection	Dusk	0.000	0.000	0.000	1.000	0.000	0.000
Not an Intersection	Other	0.000	0.000	0.500	0.500	0.000	0.000
Four-Way Intersection	Daylight	0.115	0.000	0.115	0.308	0.462	0.000
Four-Way Intersection	Dark - Not Lighted	0.000	0.000	0.667	0.167	0.000	0.167
Four-Way Intersection	Dark - Lighted	0.000	0.080	0.320	0.320	0.240	0.040
Four-Way Intersection	Dusk	0.000	0.000	0.000	0.000	1.000	0.000
Four-Way Intersection	Other	0.000	0.000	0.500	0.500	0.000	0.000
T-Intersection	Daylight	0.000	0.053	0.000	0.263	0.684	0.000
T-Intersection	Dark - Not Lighted	0.000	0.143	0.000	0.286	0.571	0.000
T-Intersection	Dark - Lighted	0.000	0.250	0.000	0.750	0.000	0.000
T-Intersection	Dusk	0.000	1.000	0.000	0.000	0.000	0.000
T-Intersection	Other	0.000	0.000	1.000	0.000	0.000	0.000
Others	Daylight	0.167	0.167	0.167	0.167	0.167	0.167
Others	Dark - Not Lighted	0.167	0.167	0.167	0.167	0.167	0.167
Others	Dark - Lighted	0.000	0.000	1.000	0.000	0.000	0.000
Others	Dusk	0.167	0.167	0.167	0.167	0.167	0.167
Others	Other	0.167	0.167	0.167	0.167	0.167	0.167

Table 3 shows the conditional probability scores for intersection types by different lighting conditions and seasons. Probability scores of 1.000 occurred during the winter, with other lighting conditions, and not at an intersection. During the autumn, at dusk, and at a T-intersection, in the spring, with other lighting conditions, and a T-intersection, and during autumn, with other lighting conditions, and at 'other' intersection type. The only season with no conditional probability scores of 1.000 was the summer.

Table 3. Conditional Probability Table for Intersection Type by Lighting Condition by Season

Intersection Type	Lighting	Winter	Summer	Autumn	Spring
Not an Intersection	Daylight	0.174	0.391	0.326	0.109
Not an Intersection	Dark - Not Lighted	0.185	0.444	0.333	0.037
Not an Intersection	Dark - Lighted	0.120	0.240	0.520	0.120
Not an Intersection	Dusk	0.000	0.500	0.500	0.000
Not an Intersection	Other	1.000	0.000	0.000	0.000
Four-Way Intersection	Daylight	0.231	0.269	0.423	0.077
Four-Way Intersection	Dark - Not Lighted	0.167	0.833	0.000	0.000
Four-Way Intersection	Dark - Lighted	0.120	0.280	0.400	0.200
Four-Way Intersection	Dusk	0.500	0.000	0.500	0.000
Four-Way Intersection	Other	0.000	0.500	0.500	0.000
T-Intersection	Daylight	0.211	0.263	0.263	0.263
T-Intersection	Dark - Not Lighted	0.714	0.000	0.286	0.000
T-Intersection	Dark - Lighted	0.500	0.000	0.500	0.000
T-Intersection	Dusk	0.000	0.000	1.000	0.000
T-Intersection	Other	0.000	0.000	0.000	1.000
Others	Daylight	0.250	0.250	0.250	0.250
Others	Dark - Not Lighted	0.250	0.250	0.250	0.250
Others	Dark - Lighted	0.000	0.000	1.000	0.000
Others	Dusk	0.250	0.250	0.250	0.250
Others	Other	0.250	0.250	0.250	0.250

Table 4 shows the conditional probability scores for intersection types by rider age and the number of vehicles. Overall, higher probability scores occurred when there was a single-vehicle collision. For single-vehicle crashes, when not at an intersection, the rider ages of children, young, and unknown all had probability scores 1.000. At a four-way intersection, the riders ages of children, young, and 25 to 45 years old had probability scores of 1.000. When at a T-intersection, the rider age group of 25 to 46 had a probability score of 1.000. At ‘other’ type of intersection, the rider age group of 25 to 46 years old had a probability score of 1.000.

Counterfactual Scenarios

Netica facilitates the creation and analysis of models like Bayesian Networks (BNs) and Markov networks. Applied in AI, machine learning, decision analysis, and data mining, Netica utilizes counting-learning for parameter learning in Conditional Probability Tables (CPTs) within TAN models. It then computes posterior probabilities, providing valuable insights for safety enhancement and crash prevention. The insights gained from Netica's analysis have great potential for informing proactive measures and strategies to mitigate risks and promote a safer environment.

Figure 2 shows the BN of the counterfactual population considering all crashes occurred not at an intersection. There was a 6.3% increase in crashes that occurred in dark but not light

conditions and a 6.7% increase in drivers aged 46 to 65 years. Interestingly, the other variables in this counterfactual did not see any significant changes.

Table 4. Conditional Probability Table for Rider Age and Number of Vehicles

Intersection Type	Rider Age	Single	Multiple
Not an Intersection	Children	1.000	0.000
Not an Intersection	Young	1.000	0.000
Not an Intersection	25-45 years	0.889	0.111
Not an Intersection	46-65 years	0.977	0.023
Not an Intersection	Older than 65 years	0.964	0.036
Not an Intersection	Unknown	1.000	0.000
Four-Way Intersection	Children	1.000	0.000
Four-Way Intersection	Young	1.000	0.000
Four-Way Intersection	25-45 years	1.000	0.000
Four-Way Intersection	46-65 years	0.889	0.111
Four-Way Intersection	Older than 65 years	0.950	0.050
Four-Way Intersection	Unknown	0.500	0.500
T-Intersection	Children	0.500	0.500
T-Intersection	Young	0.750	0.250
T-Intersection	25-45 years	1.000	0.000
T-Intersection	46-65 years	0.900	0.100
T-Intersection	Older than 65 years	0.941	0.059
T-Intersection	Unknown	0.500	0.500
Others	Children	0.500	0.500
Others	Young	0.500	0.500
Others	25-45 years	1.000	0.000
Others	46-65 years	0.500	0.500
Others	Older than 65 years	0.500	0.500
Others	Unknown	0.500	0.500

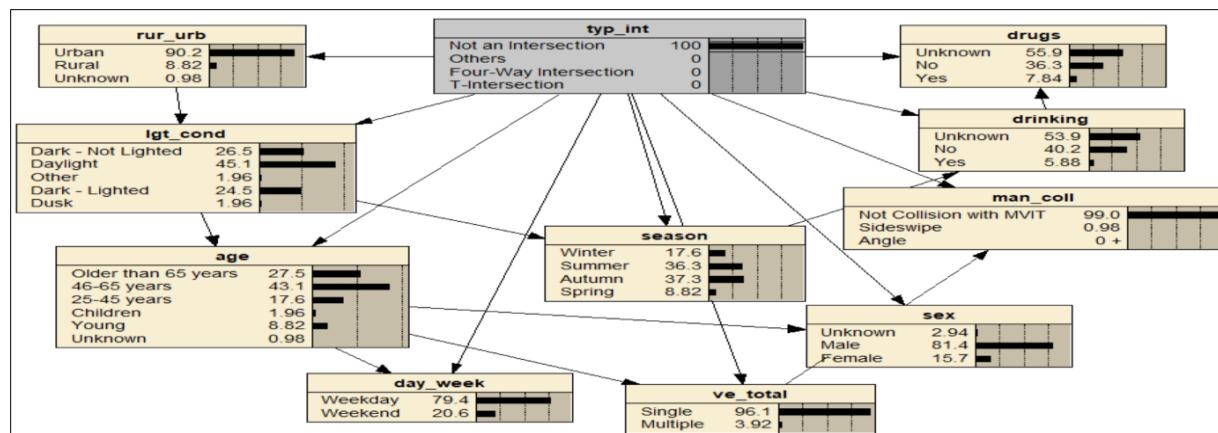


Figure 2. BN of counterfactual scenario considering all crashes occurred not at an intersection

Figure 3 illustrates a counterfactual scenario for a four-way intersection crashes. Regarding land use type, urban areas encountered a 6.1% increase. Collisions that resulted in a sideswipe decreased by 0.51%, while angle collisions increased by 1.64%. Regarding lighting conditions, there was a 10.36% decrease in dark and not lighted conditions, and a 13.2% increase in dark but lighted conditions. The recorded number of drivers using alcohol and drugs increased by 3.91% overall. In terms of driver age, the 25 to 45 years category increased by 8%, while the 46 to 65 years category decreased by 6.9%. Additionally, on weekdays there was a 5.4% increase, while on the weekends, there was a 5.4% decrease. Other variables did not show significant changes in this counterfactual scenario.

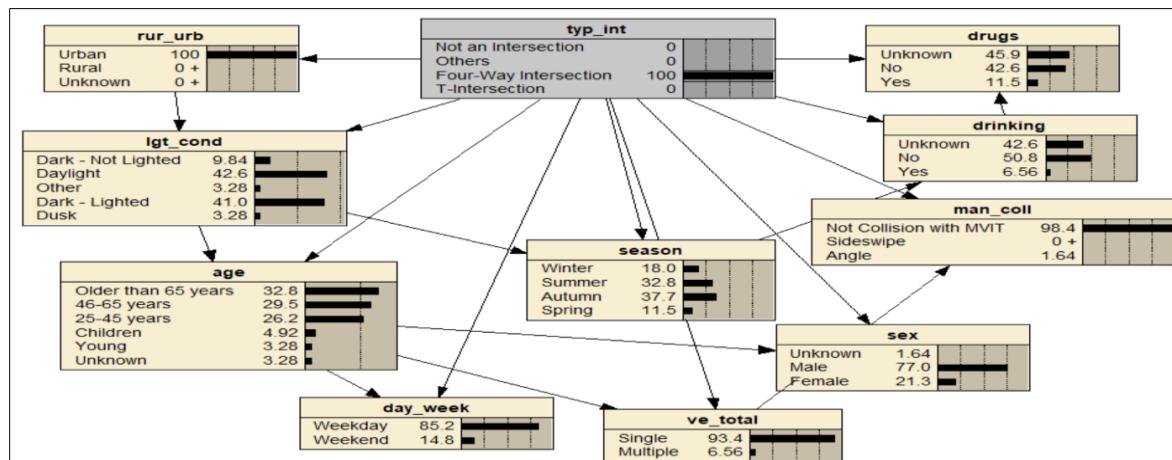


Figure 3. BN of counterfactual scenario considering all crashes occurred at a four-way intersection

Figure 4 presents the results of analyzing a counterfactual scenario where all crashes occurred at a T-intersection. There was a 2.12% increase in angle collisions in these crash incidents. The recorded number of drivers using alcohol and drugs decreased by 4.77% overall. The category of drivers aged 25 to 45 decreased by 15.07%, while those aged 65 and older increased by 16.7%. Regarding lighting conditions, there was a 13.4% increase in daylight and a 15.3% decrease in dark but lighted conditions. There was a 14.2% and 7.7% increase, respectively, in the winter season and spring seasons, while on the other hand, there was a 16.7% and 5.2% decrease, respectively, in the summer and autumn seasons. Additionally, on weekdays there was a 4.8% decrease, while on the weekends, there was a 4.8% increase. Other factors did not show significant changes in this counterfactual scenario.

The findings of this study provide comprehensive insights into non-motorist collisions using a BN model. The initial data analysis indicates that most crashes occur not at an intersection, with urban areas being the most common location. Crashes are more frequent during autumn and summer, in daylight, and on weekdays. The majority of those involved in crashes are males aged between 45 to 65 years, and alcohol and drug involvement in crashes is relatively low. Conditional probability tables also revealed crashes are more likely for ages 25 to 45 and 46 to 65, particularly during dusk and not at intersections. During the winter, crashes not at an intersection with other lighting conditions are most likely, while the summer has lower probabilities. Counterfactual scenarios showed how crash-related variables such as lighting

conditions, demographics, and seasonal characteristics differ between crashes at different intersections.

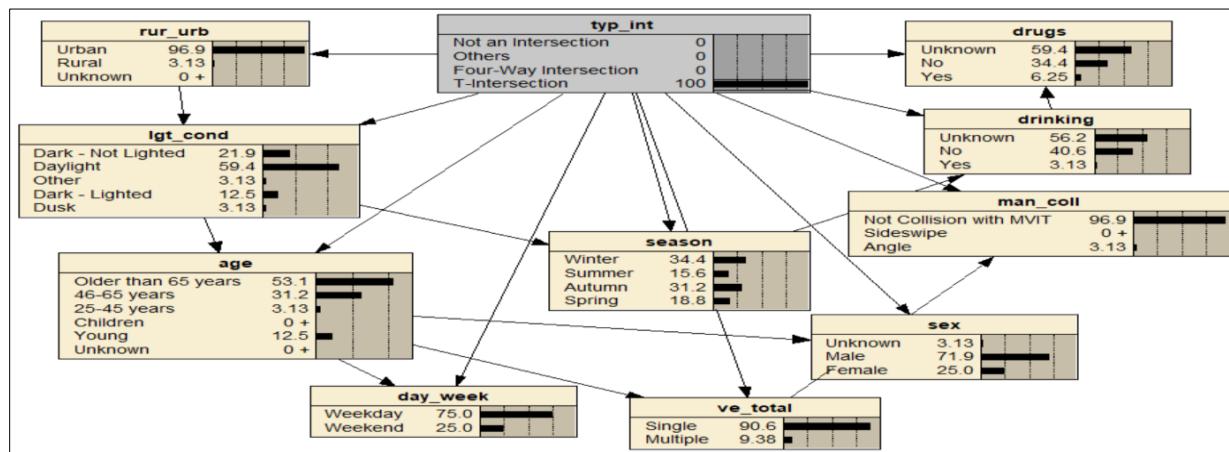


Figure 4. BN of counterfactual scenario considering all crashes occurred at a T-intersection

Based on the results, several policy-related guidelines can be recommended to improve road safety. Measures such as clear signage, designated pedestrian crossings, and advanced traffic signals can be introduced for tackling the 6.1% rise in urban crashes. Angle collision can be reduced by improving visibility at T-intersection. Increased law enforcement during high risk hours can be introduced. Targeted campaigns to reduce alcohol and drug use among drivers, especially in urban areas and on weekends can be implemented. Additionally, better street lighting should be installed, particularly at T-intersections, to address the 15.3% decrease in dark but lighted conditions. Implementation of traffic calming measures and increased police at non-intersection zones can resolve the 4.8% increase in weekend crashes. During the summer season, intensified safety campaigns can be held and increased law enforcement efforts can be made to tackle the 12.1% rise in incidents related to summer risks.

CONCLUSIONS

This study explored the contributing factors of non-motorists on personal conveyance related fatal crashes in the U.S. by utilizing data from the 2020-2021 FARS crash database. By developing a BN model, the study successfully uncovered relationships and dependencies between variables related to non-motorist crashes, such as temporal, demographic, and locational characteristics. Additionally, CPTs for different intersections involved in non-motorist collisions were displayed based on various factors such as age groups, seasons, and number of vehicles involved. The probabilities derived from this analysis offer valuable information for understanding the likelihood of non-motorist collisions at different intersections. Furthermore, the study conducted counterfactual analyses to assess the potential impact of considering all crashes occurring at different intersections type. These analyses demonstrate the changes in various factors under these counterfactual scenarios, such as lighting conditions, demographics, and seasonal characteristics.

The study unveiled nuanced patterns, revealing peak vulnerability times, high-risk demographics, and crash-prone locations. These insights provide a basis for targeted

interventions in non-motorist safety. Additionally, the research dissected interactions between non-motorists and road users, identifying conflict points and risky pedestrian behaviors, emphasizing the importance of policy interventions and awareness campaigns. Leveraging Bayesian Network analysis, the study proposed evidence-based policy interventions, recommending specific intersection designs, dedicated pedestrian crossings, bike lanes, and improved street lighting to enhance visibility, particularly at night.

Despite the valuable contributions of this study, it is crucial to acknowledge and address its limitations for future research improvement. Reliance on FARS data poses a risk of unreported incidents, limiting generalizability to the U.S. and a two-year period. Future studies could enhance data analysis by including more years and global trends, employing advanced statistical techniques and experimental designs for causal insights. Expanding variables, including weather and rider behavior, can lead to a more comprehensive understanding of non-motorist safety.

REFERENCES

- Azimian, A., and Jiao, J. 2022. Modeling factors contributing to dockless e-scooter injury accidents in Austin, Texas. *Traffic Injury Prevention* 23, pp-107-111.
- Brunner, P., von dem Bussche-Hünnefeld, T., Denk, F., Huber, W., Bogenberger, K., Kates, R. 2022. *An E-Scooter Safety Experiment – Design, Methodology and Results*. p. 18p.
- Clewlow, R., Foti, F., Seki, S., Muetting, E., Populus Technologies, I., and Office of the Secretary of Transportation. 2022. Final Report: Developing Scalable Models for Safety Insights and Improvements Using E-Scooter Exposure Data (Digital/other).
- Davis, G. A. 2003. Bayesian reconstruction of traffic accidents. *Law, Probability and Risk* 2, 69–89.
- De Oña, J., López, G., Mujalli, R., and Calvo, F. J. 2013. Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. *Accident Analysis & Prevention* 51, 1–10.
- Fabian, L. A., Thygerson, S. M., and Merrill, R. M. 2014. Boarding Injuries: The Long and the Short of It. *Emergency Medicine International* 2014, 1–7.
- Fang, K., and Handy, S. 2019. Skateboarding for transportation: exploring the factors behind an unconventional mode choice among university skateboard commuters. *Transportation* 46, 263–283.
- Fang, K., and Handy, S. 2017. Skate and die? The safety performance of skateboard travel: A look at injury data, fatality data, and rider behavior. *Journal of Transport & Health* 7, 288–297.
- Arbelaez Garces, G., Rakotondranaivo, A., and Bonjour, E. 2016. *Improving users' product acceptability: an approach based on Bayesian networks and a simulated annealing algorithm*.
- Haworth, N., Schramm, A., and Twisk, D. 2021. Changes in shared and private e-scooter use in Brisbane, Australia and their safety implications. *Accident Analysis & Prevention* 163, 106451.
- Howard, R., and Matheson, J. 1984. *Influence diagrams. the principles and applications of decision analysis*. Strategic Decisions Group.
- Huemer, A. K., Banach, E., Bolten, N., Helweg, S., Koch, A., and Martin, T. 2022. Secondary Task Engagement, Risk-Taking, and Safety-Related Equipment Use in German Bicycle and E-Scooter Riders – An Observation. *Accident Analysis & Prevention* 172, 106685.

- Kitali, A. E., Kidando, E., Kutela, B., Kadeha, C., Alluri, P., and Sando, T. 2021. Safety Evaluation of High-Occupancy Toll Facilities Using Bayesian Networks. *Journal of Transportation Engineering, Part A: Systems* 147, 04021018.
- Lauritzen, S. L. 1995. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis* 19, 191–201.
- Ma, Q., Yang, H., Mayhue, A., Sun, Y., Huang, Z., and Ma, Y. 2021. E-Scooter safety: The riding risk analysis based on mobile sensing data. *Accident Analysis & Prevention* 151, 105954.
- Ma, X., Xing, Y., and Lu, J. 2018. Causation Analysis of Hazardous Material Road Transportation Accidents by Bayesian Network Using Genie. *Journal of Advanced Transportation* 2018, 1–12.
- Mayer, E., Breuss, J., Robatsch, K., Salamon, B., and Soteropoulos, A. 2020. E-Scooter: Was bedeutet das neue Fortbewegungsmittel für die Verkehrssicherheit? *Zeitschrift für Verkehrssicherheit* 66, pp-153-64.
- National Highway Traffic Safety Administration. 2023. 2021 FARS/CRSS Coding and Validation Manual (No. DOT HS 813 426). National Highway Traffic Safety Administration.
- Pearl, J. 1988. Morgan Kaufmann series in representation and reasoning. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann, San Mateo, CA, US.
- Pobudzei, M., Tießler, M., Sellaouti, A., Hoffmann, S., and Transportation Research Board. 2023. *E-Scooter and Bicycle Accidents: Spatial, Temporal, and Demographic Characteristics in Munich, Germany*. p. 14p.
- Salas-Niño, L. 2022. Analysis of Current E-Scooter Safety Regulation in a Large U.S. City Using Epidemiological Components as a Framework. *Transportation Research Record: Journal of the Transportation Research Board* 2676, pp-163-172.
- Sandt, L., et al. 2022. *E-Scooter Safety: Issues and Solutions*. Transportation Research Board.
- Shachter, R. 2007. Model building with belief networks and influence diagrams. *Adv Decis Anal Found Appl* 177–207.
- Shah, N. R., Aryal, S., Wen, Y., and Cherry, C. 2021. Comparison of motor vehicle-involved e-scooter and bicycle crashes using standardized crash typology. *Journal of Safety Research* 77, pp-217-228.
- Yasmin, S., Bhowmik, T., Rahman, M., and Eluru, N. 2021. Enhancing non-motorist safety by simulating trip exposure using a transportation planning approach. *Accident Analysis & Prevention* 156, 106128.
- Das, S., Vierkant, V., Gonzalez, J. C., Kutela, B., and Sheykhfard, A. 2023. Bayesian Network for Motorcycle Crash Severity Analysis.
- Valdez, A. M. 2016. Playing It Safe: Injury Prevention for Self-Balancing Motorized Boards. *Journal of Emergency Nursing* 42, 269–271.
- Wu, J., Rasouli, S., Zhao, J., Qian, Y., and Cheng, L. 2023. Large truck fatal crash severity segmentation and analysis incorporating all parties involved: A Bayesian network approach. *Travel Behaviour and Society* 30, 135–147.

Understanding How We Can Make Commercial Driveways Safer: Exploratory Analysis through Case Studies

Tia Boyd, AICP¹; Kristine Williams²; Pei-Sung Lin, Ph.D.³; Cong Chen, Ph.D.⁴;
Elzbieta Bialkowska-Jelinska⁵; and Gina Bonyani⁶

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.

Email: tiaboyd@usf.edu

²Retired; formerly, Univ. of South Florida; Principal, KMW Associates LLC, Garden, MI.

Email: KMWAssociates@outlook.com

³Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.

Email: lin@usf.edu

⁴Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.

Email: congchen1@usf.edu

⁵Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.

Email: elzbieta@usf.edu

⁶Florida Dept. of Transportation, Tallahassee, FL. Email: Gina.Bonyani@dot.state.fl.us

ABSTRACT

Commercial development typically includes offices, shopping centers, gas stations, restaurants, and other uses that generate turning movements and traffic. As a result, commercial driveways and access streets connecting to major roadways are a key source of vehicular, bicycle, and pedestrian conflicts. Case studies are one of many methods for understanding the safety effects of commercial driveways. This paper includes six exploratory case studies conducted as a part of a statewide study of commercial driveway crashes in Florida. Case study applications for this safety analysis included: (1) illustrating safety issues with high levels of detail, (2) verifying research findings from high-level analysis, (3) exploring the root causes of identified safety issues, and (4) identifying potential countermeasures for specific safety issues. The case study sites represent a subsection of sites identified as high-crash commercial corridors. The results reveal several safety issues resulting from commercial driveway type and location in relation to intersections and interchange ramps, driveway alignment with median openings, driveway density, and traveler behavior (including motorists, bicyclists, and pedestrians). Although conducted in Florida, the approaches, lessons learned, and findings of this study are relevant to understanding commercial driveway safety across the US and internationally.

INTRODUCTION

Driveway design and location influence safety for motorists, bicyclists, and pedestrians. Studies evaluating the safety effects of driveways use a combination of analytical methods, including summary statistics, statistical analysis, and generalized linear modeling techniques. Exploratory case studies support these analysis methods through an in-depth assessment of the existing roadway design and site access conditions in a specific area. This granular level analysis, studying representative sites, is particularly beneficial for large statistical studies, informing how we interpret interactions between various elements and draw conclusions as to why crashes occur.

This paper describes case studies that were conducted to supplement the quantitative and qualitative analysis of crashes related to the type and location of driveway access in Florida. The study examined how commercial driveway access location and design interact with roadway and interchange characteristics to influence vehicular, bicycle, and pedestrian safety between 2015 and 2019. In this research, case studies were used to (1) illustrate safety issues with high levels of detail, (2) verify research findings from high-level analysis, (3) explore the root causes of identified safety issues, and (4) identify potential countermeasures for specific safety issues.

This paper begins with a description of the methodology to identify and assess the case study sites as well as illustrate the findings. The methodology is followed by descriptions of six case study sites, including a brief description of the roadway characteristics, driveway characteristics, crash frequency, crash types, a representative crash diagram, and potential countermeasures. The paper concludes with a discussion of the case study findings and a description of the outcomes and lessons learned.

METHODS

Case study sites were selected from a geographically diverse subset of potential study areas with a high proportion of commercial driveway-related crashes in Florida between 2015 and 2019. The broader study sample included: (a) 192 roadway segments with 9,889 commercial driveways and 10,596 driveway-related crashes; and (b) 69 interchanges with 832 commercial driveways and 853 driveway-related crashes in the vicinity of the interchanges. Details on the selection of the corridors and interchange areas for analysis are provided in the final report, available online at https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/research/reports/fdot-bdv25-977-76-rpt.pdf?sfvrsn=101903a0_1. The sample was then reduced based on evidence of crash clusters at commercial driveways in areas having different corridor or interchange area design and volume characteristics. A representative cross-section of crash clusters at specific commercial driveways along corridors and near interchanges as indicated by crash data was then identified for further analysis.

Crash locations were verified based on crash reports and relocated in ArcGIS if incorrectly geolocated in the data set. For commercial sites with more than one driveway, the research team made informed decisions as to which driveway was implicated in the crash based on the crash report narrative and crash diagram.

A total of six sites were selected for exploratory case studies. The study areas represent crash clusters on high-crash corridors and in interchange areas in different regions of Florida with distinctive design characteristics. The crash types and severity were documented in the high-crash locations, along with the interactions with other corridor and/or interchange characteristics through GIS analysis, crash report narrative review, and site observations using Google Earth. The characteristics of importance to the study included (1) roadway characteristics, (2) Average Annual Daily Traffic (AADT), (2) context, functional class, access class, (3) number of lanes, (4) driveway geometrics, (5) Driveway volume (i.e., number of parking spaces or Florida Department of Transportation (FDOT) driveway category), (6) driveway entry and exit movements (i.e., one-way, two-way, right-in/right-out), (7) driveway location in relation to signalized intersections and interchange ramps, (8) median type, (9) median opening type (i.e., full opening, directional opening, no opening), (10) median end treatment (i.e., no left-turn lane, one left-turn lane, two left-turn lanes), (11) presence of pedestrian or bicycle facilities (i.e., sidewalk, crosswalk, bike lane) and bus stops, (12) interchange type (i.e., diamond, partial

cloverleaf, full cloverleaf, trumpet, etc.), and (14) type of ramp control (i.e., signalized, stop controlled, free flow).

Crash diagrams were developed to illustrate the safety issues identified in the context of the study site. Crash types coded for this study included rear end, head on, angle, sideswipe same direction, sideswipe opposite direction, backed into, other/unknown, and bicycle/pedestrian crash. Based on the safety issues identified as well as the driveway and corridor/interchange area characteristics, a set of potential countermeasures were identified for the study site. The case studies are described in the next section.

ANALYSIS RESULTS

This section presents case studies, including an illustrative overview of crashes in the study site, and a representative example of a crash diagram for each study site. Case study sites include (1) John Young Parkway at West Colonial Drive in Orlando, (2) East Bay Drive (State Road 686) in Largo, (3) West Tennessee Street in Tallahassee, (4) North West 103rd Street and West 49th Street (State Road 932) in Hialeah, (5) West Hallandale Beach Boulevard at I-95 Interchange area in Hallandale Beach, and (6) Scenic Highway at I-10 Interchange area in Escambia County. Florida DOT uses a context classification system to help guide roadway design based on land use context. The context classification for these case examples were C3C-Suburban Commercial and C4-Urban General. According to the FDOT Context Classification Guide, these context classifications are defined as follows (FDOT, 2020, p. 8):

- C3C-Suburban Commercial - Mostly non-residential uses with large building footprints and large parking lots within large blocks and a disconnected or sparse roadway network
- C4-Urban General - Mix of uses set within small blocks with a well-connected roadway network. May extend long distances. The roadway network usually connects to residential neighborhoods immediately along the corridor or behind the uses fronting the roadway.

Florida DOT also uses access classifications to guide access design and permitting on the state highway system. The FDOT access class for each of these roadways is either Access Class 3, Access Class 5, Access Class 6, or Access Class 7. These access classes are described in the FDOT Multimodal Access Management Guidebook as follows (FDOT, 2023, p. 35):

- Access Class 3 roadways are controlled access facilities where direct access to abutting land is controlled to maximize the operation of the through traffic movement.
- Access Class 5 roadways are controlled access facilities where adjacent land has been extensively developed and where the probability of major land use change is not high. These roadways are distinguished by existing or planned restrictive medians.
- Access Class 6 roadways are controlled access facilities where adjacent land has been extensively developed, and the probability of major land use change is not high. These roadways are distinguished by existing or planned non-restrictive medians or centerlines.
- Access Class 7 roadways are controlled access facilities where adjacent land is generally developed to the maximum feasible intensity and roadway widening potential is limited. This classification is assigned to roadway segments where there is little intent or opportunity to provide high-speed travel.

John Young Parkway at W. Colonial Drive (Orlando)

John Young Parkway (State Road 423), in the City of Orlando, is a 6-lane divided highway. The study segment includes a directional median opening directly adjacent to a driveway for a

shopping center. The roadway is classified as an urban principal arterial with an average annual daily traffic (AADT) of 48,000 vehicles per day. It has sidewalks on both sides and no bicycle lanes. John Young Parkway is classified as an Access Class 3 and has a posted speed of 45 mph. The FDOT access connection spacing requirement for this access classification is 440 feet and one-half mile for full median openings and signals. The context classification for this segment is C3C-Suburban Commercial.

This high-volume major intersection of two six-lane arterial roadways has significant queues. Both driveways are disrupted by queueing traffic, with the connection closest to the signalized intersection with West Colonial Drive (within 240 ft. of the intersection) completely blocked by the queues. This decreases the egress capacity of the driveways and poses potential safety hazards if exiting vehicles attempt to cross several lanes of queueing traffic to turn left at the intersection.

Two radial design commercial driveways are of interest on the east side of the segment, shown as driveway “a” and “b” in Figure 1 along with an example crash diagram. They are unchannelized driveways serving more than 50 parking spaces and comparable to FDOT Category D driveways (shopping center, vehicle trips/hr. 121-400) per the FDOT Access Management Guide (FDOT, 2019). One commercial driveway (driveway “a”) is directly served by a directional median opening with a 290-ft-long left-turn lane. A transit stop is located directly across the road from this connection. Driveway “b” is about 309 feet north of driveway “a” and within 240 feet of the signalized intersection.

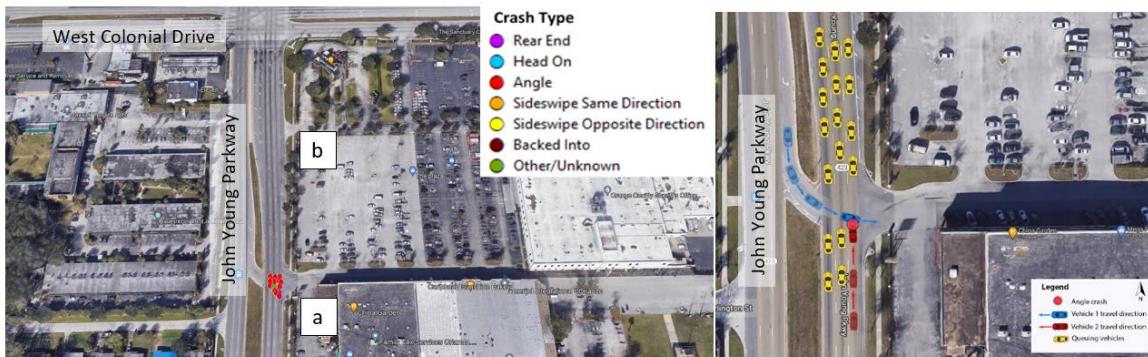


Figure 1. Overview of John Young Parkway study area.

The focus of the crash analysis was driveway “a,” which is served by the directional median opening. This driveway has a throat length of about 35 feet and is located approximately 633 feet south of the signalized intersection of John Young Parkway and West Colonial Drive. A total of 15 crashes were reported between 2015 and 2019. Most of these crashes involved left turns from the directional opening into the driveway of a shopping center and one involved a crossing maneuver turning right out of the driveway into the far left-turn lane.

The safety issues associated with the design of commercial access at this include (1) lengthy queues at the signalized intersection disrupted driveway operations and reduced visibility of vehicles crossing the through lanes to enter and exit the connection, (2) conflicts occurring between through moving vehicles and those attempting left-in entry to driveway from directional opening or right turn exit, and (3) queueing vehicles allows the turning vehicles to cross through lanes resulting in crashes in one of the through lanes, sometimes involving more than one vehicle, as the vehicles proceed into or out of the driveway.

East Bay Drive (Largo)

East Bay Drive (State Road 686) is in Largo (Pinellas County). The area of interest for this study was a 0.35-mile segment of East Bay Drive west of Belcher Road. The segment is a 6-lane divided highway with two median openings — a directional and a full median opening. The posted speed limit is 45 mph, and the AADT is 56,000 vehicles per day. East Bay Drive has sidewalks on both sides, no bicycle lanes, and three bus stops, two to the north of the segment and one to the south of the segment.

East Bay Drive is classified as an Access Class 7. The FDOT access connection spacing requirement for this access classification is 125 feet, 330 feet for directional median openings, 660 feet for full median openings, and a minimum of one-quarter mile for signal spacing. There are approximately 19 access connections along this .35-mile segment. The functional classification for this roadway is urban minor arterial and the context classification is C3C-Suburban Commercial.

An overview of the study area crash locations and an illustration of a head on crash in the median opening involving two vehicles exiting opposing driveways are shown in Figure 2. Driveway “a” is a full-movement driveway with a physical separator and provides primary access to a shopping center on the northern side of East Bay Drive. The driveway width is approximately 40 feet, and the throat length is approximately 77 feet. The nearest driveway is approximately 142 feet to the east.

Driveway “b” is a flared driveway providing access to a shopping center on the southern side of East Bay Drive. It is a large shopping center and can be categorized as FDOT driveway category “D,” which is meant to serve 1,201 vehicle trips per day and 121-400 vehicle trips per hour (FDOT, 2019). Driveway “b” is a full-movement driveway with a painted separator and provides direct access to the shopping center. The driveway width is approximately 30 feet, and the throat length is approximately 55 feet. The nearest driveway is approximately 120 feet to the west.

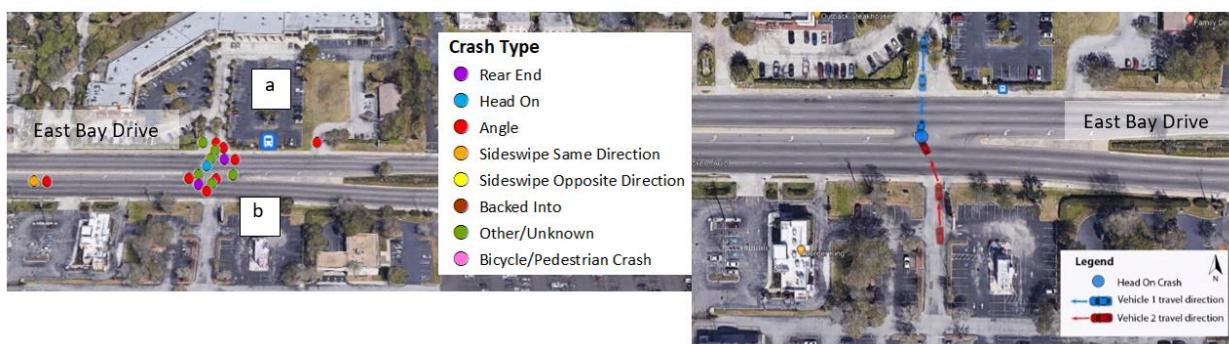


Figure 2. Overview of East Bay Drive study area.

Between 2015 and 2019, 40 driveway-related crashes were recorded along this segment. A mix of crash types was identified along this segment, including rear-end, head-on, angle, sideswipe same direction, and other/unknown. As shown in Figure 2, many of these crashes are clustered near the driveway of the two adjacent shopping centers.

The safety issues associated with the design of commercial access at this location may be summarized to include (1) the location of the median opening in relation to the shopping center

driveways results in conflicts between through moving vehicles and those attempting to enter or exit the driveways from/to the full median opening, (2) the median opening poses concerns when multiple vehicles are attempting to exit the adjacent driveways and maneuver into the median opening simultaneously, as illustrated in Figure 2, (3) conflicts occur between vehicles entering the driveway from the median opening while pedestrians attempt to cross the driveway, resulting in bicycle/pedestrian crashes and/or additional conflicts with through moving vehicles, and (4) drivers looking left while attempting to exit the driveways do not notice bicyclists and/or pedestrians crossing the driveway resulting in bicycle/pedestrian crashes.

Tennessee Street (Tallahassee)

West Tennessee Street in Tallahassee (Leon County) is a 6-lane divided highway with non-traversable medians with turn lanes at the intersections. The posted speed limit is 35 mph, and the AADT is 37,000. Figure 3 shows a segment of West Tennessee Street west of Stadium Drive that was of interest to this case study and an illustration of an angle crash on West Tennessee Street that was the result of queuing traffic from the coffee shop to the east of Caliark Street (driveway “b”). The segment has sidewalks on both sides of the roadway and sharrows on the south side of the roadway. There are two bus stops to the south of the segment.

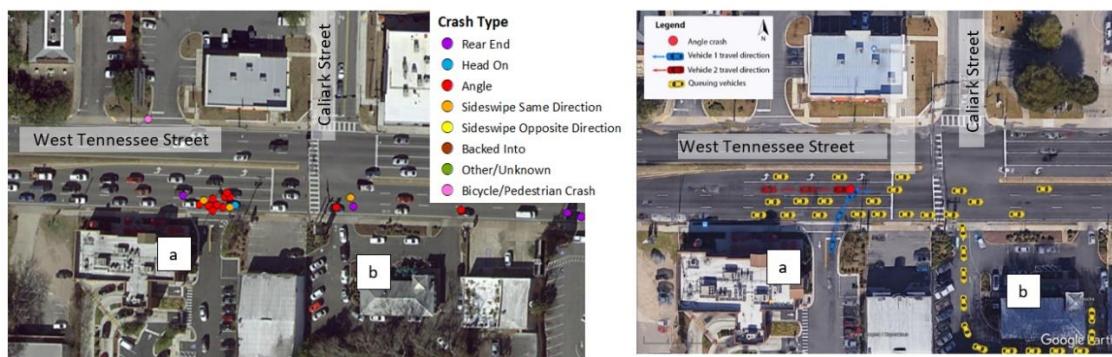


Figure 3. Overview of crashes on W. Tennessee Street study area.

This roadway is classified as Access Class 7 which has a minimum access connection spacing requirement of 125 feet, 330 feet for directional median openings, 660 feet for full median openings, and a minimum of one-quarter mile for signal spacing. There are approximately 22 driveways/access points along this segment. The functional classification for this roadway is urban minor arterial and the context classification is C4- Urban General.

Between 2015 and 2019, 30 driveway-related crashes were recorded along the segment. Crash types included rear-end, angle, head-on, sideswipe same direction, and other/unknown. Most crashes were identified as being angle or rear-end crashes.

The case study focused on a cluster of crashes at the intersection of West Tennessee Street and Caliark Street. These crashes are located at the fast-food restaurant (Chick-fil-A, driveway “a”) and are influenced by queues at the signalized intersection as well as backups due to inadequate on-site circulation at the upstream coffee shop (Starbucks, driveway “b”). A lounge with wide-open access separates the two uses.

The spacing between these three driveways, the volumes associated with the land uses, and their proximity to the intersection of West Tennessee Street and Caliark Street result in multiple

conflicts. The safety issues associated with the design of commercial access at this location include (1) lengthy queues at the coffee shop interfere with through traffic movement and disrupt nearby driveway operations, and (2) conflicts occur between through-moving vehicles and those exiting the fast-food restaurant, attempting to access the left turn lane at the intersection.

State Road 932 (Hialeah)

A 1.5-mile segment of North West 103rd Street and West 49th Street (State Road 932) in Hialeah (Miami-Dade County) between North West 77th Court and Ludlam Road was included as a case study site. This roadway segment is a 6-lane divided highway and passes under a flyover for the Palmetto Express Lane. The roadway has an AADT of 45,000 vehicles per day. Sidewalks are continuous and are visible along the north and south sides of the segment and there are no visible bike lanes. There are 13 bus stops along the roadway segment.

West 49th Street east of the Palmetto Express Lane is an Access Class 5 (restrictive) and has a posted speed limit of 40 mph. The FDOT access connection spacing requirement for this access classification is 245 feet, 660 feet for directional median openings, 1,320 feet for full median openings, and a minimum of one-quarter mile for signal spacing. North West 103rd Street west of the Palmetto Express Lane is an Access Class 6, non-restrictive, and has a posted speed limit of 40 mph. The FDOT access connection spacing requirement for this access classification is 245 feet and a minimum of one-quarter mile for signal spacing.

There are approximately 88 access connections along the observed segment of North East 103rd Street and West 49th Street. Between 2015 and 2019, 189 driveway-related crashes were recorded along this segment of the corridor (30 to 50 crashes per mile). Crash types included rear-end, angle, head on, sideswipe same direction, sideswipe opposite direction, and other/unknown. Several of the crashes (all crash types) were recorded in the same location and are therefore layered on the diagrams in Figure 4, which also shows an illustration of a rear-end crash in a driveway on West 49th Street.

The driveway density and volume of crashes along this roadway segment demonstrate how driveway density and proximity to entrance ramps influence crash rates.



Figure 4. Overview of N.E. 103rd Street and W. 49th Street.

The case study focused on a cluster of crashes near a driveway for a mall to the east of the entrance ramp for the Palmetto Express Lane. This land use associated with this driveway is most closely aligned with FDOT driveway category F or G, which are meant to serve up to 30,001 vehicle trips per day or more and up to 3,001 vehicle trips per hour (FDOT, 2019). It is a

full-movement radial design driveway with a physical separator. The driveway width is approximately 50 feet, and the throat length is approximately 140 feet. The nearest driveway is approximately 400 feet to the east.

Collisions near this driveway largely included rear-end and angle crashes. The rear-end crashes typically involved two vehicles exiting the driveway. The angle crashes involved vehicles exiting the driveway and those traveling west in the through lane or maneuvering into the entrance ramp for the Palmetto Express Lane.

The safety issues associated with the design of commercial access at this location include (1) vehicles exiting the driveway suddenly stopping to yield to through traffic or those accessing the entrance ramp for the express lane. In some instances, drivers back up to remove themselves from the roadway and back into the vehicle behind them, and (2) vehicles exiting the mall attempt to access the left turn lane and collide with through traffic traveling west and/or vehicles accessing the entrance ramp for the express lane.

West Hallandale Beach Boulevard at I-95 Interchange (Hallandale Beach)

West Hallandale Beach Blvd is a 6-lane divided highway east of the I-95 interchange and a 4-lane divided highway west of the interchange. The I-95 interchange is a diamond interchange with yield-controlled on-ramps and signalized off-ramps. Hallandale Beach Blvd is a 6-lane divided highway east of the interchange and a 4-lane divided highway west of the interchange. It is classified as a state urban principal arterial with an AADT of 47,532 vehicles per day. There are sidewalks and bicycle lanes on both sides of the roadway.

The context classification for this area is C4-Urban Commercial and C3C-Suburban Commercial. The commercial area in the immediate vicinity of the interchange includes gas stations, fast food restaurants, a hotel, auto parts, and a variety of other low-density commercial uses. Railroad tracks are located west of the interchange with a frontage road providing access to commercial uses between the railroad tracks and I-95. Due to traffic delays as well as safety issues in the area, the interchange was the subject of an interchange operational analysis report and subsequent improvements to the off-ramp configurations and signalization (FDOT, 2016).

The segment of West Hallandale Beach Blvd that intersects with the interchange is classified as an Access Class 5 and has a posted speed of 40 mph. The FDOT access connection spacing requirement for this access class is 245 feet and $\frac{1}{4}$ mile for full median openings and signals. Because this is an interchange area, FDOT Rule 14-97.003(3)(h)2 calls for a minimum of $\frac{1}{2}$ mile between the first full median opening and the end of the off-ramp taper and connection spacing of 440 feet where the posted speed is 45 mph or less. Interchange areas are defined in the rule as the lesser of $\frac{1}{4}$ mile from an interchange facility or up to the first intersection with an arterial road. The driveway access around the interchange is below FDOT access spacing standards.

Between 2015 and 2019, approximately 14 crashes were reported involving vehicles exiting or entering commercial driveways near the interchange ramp tapers (see Figure 5). Crash types included angle crashes, rear-end collisions, and bicycle/pedestrian crashes. Typical bicycle crashes occur as drivers look left while attempting to exit the driveways and fail to notice the cyclists crossing the driveway from the opposing direction as illustrated in Figure 5. In both observed examples, the cyclists were traveling the wrong way on the bicycle lane.

The location of driveways in relation to the interchange ramp results in conflicts between vehicles entering and exiting driveways closest to the interchange ramps. The safety issues associated with the design of commercial access at this location include (1) drivers looking left

to exit the driveway failing to notice passing vehicles in the right lane braking too slow in advance of the interchange signal or possibly turning into a driveway/street connection or enter the on-ramp, (2) drivers looking left to exit the driveway fail to notice cyclists traveling the wrong way on the bicycle lane, (3) vehicles traveling toward the interchange collide with vehicles turning to enter driveways. Following drivers expect the vehicles to continue onto the on-ramp rather than suddenly slow to turn into a driveway, and (4) vehicles merge suddenly to enter the on-ramp or a driveway causing rear-end collisions with the following vehicle.

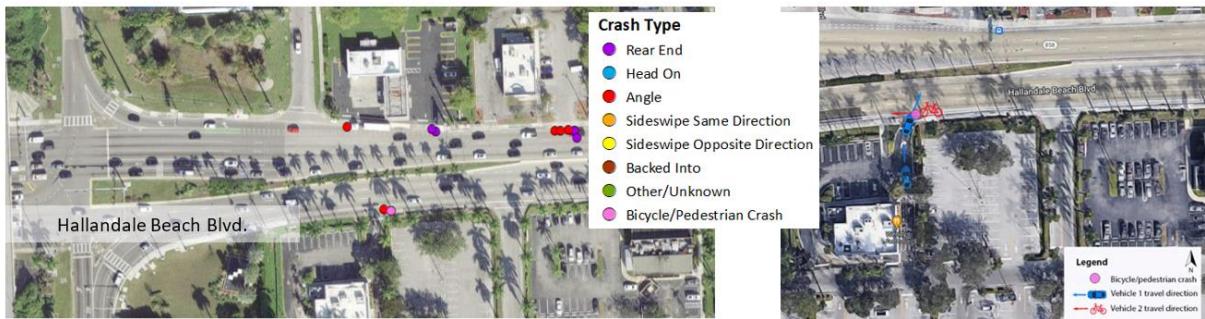


Figure 5. East side of I-95 and Hallandale Beach Blvd interchange.

Scenic Highway at I-10 Interchange (Escambia County)

Scenic Highway is a three-lane divided highway at Interstate Highway 10 (I-10) with dual left turn lanes for the southernmost on-ramps and on one left-turn lane for the northernmost on-ramp. The I-10 interchange at this location is a partial cloverleaf interchange with signalized on and off ramps to the south and a combination of signalized with yield control on- and off-ramps to the north. Scenic Highway is a three-lane divided highway at the interchange with dual left-turn lanes for the southernmost on-ramps and on one left-turn lane for the northernmost on ramp. It is classified as an urban principal arterial-other south of the interchange and an urban minor arterial north of the interchange with an AADT of 17,400 vehicles per day. Sidewalks and bicycle lanes are present in the vicinity of the interchange. Ramp AADT ranges between 5500 and 6800 vehicles per day.

Scenic Highway is classified as an Access Class 6 (no restrictive median) near the interchange with a posted speed of 45 mph. The FDOT access connection spacing requirement for this access class is 245 feet. Because this is an interchange area, FDOT Rule 14-97.003(3)(h)2 calls for a minimum of $\frac{1}{2}$ mile between the first full median opening and the end of the off-ramp taper and connection spacing of 440 feet where the posted speed is 45 mph or less. Interchange areas are defined in the rule as the lesser of $\frac{1}{4}$ mile from an interchange facility or up to the first intersection with an arterial road. The context classification for this area is C3C-Suburban Commercial.

The interchange area was reconstructed in late 2016. The timing of this reconstruction overlaps with the crash analysis period, offering insight into how the access changes impacted safety. Before reconstruction, the first access to the plaza from the interchange was offset slightly from the on-ramps to I-10. No median was present at this time, thereby allowing left turns into and out of the site, as well as crossing maneuvers at the hotel driveway. Reconstruction also changed the access design north of the interchange. A median was installed, eliminating left

turns into the gas station driveway, and requiring drivers to make a U-turn at the intersection with Northpointe Parkway just north of the gas station driveway.

Between 2015 and 2019, a total of 19 commercial driveway-related angle crashes were recorded in the immediate vicinity of the interchange illustrated in Figure 6, which includes an example of a crash involving left-turn ingress. Only one of these crashes (right-turn out of the plaza) occurred following reconstruction in 2016 and the installation of the median. Some lingering construction may, however, have been an issue in this crash due to barrels impeding sight distance. This crash type is also mitigated by the median.



Figure 6. Configuration of access and crashes before median installation in 2016.

The safety issues associated with the design of commercial access at this location include (1) closely spaced commercial driveway access in proximity to interchange ramps and left turn access in the functional area of an interchange creates overlapping conflicts, and (2) queues at the intersection with the interchange obscure visibility as queueing vehicles allow vehicles exiting right out of the commercial plaza driveway to enter the left-turn lanes for the interchange ramp.

CONCLUSIONS

This paper discusses how case studies can be used to gain a better understanding of safety issues identified through high-level analysis. In the study described in this paper, tools such as GIS and Google Earth were used to identify driveway and roadway characteristics influencing crash frequency and severity by assessing individual crashes on high-crash corridors and in crash clusters. Crash reports and other relevant documents were consulted to verify findings to ensure that the most up-to-date information on the study area was being used. For example, through case studies, the project team was able to identify reconstruction in the Scenic Highway at the I-10 Interchange case study. This finding provided an opportunity to expand the study scope and evaluate the before and after effects of reconstruction within the project time period.

The level of analysis described in this paper is useful for cross-checking large datasets with the potential for inaccuracies. The review of crash reports for each of the case studies revealed that some of the crashes were improperly geolocated, and some crash reports did not identify which commercial driveway locations were involved in the crash. For the case study analysis, these outliers were manually relocated in the ArcGIS database, given their limited numbers. Due to the size of the full dataset (over 9,000 crashes), all of the crashes were not geolocated.

Considering the large number of crashes included in the analysis, these inaccuracies impose only a minor or negligible influence on the research findings. However, the improper location of these crashes masked the intensity of crash clusters at certain driveways.

Although conducted in Florida, the approaches, lessons learned, and findings of this study are relevant to understanding commercial driveway safety in the U.S. and internationally. Broadly applicable strategies for project development such as ensuring that adequate space is provided on-site for circulation and queueing, avoiding approval of driveways near an intersection where the establishment creates a potential for long queues that may spill back onto roadways and create safety issues, identifying additional measures are needed to ensure interparcel cross access is provided to increase corner clearance and/or allow left-turn access at the signal, mitigating commercial driveway access near interchange ramps through medians and other measures, and providing, additional educational measures and signage could help to mitigate this issue, as would painted bike lanes and other measures to make it more apparent that bicyclists may be present. Furthermore, post-crash strategies such as providing appropriate training for those involved in obtaining and geolocating crash report data support effective safety analysis. The Access Management Manual, state and local access management guides, and other sources can be used to provide countermeasures on a case-by-case basis.

REFERENCES

- FDOT. (2016). I-95 at Hallandale Beach Boulevard: Interchange Operational Analysis Report (IOAR). District 4, Ft. Lauderdale, Florida. <https://fip.fdot.gov/Studies/Details/98>.
- FDOT. (2019). Access management guidebook. Tallahassee, Florida.
https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/planning/systems/documents/fdot-access-management-guidebook---nov-19.pdf?sfvrsn=c5aa6e5_4.
- FDOT. (2020). FDOT context classification guide.
<https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/roadway/completestreets/files/fdot-context-classification.pdf>.
- FDOT. (2023). Multimodal access management guidebook.
https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/planning/systems/systems-management/document-repository/access-management/fdot-multimodal-access-management-guidebook_oct2023.pdf?sfvrsn=95abdf85_1.

Effect of Crash Barriers on Driver Behavior

Hasara Senevirathna¹; Nipuna Seneviratne²; Vasantha Wickramasinghe³;
and Sunanda Dissanayake⁴

¹Dept. of Civil Engineering, Univ. of Peradeniya, Sri Lanka. Email: e17322@eng.pdn.ac.lk

²Dept. of Civil Engineering, Univ. of Peradeniya, Sri Lanka.

Email: 0713992924nipuna@gmail.com

³Dept. of Civil Engineering, Univ. of Peradeniya, Sri Lanka. Email: vskw@eng.pdn.ac.lk

⁴Kennesaw State Univ., Kennesaw, GA. Email: sunanda@k-state.edu

ABSTRACT

Road accidents pose public health concerns, incurring substantial costs. While crash barriers are recognized for preventing accidents, their impact on driver behavior in mild road curves remains understudied. This study addresses the gap by investigating crash barrier influence on vehicle speed and lateral displacement. A Class “AB” road with consecutive mild curves, one with a crash barrier and one without, underwent controlled analysis. Drones captured speeds and lateral positions. SPSS t-tests revealed significant differences ($P < 0.05$) in speeds and displacements at curves with and without barriers. Regression analysis showed road conditions did not significantly affect speeds; middle and exit speeds depended on entering speed and crash barrier presence. Vehicle speed at curves with barriers decreased, indicating heightened safety perception. Vehicles approached the centerline with barriers, posing head-to-head collision risks. Recommendations include widening lanes near crash barriers. In conclusion, crash barriers can have both positive and negative effects on driver behavior, requiring careful consideration.

INTRODUCTION

Road traffic injuries pose a significant global threat, claiming around 1.35 million lives in 2020 alone and standing as the leading cause of death for those aged 5-29 years (World Health Organization, WHO). In the quest for safer roads, crash barriers play a crucial role by preventing head-on collisions, sideswipes, and rollovers, thereby reducing the severity of injuries.

However, the effectiveness of crash barriers is closely tied to various factors, especially driver behavior. Antonson et al. (2013) discovered that the presence of crash barriers can lead to increased speeds and altered positioning on curves. Driver distractions and fatigue, highlighted by Ben-Bassata and Shinar (2011), add complexity, emphasizing the need to address both physical and psychological aspects.

Jin et al. (2018) brought another layer to the discussion, revealing that crash barriers might unintentionally contribute to increased driver distraction. Understanding this intricate relationship is crucial for refining crash barrier designs and implementing comprehensive safety measures. Beyond preventing collisions, crash barriers strategically positioned along roadways act as protectors, creating a buffer against potential hazards and committing to broader public safety goals.

Different types of crash barriers, like W-beam guardrails, cable barriers, and Jersey barriers, offer unique strengths and weaknesses. Selecting the right design involves considering specific roadway conditions and safety goals. However, the mere presence of crash barriers can influence

driver behavior, necessitating a deeper understanding to optimize their effectiveness. This research aims to fill this gap by investigating how crash barriers impact driver behavior, particularly in mild curves on single-path two-way roads. The study seeks to analyze acceleration/deceleration behavior and lateral displacement of vehicle trajectories, providing valuable insights to enhance overall road safety.

LITERATURE REVIEW

Road traffic injuries pose a significant global threat, resulting in a considerable number of fatalities and disabilities each year. In 2020 alone, an estimated 1.35 million lives were lost in road traffic crashes, with individuals aged 5-29 years facing the highest risk (World Health Organization, 2020).

Crash barriers, designed to prevent vehicles from breaching road boundaries, play a pivotal role in bolstering road safety. The effectiveness of these barriers is closely tied to various factors, notably driver speed. Studies by Antonson et al. (2013) reveal that crash barriers may lead to increased speeds on curves, underscoring the importance of comprehending the relationship between speed and crash barriers for a holistic road safety approach.

Beyond speed, driver distractions and fatigue present additional challenges. Ben-Bassata and Shinar (2011) highlight that crash barriers could contribute to reduced driver attention and increased risk-taking behavior. This emphasizes the necessity of addressing not only the physical aspects of barriers but also psychological factors to optimize road safety measures.

Interestingly, research by Jin et al. (2018) suggests that crash barriers might unintentionally heighten driver distraction. This underscores the interconnected nature of road safety elements and emphasizes the need for a comprehensive approach to balance protective benefits with potential contributing factors to distraction.

Strategically positioned along roadways, crash barriers act as crucial guardians, absorbing impact forces, and redirecting vehicles to prevent severe accidents. Serving as a first line of defense, these barriers shield motorists, passengers, pedestrians, and cyclists from the full impact of collisions and road departures.

Concrete barriers and steel guardrails have demonstrated positive effects in reducing speed and lateral displacement in horizontal curves. However, cable barriers showed limitations, emphasizing the importance of selecting barrier types based on their proven effectiveness for enhancing driver behavior and overall road safety.

Comparative studies consistently demonstrate the positive impact of crash barriers on reducing speed variation, improving lateral positioning, and decreasing crash rates in horizontal curves. Methodologies employed in these studies, including field research, driving simulations, and naturalistic observations, provide a comprehensive understanding of how crash barriers influence driver behavior.

Despite these valuable insights, research gaps persist. The lack of comprehensive studies addressing speed and lateral displacement simultaneously and limitations such as small sample sizes emphasize the need for larger-scale, real-world studies. These would enhance our understanding of the influence of crash barriers on driver behavior in horizontal curves.

In conclusion, crash barriers are instrumental in mitigating the severity of accidents and protecting road users. However, a nuanced understanding of their interaction with driver behavior, speed, and other factors is crucial for optimizing their effectiveness and improving overall road safety outcomes.

METHODOLOGY

The study aimed to analyze how crash barriers impact driver perception, reaction, vehicle speed, and lateral position. It consisted of four phases: site selection, data collection (road measurements and external factors), manual data extraction aided by software, and data analysis involving hypothesis testing and regression model development. The goal was to assess factors influencing vehicle speed and lateral position with and without crash barriers for different vehicle types.

Site selection involved two phases: a general inspection and a final selection. The general inspection considered crucial road characteristics, and a checklist guided site selection. Multiple sites were assessed in Sri Lanka's Central Province, with key criteria including two bends in the same direction, one with crash barriers and one without. The selected site along the Gampola Nawalapitiya road met these criteria, allowing for an effective comparison of vehicle behavior in bends with and without crash barriers. Criteria considered included minimal curvature variation, consistent super elevations, and uniform external factors like street lamps or signboards on both bends. Both curves have a radius of 36 meters, and the width of each lane is 3 meters. All the data utilized for constructing these models was collected during daytime hours, specifically between 11:00 AM and 1:00 PM on two weekdays. It is essential to note that the model may exhibit some deviations when applied to nighttime conditions, weekends, or public holidays. Figure 1 shows the selected site location with coordinates 7.091761 latitude and 80.555533 longitude.



Figure 1. Selected site location for the trial survey

Then, Data extraction began with synchronizing drone and camcorder footage, identifying a vehicle, and taking snapshots. Using Photoshop for alignment and AutoCAD for tracing, a clear vehicle trajectory was obtained as shown in Figure 2.

Then, 600 meters of road sections were selected from both curves and the area between the roadside line and the road center line was taken to find out the vehicle lateral displacement. Also, the area of the considered lane was considered.

Speed at the entry of the curve, at the middle of the curve, and the exit of the curve for both curves with and without crash barriers were taken as speeds. Also, the lowest speed that the

vehicle moves after entering the curve was taken as the middle speed of that particular curve. Since the trajectory was taken by the snapshots at a rate of one second, the distance between two adjacent vehicle tracers straight away gives the average speed in between those traces in ms^{-1} .

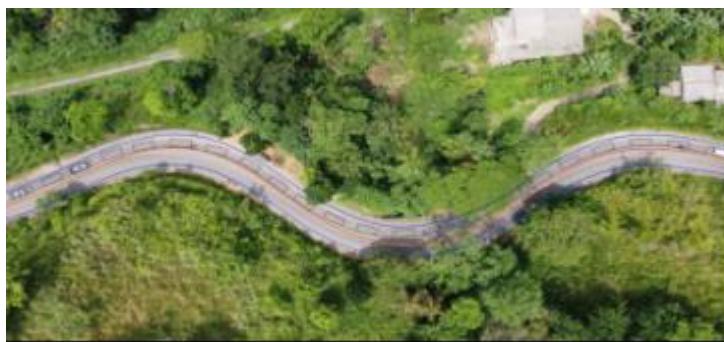


Figure 2. Vehicle Trajectory Mapping

A hypothesis test is a statistical method that is used to determine whether there is a significant difference between the two groups. The significant level was taken as 5%. It was conducted to examine the significance of the collected data from the two locations.

Once the significance check was performed a regression analysis was done to generate a model. The regression analysis was done to develop a linear regression model from the obtained variables. Model fitting was done for the three models by taking middle speed, exit speed, and cutting area percentage as independent variables and entry speed, vehicle type, presence or absence of crash barrier, road condition, and opposite lane vehicles as dependent variables. The coefficients of those equations could be found by the regression analysis. Then individual P values, multiple R^2 values, final P value of the models were checked for the three models in the model fitting step. Once the above parameters were checked the model fitting was completed.

After that, Model validation was done for the test data set to check if the generated model was acceptable and didn't violate any statistical principles. The Durbin-Watson Linearity check was conducted to check if the model can be validated. Once the model validation was completed, the model was introduced.

RESULTS AND DISCUSSION

Speed data for vehicles at entry, middle, and exit points in both curves, with and without crash barriers, were collected. The assumption was made that there were no external influences on driver behavior as the vehicles traveled through the two curves. In analyzing vehicle movement through the curves, uniformity was maintained in road conditions, vehicle category, vehicle conditions, and curve and road characteristics. This consistency was achieved because the curves were located near each other, and one vehicle traversed both similar curves in the same direction.

For 97 vehicles, graphs were created depicting vehicle speed at entry, middle, and exit points for both scenarios with and without crash barriers. Subsequently, the following graphs (Figure 3, Figure 4, Figure 5, and Figure 6) were examined to identify any clear differences between the two scenarios.

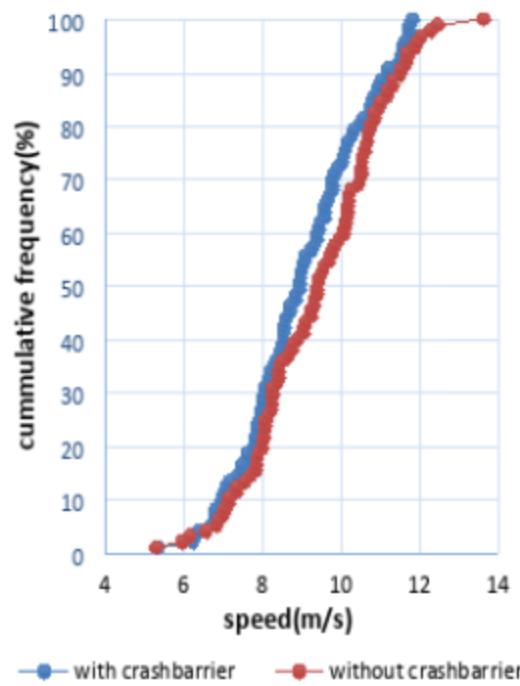


Figure 3. Cumulative frequency Vs speed at middle of curve

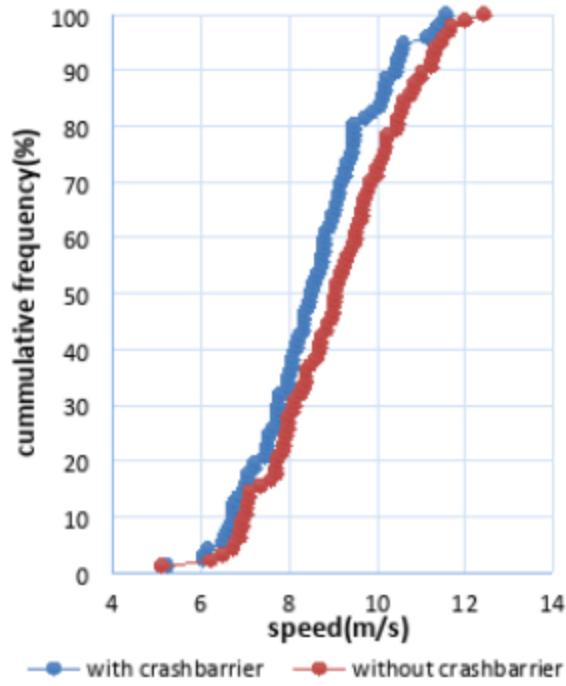


Figure 4. Cumulative frequency Vs speed at enter of curve

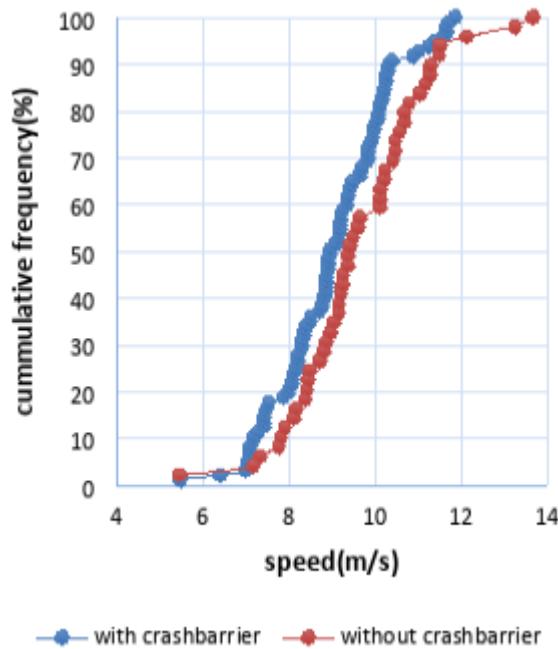


Figure 5. Cumulative frequency Vs speed at exit of curve.

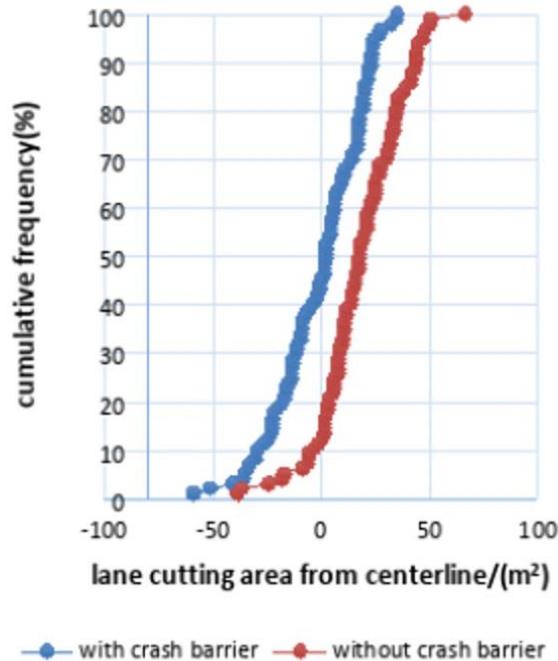


Figure 6. Cumulative frequency Vs opposite lane cutting area from centerline

Table 1. T-test results for vehicle speeds

Location of the curve	At entry		At middle		At exit	
	Yes	No	Yes	No	Yes	No
15 th percentile speed(m/s)	7.3	7.8	7.0	7.2	7.4	8.2
50 th percentile speed(m/s)	9.0	9.4	8.5	9.0	9.0	9.3
85 th percentile speed(m/s)	10.8	11.1	10.1	10.6	10.2	11.2

According to the results in Table 1, the presence of a crash barrier on the curve tends to reduce vehicle speed compared to the curve without it. Graphs were also generated for the lateral displacement positions of 97 vehicles in both scenarios.

Table 2. T-test results for vehicle speeds

The existence of a crash barrier	Lane cutting area/(m ²)	
	Yes	No
15 th percentile	-22	2
50 th percentile	3	18
85 th percentile	22	41

As shown in Table 2, Positive values indicate the vehicle traveled in the right lane without crossing the centerline into the opposite lane. Conversely, negative values indicate the vehicle entered the opposite lane by cutting across the centerline of the road. Analyzing this table reveals the point along the vehicle's path that is either closest to the centerline or farthest away from it after crossing. Out of 97 vehicles, 43% tend to veer into the opposite lane when there is a crash barrier. Conversely, when there is no crash barrier, 14% of the vehicles from the same pool demonstrate a tendency to move into the opposite lane. Figure 6 visually represents these findings. The evident distinction between scenarios with and without crash barriers in terms of speed and lateral displacement positions prompted the application of a paired two-tail t-test to determine the significance of these differences.

Hypothesis test

Hypothesis tests were individually conducted for the data sets extracted for vehicle speed at entry, middle, and exit points in curves, as well as for vehicle centerline cutting area. Utilizing the SPSS software, a two-tailed paired T-test was performed to assess the presence of a significant difference between the two scenarios. T-tests were performed for the extracted data of the vehicle's lateral position and vehicle speeds at the entry, middle, and exit of the curves to check the significant difference between the two scenarios. The results obtained from the software are as below in Table 3.

It was observed that the paired two-tailed p-value is lower than 0.001, implying that the significant value of the T-test is in the acceptable range which is less than 0.05. Hence null hypothesis was rejected and there is a significant difference between the vehicle speeds at entry,

middle, and entry of the curve concerning the two locations and there is a difference between the vehicle lateral position concerning the two locations.

Table 3. T-test results for vehicle speeds

Paired samples	P value
Vehicle speed at entry of curve	< 0.001
Vehicle speed in the middle of the curve	< 0.001
Vehicle speed at exit of curve	< 0.001
Cutting area from the centerline	< 0.001

Regression analysis

The analysis involved Multiple Linear Regression, considering four fundamental assumptions: (1) Linearity, (2) Homoscedasticity, (3) Independence, and (4) Normality. The results obtained from the software are detailed below in Table 4.

Vehicle speed at the middle of the curve

Table 4. Correlation between the variables

		Middle speed	Entry speed	Presence of crash barrier
Personal correlation	Middle speed	1.000	0.930	-0.181
	Entry speed (x_2)	0.930	1.000	-0.118
	Presence of crash barrier(x_1)	-0.118	-0.118	1.000
Sig. (1-tailed)	Middle speed		< 0.001	0.006
	Entry speed (x_2)	0.000		0.051
	Presence of crash barrier(x_1)	0.006	0.051	
N	Middle speed	194	194	194
	Entry speed (x_2)	194	194	194
	Presence of crash barrier(x_1)	194	194	194

As per Table 4, the p-value between the presence of a crash barrier (x_1) and the vehicle speed at the entry of the curve (x_2) is greater than 0.05. Additionally, the p-value between x_1 and x_2 is also greater than 0.05, indicating that x_1 and x_2 are independent of each other.

Table 5. Model Summary

R	R square	Adjusted R Square	Std.Error of the Estimate	R square change	F change	df1	df2	Sig,F change	Durbin-Watson
0.993	0.870	0.869	0.52544	0.870	640.615	2	191	<0.001	2.014

As shown in Table 5, the R-square value for the model is 0.87, surpassing the threshold of 0.5, indicating a robust model fit. Additionally, the model proves to be significant with a p-value

less than 0.05. The Durbin-Watson value for the linearity check falls within the acceptable range. A concise summary of the estimation results is provided in Table 5.

Table 6. Model Estimation

Model	Unstandardized coefficients		Standardized coefficients Beta	t	Sig	Collinearity Statistics	
	B	Std.Error				Tolerance	VIF
(constant)	1.122	0.233		4.822	<0.001		
Entry speed	-0.210	0.076	-0.073	-2.770	0.006	0.986	1.014
Presence of crash barrier	0.846	0.024	0.922	35.112	<0.001	0.986	1.041

As per Table 6, all the absolute t-values exceed 1.96, and the p-values are below 0.05, indicating that all model estimates are statistically significant at a confidence level exceeding 95%. Notably, the coefficient for x_1 is a negative value, signifying that the presence of a crash barrier results in a slowdown of the vehicle at the middle of the curve. Based on these findings, the final linear model can be calibrated as shown in the below Equation.

Vehicle speed at the middle of the curve (S_M),

$$S_M = 1.122 - 0.210 \times x_1 + 0.846 \times x_2$$

x_1 - Presence of crash barrier. (When Presence=1, Absent =0)

x_2 - Vehicle speed at entry of the curve(m/s)

Vehicle speed at the exit of the curve

Table 7. Correlation between the variables

		Exit speed	Entry speed	Presence of crash barrier
Personal correlation	Exit speed	1.000	0.905	-0.228
	Entry speed (x_2)	0.905	1.000	-0.118
	Presence of crash barrier(x_1)	-0.228	-0.118	1.000
Sig. (1-tailed)	Exit speed		<0.001	<0.001
	Entry speed (x_2)	0.000		0.051
	Presence of crash barrier(x_1)	0.001	0.051	
N	Exit speed	194	194	194
	Entry speed (x_2)	194	194	194
	Presence of crash barrier(x_1)	194	194	194

As per Table 7, the p-value between the vehicle speed at the entry of the curve (x_2) and the presence of a crash barrier (x_1) is greater than 0.05. Additionally, the p-value between x_1 and x_2 is also greater than 0.05, implying that x_1 and x_2 are independent of each other. Subsequently, Table 8 presents the model summary for the linear regression analysis.

Table 8. Model Summary

R	R square	Adjusted R Square	Std.Error of the Estimate	R square change	F change	df1	df2	Sig,F change	Durbin-Watson
0.913	0.833	0.832	0.60310	0.833	477.320	2	191	<0.001	2.106

As shown in Table 8, the R-square value for the model is 0.833, surpassing the threshold of 0.5, indicating a respectable model fit. The model is deemed significant with a p-value less than 0.05. Additionally, the Durbin-Watson value for the linearity check falls within the acceptable range, albeit the p-value is slightly less. A succinct summary of the estimation results is provided in Table 9.

Table 9. Model Estimation

Model	Unstandardized coefficients		Standardized coefficients	t	Sig	Collinearity Statistics	
	B	Std.Error	Beta			Tolerance	VIF
(constant)	1.929	0.267		7.219	<0.001		
Entry speed	0.827	0.028	0.890	29.914	<0.001	0.986	1.014
Presence of crash barrier	-0.362	0.087	-0.124	-4.154	<0.001	0.986	1.014

As per Table 9, all absolute t-values exceed 1.96, and all p-values are lower than 0.001, signifying that all model estimates are highly significant with a significance threshold of more than 95%. Moreover, the coefficient for x_1 is a negative value, indicating that the presence of a crash barrier results in a slowdown of the vehicle at the exit of the curve. Based on these results, the final linear model can be calibrated as shown in this equation.

Vehicle speed at the middle of the curve (S_E),

$$S_E = 1.929 - 0.362 \times x_1 + 0.827 \times x_2$$

x_1 - Presence of crash barrier. (When Presence=1, Absent =0)

x_2 - Vehicle speed at entry of the curve(m/s)

There is no significant effect from vehicle category, opposite lane traffic condition, and wet or dry road surface condition for vehicle speed at the middle and exit of the curve.

Vehicle centerline cutting area

According to Table 10, the presence of a crash barrier (x_1) and vehicle category (x_3) exhibit a significant linear relationship with the dependent variable, given that the p-value is less than 0.05. However, the p-value between x_1 and x_3 is greater than 0.05, indicating that x_1 and x_3 are independent of each other. Subsequently, Table 10 provides the model summary for the linear regression analysis.

The R-square value for the model is 0.575, indicating a decent model fit as it surpasses the 0.5 threshold. The model is considered significant with a p-value less than 0.05. Additionally, the

Durbin-Watson value for the linearity check falls within the acceptable range, although the p-value is slightly less. Table 11 provides a comprehensive summary of the estimation results for the model.

Table 10. Correlation between the variables

		Centerline cutting area	Vehicle category	Presence of crash barrier
Personal correlation	Centerline cutting area	1.000	-0.551	-0.524
	Vehicle category (x_3)	-0.551	1.000	0.005
	Presence of crash barrier(x_1)	-0.524	0.005	1.000
Sig. (1-tailed)	Centerline cutting area		<0.001	<0.001
	Vehicle category (x_3)	0.000		0.473
	Presence of crash barrier(x_1)	0.000	0.473	
N	Centerline cutting area	167	167	167
	Vehicle category (x_3)	167	167	167
	Presence of crash barrier(x_1)	167	167	167

Table 11. Model Summary

R	R square	Adjusted R Square	Std.Error of the Estimate	R square change	F change	df1	df2	Sig,F change	Durbin-Watson
0.758	0.575	0.569	13.77544	0.575	110.749	2	164	<0.001	1.999

Table 12. Model Estimation

Model	Unstandardized coefficients		Standardized coefficients Beta	t	Sig	Collinearity Statistics	
	B	Std.Error				Tolerance	VIF
(constant)	38.456	2.368		16.238	<0.001		
Entry speed	-9.561	0.889	-0.548	-10.759	<0.001	1.000	1.000
Presence of crash barrier	-21.802	2.132	-0.521	-10.226	<0.001	1.000	1.000

As indicated in Table 12, all the absolute t-values exceed 1.96, and all the p-values are lower than 0.001. This implies that all the model estimates are highly significant, surpassing a 95% significance threshold. Notably, the negative coefficient value of x_1 suggests that when there is a crash barrier, vehicles tend to slow down compared to scenarios without a crash barrier. Additionally, the negative coefficient for the vehicle category implies that heavy vehicles are more prone to cut the road centerline compared to light vehicles.

Vehicle centerline cutting area (C_U),

$$C_U = 38.456 - 21.802 \times x_1 - 9.561 \times x_3$$

x_1 - Presence of crash barrier. (When Presence=1, Absent = 0)

x_3 -Vehicle category (1-Car,2-LGV,3-HGV,4-Bus)

Sign convention for C_U ,

Positive values indicate that the vehicle traveled in the right lane and did not cross the centerline into the opposite lane. Negative values indicate that the vehicle entered the opposite lane by cutting across the centerline of the road.

CONCLUSION

This study examined the effect of the presence of a crash barrier on driver behavior in terms of vehicle speed and vehicle lateral position on a selected suburban road with other related road parameters either equal between the two locations or nearly insignificant. It was assumed that the driver's behavior was only affected by the presence of the crash barrier.

The study developed regression models to predict vehicle speed and vehicle lateral position as a function of the presence of a crash barrier. Two Regression models were developed to assess the relationship between vehicle speeds and crash barrier status. The models incorporated vehicle speed at the curve entry and crash barrier status as independent variables. The findings revealed a significant influence of both variables on vehicle speed, indicating that driver behavior in terms of speed is indeed influenced by the presence of crash barriers.

Similarly, a regression model was constructed to examine the association between vehicle lateral position and crash barriers. This model utilized vehicle type and crash barrier status as independent variables. The results demonstrated that both vehicle type and crash barrier status significantly impact vehicle lateral position, suggesting that drivers adjust their lateral position based on the presence of crash barriers. The models were found to be highly significant, suggesting that the presence of a crash barrier is a strong predictor of driver behavior.

The study also found that the presence of a crash barrier has a significant impact on both vehicle speed and vehicle lateral position. Vehicle speeds were found to be lower when a crash barrier was present, while light vehicle lateral positions were found to be closer to the center of the carriageway when a crash barrier was present. More heavy vehicles tend to the opposite direction when having crash barriers.

The study has several limitations. The data was collected manually, which could introduce errors. Also, the study only considered two weekdays from 11.00 am to 01.00 pm, which may not be representative of all days and all times. The study only considered one type of road and having 36 m radius of both curves, which may not be generalizable to other types of roads.

Despite these limitations, overall the study suggests that the presence of a crash barrier can have both positive and negative effects on driver behavior. While it can help to reduce vehicle speed and encourage drivers to center their vehicles, it can also lead to drivers veering into the opposite lane. So, it is important to weigh the potential benefits and risks of crash barriers carefully before installing them on a road.

RECOMMENDATIONS

In areas equipped with crash barriers, a recommendation is made for the widening of lanes to address the safety implications identified in our research. While crash barriers serve the purpose of slowing down vehicles and enhancing safety, our findings suggest that they concurrently contribute to increased lateral displacement, posing a potential safety challenge. The proposal for

lane widening aims to mitigate this heightened lateral displacement, ultimately fostering an improvement in overall road safety.

The study's regression model, which focused on a specific set of variables, emphasizes the need for an enhanced approach to understanding driver behavior. It is suggested that future models consider additional variables to comprehensively capture the multifaceted influences on driver behavior. To achieve this, an increase in sample size is recommended, allowing for a more expansive analysis of the various factors that may play a role in shaping driver behavior.

To achieve a more nuanced comprehension of driver behavior, particularly in curved road sections with crash barriers, a recommendation is made to conduct a questionnaire survey. This survey would delve into elements such as driver emotions, experience levels, gender differences, and other relevant factors. The exploration of these aspects through a comprehensive survey is expected to yield valuable insights into the intricacies of driver behavior in real-world scenarios, contributing to a more holistic understanding of the subject matter.

REFERENCES

1. Antonson, E. K., Tingvall, A., and Haworth, N. (2013). The effect of roadside barriers on driver behavior in curves. *Accident Analysis & Prevention*, 50, 237-244.
2. Ben-Bassata, E., and Shinar, D. (2011). The effect of roadside barriers on driver behavior in curves. *Accident Analysis & Prevention*, 43, 1538-1545.
3. Jin, L., Zhang, P., Chen, M., and Liu, Y. (2018). The effect of roadside barriers on driver distraction. *Transportation Research Part F: Traffic Psychology and Behavior*, 59, 202-211.
4. Montella, A., Aria, M., D'Ambrosio, A., Galante, F., Mauriello, F., and Pernetti, M. (2011). 'Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections about different perceptual cues', *Accident Analysis and Prevention*, 43(6), pp. 2072–2084.
5. Rajapaksha, D., and Wickramasinghe, W. (2020). The effect of crash barriers on driver behavior on curves. *Accident Analysis and Prevention*, 138, 105533.
6. Yang, J., Nie, Y., Zhang, J., and Ran, L. (2014). Driver behaviors on rural highways with and without curbs - a driving simulator-based study. *International Journal of Injury Control and Safety Promotion*, 21(2), 111-117.

Assessment of Stopping Sight Distance Restrictions on Vertical Alignments: A Case Study

Samantha Islam, Ph.D., P.E.¹; and Benjamin Burton²

¹Dept. of Civil, Coastal, and Environmental Engineering, Univ. of South Alabama, Mobile, AL.
Email: sislam@southalabama.edu

²Dept. of Civil, Coastal, and Environmental Engineering, Univ. of South Alabama, Mobile, AL.
Email: bdb1321@jagmail.southalabama.edu

ABSTRACT

In the study, a methodology was developed for identifying stopping sight distance (SSD) deficient vertical alignments on highways in Alabama and for assessing the safety impact of such deficiencies. The roadway geometric data were obtained from a database developed by Reiker Incorporated using their Curve Advisory Reporting Service (CARS) system. In addition to using CARS, this study used Google Earth and Google Maps to collect roadway geometric data. The geometric data from CARS were processed and exported to a computer-aided design (CAD) program to graphically measure available SSDs on vertical roadway alignments. The available SSDs were then compared with the recommended SSDs based on the posted speeds. A total of 29 roadway segments were identified that did not meet the minimum SSD requirements. The method developed in this study is expected to assist the responsible agencies in prioritizing deficient vertical alignments for upgrading to current design standard.

Keywords: Computer-aided Design, Highway Safety, Sopping Sight Distance, Vertical Curve.

INTRODUCTION

Many highways maintained by the state agencies were built prior to 1965. Since 1965, many roadway geometric design standards have changed due to reasons such as the change in the design vehicle dimensions. One of the major examples of the impacts of such changes is the change in the sight distance criteria for vertical curves. A driver needs sufficient sight distance to safely maneuver on a roadway (5). If a vertical curve does not meet the sight distance requirements for the design speed of the highway, it is necessary for the responsible agency to determine if this deficiency needs to be corrected since it may have impact on safety. Upgrading or rehabilitating highways that were built many years ago sometimes involves decisions regarding upgrading vertical alignments to meet the current design standard. However, available right-of-way and other factors may make upgrading such vertical curves to current design standards very expensive. The primary measure of design adequacy for vertical curves on a highway is the amount of sight distance provided in relation to the design speed of the highway. For example, the crest vertical curves are designed based on adequate stopping sight distance (SSD). Amount of stopping sight distance required on a crest vertical curve, for a given design speed, depends on driver's eye height and the height of the object that must be detected (1). In the 1940s, these heights were set at 4.5 ft and 4.0 inches respectively (1). In 1965, these values were changed to 3.75 ft and 6.0 inches to account for the decrease in the size of the vehicles (4). In 1984, decreasing vehicle sizes again prompted further reduction in driver height to 3.5 ft (2). As these design values were changed, the length necessary to provide recommended sight

distance also changed and as a result, vertical curves on many of the highways failed to meet the current standards. Currently, Alabama Department of Transportation (ALDOT) does not have a specified method to determine if a vertical curve meets current geometric design standard. In the study described in this paper, a methodology was developed for identifying geometrically deficient vertical curves needing upgrades to current design standard and for assessing the safety impact of such deficiencies. This methodology is expected to assist Alabama Department of Transportation (ALDOT) and other responsible agencies in prioritizing deficient vertical alignments for upgrading to current design standard.

METHODOLOGY

In this study, we developed a method for identifying vertical roadway alignments in Alabama with insufficient stopping sight distance and with at least one crash occurring on that roadway segment. The methodology for this research involved several steps, which are discussed in the following sub-sections.

Geometric Data Collection

In this study, we assembled a comprehensive dataset of geometrically deficient vertical roadway alignments and associated crash data in the state of Alabama. We used the Curve Advisory Reporting Service (CARS) System and Google Map (and Google Earth) to collect highway geometric data. CARS is a road survey system, developed by Reiker Inc., that automatically records vehicle activity as the test vehicle makes test runs on the roadway. The system consists of a GPS-based configuration that can be used to perform continuous road survey. The data sessions are created when the test vehicle equipped with GPS instrumentation makes a pass on a selected segment of roadway and collects information such as velocity, latitude, longitude, and altitude. The vehicle requires only one pass to collect the necessary data as the system is mounted on the dashboard (6). Alabama Department of Transportation (ALDOT) assigned Reiker Incorporated to collect CARS data on Highways in Alabama and record those data in a database. By the time this study was completed, the CARS database did not include all the state-maintained highways in Alabama. ALDOT allowed us to access the existing CARS database to upload data sessions for horizontal and vertical curves on a web portal for analyses. Therefore, this study does not include all the vertical alignments on the state-maintained highways. It used the data that were available in the CARS database. The CARS database displays all recorded data session geographically on a map. Each data session is depicted on the map as a series of uniformly sized circles. Changes in elevation are illustrated for each vertical curve data session by using circles of different colors, where green circles represent less drastic changes in elevation and red circles represents more drastic changes in elevation, as shown in Figure 1. In addition to using CARS data, the Google Map and Google Earth were used to identify vertical alignments and additional geometric information. When a segment of roadway was selected for analysis, the raw data were exported to Excel and processed. The processed data were imported to Power Inroads, a computer-aided design (CAD) program, to create the vertical profiles and measure Stopping Sight Distances (SSDs) as described later in this paper.



Figure 1: CARS Illustration of elevation changes on a roadway

Crash Data Collection

In this study, the historical crash data were obtained from the Critical Analysis Reporting Environment (CARE) software developed by the Center for Advanced Public Safety of the University of Alabama. The software compiles all police-reported accidents in the state of Alabama and provides individual incident information such as geographic location, contributing factors, roadway configurations, crash characteristics, etc. In this study, historical crash data for the years 2010-2017 were used. The crash data used in this study consisted of five injury severity categories: fatal injury (F), incapacitating injury (IN), non-incapacitating injury (NIN), no injury (NI), and property damage only (PDO). The fatal injury category includes any injury that resulted in death within 30 days of the crash. The incapacitating injury category includes any injury other than a fatal injury, which prevented the injured person from walking, driving, or normally continuing the activities the person was capable of performing before being injured. The non-incapacitating injury category includes any injury other than a fatal injury or an incapacitating injury, which was apparent to observers at the scene of the crash. The no injury category includes crashes with no documented or apparent injuries resulting from the crash. The property damage only category includes crashes that only result in property damage.

In this study, the crash data from CARE database were exported to ArcMap, a GIS program, so that crashes occurring on roadway segments with insufficient SSD could be identified and analyzed.

Identification of SSD Deficient Vertical Alignments

As stated earlier in this paper, the processed data from CARS were imported to Power Inroads, a computer-aided design (CAD) program, to create vertical profiles of the selected roadway segments. The available stopping sight distance (SSD) on each crest vertical alignment was measured using a graphical method provided in the AASHTO Green Book (AASHTO 2018). Figure 2 provides the concepts based on which the available stopping sight distances were

measured for crest vertical curves. This concept is based on 3.5 ft driver's eye height and 2.0 ft object height. Figure 3 demonstrates how we applied the aforementioned concept in Power Inroads to graphically determine SSD at a location on a crest vertical roadway profile. In Figure 3 the object line is a line that is parallel to the profile line and at a vertical distance 2 feet from the profile line. The sight line in the figure is drawn from driver's eye (i.e., 3.5 ft from the profile line) and which is tangent to the profile line. As shown in the figure, the SSD is the distance between the driver and where the sight line intersects the object line. We repeated the process shown in the Figure 3 at different intervals on the profile to obtain the minimum SSD. This minimum SSD is the available stopping sight distance for the crest vertical alignment for a direction of travel. Once the available SSD was determined using the above mentioned graphical methods, it was compared with the required stopping sight distance for the vertical alignment based on the posted speed limit. If the available sight distance was smaller than the required sight distance, the vertical curve was identified as a deficient curve.

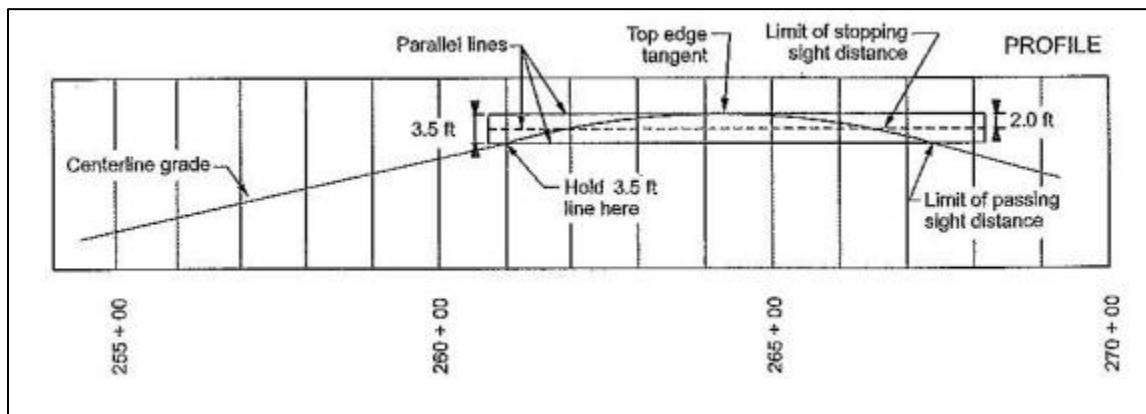


Figure 2: Concept of measuring SSD of a crest vertical curve (Source: AASHTO 2018)

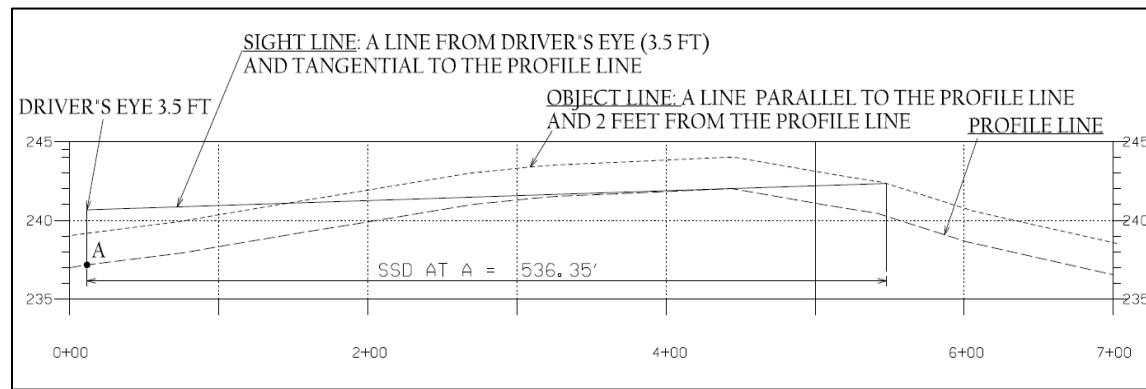


Figure 3: Measurement of SSD of an existing crest vertical curve using Power Inroads

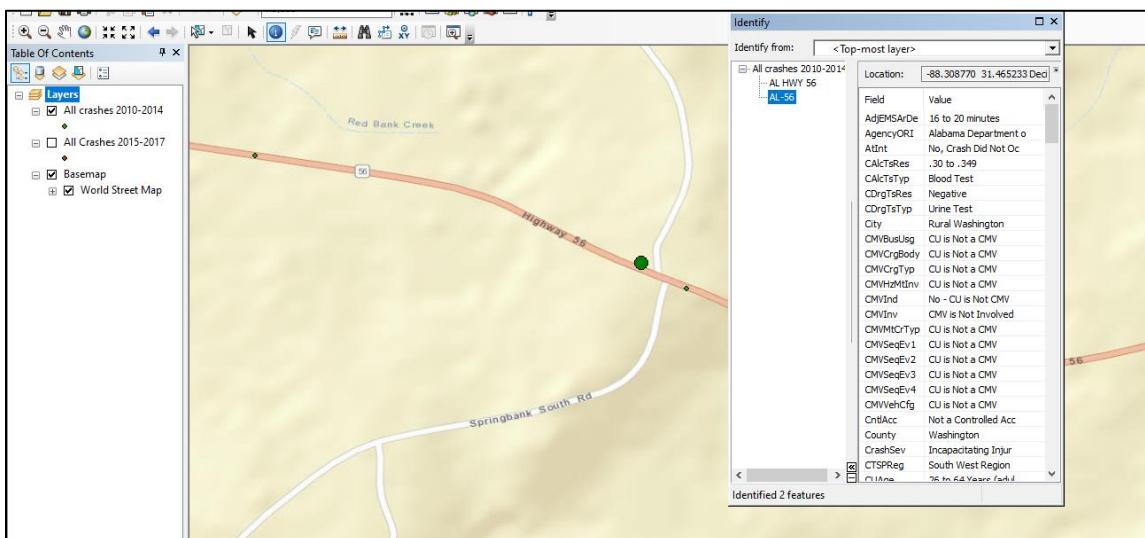
Identification of Crashes Occurring on SSD Deficient Vertical Alignments

As stated earlier in this paper, once the vertical roadway alignments with insufficient SSD were identified, the historic crash data from CARE were geographically mapped using ArcMap, a GIS program, to identify crashes on each SSD deficient roadway segment. Figure 4 provides an

example, which shows crashes on a SSD deficient vertical alignment on Alabama Highway 56 (AL 56). Figure 4 (a) provides a Google Map view of the roadway segment. Figure 4 (b) shows ArcMap view of the highway segment with crashes imported from CARE database. In Figure 7 (b), a green circle/dot represents a crash. The figure shows that 3 crashes occurred on this vertical alignment during 2010-2017. The larger circle/dot represent the crash that was selected to view detailed information (e.g., crash type, injury type, driver information, etc.) related to it. In this study, any SSD deficient vertical curve with at least 1 crash during the period 2010-2017 was selected for further analysis.



(a) Google Street view of a segment of a SSD deficient vertical curve on AL 56



(b) ArcMap interface showing the location and detail of a crash on the above vertical curve on AL 56

Figure 4: Identification of crashes on an SSD deficient vertical alignment on AL 56

RESULTS AND DISCUSSIONS

Evaluation of SSD Deficiencies

The method developed in this study used CARS database to identify the SSD deficient vertical alignments in Alabama and 35 sites were identified within the state highway system that

did not meet the minimum SSD requirement. After a roadway segment was found to be SSD deficient, the historic crash data from CARE was geographically mapped using ArcMap and crashes occurring on the insufficient segments were analyzed. The method identified 29 such vertical highway curves in the CARS database with at least 1 crash occurring on them. Based on the variations in geometric properties, the selected segments were subdivided into four categories for evaluation of SSD deficiencies: roadway segments containing horizontal curves and vertical curves, roadway segments hosting intersections on vertical curves, bridge approaches on vertical curves and roadway segments containing driveway intersections on vertical curves. The sites identified with insufficient SSD in each of these subcategories in this study are referenced as X1 through X29 and relevant information about the respective sites are summarized in Tables 1 and 2.

TABLE 1: List of Identified Sites on the State Routes with Insufficient SSD

Category	Sites with Insufficient SSD	State Route	County	Number of Lanes
1. <i>Roadway segments containing horizontal curve and a vertical curve</i>	4	X1 X2 X3 X4	Tuscaloosa Wilcox Clay Marion	4/5/2*
2. <i>Roadway segments containing intersections on vertical curves</i>	10	X5 X6 X7 X8 X9 X10 X11 X12 X13 X14	Washington Clay Coosa Etowah Winston Marion Marion Marion Limestone Mobile	2 2 4 2 3 2 2 2 5
3. <i>Bridge approaches on vertical curves</i>	5	X15 X16 X17 X18 X19	Winston Autauga Cullman Coosa Tallapoosa	2 2 2 4 2
4. <i>Roadway segments containing driveway interfaces on vertical curves</i>	10	X20 X21 X22 X23 X24 X25 X26 X27 X28 X29	Greene Fayette Marion Monroe Autauga Marengo Crenshaw Winston Colbert Colbert	2 2 2 2 2 2/5* 2 2 2 2
Total	29			

* Note: The number of lanes are different for multiple segments

TABLE 2: Evaluation of Available SSDs of the Identified Sites

Category	State Route (with segments with insufficient SSD)	Posted Speed Limit (mph)	Required SSD(ft)	Available SSD (ft) (Assessed)	Recommended Speed Limit (mph)
<i>1. Roadway segments containing horizontal curve and a vertical curve</i>	X1	40	305	291	35
	X2	35	250	227	30
	X3	55	553	521	55
	X4	55	495	228	35
<i>2. Roadway segments containing intersections on vertical curves</i>	X5	55	450	422	45
	X6	55	520	483	50
	X7	65	645	611	60
	X8	20	115	88	20
	X9	55	495	302	35
	X10	45	360	336	40
	X11	45	360	322	40
	X12	45	360	341	40
	X13	45	360	332	40
	X14	35	250	229	30
<i>3. Bridge approaches on vertical curves</i>	X15	55	495	438	50
	X16	55	495	350	40
	X17	45	360	309	40
	X18	65	682	553	55
	X19	55	520	462	50
<i>4. Roadway segments containing driveway interfaces on vertical curves</i>	X20	55	495	421	45
	X21	55	495	442	50
	X22	55	495	380	45
	X23	55	495	470	50
	X24	55	469	395	45
	X25	55	495	328	40
	X26	55	450	370	45
	X27	55	495	410	45
	X28	45	344	290	35
	X29	40	278	244	30

Safety Assessment

As stated earlier, after a roadway segment with insufficient stopping sight distance was identified, the historic crash data from CARE database were geographically mapped using ArcMap and crashes occurring on the insufficient segments were located. The summary of all crashes occurring on the identified segments between 2010 and 2017 are shown in Table 3. The table shows the number of crashes in each injury category for each identified site. As stated earlier, the crash data were categorized into five injury severity categories: fatal injury (F), incapacitating injury (IN), non-incapacitating injury (NIN), possible/no injury (NI), and property damage only (PDO).

TABLE 3: Crash Summary of the Identified Sites with Insufficient SSDs

Category	State Route (Segments with insufficient SSD)	Posted speed Limit (mph)	Crash Counts by Injury Type (2010-2017)					Total Number of Crashes
			F	IN	NIN	NI	PDO	
1. Roadway segments containing horizontal curves on a sag curve	X1	40	0	1	3	1	5	10
	X2	35	0	0	0	0	1	1
	X3	55	0	3	0	0	5	8
	X4	55	0	1	0	0	0	1
Total			20					
2. Roadway segments hosting intersections on vertical curves	X5	55	0	1	1	0	1	3
	X6	55	0	0	0	0	1	1
	X7	65	0	0	2	0	3	5
	X8	20	0	1	0	0	0	1
	X9	55	0	3	0	0	1	4
	X10	45	0	0	0	0	2	2
	X11	45	0	0	0	0	1	1
	X12	45	0	0	0	1	0	1
	X13	45	0	0	0	0	1	1
	X14	35	0	0	0	0	1	1
Total			20					
3. Bridge approaches	X15	55	0	1	0	1	2	4
	X16	55	0	0	0	0	3	3
	X17	45	0	1	0	0	1	2
	X18	65	0	1	2	0	5	8
	X19	55	0	0	0	0	3	3
Total			20					
4. Roadway segments containing driveway interfaces on vertical curves	X20	55	0	2	1	1	4	8
	X21	55	0	0	0	0	2	2
	X22	55	0	0	0	0	1	1
	X23	55	1	0	0	0	0	1
	X24	55	1	1	2	0	5	9
	X25	55	0	0	1	0	2	3
	X26	55	1	0	0	0	0	1
	X27	55	0	0	0	0	1	1
	X28	45	0	0	0	1	0	1
	X29	40	0	0	0	0	1	1
Total			28					
Grand Total			88					

To further explore the impacts of SSD deficiencies of vertical alignments on safety, the crash records were observed to identify any possible link between crashes and demographics of drivers involved in crashes at the identified sites. It is important to note that relationship between SSD and occurrence of crashes depends on a variety of different variables, which can be explored with statistical modeling and analysis, and which is beyond the scope of this paper. 56% of the drivers involved in crashes on the SSD deficient vertical curves were male. 26% of the drivers

were 16-24 years of age and 7% of the drivers were 65 years of age and older. The age distribution of the drivers involved in SSD deficient sites can present useful information to responsible agencies. For example, if there is a significant percentage of older drivers in the driving population, they will be at a higher risk of getting involved in crashes at the SSD deficient sites. In such a case, agencies need to consider issues with older drivers when they are updating design standards for existing or new highways.

The crashes identified at the SSD deficient sites were also examined to find out if any specific types of crashes were dominating. 25% of the crashes were single vehicle crashes, 57% of them were run-off the road crashes and 23% of them were rollover crashes.

CONCLUSIONS

This study developed an evaluation approach for identifying stopping sign distance (SSD) deficient vertical alignments on highways in Alabama and for assessing the safety impact of such deficiencies. Highway geometric data for this evaluation method were obtained from the Curve Advisory Reporting Service (CARS) System database developed by Rieker Incorporated. The segments within the CARS database were selected for SSD evaluation if the data sessions for the segments within the database demonstrated significant elevation change. The geometric data extracted from CARS database were processed and exported to a computer-aided design (CAD) program to graphically measure the available SSDs on vertical roadway segments. The available SSDs were compared with the theoretical AASHTO recommended SSDs. If the available SSD was less than the required SSD, the segment was considered as SSD deficient. Initially, all possible sites containing vertical curves that were available through the CARS database were screened to identify SSD deficiencies. This was in line with the safe systems approach that FHWA is currently recommending, by which locations with deficient SSD's were identified no matter crashes had already occurred or not.

In order to conduct a safety assessment of the roadway segments with insufficient SSDs, crash data for the years 2010-2017 occurring on the identified SSD deficient roadway segments were extracted from the Critical Analysis Reporting Environment (CARE) database. A total of 35 roadway segments were identified that did not meet the minimum SSD requirements and 29 of the sites had at least 1 accident occurring on them during 2010-2017. The 29 identified road segments had a total of 88 crash records from 2010 to 2017. This study also attempted to explore the relationship between SSD deficiencies and crash occurrences by gender and age group.

The state agencies can use the methodology developed in this study to address roadway limitations involving vertical alignments in order to improve safety for existing roadways and for designing new roadways. As a future direction of work, a larger set of roadway data containing vertical alignment information should be collected and evaluated so that the results from the study can provide more meaningful direction towards what design guidelines should be updated or what changes should be adopted to accommodate the driving population of a wider range.

ACKNOWLEDGMENTS

The authors would like to thank Alabama Department of Transportation (ALDOT) for providing the data used in this study. Acknowledgments are also extended to ALDOT for providing the financial support.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Samantha Islam; data reduction and analysis: Benjamin Burton; interpretation of results: Samantha Islam; draft manuscript preparation: Samantha Islam and Benjamin Burton. All authors reviewed the results and approved the final version of the manuscript.

DISCLAIMER

The findings and conclusion of this study are those of the author and do not necessarily represent the views of ALDOT, the Transportation Research Board, or the National Academies.

REFERENCES

1. AASHTO. (2018). *A Policy on the Geometric Design of Highways and Streets*. Washington, D.C.
2. AASHTO. (1940). *A Policy on Sight Distance for Highways*. Washington D.C.
3. AASHTO. (2011). *A Policy on the Geometric Design of Highways and Streets*. Washington, D.C.
4. *A Policy on Geometric Design of rural Highways*. (1965). American Association of State Highway and Transportation Officials Washington D.C.
5. Gargoum, S. A., and Karsten, L. (2021). "Virtual assessment of sight distance limitations using LiDAR technology: Automated obstruction detection and classification". *Automation in Construction*. Volume 125.
6. Green, E., Agent, K., and Lammers, E. (2016). "Development of an Improved Method for Determining Advisory Speeds on Horizontal Curves", Report No. KTC-16-14/SPR15-507, Kentucky Transportation Center.
7. Mannering, F. L., and Washburn, S. S. (2020). *Principles of Highway Engineering and Traffic Analysis*. Wiley. 7th Edition.

COVID-19 Pandemic and Distraction-Related Motorcycle Crashes in Kentucky

Bharat Kumar Pathivada, Ph.D.¹; Noelle Buhay²; Dylan Justice³; Arunabha Banerjee, Ph.D.⁴;
and Kirolos Haleem, Ph.D., P.E., M.ASCE⁵

¹Postdoctoral Research Associate, Transportation Safety & Crash Avoidance Research Lab, School of Engineering and Applied Sciences, Western Kentucky Univ., Bowling Green, KY. Email: bharatkumar.pathivada@wku.edu

²Undergraduate Student Researcher, School of Engineering and Applied Sciences, Western Kentucky Univ., Bowling Green, KY. Email: noelle.buhay318@topper.wku.edu

³Undergraduate Student Researcher, School of Engineering and Applied Sciences, Western Kentucky Univ., Bowling Green, KY. Email: dylan.justice674@topper.wku.edu

⁴Postdoctoral Research Associate, Transportation Safety & Crash Avoidance Research Lab, School of Engineering and Applied Sciences, Western Kentucky Univ., Bowling Green, KY. Email: arunabha.banerjee@wku.edu

⁵Associate Professor of Civil Engineering, School of Engineering and Applied Sciences, Western Kentucky Univ., Bowling Green, KY (corresponding author). Email: kirolos60@hotmail.com

ABSTRACT

This study investigates the impact of COVID-19 pandemic on distraction-related motorcycle crashes in Kentucky while comparing pre-pandemic (2015–2022) and post-pandemic (2020–2022) periods. Recent eight years (2015 through 2022) of distraction-related motorcycle crash records and associated roadway information were retrieved from the Kentucky Transportation Cabinet (KYTC). Statistical tests including the chi-square test of independence, Z-test of proportions, and odds ratio (OR) were used to identify the factors affecting the severity of distraction-related motorcycle crashes pre- and post-pandemic. The chi-square test indicated that crash characteristics (e.g., manner of collision and speeding-related), at-fault driver characteristics (driver age and DUI), roadway characteristics (area type, crash location, road functional class, presence of horizontal curve, and vertical gradient), and environmental characteristics (weather and lighting conditions) were significantly associated with the severity of distraction-related motorcycle crashes pre- and post-pandemic. In addition, the presence of horizontal curve significantly affected the severity of distraction-related motorcycle crashes post-pandemic. Results from the Z-test comparing the proportions of severe crashes pre- and post-pandemic revealed that the percentage of severe distraction-related motorcycle crashes significantly increased post-pandemic. Furthermore, the proportion of severe distraction-related motorcycle crashes for very young at-fault drivers (16–20 years old), presence of horizontal curves, speeding, clear/cloudy weather, and at intersections significantly increased post-pandemic. The OR results revealed that the odds of severe distraction-related motorcycle crashes involving speeding, with presence of horizontal curves, and for very young drivers increased by 67%, 64%, and 48%, respectively, post-pandemic. Based on the study findings, removal of distraction-related road elements (e.g., billboards) at high motorcycle crash risk locations, and ensuring installations of rumble strips on edge line, could help reduce the severity of distraction-related motorcycle crashes on Kentucky roads.

Keywords: Distracted Driving, COVID-19 Pandemic, Motorcycle Crashes, Safety, Severe, Speeding, Kentucky

INTRODUCTION

The severe acute respiratory syndrome (SARS-CoV-2), also known as the 2019 novel coronavirus (COVID-19), had significant global consequences and was declared a pandemic by the World Health Organization (WHO 2021). Several studies have highlighted that traffic conditions and patterns have significantly changed worldwide since the COVID-19 pandemic (Chand et al. 2021; Yasin et al. 2021). The public safety precautions during the pandemic led to reduced traffic movements, which in turn increased unsafe road user activity, including speeding and distracted driving behavior (Shaik and Ahmed 2022).

Evidence from past studies indicates that driver distraction has a greater detrimental effect on driving safety than other factors, such as alcohol intoxication and fatigue (Qin et al. 2019). According to the National Highway Traffic Safety Administration (NHTSA 2023a), distracted driving is any activity that diverts the driver's attention from the primary task of navigating the vehicle and responding to critical events. Recent technological advances including smartphones, wearable devices, and in-vehicle information systems have considerably increased the rate of driver distractions (Qin et al. 2019). Moreover, distracted driving has significantly increased post-COVID-19 pandemic. For instance, a recent report from Cambridge Mobile Telematics (CMT 2023) highlighted that distracted driving in the U.S. has increased by more than 20%, and fatalities due to distraction increased by 10.4% post the pandemic.

The risk of distraction-related casualties exponentially increases with motorcycles due to the lack of vehicle stability and in-vehicle protection for motorcyclists. Motorcyclists are among vulnerable road users and are highly prone to severe injury crashes. For instance, per 100 million vehicle miles traveled (VMT) in 2021 in the U.S., the fatality rate for motorcyclists (30.20) was almost 24 times the passenger car occupant fatality rate (1.26) and 40 times the fatality rate for light-truck occupants (0.76). According to NHTSA (2023b), motorcycles accounted for 14% of total traffic fatalities in 2021, while only representing 3.5% of all registered vehicles. In the state of Kentucky, motorcycles contributed to less than 1% of total crashes in 2021, but accounted for more than 7% of fatal crashes (Kentucky Traffic Collision Facts 2021).

These statistics emphasize the need to investigate the causes and circumstances resulting in distraction-related motorcycle crashes and the changes in crash patterns post-pandemic. Notably, limited research efforts have been made to examine distracted-related motorcycle crashes. Specifically, the impact of COVID-19 pandemic on distraction-related motorcycle crashes has not been widely investigated. Thus, this study takes the initiative and investigates the impact of the COVID-19 pandemic on distraction-related motorcycle crashes through a comparative analysis of pre-COVID-19 (2015-2019) and post-COVID-19 (2020-2022) periods in Kentucky.

LITERATURE REVIEW

Several studies investigated the impact of COVID-19 pandemic on different aspects of road safety. For example, Islam et al. (2022) investigated the risk compensation trends in road safety during COVID-19 on Interstate 4 in Orlando, Florida. Their results showed that during the pandemic, traffic volume decreased by 13.6%, contributing to decreased crash frequency, while average speed increased by 11.3%, contributing to increased crash severities. Furthermore, during the pandemic, alcohol-related crashes decreased by 22%

Javid et al. (2022) investigated the impact of COVID-19 on distracted driving using a stated preference survey. Risky driving behaviors and self-reported distractions were recorded from

158 participants in the state of Maryland. Their results showed that hands-free cell phones, GPS, and eating/drinking were primary sources of distraction during the pandemic, while the odds of being distracted by a handheld cell phone during the pandemic were 1.45 times higher than the others. Muley et al. (2021) investigated the impact of COVID-19 preventive measures on traffic mobility in Qatar. Their results showed that traffic violations and crashes dropped by 73% and 37%, respectively during the COVID-19. Chand et al. (2021) studied the impact of COVID-19 lockdowns on road traffic incidents in Sydney, Australia. Their results showed that there was a shift in traffic incidents during the pandemic with more incidents recorded in suburban areas.

Nadimi et al. (2023) investigated the impact of COVID-19 on transport behavior in Iran. Their results showed that total crash fatalities marginally decreased during the restrictions, while motorcycle fatal crashes increased due to the increased traffic speeds. Shaaban and Mortimer (2022) investigated the impact of COVID-19 pandemic on traffic fatalities in Utah. Their results showed that there was less traffic and fewer crashes, but the number of fatal crashes increased. The factors that contributed to a significant increase in fatal crashes included intersection-related, older driver-related, motorcycle-related, and speed-related.

Vanlaar et al. (2021) investigated the impact of COVID-19 on road safety in Canada and the United States using self-reported data on risky road user behaviors. Using 3,001 responses from the public opinion survey on key road safety issues, their results showed that drivers were more likely to engage in risky driving behaviors during the pandemic. The most frequently reported risky behaviors included speeding, drinking and driving, and distracted driving.

The review of existing literature revealed that several safety studies have highlighted the reduction of traffic volumes and crash frequencies, while there was an increase in crash injury severities, fatalities, and risky driving behaviors such as speeding and distracted driving during the COVID-19 pandemic. However, limited research efforts have been made to examine the detrimental effects on motorcycle crashes, specifically, the impact of COVID-19 pandemic on distraction-related motorcycle crashes has not been widely investigated. Therefore, this study aims to fill this gap by investigating the impact of COVID-19 pandemic on distraction-related motorcycle crashes through a comparative analysis of pre-COVID-19 (2015-2019) and post-COVID-19 (2020-2022) periods in Kentucky using comprehensive crash and roadway data. The study findings would help propose potential countermeasures to enhance motorcyclists' safety.

DATA COLLECTION AND PREPARATION

Recent eight years (i.e., 2015 through 2022) of motorcycle-related crash records involving at least one motorcycle were collected from the Kentucky Transportation Cabinet (KYTC). The crash severity level in the database was recorded on the KABCO scale: fatal injury or killed "K", serious injury "A", suspected minor injury "B", possible injury "C", and no injury "O" (or property damage only "PDO"). In this study, crashes were grouped into two categories based on the injury severity level; severe crashes (i.e., K+A) and non-severe crashes (i.e., B+C+O). The crash database also included factors associated with at-fault driver behavior, such as distracted, speeding, impaired driving, and driver aggressiveness. The distract "Y/N" indicator was used to screen distraction-related motorcycle crashes. Furthermore, to investigate the impact of COVID-19 on distraction-related motorcycle crashes, the eight-year crash records were split into pre-pandemic (2015-2019) and post-pandemic (2020-2022) periods. It is important to note that both "alcohol" and "drugs" were merged into a single category to represent "driving under the influence (DUI)".

The screened distraction-related motorcycle crash database was then cleaned and merged with the roadway information into a single database. In total, there were 2,440 and 1,210 distraction-related motorcycle crashes for pre- and post-pandemic periods, respectively. Figure 1 shows the percentage distribution of distraction-related motorcycle crashes across the five injury severity levels in each of the pre- and post-pandemic periods. To evaluate the impact of COVID-19 pandemic on distraction-related motorcycle crashes, crash characteristics (manner of collision, work zone-related, weekend-related, and speeding-related), at-fault driver characteristics (driver age and DUI), roadway characteristics (area type, crash location, road functional class, presence of horizontal curve, and vertical gradient), and environmental characteristics (weather and lighting conditions) were analyzed, as shown next.

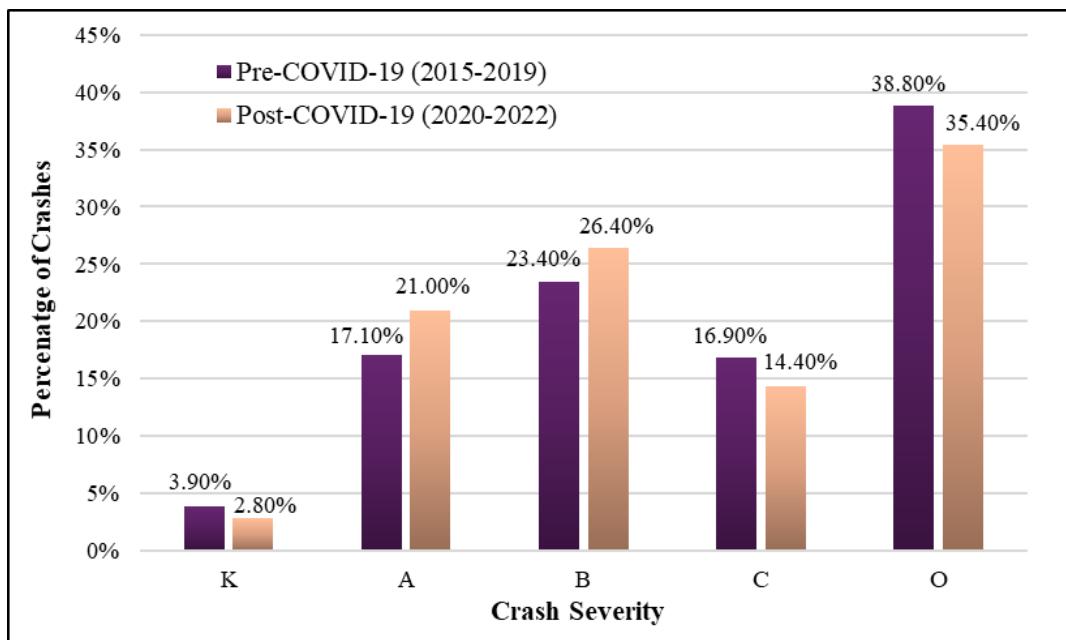


Figure 1. Distribution of distraction-related motorcycle crashes pre- and post-COVID-19 pandemic.

METHODOLOGICAL APPROACHES

Chi-Square (χ^2) Test

The chi-square test of independence was used in each of the pre- and post-pandemic periods to assess the association between distraction-related motorcycle crash injury severity and other categorical variables. The chi-square test statistic is computed using Equation (1) as follows:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

where: χ^2 is the chi-square test statistic, O is the observed crash frequency, E is the expected crash frequency, and n is the sample size. Note that a significant value for the chi-square test

indicates that the tested factor is significantly associated with the distraction-related motorcycle crash injury severity.

Odds Ratio (*OR*)

The odds ratio (*OR*) is a relative measure of risk that indicates how likely it is for an event to occur in one group compared to the other group. In the present study, OR was used to assess the relative risk of severe injuries for the studied factor in each of the pre- and post-pandemic periods. For example, OR indicates how likely it is that a severe injury happens in crashes affected by DUI compared to those non-affected. OR be calculated using Equation (2) as follows:

$$OR_{DUI} = \frac{\frac{N_{Severe,DUI}}{N_{Non-Severe,DUI}}}{\frac{N_{Severe,Non-DUI}}{N_{Non-Severe,Non-DUI}}} \quad (2)$$

where: $N_{Severe,DUI}$ and $N_{Severe,Non-DUI}$ are the number of severe distraction-related motorcycle crashes caused by DUI and non-DUI violations, respectively; and $N_{Non-Severe,DUI}$ and $N_{Non-Severe,Non-DUI}$ denote the number of non-severe distraction-related motorcycle crashes caused by DUI and non-DUI violations, respectively. Note that in a similar manner OR for other factors can be computed in each of the pre- and post-pandemic periods.

Z-Test of Proportions

The Z-test of proportions was used to examine the presence of significant difference between the proportions of severe injuries in pre- and post-pandemic distraction-related motorcycle crashes. For a specific categorical factor, the Z-statistic can be calculated using Equations (3) and (4).

$$Z = \frac{\bar{P}_{post,severe} - \bar{P}_{pre,severe}}{\sqrt{\bar{P}(1 - \bar{P})\left(\frac{1}{n_{post}} + \frac{1}{n_{pre}}\right)}} \quad (3)$$

$$\bar{P} = \frac{n_{pre,severe} + n_{post,severe}}{n_{pre} + n_{post}} \quad (4)$$

where: $\bar{P}_{pre,severe}$ and $\bar{P}_{post,severe}$ are the proportions of severe injuries in pre- and post-pandemic distraction-related motorcycle crashes attributed to the factor under investigation, respectively; n_{pre} and n_{post} are the sample sizes for pre- and post-pandemic distraction-related motorcycle crashes, respectively; and $n_{pre,severe}$ and $n_{post,severe}$ are the sample sizes for severe pre- and post-pandemic distraction-related motorcycle crashes for the studied factor, respectively. Note that a significant value for the Z-test indicates that for the tested factor there exists a significant difference in the percentage of severe injury crashes in the pre- and post-pandemic periods.

RESULTS & DISCUSSION

Chi-Square Test of Independence and Odds Ratio (*OR*) Results

Table 1 shows the results of the chi-square test of independence and odds ratios (OR) for severe distraction-related motorcycle crashes in each of the pre-pandemic and post-pandemic periods. An *OR* greater than one indicates that severe distraction-related motorcycle crashes are more likely to occur, and vice versa.

Crash Characteristics

From Table 1, the chi-square test results showed that the manner of collision and speeding-related collisions had significant association with the severity of distraction-related motorcycle crashes pre- and post-pandemic. Furthermore, angle, head-on, single-vehicle, and left-turn collisions were associated with higher severe injury risks in both pre- and post-pandemic periods (with $OR > 1$). The *OR* results showed that the odds of severe distraction-related motorcycle left-turn collisions decreased 0.6 times post-pandemic ($OR = 1.731$) compared to pre-pandemic ($OR = 2.695$). This could be possibly attributed to the reduced traffic volumes during the pandemic (Islam et al., 2022). Interestingly, weekend-related crashes were significantly associated with the severity of distraction-related motorcycle crashes pre-pandemic, but not post-pandemic. This could be since the pattern of traffic movements significantly changed post-pandemic, which is consistent with the findings of Sosik-Filipiak et al. (2023).

At-Fault Driver Characteristics

With regard to at-fault driver characteristics, the chi-square test results indicated that driver age was significantly associated with the severity of distraction-related motorcycle crashes pre- and post-pandemic. Specifically, middle-age drivers, old-age drivers, aggressive driving, speeding, and DUI were associated with higher severe injury risks in both pre- and post-pandemic periods (with $OR > 1$). Moreover, the odds of severe injuries involving very young at-fault drivers increased about 1.5 times during post-pandemic ($OR = 1.043$) compared to pre-pandemic ($OR = 0.704$). This is expected since very young drivers are prone to distractions and could be associated with risky driving behaviors; increasing their likelihood of being involved in severe crashes post-pandemic. This is consistent with the findings of Nooshabadi et al. (2023) and Lopetrone and Binodi (2023).

Interestingly, DUI did not have any significant impact on the severity of distraction-related motorcycle crashes post-pandemic. Moreover, the odds of severe injuries involving DUI decreased about 0.4 times during post-pandemic ($OR = 1.425$) compared to pre-pandemic ($OR = 4$). This could be likely attributed to the shift in drinking patterns during the pandemic, where alcohol was mostly consumed at home decreasing the propensity to drive afterward. This is consistent with the findings of Islam et al. (2022).

Roadway Characteristics

The presence of horizontal curve was found to be significantly associated with the severity of distraction-related motorcycle crashes post-pandemic. This could be attributed to the increased motorcycle speeds post-pandemic resulting in loss of vehicle control (Islam et al. 2022; Nadimi et al. 2023). Specifically, the odds of severe distraction-related motorcycle injuries involving

horizontal curves increased 1.6 times post-pandemic ($OR = 1.536$) compared to pre-pandemic ($OR = 0.937$). These findings are also consistent with the findings of Shaaban and Mortimer (2022).

Table 1. Chi-Square Test and Odds Ratio Results for Distraction-Related Motorcycle Crashes for Pre- and Post-Pandemic Periods.

Variable	Pre-Pandemic (2015-2019)		Post-Pandemic (2020-2022)	
	Odds Ratio (<i>OR</i>)	P-Value (χ^2)	Odds Ratio (<i>OR</i>)	P-Value (χ^2)
Crash Characteristics				
Manner of Collision				
Angle Collision	1.498	<0.001	1.480	<0.001
Head-On Collision	2.078		2.058	
Rear-End Collision	0.360		0.414	
Sideswipe Collision	0.655		0.458	
Single-Vehicle Collision	1.437		1.494	
Left-Turn Collision	2.695		1.731	
Work Zone-Related Crash	0.941		1.166	
Weekend-Related Crash	1.189		1.028	
Speeding-Related Crash	2.218	<0.001	3.713	<0.001
At-Fault Driver Characteristics				
DUI (Driving under the Influence)	4.000	<0.001	1.425	0.334*
Driver Age				
Very Young (16 to 20 years)	0.704	<0.001	1.043	0.018
Young (21 to 30 years)	0.829		0.821	
Middle (31 to 59 years)	1.227		1.160	
Old (≥ 60 years)	1.324		1.268	
Roadway Characteristics				
Presence of Horizontal Curve	0.937	0.624*	1.536	0.014
Presence of Vertical Gradient	1.065	0.552*	1.081	0.579*
Rural Area Type	1.874	<0.001	1.612	<0.001
Crash Location				
Segment	0.996	0.998*	0.902	0.242*
Intersection	1.003		1.167	
Ramp	1.027		0.372	
Road Functional Class				
Interstate/Principal Arterial	0.925	0.028	0.825	0.297*
Minor Arterial	1.365		1.055	
Major Collector	0.904		1.299	
Minor Collector/Local	0.818		0.885	
Environmental Characteristics				
Nighttime Lighting Condition	1.498	0.001	1.554	0.007
Peak Morning/Afternoon Crash	0.924	0.472*	1.032	0.836*
Weather				
Clear/Cloudy	0.938	0.369*	1.574	0.429*
Foggy	0.000		3.209	
Rainy	1.331		0.372	

* Not significant at 10% level of significance.

Rural area was found to be positively associated with a higher risk of severe injuries in both pre- and post-pandemic periods involving distraction-related motorcycle crashes (with $OR > 1$). This could be likely attributed to the relatively higher speeds in rural areas compared to the urban areas; hence, resulting in an increased severe injury risk. Furthermore, the chi-square test results suggested that crash location and the presence of vertical gradient did not affect the injury severity of distraction-related motorcycle crashes.

Interestingly, road functional class had a significant effect on crash severity pre-pandemic, but did not have significant effect on crash severity post-pandemic. This could be attributed to the changes in traffic patterns and lower traffic volumes post-pandemic (Shaaban and Mortimer 2022). Furthermore, the OR results showed that the odds of severe distraction-related motorcycle injuries on major collector roads increased 1.4 times during post-pandemic ($OR = 1.229$) compared to pre-pandemic ($OR = 0.904$). This is consistent with the findings of Chand et al. (2021), who could report more incidents and crashes away from the central business areas.

Environmental Characteristics

The nighttime lighting condition was found to be significantly associated with the severity of distraction-related motorcycle crashes pre- and post-pandemic periods. Specifically, nighttime lighting condition had a higher severe injury risk involving distraction-related motorcycle crashes (with $OR > 1$). This is possibly due to insufficient sight distance and reduced reaction time caused by driver distraction at night.

Weather conditions did not affect the crash severity of distraction-related motorcycle crashes pre- and post-pandemic. The OR results showed that the odds of severe distraction-related motorcycle injuries in clear/cloudy weather increased 1.7 times during post-pandemic ($OR = 1.574$) compared to pre-pandemic ($OR = 0.938$). This is expected since risk-compensating behavior could become apparent in clear weather conditions, particularly during the pandemic with reduced road traffic. This finding is consistent with that of Dong et al. (2023).

Z-Test of Proportion Results

Table 2 shows the significant Z-test of proportion results for distraction-related motorcycle crashes during the pre- and post-pandemic periods. The positive sign for the Z-test statistic indicates that the proportion of severe injury crashes during post-pandemic was higher than the proportion of severe injury crashes during pre-pandemic, and vice-versa.

The overall results of the Z-test comparing the proportions of severe crashes pre- and post-pandemic ($Z\text{-test} = 1.937$; $P\text{-value} = 0.053$) showed that the percentage of severe distraction-related motorcycle injuries significantly increased post-pandemic. Furthermore, the Z-test results in Table 2 showed that crashes involving very young at-fault drivers (16 to 20 years old), presence of horizontal curves, speeding, clear weather, and at intersections significantly contributed to the increased percentage of severe distraction-related motorcycle injuries post-pandemic. These findings are intuitive since the reduced roadway traffic post-pandemic could contribute to risky driving behaviors (e.g., speeding); hence, resulting in increased percentage of severe distraction-related motorcycle crashes. These findings are consistent with those studies that reported a significant increase in unsafe road user activity post-pandemic (Yasin et al. 2021; Shaik and Ahmed 2022; Dong et al. 2022). Interestingly, the Z-test results showed that the proportion of severe distraction-related motorcycle injuries associated with DUI decreased post-

pandemic. As previously discussed, this could be attributed to the change in drinking patterns during the pandemic.

Table 2. Significant Z-Test of Proportion Results for Distraction-Related Motorcycle Crashes Pre- and Post-Pandemic.

Variable	Pre-Pandemic (2015-2019)		Post-Pandemic (2020-2022)		Z-Test Stat.	P-Value
	Severe Crashes	Total Crashes	Severe Crashes	Total Crashes		
Very Young at-Fault Driver (16 to 20 years old)	37	229	25	102	1.798	0.072
DUI (Driving under the Influence)	40	80	11	36	-1.952	0.051
Presence of Horizontal Curve	84	418	58	188	2.892	0.004
Speeding-Related Crash	38	108	26	50	2.002	0.045
Intersection-Related Crash	153	728	108	421	1.807	0.071
Major Collector Road Functional Class	86	438	69	249	2.434	0.015
Clear/Cloudy Weather	499	2381	284	1186	2.031	0.042

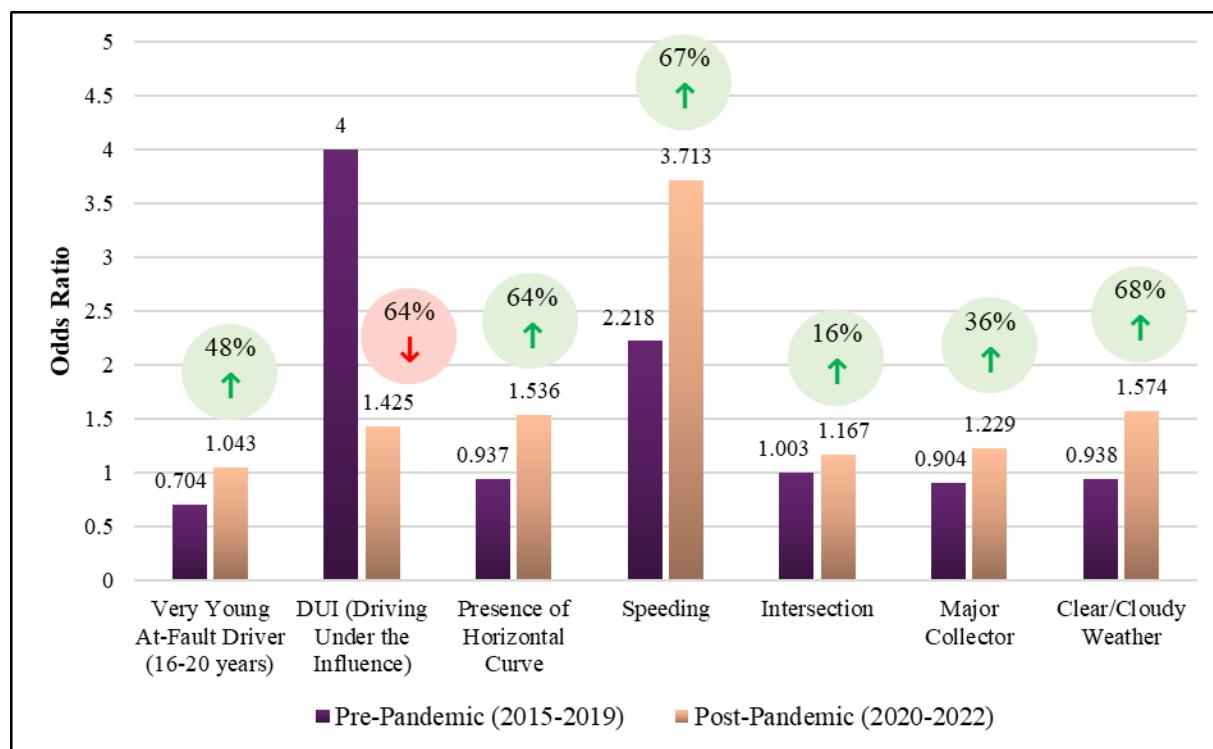


Figure 2. Odds ratios of contributing factors with significant Z-test of proportions for pre- and post-pandemic periods.

To better visualize the individual effects of each contributing factor that had a significant change in the percentage of severe distraction-related motorcycle injuries, the odds ratios for

those significant contributing factors pre- and post-pandemic are shown in Figure 2. The results revealed that the odds of severe distraction-related motorcycle crashes involving very young drivers, with presence of horizontal curves, speeding-related, at intersections, along major collector roads, and in clear/cloudy weather increased by 48%, 64%, 67%, 16%, 36%, and 68%, respectively post-pandemic. On the other hand, the odds of severe distraction-related motorcycle injuries involving DUI decreased by 64% post-pandemic.

CONCLUSIONS AND RECOMMENDATIONS

To the authors' knowledge, limited research has been conducted to examine distracted-related motorcycle crashes, specifically when it comes to investigating and contrasting the impact of COVID-19 pandemic on such crash types. This study took the initiative and investigated the impact of COVID-19 pandemic on the injury severity of distraction-related motorcycle crashes along Kentucky's roads. Recent comprehensive eight years (2015 through 2022) of distraction-related motorcycle crash records were used while comparing pre-COVID-19 pandemic (2015-2022) and post-COVID-19 pandemic (2020-2022) periods. Statistical tests including, the Z-test of proportions, chi-square test of independence, and odds ratio (*OR*) were used to investigate the circumstances resulting in changes in crash patterns post-COVID-19 pandemic and to identify the factors affecting the severity of distraction-related motorcycle crashes pre- and post-pandemic.

Results from the Z-test of proportions comparing the proportions of severe motorcycle crashes pre- and post-pandemic showed that the percentage of severe distraction-related motorcycle crashes significantly increased post-pandemic. Furthermore, the Z-test results showed that crashes involving very young drivers (16 to 20 years), with aggressive driving (i.e., speeding), with presence of horizontal curves, along major collector roads, in clear/cloudy weather, and at intersections significantly increased the percentage of severe distraction-related motorcycle injuries post-pandemic. The chi-square test indicated that crash characteristics (manner of collision), at-fault driver characteristics (driver age, speeding), roadway characteristics (area type), and environmental characteristics (nighttime) were found to be significantly associated with the severity of distraction-related motorcycle crashes pre- and post-pandemic. In addition, the presence of horizontal curve was found to significantly affect the severity of distraction-related motorcycle crashes post-pandemic. The odds ratios revealed that the odds of severe distraction-related motorcycle crashes involving very young drivers, with presence of horizontal curves, speeding-related, and in clear/cloudy weather increased by 48%, 64%, 67%, and 68%, respectively post-pandemic. On the other hand, the odds of severe distraction-related motorcycle injuries involving DUI decreased by 64% post-pandemic (possibly due to the shift in drinking patterns during the pandemic, where alcohol was mostly consumed at home).

The study findings could help suggest several countermeasures to reduce the frequency and severity of distraction-related motorcycle crashes on Kentucky roads. These could include launching safety awareness campaigns and training programs targeting younger motorcyclist age groups to improve practices of safe driving, ensuring installations of edge line and centerline rumble strips to alert drivers drifting from their driving lane, and removal of distraction-related road elements (e.g., billboards) at high motorcycle crash risk locations.

Further follow-up studies could compare the distinct effects of various distraction sources post-COVID-19 on motorcycle-related crashes. Furthermore, the safety effectiveness of different

countermeasures in reducing traffic crashes affected by distracted behaviors could be explored through before-and-after crash analysis.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Kentucky Transportation Cabinet (KYTC) for the grant provided to conduct this research. The authors also like to extend their gratitude to the KYTC for providing the necessary crash and roadway data used in this research. The opinions, findings, and conclusions in this paper are those of the authors and not necessarily those of the KYTC.

REFERENCES

- CMT (Cambridge Mobile Telematics). <CMT 2023 Distracted Driving Report.pdf (cmtelematics.com)>(September 15, 2023).
- Chand, S., Yee, E., Alsultan, A., and Dixit, V. V. (2021). "A Descriptive Analysis on the Impact of COVID-19 Lockdowns on Road Traffic Incidents in Sydney, Australia." *International Journal of Environmental Research and Public Health*, 18(21), 11701.
- Dong, N., Zhang, J., Liu, X., Xu, P., Wu, Y., and Wu, H. (2022). "Association of Human Mobility with Road Crashes for Pandemic-ready Safer Mobility: A New York City Case Study." *Accident Analysis & Prevention*, 165, 106478.
- Islam, M. R., Abdel-Aty, M., Islam, Z., and Zhang, S. (2022). "Risk-Compensation Trends in Road Safety During COVID-19". *Sustainability*, 14(9), 5057.
- Javid, R., Sadeghvaziri, E., and Jeihani, M. (2022). "The Effect of COVID-19 on Distracted Driving: A Survey Study". *MedRxiv*.
- KTC (Kentucky Transportation Center). Traffic Collision Facts: 2021 Report. <CollisionFacts2021.pdf (ky.gov)>(June 10, 2023).
- Lopetrone, E., and Biondi, F. N. (2023). "On the effect of COVID-19 on Drivers' Behavior: A Survey Study." *Transportation Research Record: Journal of Transportation Research Board*, 2677(4), 742-750.
- Muley, D., Ghanim, M. S., Mohammad, A., and Kharbeche, M. (2021). "Quantifying the Impact of COVID-19 Preventive Measures on Traffic in the State of Qatar". *Transport Policy*, 103, 45-59.
- Nadimi, N., Zayandehroodi, M. A., Rahmani, F., Asadamraji, M., and Litman, T. (2023). "Evaluation of the Impact of COVID-19 on Transport Sustainability in Iran." *In Proceedings of the Institution of Civil Engineers-Engineering Sustainability*, 40, 1-25.
- NHTSA (National Highway Traffic Safety Administration). <Distracted Driving | NHTSA>(September 30, 2023a).
- NHTSA (National Highway Traffic Safety Administration). <2021 Data: Motorcycles (dot.gov)>(September 30, 2023b).
- Nooshabadi, M. H., Vasquez, H. M., and Donmez, B. (2023). "Targeting Young Driver Emotions Can Reduce their Cellphone Distractions." *Accident Analysis & Prevention*, 192, 107202.
- Qin, L., Li, Z. R., Chen, Z., Bill, M. A., and Noyce, D. A. (2019). "Understanding Driver Distractions in Fatal Crashes: An Exploratory Empirical Analysis." *Journal of Safety Research*, 69, 23-31.

- Shaaban, K., and Mortimer, M. (2022). "The COVID-19 Pandemic's Impact on Traffic Fatalities in Utah." In *2022 Intermountain Engineering, Technology and Computing (IETC)*, 1-4.
- Shaik, M. E., and Ahmed, S. (2022). "An Overview of the Impact of COVID-19 on Road Traffic Safety and Travel Behavior." *Transportation Engineering*, 9, 100119.
- Sosik-Filipiak, K., Ostrowski, P., and Iwan, S. (2023). "Pedestrian Safety in Road Traffic in the Era of the SARS-CoV-2 Pandemic in the Example of Szczecin." *Sustainability*, 15(14), 11000.
- Vanlaar, W. G., Woods-Fry, H., Barrett, H., Lyon, C., Brown, S., Wicklund, C., and Robertson, R. D. (2021). "The Impact of COVID-19 on Road Safety in Canada and the United States." *Accident Analysis & Prevention*, 160, 106324.
- World Health Organization. (2021). <COVID-19 weekly epidemiological update> (July 30, 2023).
- Yasin, Y. J., Grivna, M., and Abu-Zidan, F. M. (2021). "Global Impact of COVID-19 Pandemic on Road Traffic Collisions." *World Journal of Emergency Surgery*, 16(1), 1-14.

Bridge Health-Informed Route Planning: Challenges and Promises

Yonas Kassa¹; Will Heller²; Brandon Lacy³; Brian Ricks⁴; and Robin Gandhi⁵

¹Univ. of Nebraska Omaha, Omaha, NE. Email: ykassa@unomaha.edu

²Univ. of Nebraska Omaha, Omaha, NE. Email: wheller@unomaha.edu

³Univ. of Nebraska Omaha, Omaha, NE. Email: bclacy@unomaha.edu

⁴Univ. of Nebraska Omaha, Omaha, NE. Email: bricks@unomaha.edu

⁵Univ. of Nebraska Omaha, Omaha, NE. Email: rgandhi@unomaha.edu

ABSTRACT

Ensuring bridge safety is crucial for transportation in many countries, including the US. In fact, the ASCE gave an overall cumulative grade of C– for US infrastructure, in large part due to deteriorating bridges. Unsafe bridges within transportation routes pose significant challenges to stakeholders including traffic safety concerns, environmental hazards, and economic implications. The collapse of Minneapolis I-35W bridge highlights these issues. Our research encourages stakeholders to make informed decisions by taking into account bridge risk factors and the sensitivity of transported materials (e.g., hazardous materials, heavy equipment). This approach enables strategies like load balancing, rerouting enforcement, and prioritizing maintenance tasks to be applied effectively. We generated a bridge-health enriched geographic information system (GIS) dataset that integrates bridge information from the National Bridge Inventory (NBI) with GIS data from OpenStreetMap. We analyzed example origin-destination (O-D) pairs in Nebraska to demonstrate how different routes in Nebraska exhibit, besides differences in travel distance, varying statistics on bridge health scores. We also introduce a simulation platform for route planning and visualization currently in active development. We finally discuss the challenges we faced and opportunities ahead in combining these distinct datasets for sustainable route planning and infrastructure maintenance.

1. INTRODUCTION

Infrastructure development has been an integral part of human progress throughout history. From ancient civilizations to modern-day transportation systems, infrastructure has played a vital role in shaping society and driving economic growth [1]. Multiple reports have highlighted the strong correlation between infrastructure development and the economic development of countries [2, 3, 4]. Transportation infrastructures in particular, bridges specifically, are vital components in facilitating the movement of goods and people across different locations. However, the safety and sustainability of bridge infrastructures face multiple challenges such as deterioration and eventual collapse due to factors like aging infrastructure, lack of maintenance, and design flaws [5, 6, 7, 8]. These challenges have significant consequences including traffic accidents, environmental hazards, and economic implications. For instance, the collapse of the I-35W Bridge in Minneapolis in 2007 claimed the lives of 13 people, injuring more than hundred, and caused widespread damage to vehicles and the surrounding area [6]. Another recent example is the collapse of Fern Hollow Bridge, a 52-year-old bridge in Pittsburgh that had a poor condition rating [6]. These incidents highlight the need for developing innovative solutions to address the challenges posed by bridge health concerns. However, for a large country like the US

with over 600,000 bridges spread across the country including in rural areas with a median age of approximately 44 years, keeping all of them in optimal condition is challenging [9]. The ASCE further reports that 7.5% of the nation's bridges are considered structurally deficient [9].

One way to tackle these challenges is by integrating bridge health information into route planning systems. By utilizing bridge health condition data and rerouting enforcement strategies, we can reduce the likelihood of bridge collapse incidents, facilitate informed maintenance, and ensure safer and more efficient transportation systems. Unfortunately, most popular route planning systems including Google Maps do not offer bridge health condition information in route planning or in suggesting alternative routes. We propose that incorporating such vital bridge health information in route planning is crucial in various settings. These include maintaining a sustainable road infrastructure, transportation of goods and services that require high attention including transportation of hazardous materials, or the movement of heavy and military equipment, or in disaster scenarios with a requirement for offline accessibility. In light of this, we leveraged a GIS data and a separate bridge health condition information from two distinct sources to generate a bridge health condition enriched GIS dataset that can be used in bridge health-informed route planning. We look at the state of Nebraska as a case study and used OpenStreetMap as our source of GIS information. The National Bridge Inventory (NBI) [10], an open data from a database maintained by the Federal Highway Administration (FHWA) with a rich information about bridges in the US was used as our source of bridge health information. The NBI dataset includes health condition ratings of different bridge components such as superstructure, sub-structure, and bridge deck, among others [10]. Using this data, we demonstrate how different routes exhibit varying statistics on bridge health scores and distances. Our results show that shortest route is not necessarily safest route and vice versa.

The remainder of this paper is organized as follows: in the next section, we discuss related work about bridge health, transportation networks, and bridge collapse of investigation. In Section 3, we discuss our approach of enriching the GIS data with bridge health condition information from NBI and the challenges in merging these distinct datasets. In Section 4 we demonstrate the usability of the generated dataset by taking sample Origin-Destination (O-D) pairs and generating alternative routes using the Dijkstra algorithm by iteratively taking out bridges from previous best routes.

The case-studies presented in this paper include O-D pairs that represent routes between two cities and a remote town in Nebraska: Omaha, Lincoln, and Scottsbluff, NE. We will show that routes between cities and to a remote town involve bridges with a poor deck condition rating of 4.

In Section 5 we discuss implications of our results and their potential applications while also stating limitations of our work. In Section 6 we highlight a visualization platform being developed that enables simulation of route planning in presence of various scenarios including bridge health condition, convoy capacity, and presence of adversaries. Finally, Section 7 concludes the paper.

2. RELATED WORK

The study of bridge health and its impact on transportation networks has been a focus of extensive research over the years [5, 6]. One such paper is a recent work by Crawford [6], which provides a comprehensive overview of bridge deterioration and failures. This work highlights the importance of understanding the causes that contribute to bridge collapse and the need for

ongoing maintenance and monitoring to prevent such incidents and bridge failures in general. The study also emphasizes the economic and social consequences of bridge collapses, including disruption to transportation networks, loss of life, and damage to property. Another related study by Morgese et al. [7], conducted a post-collapse analysis of the Ponte Morandi bridge in Genoa, Italy which showed that the bridge did not receive proper maintenance despite being in service for over 50 years. These studies highlight the importance of understanding conditions of bridges and their impact on transportation networks.

Various studies have investigated impact of bridges on transportation networks [11, 12, 13, 14]. In [11] Researchers framed transportation networks as lifeline systems and studied network level consequences of bridge closures with a framework and resilience index they proposed. Their results showed that closure scenarios on 10 bridges incurred significant regional resilience losses. Zhu et al. [12] studied the travel impact of the I-35 bridge on the society assessing its impact both during the collapse and reopening of the bridge. Other works have shown that machine learning algorithms can be used to understand and predict the likelihood of a bridge needing maintenance based on factors such as age, type, and usage [15, 16, 17]. In a loosely related work to this paper, situation aware military convoy routing algorithms for military applications were also proposed [18]. The algorithms proposed considered 10 input parameters including route width and length, threat, hostility, and infrastructure type [18]. Simulation results on a grid with UAV threats and up to 18 edges show how to avoid threats using multiple parameter configurations. This body of literature shows the importance of detailed route information in planning and maintenance activities and the impact of bridge health condition on overall routing.

3. ENRICHING OSM WITH NBI DATA

Our aim in this paper is to show the importance of incorporating bridge health information in route planning. With this in mind, we first need to find a large, openly available, offline accessible, and detailed GIS dataset that can ideally provide detailed bridge information including its name, health condition, and coordinates providing a comprehensive coverage.

3.1 Openstreetmap Gis Data

OpenStreetMap [19] is the largest open geographic information dataset that can satisfy most of our requirements. Various research works and applications have leveraged this large geographic database for a range of problem domains [20, 21, 22]. Despite its limitations, such as presence of incomplete or inaccurate data [23], the OpenStreetMap project has seen exponential growth since its inception, providing a free and global map coverage that can serve as alternative to its commercial counterparts like Google Maps.

However, bridge information is one of the least covered pieces of information on OpenStreetMap (OSM). Although most bridges are labeled correctly, these bridges may miss fields such as type of bridge, construction material, year built etc. These fields are superfluous to the average OSM user, but are vital when performing bridge-aware routing. Besides, more authoritative information such as health condition rating of bridge components is not available. Despite the shortcomings, the ease of access and comprehensive GIS coverage make OSM the most suitable data source for GIS research and offline applications.

When working with OSM data, there are several different file formats available. Two formats in particular, OSM XML and PBF (Protocol buffer Binary Format), are most common. The OSM

XML format can be considered a “raw data” format, listing all entries in a human-readable XML format. OSM in this format is generally larger in size compared to other OSM formats and is slower to work with due to the nature of its XML structure. To increase performance when reading and writing OSM data, we use the PBF format. The PBF format is much faster and more compact than XML format but it is not human readable [24]. Therefore, we opt to use the OSM XML format for correctness checking and data analysis, and the PBF format for data parsing and editing when possible, with tools.

OSM data, regardless of the file format, comprises of a collection of objects: Nodes, Ways, and Relations. In OSM, Nodes are fundamental data items in OSM. A Node has latitude and longitude coordinates and an ID, describing a single point on a map. Ways are a collection of Nodes, that can describe objects such as roads, buildings, highways, and importantly, bridges. Relations can contain Ways and Nodes and are used to define geographical relations between OSM objects. All OSM objects can have tags with a key-value pair used to define metadata for the object such as street name, highway number, bridge information, etc. In OSM, bridges can be identified as a Way object with a bridge tag, and key-value pair "bridge"="yes". Name tags are applicable to most OSM objects and are stored similarly, e.g. "name"="Highway 72".

3.2 National Bridge Inventory Data

The National Bridge Inventory (NBI) is a database created and managed by the Federal Highway Administration that contains a detailed record of bridges and bridge repairs in the United States. Though the dataset is publicly available, only the Federal Highway Administration can update it, and most bridges are only updated biannually or longer.

The NBI database is vast, containing data on nearly all bridges in the United States over 20 feet long used for traffic [25]. It contains over 600 thousand entries, having over 100 fields. Each entry describes a single real-world bridge including year constructed, health ratings, material types, etc. Since bridges will be matched between the datasets by location and name, the most crucial matching fields are latitude and longitude for location and carried-by for street name. It is important to note that in NBI, a bridge is represented by a single point coordinate only. This record could be at any point on or near the real-world bridge. NBI entries are provided by the bridge's owning agency and several standards put in place by the Federal Highway Administration must be followed when collecting bridge data. Unfortunately, these standards enforce little when collecting a bridge's location; It is fully up to the bridge's owning agency to best decide how the bridge's latitude and longitude should be measured [26, 27]. A standard from 1995 indicates a required amount of precision, but does not describe which point of the bridge to use for data collection [25]. This means it is difficult to be certain how this data is collected and consequently, how to correct any inconsistencies in latitude and longitude. A bridge's positional data is collected at the time of the bridge's initial inspection, and is rarely updated afterward [26].

Though NBI data is frequently used for bridge health analysis, inaccuracies occur frequently in the data due to the lack of standards during the data collection process. This is especially apparent when inspecting NBI positional data against the OSM dataset. In some cases, the positional data of NBI was found to be invalid in 28% of entries [28]. Inaccuracies in NBI have even led to the collapsing of some bridges that may not have been properly inspected [29]. Our implementation must find a novel method of avoiding these inaccuracies. We do so by relying on the strength and accuracy of OSM to find placement errors in NBI and determine where they lay.

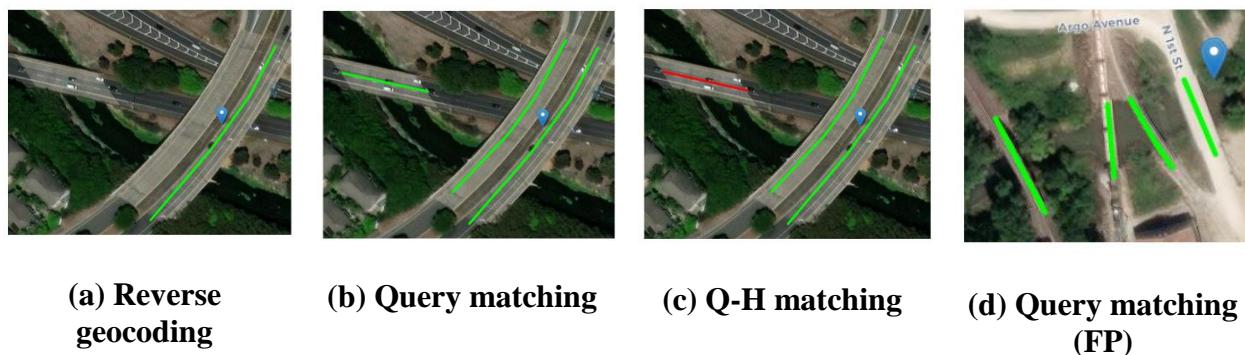


Figure 1. Example scenario and approaches, where a bridge coordinate from NBI (blue) is matched against bridges from OSM using different approaches: (a) Reverse geocoding mismatch with parallel ways in OSM, (b) Query matching using Overpass with false positives (FP), (c) Q-H matching method yielding multiple matches, filtering out FP, (d) Multiple FPs for a single NBI entry when query matching alone is used.

3.3 Challenges Combining the two Datasets

As described above, the two datasets (i.e. NBI and OSM) are of varying quality and format. Hence, it is important to understand the nuances of each. As the datasets are disparate with distinct coding techniques the process of adding a highly accurate NBI data to the OSM should not be considered “merging” or “combining” in a traditional sense, rather these words imply all data across each dataset is put together to create a new dataset via a certain key. This also generally means that the two datasets are similar in format and content. In our case, not all data can be used, nor are the two datasets similar. The process here should be considered “data enrichment,” where the OSM data is selectively enriched with bridge information from a vastly different data source, NBI. Via this data enrichment process, detailed bridge data is added to corresponding OSM bridges in the form of tags. Matching is confirmed if and only if an OSM way has a bridge tag and is a roadway. NBI data is added as tags. For example the tag “nbi:sub-cnd=5” can be added to denote a score of 5 for a bridge’s NBI sub-structure rating. Given that Openstreetmap is a community effort, uniform coverage of areas is inherently a challenge, especially for areas like Nebraska where 72% of its communities have populations below 800 people [30].

3.4 Q-H Approach to Reduce False Positives

As shown in Figure 1, multiple way objects are in proximity to an NBI point coordinate (colored blue). A successful query using different methods results in different possible matches. Using the popular reverse-geocoding approach with Nominatim API results in a match of just one way object closest to the point coordinate ignoring the other valid way point, as shown in figure 1.a. This approach resulted in only 4,697 matched roadway bridges out of 15336 NBI and 10384 OSM bridges in Nebraska, i.e. less than a third (30.6%) of NBI entries. The second approach used is query matching using overpass API, this approach returns more than one match giving the possibility of to consider multiple matches even when the closest way object is invalid. The disadvantage of this approach is the fact that it returns false positives (fig 1.d).

To address the above challenges a heuristics-based query matching referred to as query-heuristics (Q-H) was introduced. With this approach a scoring mechanism is used to pick OSM objects that maximize confidence and reduce false positives. This approach enabled us to match 56.9% of entries from NBI onto the entries in OSM, excluding culverts. The heuristic part involves average of two scoring mechanisms: distance scoring, and name scoring. With distance (geodesic distance) scoring we measure the proximity of the candidate objects with the following formulation.

$$dist_score = \frac{t}{d + t}$$

Where d is the closest geodesic distance between an NBI point and any point within the bridge polygon, as opposed to centroid of the polygon, and t is a threshold value (we found t=20 a suitable threshold). We considered matches with a distance score above 0.5 as candidates prioritizing bridges found within 20 meters of the NBI entry.

Name scoring uses Sørensen-Dice string matching [31] to measure similarity between a bridge's name in OSM vs a name in NBI (e.g. Sa = "654 AV-707 / 708 R" in OSM vs in Sb= "654 Avenue" in NBI) as follows.

$$sd_score = \frac{2|A \cap B|}{|A| + |B|}$$

Where A and B are the sets of characters found in strings Sa and Sb respectively, and |A| and |B| are the cardinalities of sets A and B respectively. Table 1 summarizes the results.

4. ROUTE GENERATION AND ANALYSIS USING THE ENRICHED DATASET

We now demonstrate the use of NBI-enriched GIS data in route generation. For this purpose we first selected three sample points in Nebraska representing the cities Omaha (OM) and Lincoln (LN), and a remote location called Scottsbluff (SF). After identifying the coordinates of these locations we generated a graph with the GIS data, which gives us vertices (V) and edges (E) as is common in traditional graph theory G(V,E). We then proceed to generate ten iterations of finding the shortest route between pairs OM-LN and OM-SF, identifying and extracting tagged NBI bridges, taking out all these identified bridges, and recomputing the graph G. The iterative methodology is diagrammatically described in Figure 2.

Figure 2 illustrates the Iterative process of route planning, in which the best route is determined at each stage. Figure 2.a displays 5 different routes that can be taken as alternatives. Applying Dijkstra shortest path algorithm between s and t results in s->A0->t as best route, after taking out bridge A0, i.e. removing p2 from a way point (P1, P2), and recomputing the graph shows the previous route is no longer viable and results in the route shown on figure 2.b, iterating over this process until the last bridge on E0 as shown on Figure 2.f. In our experiment on the actual GIS data, we performed ten iterations for the two O-D pairs. In the following we will discuss results in applying our method on the actual GIS data from Nebraska.

Table 1: Comparison of the matching techniques on 11,122 non-culvert NBI entries and 10,384 OSM bridges. The table shows the improvement in Q-H matching over RG matching (4,483 added vs 6,329), while reducing false positives introduced in Query matching (7,905 vs 6,509 received).

Iteration	Match Rates			
	NBI		OSM	
	Entries	Added	Entries	Received
RG Matching	11,122	4,483	10,384	4,483
Query Matching	11,122	6,329	10,384	7,905
QH Matching	11,122	6,329	10,384	6,509

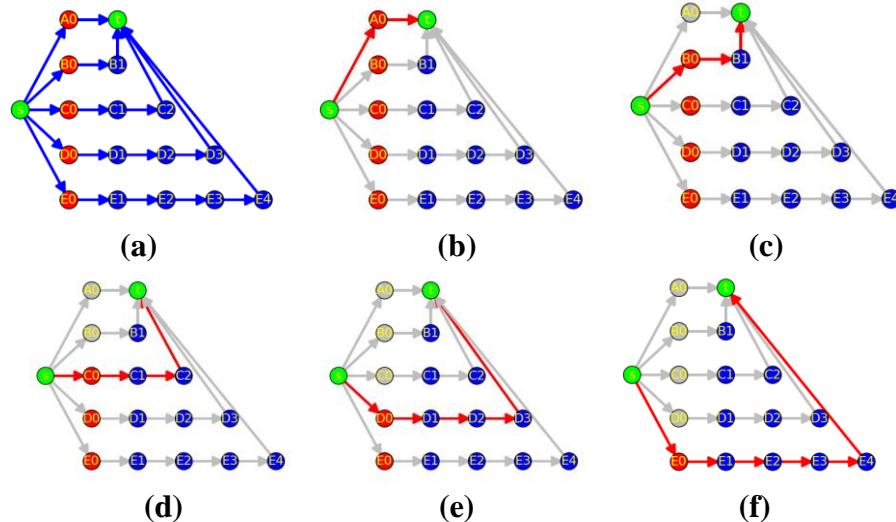


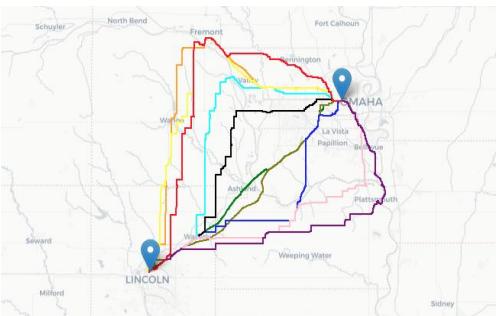
Figure 2. Route planning with iterative graph regeneration. The green circles labeled s and t represent source and destination nodes (S-D), the red circles represent bridges, blue arrows represent possible routes forming edges, while red arrows show selected best routes at each iteration using Dijkstra algorithm. Gray circles represent taken-out bridges after each iteration.

4.2 Results with Sample O-D Pairs

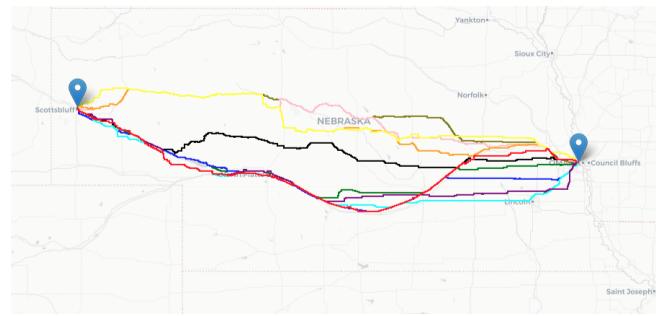
The generated ten alternative routes between OM-LN and OM-SF pairs are shown on Figure 3. As shown on Figure 3 the shortest routes tend to drastically change as the bridges are removed from the graph. We performed preliminary analysis of these routes consisting bridges to study the length of the routes, number of identified bridges, and their health condition rating taking the deck condition rating (nbi:deck-rating) from our enriched GIS dataset. Note that this method can be extended to other components of the bridge such as super-structure and sub-structure as they

are already integrated. From the results in Figure 3, we can visually inspect that bridges alter routes. Further analysis on Figure 4.a shows that the length of routes between OM-LN pairs (nearby and two most populous cities in the state) also drastically increases as we iteratively remove bridges.

An interesting observation here is that the condition rating denoted by the minimum deck condition rating of bridges along the routes does not follow a predictable pattern where the best route and the tenth route both involve a deck condition rating of 5 (FAIR CONDITION - all primary structural elements are sound but may have minor section loss, cracking, spalling or scour), while intermediary routes have a condition rating as low as 4 (POOR CONDITION - advanced section loss, deterioration, spalling or scour). Given the distance between OM-SF pairs is significantly longer than the previous pairs there seem to be more alternatives with fairly more equivalent distance. An interesting observation in the second pair is that the best route involves a bridge with a POOR condition rating. In both cases we can see that bridges are critical components in creating connections while best routes did not show better condition rating. Figure 5 shows a more detailed result on the distribution of deck-condition rating of routes with corresponding colors on Figure 3.



(a) Set of iterative shortest routes between Omaha and Lincoln.



(b) Set of iterative shortest routes between Omaha and Scottsbluff.

Figure 3. Results of iterative graph generation and route planning taking out bridges identified in each shortest route. The figure shows the shortest path in the first iteration (green) tends to change gradually, finally taking the routes colored red on the tenth iteration.

5. DISCUSSION

The results from the generated ten alternative routes with between the OM-LN and OM-SF pairs mainly demonstrate the importance of bridge health information. A notable observation from Figure 3 is that the computed shortest routes undergo significant changes as bridges are removed from the graph. Our results further acknowledged the fact that bridges play important roles in connecting communities. We have also demonstrated the ease of identifying the health condition rating of bridges in route computations, this demonstrates its potential for bridge health-informed route planning applications. As results on Figure 4 suggest best routes in terms of distance are not safest nor did receive special treatment to improve their condition rating. This is further demonstrated in Figure 5 where we can observe condition rating as low as 4 in some of the best routes, while also observing a top condition rating among the last routes. While the

results are preliminary, we can see the potential of enriching GIS dataset with highly up to date and highly accurate bridge health information for route planning. This will have a profound impact in special transportation scenarios such as transportation of hazardous materials [8, 32] or designing routes for military convoys. For example, a route planner with the objective of avoiding bridges based on their condition rating can easily incorporate condition rating of the desired component(s) as a weight (cost) function in route computation [33]. This information is also valuable for policy makers. Taking the Omaha-Lincoln pair as an example, a load balancing recommendation system can be implemented to alternate traffic between route 0 and route 1, which have fairly closer travel distances. A maintenance prioritization system can also be developed to identify the risk level of a route based on assessment of involved bridges.

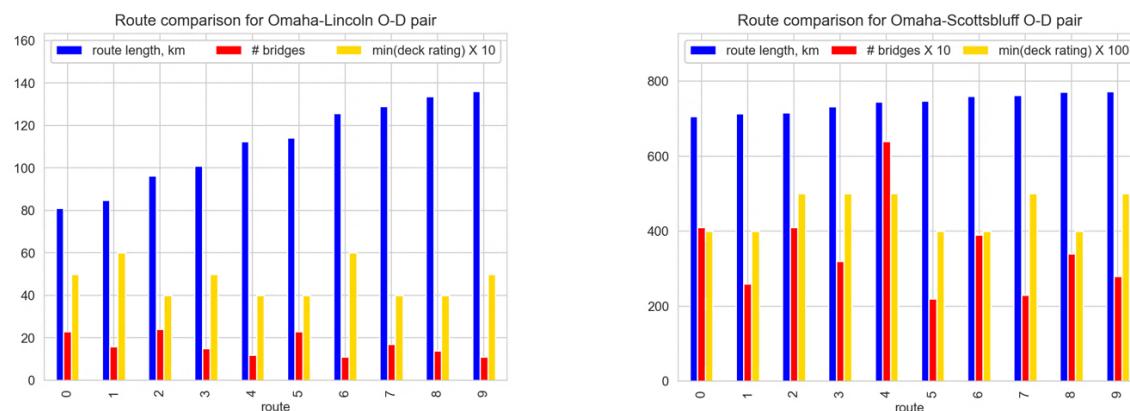


Figure 4. Comparison of identified routes on each iteration. The blue color shows the length of route in Kilometers, while the red and yellow colors represent number of bridges and minimum deck condition rating of bridges along the routes. The red and yellow bars are scaled for ease of visualization as shown in the legend boxes.

However, our proposed approach is not without limitations. First, the bridge health data from NBI is a less frequently updated inspection database. This brings timeliness limitations. Second, matching the two disparate datasets might not be reliable because of multiple reasons. These include unavailability of datasets on remote locations, and the problems due to inconsistencies we already discussed. The first limitation of our approach can be addressed by using a more reliable and frequently updated bridge health dataset such as continuous measurement-based approaches with sensors, though this solution might bring scalability issues. Future work can address second limitation of our approach with different approaches, such as crowdsourcing to properly tag remote bridges in OSM and using more finetuned matching approaches that take advantage of computer vision, NLP, and related machine learning techniques.

6. VISUALIZATION AND SIMULATION PLATFORM

With the objective of advancing bridge-health informed routing methods, we have been developing a route planning and visualization system that can work offline (in lieu of Google maps). This standalone system takes convoy information with desired source and destination locations as inputs and visually shows the selected routes including the health condition of each bridge along the computed routes. Future iterations will include presence of adversarial

conditions [18] which will affect computation of route costs. The UI of current status of this project developed with react framework is shown on Figure 6.

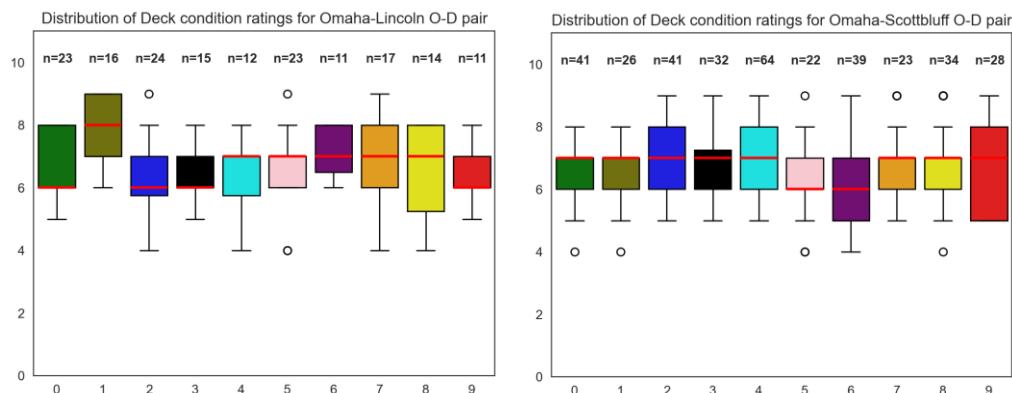


Figure 5. Distribution of condition rating of the decks of bridges along the selected best routes in each iteration ranging from zero to nine. The figure on the left shows the distribution for Omaha-Lincoln pair, while distribution for Omaha-Scottsbluff is shown on the right. The number of bridges is denoted by n on top of each distribution. The colors correspond to the routes shown on figure 3.

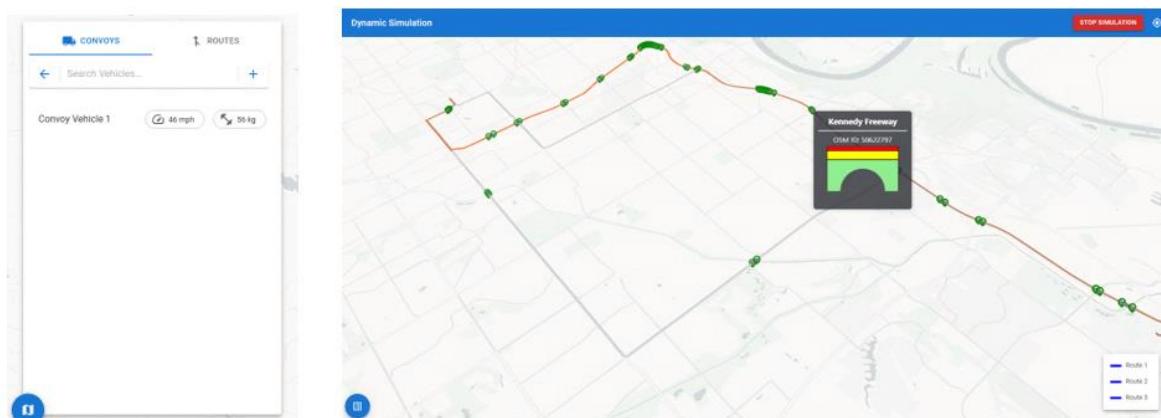


Figure 6. screenshots showing the simulation platform in development. The left figure shows the specification of a convoy for simulation, the right figure shows the alternative routes for an S-D pair of the convoy providing a detailed bridge information.

7. CONCLUSION

In this study we call for improving route planning and maintenance in transportation networks via enriching GIS data with bridge health information from external authoritative sources such as NBI. With real experiments we show that best routes computed with common shortest path algorithms do not necessarily mean safest routes. We showed promising opportunities of leveraging these distinct datasets to improve safety and logistics of transportation networks. We also showed limitations of our proposed approach which include the availability of datasets on remote locations, inconsistencies in data quality, and reliability of

Downloaded from ascelibrary.org by Texas State University on 06/14/24. Copyright ASCE. For personal use only; all rights reserved.

health information sources. We also indicated how future work can address these limitations with approaches such as using alternate bridge health datasets, crowdsourcing to properly tag remote bridges in OSM, and improving the matching approach with emerging AI techniques including computer vision. Our study contributes to the existing literature by providing a methodology for enriching specific elements of GIS graph with bridge health information for transportation networks. We hope this research lays the groundwork for a safer and more efficient route planning, offering potential applications in advanced traffic management systems and optimized infrastructure investment plans.

ACKNOWLEDGEMENT

This work is partially supported by contracts W912HZ21C0060 and W912HZ23C0005, US Army Engineering Research and Development Center (ERDC), and Award Number 1762034 from the National Science Foundation.

REFERENCES

- [1] Agbelie, B. R. D. K. "Economic impacts of transportation infrastructure investment across countries: An empirical analysis." (2013).
- [2] Hooper, E., S. Peters, and P. A. Pintus. "The impact of infrastructure investments on income inequality: Evidence from US states." *Economics of Transition and Institutional Change* 29.2 (2021): 227-256.
- [3] Munnell, A. H. "Policy watch: infrastructure investment and economic growth." *Journal of economic perspectives* 6.4 (1992): 189-198.
- [4] Seidu, R. D., et al. "The impact of infrastructure investment on economic growth in the United Kingdom." *Journal of Infrastructure, Policy and Development* 4.2 (2020): 217-227.
- [5] Deng, L., W. Wang, and Y. Yu. "State-of-the-art review on the causes and mechanisms of bridge collapse." *Journal of Performance of Constructed Facilities* 30.2 (2016): 04015005.
- [6] Crawford, K. C. "Bridge Deterioration and Failures." *Failure Analysis*. IntechOpen, 2023.
- [7] Morgese, M., et al. "Post-collapse analysis of Morandi's Polcevera viaduct in Genoa Italy." *Journal of Civil Structural Health Monitoring* 10 (2020): 69-85.
- [8] Press, T. A. (2023). A train carrying hazardous materials plunges into Yellowstone River after bridge fails. Available at: <https://www.npr.org/2023/06/24/1184189040/a-train-carrying-hazardous-materials-plunges-into-yellowstone-river-after-bridge> (Accessed: 10 December 2023).
- [9] ASCE. (2023). Infrastructure report card, American Society of Civil Engineers. <https://www.infrastructurereportcard.org/cat-item/bridges-infrastructure/>. Accessed Dec 2023.
- [10] National Bridge Inventory, Federal Highway Administration. Available online: <https://www.fhwa.dot.gov/bridge/nbi/ascii2023.cfm> (accessed on 1 Dec 2023).
- [11] Twumasi-Boakye, R., and J. O. Sobanjo. "Resilience of regional transportation networks subjected to hazard-induced bridge damages." *Journal of Transportation Engineering, Part A: Systems* 144.10 (2018): 04018062.
- [12] Route planning and analysis: Zhu, S., et al. "Travel Impacts and Adjustment Strategies of the Collapse and the Reopening of the I-35W Bridge." *Network Reliability in Practice: Selected Papers from the Fourth International Symposium on Transportation Network Reliability*. Springer New York, 2012.

- [13] Faturechi, R., and E. Miller-Hooks. "Travel time resilience of roadway networks under disaster." *Transportation research part B: methodological* 70 (2014): 47-64.
- [14] Capacci, L., F. Biondini, and D. M. Frangopol. "Resilience of aging structures and infrastructure systems with emphasis on seismic resilience of bridges and road networks." *Resilient Cities and Structures* 1.2 (2022): 23-41.
- [15] Masum, M., et al. "A Novel Machine Learning Based Framework for Bridge Condition Analysis." *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022.
- [16] Kale, A., Y. Kassa, B. Ricks, and R. Gandhi. (2023). A Comparative Assessment of Bridge Deck Wearing Surfaces: Performance, Deterioration, and Maintenance. *Applied Sciences* (2076-3417), 13(19).
- [17] Ardani, S., S. Eftekhar Azam, and D. G. Linzell. (2023). Bridge Health Monitoring Using Proper Orthogonal Decomposition and Transfer Learning. *Applied Sciences* (2076-3417), 13(3).
- [18] Ruuben, T., and O. Kreison. "Route planning in asymmetric military environments." *Second International Conference on Future Generation Communication Technologies (FGCT 2013)*. IEEE, 2013.
- [19] Bennett, J. *OpenStreetMap*. Packt Publishing Ltd, 2010.
- [20] Feldmeyer, D., et al. (2020). Using OpenStreetMap data and machine learning to generate socio-economic indicators. *ISPRS International Journal of Geo-Information*, 9(9), 498.
- [21] Cohen, A., and S. Dalyot. "Route planning for blind pedestrians using OpenStreetMap." *Environment and Planning B: Urban Analytics and City Science* 48.6 (2021): 1511-1526.
- [22] Grinberger, A. Y., et al. "OSM Science—The Academic Study of the OpenStreetMap Project, Data, Contributors, Community, and Applications." *ISPRS International Journal of Geo-Information* 11.4 (2022): 230.
- [23] Haklay, M. "How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets". en. In: *Environment and Planning B: Planning and Design* 37.4 (Aug. 2010), pp. 682–703. issn: 0265-8135, 1472-3417. url: <http://journals.sagepub.com/doi/10.1068/b35097> (visited on 08/24/2023).
- [24] Owusu, C., et al. "Geocoding Fundamentals and Associated Challenges." *Geospatial data science techniques and applications* 118 (2017): 41-62.
- [25] "Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges". en. In: (May 1995). URL: <https://www.fhwa.dot.gov/bridge/mtguide.pdf>.
- [26] BIRM (Bridge Inspector's Reference Manual). en. (Mar. 2023). url: www.fhwa.dot.gov/bridge/nbis/pubs/nhi23024.pdf. (visited on 08/04/2023).
- [27] National Bridge Inspection Standards. (May 2022). url: www.federalregister.gov/documents/2022/05/06/2022-09512/national-bridge-inspection-standards (visited on 08/04/2023).
- [28] Chase, S. B., C. Nutakor, and E. P. Small. "An in-depth analysis of the National Bridge Inventory database utilizing data mining, GIS and advanced statistical methods". In: 2000. url: <https://api.semanticscholar.org/CorpusID: 54968322>.
- [29] Dedman, B. Bridge collapse revealed holes in federal data. en. Aug. 2008. url: www.nbcnews.com/id/wbna25956395 (visited on 08/04/2023).
- [30] Drozd, D., and J. Deichert. Nebraska County Population Projections: 2010 to 2050. en. Tech. rep. 258. Dec. 2015.
- [31] Dice, L. R. "Measures of the Amount of Ecologic Association Between Species". In: *Ecology* 26.3 (1945), pp. 297–302.

- [32] Ovidi, F., V. van der Vlies, S. Kuipers, and G. Landucci. (2020). HazMat transportation safety assessment: Analysis of a “Viareggio-like” incident in the Netherlands. *Journal of Loss Prevention in the Process Industries*, 63, 103985.
- [33] Boeing, G. (2017). “OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks.” *Computers, Environment and Urban Systems*. 65, 126-139.

Time Series Analysis of Highway Work Zone Crashes in North Carolina

Kazi Tahsin Huda¹; Yuting Chen²; Don Chen³; and Srinivas S. Pulugurtha⁴

¹Dept. of Engineering Technology and Construction Management, Univ. of North Carolina at Charlotte, NC. Email: khuda@charlotte.edu

²Dept. of Engineering Technology and Construction Management, Univ. of North Carolina at Charlotte, NC (corresponding author). Email: ychen106@charlotte.edu

³Dept. of Engineering Technology and Construction Management, Univ. of North Carolina at Charlotte, NC. Email: dchen9@charlotte.edu

⁴Dept. of Civil and Environmental Engineering, Univ. of North Carolina at Charlotte, NC. Email: sspulugurtha@charlotte.edu

ABSTRACT

Highway work zones present significant safety concerns. This study aims to identify spatial and temporal factors that affect the number of highway work zone crashes in North Carolina (NC). The data from 2000 to 2020 were provided by the NC Department of Transportation (NCDOT). Descriptive statistics, exploratory analysis, and time series analysis were employed. It was found that crashes occurred most frequently in construction-related work zones, work zones with ongoing activities, during clear weather conditions, dry surface conditions, daytime, and right next to the work area. The time series analysis identified multiple trends and a seasonal pattern for monthly crash frequencies, which were used to develop a Holt-Winters time series crash prediction model. The model findings indicate that the crash trend and seasonality are important factors in highway work zone crash modeling along with geometric (e.g., lane width) and traffic variables (e.g., volume).

INTRODUCTION

Highway work zones are defined as areas on roadways that are undergoing repair, maintenance, utility, or construction work. These work zones are separated from the passing traffic using traffic cones or barriers (e.g., concrete barriers). Speed limits are generally lower in work zones with higher fines imposed for speeding offenses. Despite these precautions, work zones still experience a good number of crashes each year. Some researchers have found that highway work zone crashes are more injurious than non-work zone crashes, especially those involving large trucks (Khattak and Targa, 2004). With the large number of fatal crashes and fatalities that occurred at highway work zones in the United States (774 fatal crashes and 857 fatalities in 2020) (FHWA, 2020), it is necessary to understand the crash trends in work zones and identify factors related to these crashes, such as weather conditions, road features, and light conditions, for improving highway work zone incident prevention and response.

A few studies (Daniel et al., 2000; Osman et al., 2016; Weng et al., 2016; Yu et al., 2022; Zhang & Hassan, 2019) have identified different spatial factors that influence work zone crashes. For example, Weng et al (2016) identified fog/snow to be associated with higher number of crashes in Michigan. Time series analysis was also used in a few traffic safety studies (Deretić et al., 2022; Khan et al., 2022; Theofilatos et al., 2018) as well. For example, Deretić et al. (2022) used the seasonal crash patterns to develop crash prediction models for Belgrade Serbia.

However, no study specifically identified spatio-temporal factors affecting highway work zone crashes for the state of North Carolina (NC). Also, no studies were found to utilize time-series prediction models for highway work zone crash forecasting for NC. This study, therefore, aims to identify different spatial and temporal factors that affect the occurrences of highway work zone crashes in NC and to develop time-series crash prediction model to predict future highway work zone crashes in the state.

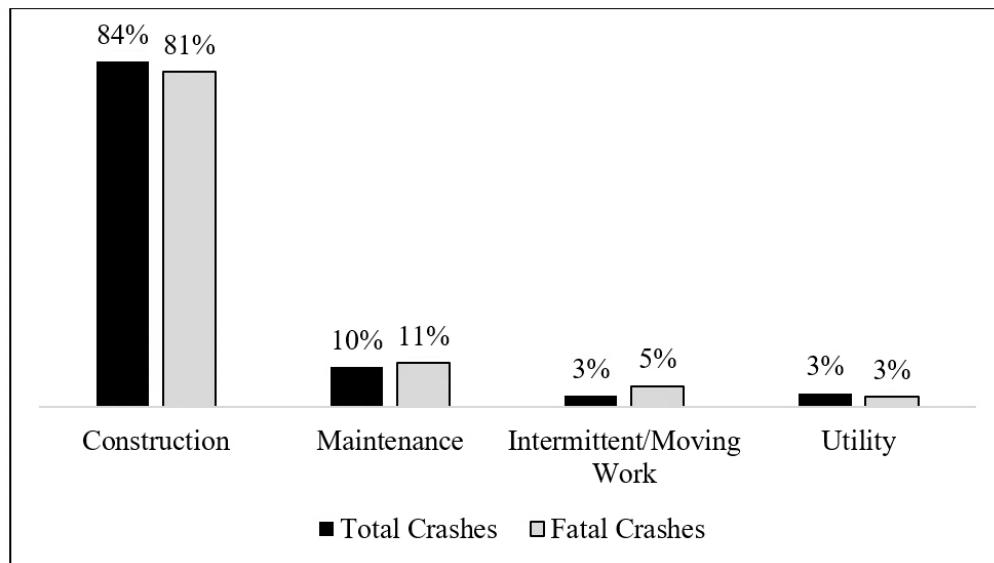
METHOD

Detailed work zone crash data from 2000 to 2020 were used for analysis in this study. Data from 2012 to 2020 were collected from a NC Department of Transportation (DOT) website (NC DOT, 2020). Data from 2000 to 2011 were provided by the Transportation Mobility and Safety Division of NC DOT on request. The study used descriptive analytics, exploratory data analysis, and time-series analysis to analyze the data. Descriptive analytics is a process of identifying trends and relationships of different factors in a dataset. Exploratory data analysis is the process of summarizing the main characteristics of a dataset by using data graphs and visualization. Time series is data collected over a specific time interval and time series analysis is the process of predicting future values using the time series (Nugus, 2009). Furthermore, a Holt-Winters time series crash prediction model was developed for a preliminary analysis to investigate whether the trend and seasonal pattern of the monthly highway work zone crash data could be used for predicting future crash frequencies.

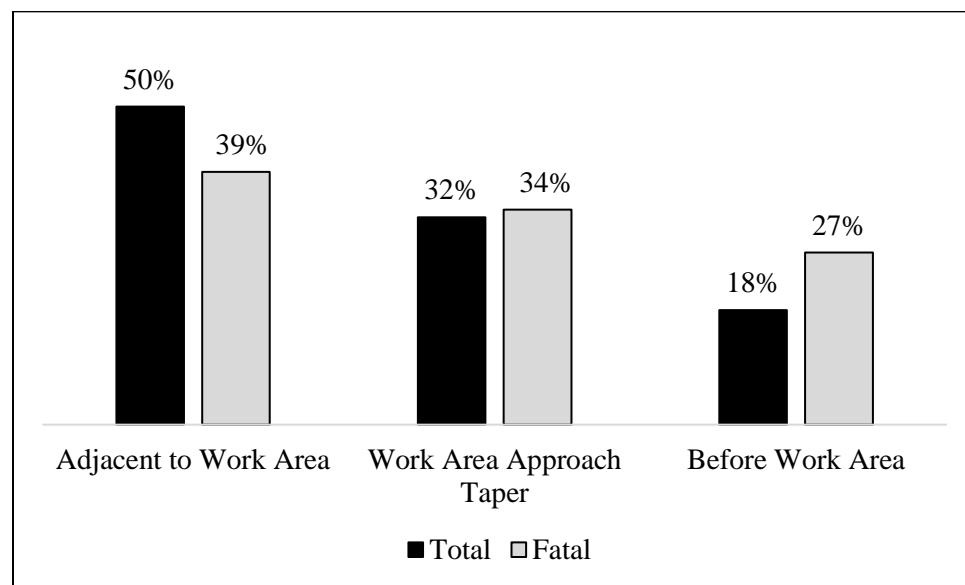
RESULTS

This section presents the results from the analysis, which are divided into five sub-sections including types of work zone, work zone activity and location, weather and lighting conditions, surface conditions and road features, and time-series analysis.

Types of Work Zone. There are four types of NC highway work zones: construction, maintenance, intermittent/moving work, and utility. Figure 1 shows the number of total and fatal crashes by different work zone types. There were 109,716 total crashes and 577 fatal crashes from 2000 to 2020. About 84 percent (92,030) of the total and 81 percent (470) of the fatal work zone crashes occurred in a construction related work zone. The numbers in parenthesis are the corresponding number of crashes. Construction work zones are work zones where there is construction of a new facility (e.g., building a new bridge, adding a new lane, etc.). These work zones are set up for longer times as construction of new facilities generally takes a much longer time. The highest frequency of crashes for construction work zones could be caused by this longer exposure time of vehicle traffic. Maintenance related work zones followed second with about 10 percent (10,783) of the total and 11 percent (62) of the fatal crashes. Maintenance related work zones are set up due to some maintenance work on the existing infrastructure. Intermittent or moving work zones and utility work zones both accounted for about 3 percent of the total crashes (intermittent/moving work: 3,159 crashes; utility: 3,744 crashes). Intermittent work zones are set up for tasks such as lane markings, whereas utility related work zones are set up for any utility related work. Both work zone types accounted for up to 5 percent of the fatal crashes. These work zones are set up for the shortest amount of time and this could be the reason for the least percentages of total and fatal crashes.

**Figure 1: Crashes by Work Zone Types**

Work Zone Activity and Location. The data showed that 60 percent (66,121) of the total and about 50 percent (286) of the fatal crashes happened in an active work zone. These high percentages could possibly be due to drivers being distracted by the construction activity in the work zones. This could be further supported by Figure 2, which illustrates that about 50 percent (54,541) of the total and about 39 percent (94) of the fatal crashes occurred right adjacent to the work area. Other work zone areas, namely the work area approach taper (total: 32 percent; fatal: 34 percent) and before-work area (total: 18 percent; fatal: 27 percent), had much fewer crashes. The data also shows that the work area approach taper and the before-work area had higher percentages for fatal crashes than for total crashes.

**Figure 2: Crashes by Work Zone Areas**

Weather and Lighting Condition. Table 1 summarizes the total and fatal work zone crashes caused by weather conditions. About 75 percent (81,911) of the total and 78 percent (452) of the fatal crashes occurred in clear conditions, followed by cloudy conditions with about 16 percent of the total and 17 percent of the fatal crashes, and rainy conditions with about 8 percent of the total and 4 percent of the fatal crashes. The highest percentages of total and fatal crashes in clear conditions indicate the possibility that drivers might pay less attention to road conditions, and thus not closely follow the posted work zones speed limits, resulting in an increased chance of crashes.

Table 1: Work Zone Crashes by Weather Conditions

Weather	Total	Fatal
Clear	81,911	452
Cloudy	17,476	97
Rain	9,157	22
Snow	461	1
Fog, Smog, Smoke	369	5
Sleet, Hail, Freezing Rain/Drizzle	238	0
Other	64	0
Severe Crosswinds	33	0
Blowing Sand, Dirt, Snow	7	0

Table 2 summarizes that 75 percent (82,450) of the total crashes and 55 percent (319) of the fatal crashes occurred during daylight conditions, followed by not-lighted dark conditions with about 15 percent (15,787) of the total and about 33 percent (188) of the fatal crashes. Lighted dark conditions had about 7 percent (7,534) of the total and about 9 percent (49) of the fatal crashes. Daylight generally gives better visibility to read the work zone warning and speed limit signs. Therefore, the high percentage of total (75 percent) and fatal (55 percent) crashes in daylight indicate the possibility of a lack of work zone regulation compliance by NC drivers. The higher percentage of fatal crashes (33 percent) compared to total crashes (15 percent) occurring in not lighted dark conditions indicates a need of improving lighting conditions and stricter law enforcement during nighttime on work zones. The difference between the percentages of fatal crashes and total crashes for not lighted dark conditions (total: 15 percent; fatal: 33 percent) is much higher compared to lighted dark conditions (total: 7 percent; fatal: 9 percent). This further indicates the possibility of good work zone lighting being able to reduce fatal crash occurrences.

Table 2: Work Zone Crashes by Lighting Conditions

Lighting Condition	Total	Fatal
Daylight	82,450	319
Dark-Not Lighted	15,787	188
Dark-Lighted	7,534	49
Dusk	2,105	7
Dawn	1,543	9
Dark-Unknown	192	1
Unknown	62	2
Other	43	2

Table 3: Work Zone Crashes by Surface Conditions

Surface Conditions	Total	Fatal
Dry	92,617	516
Wet	14,300	44
Water (Standing, Moving)	900	5
Ice	722	0
Sand, Mud, Dirt, Gravel	451	4
Snow	349	0
Slush	146	3
Unknown	139	2
Other	80	3
Fuel, Oil	12	0

Surface Conditions and Road Features. From Table 3, it can be observed that about 85 percent of total (92,617) and 90 percent of the fatal (516) crashes occurred on dry surface conditions followed by wet conditions at 13 percent of the total and about 8 percent of the fatal work zone crashes. Similarly, an analysis based on road features (Table 4) indicates that about 70 percent of the total (75,592) and 66 percent (381) of the fatal work zone crashes occurred on roads with no special features, i.e., no intersection, bridge approach, or merge lanes. Dry surface conditions and no special features are expected to have lower crashes as they do not require much effort from drivers to navigate. However, the highest percentages for these groups also indicate lack of work zone law compliance by NC drivers.

Time-Series Analysis. Figure 3 shows the monthly total work zone crashes in NC from 2000 to 2020. The graph shows multiple trends (e.g., increase from 2000 to 2004, decrease from 2005 to 2010) and a seasonal pattern. For further investigation, the time series was decomposed to investigate the trend and seasonal patterns separately (Figure 4).

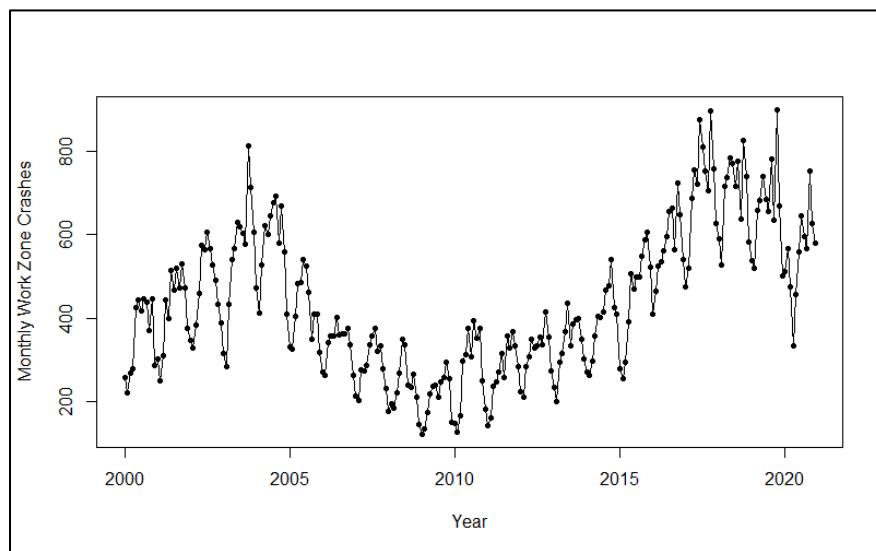
**Figure 3: Monthly Work Zone Crashes**

Table 4: Work Zone Crashes by Road Features

Road Feature	Total	Fatal
No Special Feature	75,592	381
Four-Way Intersection	8,561	51
T-Intersection	6,071	29
No Information	4,269	25
Bridge	2,390	17
Public Driveway	1,516	7
Related To Intersection	1,461	2
On Ramp Entry	1,394	10
Off Ramp Proper	1,204	3
Off Ramp Entry	1,192	8
Private Driveway	902	12
Bridge Approach	762	1
Other	673	4
On Ramp Proper	672	2
Underpass	570	5
Off Ramp Terminal On Crossroad	569	1
On Ramp Terminal On Crossroad	523	5
Merge Lane Between On And Off Ramp	360	1
Y-Intersection	321	3
End Or Beginning-Divided Highway	176	3
Non-Intersection Median Crossing	169	4
Traffic Circle/Roundabout	155	1
Railroad Crossing	97	2
Alley Intersection	56	0
Five-Point, Or More	43	0
Tunnel	11	0
Shared-Use Paths Or Trails	7	0

The trend plot in Figure 4 shows that from 2000 to 2004 the total crashes increased followed by a decrease from 2005 to 2009, which then increased from 2010 to 2017 followed by a decrease again from 2018 to 2020. The seasonal plot in Figure 4 shows that work zone crashes are the most frequent from June to October every year, with the highest occurrence consistently observed in October every year. To investigate whether the trend and seasonal pattern are useful for predicting the monthly work zone crash frequencies, a Holt-Winters time series prediction model was developed.

Holt-Winters is an exponential smoothing time-series prediction technique that utilizes the trend and seasonality of the data (Chatfield, 1978). In this paper, we compared two commonly used Holt-Winters modelling methods: multiplicative and additive. Data from 2000 to 2018 were used to train the models and data from 2019 to 2020 were used to validate the models. Three performance indexes including the mean absolute percentage error (MAPE), mean absolute error (MAE), and mean squared error (MSE) were chosen (Hyndman, 2006). Many researchers have

used these metrics to assess the accuracy of models. When the values of these indexes of a model are smaller, the fitting degree of the model is better. If the index values of a model are lower than another model, then the model is more superior (Xian et al. 2023). For the multiplicative model, we obtained a MAPE of 16.5 percent, MAE of 101.94, and MSE of 18,040.24. For the additive model, we obtained a MAPE of 38.42 percent, MAE of 284.13, and MSE of 114,691.30. Because the smaller performance index values of the multiplicative model, the multiplicative model was selected. Figure 5 shows the actual and multiplicative model predicted test data. The model predictions were much closer to the actual data points in 2019 (MAPE: 13.2 percent), compared to the predictions for 2020 (MAPE: 19.8 percent). The reduced model performance observed in 2020 might be attributed to the influence of the COVID-19 pandemic on traffic patterns, which began in the United States during the spring of that year. This discrepancy could stem from the fact that the model was trained using data spanning from 2000 to 2018, potentially failing to account for the unprecedented changes induced by the pandemic.

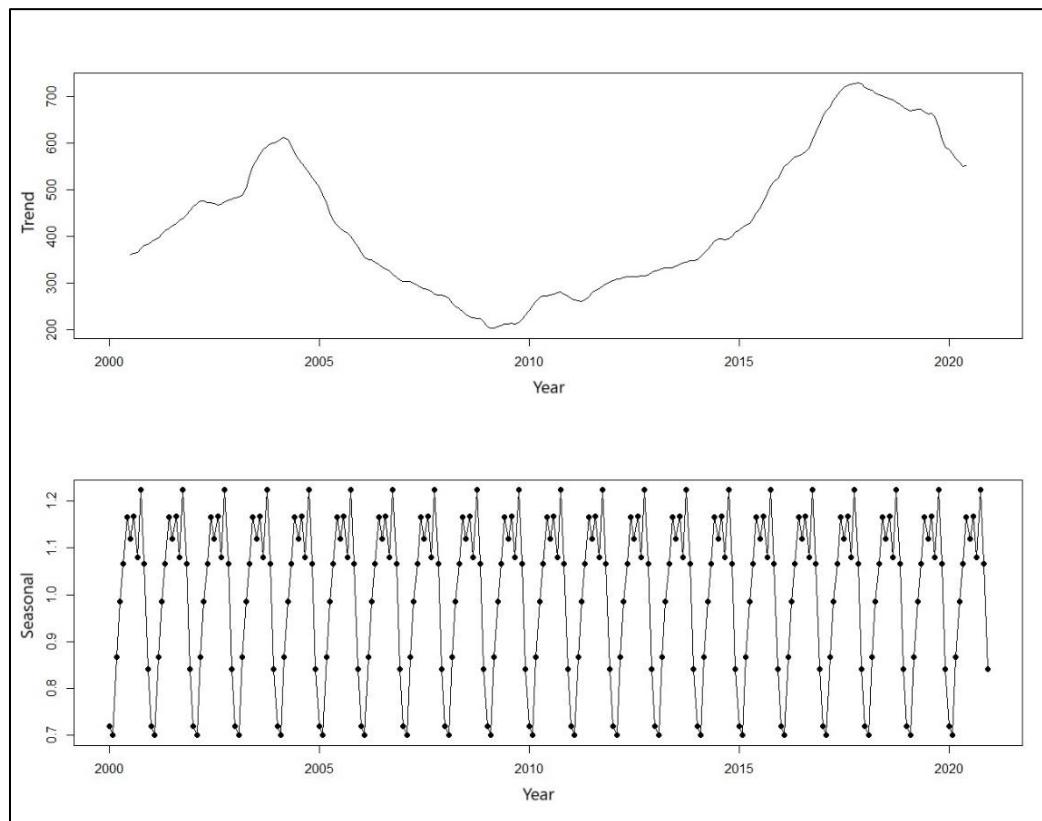


Figure 4: Trend and Seasonal Pattern of Monthly Work Zone Crashes

We also excluded the data in 2020 and ran a new model using the data from 2000 to 2017 for model training and the data from 2018 to 2019 for model validation. The new model had a MAPE of 22.2 percent, MAE of 112.74, and MSE of 18,699.37, which was inferior to the multiplicative model above. Therefore, we decided to proceed with the multiplicative model with all the data. Further analysis with supplemental data from other sources will be needed to verify whether the pandemic has any impact on the traffic patterns and the modelling results, which is beyond the scope of this paper.

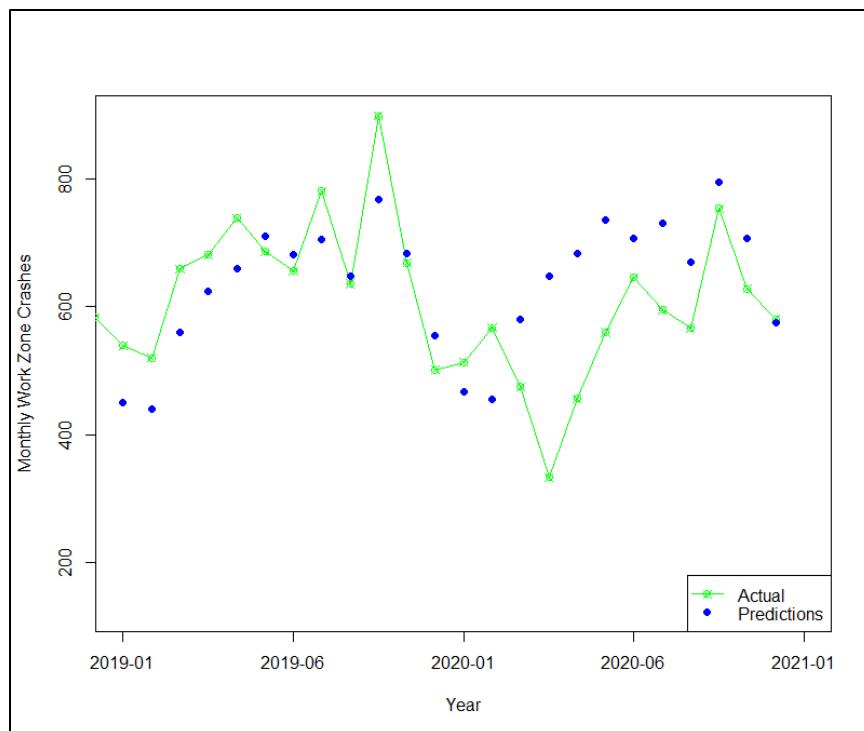


Figure 5: Actual and Holt-Winters Multiplicative Model Predicted Data

DISCUSSION

This study analyzed the trends of NC highway work zone crashes using descriptive analytics, exploratory data analysis, and time-series analysis.

Construction Related Work Zones. Construction related highway work zone crashes are the major concern for NC since they made up 84% of total crashes, while the other types combined contributed about 16% of total crashes. Similar trend was observed for fatal crashes as well, i.e., 81% occurred at construction related highway work zones, while 19% in the other three categories combined. This is consistent with findings from a previous study focusing on Georgia work zones (Daniel et al., 2000). This could be because construction related work zones require the highest amount of time to finish the work, exposing traffic to the work zone for a longer time, and therefore increasing the crash probability.

Work Zone Activity and Work Zone Location. Active work zones are generally expected to have a higher number of total and fatal crashes, because of the increasing uncertainties and driver distractions in the work environment, e.g., mobile equipment and construction workers. In this study, it was observed that 60% of total crashes and 50% of fatal crashes occurred at active highway work zones; and 50% of total crashes and 39% of fatal crashes occurred adjacent to work areas. These findings are consistent with the expectation that active work zones/areas tend to lead to more crashes. However, the finding that more than half of the fatal crashes occurred at non-active work zones also indicate that the causes of work zone crashes are complex and factors beyond work zone themselves, such as driver behaviors, could also contribute to crash occurrences. A further analysis of the fatal crashes using other external factors would help to identify key risk factors for work zone crashes.

It is also worth mentioning that 32% of total crashes and 34% of fatal crashes occurred at work areas approaching tapers. Tapers are the locations in work zones where a driver encounters an unexpected change and therefore must adjust their travel path. This finding highlights the importance of proper design and placement of tapers for highway work zones. Proper installation of pre-work zone warning signs would also help to reduce taper crashes.

Weather and Light Conditions. The highest percentages of total and fatal work zone crashes happened during clear weather conditions and daylight conditions. These observations could mean that a higher number of NC drivers does not comply with work zone speed limits and hence result in the higher number of total and fatal crashes during clear weather and daytime conditions. However, it should be noticed that results are aggregated results at the state level and hence different results could be found if a highway specific analysis was conducted. For example, 326 out of 371 crashes on Michigan M-94\I-94\I-94BL\I-94BR highway from 2004 to 2008 occurred in fog/snow/severe wind weather (Weng et al., 2016).

Three quarters of total crashes occurred in daylight conditions, followed by 14% in not lighted dark conditions. For fatal crashes, 55% occurred in daylight conditions, and 33% occurred in not lighted dark conditions. Zhang and Hassan (2019) found that significant differences exist between factors contributing to injury severity for daytime and nighttime highway work zone crashes in Egypt, based on data from Egyptian long-term highway work zone projects between 2010 and 2016. They reported that older male drivers, the number of lane closures, sidewise crashes, and rainy weather have opposite effects on injury severity in nighttime and daytime crashes. A similar investigation could be carried out to analyze the NC highway work zone crashes to investigate whether there are different predictors for daytime and nighttime crashes to predict crash severity.

Trend and Seasonal Patterns. The time series analysis revealed that NC work zone crashes have a trend and a seasonal pattern. The trend showed that the annual work zone crashes increased from 2000 to 2004 followed by a significant decrease from 2005 to 2009. In September 2004, the Federal Highway Administration (FHWA) updated the work zone regulations at “23 CFR 630 Subpart J.” (Code of Federal Regulations, 2004). Further in 2005 the US Department of Transportation (USDOT) and Cambridge Systematics published guidance documents to help agencies adopt those new rules. The decrease in annual work zone crashes could be due to the adoption of these rules by agencies in NC. The continuation of this downward trend could be because of the 2007-08 recession. Vehicle miles traveled decreased during that economic recession (Statista, 2021) and therefore, fewer vehicles were exposed to work zones, and hence the decrease in work zone crashes. The trend then moved upward from 2010 to 2018, which could be due to the increase in vehicle miles traveled both nationally (Statista, 2021) and in NC (Bureau of Transportation Statistics, 2020). An increase in vehicle miles traveled is generally associated with an increase in traffic crashes. After 2018, the trend was again observed to take a sharp downturn which could also be due to the decreasing rate of vehicle miles traveled (2017-2016: 2.1 percent; 2018-2017: 1.6 percent; 2019-2018: 1.1 percent) in the state during this period (FHWA, 2020).

The work zone crashes were also observed to have a seasonal pattern. The months of June to October had the highest work zone crashes, with the peak occurring in October every year. NC attracts the highest number of visitors during the summer months (June to August) and during peak fall color season (mid-September to October) (Economic Development Partnership of North Carolina, 2021). These increased visitors lead to an increase in traffic volume which increases traffic exposure to the work zones, and thus increases the possibility of crash

occurrences. The results from the Holt-Winters time series prediction model indicate that the trend and seasonal patterns could be used for predicting future work zone crash frequencies. This also shows that the general crash trend and crash month could also be a potential predictor for location-specific crash prediction models along with traffic and geometric variables.

CONCLUSION

An analysis of the NC highway work zone crashes from 2000 to 2020 was conducted in this study. It was found that NC highway work zone crash occurrences are significantly associated with the following conditions: active work zones, adjacent to work areas and work areas approaching tapers, clear weather, dry road surface, and daylight conditions. Further, it was found that NC work zone crashes had a trend and a seasonal pattern, which could be leveraged to predict future crash frequencies. The findings of this study serve as vital insights for NCDOT in the development of efficient policies and countermeasures to improve safety in highway work zones. One potential extension of this work could be the development of micro-level highway-specific work zone crash prediction models using temporal, traffic, and geometric variables.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to Dr. Joseph E Hummer and Timothy S. Nye from the North Carolina Department of Transportation for their kind support with the data collection for this research.

REFERENCES

- BTS (Bureau of Transportation Statistics). (2020). North Carolina Transportation by Numbers.https://www.bts.dot.gov/sites/bts.dot.gov/files/states2020/North_Carolina.pdf.
- Chatfield, C. (1978). "The Holt-Winters Forecasting Procedure". *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 27(3), 264-279.
- Code of Federal Regulations. (2004). Subpart J- Work Zone Safety and Mobility. <https://www.ecfr.gov/current/title-23/chapter-I/subchapter-G/part-630/subpart-J>.
- Daniel, J., Dixon, K., and Jared, D. (2000). "Analysis of Fatal Crashes in Georgia Work Zones". *Transportation Research Record: Journal of the Transportation Research Board*, 1715(1), 18-23.
- Deretić, N., Stanimirović, D., Al Awadh, M., Vujanović, N., and Djukić, A. (2022). SARIMA Modelling Approach for Forecasting of Traffic Accidents. *Sustainability (Switzerland)*, 14(8).
- Economic Development Partnership of North Carolina. (2021). *A Publication of Visit North Carolina*.
- FHWA (Federal Highway Administration). (2020). FHWA Work Zone Facts and Statistics. https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm.
- FHWA (Federal Highway Administration). (2020). Highway Statistics 2016. <https://www.fhwa.dot.gov/policyinformation/statistics/2016/vm2.cfm>.
- Hyndman, R. J. (2006). "Another look at Forecast Accuracy Metrics for intermit-tent demand". *Foresight: The International Journal of Applied Forecasting*.

- Khan, M. A., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. (2022). Do ridesharing transportation services alleviate traffic crashes? A time series analysis. *Traffic Injury Prevention*, 23(6), 333–338.
- Khattak, A. J., and Targa, F. (2004). “Injury Severity and Total Harm in Truck-Involved Work Zone Crashes”. *Transportation Research Record: Journal of the Transportation Research Board*, 1877(1), 106-116.
- Nugus, S. (2009). “Time Series Analysis” Chapter 6 of *Financial Planning Using Excel*, Second Edition: 77-84.
- Osman, M., Paleti, R., Mishra, S., and Golias, M. M. (2016). Analysis of injury severity of large truck crashes in work zones. *Accident Analysis and Prevention*, 97, 261–273.
- Statista. (2021). Number of vehicle-miles traveled on all roads in the United States from 1993 to 2020. <https://www.statista.com/statistics/185579/us-vehicle-miles-in-transit-since-1960/>.
- Theofilatos, A., Yannis, G., Antoniou, C., Chaziris, A., and Sermpis, D. (2018). Time series and support vector machines to predict powered-two-wheeler accident risk and accident type propensity: A combined approach. *Journal of Transportation Safety and Security*, 10(5), 471–490.
- Weng, J., Zhu, J. Z., Yan, X., and Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis and Prevention*, 92, 43–52.
- Xian, X., Wang, L., Wu, X., Tang, X., Zhai, X., Yu, R., Qu, L., and Ye, M. (2023). “Comparison of SARIMA model, Holt-winters model and ETS model in predicting the incidence of foodborne disease”. *BMC Infectious Diseases*, 23, 803.
- Yu, M., Ma, C., Zheng, C., Chen, Z., and Yang, T. (2022). “Injury severity of truck-involved crashes in work zones on rural and urban highways: Accounting for unobserved heterogeneity”. *Journal of Transportation Safety and Security*, 14(1), 83–110.
- Zhang, K., and Hassan, M. (2019). “Crash severity analysis of nighttime and daytime highway work zone crashes”. *PLoS ONE*, 14(8).

Project-Based Crash Analysis for Crash Risk Reduction during Pavement Preservation

Sirisha Gangireddy¹; Ruijie Bian, Ph.D., P.E., M.ASCE²; Tara Tolford, AICP³;
and Hany Hassan, Ph.D., P.E., M.ASCE⁴

¹Dept. of Civil and Environmental Engineering, Louisiana State Univ., Baton Rouge, LA.
Email: sgangi1@lsu.edu

²Louisiana Transportation Research Center, Baton Rouge, LA. Email: rbian1@lsu.edu

³UNO Transportation Institute, Univ. of New Orleans, New Orleans, LA.
Email: tmtolfor@uno.edu

⁴Dept. of Civil and Environmental Engineering, Louisiana State Univ., Baton Rouge, LA.
Email: hassan1@lsu.edu

ABSTRACT

Low-cost safety improvements, road diets, and project upgrades are three typical solutions to integrate the concept of Complete Streets into pavement preservation programs of state departments of transportation (DOTs). Speed management, as a low-cost safety improvement method, could be implemented during pavement preservation to reduce crash risk due to speeding issues. The current study collected a list of preservation projects funded by Louisiana State DOT from 2018 to 2020 and conducted a project-based crash analysis to examine the change of crash rate before and after pavement preservation to evaluate the needs and feasibility of applying speed management. The before-and-after comparison and the regression analysis results revealed a significant increase of crashes after pavement preservation works started, particularly on rural collectors. Radar speed feedback signs (RSFS) were identified as a potential countermeasure to reduce crash risk during pavement preservation on rural collectors. The subsequent benefit-cost analysis results show that implementing speed management within two years after pavement preservation work starts would yield large positive returns on investment as evidenced by a benefit-cost ratio (BCR) over three. The study findings, along with the speed management technology reviewed, are expected to be useful for state DOTs in reducing crash risk related to speeding, improving work zone safety, and better integrating the concept of Complete Streets into their pavement preservation programs.

INTRODUCTION

As the Federal Highway Administration (FHWA) adopts a Complete Streets design model, (FHWA 2022a), transportation agencies are actively looking for approaches to better integrate the concept of Complete Streets into their institutional processes (Bian and Tolford 2022; Jordan and Ivey 2021). Complete Streets emphasize providing safe accommodation to all road users (irrespective of their age and abilities) (SGA 2023), but sometimes are narrowly interpreted as providing facilities like sidewalks and bicycle lanes. Such limited-scope interpretations mislead some stakeholders to consider Complete Streets are not applicable to certain project types (like pavement preservation) or project delivery phases (like construction). It is important to note that Complete Streets do not imply a one-size-fits-all approach for every street. In certain contexts, implementing Complete Streets can be both cost-effective and achievable without the need for expanding existing infrastructure, making this concept applicable for all project types and phases

(SGA 2016). Safety enhancements which focus on speed management can improve outcomes for all road users—and particularly vulnerable road users like pedestrians and bicyclists—even where dedicated facilities for active modes are lacking.

The motivation to perform this study emerged from the above-mentioned background. The author's previous practice review study has identified three general solutions of integrating Complete Streets opportunities into pavement preservation programs: low-cost safety improvements, road diets, and project upgrades (Gangireddy et al. 2023). Meanwhile, a study conducted to evaluate project outcomes raised potential concerns regarding speeding issues and increasing crash risk that need to be addressed during project construction (Bian et al. 2023). Thus, the current study investigated the needs and feasibility of applying speed management (as a low-cost safety improvement method) in pavement preservation as a step moving towards institutional integration of Complete Streets. Specifically, the current study: 1) conducted in-depth statistical analysis to determine at which project stage/location pavement preservation works are more likely to be associated with higher number of crashes and 2) performed benefit-cost analysis to identify appropriate speed management techniques that maximize safety benefits without inflating project costs.

The paper starts by discussing the importance of speed management and work zone safety, conducts a project-based crash analysis comparing crashes before and once preservation works started, and statistically identifies which preservation work types are more likely to be associated with increased crashes. The paper then focuses on reviewing Intelligent Transportation System (ITS) technologies to select an appropriate one for speed management in this case. Subsequently, Benefit-Cost Analysis (BCA) is conducted for a selected set of projects to quantify safety benefits in monetary while considering associated costs of speed management. Finally, the paper concludes by summarizing key findings, providing practical recommendations, and acknowledging study limitations.

LITERATURE REVIEW

Need for Speed Management

In 2020, the National Highway Traffic Safety Administration (NHTSA) reported that speeding was a contributing factor in about 28% of fatal crashes, 13% of injury crashes, and 10% of property-damage-only crashes (National Center for Statistics and Analysis 2022). Prioritizing safety, FHWA has launched the National Road Safety System (NRSS), incorporating the Safe System Approach (SSA) as a key component (FHWA 2021a). The SSA is designed to develop and manage road infrastructure in a way that minimizes crash risk, focusing on five key elements: safe vehicles, safe speeds, post-crash rate, safe roads, and safe road users (FHWA 2021a). Among these elements, safe speed is considered as a fundamental aspect contributing to reduced crash severity and safer streets, leading to the introduction of speed management as one of the FHWA's proven safety countermeasures in 2021 (FHWA 2021b).

FHWA's speed management countermeasures include three strategies (FHWA 2021b). First, setting appropriate speed limits is crucial to reduce the risks for both drivers and vulnerable road users. (FHWA 2012). Setting a speed limit involves considering factors like land use, crash rates, pedestrian and cyclists' activity, road geometry, and road conditions. Second, speed safety cameras (SSCs) are considered a reliable and effective technology to enforce speed limits in order to promote safer driving behaviors (FHWA 2021c). Agencies need to consider USDOT

SSCs guidelines for planning, public involvement, stakeholder coordination, implementation, maintenance, and evaluation (FHWA 2021c). Third, variable speed limits (VSLs) use real-time information on the roadway (like traffic volume, traffic speed, weather and roadway conditions) to dynamically adjust speed limits displayed to drivers (FHWA 2014). VSLs are often integrated into Active Transportation Management (ATM) or Road Weather Information System to address congestion, incidents, work zones, and inclement weather (FHWA 2014).

Work Zone Safety

According to the NHTSA, speeding was involved in 37% of the fatal crashes that occurred in work zones, and 28% for crashes occurred outside of work zones (National Center for Statistics and Analysis 2022). The increased concern about speeding within work zones is related to the risk posed by construction equipment, changes in roadway design and markings, and increased pedestrian activity (National Center for Statistics and Analysis 2022).

Maintenance/preservation activities are usually smaller in scale compared to construction projects because of their short-term nature and lower intensity (FHWA 2022b). However, maintenance/preservation activities still cause work zone impacts like merging and lane changing conflicts, distracted driving, reduced lane width, driver non-compliance with reduced speed limits, and even large impacts, depending on the activity type and location (e.g., daytime maintenance work on a busy urban corridor can result in substantial traffic slowdowns) (FHWA 2022b). It is crucial to address work zone impacts of maintenance/preservation activities from an overall system management perspective, considering both the overall system and specific activity or corridor involved. Assessing and managing maintenance work zone impacts involves the following key activities (FHWA 2022b): 1) improving agency procedures to minimize direct safety and mobility impacts from maintenance activities, such as improving Temporary Traffic Control (TTC) procedures, providing better public information through advance warning systems, and establishing permitted lane-closure times; 2) planning and coordinating maintenance activities to reduce system-wide impacts, such as integrating maintenance activities into planned construction projects to reduce separate disruptions; 3) implementing and managing maintenance activities to minimize impacts on other construction projects and vice-versa, such as considering detour routes and diversions; and 4) incorporating features like wider shoulders, maintenance turnabouts, and designated pullouts into construction projects to facilitate minimum disruption in future maintenance projects.

PROJECT-BASED CRASH ANALYSIS

Data Collection and Description

The initial phase of this study involved conducting a project-based crash analysis to identify any possible correlations between the specific characteristics of pavement preservation projects and higher observed numbers of crashes. A comprehensive list of pavement preservation projects funded by the Louisiana Department of Transportation and Development (DOTD) within the time frame of 2018 to 2020 was compiled. This data collection yielded 145 pavement preservation projects, with 103 in rural areas and 42 in urban areas. Basic project details were collected from DOTD project management systems including: project ID, project time (letting date), project category, project location (including parish, route, control section, and begin/end

log mile), and project cost. More detailed project-level information was also gathered from various DOTD systems, including the Preservation, Rehabilitation and Replacement (PRR) report, Baseline Safety Checklists, and project plan sheets (DOTD 2023). The following information was collected: preservation work types, specific project location (road segments), project start date, rural or urban categorization, changes in shoulder width, and a list of 28 safety countermeasures (which do not include speed management) considered by each project.

Mapping all projects in ArcGIS (as shown in Figure 1) facilitates spatial join of crash data to each project. The following spatial characteristics were collected in this process: posted speed limits (LADOTD 2020a), number of lanes (LADOTD 2020b), and roadway functional classification (LADOTD 2021). The enriched dataset allows the authors to better understand the relationship between preservation activities, crash data, and built environment.

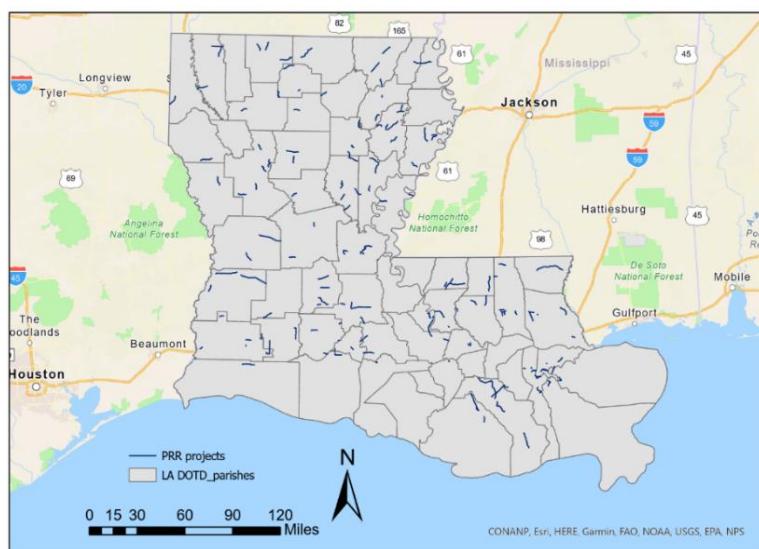


Figure 1. A map of pavement preservation projects in Louisiana (2018-2020)

Before-and-After Analysis

First, the annual average number of crashes for each project (within a distance threshold of 50 feet) was calculated based on each project's status (Before and During/After pavement preservation). Then the annual average number of crashes were averaged across all the projects (Before and During/After pavement preservation) and presented in Figure 2(a). The results show that the average number of crashes increased in general once preservation work started. Second, the average number of crashes was determined for each work type and presented in Figure 2(b). It is observed that the average number of crashes decreased in the case of major rehabilitation work but increased in the other three pavement preservation work types. This finding indicates that it is likely that the average number of crashes changed in a different manner during different types of preservation works, which requires further analysis and consideration of additional factors to identify target projects that need speed management for crash risk reduction.

Second, paired t-tests were conducted to statistically compare the average number of crashes before and during/after pavement preservation. This is to statistically determine whether there is a significant increase or decrease in the number of crashes after pavement preservation works

started (irrespective of each work type). The first paired t-test examines whether the average number of crashes are equal. The p-value is 0.04, which rejects the null hypothesis and means that the average number of crashes before and during/after pavement preservation are not equal. The second paired t-test evaluates whether the average number of crashes increased after pavement preservation started. The p-value is 0.02, which rejects the null hypothesis and means that the average number of crashes increased after pavement preservation started. This preliminary finding highlighted the need to apply speed management for crash risk reduction during pavement preservation.

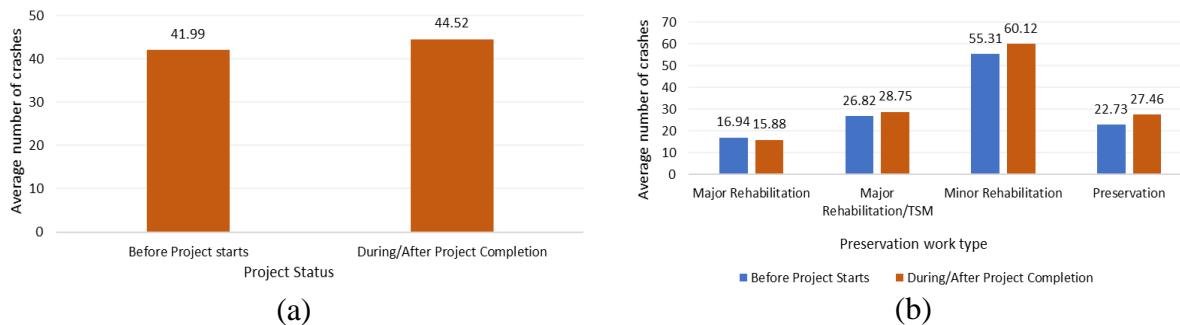


Figure 2. Change in the average number of crashes before and during/after pavement preservation: (a) Average number of crashes with respect to project status; (b) Comparison of average number of crashes for different preservation work types

Regression Analysis

This section examines the significance of associations between the number of crashes and various pavement preservation works via statistical modeling. In this analysis, the dependent variable was the change in the number of crashes before and during/after pavement preservation ($1 - \frac{\text{Crash}_{\text{after}}}{\text{Crash}_{\text{before}}}$). Independent variables included the collected project information and related built environment characteristics as described in the “Data Collection and Description” section. The regression analysis was done in two steps: First, a simple linear regression was performed to assess the significance of each project-related factors on observed crash changes. Second, a multiple linear regression model was estimated with forward variable selection technique. The final model comprises variables that collectively had significant associations with crash changes.

Table 1 presents the multiple linear regression results. The results indicated that pavement preservation work on rural major/minor collectors observed a significant increase of crashes ($p\text{-value} < 0.01$). Additionally, it was found that pavement preservation projects involving narrow bridges but lacking proper delineation with signs, object markers, delineators, and/or pavement markings as per PM-01 and MUTCD (i.e., Safety10) and/or re-evaluating curve advisory speed according to DOTD policy (i.e., Safety17) also experienced a significant increase in crashes. This suggests that the absence of these two safety countermeasures on rural major/minor collectors is likely to contribute to a notable rise of crashes during preservation. In other words, it can be concluded that longitudinal sections on rural collectors without narrow bridges in their adjacent area are more likely to observe higher number of crashes during pavement preservation.

Table 1. Multiple linear regression results

Variables	Multiple regression	
	Parameter	P-value
Major collector (rural)	0.5995	0.003
Minor collector (rural)	0.5153	0.003
Safety10 (Delineate narrow bridges with signs, object markers, delineate and/or pavement markings per PM-01 and MUTCD)	-0.5283	0.051
Safety17 (Re-evaluate curve advisory speed according to DOTD policy)	-0.3518	0.023

(Note: the independent variables are binary variables, where one equals Yes and zero equals No.)

SAFETY COUNTERMEASURE APPLICATION

State DOTs have been using Intelligent Transportation System (ITS) technologies in implementing Temporary Traffic Control to improve traffic flow and enhance work zone safety (FHWA 2008). A comprehensive before-and-after analysis study, conducted by FHWA across North Carolina, Arkansas, Michigan, Texas, and the District of Columbia, demonstrated the effectiveness of ITS technologies in work zones (FHWA 2008). The study revealed significant benefits, including improved driver reaction to traffic changes (reported by 82% surveyed drivers) reduced congestion through real-time information and alternate routes, decreased aggressive driving behaviors, and an overall enhancement of safety for both motorists and workers (FHWA 2008). Furthermore, the driver perception surveys conducted in Arkansas indicated that the presence of ITS applications in work zones positively influenced the perceived level of safety among drivers (FHWA 2008). All these findings suggest that ITS interventions not only contribute to actual safety improvements but also enhance drivers' confidence and satisfaction when navigating through work zones.

Table 2 is a summary of typical ITS technologies for work zone safety improvement. The summary primarily focused on their safety benefits, crash modification factor (CMFs), costs associated with their deployment, and installation/placement guidance provided in the Manual on Uniform Traffic Control Devices (MUTCD). This information enriches our understanding of factors influencing the selection of ITS technologies for speed management during/after pavement preservation and provides parameter values supporting the benefit-cost analysis conducted in this study. In addition, it is worth noting that using some of the technologies (e.g., Speed Safety Cameras) is often more of a political issue in practice that requires an integrated agency approach for successful deployment. Some non-ITS approaches, such as installing rumble strips (which is suitable for rural two-lane highways), are also available for speed management.

BENEFIT-COST ANALYSIS

Performing benefit-cost analysis (BCA) allows us to quantify the effectiveness of the implemented countermeasure to reduce crash risk during pavement preservation in monetarized values. Contrary to common belief, a review study assessing the use of BCA by state DOTs revealed that BCA application is not widespread in practice, especially in the context of preservation projects (FHWA 2015). The report highlighted that BCA is primarily used for

specific projects or when federal funding mandates it (FHWA 2015). Quality issues, such as improper baselines, speculative benefits, and a lack of transparency were highlighted (FHWA 2015). Furthermore, the complexities of BCA and a shortage of staff capacity were cited as challenges for some agencies. The report also revealed that many states opt for a multi-criteria scoring approach over BCA for simplicity in project prioritization (FHWA 2015). However, it is worth noting that states have successfully employed BCA. For instance, Iowa DOT applied BCA in a revitalization project, revealing positive returns on investment across safety, livability, economy, and environmental metrics, including crash reduction, increased property value, business/tourism benefits, reduced speeding incidents, and decreased average daily traffic for environmental gains (Iowa DOT 2016).

The above-mentioned challenges of conducting BCA in practice motivated the authors to review how transportation agencies evaluated their adoption of speed management techniques. Iowa DOT conducted a study on a speed feedback sign loan program for local agencies, focusing on the program's necessity, selecting suitable sign types, and assessing the impact on vehicle speeds pre- and post-implementation (Knapp et al. 2021). Their insights on aligning sign letter size with MUTC guidelines (especially for roads with speed limit of 45 mph or higher) informed our selection of appropriate sign types in the current study (Knapp et al. 2021).

Based on the statistical analysis results, pavement preservation projects on rural collectors without narrow bridges delineated with signs, object markers, and pavement markings as per PM-01 and MUTCD (i.e., Safety10) and without re-evaluating curve advisory speed in accordance with DOTD policy (i.e., Safety17) were selected for BCA calculations. The purpose of conducting BCA is to evaluate the net safety benefits from applying speed management techniques. The ideal technique application time was also discussed through BCA in the following subsections.

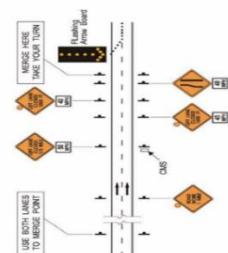
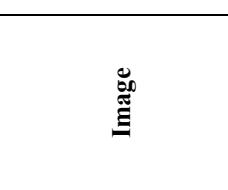
Parameter Setup: Benefits

This study focused on evaluating the safety benefits achieved through reducing fatal/injury crashes by using FHWA's BCA guidance (USDOT 2018). The following three parameters are crucial in determining safety benefits. First, the baseline risk represents the number of crashes during the years of BCA analysis. Those values are from the collected crash data.

Second, the risk reduction factor (RRF) is calculated as one minus crash modification factor (i.e., $RRF = 1 - CMF$). This study employed dynamic CMF values reflecting the decreasing effectiveness of radar speed feedback signs from year to year. The dynamic CMF values were determined based on the relationship between average traffic speed and crash risk (Forbes et al. 2012). The average traffic speed values over a two-year period were provided in a nationwide case study evaluating the effectiveness of dynamic speed feedback signs (Fisher et al. 2021). In this regard, the approximate CMF values for RSFS were 0.95 in the first year and 0.98 in the second year. It is considered ineffective to implement RSFS for over two years as the estimated CMF value is over one.

Lastly, the expected consequences of reducing injuries and fatalities were converted into monetized values using FHWA's BCA guidance, with the values adjusted to approximate 2023 dollars (USDOT 2018). The number of injuries and fatalities were also obtained from the collected crash data. Monetized values were based on the KABC level (USDOT 2018), and the inflation adjustment value from the Bureau of Economic Analysis was applied for the adjustment (BEA 2023).

Table 2. ITS technologies for work zone safety improvement

Parameter	Portable Changeable Message Signs (PCMSS)	Dynamic Lane Merge (DLM) system	Variable Speed Limit (VSL) signs	ITS Technology	Radar Speed Feedback Signs (RSFS)	Automated Speed Enforcement
Image						
Application Scenario	Facilitates standard signing for speed reduction, lane closures, ramp closures, etc. (FHWA 2020a)	Improves safety during lane closures and reduces aggressive driving in work zones (Michigan DOT 2004)	Effective in lowering average speeds in urban uncongested work zones (McMurtry et al. 2009)	Effective in areas with higher speed (Iowa DOT 2019)	Capture speed violation and effective in improving compliance (FHWA 2020b)	
Crash Modification Factor (CMF)	0.84 (CMF clearinghouse 2009)	0.54 (CMF Clearinghouse 2009a)	0.92 (CMF Clearinghouse 2012)	0.95 (Fisher et al. 2021)	0.85 (CMF Clearinghouse 2009b)	
Project Costs	\$12,000 - \$20,000 (American Traffic Safety Services Association (ATSSA) 2013)	\$24,000 (Michigan DOT 2004)	\$9,000 - \$18,000 (Edara et al. 2017)	\$2,500 - \$15,000 (Iowa DOT 2017)	\$200,000 - \$450,000 (Plan and Year 2015)	
MUTCD Installation Guidance	Mounted on a trailer or truck bed in work zone for temporary requirements (USDOT 2009)	Local traffic signs (e.g., ROAD CLOSED TO THRU TRAFFIC) are used when road users need to detour to avoid a closure (FHWA 2009)	Recommended to use when there is a speed difference of 10-25 mph between posted and advisory speed limits (Sunriseesesa 2023)	Letter size of 6" and LED digits of 18" are recommended for single or multi-lanes with posted speed limit of 45 mph or higher (USDOT 2009)	Not regulated by MUTCD (they are not considered as traffic control devices) (FHWA 2009)	

Parameter Setup: Costs

Based on our literature review, RSFS were selected for their cost-effectiveness and higher safety improvement efficacy. All identified projects from the statistical analysis had posted speed limits of 45 mph or higher. Therefore, according to the MUTCD, the optimum distance for installing radar speed feedback signs is 1,000 feet upstream of the work zone.

Project costs, including capital, operations, and maintenance were considered. The capital cost for purchasing RSFS was determined based on Iowa DOT's study, aligning with MUTCD's radar speed feedback sign guidelines and endorsing the suitability of the Traffic Logix safepace475 sign for the current case study (Knapp et al. 2021). The capital cost of this sign is approximately \$5,385 in 2023 dollars, sourced from a specialized traffic safety solutions firm ("SafePace 475 18" Radar Feedback Sign - Radar Feedback Signs | TAPCO" n.d.). For installation costs, the study examined Bellevue's stationary radar sign program, focusing on cases closely resembling our study scenarios (City of Bellevue Transportation Department 2009). The average installation cost was estimated to be about \$17,200 per sign (converted in 2023 dollars). While no specific studies addressed RSFS maintenance costs, the adoption of RSFS generally showed no issues within the first two to three years after installation (Hallmark et al. 2015), leading to an assumption of zero maintenance costs in this study.

BCA results

Table 3 shows the eight projects selected for BCA study. Two scenarios were developed: 1) applying speed management in the first year after preservation work starts and 2) applying speed management in the first two years after preservation work starts. Note that the time durations are longer than a typical preservation project and cover post-project time without work zones (when speeding is also likely to occur due to improved pavement quality). In this study, the determination of the benefit-cost ratio (BCR) primarily depends on the amount of safety benefits obtained since the cost is fixed. The BCR values for all projects are greater than one indicating a positive return on investment. The BCR values are exceptionally high for projects with a higher number of injuries and fatalities. The reason for such high BCR values is that safety benefits primarily rely on expected crash reduction outcomes. The outstanding BCR values suggest that using RSFS can generate substantial net safety benefits. Based on Table 3, two-year application (i.e., Scenario 2) seems to be more ideal since it has higher BCR values than Scenario 1. However, this result is due to our assumption of negligible maintenance cost so the costs associated with one-year and two-year deployment remain fixed but safety benefits are accumulated each year. In addition, the crash reduction effectiveness of RSFS also decreases from year to year. Therefore, it would be better to consider one-year application in practice at first; then an additional year of application will be considered based on crash rate to-be-observed in the first year.

CONCLUSIONS AND RECOMMENDATIONS

The current study provides a comprehensive examination of speed management application in pavement preservation projects and the use of ITS technologies to reduce speeding. Radar speed feedback signs were chosen as a more suitable speed management technology for preservation projects considering their crash reduction performance and relative low cost.

Subsequently, a project-based analysis of crashes and a benefit-cost analysis were conducted to determine the ideal time for implementing the selected speed management technique and quantify the safety benefits gained from its application.

Table 3. Applying Speed Management in Pavement Preservation

Project ID	Work type	Safety benefits of Scenario 1 (in 2023\$)	Safety benefits of Scenario 2 (in 2023\$)	Total cost (in 2023\$)	BCR of Scenario 1	BCR of Scenario 2
01	Preservation	94,046	94,046	22,585	4.16	4.16
02	Major Rehabilitation	70,535	70,535	22,585	3.12	3.12
03	Major Rehabilitation	117,558	117,558	22,585	5.21	5.21
04	Major Rehabilitation/TSM	105,802	246,871	22,585	4.68	10.95
05	Major Rehabilitation/TSM	5,713,295	7,029,939	22,585	252.97	311.26
06	Major Rehabilitation/TSM	28,042,393	31,639,653	22,585	1241.64	1400.91
07	Major Rehabilitation	1,481,225	2,666,205	22,585	65.58	118.05
08	Major Rehabilitation/TSM	329,161	413,802	22,585	14.57	18.32

The before-and-after study and regression analysis results revealed an increase in crashes following pavement preservation, particularly on rural collectors. A list of selected projects was used to conduct the benefit-cost analysis. It is suggested to consider one-year application in practice at first; then an additional year of application will be considered based on crash rate to-be-observed in the first year. The suggestion is based on the satisfying BCR values of one-year application (i.e., >3), decreasing crash reduction effectiveness of RSFS from year to year, and maintenance costs that are likely to occur in practice.

Based on the findings, the following are a few recommendations for state DOTs' considerations. First, speed management techniques should be prioritized in preservation projects on rural collectors to address safety concerns. Second, conducting project-based crash analysis and assessing economic feasibility of safety interventions will be helpful in improving safety during and after the construction period, which is an approach for advancing integration of the concept of Complete Streets into all project types and phases. As states move toward a Complete Streets and/or Safe Systems approach, a more holistic approach on how to improve safety for all road users is needed. For most states, preservation and/or maintenance programs constitute the bulk of roadway lane-mile improvements completed annually, and thus it is an important opportunity to affect system-wide impact. Speed is one of the most important determinants of crash outcomes: higher speed state routes in both rural and urban areas are frequently among the deadliest for people walking and bicycling. Additionally, the current study's benefit-cost analysis results, which showed a significant amount of safety benefits after applying low-cost speed management techniques, support the idea that low-cost safety interventions aimed at managing speed will have a positive impact on outcomes for vulnerable road users.

The current study has several limitations that should be noted. First, the scope of the study is limited to a dataset of 145 preservation projects in Louisiana, in which we did not observe a high number of crashes involving bicyclists/pedestrians. States with higher bicyclist/pedestrian volumes need to consider facilitating walking/biking needs during pavement preservation (e.g., setting "Sidewalk Closed Ahead" signs and providing temporary facilities), which will help

better integrate the concept of Complete Streets into the institutional process. Second, this study focused on evaluating safety benefits without considering other potential impacts (e.g., mobility and accessibility). Third, this study assumes the maintenance costs for radar speed feedback signs to be zero, which might not be realistic in places with recurring extreme weather events or frequent vandalism. States with larger geographic extents may also face challenges in maintaining the devices due to substantial logistical costs (e.g., time, labor, and equipment). Fourth, this study assumes constant Annual Average Daily Traffic (AADT) from year to year for each studied project due to data limitations. While this approach helps place all the projects on a similar scale for statistical analysis, it may overlook potential increase in traffic volumes after project completion (due to improved pavement surface). Despite these limitations, the current study provides valuable insights into potential safety benefits of implementing speed management during pavement preservation. The study also serves as a response to FHWA's call for "supporting rigorous safety assessment during project development and design to help prioritize safety outcomes across all project types" (FHWA 2022c).

REFERENCES

- ATSSA (American Traffic Safety Services Association). 2013. *Guidance for the Use of Portable Changeable Message Signs in Work Zones*.
- BEA. 2023. "Inflation adjustment| BEA." Accessed February 16, 2023.
<https://apps.bea.gov/iTable/?reqid=19&step=3&isuri=1&1921=survey&1903=13>.
- Bian, R., T. Tolford, S. Liu, and S. Gangireddy. 2023. "Lessons learned from evaluating complete streets project outcomes with emerging data sources." *Transp. Plan. Technol.*, (May).
- Bian, R., and T. M. Tolford. 2022. "Evaluating the Implementation of the Complete Streets Policy in Louisiana: A Review of Practices and Projects in the Last 10 Years." *Transp. Res. Rec. J. Transp. Res. Board*.
- City of Bellevue Transportation Department, W. 2009. *Stationary Radar Sign Program*.
- CMF clearinghouse. 2009. "Install Changeable Queue Ahead Warning Signs - CMF."
<https://www.cmfclearinghouse.org/cmfpdf.php>.
- CMF Clearinghouse. 2009a. "Dynamic Lane Merge System." Accessed July 5, 2023.
<https://www.cmfclearinghouse.org/detail.php?facid=2000>.
- CMF Clearinghouse. 2009b. *Implement mobile automated speed enforcement*.
- CMF Clearinghouse. 2012. *CMF/ Variable Speed Limit Signs*.
- DOTD. 2023. "Falcon/WebSuite." Accessed July 5, 2023.
<https://www9.dotd.la.gov/falconwebv3/>.
- Edara, P., C. Sun, and Y. Hou. 2017. "Evaluation of variable advisory speed limits in congested work zones." *J. Transp. Saf. Secur.*, 9 (2): 123–145.
- FHWA (Federal Highway Administration). 2008. "Benefits of Using Intelligent Transportation Systems in Work Zones."
https://ops.fhwa.dot.gov/wz/its/wz_its_benefits_summ/wz_its_benefits_summ.pdf.
- FHWA (Federal Highway Administration). 2009. *Automated speed enforcement MUTCD installation guidelines.pdf*.
- FHWA (Federal Highway Administration). 2012. "Appropriate Speed Limits for All Road Users." <https://highways.dot.gov/safety/proven-safety-countermeasures/appropriate-speed-limits-all-road-users>.

- FHWA (Federal Highway Administration). 2014. "Variable Speed Limits - Safety." <http://safety.fhwa.dot.gov/speedmgt/vslimits/>.
- FHWA (Federal Highway Administration). 2015. "Use of Benefit-Cost Analysis by State Departments of Transportation: Report to Congress." https://www.fhwa.dot.gov/policy/otps/pubs/bca_report/.
- FHWA (Federal Highway Administration). 2020a. "Portable Changeable Message Signs Handbook." Accessed March 7, 2023. https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpv/reports/03066/?_gl=1*13dwdd3*_ga*OTY2Njc2Mjk1LjE2NjE3ODI5MzM.*_ga_VW1SFWJKBB*MTY4ODU3ODE3Mi4yNi4xLjE2ODg1ODEyMzMzM4wLjA.
- FHWA (Federal Highway Administration). 2020b. "Automated Enforcement | NHTSA." Accessed February 5, 2022. <https://www.nhtsa.gov/book/countermeasures/countermeasures/21-automated-enforcement>.
- FHWA (Federal Highway Administration). 2021a. *Safe System Approach for Speed Management*.
- FHWA (Federal Highway Administration). 2021b. "Proven Safety Countermeasures | Federal Highway Administration - Safety | Federal Highway Administration." Accessed March 12, 2022. <https://safety.fhwa.dot.gov/provencountermeasures/>.
- FHWA (Federal Highway Administration). 2021c. "Speed Safety Cameras." https://safety.fhwa.dot.gov/provencountermeasures/pdf/PSC_New_SpeedCamera_508.pdf.
- FHWA (Federal Highway Administration). 2022a. *Moving to a Complete Streets Design Model: A Report to Congress on Opportunities and Challenges*.
- FHWA (Federal Highway Administration). 2022b. "Work Zone Impacts Assessment in Maintenance and Operations - Work Zone Impacts Assessment - FHWA Work Zone." Accessed January 5, 2023. https://ops.fhwa.dot.gov/wz/resources/final_rule/wzi_guide/sec9.htm.
- FHWA (Federal Highway Administration). 2022c. *Moving to a Complete Streets Design Model: A Report to Congress on Opportunities and Challenges*.
- Fisher, D. L., A. Breck, O. Gillham, D. Flynn, and J. A. V. N. T. S. Center, and N. H. T. S. Administration. 2021. "Effectiveness of Dynamic Speed Feedback Signs Volume I: Literature Review and Meta-Analysis." I (August): 81p.
- Forbes, G. J., T. Gardner, H. McGee, and R. Srinivasan. 2012. "Methods and Practices for Setting Speed Limits: An Informational Report." 1–120.
- Gangireddy, S., R. Bian, T. Tolford, and H. Hassan. 2023. "Integrating the Concept of Complete Streets into Pavement Preservation." *Int. Conf. Transp. Dev. - Am. Soc. Civ. Eng.*, 289–298. International Conference on Transportation and Development.
- Hallmark, S. L., N. Hawkins, and O. Smadi. 2015. "Evaluation of Dynamic Speed Feedback Signs on Curves: A National Demonstration Project." *Am. Eval. Soc.*, (January).
- Iowa DOT. 2016. "Benefit-Cost Analysis." <https://www.burlingtoniowa.org/DocumentCenter/View/1259/TIGER---Appendix-D---Benefit-Cost-Analysis>.
- Iowa DOT. 2017. *Traffic Engineering Assistance Program - Information Sheet*.
- Iowa DOT. 2019. "Speed Feedback Trailer Signs." Accessed October 26, 2022. <https://iowadot.gov/design/dmanual/09b-11.pdf>.
- Jordan, S. W., and S. Ivey. 2021. "Complete Streets: Promises and Proof." *J. Urban Plan. Dev.*, 147 (2): 1–10.

- Knapp, K., D. Veneziano, and I. L. T. A. Program, and I. D. of Transportation. 2021. *Speed Feedback Sign Loan Program*.
- LADOTD (Louisiana Department of Transportation & Development). 2020a. "Speed Limit Signs | Louisiana Department of Transportation & Development." Accessed February 1, 2023. <https://gismapping-ladotd.opendata.arcgis.com/datasets/speed-limit-signs/explore?location=31.128023%2C-92.308970%2C9.00>.
- LADOTD (Louisiana Department of Transportation & Development). 2020b. "Number of Lanes | Louisiana Department of Transportation & Development." Accessed January 7, 2023. <https://data-ladotd.opendata.arcgis.com/datasets/number-of-lanes/explore?location=30.612334%2C-92.764404%2C15.70>.
- LADOTD (Louisiana Department of Transportation & Development). 2021. "LADOTD Functional System Map." Accessed January 5, 2023. <https://ladotd.maps.arcgis.com/home/webmap/viewer.html?webmap=a37461260bec43dea7bcbf6b710a662e>.
- McMurtry, T., M. Saito, M. Iffkin, and S. Heath. 2009. "Variable Speed Limits Signs: Effects on Speed and Speed Variation in Work Zones." *Proceedings TRB 2009 Annu. Meet.*, 17.
- Michigan DOT. 2004. "Dynamic Lane Merge System." <https://ops.fhwa.dot.gov/wz/technologies/michigan/index.htm>.
- National Center for Statistics and Analysis. 2022. "Speeding: 2020 data." Traffic Saf. Facts. Rep. No. DOT HS 813 320, (June): 1–18.
- Plan, A. W., and F. Year. 2015. "Automated Speed Enforcement Implementation: Survey Findings and Lessons Learned from around the Country."
- "SafePace 475 18" Radar Feedback Sign - Radar Feedback Signs | TAPCO. n.d. Accessed April 5, 2023. <https://www.tapconet.com/product/safepace-475-18-radar-feedback-sign?sku=146757>.
- SGA. 2016. "Changing Complete Streets Policy: A Brief Guidebook." https://smartgrowthamerica.org/wp-content/uploads/2016/09/Changing-Complete-Streets-Policy_Brief-Guidebook.pdf.
- SGA. 2023. "Complete Streets policies." Accessed January 12, 2023. <https://smartgrowthamerica.org/program/national-complete-streets-coalition/policy-atlas/>.
- Sunrisesesa. 2023. "True Compliance to MUTCD on Variable Speed Limit Signs." Accessed April 7, 2023. <https://web.sunrisesesatech.com/resources/getting-to-perfect-true-compliance-to-mutcd-on-variable-speed-limit-signs>.
- USDOT. 2009. *Manual on Uniform Traffic Control Devices*.
- USDOT. 2018. "Benefit-Cost Analysis Guidance for Discretionary Grant Programs." Off. Secr. U.S. Dep. Transp., 2022 (March).

Using the Estimated Available Friction at the Posted Speed as a Pavement Safety Performance Measure on Freeways

Behrokh Bazmara¹; Gerardo W. Flintsch, Ph.D., P.E.²; Edgar de León Izeppi, Ph.D.³;
and Samer W. Katicha, Ph.D.⁴

¹Ph.D. Candidate, Dept. of Civil and Environmental Engineering, Virginia Tech Transportation Institute, Blacksburg, VA. Email: behrokhb@vt.edu

²Dan Pletta Professor of Dept. of Civil and Environmental Engineering and Director of Center for Sustainable and Resilient Infrastructure, Virginia Tech Transportation Institute, Blacksburg, VA. Email: flintsch@vt.edu

³Research Scientist, Center for Sustainable Transportation Infrastructure, Virginia Tech Transportation Institute, Blacksburg, VA. Email: edeleonizeppi@vtti.vt.edu

⁴Research Scientist, Center for Sustainable Transportation Infrastructure, Virginia Tech Transportation Institute, Blacksburg, VA. Email: skaticha@vtti.vt.edu

ABSTRACT

Highway safety is impacted by pavement surface characteristics, especially microtexture and macrotexture. This research proposed a new approach to consider the combined effects of pavement macrotexture and low slip friction (mainly impacted by microtexture) to analyze pavement safety performance functions (SPFs). The study investigated the impact of microtexture and macrotexture into a single indicator, allowing a more comprehensive assessment of pavement safety. The implementation of estimated available friction at different speeds was analyzed using the speed correction equation in ASTM E1960-07 to model the expected number of crashes on the interstate freeway facility in the United States. The analysis showed that the estimated available friction at 120 km/h (FRS120) significantly combines the effect of friction and macrotexture on freeway crashes. This index allowed developing crash modification factors (CMFs) by effectively capturing potential reduction in expected crashes due to changes in both SFN40 (friction at a low slip speed) and macrotexture.

INTRODUCTION

According to the National Highway Transportation Safety Administration (NHTSA), motor vehicle crashes in 2023 on U.S. freeways is expected to result in approximately 1.24 fatalities per 100 million vehicle miles traveled (NHTSA 2022). While serious crashes frequently occur because of the driver's behavior, the contribution of pavement surface characteristics to the crashes cannot be disregarded. According to the Federal Highway Administration (FHWA), run-off road and rear-end crashes could also be prevented or reduced by improving friction and texture properties (Cenek, et al. 2011). Studies have found a strong statistical association between pavement surface frictional properties, including friction and macrotexture, and roadway crashes in dry and wet weather conditions (Noyce, et al. 2005). Maintaining appropriate friction is significant to facilitating crash prevention in roadway safety (AASHTO 2022).

Research in the U.K. found friction measured with a Sideway-force Coefficient Routine Investigation Machine (SCRIM, which mainly measures microtexture), not to be statistically significant to the freeway's safety evaluation indicating a relatively small impact on the

prediction of crashes. Similar research in the U.S. has produced mixed results; while some studies have shown a statistically significant impact of friction on crashes of interstate highways, others have not. Yet, studies indicated a statistically significant impact of surface macrotexture on crash counts (Ivan, et al. 2010).

Research shows that the most common crashes on freeways occur as an influence of low frictional properties (Xu, et al. 2022). According to the AASHTO Guide for pavement friction, it has been confirmed that there is a demand for frictional properties (micro- and macrotexture) demonstrating the amount of friction needed for vehicles to safely navigate the roadways (AASHTO 2022). Friction demand is a result of diverse factors for every section of the roadway, and is dominated by the roadway category, average daily traffic, roadway geometry, pavement surface texture, driving speed, and the presence of intersections, ramps, and crossings (Flitsch, et al. 2023). As vehicle speed increases, the tire-pavement interface decreases and the interaction between macrotexture and friction becomes more complex. Consequently, the interaction of the two parameters on higher speed should be considered to thoroughly analyze the frictional properties of the highway facility.

The objective of this research is to investigate the effects of combining microtexture and macrotexture into a single performance indicator to comprehensive assessment of pavement safety. The study investigated the effects of the estimated available friction at different speeds, implementing the speed correction equation in ASTM E1960-07 (ASTM 2015), to model the expected number of crashes by implementing Negative Binomial regression (NB) on the Interstate freeway facility in the United States. The research considered annual average daily traffic (AADT), pavement macrotexture (MPD), sideway-force friction, and roadway surface geometric properties to determine the effects of combined macrotexture and microtexture characteristics. The data were collected from three states over a three-year period.

Background

Examining pavement texture only is insufficient to clarify the interaction between tires and the road surface. Available research on the use of frictional properties (macrotexture and microtexture) in the development of Safety Performance Functions (SPFs), showed limited studies on the interaction of friction and macrotexture characteristics in the analysis of roadway safety. A few studies have specifically looked at skid-related crashes or included macrotexture and microtexture in weather-related crashes.

Kummer (1996) conducted a study on simulating the interaction of the tire and pavement surface. The study analyzed the relationship between friction components, adhesion and hysteresis, and the properties associated with pavement texture. It was found that while tire and vehicle design contribute to frictional force, the critical role of pavement texture parameters (texture depth), on influencing the slip and friction, particularly on surfaces with a fine aggregate particle should be considered.

Corney and Corney (1997) investigated the influence of texture depth on the frictional properties of the pavement surface. The authors showed that a surface with low texture exhibits greater friction at low speeds compared to a surface with high texture, which is created using a coarse texture. Further, it was concluded that a fine texture can elevate low-speed friction when constructed with a high-texture surface.

Ergun et al. (2005) used a statistic regression model to analyze the effects of surface texture on frictional properties. The authors concluded that as macrotexture increases, friction decreases.

Further, while skid speed is assigned in low and medium ranges (equal to or less than 18 km/h), friction increases slowly as skid speed was raised to higher level.

Cairney and Germanchev (2006) investigated the correlation between stopping distance, operating speed, site condition, and the frictional properties for Australian Transport Safety (ATSB). Study showed that there is not a clear correlation between macrotexture and stopping distance for wet and dry conditions. Also, it was found that the interaction of speed and pavement condition correlates with higher frictional properties.

Hall et al. (2009) demonstrated that at higher speeds on wet pavement, a significant portion of friction originates from macrotexture. Nevertheless, at low speeds, microtexture takes rank in influencing frictional forces, whereas at high speeds, MPD becomes the dominant factor in providing frictional properties.

Wallbank et al. (2016) evaluated the influence of texture depth and friction for the United Kingdom's Strategic Road Network (SRN) based on Generalized Linear Model (GLM) regression techniques. The study confirms that macrotexture affects all types of crashes on all types of roadways. Further, research indicated that texture depth is more predictive than the low slip friction for all crash types on freeways. It was also found that friction is only a quite significant predictive of wet weather crashes than macrotexture to the analysis of the highway safety.

Flintsch et al. (2023) used a dataset to develop crash modification factors (CMF) for pavement friction and macrotexture on various roadways and site types, for instance, curves, segments, intersections, and ramps. The study found that higher friction (on all road categories and site types) and macrotexture (mainly on high-speed roadways) result in lower crash rates. Research concluded that an increase in SFN40 could potentially lead to a reduction of up to 30 percent in total crashes.

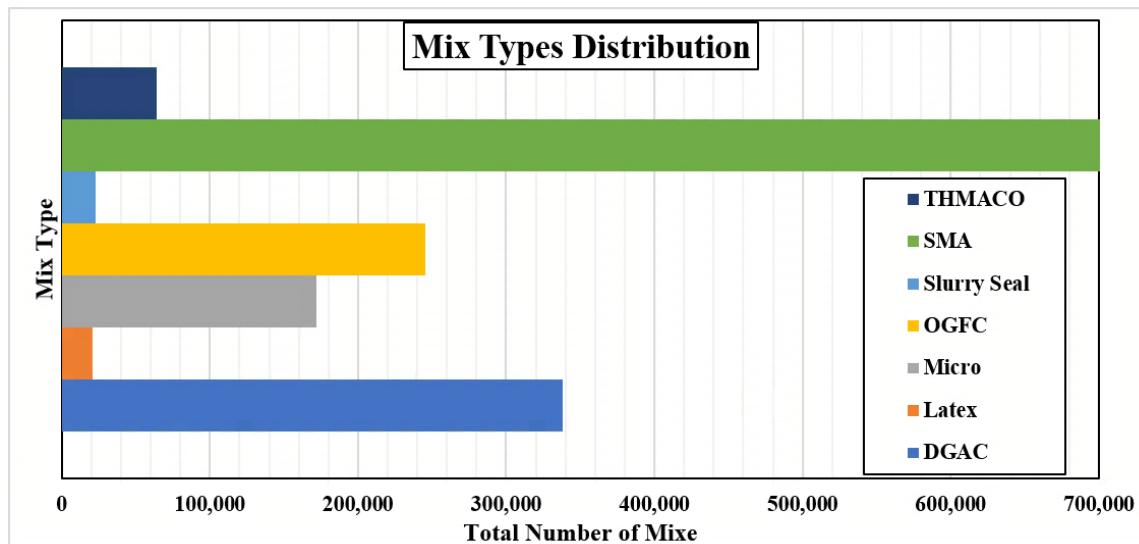
Although literature confirmed the contribution of macrotexture to all types of crashes for all types of roadways, the low slip friction was found only a significant predictor to wet crashes on the high-speed roadways. In addition, review indicates that there are many studies, but the evaluation of interplay between macrotexture and friction is missing and still needs to be considered.

Methodology

Available data

The data utilized in this study was gathered through multiple research projects employing the SCRIM machine. This device assesses wet friction at a low slip speed and macrotexture on the pavement surface. The continuous measurement of surface texture properties serves as a preventive measure against severe crashes by identifying locations where the friction demand exceeds what the pavement can provide at the given operating speed (AASHTO 2022).

The dataset comprises information on state, route type, functional classification of roadways, friction (SFN40) and macrotexture (MPD), traffic (AADT), surface mix type, road geometry, and crash data for freeway sites. The friction, texture measurements, and roadway geometry were measured with a SCRIM, while AADT and crash data were sourced from the respective Department of Transportation (DOTs) databases of each state. The data was compiled on 0.1-mile sections. The summary of distribution by different mix types in the data set is provided in Fig 1.

**Fig. 1. Mix type's distribution.**

The research considered a variety of roadway geometry such as cross-slope, gradient, horizontal curvature, lane count, route type, and presence of intersection. Regarding the functional classification of the roadway, freeway's ramp, curve, and segment were considered. Table 1 provides more information on the data set.

Table 1. Predictors Considered and Ranges.

Parameter estimates	Unit	Maximum	Minimum	Mean
Grade	Percentage	7.40	-7.99	—
Cross Slope	Percentage	10.80	0.07	2.52
Horizontal Curvature	1/meter	0.007	0.065	0.001
AADT	Count	132,000	3,100	19,394
Friction	Number	87.9	8.3	51.6
MPD	Millimeter	2.82	0.27	0.77

Notes: Divided accounts for if the roadway is divided, it is 1; otherwise= 0.

Estimating Available Friction at Different Speed

To evaluate methods of combining friction and macrotexture, the research estimated the available friction measurements employing the speed correction equation in the ASTM E1960-07 (ASTM 2015). The estimated friction, $FR_S(v)$, at speed of 60 km/hr. 100 km/hr., and 120 km/hr. were computed for all segments using Eq. (1) and Eq. (2). The slip ratio of the SCRIM survey device is 34 percent of the speed. In this equation, s_p is the speed correction factor, $FR(S) =$ Friction value for the devise used at a slip speed (s), v is speed at which the friction is being estimated (km/h), and MPD represents pavement macrotexture in millimeter.

$$FRS(v) = FR(s) \cdot \text{Exp}\left[\frac{(s-v)}{s_p}\right] \quad (1)$$

$$s_p = 14.2 + 89.7 \cdot MPD \quad (2)$$

The initial stage of the SPF analysis involved transforming friction measurements obtained from the SCRIM operating at 40 mph (SFN40) into a range of testing speeds. This conversion was accomplished by applying the speed correction technique as outlined in Eq. (1) and Eq. (2). The assessment of the impact of the joint index (FRS) at different testing speeds (60 km/hr., 100 km/hr., and 120 km/hr.) was done utilizing network-level data on freeway curves, segments, and ramps. Following this, the speed correction was implemented on the FR(S) values, Friction value for the device used at a slip speed s, leading to the conversion of friction values across various speeds.

Safety Performance Function (SPF)

The impact of roadway characteristics on crashes is typically investigated using SPFs. This paper developed various Negative Binomial (NB) regression models using different frictional properties, as recommended in the FHWA report (FHWA 2010), shown in Eq. (3). In this equation, the expected crash count (SPF) is linked with the average annual daily traffic (AADT), roadway characteristics, and frictional properties through a log-link function. NB regression was applied to generate the SPF models by evaluating the effect of various frictional properties on the expected number of crashes. In this equation β_0 is the model intercept, β_1 is regression coefficient for Average Annual Daily Traffic, and β_j is regression coefficient for the additional predictors ($j=2,.., k$).

$$SPF = \text{Exp}(\beta_0 + \beta_1 \cdot \ln(AADT) + \sum_{i=1}^k \beta_j X_{ij}) \quad (3)$$

Crash Modification Factor (CMF) and Potential Crash Reduction (CR)

To quantify the effectiveness of a specific safety improvement or countermeasure in reducing the frequency or severity of crashes the crash modification factor (CMF) was calculated using Eq. (4). CMF is a statistical measure that compares the expected number of crashes before implementing the countermeasure (Friction _{Before}) to the expected number of crashes after (Friction _{After}) its implementation using the estimate coefficient of friction from the SPF analysis (β_j). Crash reduction (CR) was computed based on the observed crashes (C) in Eq. (5) (FHWA 2014). The CMF is derived from the estimated SPFs.

$$CMF = \text{Exp}(\beta_j (\text{Friction}_{\text{After}} - \text{Friction}_{\text{Before}})) \quad (4)$$

$$CR = C \cdot (1 - CMF) \quad (5)$$

RESULTS

SPF Regression Modeling

In this section the outcomes of the SPF model results are presented. Table 2 shows the model fitted with AADT, mean profile depth (MPD), SFN40 and roadway geometry factors for all

crashes considering freeway facilities. The table displays the model without the interaction term of SFN40 and macrotexture (MPD), as well as the model with the inclusion of this interaction term. Both SPF analyses indicate the statistical significance of MPD in predicting expected crashes (p -value < 0.05), whereas SFN40 does not demonstrate to be a predictive of freeways analysis. This validates the study findings derived from the United Kingdom's Strategic Road Network (SRN), which utilized GLM and multivariate regression techniques. Consistently, the observed trend indicates that pavement texture (MPD) is a more predictive than friction (SFN40) for all freeway crashes (Wallbank, et al. 2016).

The second model, however, reveals the statistical significance of the interaction between SFN40 and MPD. The Akaike Information Criterion (AIC) supports the advantage of the model incorporating the interaction term (SFN40 and MPD) for predicting expected crashes, as evidenced by its lower AIC. Additionally, the model with the interaction demonstrates a slightly improved coefficient of determination (R^2).

Table 2. Negative Binomial Regression Output for Total Crashes, Using SFN40 and MPD.

Parameter estimates	Without interaction		With interaction	
	β	p -value	β	p -value
Intercept	-13.4	0.000	-13.5	0.000
Route Type	-0.088	0.048	-0.139	0.002
Intersection	0.459	0.000	0.453	0.000
Lane Count	-0.047	0.022	-0.020	0.320
Ln AADT	1.40	0.000	1.37	0.000
Friction	0.002	0.882	0.003	0.029
Texture	-0.336	0.000	-0.119	0.003
Grade (%)	-0.036	0.000	-0.033	0.000
Cross-slope (%)	-0.073	0.000	-0.069	0.000
Horizontal Curvature (1/m)	546.2	0.000	518.1	0.000
Interaction term (SFN40 and MPD)	—	—	0.042	0.000
R^2		0.397		0.401
Number of Parameters		11		12
AIC		53,315		53,227
Overdispersion		1.28		1.30
NLL		-26,647		-26,602

Notes: AIC = Akaike Information Criterion; NLL = Negative Log-Likelihood; Route types are Interstate and Primary; and β = Represents the estimate coefficient.

Table 3 provides a summary of the NB regression results for the SPF models, utilizing the estimated available friction (FRS60, FRS100, and FRS120). These estimates are derived by converting the measured friction (SFN40) and macrotexture values.

In this case, the combined indices exhibit statistical significance, as anticipated, showing a negative coefficient for FRS. This implies that as the combined index increases, there is a tendency for the estimated crashes to decrease. This suggests that the FRS index effectively

captures the combined influence of the friction measure at a low slip ratio (representing macrotexture estimation) and macrotexture on the likelihood of crashes.

The AIC result indicates that the model, specifically considering the estimated available friction at 120 km/hr. (FRS120), offers the optimal fit with the lowest AIC of 53,376. Therefore, the SPF model based on FRS120 was selected for the modeling of freeway facilities within the dataset. Since freeways are designed for higher speeds, FRS120 proves to be more relevant to the conditions that vehicles experience on freeway roadways.

Table 3. Negative Binomial regression output for total crashes, using FRS60, FRS100, and FRS120.

Parameter estimates	FRS 60 (km/hr.)		FRS 100 (km/hr.)		FRS 120 (km/hr.)	
	β	p-value	β	p-value	β	p-value
Intercept	-13.2	0.000	-13.1	0.000	-13.1	0.000
Route Type	-0.169	0.001	-0.143	0.001	-0.133	0.003
Intersection	0.449	0.000	0.451	0.000	0.452	0.000
Lane Count	-0.051	0.011	-0.045	0.025	-0.045	0.026
Ln AADT	1.38	0.000	1.37	0.000	1.37	0.000
Friction Index	-0.006	0.001	-0.012	0.000	-0.014	0.000
Grade (%)	-0.042	0.000	-0.04	0.000	-0.039	0.000
Cross-slope (%)	-0.089	0.000	-0.086	0.000	-0.084	0.000
Horizontal Curvature (1/m)	575.6	0.000	571.1	0.000	568.6	0.000
R2	0.394		0.395		0.395	
Number of Parameters	10		10		10	
AIC	53,410		53,385		53,376	
Overdispersion	1.25		1.26		1.26	
NLL	-26,695		-26,683		-26,678	

Notes: AIC = Akaike Information Criterion; NLL = Negative Log-Likelihood; Route types are Interstate and Primary; and β = Represents the estimate coefficient.

Mix Type Comparison

To verify the appropriateness of the proposed indicator, the research investigated the relative rank of the various pavement surfaces to provide insight into the interaction between different pavement mix types and frictional measurements, Fig 2 and Fig 3.

Referring to figures, it becomes apparent that the estimated available friction at 120 km/hr. presents a more noticeable ranking of the mixture types. Notably, in Fig 2 the OGFC stands out with the highest available friction, aligning consistently with the observed trend of the lowest crash rates on these AC pavements.

Examining the position of OGFC in Fig 3, it is evident that SFN40 alone fails to present the advantages of these surface types. The SFN40 of the OGFC mix indicates lower values compared to other surfaces because it predominantly reflects microtexture, while the dominant feature in these coarse surfaces is macrotexture. Conversely, the FRS120 index comprehensively captures the interaction between macrotexture and friction, providing a more accurate

representation of the OGFC mix's ability to maintain favorable friction characteristics at higher speed.

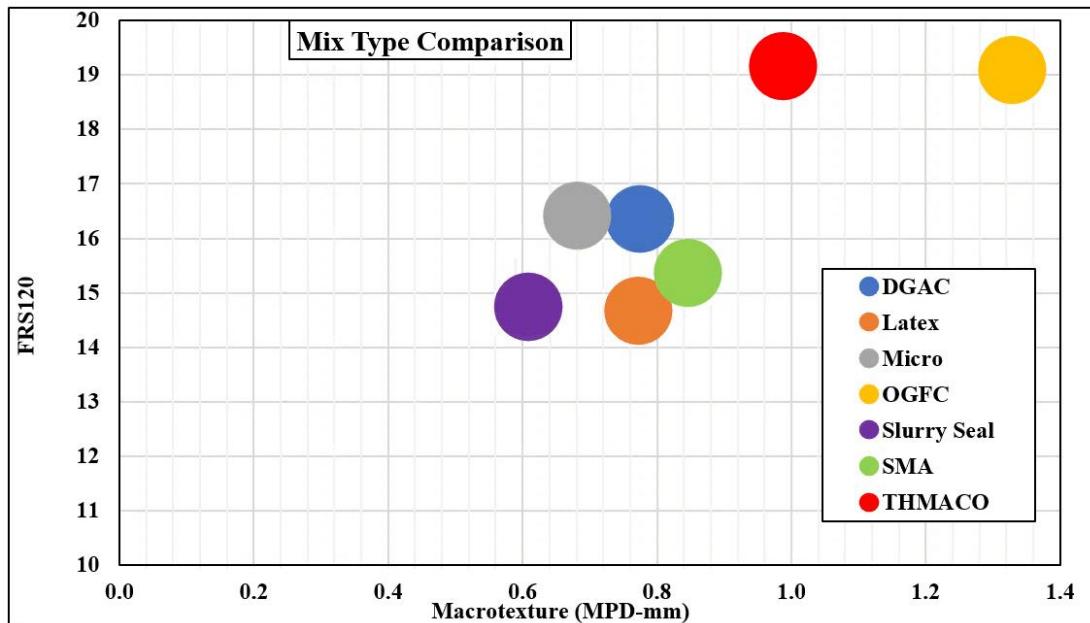


Fig. 2. Plot of FRS120 vs the macrotexture (MPD) by surface mix.

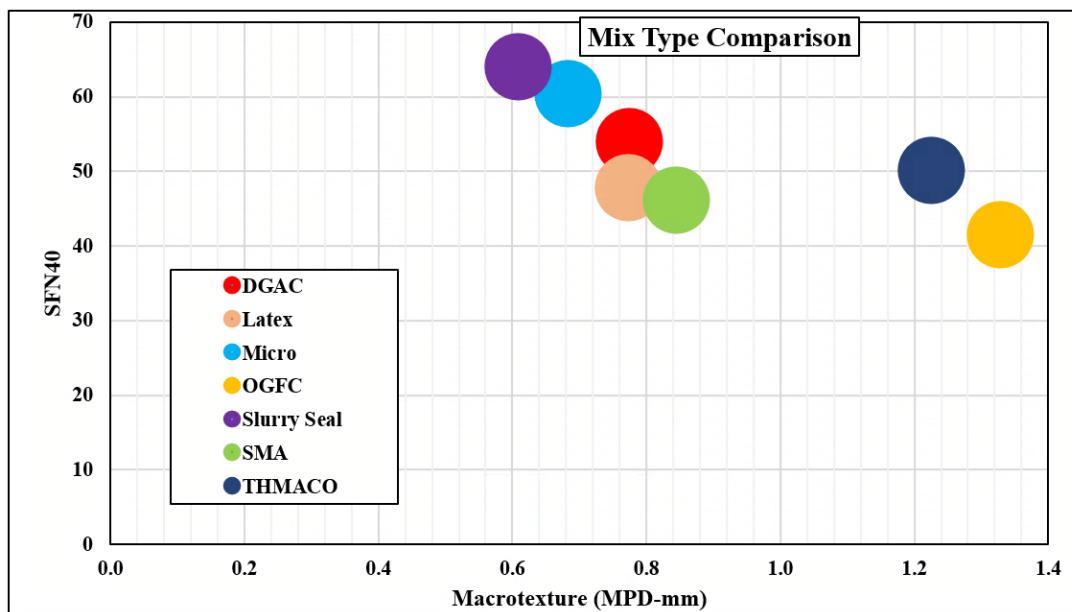


Fig. 3. Plot of SFN40 vs macrotexture (MPD) and surface mix.

Example Application

To illustrate the potential use of the proposed friction indicator, the research used the developed SPF_s to estimate the potential crash reduction of resurfacing a pavement. The

presented graph in Fig 4 depicts the Crash Modification Function (CMFx) for different friction improvements. The general trend shows lower CMF (indicating higher potential crash reductions) for those with higher friction enhancements.

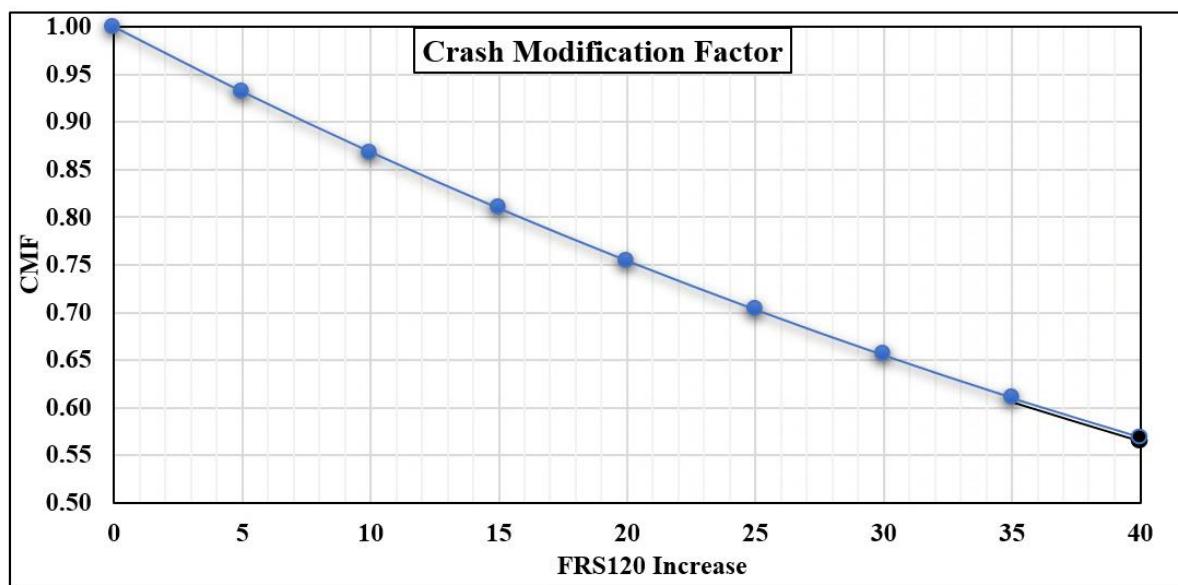


Fig. 4. Estimated Crash Modification Factor as a function of friction (FRS120) Improvement.

An illustrative Benefit-Cost analysis was conducted to demonstrate the utility of FRS120 in comparing potential friction improvement treatment. The objective was to calculate the potential benefits per dollar invested associated with two candidate friction improvement treatments. The following table, Table 4, displays the characteristics of the selected sections with DGAC mix surface, one with low macrotexture and another with low friction to evaluate in the analysis. It is considerable to note that the alternative treatments have similar FSR120.

Table 4. Cost for FRS120 Development Based on the Surface Type Treatment.

Before Treatment					
Type	FRS120	MPD-mm	SFN40	AADT	Crashes
DGAC	5	0.30	55	70,000	15
DGAC	5	0.45	30	65,000	20

Table 5, presents the assumed average cost per 0.1-mile segment by the pavement type treatment in the selected segment of Interstate and expected friction and macrotexture after the treatment. The potential efficiency of the two friction enhancement treatments, Stone-Matrix Asphalt (SMA) and High Friction Surface Treatment (HFST) were evaluated. It assesses the corresponding FRS120, the Crash Modification Factor (CMF) and potential crash reduction corresponding to the increase in FRS120, as well as the estimated socio-economic benefit using an estimated cost per crash (Flintsch, et al. 2023).

Table 5. Benefit-Cost Analysis of the FRS120 Development.

FRS120 Enhancement									
Treatment Options	FRS120	MPD	SFN40	CMF	Costs (per/lane/0.1 Segment)	CR	Benefit	B/C	
SMA	17.5	0.80	55	0.84	\$8,525	2.40	\$267,244	31.3	
HFST	35.7	1.20	80	0.65	\$19,000	5.30	\$580,824	30.6	

The analysis for the illustrative example shows potential reduction in crashes of 26-percent and 35-percent for the SMA and HFST treatment, respectively. However, given the high cost of the HFST, looking at the benefit-cost (B/C) ratio, the SMA treatment yields a slightly higher B/C, 31.3, than the HFST. It can be concluded that compared to HFST, the SMA treatment should be recommended.

CONCLUSION

This paper proposes a new pavement safety performance index for freeways, the estimated available friction at 120 km/h (FRS120), which combines the effects of friction and macrotexture on freeway crashes. The presented analysis allows to draw the following conclusions:

- The estimated available friction at 120 km/h (FRS120) allowed to develop safety performance functions (SPFx) and crash modification factors (CMFs) that effectively captures the potential reduction in expected crashes due to improvements in both, SFN40 (friction at a low slip) and macrotexture.
- Negative Binomial regression analysis confirmed that FRS120 had a statistically significant effect on estimated crash counts on freeways.
- The application of the FRS120 index on a case study illustrated how the index could be used to evaluate alternative treatments and select the most cost-effective one.

Subsequent investigations should explore the application of similar indices, correlated with designated speed, across diverse categories of roadways characterized by varying functional classifications and posted speeds. Additionally, further research should discover the generalizability of the obtained results to alternative types of pavement surface.

REFERENCES

- AASHTO. (2022). *Standard method of test for continuous measurement of sideway-force friction number for highway pavements*, TP 143-21. Washington, DC: AASHTO.
- ASTM. (2015). *Standard Practice for Calculating International Friction Index of a Pavement Surface*. ASTM E1960 – 07. West Conshohocken, PA: ASTM.
- Cairney, P., and Germanchev, A. (2006). “A pilot study of the effects of macrotexture on stopping distance”. CR 226. Australian Transport Safety Bureau.
- Cenek, P., Brodie, C. A., Davies, R. B., and Tate, F. N. (2011, May). “A prioritization scheme for the safety management of curves”. *3rd International Surface Friction Conference*.

- Croney, P., and Croney, D. (1997). *The design and performance of road pavements*. McGraw-Hill book company Europe, Maidenhead.
- Ergun, M., Iyinam, S., and Iyinam, A. F. (2005). "Prediction of road surface friction coefficient using only macro- and microtexture measurements", *Journal of Transportation Engineering*, 131(4), 311–319.
- FHWA (Federal Highway Administration). (2010). "Safety Performance Function Decision Guide." SPF Calibration vs SPF Development. (FHWA-SA-14-004)." Accessed July 20, 2023. spf_decision_guide_final_0.pdf (dot.gov).
- FHWA (Federal Highway Administration). (2014). "Crash Modification Factors (CMFs). Local and Rural Road Safety Briefing Sheets. (FHWA-SA-14-083)." Accessed July 20, 2023. Crash Modification Factors (CMFs) | FHWA (dot.gov).
- Flintsch, G., de León Izeppi, E., McCarthy, R., Katicha, S., Persaud, B., Guo, F., Medina, A., and Tobias, P. (2023). Characterizing Road Safety Performance Using Pavement Friction. FHWA SA-23-006. Washington, DC: US DOT and Federal Highway Administration.
- Hall, J., Smith, K., Titus-Glover, L., Wambold, J., Yager, T., and Rado, Z. (2009). "Guide for pavement friction". Web-Only Document 108. Washington, DC: National Cooperative Highway Research Program and Transportation Research Board.
- Ivan, J. N., Ravishanker, N., Jackson, E., Guo, S., and Aronov, B. (2010, Nov). "Incorporating wet pavement friction into traffic safety analysis". Report No. JHR 10-324. University of Connecticut, Connecticut Transportation Institute.
- Kummer, H. W. (1966). "Unified theory of rubber friction", Engineering Research Bulletin. B-94, Penn State University, State College, University Park, Pa.
- Noyce, D. A., Bahia, B., Yambó, J. M., and Kim, G. (2005). "Incorporating Road Safety into Pavement Management" *Maximizing Asphalt Pavement Surface Friction for Road Safety Improvement*, 19-110.
- NHTSA. (2022). Traffic Safety Facts: Early Estimate of Motor Vehicle Traffic Fatalities for the First Quarter of 2023. DOT HS 813 482., US Department of Transportation.
- Wallbank, C., Viner, H., Smith, L., and Smith, R. (2016). "The relationship between collisions and skid resistance on the strategic road network (Project Report PPR 806)". Transportation Research Laboratory.
- Xu, G., Zineddin, A., Atkins, R., and Abel, S. (2022). *Speed Management is Key to Road Safety*. Transportation Research Board, Washington, D.C.

Exploring Statistical Methods in Developing Safety Performance Functions

Xingjing Xu, Ph.D.¹; and Ilir Bejleri, Ph.D.²

¹Dept. of Urban and Regional Planning, Univ. of Florida, Gainesville, FL.

Email: axuxinjing@ufl.edu

²Dept. of Urban and Regional Planning, Univ. of Florida, Gainesville, FL. Email: ilir@ufl.edu

ABSTRACT

A safety performance function (SPF) is a tool to identify explanatory variables prone to traffic crashes in safety systemic analysis. This study explores advanced statistical methods for developing SPFs. Traditional negative binomial regression, while linear and interpretable, could have inaccurate performance with an increasing number of variables included in SPFs. We explore alternative methods and compare them to the traditional method, including best subset selection, shrinkage, generalized additive model, and tree-based method. Using Florida curve crash data, the study examines various risk factors such as traffic volume, demographics, curve characteristics, roadway characteristics, and spatial relationships, and develops SPFs. The study assesses the contributing risk factors, the overall SPFs performance, the methods' capabilities in handling statistical challenges, and variable interpretations. The findings show that complex methods, such as generalized additive model, outperform the traditional approach, suggesting their potential application in SPF development for more accurate crash predictions and variable interpretations.

1. INTRODUCTION

In transportation safety analysis, there are two main approaches to identify the dangerous locations on roadway networks: hot-spot analysis and systemic analysis. The hot-spot analysis mainly uses the historical crash data to identify the hazardous locations with high crash frequencies (FHWA 2015). The systemic analysis needs both historical crash data and contributing risk factors such as traffic volume, roadway characteristics and other explanatory variables. In the systemic approach, the transportation engineers and planners develop safety performance functions (SPF), screen the traffic networks, apply the SPFs to obtain the estimated crash numbers for roadway segments, and locate the most high-risk sites with the large estimated crash numbers (Preston et al. 2013; Srinivasan et al. 2013). Subsequently, a detailed examination of these high-risk sites involves a comparison between predicted and observed crash frequencies, along with an analysis of contributing risk factors at each location. These factors play an important role in a comprehensive understanding of the reasons behind high-risk locations, which helps to identify appropriate countermeasures and improvements for each hazardous site, ultimately contributing to saving lives.

The systemic analysis is highly interpretable, since it includes contributing risk factors associated with traffic crashes. It is also the only feasible approach when the historical crash data is unavailable, a common scenario for roadways where crash data is limited. Therefore, to select the well-interpreted explanatory variables and develop an accurate SPF is a key step in transportation safety systemic analysis.

Various statistical methods exist for developing SPFs. The Highway Safety Manual (HSM) (AASHTO 2010) recommends the negative binomial regression within the generalized linear

model (GLM) family, a widely applied method for its effective interpretation of linear statistical relationships. However, the negative binomial method tends to become unstable and may fail to produce accurate crash estimates, particularly as the number of variables increases (Wang et al. 2016; Xie and Zhang 2008). The SPF developed using the negative binomial method, as suggested in the HSM, only has two variables: the traffic volume and the roadway segment length (AASHTO 2010; Donnell et al. 2014). While researchers have been providing extensive information and incorporating more data in transportation safety analysis, there is a need for more advanced methods for variable selection in SPF development, beyond the traditional negative binomial regression.

In this study, various statistical methods were examined to identify contributing risk factors as explanatory variables and to develop SPFs for curves—a basic element in roadway networks. The study used horizontal curve data from across the state of Florida. Besides the traditional negative binomial regression, complex methods such as GLM with best subset selection and the shrinkage method, the generalized additive model (GAM) method and the tree-based method (Random Forests) were explored. The study examined the impacts of different variables on crashes, including traffic volume, demographic, roadway, spatial and curve characteristics. It further compared the contributing risk factors and model performance of each SPF.

The paper is organized as follows: Section 2 reviews the theoretical background of various statistical methods for variable selection in SPF development; Section 3 described the study methods, including data collection and the variables listed for analysis; Section 4 conducts the variable selection and SPF development using each method, and assesses the SPF performance; Section 5 compares the variable selection results and discusses the methods' capabilities in handling statistical challenges; Section 6 presents the conclusions and recommendations for the variable selection method in SPF development.

2. THEORETICAL BACKGROUND

Safety performance functions (SPF) aim to characterize the relationship between crash frequencies and contributing risk factors (AASHTO 2010). Initially, Poisson regression was the primary method for SPF development, as it aligns with the discrete counting nature of traffic crash data. However, researchers discovered that the crash frequencies always have a larger variance comparing to the mean, which violates the Poisson distribution properties (Miaou and Lum 1993). Therefore, negative binomial regression (NB), with the logarithm of crash frequencies as the response variable, was proposed for SPFs in safety systemic analysis, also explored in this paper.

2.1. Negative Binomial Generalized Linear Model (NBGLM)

The HSM suggests the SPF for a roadway segment in the following formula (AASHTO 2010):

$$\mu_i = e^{\beta_0 * L * AADT^{\beta_1}} \quad (1)$$

where μ_i = expected crash frequencies on roadway segment i;

L = roadway segment length;

AADT = average annual daily traffic;

β_0 = coefficient for the intercept;

β_1 = coefficient for AADT.

The manual also suggests that the SPF can be varied into different forms with more explanatory variables:

$$\mu_i = e^{\beta_0} * L * AADT^{\beta_1} * e^{(\beta_2 X_2 + \dots + \beta_p X_p)} \quad (2)$$

where X_2, \dots, X_p = value of the j^{th} explanatory variable on roadway segment i ($j = 2, \dots, p$);

β_2, \dots, β_p = coefficient for the j^{th} explanatory variables ($j = 2, \dots, p$).

Taking the logarithm on both sides of Equation 2 results in:

$$\log(\mu_i) = \beta_0 + \log(L) + \beta_1 \log(AADT) + \beta_2 X_2 + \dots + \beta_p X_p \quad (3)$$

which could be further rearranged into the form:

$$\log(\mu_i) = \beta_0 + \sum_{j=1}^p \beta_j X_{ij} \quad (4)$$

where μ_i = expected crash frequencies on roadway segment i ;

β_0 = coefficient for the intercept;

β_j = coefficient for j^{th} explanatory variable ($j = 1, \dots, p$);

X_{ij} = value of the j^{th} explanatory variable on roadway segment i ; and

p = number of explanatory variables.

The right side of Equation 4 indicates a linear relationship between the logarithms of the expected crash frequencies with the explanatory variables. Even though the relationship between the actual crash frequencies with the variables are nonlinear because of the logarithms, it can still be modelled by the generalized linear model (GLM). The method is named the negative binomial regression in the family of GLM, denoted as NBGLM (Wang et al. 2016; Xie and Zhang 2008).

NBGLM is widely used in the development of SPFs for safety systemic analysis. It is well-known for its good representation of linear relationship between the logarithms of expected crash frequencies and explanatory variables. NBGLM is often used when the variables for SPFs are already known. Commonly included variables, such as AADT and roadway segment length, as well as speed limits and various roadway design factors, are often considered to be included in the SPFs. Researchers have made great efforts to enhance the roadway database, providing more data for SPFs and informing transportation engineers and planners for improved safety systemic analysis. Despite this, the traditional NBGLM lacks a robust method for variable selection. Including all significant variables in the SPF can lead to model instability and inaccurate crash frequency estimations, particularly when dealing with a larger number of variables (James et al. 2013; Wang et al. 2016; Xie and Zhang 2008).

In contrast to the conventional NBGLM, two main methods can be integrated and are useful to perform variable selection. The best subset method involves identifying a subset of variables related to the response by fitting various combinations of variables in the model. The shrinkage method, on the other hand, is an advanced approach where coefficients of variables are shrunk towards or exactly to zero by imposing restrictions (James et al. 2013; Lord et al. 2005). Both these methods aim to select variables for developing SPFs with a reduced set of variables that ultimately result in optimal performance.

NBGLM estimates the coefficients of variables by minimizing the quantity:

$$\sum_{i=1}^n \left(\log(\mu_i) - \beta_0 - \sum_{j=1}^p \beta_j X_{ij} \right)^2. \quad (5)$$

The shrinkage method with variable selection is a powerful method to stabilize the regressions and select variables. Some coefficients of the explanatory variables would be shrunken towards or exactly to zero. This is how the method performs the variable selection (James et al. 2013; Lord et al. 2005; Zou and Hastie 2005). The way the shrinkage method estimated the coefficients of variables is similar as the Equation 5, but adding a penalty term:

$$\sum_{i=1}^n \left(\log(\mu_i) - \beta_0 - \sum_{j=1}^p \beta_j X_{ij} \right)^2 + \lambda \sum_{j=1}^p \hat{\omega}_j |\beta_j|. \quad (6)$$

where λ = the tuning parameter chosen through the cross-validation;

$\hat{\omega}$ = the adaptive weights.

The Lasso (Least Absolute Shrinkage and Selection Operator) method is one shrinkage method to select important variables. The adaptive Lasso introduces the adaptive weights $\hat{\omega}$, which ensure that the penalty imposed on each coefficient is proportional to its estimated importance. This method iteratively adapts the penalty weights to each variable's estimated coefficients, effectively enhancing the selection of relevant variables while reducing the impact of irrelevant ones in the regression model. The R package *glmnet* was used to minimize the above quantity (Wood 2017).

The above methods all belong to the family of linear methods, showing the linear relationships between variables. To account for potential nonlinear relationships among variables, more intricate statistical methods can be considered for variable selection and the development of SPF.

2.2. Negative Binomial Generalized Additive Model (N BGAM)

The Generalized additive model (GAM) introduces the smoothing functions for each variable, which allows for more flexibility in developing SPF. The following function is commonly used for the negative binomial generalized additive model (NBGAM) and was adopted for this paper:

$$\log(\mu_i) = \beta_0 + \sum_{j=1}^p f_j(X_{ij}) \quad (7)$$

where μ_i = expected crash frequencies on roadway segment i;

β_0 = coefficient for the intercept;

f_j = smooth function for the j^{th} explanatory variable ($j = 1, \dots, p$);

X_{ij} = value of the j^{th} explanatory variable on roadway segment i; and

p = number of explanatory variables.

Comparing Equation 4 and 5, the main difference between NBGLM and NBGAM exists in the smooth functions of the variables. If all the smooth functions of explanatory variables are constants, then the NBGAM can be treated as NBGLM. NBGAM is more powerful than NBGLM for its adaptiveness for nonlinearity. The smooth functions, also known as the curvy splines, can be any functions of explanatory variables, mostly the quadratic and cubic functions, allowing higher orders of polynomial relationship into modelling. For example, the smooth function for the j^{th} explanatory variable for the roadway segment i can be written as:

$$f_j(X_{ij}) = \alpha_{j0} + \alpha_{j1}X_{ij} + \alpha_{j2}X_{ij}^2 + \alpha_{j3}X_{ij}^3 \quad (8)$$

where α_{jk} = coefficients of the polynomial terms in the smooth function for the j^{th} explanatory variable on roadway segment i ($k = 0, 1, 2, 3$).

Since the NBGAM introduces more terms such as X_{ij}^2 and X_{ij}^3 , the SPF from NBGAM are much more flexible in explaining the nonlinear relationship between variables and thus have an improved performance than the SPF from NBGLM.

The issue with NBGAM is the interpretation of the variable selected in SPF. With high orders terms included in the model, it becomes much harder to interpret since the relationship between the explanatory variables with the responding crash frequencies is now modeled using a curvy spline. In an SPF from traditional NBGLM, a positive sign is interpreted as an increase, while a negative sign is understood as a decrease, in crash frequencies. While in an SPF from NBGAM, there is not a definite increase or decrease in number of crashes with the changing of the explanatory variables (James et al. 2013; Wood 2017; Xie and Zhang 2008). The R package *GAM* was used to perform a backward variable selection in NBGAM (Hastie 2023).

2.3. Tree-Based Method

The tree-based method is also a straightforward but effective method to capture nonlinear relationships among variables. It does not have a specific written formula, and instead uses a regression tree diagram to predict crash frequencies. Each branch in the tree diagram represents a combination of explanatory variables and produces a prediction value for the response—crash frequency. However, when dealing with a large number of variables, the regression tree might produce numerous branches, leading to potential inaccuracies due to overfitting. The Random Forests (RF) method's variable selection is robust, helping prune branches and obtain a more reasonable regression tree (James et al. 2013). It is a powerful algorithm for variable selection. It generates a lot of regression trees, where each tree uses a random subset of variables. Then it produces an output that ranks variables based on their importance in prediction. In the study, we identified variables with low importance in prediction, excluded them from the study. The R package *randomForest* was used to perform the modelling (Liaw and Wiener 2002).

3. METHODS

3.1. Data Collection

This paper focused on the horizontal curves in Florida as the interest of roadway network. Curves represent a series of changing directions and super-elevations in roadway alignment. They have larger number of crashes compared to the tangent sections of the same road (Albin et al. 2016; McGee et al. 2006). However, there is limited literature addressing the SPF for curves compared to the general roadway segments, primarily due to the surprising lack of availability of curve data in standard GIS roadway databases. Our research team developed a tool that automatically obtains the curve data from GIS centerline map (Bejleri et al. 2021). This dataset includes curve locations and curve characteristics such as curve radius, curve length and central angle.

The state of Florida has over 100,000 curves maintained by the Florida Department of Transportation (FDOT), providing comprehensive roadway characteristics. The study focused on

a sample of 685 curves, using a 5-year crash data (2013-2017) from Signal Four Analytics (<https://signal4analytics.com/>) to conduct the analysis.

3.2. Variables for Developing SPF

We assembled 16 explanatory variables as contributing risk factors for developing SPF: 7 categorical variables and 9 numerical variables, including the traffic volume, demographic characteristic, curve characteristics, roadway characteristics and spatial characteristics. The numerical variables were normalized and standardized. We use annual daily traffic volume (AADT) to represent traffic volume. Population serves as the demographic characteristic. Curve characteristics include curve radius, curve central angle, curve length, whether the curve is spiral, and whether the curve is on a ramp. Roadway characteristics consist of functional classification, speed, number of lanes, surface width, and median width. Spatial characteristics involve area type and the number of intersections on the curve. Table 1 provides an overview of categorical variables, while Table 2 outlines numerical variables. We checked the correlations among variables to avoid potential multicollinearity concerns in the regression models.

Table 1. Summary of categorical variables

Variable Type	Variable Name	Definition	Categories	Category Definition and Frequencies
Demographic characteristics	<i>Pop</i>	Population of the county	3	1 Large county (339); 2 Median county (326); 3 Small county (20)
Curve characteristics	<i>Spiral</i>	Whether the curve is spiral or not	2	N No (650); Y Yes (35)
	<i>Ramp</i>	Whether the curve is on a ramp or not	2	N No (681); Y Yes (4)
Roadway characteristics	<i>Func_class_diff</i>	Whether there is a changing in functional class or not	2	0 No (680); 1 Yes (5)
	<i>Speed_cat_diff</i>	Whether there is a changing in speed or not	2	0 No (657); 1 Yes (28)
	<i>Lane_cat_diff</i>	Whether there is a changing in number of lanes or not	2	0 No (663); 1 Yes (22)
Spatial characteristics	<i>Area_Type</i>	Area type	2	0 Rural (98); 1 Urban (587)

Table 2. Summary of numerical variables

Variable Type	Variable Name	Definition	Min	Max	Mean	SD
Traffic volume	<i>AADT_avg</i>	Logarithm of average of annual daily traffic volume (count)	6.62	11.81	9.45	0.81
Roadway characteristics	<i>Speed_avg</i>	Logarithm of speed limit (mile per hour)	3.00	4.25	3.76	0.25
	<i>Med_wid_avg</i>	Logarithm of median width (feet)	1.79	6.17	3.13	0.64
	<i>Lane_cnt_avg</i>	Logarithm of number of lanes (count)	1	6	2.39	0.79
	<i>Surf_wid_avg</i>	Logarithm of surface width (feet)	2.30	4.28	3.30	0.30
Curve characteristics	<i>Angle</i>	Logarithm of curve central angle (degree)	2.30	4.74	3.28	0.59
	<i>Radius</i>	Logarithm of curve radius (meter)	2.90	8.07	5.84	0.90
	<i>Curve_len</i>	Logarithm of curve length (meter)	3.19	7.14	5.31	0.71
Spatial characteristics	<i>Int</i>	Number of intersections on the curve (count)	0	5	0.68	0.85
Responding variable	<i>crash_all</i>	Crash frequencies (count)	0	89	4.09	8.46

3.3. SPF Performance Evaluation

This paper explored different variable selection methods for developing SPF. A total of 5 SPFs were developed using the methods from the GLM family, GAM and the tree-based method. We evaluated the SPF performance using the 10-fold cross-validation by obtaining average of the prediction errors for each SPF. The curve data was randomly and evenly divided into 10

folders. For each method, a looping algorithm was applied through 5 folders: for each folder, 1) the current folder was treated as the test set and the remaining 9 folders were used as the training set; 2) use the training set to develop the SPF; 3) use the resulting SPF with the test set and get the predicted values; 4) obtain the prediction errors from the test set. After the loop went through 10 folders, the average of the prediction errors was calculated. Here are the formulas for the room mean squared error (RMSE) and the mean absolute error (MAE): $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$; $MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$; where n is the sample size, y_i and \hat{y}_i are the observed and predicted value for the curve i.

RMSE and MAE were used as the merits indicating the SPF performance across the methods. The SPF with the smaller prediction error means the better performance it has. Within each method, besides the prediction error from cross-validation, other criterions such as AIC, BIC, Mallow's Cp and adjusted R² were also considered when selecting variables in SPFs.

The cumulative residual (CURE) plot is a graph of cumulative residuals against the fitted values sorted in ascending order. It allows for visual assessment of the model's fitness across the range of risk factors and can help detect potential issues with the SPF model. Ideally, the cumulative residuals should be close to zero and fluctuate within the boundaries of two standard deviations, indicating a well-fitting SPF (Lyon et al. 2018).

Figure 1 summarizes the steps in the study.

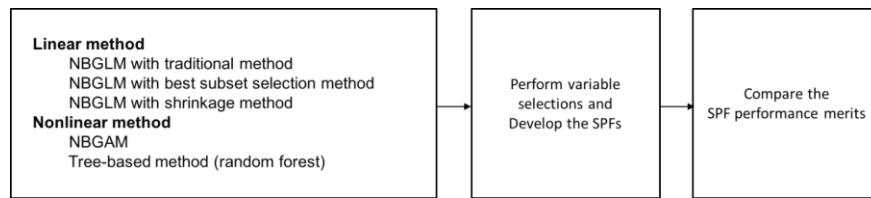


Figure 1. Methods for selecting variables and developing SPFs in the study

4. RESULTS

4.1. SPFs from NBGLM

4.1.1. NBGLM with Traditional Method

Using 16 variables to develop the SPF, 9 variables were showing the significant in the model: *Angle*, *Curve_len*, *Spiral*, *AADT_avg*, *Med_wid_avg*, *Func_class_diff*, *Speed_cat_diff*, *Area_Type*, and *Int*. The SPF with the selected variables yielded an RMSE of 9.29 and MAE of 4.03. This SPF was considered as the baseline condition for further comparison.

4.1.2. NBGLM with Best Subset Selection Method

To obtain better performance and avoid overfitting from the NBGLM, we performed the best subset selection method. The best subset selection went through every possible combination of all the explanatory variables and obtained the SPF with the cross-validated and other criterions including Cps, BIC and adjusted R². The cross-validation result showed that the SPF with 4 variables achieved the lowest prediction error (see Figure 2).

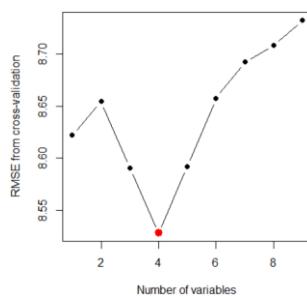


Figure 2. RMSE from cross-validation vs. number of variables in NBGLM with best subset selection

Only 4 variables, *Curve_len*, *Func_class.diff.f*, *Lane_cnt.avg*, and *Int* were selected in the SPF. The SPF with the selected variables yielded an RMSE of 8.64 and MAE of 4.00. The best subset selection method identified a significantly smaller set of variables compared to the traditional method, and it achieved a better SPF performance.

4.1.3. NBGLM with Shrinkage Method

When using cross-validation to find the best fitted λ in Equation 6, we found out the SPF didn't achieve an improvement in prediction error. After experimenting with several shrinkage methods, we obtained the SPF with adaptive Lasso method. The best fitted λ resulted in the SPF with 6 variables selected. Only *Curve_len*, *Med_wid_avg*, *Surf_wid_avg*, *Func_class_diff*, *Area_Type*, and *Int* were selected in the SPF. The SPF with the selected variables yielded an RMSE of 7.71 and MAE of 3.90. The shrinkage method produced a better performance than the SPF from NBGLM with the best subset selection.

4.2. SPFs from NBGAM

A list of variable selection scope was defined indicating which variables to explore for the modelling. An SPF with all the variables to the highest polynomial orders was obtained as the starting point of the variable selection procedure. A looping algorithm was applied: for each variable, fit a different SPF by changing the smooth function, calculate the performance criterion AIC. This algorithm continues unless the AIC of the SPF in this run cannot be smaller than the AIC from the previous SPF. After comparing the different order splines, the results from cross-validations indicated that the SPF with cubic splines had the best prediction performance. The best SPF from NBGAM had 8 variables selected, including *Radius*, *Angle*, *Curve_len*, *AADT_avg*, *Lane_cat_diff*, *Speed_avg*, *Lane_cnt_avg*, and *Int*. Among these variables, *Angle* was fitted in the quadratic spline; *Radius*, *Int* and *Curve_en* were fitted in the cubic splines. The SPF had a prediction error of RMSE 8.29 and MAE 3.40.

4.3. SPFs from Tree-Based Method (Random Forests)

Random Forests (RF) method is a vigorous tree-based method. All the explanatory variables were included in the SPF modelling in the beginning. From the report of the variable importance measurements, a subset of variables with larger %IncMSE (greater than 2 in Figure 3) were selected for the SPF. 9 variables were selected in this procedure and the SPF ended up with the

smallest prediction error among all the methods experimented in the study. *Radius*, *Angle*, *Spiral*, *Curve_len*, *AADT_avg*, *Speed_avg*, *Lane_cnt_avg*, *Surf_wid_avg*, and *Func_class_diff* were selected in the SPF. The SPF had a prediction error of RMSE 8.27 and MAE 3.38.

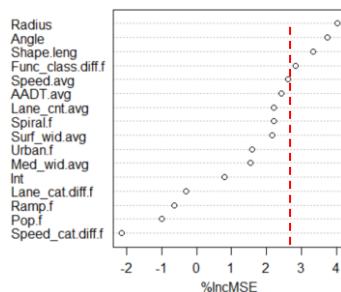


Figure 3. Variable importance in the SPF obtained by Random Forests

5. DISCUSSION

This study used different statistical methods to identify the contributing factors and develop SPFs for crashes on curves. By comparing the variables selected and the SPF performance obtained through different methods, we could evaluate the methods of developing SPFs and have better understanding of the safety issues related with curves.

Table 3 is the summary of the significant variables selected in the 5 SPFs from different statistical methods and Table 4 is the written formulas of these SPFs with coefficients.

Table 3. Summary of variable selected in the best performance SPFs from different statistical methods

Variable type	Variable name	NBGLM with traditional method	NBGLM with best subset selection	NBGLM with shrinkage method	NBGAM	Tree-based method (Random Forests)
Curve characteristics	<i>Angle</i>	+			+	+
	<i>Radius</i>				+	+
	<i>Spiral</i>	+				+
	<i>Curve_len</i>	+	+	+	+	+
Traffic volume	<i>AADT_avg</i>	+			+	+
Roadway characteristics	<i>Speed_avg</i>				+	+
	<i>Med_wid_avg</i>	+		+		
	<i>Lane_cnt_avg</i>		+		+	+
	<i>Surf_wid_avg</i>			+		+
	<i>Func_class_diff</i>	+	+	+		+
	<i>Speed_cat_diff</i>	+				
	<i>Lane_cat_diff</i>				+	
Spatial characteristics	<i>Area_Type</i>	+		+		
	<i>Int</i>	+	+	+	+	
Prediction error from cross-validation						
	Root mean squared error (RMSE)	9.29	8.64	7.71	8.29	8.27
	Mean absolute error (MAE)	4.03	4.00	3.90	3.40	3.38

In general, more complex methods tend to have smaller prediction errors (RMSE and MAE) compared to simpler methods. Specifically, the NBGLM with the shrinkage method, NBGAM, and Tree-based (Random Forests) method demonstrate superior prediction performance

compared to the NBGLM traditional method and NBGLM with best subset selection method. The explanatory variables such as the *Curve_len* appeared to be selected by all the SPFs. *Func_class_diff* and *Int* were selected by 4 SPFs. *Angle*, *AADT_avg*, and *Lane_cnt_avg* were selected by 3 SPFs. Surprisingly, in the NBGLM with best subset selection and shrinkage method, both methods exclude curve radius and curve central angle—the key variables that distinguish curve segments from roadway tangent segments, and AADT—the variable that should always be considered in SPFs, suggested in HSM. While *Curve_len* was consistently incorporated into the SPFs across all methods, *Radius* was exclusively included in SPFs derived from non-linear regressions, namely NBGAM and RF. Similarly, *Speed_avg* was only included in NBGAM and RF SPFs. This exclusion may be originated from concerns regarding model complexity, where the inclusion of too many variables could make the model unstable. Certain variables, including *Med_wid_avg*, *Surf_wid_avg*, *Speed_cat_diff*, *Lane_cat_diff*, and *Area_Type*, were selected in fewer than 3 SPFs. These variables are not typically prevalent in SPF analyses and were included in this study as factors that might have impacts but not be highly impactful. Therefore, their exclusion from most SPFs was anticipated.

Table 4. The SPFs developed from different statistical methods

Method	SPF with coefficients
NBGLM with traditional method	$\log(\mu_i) = 0.94 + 0.28 * I_{(Area_Type=1)} + 0.52 * I_{(Spiral=1)} + 1.85 * I_{(Func_class_diff=1)} + 0.38 * I_{(Speed_cat_diff=1)} + 0.24 * AADT - 0.20 * Med_wid_avg - 0.18 * Angle + 0.22 * Int - 0.22 * Curve_len$
NBGLM with best subset selection	$\log(\mu_i) = 7.63 + 1.36 * I_{(Func_class_diff=1)} + 1.10 * Lane_cnt_avg + 0.82 * Int - 1.30 * Curve_len$
NBGLM with shrinkage method	$\log(\mu_i) = -1.06 + 0.81 * I_{(Area_Type=1)} + 23.19 * I_{(Func_class_diff=1)} - 0.02 * Med_wid_avg + 1.99 * Surf_wid_avg + 0.09 * Int - 0.43 * Curve_len$
NBGAM	$\log(\mu_i) = 10.19 + 1.36 * I_{(Lane_cat_diff=1)} + 0.37 * AADT - 0.27 * Speed_avg + 1.07 * Lane_cnt_avg - 2.12 * s(Radius, df = 3) - 1.92 * s(Angle, df = 2) + 1.00 * s(Int, df = 3) + 1.27 * s(Curve_len, df = 3)$
Tree-based method (Random Forests)	No written formula with coefficients

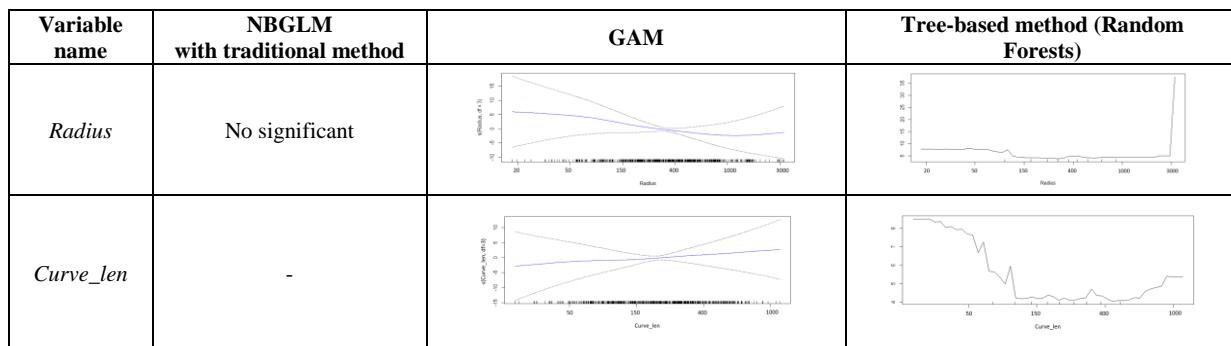
Table 5 shows the marginal effects of some selected variables in the SPFs. Specifically, we present the effects of radius and curve length, as the non-linear SPFs discovered the non-linear relationships between these factors and the crash frequencies. We selected these variables to have a close examination as they are the key variables of curve characteristics. Additionally, these variables show correlations, potentially suggesting multicollinearity issues.

The sample data reveals a correlation of approximately 0.7 between curve radius and curve length, suggesting that curves with larger radii may also have longer lengths. Consequently, all linear SPFs in the NBGLM methods exclude radius but include curve length. However, the non-linear methods such as GAM and RF have better capabilities in handling multicollinearity. Therefore, including all three variables in SPFs with these methods will not lead to an increase in prediction errors. Furthermore, these methods offer additional insights when interpreting the relationships among variables.

Based on SPF from GAM, for small curve radii, specifically those less than 1000 meters, crash frequencies decrease with increasing radius. Conversely, when the curve radius is larger than 1000 meters, crash frequencies show a slight increase with the radius. It is reasonable to comprehend that a small radius implies a sharp curve with a very tight curvature, indicating a high risk on the roadway. The RF model also identifies this initial decrease followed by an increase pattern.

Regarding the curve length, it shows a rising trend of impact on crashes in the GAM model. Even though the GAM model shows a smooth function in relation to curve length, the overall trend suggests an increase in crashes with the increasing curve length. However, the RF model shows an opposite impact, same as the findings from the NBGLM models. In the RF model, when the curve length is less than 150 meters, crash frequencies significantly decrease with length; between 150 and 400 meters, crashes slightly increase as length grows; and for lengths exceeding 400 meters, crashes significantly increase with length.

Table 5. Marginal effects of selected variables in the SPF developed from different statistical methods



In summary, although the linear SPF suggest an overall impact of variables, the non-linear SPF can reveal more detailed intricate relationships among variables, especially highlighting how the impact of variables may vary within specific ranges. To be consistent with the linear models, the non-linear SPF provide additional insights.

Table 6 presents the CURE plots with 95% confidence intervals (CI) for the cumulative residuals versus the fitted values of all 5 SPF models. The grey lines represent the cumulative residuals, the blue and red lines represent the lower and upper 95% CI levels. Four of the five models estimated have over 95% of the residuals within 95% CI levels. The CURE plots from the NBGLM using both the traditional method and the best subset selection method show the best overall fit. Additionally, CURE plots from the NBGLM with the shrinkage method and NBGAM show most points within the 95% CI, suggesting a generally appropriate fit. However, the CURE plot from the tree-based method (Random Forests) displays a majority of points outside the 95% CI. This suggests a systematic underestimation for predicted crash frequencies below 10 and an overestimation for predicted crash frequencies above 10. This suggests that the tree-based model may not be accurately capturing the relationships within the data, which is possibly due to the overfitting issue.

Table 7 summarized the advantages and disadvantages of the methods we explored in the study. It is intended to offer suggestions on selecting suitable statistical methods for SPF development. The traditional statistical method for the SPF was the NBGLM, applied in the HSM as well as in most literature. The negative binomial method provides an easy interpretation of how independent variables are linearly related to crash frequency and thus predicts the number of crashes. Various NBGLM methods may have similar limitations, particularly in detecting complex relationships among explanatory variables, especially when dealing with intricate crash data and associated variables. In contrast, NBGAM can be advantageous for developing SPF due to their capability to handle a large and complex dataset. When applying tree-based method,

despite its potential for superior prediction capability, it is essential to examine the CURE plot to detect systematic issues.

Table 6. CURE plots of SPF developed from different statistical methods

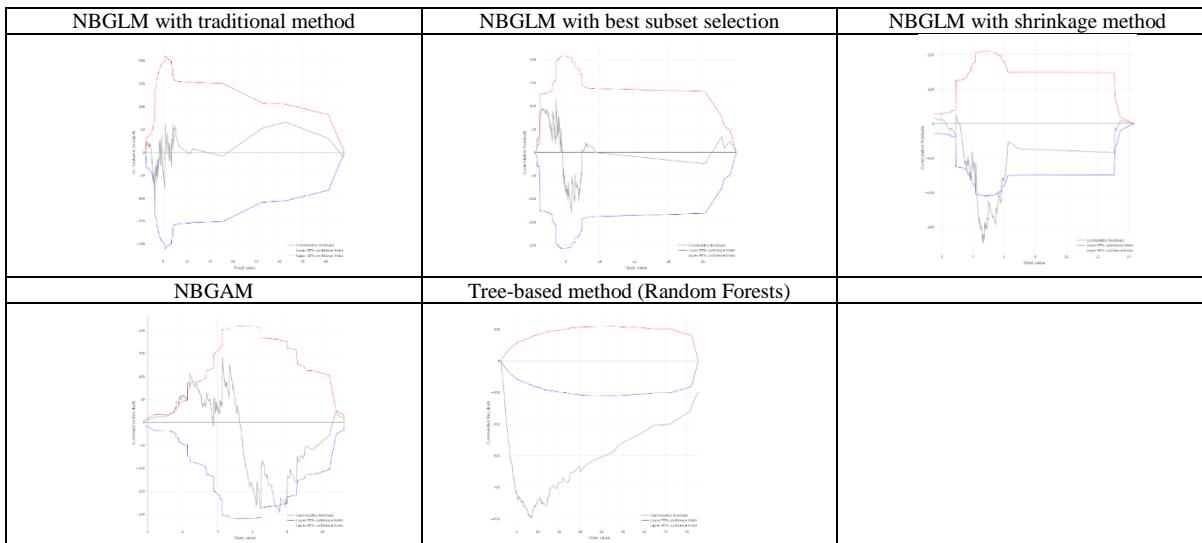


Table 7. Advantages and disadvantages of different methods of variable selection and developing SPFs

Method		Advantages	Disadvantages
Linear relationship	NBGLM with traditional method	Easy for understanding the linear relationships among explanatory variables; Stable model based on CURE plot	Lack of variable selection method; SPF prediction performance is poor when a large number of variables are included; Cannot detect the intricate relationships among explanatory variables
	NBGLM with best subset selection	Efficient variable selection method: Achieve a better SPF performance with very small amount of variables; Stable model based on CURE plot	May exclude key variables in the SPF; Cannot detect the intricate relationships among explanatory variables
	NBGLM with shrinkage method	Advanced variable selection method: Have the best SPF prediction performance among the linear methods	May exclude key variables in the SPF; Cannot detect the intricate relationships among explanatory variables
Nonlinear relationship	NBGAM	Flexible to handle the complex nonlinear relationships; Have better SPF prediction performance than the linear methods	Hard for variable interpretation
	Tree-based method (Random Forests)	Flexible to handle the complex nonlinear relationship; Have better SPF prediction performance than the linear methods	No written formula with coefficients; May have systematic issue based on CURE plot

6. CONCLUSION

This study explored various statistical methods for SPF development. Using a sample of horizontal curve data in Florida, the study analyzed the variable selection, methods' capabilities in handling statistical challenges. By analyzing the explanatory variables selected in SPF such as traffic volume, demographic characteristic, curve characteristics, roadway characteristics and spatial characteristics, and by comparing the SPFs using cross validation method, we concluded

that curve length, functional classification of roadway, and the relationship to the intersection are the highest contributing risk factors for curve safety. In general, the SPF generated by nonlinear methods (NBGAM and RF) displayed improved performance in prediction error compared to those from linear methods (NBGLM). The NBGLM with best subset selection and the shrinkage method also obtained better SPF than the traditional NBGLM. However, the tree-based method (RF) demonstrated instability based on CURE plot.

This is an important finding that can help improve the current practice that suggests the traditional NBGLM for developing SPF, which are only capable of a few contributing risk factors and can only interpret the linear relationships between the number of crashes with those factors. These findings contribute to the application of more advanced methods that can improve understanding of safety issues when detail-rich transportation data is available.

The main limitation of this study is that chose our study sample based on the data availability. Unlike the literature, we did not specify a particular facility type, which could lead to inconsistencies in the observed impact of variables compared to existing literature. For example, the curve length shows a negative impact on crashes in some SPFs, which contradicts the findings in the literature (Gayah and Donnell 2021; Gooch et al. 2018; Khan et al. 2012). This inconsistency might be attributed to a potential sampling issue. Our sample data has a high correlation between curve radius and curve length, preventing their inclusion in the same SPF in many methods. However, the abovementioned literature did include both variables within the same SPF. Nevertheless, our primary focus was to compare different statistical methods to assess their abilities to explain the relationships among variables using the same sampled data as an example. Further exploration of the regression techniques may be needed to better understand and resolve these inconsistencies.

Despite the limitation, this study provides statistical support for transportation engineers and planners to screen the roadway network using SPFs and identify the most high-risk locations for curve improvements. Findings in this study can guide to a better choice of SPF methods, which may lead to a more accurate and interpretable transportation safety analysis.

REFERENCES

- AASHTO. 2010. *Highway safety manual*. Washington DC.: American Association of State Highway and Transportation Officials.
- Albin, R., et al. 2016. *Low-Cost Treatments for Horizontal Curve Safety 2016*. Washington, D.C.
- Bejleri, I., X. Xu, D. Brown, S. Srinivasan, and N. Agarwal. 2021. "Automatic Horizontal Curve Identification for Large Areas from Geographic Information System Roadway Centerlines." *Transp. Res. Rec. J. Transp. Res. Board*, 2675 (12): 1088–1105. Sage Publications, Incorporated.
- Donnell, E., V. Gayah, J. Paul, and T. D. Larson. 2014. *Safety Performance Functions*.
- FHWA. 2015. "Highway Statistics 2015." Highw. Stat. 2015.
<https://www.fhwa.dot.gov/policyinformation/statistics/2015/>.
- Gayah, V. V., and E. T. Donnell. 2021. "Estimating safety performance functions for two-lane rural roads using an alternative functional form for traffic volume." *Accid. Anal. Prev.*, 157 (March): 106173. Elsevier Ltd.
- Gooch, J. P., V. V. Gayah, and E. T. Donnell. 2018. "Safety performance functions for horizontal curves and tangents on two lane, two way rural roads." 120 (June): 28–37.

- Hastie, T. 2023. "gam: Generalized Additive Models." R Packag. version 1.22-2. <https://cran.r-project.org/package=gam>.
- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. Springer.
- Khan, G., A. R. Bill, M. Chitturi, and D. A. Noyce. 2012. "Horizontal Curves, Signs, and Safety." (2279): 124–131.
- Liaw, A., and M. Wiener. 2002. "Classification and Regression by randomForest." R News. <https://cran.r-project.org/doc/Rnews/>.
- Lord, D., S. P. Washington, and J. N. Ivan. 2005. "Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory." *Accid. Anal. Prev.*, 37 (1): 35–46.
- Lyon, C., B. Persaud, and F. Gross. 2018. *The Calibrator - An SPF Calibration and Assessment Tool - Updated User Guide*. Washington, DC.
- McGee, H. W., F. R. Hanscom, I. Vanasse Hangen Brustlin, T. R. Corporation, and F. H. Administration. 2006. *Low-Cost Treatments for Horizontal Curve Safety*.
- Miaou, S.-P., and H. Lum. 1993. "Modeling vehicle accidents and highway geometric design relationships." *Accid. Anal. Prev.*, 25 (6): 689–709.
- Preston, H., R. Storm, J. D. Bennett, and B. Wemple. 2013. *Systemic Safety Project Selection Tool*. Wahsington, D.C.
- Srinivasan, R., D. Carter, and K. Bauer. 2013. "Safety Performance Function Decision Guide: SPF Calibration versus SPF Development." (September): 1–31.
- Wang, K., J. K. Simandl, M. D. Porter, A. J. Graettinger, and R. K. Smith. 2016. "How the choice of safety performance function affects the identification of important crash prediction variables." *Accid. Anal. Prev.*, 88: 1–8. Elsevier Ltd.
- Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Gen. Addit. Model. An Introd. with R, (J. K. Blitzstein, J. J. Faraway, M. Tanner, and J. Zidek, eds.). CHAPMAN & HALL/CRC Press.
- Xie, Y., and Y. Zhang. 2008. "Crash Frequency Analysis with Generalized Additive Models." *Transp. Res. Rec. J. Transp. Res. Board*, 39–45.
- Zou, H., and T. Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, 67 (2): 301–320.

Motorcycle Crashes in California: Analysis of Crash Severity and Contributing Factors

Amal Mamlouk¹; Ashkan Teymouri, Ph.D.²; and Masoud Ghodrat Abadi, Ph.D.³

¹Dept. of Civil Engineering, California State Univ., Sacramento, Sacramento, CA.
Email: amalmamlouk@csus.edu

²Dept. of Civil and Environmental Engineering, Univ. of California, Davis, Davis, CA.
Email: aknteymouri@ucdavis.edu

³Dept. of Civil Engineering, California State Univ., Sacramento, Sacramento, CA.
Email: abadi@csus.edu

ABSTRACT

This study delves into the causes of motorcycle crashes in California by analyzing data from the California Statewide Integrated Traffic Records System (SWITRS), comprising 111,282 motorcycle crashes between 2012 and 2022. Analysis of main violations reported for motorcycle crashes indicated the need to focus on deterring riding under the influence, managing speed, promoting safe turning and lane-changing practices, and addressing traffic signal compliance and improper passing. The study also explored the relationship between collision types and crash severity, emphasizing the significance of addressing issues related to broadside and head-on collisions in fatal crashes, overturning incidents in injury-related crashes, and sideswipe and rear-end collisions in property damage only. Utilizing an ordered logit model with random coefficient decomposition, the research identified statistically significant variables affecting crash severity. High-severity outcomes were associated with crashes occurring on highways, non-compliance with laws, collisions with other motor vehicles, and wet road surfaces. Age and gender further influenced the severity of crashes. Findings can collectively contribute to reducing the frequency and severity of motorcycle crashes, making the roads safer for motorcyclists and other road users alike.

INTRODUCTION

According to the National Highway Traffic Safety Administration, 5,579 motorcyclists were killed on US roads in 2020, constituting 14 percent of all traffic fatalities and the highest number of motorcyclists killed since the Fatality Analysis Reporting System (FARS) started operating in 1975 (NHTSA, 2022a). In recent years, California has consistently ranked as having the highest number of motorcycle fatalities and injuries. The majority of the motorcycle industry operates in California, and California has the highest number of registered motorcycles in the US. According to the 2020 Traffic Safety Fact Sheet, California had over 3.3 million registered motorcycles (11% of total registered vehicles). Florida followed with more than 2.4 million registered motorcycles, and Texas came next with more than 2 million (NHTSA, 2022b). Motorcycle crashes in California affect not only the riders but other motorists, pedestrians, and vulnerable road users. The number of traffic and motorcyclist fatalities obtained from FARS shows a trend in motorcycle crashes in California in recent years (Figure 1). The rate of motorcyclist fatalities has remained relatively stable while the total traffic fatalities have increased. Between 2012 and 2020, a total of 4,799 motorcyclist fatalities were recorded in California, compared to 5,169 fatalities in Florida and 4,236 in Texas (FARS, 2023). The motorcycle fatality dataset highlights

the urgent need to study motorcycle crashes in California to understand the contributing factors and develop effective interventions to reduce their occurrence and severity.

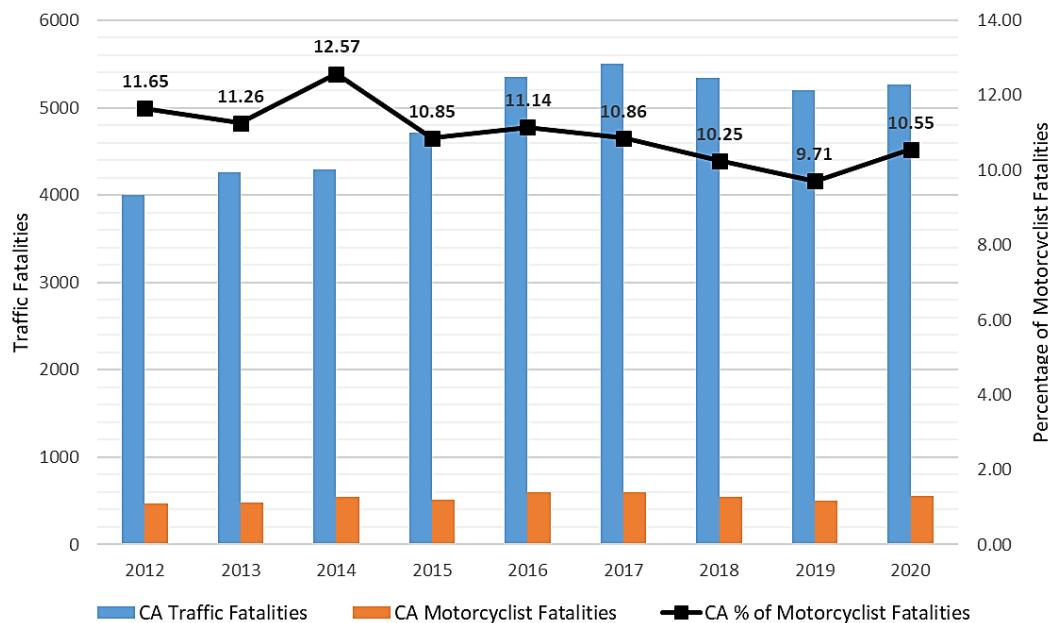


FIGURE 1. Traffic Fatalities and Motorcyclist Fatalities in California

Given the importance of motorcycles as a mode of transportation and the significance of ensuring it is accessible, safe, and easy to use, this study aimed to evaluate factors impacting the frequency and severity of motorcycle crashes in California. This objective was achieved through a detailed analysis of determinants such as temporal distributions, weather conditions, location (e.g., intersection vs. non-intersection), road surface quality, lighting condition, and types of collision associated with motorcycle crashes. The study focused on the causes of motorcycle crashes in California by investigating the California Statewide Integrated Traffic Records System (SWITRS), an electronic database of police-reported traffic collisions operated by the California Highway Patrol (CHP). The study collected data for motorcycle crashes that occurred in California between the years 2012 and 2022. The goal of this study is to identify the factors that caused an increase or decrease in motorcycle crashes using descriptive and inferential statistics.

LITERATURE REVIEW

Motorcycles have established widespread usage worldwide, serving various purposes across different regions. Notably, Asia stands out as the continent where motorcycle usage is particularly prevalent. Consequently, a significant portion of the existing literature on motorcycles originates from Asian developing countries, including Taiwan, Thailand, Vietnam, Malaysia, Macao and China (e.g., Hsu et al., 2003; Pongthanaisawan & Sorapipatana, 2010; Wong, 2013). However, it is essential to acknowledge that this literature may not directly translate to the U.S. context due to contrasting patterns of motorcycle usage in the two regions. In developing countries, motorcycles are predominantly employed as a regular and practical means of transportation for daily commutes, encompassing travel to educational institutions, workplaces, and routine destinations (Davoodi & Hossayni, 2015). Conversely, in the U.S.,

motorcycles are primarily used for recreational purposes, leading to a distinct focus on research areas. For example, a study from Florida tabulated the survey results of motorcyclists behind their reason for riding a motorcycle and found that 46.3% rode for the thrill of riding and 23.1% rode it leisurely as an expensive mode of transportation (Lee et al., 2013). This study also reported that the reason for the attraction towards motorcycles differs significantly by age.

Most U.S. states require rider training prior to licensure for at least some motorcyclists; six states and D.C. require training for all novice motorcyclists. A study conducted by Goodwin et al. (2022) showed that between 1991 and 2018, there were 103,142 younger novice riders and 98,540 older novice/returning motorcycle riders in North Carolina. This study investigated different factors, such as age, riding experience, speeding, and the presence of a work zone, and examined crash rates of novice motorcycle riders for both younger and older motorcyclists. It found that the overall driving experience had a smaller effect on the fatality rate increase in motorcycle crashes; however, the combination of experience and age among motorcyclists indicates that experience does matter. A Wyoming case study conducted by Rezapour et al. (2020) focused on predicting the severity of outcomes of motorcycle collisions. In 2017, fatal motorcycle crashes constituted 16.19% of the fatal crashes in Wyoming; however, motorcycle crashes only constituted 1.1% of the total fatal crashes. Rezapour et al. (2020) found that Some environmental factors in Wyoming, such as mountainous topography and winter weather conditions, increased the risk of collisions.

Safaei et al. (2021) investigated the relationship between fatal traffic crashes and the increase in gasoline prices in the U.S. between 2007 and 2016. The study used 6,121 observations, and it was found that fatal crash density was higher in rural areas compared to urban areas. There was a correlation between increased gas prices and motorcycle fatalities, as a one-dollar increase in gasoline price increased fatal motorcycle crashes by 24.4%. Riders' gender and transportation modes also contributed to a relationship between gasoline prices and fatal crashes.

A study conducted in Miami-Dade County in Florida (Sarmiento et al., 2020) investigated the impact of alcohol and illegal substances on motorcycle crashes by identifying the relationship between toxicology and riders' behavior. According to the data collected between 2009 and 2014 by the County Medical Examiner Office's toxicology reports and the corresponding crash reports, 44% of crash fatalities resulted from alcohol/illegal substance use. While the data showed that intoxicated individuals were less likely to wear helmets, no significant relationship was observed between speeding and alcohol use.

Nazemetz et al. (2019) investigated 500 crashes in Orange County, CA, over a period of 20 years (1994 to 2014) with the goal of identifying the reasons behind the high rate of motorcycle crashes in the area. It was found that about 12% of the crashes included fatalities and usually happened at night. Additionally, most motorcycle crashes happened on Sundays, reflecting the recreational use of motorcycles in the U.S.

A study conducted by Wali et al. (2021) focused on comparing riders' anatomical injuries to identify and predict mortality in motorcycle crashes. The findings revealed that about 9% of riders experienced fatal injuries, 45% experienced injuries, and 36% of collisions resulted in hospitalizations and/or disabilities. Researchers found that injury severity rate helped in the analysis of predicting mortality risk.

Chawla et al. (2019) investigated data from the Federal Highway Administration (FHWA) to identify the risks of fatality in motorcycle crashes in Orange County, California. Riders' age, physical status, and educational level were listed as non-random risk factors, and other variables, such as motorcycle type, helmet coverage, motorcycle ownership, speed, trip destination, and

traffic violation history, were identified as random factors. The study found that lower travel speed increased the risk of motorcycle collision. Furthermore, the study found that trips to and from work and home increased the probability of collisions versus trips made for leisure.

A study by Das et al. (2021) investigated 5,172 motorcyclists killed in traffic crashes between 2001 and 2015, representing about 17% of all traffic fatalities. This study found that 94% of motorcycle collisions resulted from human errors, with the rest resulting from environmental and vehicle-related factors. The study found that tire inflation levels, brake failures, the crossing of dedicated left lanes, passing solid yellow lines, speeding over the posted limit, and crashes on right-sided curves were all high-risk factors in fatal outcomes. Researchers noted that most “conventional” crash prevention strategies focus on larger commercial-size vehicles, not considering smaller vehicles and two-wheeled bikes, even though they account for more than half of motorized vehicle fatalities.

With a detailed understanding of the past studies conducted in the U.S. about motorcyclist safety, the present study contributes to the literature by utilizing a comprehensive dataset from motorcycle Crashes that occurred in California from 2012 to 2022. We use descriptive statistics to explore common trends in motorcycle crashes and use advanced statistical modeling to identify and explain underlying factors.

DATA

The primary objective of this research was to investigate the causes of motorcycle crashes in California. To achieve this goal, the study leveraged data from the California Statewide Integrated Traffic Records System (SWITRS), an electronic database managed by the California Highway Patrol (CHP) that records police-reported traffic collisions (CHP, 2023). The dataset covered motorcycle crashes that occurred within the state of California during the extensive period from 2012 to 2022, providing a robust and comprehensive foundation for the analysis. The researchers obtained three distinct datasets from SWITRS to conduct their investigation. The first dataset, known as Collision raw data, contained detailed information about each traffic collision, including factors such as location, date, time, and collision severity. The second dataset, Party raw data, comprised information about the parties involved in each collision, such as drivers, passengers, and pedestrians, along with their associated attributes. The third dataset, Victim raw data, captured specific details about the victims of the crashes, such as their demographics and injury severity.

Before conducting the analysis, an essential step in data preparation was undertaken, which involved extensive data reduction and cleansing. The purpose of this process was to ensure that the final dataset used for analysis was accurate, reliable, and free from inconsistencies or errors. These measures were taken to improve the quality and validity of the findings derived from the dataset. To create a unified dataset for analysis, the three raw datasets were merged using a unique case ID assigned to each observation. This unique identifier facilitated the integration of relevant information from the Collision, Party, and Victim datasets, allowing for a comprehensive examination of motorcycle crashes and associated factors. Subsequently, the researchers filtered the merged dataset to include only those crashes that involved at least one motorcyclist.

The final dataset comprised a total of 111,282 motorcycle crashes, providing a comprehensive basis for analyzing various characteristics related to these incidents. The data revealed that the majority of motorcyclists involved in the crashes were male, accounting for approximately 91.9% of the total cases. On the other hand, females constituted 7.4% of the cases.

The remaining percentage accounted for individuals whose gender identity fell outside the binary categories of male and female. Furthermore, the data indicated that the largest racial group among the motorcyclists involved in the crashes was white, making up 54.2% of the cases. Hispanic motorcyclists accounted for 25.8% of the incidents, while black motorcyclists constituted 8.0%. Additionally, Asian motorcyclists were involved in 3.7% of the crashes. The remainder of the cases represented individuals from other racial backgrounds.

An important aspect investigated in the study was the collision severity of motorcycle crashes. The severity was categorized into five levels: 1.1% of the incidents resulted in property damage only (PDO), 27.6% involved injuries with complaints of pain, 45.4% led to injuries with visible signs, 21.6% resulted in severe injuries, and 4.3% were fatal crashes. This distribution of collision severity was instrumental in understanding the overall impact and consequences of motorcycle crashes, helping to identify areas that require specific attention and interventions to improve road safety for motorcyclists.

ANALYSIS

Contributing Factors

The data on the percentage of motorcycle crashes against the day of the week revealed interesting patterns in crash occurrences. It is evident that the frequency of motorcycle crashes increases significantly over the end of the week, with Saturday, Friday, and Sunday accounting for the highest percentage of crashes at 16.8%, 15.8%, and 15.5%, respectively. This observation aligns with the hypothesis that weekends see higher crash rates due to a predominant use of motorcycles for leisure travel. During weekends, riders often take their motorcycles for recreational rides, exploring scenic routes or participating in group rides, which may increase the exposure to potential risks on the road. Conversely, weekdays exhibit relatively lower crash rates, with the percentage of crashes gradually rising from Monday to Thursday (11.5% to 14.0%), possibly indicating a higher proportion of motorcycle usage for commuting and work-related purposes on these days.

Regarding the location of motorcycle crashes, the analysis indicated that a substantial portion of crashes, 78.1%, occurred at non-intersection locations. This emphasizes the importance of understanding crash dynamics on various road segments, including highways, ramps, and loops, where the majority of crashes happen. These findings underscore the significance of targeted safety measures on these specific stretches to reduce the occurrence of motorcycle crashes. The remaining 21.6% of crashes at intersections warrant a closer examination of factors such as traffic signal timings, visibility, and intersection design to develop effective strategies for crash prevention at these critical points.

Weather conditions play a significant role in motorcycle safety. The data revealed that the vast majority of crashes, 91.1%, occurred during clear weather, with only 7.6% occurring during cloudy weather. While these results may not come as a surprise, they highlight the need for heightened caution during clear weather when motorcyclists are more likely to be on the roads. It is worth mentioning factors such as increased traffic volume or rider behavior during clear weather that may contribute to the higher crash rate during such conditions.

The analysis of motorcycle crashes based on light conditions revealed that the majority, 73.3%, happened during daylight hours, which aligns with typical riding hours. However, a noteworthy 20.2% of crashes occurred during nighttime, with 16.1% of them happening under dark conditions with street lights. This highlights the importance of street lighting in enhancing

visibility for both motorcyclists and other road users. Moreover, 6.2% of crashes occurring in the dark with no street lights and 3.9% at dusk (dawn) necessitate further investigation into the contributing factors, such as insufficient visibility and potential visibility-related hazards during these periods. Implementing improved street lighting and promoting the use of reflective gear during low-light conditions could help mitigate the risk of crashes during these times.

The analysis of main violations reported for motorcycle crashes in California provided valuable insights into the contributing factors behind these crashes, as well as how the severity of crashes relates to specific violations. Table 1 illustrates the frequency of reported causes for all crashes, with unsafe speed, improper turning, and automobile right of way being the most commonly cited violations, accounting for 30.6%, 19.4%, and 15.6% of crashes, respectively. However, the study revealed that the distribution of these violations varies significantly depending on the severity of the crashes.

TABLE 1. Reported Violations against Crash Severity

Violation	Total%	PDO	Injury (Complaint of Pain)	Injury (Other Visible)	Injury (Severe)	Fatal
Unsafe Speed	30.6%	29.7%	29.8%	31.3%	30.4%	29.9%
Improper Turning	19.4%	13.7%	17.5%	20.5%	20.3%	16.4%
Automobile Right of Way	15.6%	12.1%	17.0%	14.3%	16.8%	16.1%
Unsafe Lane Change	11.3%	13.1%	14.8%	12.1%	6.7%	2.4%
Under the Influence of Alcohol	5.3%	7.0%	2.1%	4.5%	8.3%	17.8%
Traffic Signals and Signs	3.3%	5.4%	3.7%	2.8%	3.7%	4.8%
Wrong Side of Road	2.6%	1.9%	2.0%	2.1%	3.8%	5.8%
Improper Passing	2.3%	3.0%	2.6%	2.4%	2.0%	1.2%
Other	9.6%	14.1%	10.5%	10.0%	8.0%	5.6%

For fatal crashes (compared to other severity levels), two specific violations were overrepresented: riding under the influence of alcohol or drugs and being on the wrong side of the road. Regarding injury crashes with visible signs, two violations were found to be overrepresented: unsafe speed and improper turning. These results suggest that managing speed and promoting safe turning practices could significantly reduce the number of injury-related motorcycle crashes. In injury crashes where complaints of pain were reported, automobile right of way and unsafe lane change were overrepresented. Finally, for PDO crashes, traffic signals and signs, and improper passing were overrepresented. While PDO crashes may not result in injuries or fatalities, they can still lead to property damage and inconvenience for all parties involved.

Table 2 illustrates the frequency of reported collision types, with broadside, overturned, and sideswipe being the most commonly cited, accounting for 21.3%, 20.2%, and 18.9% of crashes, respectively. However, the study revealed that the distribution of collision types varies significantly depending on the severity of the crashes. For fatal crashes, three specific collision types were overrepresented: broadside, hit-object, and head-on collisions. The prevalence of broadside collisions in fatal crashes might be related to the severity of impact in such scenarios, where the front of the motorcycle is directly hit on its side, leaving the rider exposed to significant injuries. The observation that hit object collisions were overrepresented in fatal crashes aligns with previously observed riding rates under the influence and dark conditions.

Impaired riding or reduced visibility during dark conditions could lead to a higher probability of motorcyclists colliding with objects on the road, resulting in fatal outcomes. Additionally, the overrepresentation of head-on collisions in fatal crashes could be related to incidents where riders were on the wrong side of the road, significantly increasing the risk of head-on impacts with other vehicles. Overturned collisions were found to be overrepresented in crashes classified as an injury with visible signs. Finally, sideswipe and rear-end collisions were overrepresented in PDO crashes. Sideswipe collisions typically involve the lateral impact of the motorcycle, often resulting in damage to the vehicle's side panels, while rear-end collisions occur when a vehicle hits the back of the motorcycle. While these types of crashes may not result in severe injuries, they can still lead to property damage and financial burdens for the parties involved.

TABLE 2. Type of Collision against Crash Severity

Type of Collision	Total%	PDO	Injury (Complaint of Pain)	Injury (Other Visible)	Injury (Severe)	Fatal
Broadside	21.3%	20.9%	23.0%	18.3%	23.9%	29.2%
Overturned	20.2%	3.4%	14.8%	24.0%	20.9%	15.4%
Sideswipe	18.9%	30.1%	23.6%	19.5%	13.4%	7.6%
Rear End	17.6%	30.8%	21.6%	16.5%	15.2%	12.8%
Hit Object	11.7%	7.4%	7.4%	11.4%	15.7%	22.8%
Head-On	4.1%	4.3%	3.6%	3.1%	6.0%	9.2%
Other	6.2%	3.1%	6.0%	7.2%	4.9%	3.0%

Model Specifications

In addition to exploring descriptive analysis of contributing factors, this study analyzed the severity of motorcycle crashes (response variable) using inferential statistics. Given the ordinal and discrete nature of crash severity, which is categorized into five categories (PDO, injury (complaint of pain), injury (other visible), injury (severe), and fatality), an ordered logit model has been employed. Furthermore, since the objective of this analysis was to explore heterogeneity in the influential variables affecting crash severity and to identify the source of this heterogeneity, an ordered logit model with random coefficient decomposition of mean has been calibrated (Eluru et al., 2008; Mannering et al., 2016).

Table 3 shows the statistically significant variables in the final model. The mixed logit model with heterogeneity in means has been estimated using 200 Halton draws. This approach is especially suitable for calculating multidimensional integrals or conducting Monte Carlo simulations. To examine whether the random coefficients possess a stochastic component, five distributions, namely normal, log-normal, uniform, triangular, and Weibull, have been investigated for all of them. Additionally, the presence of any heterogeneity source is explored through demographic variables.

Model results showed that if the motorcycle crash occurs on a highway rather than at a ramp or intersection, there is a higher probability of the crash resulting in more severe outcomes (as indicated by the positive coefficient of the variable "lotype1"). This could be attributed to the fact that motorcyclists on highways generally have more maneuverability compared to intersections and ramps, and they often ride at higher speeds.

It was further found that non-compliance with laws, such as alcohol consumption, unsafe lane change, and failure to yield, increases the probability of crashes. In the calibrated model, other variables related to rules of the road, such as uncertain speed, unsafe maneuvers, and inappropriate lighting, were not found statistically significant and, therefore, were not included. Among the three significant variables—alcohol use, unsafe lane change, and failure to yield—it was found that failure to yield combined with alcohol consumption leads to a higher probability of more severe crashes compared to the variable of unsafe lane change. This is in line with the statistics observed in the previous research. Therefore, it is recommended to provide better education to motorcyclists regarding yielding right of way and to implement preventive and deterrent laws related to alcohol consumption during motorcycling. Another suggestion arising from this analysis is the active review of driving regulations concerning right of way, accompanied by receiving feedback on related questions in the motor vehicle code. Additionally, the proposal for one-day free workshops with comprehensive training, including an explanation of the consequences and dangers of alcohol use while riding motorcycles, might present a practical option to lower the probability of motorcycle crashes.

TABLE 3. Mixed Ordered Logit with Decomposition of Mean Model

Variable	Description	Coefficient	p-value	Type	Source of Heterogeneity
C	Constant	4.369	0.000	Fixed	-
Lototype1	Crash occurred on a highway (1); Otherwise (0)	0.111	0.000	Fixed	-
Vcat4	Rider under alcohol influence (1); Otherwise (0)	0.366	0.000	Fixed	-
Vcat7	Unsafe lane change (1); Otherwise (0)	0.218	0.000	Fixed	-
Vcat21	Failure to yield (1); Otherwise (0)	0.394	0.000	Fixed	-
With3	Motorcycle collided with a motor vehicle (1); Otherwise (0)	Mean: 0.249	0.000	Random (Normal)	V_Gender*: 0.069 (p-value < 0.01)
		Standard Deviation: 0.632	0.000		V_age**: -0.00025 (p-value < 0.01)
		Mean: 0.341	0.000		N/A
Rsur2	Wet surface (1); Otherwise (0)	Standard Deviation: 0.425	0.000	Random (Normal)	
Rsur4	Slippery surface (oily or muddy) (1); Otherwise (0)	0.242	0.016	Fixed	-
Vequ4	Victim did not use lap belt (1); Otherwise (0)	0.990	0.067	Fixed	-
Vequ2	Victim did not use air bag (1); Otherwise (0)	0.098	0.000	Fixed	-
LL(β)		-136256.3858			
LL(0)		-179101.4698			
$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)}$		0.24			
$\chi^2_{N,1-\alpha}$		70.065			
$-2(LL(0) - LL(\beta)) > \chi^2_{N,1-\alpha}$		Yes			
Number of observations		111,282			
* Victim is a male (1); Otherwise (0)					
** Age of the victim					

If the collision of a motorcyclist involves other motor vehicles (compared to incidents involving pedestrians, trains, animals, bicyclists, etc.), the probability of a more severe crash is higher. Additionally, it is observed that this variable follows a random distribution, indicating that the severity of motorcycle crashes when colliding with other motor vehicles is not fixed. The coefficient for this variable follows a normal distribution (with a mean of 0.24886 and a standard deviation of 0.63175). The investigation of heterogeneity sources reveals that the heterogeneity in the coefficient of this variable is influenced by two factors: gender and age. Examining the coefficient related to age's impact on the mean of the normal distribution (-0.00025) indicates that as age increases, if a motorcyclist collides with other motor vehicles, the probability of more severe crashes decreases. This can be interpreted as increased caution among individuals during motorcycling as they grow older, resulting in lower crash severity when colliding with other motor vehicles compared to when younger individuals are involved in such crashes. Moreover, the examination of the gender coefficient (0.06922) suggests that if a male motorcyclist is involved in a collision with other motor vehicles, the probability of a more severe crash is higher. This may be attributed to men tending to ride more aggressively compared to women, who often ride more cautiously. Hence, it is recommended to provide additional education to young and/or male individuals to encourage more cautious motorcycling practices.

If the road surface is wet or slippery, the probability of a more severe crash is higher. Compared to the road being slippery, a wet road surface increases the probability of more severe crashes (based on the estimated coefficient in the calibrated model). Additionally, it is observed that the coefficient for the variable of a wet road surface follows a random distribution with a normal distribution, having a mean of 0.34090 and a standard deviation of 0.42520. This indicates presence of heterogeneity. In this study, all available variables in the data were used to identify the source of this heterogeneity, but none of the variables demonstrated statistical significance. This means that the source of the observed heterogeneity is not related to factors like socioeconomic variables, gender, or different age groups. In other words, there is no significant difference in crash severity between men and women or individuals of different age groups when crashes occur on wet roads. Based on these findings, it is recommended to implement appropriate road maintenance and cleaning measures to prevent crashes with more severe outcomes. Ensuring proper drainage and removing any contaminants from the road surface can contribute to reducing the occurrence of high-severity crashes.

According to the results of the calibrated model, it is observed that if the victim did not use a lap belt or an airbag, the probability of a crash with high severity increases. This could be due to the victim being thrown off and sustaining injuries as a result. Therefore, it is recommended to implement and enforce stricter measures to ensure the use of these safety devices and monitor their compliance effectively.

One of the most fundamental validation tests for the calibrated model is examining the magnitude and signs of the calculated coefficients. Typically, in model validation, there is an expectation regarding the signs and relative magnitudes of the coefficients. As discussed in this section, all the calculated coefficients exhibit signs and relative magnitudes that align with the expected outcomes. Additionally, to assess the statistical significance of the coefficients in the calibrated model, p-values have been utilized. As evident in Table 3, all the coefficients are statistically significant, with a confidence level exceeding 93%. To assess the goodness of fit of the model, ρ^2 is used. According to table 3, this index has an acceptance range. The probability ratio test is employed to determine the statistical significance of the difference between $LL(\beta)$

and $LL(0)$, using the chi-square (χ^2) distribution. According to Table 3, the difference between these two values in the final model of this study is statistically significant.

SUMMARY AND CONCLUSION

The analysis of 111,282 motorcycle crashes in California that happened between 2012 and 2022 provided valuable insights into the patterns and factors influencing crash occurrences. This study also provided a comprehensive analysis of the main violations reported for motorcycle crashes in California, shedding light on the varying patterns of violations based on crash severity. Implementing measures to deter riding under the influence, managing speed, promoting safe turning and lane-changing practices, and addressing traffic signal compliance and improper passing can collectively contribute to reducing the frequency and severity of motorcycle crashes, making the roads safer for all users. An analysis of the main types of collision reported for motorcycle crashes in California offered insights into the relationship between collision types and crash severity. Addressing issues related to broadside and head-on collisions in fatal crashes, overturning incidents in injury-related crashes, and sideswipe and rear-end collisions in PDO crashes can collectively contribute to reducing the frequency and severity of motorcycle crashes, making the roads safer for motorcyclists and other road users alike.

This study employed both descriptive and inferential statistics to analyze the severity of motorcycle crashes and explore the influential factors affecting crash outcomes. The use of an ordered logit model with random coefficient decomposition allowed for a comprehensive examination of heterogeneity in the influential variables. Several statistically significant variables were identified in the final model, providing valuable insights into the factors influencing crash severity. Notably, crashes occurring on highways were associated with a higher probability of more severe outcomes, likely due to the increased maneuverability and higher speeds of motorcyclists in such settings. Non-compliance with laws, particularly alcohol consumption, unsafe lane changes, and failure to yield, also contributed to higher crash severity. To address these issues, this study recommends implementing educational programs on yielding right of way and preventive measures against alcohol consumption while riding motorcycles. Collisions with other motor vehicles and wet road surfaces were found to increase the probability of severe crashes, with gender and age influencing the heterogeneity in the coefficient of the former. The results suggest the need for cautionary education for young and male motorcyclists. Lastly, promoting the use of lap belts and airbags, as well as road maintenance and cleaning measures, is essential to mitigate the severity of motorcycle crashes. By implementing these recommendations, it is possible to enhance motorcycle safety and reduce the occurrence of high-severity crashes.

REFERENCES

- Chawla, H., Karaca, I., and Savolainen, P. T. Contrasting Crash- and Non-Crash-Involved Riders: Analysis of Data from the Motorcycle Crash Causation Study. *Transportation Research Record: Journal of the Transportation Research Board*. 2673 (7). 2019.
- CHP (California Highway Patrol). The Statewide Integrated Traffic Records System (SWITRS). <https://iswitrss.chp.ca.gov/Reports/jsp/reportSelect.do> Accessed 7/15/23.

- Das, S., Dutta, A., and Tsapakis, I. Topic models from crash narrative reports of motorcycle crash causation study. *Transportation Research Record: Journal of the Transportation Research Board*. 2675(9), 449–462. 2021.
- Davoodi, S. R., and Hossayni, S. M. Role of motorcycle running lights in reducing motorcycle crashes during daytime. A review of the current literature. *Bulletin of Emergency & Trauma*. 3(3):73-8. 2015.
- Eluru, N., Bhat, C. R., and Hensher, D. A. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, 40(3), 1033-1054. 2008.
- FARS (Fatality Analysis Reporting System). <https://www-fars.nhtsa.dot.gov/Main/index.aspx> Accessed 7/15/2023.
- Goodwin, A. H., Wang, Y. C., Foss, R. D., and Kirley, B. The role of inexperience in motorcycle crashes among novice and returning motorcycle riders. *J Safety Res.* (82). 371-375. 2022.
- Hsu, T. P., Sadullah, E. A. F. M., and Dao, I. N. X. *A comparison study on motorcycle traffic development in some Asian countries—case of Taiwan, Malaysia, and Vietnam*. The Eastern Asia Society for Transportation Studies (EASTS), International Cooperative Research Activity. 2003.
- Lee, C., Pino, J., and Choi, K. Lessons Learned from Motorcyclist Surveys: Riders' Attitudes and Behaviors in Florida. *Transportation Research Record: Journal of the Transportation Research Board*. 2388 (1). 2013.
- Mannering, F. L., Shankar, V., and Bhat, C. R. Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic methods in accident research*, 11, 1-16. 2016.
- Nazemetz, J. W., Bents, F. D., Perry, J. G., Thor, C. P., and Mohamedshah, Y. M. Motorcycle Crash Causation Study (No. FHWA-HRT-18-064). United States. Federal Highway Administration. 2019.
- NHTSA (National Highway Traffic Safety Administration). Motorcycles: 2020 data (Traffic Safety Facts. Report No. DOT HS 813 306). National Center for Statistics and Analysis. 2022a.
- NHTSA (National Highway Traffic Safety Administration). State traffic data: 2020 data (Traffic Safety Facts. Report No. DOT HS 813 368). National Center for Statistics and Analysis. 2022b.
- Pongthanaisawan, J., and Sorapipatana, C. Relationship between level of economic development and motorcycle and car ownerships and their impacts on fuel consumption and greenhouse gas emission in Thailand. *Renewable & Sustainable Energy Reviews*, 14(9), 2966–2975. 2010.
- Rezapour, M., Nazneen, S., and Ksaibati, K. Application of deep learning techniques in predicting motorcycle crash severity. *Engineering Reports*. 2020.
- Safaei, N., Zhou, C., Safaei, B., and Masoud, A. Gasoline prices and their relationship to the number of fatal crashes on U.S. roads. *Transportation Engineering*; 2021.
- Sarmiento, J. M., Gogineni, A., Bernstein, J. I., Lee, C., Lineen, E. B., Pust, G. D., and Byers, P. Alcohol/Illlicit Substance Use in Fatal Motorcycle Crashes. *Journal of Surgical Research*; 2020.

- Wali, B., Ahmad, N., and Khattak, A. J. Toward better measurement of traffic injuries – Comparison of anatomical injury measures in predicting the clinical outcomes in motorcycle crashes. *Journal of Safety Research*; Vol.80, 175-189. 2021.
- Wong, K. I. An Analysis of Car and Motorcycle Ownership in Macao. *International Journal of Sustainable Transportation*, 7(3), 204–225. 2013.

Assessing the Impact of Cellular Coverage Areas on Distracted Driving, Crashes, and Injuries

S. Hernandez, Ph.D.¹; and Md. M. Rahman²

¹Dept. of Civil Engineering, Oregon State Univ., Corvallis, OR.

Email: sal.hernandez@oregonstate.edu

²Dept. of Civil Engineering, Oregon State Univ., Corvallis, OR.

Email: rahmamdm@oregonstate.edu

ABSTRACT

This study combines econometric and geospatial methods to analyze the impact of cellular coverage on distracted driving incidents, crashes, and injuries in Oregon from 2017 to 2020, focusing on cell phone-related crashes. Despite reduced travel in 2020, these crashes remained high. Geospatial tools identified urban hotspots like Portland and Salem. We used the mixed logit model to evaluate factors like driver demographics, vehicle characteristics, and environmental conditions, shedding light on the economic aspects of injury severity. Results highlight the crucial role of seatbelt use in reducing injury severity. The study underlines the need for comprehensive strategies to combat distracted driving in Oregon for better road safety and to lower economic costs associated with such incidents.

1. INTRODUCTION

Technological advancements in communication, coupled with the growing popularity of social media platforms like Instagram, Meta (formerly Facebook), and TikTok, have accelerated the use of cell phones in motor vehicles. As of now, the United States boasts over 300 million smartphone users, a figure projected to reach 360 million by 2040 (Statista, 2023). While smartphones serve as beneficial technological aids, offering everything from mapping directions to real-time traffic alerts, they pose significant distractions to drivers. The National Highway Traffic Safety Administration (NHTSA) defines "distraction" as a type of inattention occurring when drivers divert their attention from driving to another activity. NHTSA's (2021) report revealed that in the US in 2019, distraction-affected crashes accounted for 9% of fatal crashes, 15% of injury-related crashes, and 15% of all police-reported vehicle crashes. Additionally, 6% of drivers involved in fatal crashes were identified as distracted. Research indicates that 31.4% of individuals are distracted while on phone calls, and 16.6% while texting or dialing. Schroeder et al. (2018) conducted a comprehensive survey, finding that 56% of drivers engage in phone conversations, 9% send texts or emails, and 8% occasionally use apps while driving. Claveria et al. (2019) found that approximately 45% of truck drivers, from a sample of 515 respondents, used cell phones while driving in the Pacific Northwest Zone. Similarly, Gliklich et al. (2016) conducted a survey of 1,211 U.S. drivers, revealing that 43% frequently viewed maps. Notably, the highest percentage of distracted drivers involved in fatal crashes were those aged 15 to 20.

Between 2016 and 2020 in Oregon, distracted driving was a factor in over 15,000 crashes, resulting in 186 deaths and approximately 24,000 injuries, as ODOT (2023) reported. This trend is particularly alarming given the growing integration of social media into daily life. There's a notable correlation between the use of social media platforms and wireless connectivity

availability. According to Hersh et al. (2019), in areas with wireless connectivity, the accident rate increases by 1.1%, with most injuries being non-severe. Huisingh et al. (2019) found that the risk of a severe crash is 3.79 times higher when using a cell phone compared to not using one at all. Teenagers, particularly prone to serious injuries while driving distracted, are often affected by distractions from other drivers or their cell phones (Neyens et al. 2008). Additionally, Klauer et al. (2014) discovered that even experienced drivers are significantly more likely to have crashes or near-crashes when making phone calls.

A number of studies have explored the effects of distracted driving on road safety using naturalistic data (Owens et al., 2018; Dingus et al., 2016; Lu et al., 2020) and these have often focused on the impacts of texting, calling, or engaging with passengers while driving. And other studies have also highlighted the consequences of distractions related to both on- and off-cellular service (Qin et al. 2019; Sundfør et al. 2019). Despite extensive research, the connections between distracted driving, crash factors, crash severity, and geographical location remain unclear. This study aims to investigate these relationships, particularly examining how cellular coverage influences distracted driving and crashes. Further this study will investigate the statistical and spatial risks associated with cellphone usage, including its impact on the severity of injuries in vehicle crashes. By analyzing various contributing factors, this study seeks to understand the broader implications of distracted driving. Key questions include: *How does cellular data coverage affect driver attention? What factors contribute to injury severity in distracted driving crashes in Oregon?*

To accomplish this, the research utilized crash data obtained from the Oregon Department of Transportation (ODOT), primarily focusing on the data related to crashes caused by distracted driving between the years 2017 and 2020. This study used a mixed logit modeling framework with heterogeneity in means and variance for the analysis (Alnawmasi and Mannering 2022). Recognizing that distracted driving is a complex issue intertwined with driver behavior, it was essential to consider unobserved individual characteristics in the statistical model. To our knowledge, this is the first study to investigate the impact of cellphone coverage on distracted driving and injury severity. This research makes a significant contribution to road safety, particularly in the context of distracted driving. It introduces a mixed logit model that accounts for variations in both means and variances, incorporating random factors affecting distraction. The findings offer valuable insights for stakeholders in transportation safety, law enforcement, public health, and emergency medical services. These insights could be instrumental in developing targeted interventions to combat distracted driving.

2. LITERATURE REVIEW

2.1 Injury Severity Studies

With the rise of technology, distracted driving has become a major road safety issue. To better understand injury severity in such crashes, researchers have applied various statistical and econometric models that account for unobserved factors (unobserved heterogeneity) like the mixed logit (Alnawmasi and Mannering 2022; Fatmi et al. 2019; Wu et al. 2022; Chen et al. 2021). For example, Fatmi et al., (2019) through there application of such a model found that environmental factors like rain and road alignment impact injury severity in distracted driving incidents, with some elements like sidewalk length reducing it. Similarly, Razi-Ardakani et al. (2019) determined that cognitive and passenger distractions decrease injury severity, while cell

phone use increases it. Rain and curved roads also heighten injury risk, but morning peak hours reduce it. Islam (2023) focused on vehicle type in single-vehicle crashes, using random parameter multinomial logit models to account for heterogeneity. Restraint use emerged as a significant factor. Alnawmasi and Mannering (2022) noted a temporal decrease in injury severity, with daylight and high right shoulder indicators being significant factors in different years. Wu et al. (2022) observed a shift in significant factors like daytime and urban location in cell phone-related crashes in Pennsylvania. Meanwhile, Neyens and Boyle (2007) linked teenage drivers' rear-end collisions to cell phone distractions. García-Herrero et al. (2021) found that technological distractions almost double the risk of severe or fatal injuries in speeding scenarios. These studies emphasize the complexity of injury severity in distracted driving and the evolving use of econometric models that consider heterogeneity to better understand this issue.

2.2 Non-injury Severity Studies

This section provides a consolidated overview of various studies examining the impact of distracted driving on traffic safety and efficiency. It encompasses findings from Stavrinos et al. (2013) and Cooper et al. (2009), who identified that distracted driving leads to significant variations in lane positioning and speed, along with a tendency for riskier lane changes, adversely affecting traffic flow. Choudhary et al., (2017) and Xiao et al. (2016) further expand on this by noting behaviors such as reduced speeds and increased distances between vehicles, which contribute to lower traffic efficiency and more frequent overtaking incidents Xiao et al., (2015). Sherif et al., (2023) specifically focus on the impact of distracted driving at intersections, revealing a marked increase in the time interval between vehicles and a consequent reduction in intersection capacity.

Overall, these studies collectively illustrate the profound and varied ways in which driver distractions disrupt traffic dynamics, underscoring the critical need for continued research and targeted policy interventions to enhance road safety and maintain efficient traffic flow.

3. EMPIRICAL SETTING

For this study police-reported crash data sourced from ODOT's Crash Analysis and Reporting Unit, spanning 2017 to 2020, was collected (See Figure 1). Emphasis was placed on crash-level events specifically related to drivers' distractions. These events are characterized by several forms of distractions, including cell phone use, as documented on a Police Accident Report (PAR) or observed in use by the driver, instances where another party witnessed the driver's cell phone usage, distractions stemming from the operation of navigation systems or GPS devices, distractions attributed to other electronic devices, and incidents related to texting while driving. The comprehensive dataset identified a subset of 2,690 observations, each representing drivers involved in such distracting events. Each observation included information regarding driver, driver action, crash, roadway, temporal, environmental, and vehicle characteristics.

The study employed a modified version of the traditional KABCO injury scale to assess the severity of the outcomes stemming from these distractions. This scale was condensed into three primary categories for clarity: severe injury (comprising fatal and incapacitating outcomes, labeled as K+A), minor injury (including non-incapacitating and potential injuries, denoted as B+C), and cases where there was no injury sustained by the driver, resulting solely in property

damage (categorized as O). As shown in Table 1, a closer examination of the 2,690 observations revealed a breakdown in injury outcomes: 32 crashes (or 1.19%) led to severe injuries; 960 cases (or 35.69%) ended in minor injuries; and the majority, accounting for 1,698 crashes or 63.12%, documented instances where no injury, with damages limited to properties. The following table illustrates the descriptive statistics of the significant variables in the three injury severity models.

Table 1: Injury Severity distribution of the final dataset

Year	Severe Injury (%)	Minor Injury (%)	No Injury (%)	Total (%)
2017	6(0.87)	236(34.10)	450(65.03)	692(100)
2018	11(1.38)	292(36.64)	494(61.98)	797(100)
2019	6(0.89)	265(39.32)	403(59.79)	674(100)
2020	9(1.71)	167(31.69)	351(66.60)	527(100)
2017-2020	32(1.19)	960(35.69)	1698(63.12)	2690(100)

Table 2 illustrates the descriptive statistics of the significant variables in each of the three injury severity models. Collision type (rear-end, fixed object), airbag deployment, seatbelt use, speed greater than 55 mph, female, and driver proximity within 25 miles to the residence were the variables found to be significant for different injury severity categories.

Table 2: Descriptive Statistics of Significant Variables by Injury Severity Category

Variable	Mean	Std Deviation
Mixed Logit Model		
Airbag (1 if the airbag deployed, 0 otherwise)	0.134572	0.341288
Collision Type (1 if rear-end, 0 otherwise)	0.562082	0.496162
High Speed (1 if was greater than 55 MPH, 0 otherwise)	0.178439	0.382905
Airbag (1 if the airbag deployed, 0 otherwise)	0.134572	0.341288
Collision Type (1 if fixed-object, 0 otherwise)	0.086245	0.280743
Safety Equipment (1 if seatbelt use, 0 otherwise)	0.549814	0.497543
Low Speed (1 if speed greater than 20 MPH but Less than 40 MPH, 0 otherwise)	0.302602	0.459413
Gender (1 if female, 0 otherwise)	0.30223	0.459253
Driver Proximity to Residence (1 if within 25 Miles, 0 otherwise)	0.521933	0.49955
Age (1 if driver age is less than 25 years old)	0.526766	0.499314

In addition, a pivotal aspect of this research aimed to ascertain whether cell phone coverage, or its absence, played a role in influencing the locations of distracted driving crash clusters (see Figure 1a). Figure 2a and Figure 2b present the mobile coverage maps for Verizon and AT&T, respectively, superimposed onto the recorded crash sites from the study period. The maps employ light-colored regions to depict areas devoid of coverage, while pink (in Figure 2a for Verizon) and yellow (for AT&T in Figure 2b) shades signify areas with cellular service. Upon close examination, a notable pattern emerges: most crashes appear to be concentrated within the

cell service zones for both carriers. This suggests a potential correlation between areas with active mobile service and the incidence of distracted driving crashes, underscoring the need for further investigation into the underlying factors and drivers' behaviors in these regions.

Furthermore, central to this research was the utilization of heatmaps/hotspots within QGIS—a geospatial tool renowned for its adeptness at visualizing spatial data distributions (See Figure 1b). Through the heatmaps/hotspots, the analysis transformed discrete data points into continuous visual narratives, delineating regions experiencing elevated instances of cell phone-induced/related crashes. The subsequent urban analyses illustrated that Portland and Salem, highlighted the distracted driving scenario in Oregon (see Figure 1b).

Hence, the relationship between cell phone use, connectivity, and distracted driving crashes is both complicated and multifaceted. While regions with pronounced cell service witness a concentration of such crashes, sporadic connectivity zones present their own set of challenges, potentially diverting driver attention (this was confirmed from hot spot analysis). As this study reveals the overarching patterns in Oregon, it also underscores the importance of further research and strategic interventions to address this pressing concern. As such, this study proposes an econometric to uncovers the complex interactions.

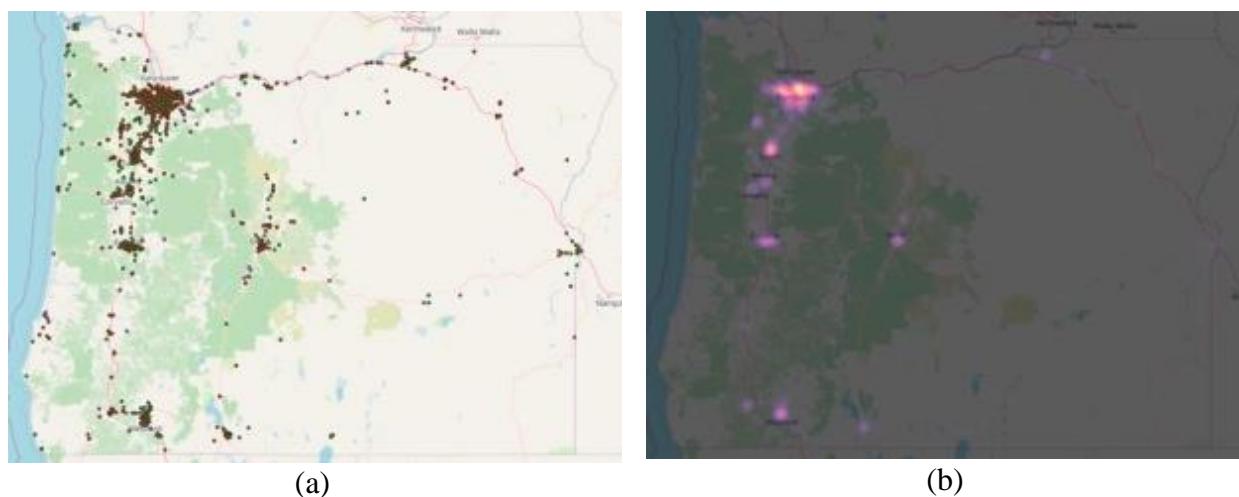


Figure 1. (a) Spatial Distribution of Distracted Driving Crashes in Oregon (2017-2020): A GIS visualization highlighting the geolocations of reported incidents over the four-year study period; (b) Heatmap of Oregon: Delineating Concentrations of Distracted Driving Crashes with Dominant Clusters in Major Urban Centers like Portland, Salem, Eugene, Medford, and Bend.

4. METHODOLOGY

In the present research, while police-reported crash data offer extensive insights, they need to capture certain details. Aspects such as the driver's physical characteristics (e.g., height, weight) or nuanced environmental conditions at the exact moment of the crash (e.g., subtle shifts in weather or lighting) remain to be determined. Such factors can introduce unobserved variations across the dataset, termed as "unobserved heterogeneity." If not addressed, this heterogeneity can skew the model's estimations, potentially leading to biased outcomes, as highlighted by (Mannering et al., 2016).

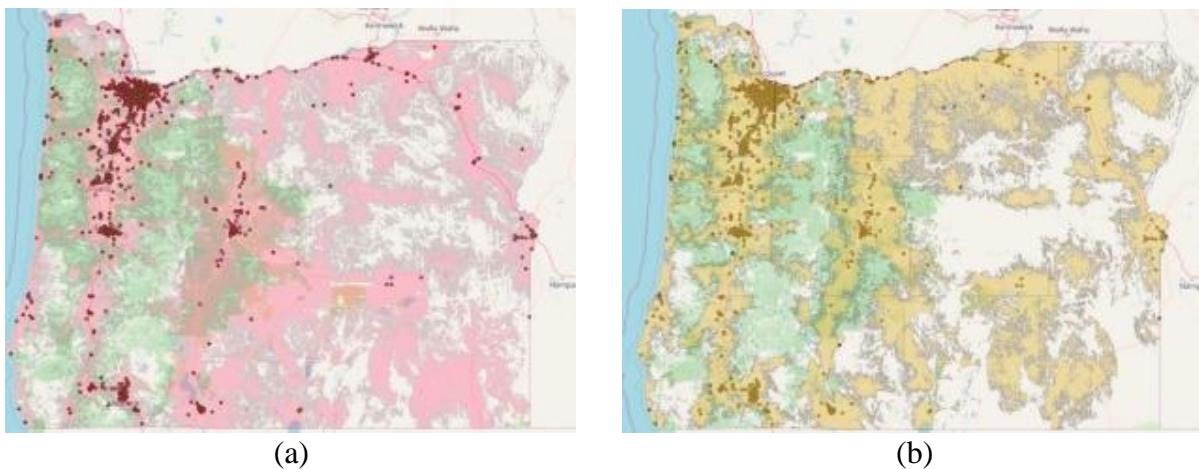


Figure 2. (a) Verizon Mobile Cell Coverage Map Superimposed onto the Recorded Crash Sites from the Study Period (2017-2020); (b) AT&T Mobile Coverage Map Superimposed onto the Recorded Crash Sites from the Study Period (2017-2020).

Current research employed the mixed logit model with possible heterogeneity in the means and variance of random parameters to mitigate the impact of this unobserved heterogeneity. This methodology stands as a cutting-edge statistical and econometric tool, with its application evident in a myriad of recent studies focused on injury severity (Alnawmisi and Mannering 2022; Al-Bdairi et al. 2020; Behnoor and Mannering 2017; Islam 2021; Zubaidi et al. 2021). Further, this econometric modeling method treats injury severity outcomes as discrete choices, enabling insights into the probability of each injury severity outcome. Using this approach, the estimated parameters of the mixed logit model highlight statistically significant factors that either elevate or reduce the likelihood of specific injury severity outcomes.

The mixed logit model starts with a linear function. Each linear function corresponds to a particular injury severity resulting from a distracted driving crash and can be represented as:

$$U_{in} = \beta_i X_{in} + \epsilon_{in} \quad (1)$$

Where U_{in} is a linear function for injury severity i and distracted driving crash n ; i represents injury severities of no injury, minor injury and severe injury; X_{in} represents the vector of explanatory variables (roadway characteristics, driver actions, driver characteristics, roadway characteristics, demographic characteristics, environment characteristics) that lead to the discrete outcome of crash due to distracted driving n ; β_i represents the vector of estimated parameters for injury severity i and ϵ_{in} is the error term that attempts to capture the unobserved factors within the model (Washington et al. 2011); but ϵ_{in} is unable to capture all the unobserved factors. Police-reported crash data often lacks certain essential variables, and the variability within the available variables can lead to unobserved heterogeneity. If this heterogeneity is overlooked, it may produce biased estimates and lead to incorrect conclusions (Mannering et al., 2016). Therefore, the mixed logit model captures this heterogeneity by allowing varying parameters. In addition, the mixed logit model (if variables are found to be random) eliminates the independence from irrelevant alternatives (IIA) property. In essence, by accounting for

variables identified as random, unobserved factors are addressed, allowing for the categorization of injury severities into three distinct groups (Geedipally et al., 2011). The mixed logit model is then formulated as follows (McFadden and Train 2000; Washington et al. 2011).

$$P_n(i|\phi) = \frac{e^{(\beta_i X_{in})}}{\sum_{\forall i} e^{(\beta_i X_{in})}} k(\beta_i|\phi) d\beta_i \quad (2)$$

where $P_n(i|\phi)$ is the weighted outcome probability of injury severity i (severe, minor and no injury) conditional on $k(\beta_i|\phi)$, where $k(\beta_i|\phi)$ is the density function of β_i and ϕ with distribution specified by the analyst—the density function is what allows the parameters to vary and is regularly specified to be normally distributed. All other variables have the same definition as the ordinary multinomial logit model (Washington et al. 2011).

To account the unobserved heterogeneity in distracted driving by incorporating heterogeneity in the means and variances of random parameters, β_{in} is modeled to be a function of additional explanatory variables that influence its mean and variance as demonstrated (Seraneeprakarn et al. 2017; Behnood and Mannering 2017)

$$\beta_{in} = \beta + \delta_{in} Z_{in} + \sigma_{in} \text{EXP}(\omega_{in} W_{in}) v_{in} \quad (3)$$

Where, β is the mean parameter estimate across all distracted driving crashes, Z_{in} is a vector of explanatory variables that captures heterogeneity in the mean that affect drivers injury-severity level i (severe, minor and no injury), δ_{in} is a corresponding vector of estimable parameters, W_{in} is a vector of explanatory variables that captures heterogeneity in the standard deviation σ_{in} with corresponding parameter vector ω_{in} , and v_{in} represents a disturbance term.

A total of 200 Halton drawings were employed in this simulation approach due to its higher effectiveness and preference over random draws (Bhat 2003). This study estimates the marginal effect for all significant explanatory variables, which enables the assessment of how individual variable estimations influence the likelihood of distracted driving injury severity outcomes. The marginal effect quantifies the effect of a one-unit change in the chosen explanatory variable on the probability of injury severity outcomes while holding all other variables constant. The marginal effect for the k th indicator variable associated with injury severity level i for driver n (X_{ikn}) can be calculated by:

$$ME_{X_{ikn}}^{P_n(i)} = [P_n(i) = 1 | X_{ikn} = 1] - [P_n(i) = 1 | X_{ikn} = 0] \quad (4)$$

5. DISCUSSION OF ESTIMATED RESULTS

From the analysis, as shown in Table 3, ten unique variables were identified as significant across three injury severity categories (severe, minor, and no injury). Notably, the variable 'airbag deployment' was significant in both the 'No injury' and 'Minor injury' categories. Out of these ten variables, two were found to be random parameters with statistically significant means and standard deviations. Specifically, as per Table 3, the random parameters were 'Airbag

deployment' for the 'Minor Injury' category and 'Driver Proximity to Residence' (within 25 miles of their home) for the 'No Injury' category.

Table 3: Estimated results of injury severity for mixed logit model

Variable	Coefficient	T-Statistic	Marginal effects		
			Severe Injury	Minor Injury	No Injury
Constant [SI]	-7.04763	-13.39***			
Constant [MI]	-2.64681	-10.11***			
Driver Characteristics					
Gender (1 if female, 0 otherwise) [NI]	-0.84311	-3.43***	0.0007	0.0205	-0.0212
Age (1 if driver age is less than 25 years old) [NI]	1.77923	7.17	-0.0010	-0.0347	0.0358
Crash Characteristics					
Collision Type (1 if rear-end, 0 otherwise) [SI]	-1.59146	-3.41***	-0.0049	0.0038	0.0011
Airbag (1 if the airbag deployed, 0 otherwise) [SI]	2.35322	4.15***	0.0124	-0.0092	-0.0032
Collision Type (1 if fixed-object, 0 otherwise) [MI]	1.76798	3.88***	-0.0012	0.0107	-0.0096
Airbag (1 if the airbag deployed, 0 otherwise) [MI]	1.91106	4.18***	0.0008	0.0130	-0.0122
(Standard Deviation of Parameter, Normally Distributed)	(1.94798)	(2.18) **			
Safety Equipment (1 if seatbelt use, 0 otherwise) [NI]	-2.58226	-9.25***	0.0052	0.1256	-0.1308
Accident-Specific Characteristics					
High Speed (1 if speed was greater than 55 MPH, 0 otherwise) [SI]	2.62176	4.97***	0.0139	-0.0102	-0.0036
Low Speed (1 if speed greater than 20 MPH but Less than 40 MPH, 0 otherwise) [NI]	0.73945	3.12***	-0.0002	-0.0124	0.0127
Driver Proximity to Residence (1 if within 25 Miles, 0 otherwise) [NI]	-2.12281	-4.82***	0.0008	0.0217	-0.0225
(Standard Deviation of Parameter, Normally Distributed)	(3.01796)	(4.48)			
Heterogeneity in the means of random parameter					
Airbag [MI]: Age greater or equal to 35 and less than 45	-1.57362	-2.22**	-	-	-
Driver Proximity to Residence [NI]: Male	1.31031	2.48**	-	-	-
Heterogeneity in the variance of random parameter					
Driver Proximity to Residence [NI]: Rear-end crash	0.74355	3.38***	-	-	-
Model Statistics					
Number of Observations	2690				
Restricted Log-Likelihood	-2955.267				
Log-Likelihood at Convergence	-1246.0096				
McFadden pseudo-R-squared (ρ^2)	0.5784				

Note: *Italic* value: Random parameter. [SI]: Severe injury, [MI]: Minor injury, [NI]: No injury. ***, **, * denotes significance at 1%, 5%, 10% level.

5.1 Random Parameters

The variable 'airbag deployment' in the 'Minor injury' category was found to be a random parameter that followed a normal distribution with a mean of 1.91106 and a standard deviation of 1.94798 (see Table 3). This indicates that in approximately 16.33% of the cases where airbags were activated during distracted driving events, the average effect of the parameter was negative. Conversely, the average effect was positive for 83.67% of the cases. Therefore, for 16.33% of drivers, airbag deployment reduced the likelihood of incurring a minor injury during distracted driving incidents. However, for the remaining 83.67%, airbag deployment had the inverse effect.

Similarly, the variable 'Driver Proximity to Residence' (within 25 miles of their home) in the 'No Injury' category exhibited characteristics of a random parameter. It was found to be random and normally distributed, with a mean of -2.12281 and a standard deviation of 3.01796. This distribution suggests that in cases where drivers were within 25 miles of their residence during distracted driving crashes, the average effect of the parameter was positive for a certain percentage of observations and harmful for the rest. Specifically, for approximately 24.09% of such cases, being close to one's residence increased the likelihood of sustaining no injuries during distracted driving events. Conversely, for the remaining 75.91%, being near one's home had the opposite effect, suggesting these drivers were more prone to sustaining injuries (the negative sign).

5.2 Heterogeneity in Means and Variance

The data in Table 3 reveals interesting insights into the impact of various explanatory variables on the mean and variance of random parameters in distracted driving crashes. Specifically, it was noted that individuals aged 35 to 45 and males had a sequential effect on the mean values of the Airbag and Driver Proximity to Residence random parameters. The variable "Age 35 to less than 45" was associated with a decrease in the mean of the airbag variable for minor injuries. This may indicate a correlation where drivers in this age group are less likely to sustain minor injuries in such events. The variable 'Male' was observed to be positively associated with the 'Driver Proximity to Residence' variable in instances of no-injury crashes, suggesting a correlation where male drivers are more often involved in no-injury crashes when these occur closer to their residence. However, it's important to note that this association does not imply causality. This could be capturing driving behaviors associated with male drivers especially in closer distances to their residence in comparison to female drivers.

The "Rear-end" variable was the only explanatory factor found to be significant in accounting for the variability in the variance of the random parameter related to the driver's proximity to residence in no-injury scenarios. This variable contributes to an increased variability of the 'Driver's Proximity to Residence' parameter. Specifically, this suggests a correlation where the likelihood of a driver avoiding injury in a rear-end collision seems to increase when the incident occurs closer to their residence. This could be capturing the risk taking behaviors of drivers close to home which is consistent with Burdett, Starkey, and Charlton (2017).

5.3 Driver Characteristics

The data reveals a negative correlation (-0.84311) between the presence of female drivers (Gender Variable) and the occurrence of no-injury crashes, suggesting a tendency for female

drivers to be less frequently involved in crashes that do not result in injuries. However, the same data shows that female drivers have a higher likelihood of being involved in crashes resulting in severe or minor injuries. As indicated by the marginal effects (See Table 3), being a female driver increases the probability of experiencing serious injuries by 0.0007 and minor injuries by 0.0205. These findings, particularly concerning the context of cell phone use during driving, align with the results reported by Russo et al. (2014), although it's crucial to consider the differences in study design, population, and variables when making such comparisons. The observed trends warrant a cautious interpretation and highlight the need for a more in-depth understanding of the underlying factors that contribute to these gender differences in crash outcomes, especially in the context of distracted driving.

Conversely, for drivers under 25 years old, there is a positive correlation (1.77923) with no-injury crashes. This implies that the likelihood of these younger drivers not sustaining injuries in an accident increases by 0.0358, while the chances of incurring minor and severe injuries decrease. Factors contributing to this trend may include a lack of driving experience and a tendency for reduced risk-taking due to lower confidence. Additionally, greater physical resilience in younger individuals might lead to a higher incidence of non-injury outcomes compared to severe or minor injuries.

5.4 Crash Characteristics

Rear-end collisions are associated with a decrease in the probability of severe injuries but show an increase in the likelihood of minor and no injuries. This indicates that while rear-end collisions are common, they often result in less severe injuries. This is likely because such crashes typically occur in congested conditions where vehicles maintain lower speeds, leading to less severe impacts.

In contrast, collisions with fixed objects, increase the likelihood of minor injuries while reducing the chances of severe injuries and no injuries. This finding is consistent with Alnawmasi and Mannering (2022) who similarly identified the significance of fixed object collisions in causing minor injuries.

When an airbag is deployed during a distracted driving crash, the probability of severe injuries increases as indicated for the severe injury category (see Table 3). However, this same deployment slightly reduces the chances of minor injuries and no injuries, respectively.

Additionally, for the minor injury category, airbag deployment is associated with slight increase in minor injuries and a decrease in no-injury crashes. These findings imply that while airbag deployment can mitigate some injuries, it is also linked to a notable increase in severe injuries. The airbag's effect appears to have a degree of randomness, which can lead to minor injuries, as evidenced by the significant variation in the standard deviation of this parameter, suggesting it follows a normal distribution (See Table 3).

Wearing a seatbelt significantly reduces the risk of sustaining severe injuries in a crash and increases the likelihood of surviving without any injuries. This underscores the critical role of seatbelts in enhancing passenger and driver safety during vehicular accidents. However, there is a possibility that some drivers might develop a false sense of security when wearing a seatbelt, potentially leading to more aggressive driving behaviors. This could inadvertently result in more severe collisions as seen from the marginal effects.

5.5 Accident-Specific Characteristics

Driving at speeds greater than the posted speed limit of 55 MPH is associated with a higher probability of severe injuries, while it slightly decreases the likelihood of minor and no injuries (see marginal effects). This correlation between higher speeds and more severe outcomes is consistent with expectations, as higher velocities can worsen the impact and consequences of a crash. Conversely, adhering to posted speed limits, especially in zones with limits between 20 to 40 MPH, increases the probability of emerging from a crash without injuries and decreases the likelihood of severe and minor injuries. This outcome aligns with general traffic safety principles, as driving at lower, regulated speeds typically reduces the force of impact and the potential for injury severity in the event of an accident.

Being within 25 miles of one's residence slightly diminishes the probability of no injuries, yet it increases the odds for minor and severe injuries. Nevertheless, it is essential to acknowledge that this variable produced unpredictable parameters for the model, suggesting the presence of substantial variations in its effects on distracted driving crashes, as shown earlier.

6. CONCLUSION

This study investigated crash factors linked to distractions like cell phone use, utilizing data from the Oregon Department of Transportation from 2017 to 2020. It reviewed 2,690 distracted driving crashes and found 1.19% led to severe injuries, 35.69% to minor injuries, and 63.12% resulted in property damage only. Hot spot analysis was conducted and confirmed the visual inspection of crashes to cell coverage. In addition, the study acknowledges certain limitations, such as the absence of detailed information on the driver's physical condition or minor environmental variations at the time of the crash, which introduces "unobserved heterogeneity." To mitigate this, a mixed logit model for injury severity was estimated. The analysis shed light on several interesting findings. For instance, airbag deployment during a distracted driving incident heightens the chance of severe injuries by 0.99%. Conversely, rear-end collisions, despite being frequent, often culminate in less severe injuries. Safety equipment usage, particularly seatbelts, substantially mitigates injury, emphasizing their critical importance. Furthermore, younger drivers, those below 25 years, exhibited a higher likelihood of less severe injuries. These insights offer a comprehensive view of distraction-induced crashes, highlighting the need for targeted safety interventions.

REFERENCES

- Al-Bdairi, N. S. S., A. Behnood, and S. Hernandez. 2020. "Temporal Stability of Driver Injury Severities in Animal-Vehicle Collisions: A Random Parameters with Heterogeneity in Means (and Variances) Approach." *Analytic Methods in Accident Research* 26 (June): 100120.
- Alnawmasi, N., and F. Mannering. 2022. "A Temporal Assessment of Distracted Driving Injury Severities Using Alternate Unobserved-Heterogeneity Modeling Approaches." *Analytic Methods in Accident Research* 34 (June).
- Behnood, A., and F. Mannering. 2017. "Determinants of Bicyclist Injury Severities in Bicycle-Vehicle Crashes: A Random Parameters Approach with Heterogeneity in Means and Variances." *Analytic Methods in Accident Research* 16 (December): 35–47.

- Bhat, C. R. 2003. "Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences." *Transportation Research Part B: Methodological* 37 (9): 837–55.
- Burdett, B. R. D., N. J. Starkey, and S. G. Charlton. 2017. "The Close to Home Effect in Road Crashes." *Safety Science* 98 (October): 1–8.
- Chen, Z., and Y. Lym. 2021. "The Influence of Built Environment on Distracted Driving Related Crashes in Ohio." *Transport Policy* 101 (February): 34–45.
- Choudhary, P., and N. R. Velaga. 2017. "Modelling Driver Distraction Effects Due to Mobile Phone Use on Reaction Time." *Transportation Research Part C: Emerging Technologies* 77 (April): 351–65.
- Claveria, J. B., S. Hernandez, J. C. Anderson, and E. L. Jessup. 2019. "Understanding Truck Driver Behavior with Respect to Cell Phone Use and Vehicle Operation." *Transportation Research Part F: Traffic Psychology and Behaviour* 65 (August): 389–401.
- Cooper, J. M., I. Vladisavljevic, N. Medeiros-Ward, P. T. Martin, and D. L. Strayer. 2009. "An Investigation of Driver Distraction Near the Tipping Point of Traffic Flow Stability." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51 (2): 261–68.
- Dingus, T. A., F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King, and J. Hankey. 2016. "Driver Crash Risk Factors and Prevalence Evaluation Using Naturalistic Driving Data." *Proceedings of the National Academy of Sciences* 113 (10): 2636–41.
- Fatmi, M. R., and M. Ahsanul Habib. 2019. "Modeling Vehicle Collision Injury Severity Involving Distracted Driving: Assessing the Effects of Land Use and Built Environment." *Transportation Research Record: Journal of the Transportation Research Board* 2673 (7): 181–91.
- García-Herrero, S., J. Diego Febres, W. Boulagouas, J. Manuel Gutiérrez, and M. Ángel Mariscal Saldaña. 2021. "Assessment of the Influence of Technology-Based Distracted Driving on Drivers' Infractions and Their Subsequent Impact on Traffic Accidents Severity." *International Journal of Environmental Research and Public Health* 18 (13): 7155.
- Geedipally, S. R., P. A. Turner, and S. Patil. 2011. "Analysis of Motorcycle Crashes in Texas with Multinomial Logit Model." *Transportation Research Record: Journal of the Transportation Research Board* 2265 (1): 62–69.
- Gliklich, E., R. Guo, and R. W. Bergmark. 2016. "Texting While Driving: A Study of 1211 U.S. Adults with the Distracted Driving Survey." *Preventive Medicine Reports* 4 (December): 486–89.
- Hersh, J. S. H., B. J. Lang, and M. Lang. 2019. "Digitally Distracted at the Wheel: Car Accidents and Smartphone Coverage (Authors' Names Blinded for Peer Review)."
- Huisingsh, C., C. Owsley, E. B. Levitan, M. R. Irvin, P. MacLennan, and G. McGwin. 2019. "Distracted Driving and Risk of Crash or Near-Crash Involvement Among Older Drivers Using Naturalistic Driving Data With a Case-Crossover Study Design." *The Journals of Gerontology: Series A* 74 (4): 550–55.
- Islam, M. 2021. "The Effect of Motorcyclists' Age on Injury Severities in Single-Motorcycle Crashes with Unobserved Heterogeneity." *Journal of Safety Research* 77 (June): 125–38.
- Islam, M. 2023. "Unraveling the Differences in Distracted Driving Injury Severities in Passenger Car, Sport Utility Vehicle, Pickup Truck, and Minivan Crashes." <https://ssrn.com/abstract=4606333>.
- Klauer, S. G., F. Guo, B. G. Simons-Morton, M. Claude Ouimet, S. E. Lee, and T. A. Dingus. 2014. "Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers." *New England Journal of Medicine* 370 (1): 54–59.

- Lu, D., F. Guo, and F. Li. 2020. "Evaluating the Causal Effects of Cellphone Distraction on Crash Risk Using Propensity Score Methods." *Accident Analysis & Prevention* 143 (August): 105579.
- Mannering, F. L., V. Shankar, and C. R. Bhat. 2016. "Unobserved Heterogeneity and the Statistical Analysis of Highway Accident Data." *Analytic Methods in Accident Research* 11 (September): 1–16.
- McFadden, D., and K. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15 (5): 447–70.
- Neyens, D. M., and L. Ng Boyle. 2007. "The Effect of Distractions on the Crash Types of Teenage Drivers." *Accident Analysis & Prevention* 39 (1): 206–12.
- Neyens, D. M., and L. Ng Boyle. 2008. "The Influence of Driver Distraction on the Severity of Injuries Sustained by Teenage Drivers and Their Passengers." *Accident Analysis & Prevention* 40 (1): 254–59.
- Owens, J. M., T. A. Dingus, F. Guo, Y. Fang, M. Perez, and J. Mcclafferty. 2018. "Crash Risk of Cell Phone Use While Driving: A Case-Crossover Analysis of Naturalistic Driving Data." www.aaafoundation.org.
- Qin, L., Z. R. Li, Z. Chen, M. S. Andi Bill, and D. A. Noyce. 2019. "Understanding Driver Distractions in Fatal Crashes: An Exploratory Empirical Analysis." *Journal of Safety Research* 69 (June): 23–31.
- Razi-Ardakani, H., A. Mahmoudzadeh, and M. Kermanshah. 2019. "What Factors Results in Having a Severe Crash? A Closer Look on Distraction-Related Factors." *Cogent Engineering* 6 (1).
- Russo, B. J., J. J. Kay, P. T. Savolainen, and T. J. Gates. 2014. "Assessing Characteristics Related to the Use of Seatbelts and Cell Phones by Drivers: Application of a Bivariate Probit Model." *Journal of Safety Research* 49 (June): 137.e1-142.
- Schroeder, P., M. Wilbur, and R. Peña. 2018. "National Survey on Distracted Driving Attitudes and Behaviors-2015." www.ntis.gov.
- Seraneeprakarn, P., S. Huang, V. Shankar, F. Mannering, N. Venkataraman, and J. Milton. 2017. "Occupant Injury Severities in Hybrid-Vehicle Involved Crashes: A Random Parameters Approach with Heterogeneity in Means and Variances." *Analytic Methods in Accident Research* 15 (September): 41–55.
- Sherif, B., H. Abou-Senna, and E. Radwan. 2023. "Investigating the Effects of Left-Turn Distracted Drivers on Signalized Intersections' Traffic Operations." *Transportation Research Part F: Traffic Psychology and Behaviour* 96 (July): 111–32.
- Stavrinos, D., J. L. Jones, A. A. Garner, R. Griffin, C. A. Franklin, D. Ball, S. C. Welburn, K. K. Ball, V. P. Sisiopiku, and P. R. Fine. 2013. "Impact of Distracted Driving on Safety and Traffic Flow." *Accident Analysis & Prevention* 61 (December): 63–70.
- Sundfør, H. B., F. Sagberg, and A. Høye. 2019. "Inattention and Distraction in Fatal Road Crashes – Results from in-Depth Crash Investigations in Norway." *Accident Analysis & Prevention* 125 (April): 152–57.
- Washington, S., M. Karlaftis, F. Mannering, and P. Anastasopoulos. 2011. "Statistical and Econometric Methods for Transportation Data Analysis." <https://www.crcpress.com/go/ids>.
- Wu, P., L. Song, and X. Meng. 2022. "Temporal Analysis of Cellphone-Use-Involved Crash Injury Severities: Calling for Preventing Cellphone-Use-Involved Distracted Driving." *Accident Analysis & Prevention* 169 (May): 106625.

- Xiao, Y., and J. Shi. 2015. "Analyzing the Influence of Mobile Phone Use of Drivers on Traffic Flow Based on an Improved Cellular Automaton Model." *Discrete Dynamics in Nature and Society* 2015: 1–9.
- Xiao, Y., and J. Shi. 2016. "Study on the Influence of Driving Distraction on Traffic Flow Considering the Stochastic Duration Time of Distraction." *Modern Physics Letters B* 30 (32n33): 1650380.
- Zubaidi, H., I. Obaid, A. Alnedawi, S. Das, and M. M. Haque. 2021. "Temporal Instability Assessment of Injury Severities of Motor Vehicle Drivers at Give-Way Controlled Unsignalized Intersections: A Random Parameters Approach with Heterogeneity in Means and Variances." *Accident Analysis & Prevention* 156 (June): 106151.

Estimating Driver Behaviour at Unsignalized Crosswalks

H. M. N. S. Kumari¹; I. P. W. N. Thathsara²; W. M. V. S. K. Wickramasinghe³;
and Sunanda Dissanayake⁴

¹Dept. of Civil Engineering, Faculty of Engineering, Univ. of Peradeniya, Sri Lanka.
Email: nawodaherath1998@gmail.com

²Dept. of Civil Engineering, Faculty of Engineering, Univ. of Peradeniya, Sri Lanka.
Email: nthathsara1@gmail.com

³Dept. of Civil Engineering, Faculty of Engineering, Univ. of Peradeniya, Sri Lanka.
Email: vskw@pdn.ac.lk

⁴Dept. of Civil Engineering, Kennesaw State Univ., Kennesaw, GA. Email: sunanda@k-state.edu

ABSTRACT

This research investigates driver-yielding behaviour at unsignalized crosswalks, focusing on pedestrian characteristics, and oncoming vehicle factors. The study uses data from two-lane, two-way divided roads, and drone video footage to capture vehicle trajectories and pedestrian behaviour at the median or further lanes. Driver-yielding rates are accurately captured using TRACKER 5.0 software, and the behavior is modelled using binary logistic regression. The study found that yielding patterns are similar across different vehicle categories, but factors influencing drivers' willingness include the number of female pedestrians, the number of pedestrians in the median, the number of injured pedestrians, and the vehicle position/vehicle approaching the lane. The research contributes to understanding and addressing the complexities of unsignalized pedestrian crossings, laying the groundwork for future traffic management strategies to create safer, more efficient urban environments for pedestrians and drivers.

INTRODUCTION

This research focuses on pedestrian safety at unsignalized crosswalks in urban environments, highlighting the complex interactions between pedestrians and vehicular traffic. It aims to understand driver behaviour and estimate the same for enhancing pedestrian safety and developing effective traffic management strategies. The study focuses on two-lane two-way divided roads, a context often overlooked in previous research. It aims to clarify driver-yielding patterns based on vehicular velocities examining factors influencing yielding behaviour based on pedestrian attributes. The methodology is designed to explore driver behaviour during and after pedestrian crossings, identifying yielding behaviour across various vehicle classes and understanding factors influencing drivers' willingness based on pedestrian attributes. The research contributes to understanding complex driver-pedestrian dynamics at unsignalized crosswalks, offering insights for refined traffic management strategies and calibrated velocity control measures.

LITERATURE REVIEW

For the data collection methods, Bertulis et al. (2014) used a radar gun to collect vehicle speed data, focusing on free-flowing vehicles within a 2 m/s range. Dileep et al. (2016) collected data using videographic surveys, radar guns, and manual observations, capturing the interaction

between pedestrians and vehicles at unsignalized crosswalks. Schroeder et al. (2011) used a laser speed gun with light detection and ranging (LIDAR) technology to collect data on vehicle speeds and their relative positions on the crosswalk. Yang et al. (2022) used a purposive sampling approach to select participants for semi-structured interviews, obtaining in-depth qualitative data. Lobjois et al. (2007) used participants to decide whether to traverse the space between two cars, with the response button representing acceptance or rejection. Fu et al. (2019) used GoPro's Hero and Hero 3 Edition cameras for video data collection, capturing data from an elevated vantage point. Bella et al. (2021) conducted a study on two pedestrian crossings in Rome, comparing driver behaviour at the approach to the conflict point. The recordings were conducted during off-peak hours, minimizing disruptions caused by high traffic volume.

For the data extraction methods, Schroeder et al. (2011) and Bertulis et al. (2014) conducted data collection and analysis on yielding and not yielding behaviours at crosswalks. Schroeder et al. (2011) recorded time stamps for pedestrian and vehicle arrivals at the crosswalk, while Bertulis et al. collected data from nine locations, including marked intersections in Boston and Brookline. Yang et al. (2022) used a purposeful selection method to collect primary data, which was processed and analyzed using NVIVO11 software. Fu et al. (2019) used the BriskLUMINA 0.1 computer-vision platform to extract trajectory data from video data. Lobjois et al. (2007) conducted an experiment where participants were asked to decide whether to cross the road, but they were not physically required to cross. The study recorded participants' choices regarding crossing and response times for each trial. The results provide valuable insights into yielding behaviours at crosswalks.

For the data analysis methods, Dileep et al. (2016) employed descriptive and inferential statistics, utilizing the Chi-square test to examine driver-pedestrian interactions at unsignalized crosswalks. They explored dependencies between various variables, including vehicular and pedestrian characteristics. Bertulis et al. (2014) focused on sample size adequacy, considering the necessary participants for their study. Schroeder et al. (2011) collected data on driver-yielding behaviour, introducing variables like "soft yields" and "hard yields." Yang et al. (2022) used grounded theory analysis with 22 interview materials to explore drivers' concerns, transcribing over 11,000 words and employing NVivo11 coding software for inductive coding, resulting in 316 original codes.

METHODOLOGY

Site Selection

In the design of four-lane divided roads, one of the paramount considerations is the establishment of specific headway guidelines for vehicles. Headway, which refers to the distance or time interval between successive vehicles travelling in the same direction, plays a pivotal role in ensuring the safety and efficiency of traffic flow on these roadways. To comprehensively analyze the patterns of driver yielding, particularly about vehicle velocity and pedestrian characteristics, the choice of location is of utmost importance. Opting for a road configuration where the path is nearly straight on both sides of the crosswalk as shown in Figure 1, is the key to achieving this goal. The advantages of such a straight-path design are multifold. Firstly, it significantly enhances visibility for drivers, enabling them to have a clear line of sight to the crossing pedestrians. This heightened visibility empowers drivers to maintain appropriate following distances between vehicles, mitigating the risks associated with tailgating and rear-end

collisions. Additionally, on a straight road, drivers can more accurately perceive changes in vehicle velocity, facilitating smoother and more predictable traffic flow. The efficient traffic flow, in turn, reduces congestion and the likelihood of delays, thus improving the overall driving experience for commuters. Ultimately, a straight path configuration contributes to road safety by reducing collision risks, especially at pedestrian crosswalks. Consequently, the crosswalk located near the Royal Mall on William Gopallawa Mawatha, Sri Lanka which is shown in Figure 2, has been chosen as the designated site.



Figure 1: Drone image showing the straight upstream to the crosswalk

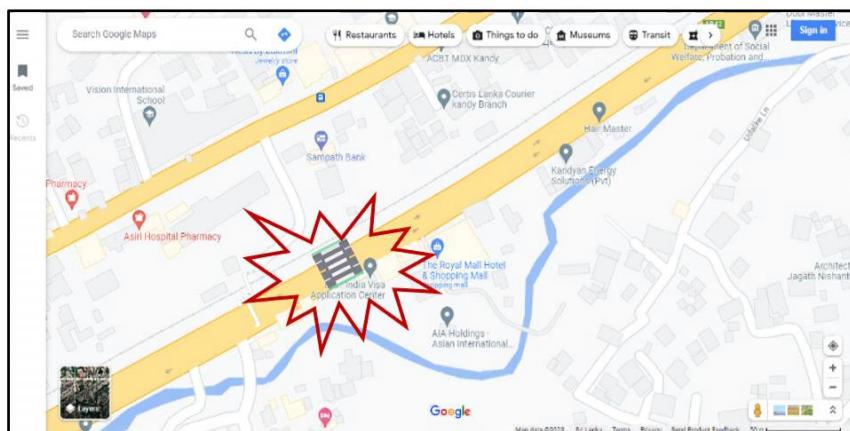


Figure 2: Location of the selected crosswalk (7.2742005 °N, 80.6125394 °E)

Data Collection

For data collection, an innovative approach involving the use of video graphics was employed. This method harnessed the capabilities of drone technology in conjunction with video camera camcorders. Drones equipped with high-resolution video cameras were deployed to capture a comprehensive and highly detailed view of the surveyed area. The aerial perspective offered by the drones proved instrumental in conducting an in-depth analysis of the road infrastructure and traffic dynamics. Simultaneously, video cameras recorded real-time footage, capturing crucial

information, including vehicle trajectories, lane utilization, and pedestrian activities. This approach to data collection facilitated a thorough evaluation of road conditions and served as a valuable resource for the analysis of driver behaviour, while also highlighting opportunities for traffic management enhancements.

The utilization of videos as a data source offered a holistic and real-time perspective of traffic and pedestrian interactions. This method significantly mitigated observer bias, as it provided an objective and impartial data stream. Further, the video footage could be repeatedly reviewed for data extraction and analysis, ensuring the precision and trustworthiness of the findings. Additionally, video camera camcorders established at both the curb and the median as shown in Figure 3, enabled the identification of distinct vehicle categories and attributes, as well as the recording of pertinent pedestrian characteristics. This comprehensive data contributed to a more thorough investigation into traffic flow, pedestrian conduct, and broader transportation patterns.



Figure 3: Camera camcorder mounted at the curb

Drones, equipped with cameras, possessed the distinct advantage of covering a larger area and capturing concurrent activities. This aerial viewpoint provided supplementary context and insights into traffic behaviours, road network configurations, and pedestrian movements. In sum, the incorporation of video footage in the research process extended beyond initial data collection, enabling comprehensive, unbiased, and detailed scrutiny of traffic and pedestrian actions. This, in turn, has the potential to drive advancements in safety, efficiency, and transportation planning.

Data Extraction

Vehicle Velocity Data

The data extraction process during the planned video surveys is centred around capturing vehicle velocity patterns, particularly those of vehicles that yield to pedestrians. To ensure precision and reliability in depicting the velocity variation graph, it is imperative to take velocity measurements at multiple gaps. This approach allows for a more accurate representation of the velocity fluctuations exhibited by vehicles during yielding events. Additionally, vehicle velocity measurements are specifically obtained 80 meters upstream from the crossing point, as illustrated in Figure 6 there is an overhead crosswalk disturbing the view to the drone in one direction so the same upstream distance for either direction. This predetermined location serves the purpose of evaluating vehicle velocity immediately before their approach to the pedestrian crossing. The sufficiency of this 80-meter upstream location will be elaborated upon in subsequent sections.

The study aims to acquire precise and relevant velocity data, which in turn enables a comprehensive analysis of the velocity patterns and deceleration behaviours of vehicles near the

pedestrian crossing. The consistent measurement of vehicle velocity at various junctures ensures the collection of velocity patterns that accurately reflect the free flow velocity of vehicles up to the point of complete yielding to pedestrians. The precise tracking of vehicle velocity variations commences from the moment pedestrians enter the curb, allowing for the identification and examination of deceleration patterns exhibited by different vehicle types as they approach the crosswalk.

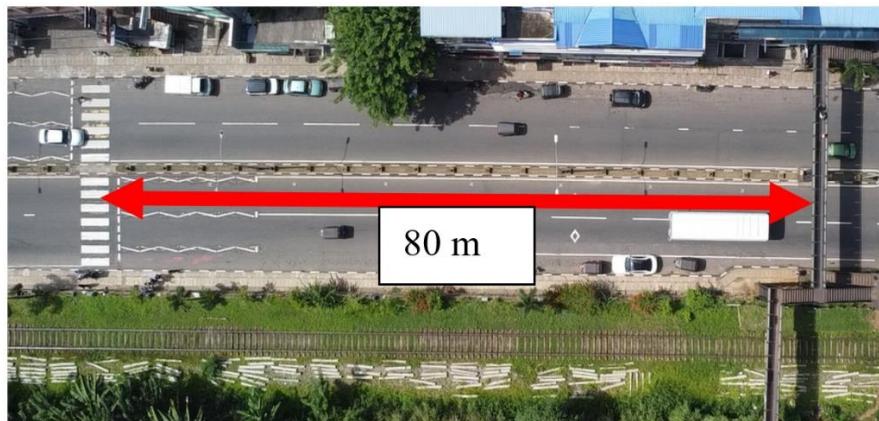


Figure 4: Upstream length and Overhead Crosswalk

The vehicle velocity variations were obtained using TRACKER 5.0, with the x-axis aligned along the median. The software was utilized to track a specific vehicle, automatically generating a Velocity (V_x /(m/s)) vs. Distance(x /(m)) graph. This approach provided a detailed representation of the vehicle's velocity patterns, shown in Figure 5, as it approached the pedestrian crossing and also we recorded the V_x vs x values separately. In this configuration, the x-axis represents the distance (x /(m)) where the vehicle moves along, as illustrated in Figure 6.

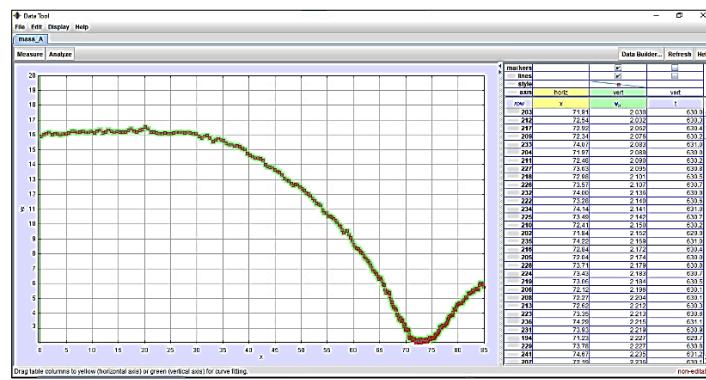


Figure 5: Velocity vs distance graph example

By employing these precise measurement points and distances, the study aimed to obtain accurate and relevant velocity data, facilitating a comprehensive analysis of velocity patterns and deceleration tendencies exhibited by vehicles near the pedestrian crossing. This data provided insights into the deceleration patterns of vehicles during yielding manoeuvres, contributing to a better understanding of the dynamics and behaviour of vehicles in pedestrian interactions.

Understanding these patterns contributed to enhancing pedestrian safety and informing traffic management strategies.

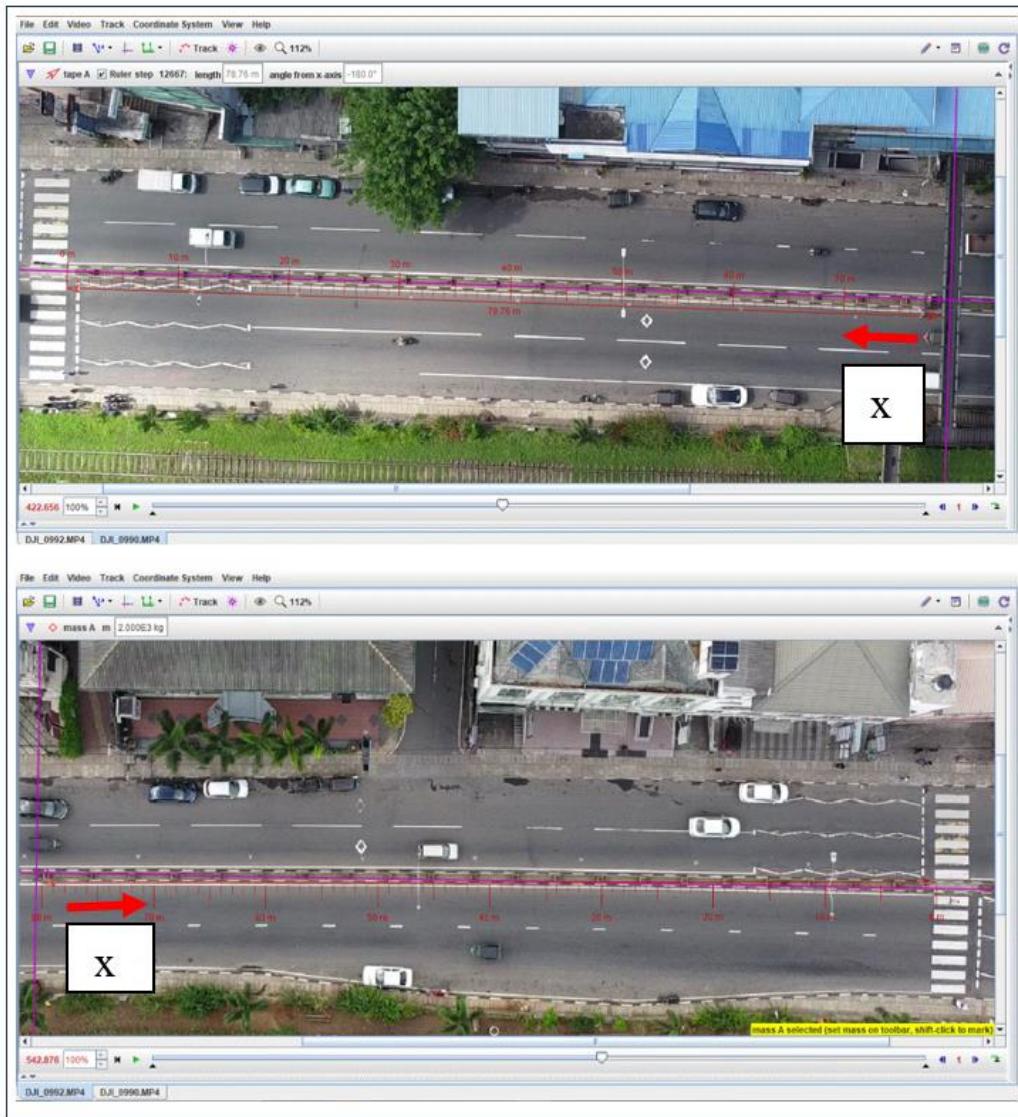


Figure 6: Axis markings in either direction

Vehicle and Pedestrian Characteristics Data

To achieve the second objective of identifying the pedestrian characteristics that influence driver-yielding behaviour, manual data entry was performed while referring to the video footage taken from the camera camcorders mounted beside the crosswalk. All the pedestrian data was collected upon selecting one vehicle to which the pedestrians crossed. These characteristics were taken as the independent variables to build the model. As the dependent variable, we considered whether the vehicle yielded or not to the crossing pedestrians. All the events were considered only when there was a pedestrian occurrence.

The independent data were tabulated as shown in Table 1, including characteristics such as the number of pedestrians crossing, the location of crossing marked in Figure 3.7 (starting to cross from the median or curb), the gender distribution of crossing pedestrians, the presence of disturbances (such as carrying a child or having an injury/disability), vehicle type, and the lane from which the vehicle approached the pedestrian/s, illustrated in Figure 7 as Near Lane (NL) and Far Lane (FL, the lane close to the median). As shown in Table 1, we have categorized the vehicles into six distinct buckets: Van, Bus, Lorry (all heavy vehicles), Car (including jeeps and cabs), Motorbike (MB) and Three-Wheeler (TW) for analysis and representation. These characteristics were considered as independent variables for building the model, with the dependent variable being whether the vehicle yielded to the crossing pedestrians. The determination of whether a vehicle yielded or not was made by examining the velocity variation graphs. A significant drop in velocity was considered an indicator of a Yielded (Y) event, while the absence of such a drop signified a Non-Yielded (NY) event. Data collection events were recorded only when pedestrians were present.

Table 1: Example of Recorded Dependent and Independent Variable

Yield (Y) or Not Yield (NY)	Vehicle Type	Vehicle Position	No Of Pedestrians at the;		Gender Distribution		Disturbances Count	
			Curb	Median	Male	Female	Injured	Children
Y	MB	NL	1	1	2	0	0	0
NY	Bus	FL	1	4	5	0	0	1
NY	Car	FL	0	2	0	2	0	0
NY	TW	NL	3	0	2	1	2	0
Y	Lorry	FL	3	0	2	1	0	0
NY	Van	NL	0	2	1	1	0	1



Figure 7: Vehicle Positions (NL/FL) and Pedestrian Positions (Curb/Median)

Data Analysis

Identify the yielding behaviour among various classes of vehicles

In our analysis, we observed that the velocity values obtained by TRACKER were not specific to certain gaps of distance values(x). To address this, we utilized PYTHON to calculate vehicle velocity (V_x) values corresponding to the same set of distances (x) from the collected data. We obtained Velocity vs Distance graphs for all vehicles as shown in Figure 8. By doing so, we determined the range where vehicle approaching velocity and vehicle yielding velocity occurred for each vehicle type. Subsequently, we created envelope graphs for each vehicle category to visualize this information better.

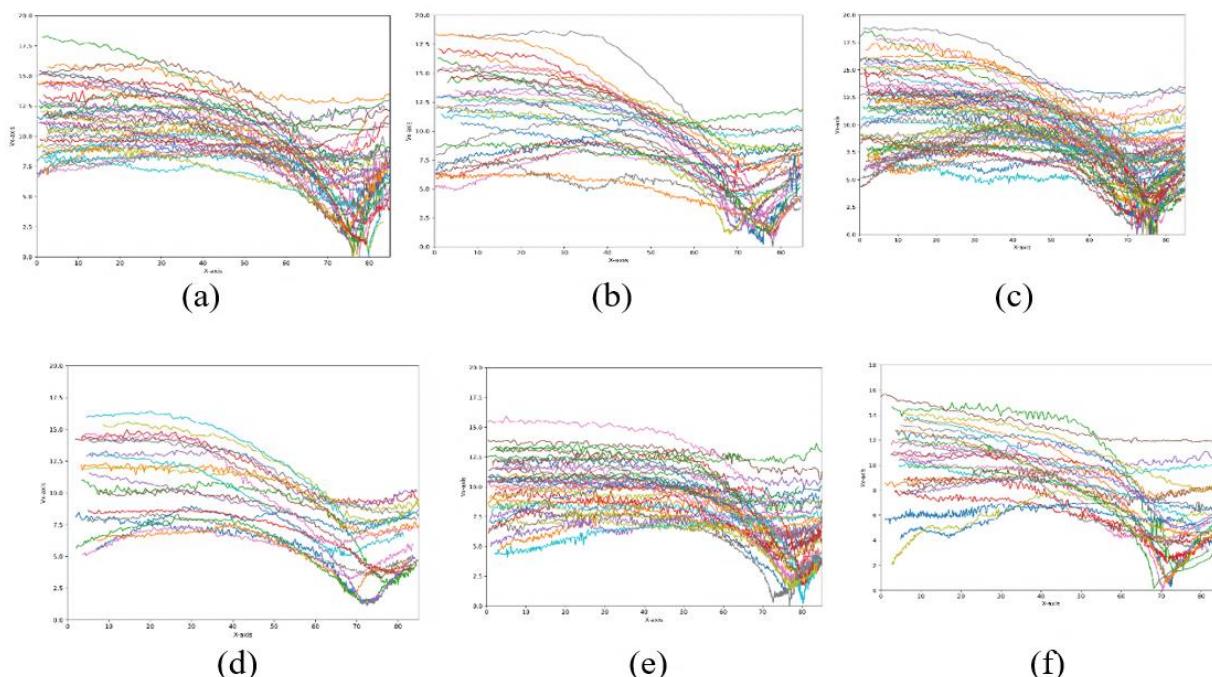


Figure 8: Velocity vs Distance Graphs (a) For Motor bicycles (b) For Vans (c) For Cars (d) For Lorries (e) For Three Wheelers (f) For Buses

Our analysis focused on examining the variations in vehicle velocity during the deceleration process. To achieve this, we plotted the data points and closely examined the emerging trends and patterns. Based on the results of our analysis, we constructed average vehicle deceleration curves for each vehicle category. These curves provide a clear representation of the relationship between the distance from the crosswalk and the corresponding vehicle velocities, shedding light on the deceleration behaviour of different types of vehicles. These deceleration curves serve as a useful basis for comparing various vehicle types. They offer valuable insights into how different vehicles behave during interactions with pedestrians. By analyzing these curves, we gain a deeper understanding of driver-yielding patterns at crosswalks. This knowledge is instrumental in the development of effective strategies and interventions aimed at improving pedestrian safety and optimizing traffic flow in these critical areas.

Identify the factors that influence drivers' yielding willingness

In our analysis of factors influencing drivers' yielding willingness, we considered several independent variables, including vehicle type, number of pedestrians at the median, number of pedestrians at the curb, number of female pedestrians, number of male pedestrians, number of child pedestrians, number of injured pedestrians (number of pedestrians having any disability/injury/differently abled from the total number of crossing pedestrians) and vehicle approaching lane. The primary focus was on the dependent variable whether the vehicle yielded or not.

To model this binary outcome (yes/no, 1/0), we employed logistic regression analysis to determine the probability of yielding. The logistic regression equation allowed us to assess the impact of these independent variables on the likelihood of vehicles yielding to pedestrians, providing valuable insights into the factors influencing driver behaviour at the selected crosswalk. The logistic regression equation that describes the model's intricate links is shown below.

$$P(z) = \frac{1}{1 + e^{-z}}$$

$$Z = a_0 + a_1 * \text{Number of pedestrians at the median} + a_2 * \text{Number of pedestrians at the curb} + a_3 * \text{Vehicle Type} + a_4 * \text{Vehicle Position} + a_5 * \text{Number of Male} + a_6 * \text{Number of Female} + a_7 * \text{Number of Injured} + a_8 * \text{Number of Children}$$

We conducted a stepwise regression analysis to identify the independent variables that exert a significant impact on the dependent variable. The results, obtained from the analysis of a dataset comprising 916 samples using R Studio software, provide valuable insights into the factors influencing driver-yielding willingness at a crosswalk, with the sample size carefully balanced through class balancing procedures.

RESULTS AND DISCUSSION

Yielding Patterns of Vehicle According to The Vehicle Class

The analysis reveals a consistent yielding behaviour across various vehicle types, characterized by a uniform and gradual decrease in average rate. Notably, no significant discrepancies were observed among the six distinct vehicle categories while looking at the average deceleration graph illustrated in Figure 9. This investigation demonstrates the yielding patterns in tandem with vehicle velocity, emphasizing the commonality in the deceleration patterns across the diverse spectrum of vehicles considered.

In this study, we examined vehicle deceleration patterns across six distinct categories including 40 motorcycles, 40 three-wheelers, 30 vans, 60 cars, 20 lorries, and 30 buses, providing a diverse and representative sample within each vehicle category. This comprehensive approach ensures a nuanced understanding of the yielding behaviour across various vehicle types. The following Figure 10 presents the outcomes of the vehicle deceleration patterns, offering insights into the unique characteristics exhibited by each vehicle class.

Factors That Influence the Yielding Behaviour

The investigation into the factors influencing drivers' yielding willingness, particularly about pedestrian characteristics, employed stepwise logistic regression analysis. The outcomes of this

analysis are detailed in Table 2. Utilizing a binary logistic regression approach, where 1 is assigned to yielded events and 0 to non-yielded events, was a logical choice for the study. Based on the outcomes of the regression analysis refinements can be made to the regression equation to enhance its predictive accuracy

Residual standard error: 0.2687 on 911 degrees of freedom

Multiple R-squared: 0.4288,

Adjusted R-squared: 0.4263 F-statistic: 171 on 4 and 911 DF, p-value: < 2.2e-16

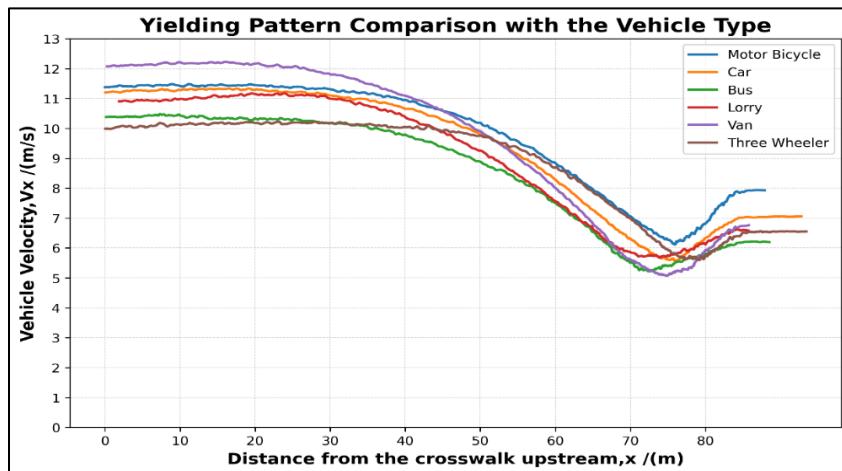
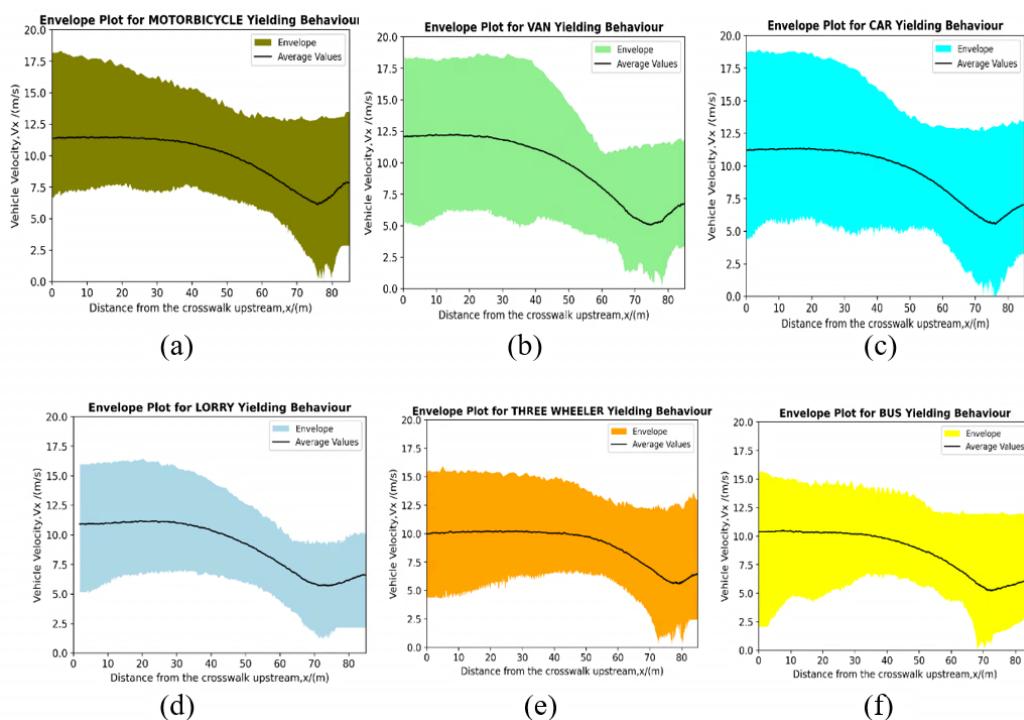


Figure 9: Average Velocity vs Distance graphs



**Figure 10: Average Velocity vs Distance graphs and Envelope (a) For Motor bicycles
(b) For Vans (c) For Cars (d) For Lorries (e) For Three Wheelers (f) For Buses**

Table 2: Results of the logistic regression analysis (after doing stepwise analysis)

	Estimate	t value	Pr (> t)
Intercept	0.685064	37.465	<2e-16 ***
Number of pedestrians at the median	0.016426	2.173	0.03 *
Number of female pedestrians	0.335825	12.872	<2e-16 ***
Number of injured pedestrians	-0.701490	-16.108	<2e-16 ***
Vehicle position (Near Lane)	0.218798	9.432	<2e-16 ***

The updated regression equation is as follows:

$$Z = 0.685064 + 0.016426 * \text{Number of pedestrians at the median} + 0.218798 * \text{Vehicle Position} \\ + 0.335825 * \text{Number of Female} - 0.701490 * \text{Number of Injured}$$

$$P(Z) = \frac{1}{1+e^{-(0.016426 * \alpha + 0.335825 * \beta - 0.701490 * \gamma + 0.218798 * \delta + 0.685064)}}$$

P(Z) = Probability of Yielding	γ = No of Injured Pedestrians
α = No of pedestrians in the median	δ = Vehicle Position
β = No of Female Pedestrians	1 ; at the Near Lane
	0 ; otherwise

In this revised equation, if the vehicle is approaching the pedestrian from the near lane, the substitution is 1; otherwise, it is 0. For other variables, the pedestrian count is directly substituted. This model generates values within the range of 1 to 0. This process provides the probability of vehicle yielding by substituting the obtained values. To ascertain whether a driver yields or not, a 50% threshold is employed. This logistic regression function estimates the probability (P) of a vehicle yielding at the crosswalk based on the specified independent variables. The sigmoid-shaped curve generated by this function facilitates a clear interpretation of the impact of each variable on the likelihood of yielding. The refined equation provides a more nuanced understanding of the complex interplay between pedestrian characteristics, vehicle attributes, and driver-yielding behaviour, thereby contributing to the development of targeted interventions for improving pedestrian safety. Enhancing the data collection process by expanding the sample size and including diverse sites with various weather conditions could have potentially improved the robustness and generalizability of the results. This approach would have provided a more comprehensive understanding of driver-yielding behaviour across different contexts, contributing to a more nuanced and reliable analysis.

CONCLUSION

This study explores yielding patterns at unsignalized pedestrian crossings, focusing on vehicle classes and their interactions with pedestrian characteristics. It found consistent and uniform yielding behaviour across six categories, including motorcycles, three-wheelers, vans, cars, lorries,

and buses. Despite the diversity in vehicle types, yielding patterns exhibited remarkable similarities, providing valuable insights for traffic management and pedestrian safety. The study also explored factors influencing driver-yielding willingness, particularly pedestrian characteristics, using stepwise logistic regression analysis. The results highlighted significant variables, including the number of pedestrians at the median, the number of female pedestrians, the number of injured pedestrians, and the vehicle position in the near lane. The refined logistic regression equation provided a predictive model for estimating vehicle-yielding probability based on independent variables. This model contributes to developing targeted interventions and strategies to enhance pedestrian safety and optimize traffic flow at unsignalized crosswalks.

REFERENCES

- Bella, F., and Ferrante, C. 2021. Drivers' Yielding Behavior in Different Pedestrian Crossing Configurations: A Field Survey'. *J Adv Transp* 2021.
- Bertulis, T., and Dulaski, D. M. 2014. 'Driver approach speed and its impact on driver yielding to pedestrian behaviour at unsignalized crosswalks'. *Transp Res Rec*.
- Dileep, R., Koshy, B. I., and Sam, E. 2016. 'Study on Driver Yielding to Pedestrians at Unsignalized Crosswalks'. *Int J Sci Eng Res* 7.
- Fu, T., Hu, W., Miranda-Moreno, L., and Saunier, N. 2019. 'Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data'. *Transp Res Part C Emerg Technol* 105, 222–240.
- Lobjois, R., and Cavallo, V. 2007. 'Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task'. *Accid Anal Prev* 39, 934–943.
- Schroeder, B. J., and Roushail, N. M. 2011. 'Event-Based Modeling of Driver Yielding Behavior at Unsignalized Crosswalks'. *J Transp Eng* 137, 455–465.
- Fu, T. 2018. *A novel approach to investigate pedestrian safety in non-signalized crosswalk environments and related treatments using trajectory data*.
- Yang, Y., Wang, Y., Easa, S. M., and Zheng, X. 2022. 'Analyzing Pedestrian Behavior at Unsignalized Crosswalks from the Drivers' Perspective: A Qualitative Study'. *Applied Sciences (Switzerland)* 12.

Comprehensive Analysis of Red Light Running in the United States: Strategies, Successes, and Future Directions

Saquib M. Haroon¹; Cipriano De Luna Gutierrez²; Sushmita Bhandari³; Alyssa Ryan⁴; and Henrick Haule⁵

¹Dept. of Civil and Architectural Engineering and Mechanics, Univ. of Arizona, Tucson, AZ.
Email: saquib@arizona.edu

²Dept. of Civil and Architectural Engineering and Mechanics, Univ. of Arizona, Tucson, AZ.
Email: cdelunagutierrez@arizona.edu

³Dept. of Civil and Architectural Engineering and Mechanics, Univ. of Arizona, Tucson, AZ.
Email: sushmitab@arizona.edu

⁴Dept. of Civil and Architectural Engineering and Mechanics, Univ. of Arizona, Tucson, AZ.
Email: alyssaryan@arizona.edu

⁵Dept. of Civil and Architectural Engineering and Mechanics, Univ. of Arizona, Tucson, AZ.
Email: hhaule@arizona.edu

ABSTRACT

This study provides a comprehensive review regarding the issue of red-light running in the US, highlighting its significant impact on road safety. Through a thorough examination of pertinent literature, the research sheds light on the challenges posed by red-light running and explores diverse strategies adopted by major US cities to curb this behavior, including through the deployment of red-light cameras. This study discusses the effectiveness of measures like red-light cameras in reducing intersection crashes, notably right-angle collisions, despite potential concerns like a slight increase in rear-end collisions. Emphasizing a comprehensive approach, it reviews punitive measures for red light running across various US cities. Additionally, the research investigates Vision Zero action plans in different cities, examining their strategies for reducing red light running crashes. By providing these valuable insights, this study aims to guide researchers, practitioners, and policymakers in developing holistic solutions to address and mitigate the challenges associated with red-light running.

INTRODUCTION

The act of running red lights presents a substantial hazard to road safety, posing a notable risk to the well-being of drivers and other vulnerable road users (Hussain et al., 2020). Red lights on signal displays are designed to control traffic and prevent collisions at intersections, where multiple lanes of traffic cross paths. When drivers "run" red lights, they increase the chances of causing a collision with another vehicle, a pedestrian, or cyclist (Retting et al., 1999). In 2021, 1,109 people were killed in crashes that involved red light running while 127,000 were injured in the United States (*Red Light Running*, 2023). Recognizing the pressing need to address this issue, many cities have implemented safety programs to mitigate the impact of such incidents (Lee et al., 2014). A prevalent initiative gaining traction is "Vision Zero," a strategy aimed at eliminating all traffic fatalities and severe injuries while promoting safe, healthy, and equitable mobility. This strategy acknowledges that eliminating all injuries may not be possible but rather recognizes that severe injuries and fatalities can be greatly reduced. This strategy sets goals and expectations for the cities that are interested in applying it (Ferenchak, 2023).

In addition to the inherent physical risks, red light running entails significant legal consequences, constituting a traffic violation that may lead to fines, points on the driver's license, and potential license suspension (Baratian-Ghorghi et al., 2016). It can also increase a driver's insurance premiums and make it more difficult to obtain auto insurance in the future (Tay & de Barros, 2011). Notably, not all red-light running results in legal action, which may result in a change to the driver's behavior in the future. Due to limited resources, it is impossible to have an enforcement officer stationed at every intersection at every hour to respond to potential violators. In response to this problem, many jurisdictions have decided to install red light cameras as a solution to mitigate red light running (Llau & Ahmed, 2014). These cameras, strategically connected to traffic signals and equipped with sensors monitoring traffic flow, capture images of vehicles proceeding through red lights. Positioned just before crosswalks or stop lines, the cameras provide a snapshot of any vehicle failing to halt during the red phase (Retting et al., 2008). According to the Insurance Institute for Highway Safety, red light camera programs generally provide motorists with a grace period of about half a second after the light switches to red (*Red Light Running*, 2023). Based on the automated enforcement system in use, photographic evidence capturing the violating driver is either electronically or physically transmitted to a photo-processing center. At this center, each photo is scrutinized by a police officer. Should the officer determine that a moving violation has taken place, the owner of the implicated vehicle will receive a moving citation, municipal citation, or traffic citation via mail, contingent on the specific legislation of the state (Fitzsimmons et al., 2007).

Although red light cameras provide the appropriate countermeasures in reducing red light running, many citizens and jurisdictions do not agree with its implementation. Citizens have assumed red light cameras are mainly used as a revenue generation mechanism by the government and have raised concerns regarding privacy (Yang et al., 2013). Others have worried that it may lead to racial profiling if it is not automatically enforced (Eger et al., 2015). Lastly, researchers have noticed that although right angle crashes have decreased, the amount of rear end crashes has been on the rise at locations with these systems (Retting et al., 2003). Due to this backlash, many states have opted to not use red light cameras. It's essential to emphasize that fundamental rights are not infringed upon, as underscored by the Seventh Circuit, which emphasized the absence of an inherent right to run a red light or avoid surveillance by a public street camera. Consequently, when evaluating the constitutionality of red-light camera regulations, courts apply the rational-basis test, providing constitutional support for red light camera programs. (Letizia, 2009).

Given the current lack of comprehensive studies on red light running in the U.S., this research takes a multi-faceted approach to assess this issue. The primary goal of this study is to gain an in-depth understanding of how cities perceive and address red light running issues, assess the effectiveness of different practices in mitigating red light violations, and scrutinize the implementation and success of these strategies across the U.S. Key considerations include whether red light running prevention is integrated into city planning, the utilization of relevant studies, existing red light camera policies, alternative prevention methods, and the comparative effectiveness of these approaches. The study's primary goals are as follows:

1. Conduct a comprehensive review of the outcome of implementing red light running cameras and other countermeasures across different cities in the United States.
2. Examine the penalty measures implemented by these cities in response to red light violations, providing insights into the enforcement mechanisms in place.

3. Analyze the future safety goals related to red light running by investigating the safety programs and Vision Zero initiatives embraced by 15 major cities across United States.

By comprehensively reviewing penalty measures and future safety goals, the study aims to contribute actionable data for enhancing road safety practices. Ultimately, these findings have the potential to inform evidence-based decision-making, improve existing practices, and contribute to the overarching goal of minimizing road accidents and ensuring safer urban mobility.

RED LIGHT RUNNING PREVENTION: STRATEGIES AND OUTCOMES

Red light running cameras are currently used in 22 states and in Washington, D.C. (*Red Light Running*, 2023). Several cities have employed red light running cameras along with other countermeasures to reduce the impact of red light running. The effectiveness of these countermeasures is summarized below.

The city of Chicago conducted a two-stage safety analysis where crash frequency and the frequency of red-light camera violations were analyzed. Notably, the implementation of red-light cameras led to a 19% decrease in angle and turn crashes but also resulted in a 14% increase in rear-end collisions at the targeted intersections. Overall, it was found that there was a reduction in crashes in general at these locations by 10%. It was noted that intersections with higher traffic volume, more approach lanes, higher speed limits, wider intersections, longer cycles, not allowing left turns, and longer all-red phases were found to have a higher frequency of red-light running incidents. It was also found that intersections which had presence of a physical median, a “No turn on red” sign, a longer yellow phase duration, and a protected left bay decreased red light camera violation frequency (*Chicago Red Light Camera Enforcement*, 2017).

New York City's utilization of red-light cameras has yielded positive outcomes, as evidenced by the New York City Department of Transportation's thorough review of the program. Comparing collected data to previous years, they observed a notable 80% decrease in right-angle crashes and a 61% reduction in rear-end collisions at intersections equipped with red light cameras since the program's inception in 1994. Despite these achievements, the incidence of red light running escalated to seven violations per camera day in 2021, up from 4.79 in 2020. To address this surge, New York City is contemplating the implementation of escalated fines for repeat offenders (*New York City Red Light Camera Program Review - 2022 Report*, 2022).

The city of Philadelphia initiated a red-light running camera program in 2005 and has seen a reduction in red light running in all intersections where the red-light running camera has been installed. As of 2021, red light running cameras have been installed in thirty-two intersections. The location in which the first red light running camera was installed saw a 55% decrease in red light violation from 2006 to 2020. Similar results were observed in almost all intersections. However, the city of Philadelphia saw a 10.47% of increase in overall violation in 2020 compared to 2019 which has been attributed to behavioral changed during the COVID-19 pandemic (Petri, 2021).

California has employed diverse strategies in red light running prevention. San Francisco Municipal Transportation Agency (SFMTA) analyzed the effectiveness of San Francisco's red light camera program based on current and past collision trends. Between the years 1995 to 2018, they discovered that their red-light running crashes were a third of what they were in 1995. One of the contributing factors to the improvement in collision prevention was the signal visibility upgrades they made. SFMTA found these helped prevent an average of 52% of

collisions in the four intersections studied (SFMTA Board of Directors Policy and Governance Committee, 2019). In a unique approach, the City of San Jose utilized "rat boxes" instead of red-light cameras, allowing police officers to position themselves strategically for enforcement (San Jose Police Department, n.d.).

Table 1: Overview of the effectiveness of red-light running countermeasures in major US cities.

City	Counter measure	Effect
Chicago	RLR Camera	19% decrease in angle crashes; 14% increase in rear-end crashes
New York City	RLR Camera	19% decrease in angle crashes; 61% decrease in rear-end crashes
Philadelphia	RLR Camera	55% decrease in red light running violations
San Francisco	Signal Visibility Increase	52% decrease in red light running collisions
San Jose	Rat Boxes	31% reduction in injuries due to red light running
Phoenix	RLR Camera	14% decrease in angle crashes; 20% increase in rear-end crashes

The Florida Highway Safety and Motor Vehicles conducted a thorough examination of the impact of red-light camera programs in cities across the state. The findings revealed a 4.69% reduction in instances of red light running following the implementation of the camera program. While the overall results indicated a decrease in angle crashes, there was a concurrent increase in rear-end crashes (Terry L Rhodes, 2022). In the examination of red-light running cameras in the City of Phoenix through the comparison group method, the assessment revealed a 42% decline in angle crashes and a 10% reduction in left-turn crashes across all approaches. Conversely, there was a notable 20% upswing in rear-end crashes observed across all approaches, with particularly pronounced effects on target approaches (Shin & Washington, 2007).

The Texas Traffic Operations Division conducted a study examining the effects of red-light cameras, uncovering a substantial reduction of 20% in all crash types and 24% in right-angle crashes. However, the study also noted a significant 37% increase in rear-end crashes (Ko et al., 2013). Despite these advantages, the State of Texas implemented a ban on red light cameras in 2019. Nonetheless, specific communities continue to utilize them through existing contracts (Aldossari et al., 2023).

RED LIGHT RUNNING PREVENTION: PENALTIES

Enforcement for red light running involves both in-person enforcement as well as camera enforcement. There is no clear distinction in how the enforcement differs. Camera enforcement involves the use of surveillance cameras to capture instances of vehicles running red lights at intersections. Authorities then review the recorded footage to impose fines or penalties on violators. Conversely, in-person enforcement relies on human intervention by law enforcement officers or authorized personnel. Although the literature does not identify clear differences between the fines associated with in person enforcement. The search for information on penalties for red light running involved examining specific cities. However, it is essential to recognize that

not all cities provided publicly searchable data on these penalties. The accessibility of such information may differ by jurisdiction, and the absence of searchable data in certain cases could be attributed to factors such as local policies, data availability, or discrepancies in reporting practices among municipalities.

Cities across the United States have embraced the use of red-light camera systems to enhance traffic safety, each implementing distinct violation fees, enforcement practices, and potential implications on driving records. In Portland, Oregon, running a red light triggers a \$265 fine, categorized as a class B violation. The presence of graduated fines remains uncertain, but eligible drivers may opt for dismissal through traffic school (*Portland.Gov*, 2023).

Seattle, Washington handles red light violations through photo enforcement, incurring a \$139 fine categorized as a parking infraction. Importantly, this infraction does not tarnish the driver's license with points or impact insurance (David Kroman, 2023).

Chicago, Illinois levies a \$100 fine for red light camera tickets, not classifying them as moving violations (David Kidwell and Alex Richards, 2014). This distinction spares drivers from accruing points on their records or facing heightened insurance rates. In more severe instances, however, 20 demerit points and potential reckless driving charges may be applied (Brian J Mirandola, 2023). While reports mention graduated fines, specifics remain unclear.

In New York City, a \$50 fine is imposed for red light camera violations, and vehicle owners are held responsible regardless of the driver (Baratian-Ghorghi et al., 2016). Encouragingly, such violations do not result in points being added to the driver's record, though details on graduated fines remain elusive.

Table 2: Penalties imposed by major cities for red light running and their effect to driving record.

City	Violation fees	Effect to Driving Record
Portland	\$265	No effect on points in driving record
Seattle	\$139	No effect on points in driving record
Chicago	\$100	No effect on points in driving record
New York City	\$50	No effect on points in driving record
Houston	\$155	2 points
San Antonio	\$285	2 points
San Francisco	\$490	1 point
Los Angeles	\$490	1 point

Houston, Texas faced a red-light camera ban in 2019 because of statewide red light camera ban in Texas, but the city of Humble in Houston continues to operate red light cameras until June 2024. Running a red light in Texas is considered a Class C misdemeanor, resulting in two points on the offender's driving record. Notably, there is no information available regarding graduated fines (Jessica Zimmer, 2023). In San Antonio, Texas, a \$285 fine is enforced for red light violations, with additional points added to the driving record based on the severity of accidents. Accumulating more than six points may incur surcharges (Bexar County, 2023).

On the West Coast, both San Francisco and Los Angeles in California impose a \$490 fine for red light violations, categorizing them as moving violations with one point added to the driver's record. While information on graduated fines is lacking, these violations may trigger heightened insurance premiums (Baratian-Ghorghi et al., 2016).

RED LIGHT RUNNING PREVENTION: VISION ZERO

Vision Zero is based on the fundamental belief that everyone deserves safe mobility, and it underscores the shared responsibility of planners, engineers, and policymakers in ensuring the safety of all road users (Belin et al., 2012). However, the methods to achieve Vision Zero vary widely, and the techniques promoted by national Vision Zero organizations, like incorporating human fallibility and adopting a systemic approach, are broad and susceptible to diverse interpretations (Evenson et al., 2023). This study investigated the safety plans for the 15th largest city in terms of population in the United States to understand if these cities integrate vision zero goals in their future safety plans,

In this study, an analysis of various cities was conducted to explore their incorporation of Vision Zero goals into future safety plans. The cities under consideration include Phoenix, AZ; Los Angeles, CA; San Diego, CA; San Jose, CA; San Francisco, CA; Jacksonville, FL; Chicago, IL; Indianapolis, IN; New York City, NY; Columbus, OH; Philadelphia, PA; Houston, TX; San Antonio, TX; Dallas, TX; and Austin, TX. Among the cities analyzed, all but two—Indianapolis and Jacksonville—incorporate Vision Zero goals into their future safety plans. In contrast to the state-wide Vision Zero goal in Indiana, where the state expresses its aim for zero deaths and serious injuries (INDOT, 2022), it's important to clarify that Indianapolis, among the cities analyzed, does not have its own Vision Zero plan. Meanwhile, Jacksonville, located in Florida, is in the process of developing its own Vision Zero action plan (Scanlan, 2023). Notably, cities like Phoenix, San Diego, San Jose, San Francisco, New York, and Philadelphia address red light running as a safety concern within their Vision Zero plans. However, the specific countermeasures vary, ranging from documented action plans to combat red light running, increased camera programs, and targeted hotspot enforcement.

Conversely, some cities, despite embracing Vision Zero, do not explicitly address red light running in their action plans. Dallas acknowledges red light running as a major factor in severe crashes but lacks specific measures to reduce it (City of Dallas, 2022). Austin initially set plans to increase red-light camera programs (City of Austin, 2016) but faced challenges following the 2019 red light camera ban in Texas, resulting in a lack of reported data and action plans. Houston commits to enhancing penalties for red light running violations as part of its Vision Zero goals (City of Houston, 2020), while San Antonio does not outline any specific actionable plans for addressing red light running within its safety framework (City of San Antonio, 2016). Chicago, while pledging for safer streets in its Vision Zero initiative, does not provide a detailed action plan to combat red light running (City of Chicago, 2021). Table 3 provides a summary of the red-light running action plans identified in various Vision Zero initiatives across different cities.

SUMMARY AND DISCUSSION

The study provides an overview of red-light running prevention strategies predominantly focusing on red light cameras as a countermeasure. The effectiveness of red-light camera programs varies, with each city experiencing different outcomes in terms of reducing angle crashes, rear-end collisions, and overall violations.

Examining the case studies of Chicago, New York City, Philadelphia, San Francisco, Phoenix, and Texas provides insight into the effects of red-light camera programs. Chicago's study emphasizes a decrease in angle crashes but notes an increase in rear-end collisions, underscoring the intricate nature of the problem. A recurring observation is that the

implementation of red-light cameras consistently leads to a substantial decrease in right-angle crashes, which are often more severe and result in serious injuries. However, this positive impact is frequently accompanied by a trade-off: an increase in rear-end collisions. It is crucial to note that the Vision Zero initiative aims specifically at reducing serious injuries, not merely overall crash occurrences. Given that right-angle crashes tend to be more injurious (Jin et al., 2010), the effectiveness of red-light camera programs becomes evident in contributing to the Vision Zero goal by mitigating the severity of accidents and enhancing overall road safety. New York City's sustained success in reducing both right-angle and rear-end collisions over the long term highlights the potential advantages of consistently applied red light camera programs. Philadelphia's experience reinforces the positive influence of red-light cameras in most intersections, clearly demonstrating a significant reduction in violations.

Table 3: Vision Zero adoption across cities and action plans for red light running prevention.

City	Vision Zero Action Plan	Plan for reducing red light running	Action type
Phoenix, AZ	✓	✓	Awareness and Education Campaign
Los Angeles, CA	✓	✗	Acknowledged red light running as concern
San Diego, CA	✓	✓	Targeted hotspots enforcement
San Jose, CA	✓	✓	Run targeted enforcement
San Francisco, CA	✓	✓	Increase red light running camera program
Jacksonville, FL	✗	-	
Chicago, IL	✓	✗	Pledge to stop at red lights
Indianapolis, IN	✗	-	
New York City, NY	✓	✓	Increase red light running camera program, Work with state to solve legislation issues
Columbus, OH	✓	✗	
Philadelphia, PA	✓	✓	Increase red light running program at high crash locations
Houston, TX	✓	✗	Analyze data and enhance penalty
San Antonio, TX	✓	✗	
Dallas, TX	✓	✗	Acknowledges red light running as an issue. No action plan
Austin, TX	✓	✗	Increase red light running camera program. (No mention after 2019 ban)

This review also provides an insight into other countermeasures applied by other cities. San Francisco's use of strategies such as signal visibility enhancement and the use of rat boxes by

city of San Jose may be adopted by cities as countermeasures especially in states where red light camera programs are banned by state legislature. The issue of red light camera program tends to be more political with communities believing these programs are cash cows for cities and with rights organizations like American Civil Liberties Union (ACLU) often backing bills to ban traffic cameras (Irwin David, 2018).

The examination of penalties reveals the disparities in fines and their consequences on driving records across various cities. Portland's \$265 fine and Seattle's \$139 fine, both without repercussions on driving records, highlight the range of penalty structures in place. This divergence is frequently shaped by the classification of camera citations as civil offenses rather than moving violations in many states. In such instances, driver license points are usually not assigned, and there are no associated insurance implications. It's noteworthy that certain states treat red light safety camera citations on par with those issued by police officers during traffic enforcement, adding a layer of complexity to the regulatory landscape (*Red Light Running*, 2023).

The integration of Vision Zero initiatives into the conversation introduces an additional layer of intricacy. While most cities align with the Vision Zero philosophy, the strategies employed to tackle red light running differ significantly. Some cities explicitly include red light running in their action plans, outlining specific measures, while others acknowledge it as a contributing factor to overall road safety without providing detailed strategies. Despite red light running being a significant contributor to severe injuries and fatalities, it often occupies a lower priority in the action plans of most major cities.

CONCLUSION

The research investigates a critical safety concern—red light running—across various cities in the United States. It conducts a comprehensive review of prior studies that assess the efficacy of red-light running programs implemented in different cities across the country. The findings consistently indicate promising results, especially in reducing severely injurious crash types, such as right-angle crashes, when red light camera programs are introduced. However, the acceptance of these programs within different communities becomes a notable challenge. Moreover, the issue often takes on a political dimension in several states, with advocacy groups opposing the implementation of camera programs.

Additionally, the study examines the Vision Zero action plans of the largest cities in the U.S. While there is a unanimous acknowledgment among major cities that red light running poses a significant safety concern, the Vision Zero action plans generally lack specific strategies to address this issue. The limited countermeasures for red light running primarily involve the implementation of camera programs in cities where action plans exist. The study highlights the pressing need for innovative solutions that are not only acceptable to the community but are also effective in reducing red light running violations. These strategies should be integrated as actionable plans within the Vision Zero initiatives of various cities, emphasizing the importance of comprehensive and community-supported approaches to enhance road safety.

REFERENCES

- Aldossari, M., Bandara, N., and Al-Werfalli, D. (2023). *Implications of Resistance to Automated Speed Enforcement and Red-Light Camera Implementation*. 1–9.

- Baratian-Ghorghi, F., Zhou, H., and Zech, W. C. (2016). Red-light running traffic violations: A novel time-based method for determining a fine structure. *Transportation Research Part A: Policy and Practice*, 93, 55–65.
- Belin, M.-Å., Tillgren, P., and Vedung, E. (2012). Vision Zero – a road safety policy innovation. *International Journal of Injury Control and Safety Promotion*, 19(2), 171–179.
- Bexar County. (2023). Ticket Fine Schedule | Bexar County, TX - Official Website. <https://www.bexar.org/3165/Ticket-Fine-Schedule>.
- Brian J Mirandola. (2023, September 12). What is a Red Light Camera Ticket and How Can I Fight It? | IL. What Is a Red Light Camera Ticket and How Can I Fight It? | IL. <https://www.mirandalacriminaldefense.com/elgin-criminal-defense-lawyer/what-is-a-red-light-camera-ticket-and-how-can-i-fight-it>.
- Chicago Red Light Camera Enforcement: Northwestern University Transportation Center - Northwestern University. (2017, March 17). <https://transportation.northwestern.edu/research/featured-reports/chicago-red-light-camera-report.html>.
- City of Austin. (2016). 2016-2018 Action Plan. <https://www.austintexas.gov/department/vision-zero>.
- City of Chicago. (2021). Vision Zero Chicago Downtown Action Plan. https://www.chicago.gov/content/dam/city/depts/cdot/CDOT%20Projects/VisionZero/2022/VZDT_PlanDocument_online.pdf.
- City of Dallas. (2022). Vision Zero Action Plan. [https://dallascityhall.com/departments/transportation/Documents/FINAL-Vision%20Zero%20Action%20Plan%20\(high%20res\).pdf](https://dallascityhall.com/departments/transportation/Documents/FINAL-Vision%20Zero%20Action%20Plan%20(high%20res).pdf).
- City of Houston. (2020). Houston Vision Zero Action Plan. https://www.houstontx.gov/visionzero/pdf/VZAP_Final%20Report.pdf.
- City of San Antonio. (2016). Vision Zero San Antonio. <https://www.visionzerosa.com/Portals/38/Images/Resources/Vision-Zero-Action-Plan.pdf>.
- Kidwell, D., and Richards, A. (2014, July 18). Red light cameras tag thousands for undeserved tickets. <https://www.chicagotribune.com/news/ct-red-light-camera-ticket-spikes-met-20140717-story.html>.
- Kroman, D. (2023, October 9). Why Seattle can't collect on \$4.3 million worth of traffic tickets. The Seattle Times. <https://www.seattletimes.com/seattle-news/transportation/100000-expired-traffic-tickets-take-bite-out-of-seattles-safety-budget/>.
- Eger, R. J., Fortner, C. K., and Slade, C. P. (2015). The Policy of Enforcement: Red Light Cameras and Racial Profiling. *Police Quarterly*, 18(4), 397–413.
- Evenson, K. R., LaJeunesse, S., Keefe, E., and Naumann, R. B. (2023). Mixed-methods approach to describing Vision Zero initiatives in United States' municipalities. *Accident Analysis & Prevention*, 184, 107012.
- Ferenchak, N. N. (2023). U.S. Vision Zero Cities: Modal fatality trends and strategy effectiveness. *Transportation Letters*, 15(8), 957–968.
- Fitzsimmons, E., Hallmark, S., McDonald, T., Orellana, M., and Matulac, D. (2007). The Effectiveness of Iowa's Automated Red Light Running Enforcement Programs. Final Report. Center for Transportation Research and Education Iowa State University.
- Hussain, Q., Alhajyaseen, W. K. M., Brijs, K., Pirdavani, A., and Brijs, T. (2020). Innovative countermeasures for red light running prevention at signalized intersections: A driving simulator study. *Accident Analysis & Prevention*, 134, 105349.

- INDOT. (2022, July 25). Strategic Highway Safety Plan (SHSP). INDOT. <https://www.in.gov/indot/traffic-engineering/traffic-safety-office/strategic-highway-safety-plan-shsp/>.
- David, I. (2018, January 20). Legislation to Ban Traffic Cameras Creates Odd Political Alliances. <https://www.planetizen.com/news/2018/01/96800-legislation-ban-traffic-cameras-creates-odd-political-alliances>.
- Zimmer, J. (2023, March 1). How to Fight a Red Light Camera Ticket in Texas | Legal Beagle. <https://legalbeagle.com/6386670-fight-light-camera-ticket-texas.html>.
- Jin, Y., Wang, X., and Chen, X. (2010). Right-Angle Crash Injury Severity Analysis Using Ordered Probability Models. *2010 International Conference on Intelligent Computation Technology and Automation*, 3, 206–209.
- Ko, M., Geedipally, S. R., and Walden, T. D. (2013). Effectiveness and Site Selection Criteria for Red Light Camera Systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2327(1), 53–60.
- Lee, Y., Li, Z., Zhang, S., Roshandeh, A. M., Patel, H., and Liu, Y. (2014). Safety impacts of red light running photo enforcement at urban signalized intersections. *Journal of Traffic and Transportation Engineering (English Edition)*, 1(5), 309–324.
- Letizia, K. (2009). The Seventh Circuit Gives the Green Light to Red Light Cameras: An Analysis of the Court's Application of the Rational-Basis Test to Red Light Camera Laws. *Seventh Circuit Review*, 4(2), 338.
- Llau, A. F., and Ahmed, N. U. (2014). The Effectiveness of Red Light Cameras in the United States—A Literature Review. *Traffic Injury Prevention*, 15(6), 542–550.
- New York City Red Light Camera Program Review—2022 Report. (2022).
- Petri, S. A. (2021). RED LIGHT PROGRAM Annual Report. <https://philapark.org/red-light-cameras/>.
- Portland.gov. (2023, October 13). <https://www.portland.gov/transportation/vision-zero/safety-cameras>.
- Red light running. (2023). IIHS-HLDI Crash Testing and Highway Safety. <https://www.iihs.org/topics/red-light-running>.
- Retting, R. A., Ferguson, S. A., and Farmer, C. M. (2008). Reducing red light running through longer yellow signal timing and red light camera enforcement: Results of a field investigation. *Accident Analysis & Prevention*, 40(1), 327–333.
- Retting, R. A., Ferguson, S. A., and Hakkert, A. S. (2003). Effects of Red Light Cameras on Violations and Crashes: A Review of the International Literature. *Traffic Injury Prevention*, 4(1), 17–23.
- Retting, R. A., Ulmer, R. G., and Williams, A. F. (1999). Prevalence and characteristics of red light running crashes in the United States. *Accident Analysis & Prevention*, 31(6), 687–694.
- San Jose Police Department. (n.d.). Traffic Enforcement Unit. Retrieved December 4, 2023, from <https://www.sjpd.org/about-us/organization/bureau-of-field-operations/traffic-enforcement-unit>.
- Scanlan, D. (2023, February 7). Feds help Jacksonville try to end bicyclist and pedestrian deaths. Jacksonville Today. <https://jaxtoday.org/2023/02/07/feds-help-jacksonville-end-bicyclist-and-pedestrian-deaths/>.

- SFMTA Board of Directors Policy and Governance Committee. (2019, October 22). Red light running collisions and the red light camera program.
https://www.sfmta.com/sites/default/files/reports-and-documents/2019/10/10-22-19_red_light_camera_presentation.pdf.
- Shin, K., and Washington, S. (2007). The impact of red light cameras on safety in Arizona. *Accident Analysis & Prevention*, 39(6), 1212–1221.
- Tay, R., and de Barros, A. (2011). Should traffic enforcement be unpredictable? The case of red light cameras in Edmonton. *Accident Analysis & Prevention*, 43(3), 955–961.
- Rhodes, T. L. (2022). Red Light Camera Summary Report.
<https://www.flhsmv.gov/resources/cabinet-and-legislature-reports/red-light-camera-program-analysis/>.
- Yang, Q., Han, L. D., and Cherry, C. R. (2013). Some measures for sustaining red-light camera programs and their negative impacts. *Transport Policy*, 29, 192–198.

Assessing the Influence of Various No-Right-Turn-on-Red Signs: A Driving Simulator Study

Parisa Masoumi¹; Abolfazl Taherpour²; Mansoureh Jeihani, Ph.D.³; Anam Ardestiri, Ph.D.⁴;
and Samira Ahangari, Ph.D.⁵

¹Ph.D. Student, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: pamas2@morgan.edu

²Ph.D. Student, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: abtah3@morgan.edu

³Dept. of Transportation and Urban Infrastructure, National Transportation Center, Morgan State Univ., Baltimore, MD. Email: mansoureh.jeihani@morgan.edu

⁴Traffic and ITS Engineer, Mead & Hunt, Inc., Columbia, MD. Email: anam.ardeshiri@meadhunt.com

⁵Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: samira.ahangari@morgan.edu

ABSTRACT

This study explores how “No Turn on Red” (NTOR) signs impact driver compliance and behavior at signalized intersections using a driving simulator. Data from 67 participants, navigating three scenarios (no sign, static sign, and dynamic sign), were analyzed to assess driver performance. The primary aim was to understand driver behavior concerning various NTOR signs, evaluating their speed profiles when approaching and turning at signalized intersections. Statistical and compliance analyses were conducted to assess driver performance. Results indicated that male drivers exhibited lower speeds in the presence of dynamic signs, suggesting heightened attention and compliance. Compliance analysis revealed a higher adherence rate in the dynamic sign scenario, emphasizing the effectiveness of its clear and easily understandable message in influencing driver behavior. These findings have implications for road safety, NTOR sign design, and educational strategies, offering valuable insights for future research and decision-making in these domains.

Keywords: Safety, NTOR Sign, Driver Behavior, Driving Simulator, Signalized Intersection

INTRODUCTION

Careful consideration is necessary when accommodating motor vehicles making turns at signalized intersections. Right-turn movements play a crucial role, and the appropriate design of roadway signage significantly impacts the efficiency and safety of road users. Turning vehicles are the primary collision risk for vulnerable road users. Ensuring the safety of right-turn movements at signalized intersections is therefore of paramount importance on improving traffic safety. Improving the operation of right-turn movements involves various factors such as the geometric layout of roads, signage, and traffic signals. In the 1970s, many states implemented the Right-Turn-On-Red (RTOR) policy at signalized intersections as an energy-saving measure aimed at reducing fuel consumption by vehicles in the right turn lane (P. L. Zador, 1984), reducing signal delay, and enhancing traffic flow efficiency. This policy allows motorists to

make a right turn after coming to a complete stop at the intersection and yielding to approaching traffic/pedestrian when the traffic signal is red, unless prohibited by signage. Control strategies permitting RTOR are widely implemented due to their numerous advantages, such as reducing delays, emissions, and fuel consumption, improving level of service (LOS), and increasing capacity at signalized intersections (Compton et al., 1994). However, studies have indicated that implementing RTOR may lead to an increase in crash rates (approximately 23%) and other safety concerns, particularly conflicts with cyclists and pedestrians (Abu-Lebdeh et al., 1997; Compton et al., 1994; Parker et al., 1975; Wagoner, 1992; P. Zador et al., 1982). Pedestrians and cyclists approaching from the right side of drivers are particularly vulnerable to crashes, as RTOR drivers tend to focus on left-side crossing traffic while searching for a gap to make their right turn, potentially overlooking pedestrians and cyclists on their right (Preusser et al., 1981).

Currently, drivers are allowed to make right turns on red at many intersections. However, this permission is restricted at certain intersections, leading to a confusion. Consequently, researchers have focused on evaluating the effectiveness of traffic signs in influencing driver behavior and enhancing safety (Pulugurtha et al., 2010; Zegeer & Cynecki, n.d.). Notably, studies have shown a higher violation rate for No Turn On Red (NTOR) signs compared to other signs, prompting researchers to propose a new NTOR sign design aimed at improving intersection safety (Podany et al., 2004). Additionally, other studies have suggested the use of Flashing Yellow Arrows (FYA) to enhance safety for turning approaches at intersections (Brehmer et al., 2003; Casola, 2018; Jashami et al., 2019; Tipples, 2002).

Despite the extensive literature on this subject, there remains a dearth of information and studies regarding drivers' behavior, comprehension, understanding, and adherence to NTOR signs. The main objective of this research is to assess the impact of Dynamic and Static NTOR signs on drivers' performance as they approach signalized intersection, comparing them to a baseline scenario in the absence of any signage. The change in speed will serve as an indicator of drivers' compliance when NTOR signs, thereby evaluating their adherence to different types of NTOR signs. Consequently, the influence of various signs on drivers' compliance can be quantified. The findings from this study can be leveraged to enhance intersection safety, improve sign design, and contribute to educational initiatives.

METHODOLOGY

Driving Simulator

A driving simulator at Morgan State University was used in this research to assess the impact of the NTOR signs on drivers' behavior and their compliance at a signalized intersection. The driving simulator consisted of an open cabin containing a steering wheel, gas pedal, brake pedal, and driving engine sound simulation (Figure 1). There were three 40-inch LCD screens, which positioned in front of the cabin, to display the simulated scenarios. The driver front and side windows, plus rear and side mirrors were displayed on the screen (Swake et al., 2013).

Study Scenarios

The virtual environment was designed using UC-WinRoad software packages. Intersection approaches with light traffic volume were characterized by a one-way, three-lane roadway lacking an exclusive right-turn lane and median. This study involved three distinct scenarios

designed for a single intersection. The first scenario, **No Sign Scenario**, served as the base scenario without any signage, allowing right turns on red. The second scenario featured a **Static NTOR Sign** allowing right turns on red only between 7 AM and 7 PM. The third scenario included a **Dynamic NTOR Sign** that prohibited right turns on red when the screen is on and shows the NTOR message.

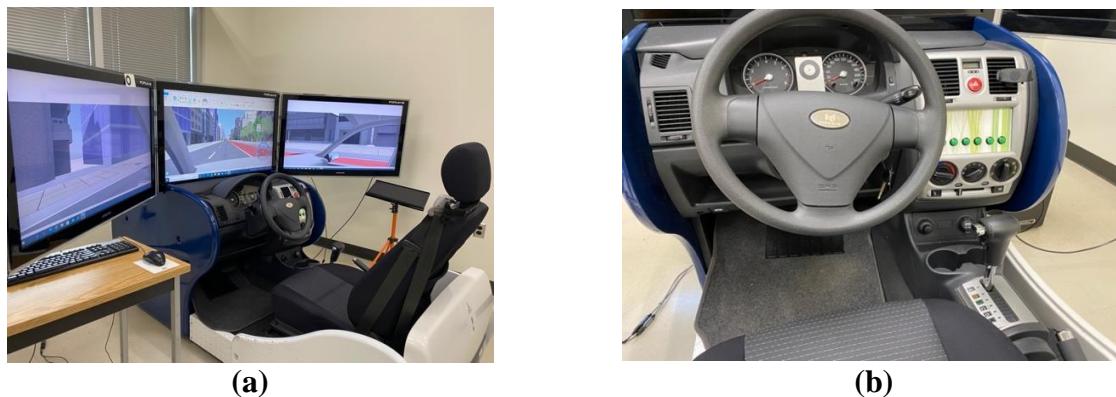


Figure 1. Views from (a) outside and (b) inside the driving simulator

Data Collection

In total, 67 persons (38 Female, 29 Male) of different age groups took part in this experiment. Participants were recruited by distributing flyers with a short, detailed description of the experiment via email in Baltimore, Maryland. All participants included in this experiment had a valid U.S. driver's license. Each scenario started one after the other, with each taking about 8-10 minutes to complete. The socio-demographic characteristics of the participants are presented in Table 1.

TABLE 1. Socio-demographic Characteristics of Participants

Variables		Percentage
Gender	Male	43%
	Female	57%
Age	18 to 25	34%
	26 to 35	27%
	36 to 45	12%
	46 to 55	13%
	56 to 65	13%
Education	High School or Below	19%
	College Student	10%
	Associate Degree	6%
	Bachelor's Degree	24%
	Graduate Degree	36%
	Professional Degree	4%

Due to the purpose of this research, only the data of participants who approached the intersection while the traffic signal was red were considered. The final sample comprised of 53 individual drivers for scenario 1, 60 participants for scenario 2, and 62 drivers for scenario 3.

Two primary dependent variables were extracted from the simulator during the experiment, including the driver's speed and the distance along the road. The average speed of the participants in each scenario was calculated. Drivers' responses to different NTOR signs were analyzed by comparing their speeds in different scenarios along the intersection.

ANALYSIS AND RESULTS

Many studies use statistical analysis to develop policies that improve traffic safety, investigate and forecast travel behavior, and pinpoint deficiencies in transportation policy (Ansariyar & Taherpour, 2023; Mehryaar & Bandelt, 2022; Sadeghvaziri et al., n.d.). A statistical analysis was conducted to evaluate drivers' performance when confronted with different NTOR signs. Average speed and lateral movements were calculated to find out how speed patterns changed to comply with different NTOR sign designs in various scenarios. Moreover, male and female drivers' behavior was analyzed individually to differentiate drivers' performance by gender under different scenarios.

Speed Profile Analysis

All participants

The study examined how different NTOR signs affected average speed changes near a constant 280-meter distance to an intersection. The analysis focused on the 100 to 500-meter range around the intersection. In Figure 2, **without signage**, participants showed a smooth speed drop and acceleration. In the **Static NTOR Sign Scenario**, there was a sharp speed reduction and quick acceleration. The **Dynamic NTOR Sign Scenario** had a sharp speed reduction, similar to the **static** scenario, followed by a gradual increase in speed. Both **static** and **dynamic** scenarios resulted in lower average speeds after turning compared to before approaching the intersection. Participants with **static** signage had higher average speeds than those without, but lower than those with **dynamic** signage. The study also noted that drivers in the **Dynamic NTOR Sign Scenario** complied with the sign earlier, proactively slowing down at the intersection.

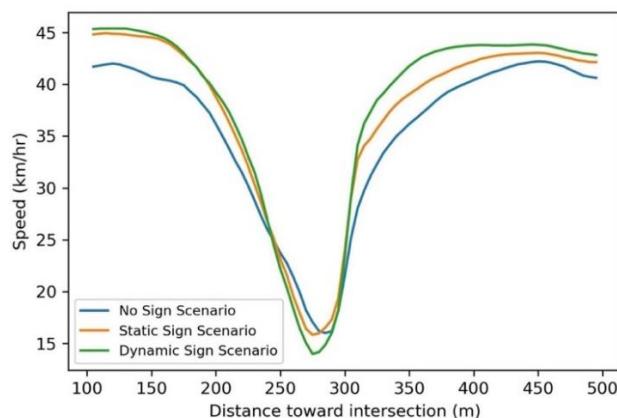


Figure 2. Average Speed Curve Under Different NTOR Signs for All Drivers

Female drivers

Figure 3 illustrates the average speed of female drivers in different scenarios. In the baseline scenario with **no signs**, female drivers had higher speeds before the intersection due to the absence of restrictions. In the **Static NTOR Sign Scenario**, where a time-of-day restriction was indicated, female drivers had lower speeds before the intersection, potentially influenced by their awareness of the restriction. However, at the intersection, their speeds unexpectedly increased, possibly due to visibility challenges and sign comprehension issues. After the intersection, they surpassed the **No Sign Scenario** but remained below the **Dynamic NTOR Sign Scenario**.

In the **Dynamic NTOR Sign Scenario**, female drivers consistently maintained higher speeds before the intersection, attributed to the clear visibility and comprehensibility of the sign. At the intersection, their speed dropped below the **Static NTOR Sign Scenario**, indicating compliance with the "no turn on red" instruction. After the intersection, they sharply increased their speed, remaining lower than the **No Sign Scenario** but higher than the **Static NTOR Sign Scenario**. The **Dynamic NTOR Sign** effectively influenced speed reduction at the intersection but did not result in a prolonged decrease in speed after passing, possibly due to perceived clearance or guidance provided by the sign. The ease of understanding the **Dynamic NTOR Sign** likely contributed to prompt and appropriate speed adjustments.

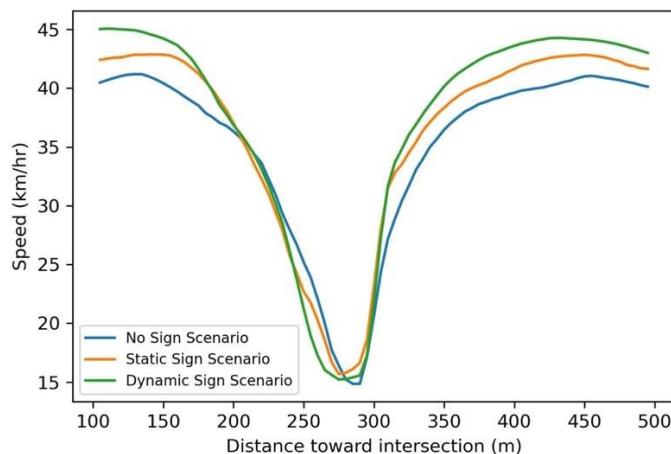


Figure 3. Average Speed Curve Under Different NTOR Signs for Female Drivers

Male drivers

The analysis of average speeds in male drivers (Figure 4) reveals noteworthy trends in different scenarios. In the **No Sign Scenario**, males had lower speeds before and after the intersection, but surprisingly higher speeds at the intersection, indicating a potentially less cautious approach without signage.

In the **Static NTOR Sign Scenario**, males had higher speeds before the intersection than in the **Dynamic NTOR Sign Scenario**. They sharply reduced speed approaching the intersection, likely due to awareness of the right turn restriction. At the intersection, they had higher speeds than the **Dynamic NTOR Sign Scenario** but lower than the **No Sign Scenario**, suggesting effective speed control influenced by the **static sign**. After the intersection, males increased

speed, surpassing the **No Sign** Scenario but remaining slower than the **Dynamic NTOR Sign** Scenario, indicating an impact of the **Static NTOR Sign** on speed behavior.

In the **Dynamic NTOR Sign** Scenario, males had higher speeds before the intersection compared to the **Static NTOR Sign** Scenario. At the intersection, they had lower speeds, possibly influenced by the clear message on the **Dynamic NTOR Sign**. After passing the intersection, males sharply increased speed.

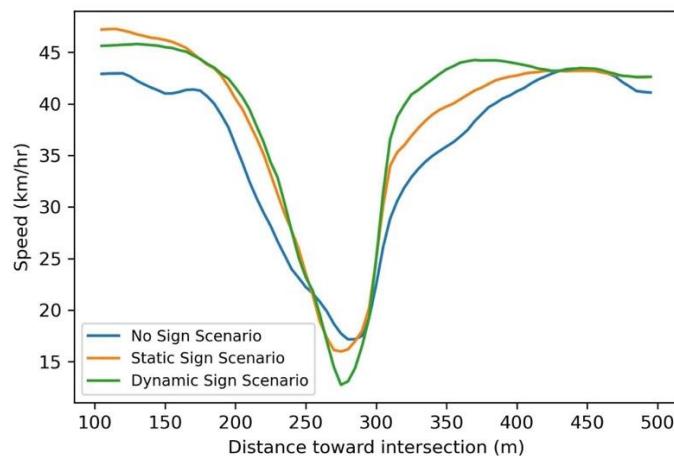


Figure 4. Speed Curve Under Different NTOR Signs for Male Drivers

Compliance Analysis

This study assessed driver compliance with NTOR regulations using **Static** and **Dynamic NTOR Signs**. Figures 2, 3, and 4 show that most drivers did not fully stop at intersections, with speeds exceeding 10 km/h. Compliance was evaluated against an 8 km/h baseline, with speeds below considered compliant. The **Dynamic NTOR Sign** Scenario achieved 55% compliance, indicating its effectiveness in influencing driver behavior. This exceeded the compliance rates of 45% for the **Static NTOR Sign** Scenario and 41% for the **No Sign** Scenario. The study underscores the importance of well-designed signage, such as **Dynamic NTOR Signs**, in conveying regulatory information and promoting compliance among drivers.

Statistical Analysis

t-test-All Participants

A study used a t-test to compare each pair of scenarios at the 280-meter intersection point, assessing differences in minimum speed among all drivers in all three scenarios (Table 2). The analysis aimed to understand how different **NTOR sign** designs influence driving behavior. Results showed a statistically significant difference in minimum speeds between the **No Sign** and **Dynamic NTOR Sign** Scenarios ($p=0.008$), indicating the **Dynamic NTOR Sign** had a noticeable impact with significantly lower minimum speeds compared to the **No Sign** Scenario. However, minimum speed comparisons between the **Static NTOR Sign** and **Dynamic NTOR Sign** Scenarios, as well as between the **No Sign** and **Static NTOR Sign** Scenarios, did not yield statistically significant differences. This suggests that the **Static NTOR Sign** Scenario did not

result in significantly different minimum speeds compared to either the **Dynamic NTOR Sign** or the **No Sign** Scenario.

t-test-Female Drivers

To examine the behavior of female drivers in response to different NTOR signs at 280 meters, t-tests were conducted for each pair of scenarios (Table 2). The analysis revealed that there were no statistically significant differences in the minimum speed between the **No Sign** and **Dynamic NTOR Sign Scenarios**, between **Static NTOR Sign** and **Dynamic NTOR Sign Scenarios**, as well as between **No Sign** and **Static NTOR Sign Scenarios**. These findings indicate that, among female drivers, the presence of different NTOR signs did not lead to statistically significant variations in the minimum speed observed at the 280 meters (Female drivers' result excluded from Table 2). It suggests that female drivers in these scenarios exhibited similar behavior in terms of speed reduction when approaching the intersection, regardless of the type of NTOR signage present.

TABLE 2. T-test Analysis of All Participants, Female Driver's, and Male Driver's Min Speed in Different Scenarios

	No Sign	Static NTOR Sign	No Sign	Dynamic NTOR Sign	Static NTOR Sign	Dynamic NTOR Sign
All Participants						
Mean	16.460	15.821	16.460	13.500	15.821	13.500
Variance	26.544	57.849	26.544	39.176	57.849	39.176
Observations	28	33	28	33	33	33
df	111		112		119	
t Stat	0.512		2.708		1.819	
P-value	0.610		0.008		0.071	
Male Drivers						
Mean	16.986	16.032	16.986	11.533	16.032	11.533
Variance	30.059	68.232	30.059	42.498	68.232	42.498
Observations	28	33	28	33	33	33
df	50		51		53	
t Stat	0.477		3.212		2.206	
P-value	0.635		0.002		0.032	

t-test-Male Drivers

T-tests were conducted to analyze male drivers' responses to different **NTOR signs** at 280 meters. Results in Table 2 show significant differences in minimum speed between **No Sign** and **Dynamic NTOR Sign Scenarios** ($p=0.002$) and **Static NTOR Sign** and **Dynamic NTOR Sign Scenarios** ($p=0.032$). However, there was no significant difference in minimum speed between **No Sign** and **Static NTOR Sign Scenarios**.

These findings highlight that the presence of different **NTOR signs** significantly influenced minimum speeds among male drivers at 280 meters. The **Dynamic NTOR Sign Scenario** led to lower minimum speeds compared to both the **No Sign** and **Static NTOR Sign Scenarios**, suggesting a greater impact on speed reduction. This underscores the effectiveness of the **Dynamic NTOR Sign** in influencing male drivers' behavior compared to **no sign** or a **static sign**.

ANOVA and Post Hoc Tukey

An ANOVA and Post Hoc Tukey tests were conducted to assess minimum speed differences at the intersection among scenarios for all, female, and male drivers. For all participants, a significant difference was found ($p = 0.038$), while for females, no significant difference was observed ($p = 0.865$). However, for males, a significant difference was present ($p = 0.012$).

Post Hoc Tukey analysis revealed no significant difference in minimum speed between **No Sign** and **Static NTOR Sign Scenarios** for all drivers ($p = 0.8616$). The **Dynamic NTOR Sign** led to a significantly lower minimum speed compared to the **No Sign Scenario** ($p = 0.0437$). There was no significant difference between **Static NTOR Sign** and **Dynamic NTOR Sign Scenarios** ($p = 0.1259$).

For male drivers, no significant difference was found between **No Sign** and **Static NTOR Sign Scenarios** ($p = 0.8762$). The **Dynamic NTOR Sign** resulted in a significantly lower minimum speed than the **No Sign Scenario** ($p = 0.0163$). There was no significant difference between **Static NTOR Sign** and **Dynamic NTOR Sign Scenarios** for male drivers ($p = 0.0512$).

These findings highlight the **Dynamic NTOR Sign's** significant impact on reducing minimum speed, especially for male drivers, emphasizing its effectiveness in influencing driver behavior and promoting compliance.

TABLE 3. ANOVA Analysis of Driver's Min Speed for all three Scenarios

Data	F-Statistic	P-Value
All Participants	3.336	0.038
Female Participants	0.145	0.865
Male Participants	4.672	0.012

TABLE 4. Post Hoc Tukey Analysis of All Driver's and Male Driver's Min Speed for Different Scenarios

	No-Sign and Static NTOR Sign	No Sign and Dynamic NTOR Sign	Static NTOR Sign and Dynamic NTOR Sign
All Participants			
Meandiff	-0.6393	-2.9601	-2.3209
P-Value	0.8616	0.0437	0.1259
Male Drivers			
Meandiff	-0.9535	-5.4531	-4.4997
P-Value	0.8762	0.0163	0.0512

Survey Analysis

Post Survey

In the post-survey, the result of the question "Which right turn sign do you prefer better?" provides insight into participants' preference about different NTOR signs. The post survey results indicate that most participants (61%) expressed a preference for the **Dynamic NTOR Sign** over the **Static NTOR Sign** for right turn maneuvers. This suggests that the **Dynamic NTOR Sign** was more visually appealing or perceived as more effective in conveying the message compared to the **Static NTOR Sign**. The higher preference for the **Dynamic NTOR Sign** aligns with its higher compliance rate observed in the study, indicating that participants not only found it preferable but also perceived it as a more suitable and efficient signage option. These findings highlight the importance of considering driver preferences and perceptions when designing and implementing traffic signage systems to promote compliance and enhance road safety.

CONCLUSIONS AND DISSCUSSION

In conclusion, this investigation illuminates' aspects of driver compliance and conduct concerning **NTOR** signs at signalized intersections within the framework of a driving simulator. The results underscore the efficacy of **Dynamic NTOR Signs** in steering drivers toward adherence to the "No Turn On Red" directive. Nevertheless, the study's limitations and avenues for subsequent research warrant consideration.

Primarily, the study concentrated exclusively on right turns at a signalized intersection, constraining the applicability of the findings to other driving scenarios. Additionally, gender-based disparities in compliance, as observed, necessitate further exploration to comprehend broader implications and ensure an all-encompassing approach to traffic signage design.

A salient implication for regulatory enforcement and engineering interventions lies in the conspicuous influence of **Dynamic NTOR Signs** on compliance. Authorities in transportation stand to gain by prioritizing the integration of such signs to amplify the transmission of regulatory information and foster adherence. Furthermore, the recognition of driver preferences, as indicated by post-survey outcomes, underscores the significance of incorporating user perceptions into the design of traffic signage.

Nevertheless, it is imperative to acknowledge the study's constraints. The analysis centered on a specific distance range and intersection context, potentially neglecting variations in driver behavior across diverse conditions. Subsequent research endeavors should scrutinize the enduring nature of the identified disparities at various distances and within different driving scenarios, facilitating a more holistic comprehension of driver responses to **NTOR** signs.

To summarize, although this study furnishes valuable insights into **NTOR** sign compliance, acknowledging its limitations and addressing areas for further exploration will refine our comprehension of driver behavior. This refinement, in turn, will inform the development of more efficacious enforcement strategies and engineering interventions within the domain of road safety.

AUTHOR CONTRIBUTIONS

Study conception and design: P. Masoumi, A. Taherpour, M. Jeihani, A. Ardeshtiri, S. Ahangari; data processing: P. Masoumi, A. Taherpour, M. Jeihani, A. Ardeshtiri, S. Ahangari;

analysis and interpretation of results: P. Masoumi, A. Taherpour, M, Jeihani, A. Ardeshtiri, S. Ahangari; draft manuscript preparation: P. Masoumi, A. Taherpour, M, Jeihani, A. Ardeshtiri, S. Ahangari. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Abu-Lebdeh, G., Benekohal, R., and Al-Omari, B. (1997). Models for Right-Turn-on-Red and Their Effects on Intersection Delay. *Transportation Research Record Journal of the Transportation Research Board*, 1572, 131–139.
- Ansariyar, A., and Taherpour, A. (2023). Statistical analysis of vehicle-vehicle conflicts with a LIDAR sensor in a signalized intersection. *Advances in Transportation Studies*, 60, 87–106.
- Brehmer, C. L., and American Association of State Highway and Transportation Officials, & United States (Eds.). (2003). Evaluation of traffic signal displays for protected/permissive left-turn control: NCHRP report 493. Transportation Research Board of the National Academies.
- Casola, E. (2018). Driver Understanding of the Flashing Yellow Arrow and Dynamic No Turn on Red Sign for Right Turn Applications. <https://www.semanticscholar.org/paper/Driver-Understanding-of-the-Flashing-Yellow-Arrow-Casola/9483211582c21f77f92dc48c3d09ab8897cb3b56>.
- Compton, R. P., Milton, E. V., and United States National Highway Traffic Safety Administration, Office of Program Development and Evaluation, Evaluation Division. (1994). Safety Impact of Permitting Right-Turn-on-Red: A Report to Congress (DOT-HS-808).
- Jashami, H., Hurwitz, D. S., Monsere, C., and Kothuri, S. (2019). Evaluation of driver comprehension and visual attention of the flashing yellow arrow display for permissive right turns. *Transportation Research Record*, 2673(8), 397–407.
- Mehryaar, E., and Bandelt, M. J. (2022). Geological Modeling Along Tunnel Projects Using Machine Learning Techniques. *ARMA/DGS/SEG International Geomechanics Symposium*.
- Parker, M. R., and Virginia Highway Safety Division, Virginia Highway & Transportation Research Council, & Virginia Dept. of Highways and Transportation. (1975). Right turn on red: A report to the Governor and General Assembly of Virginia. (VHTRC 76-R9). <https://rosap.ntl.bts.gov/view/dot/18802>.
- Podany, K. I., Wogalter, M. S., and Mayhorn, C. B. (2004). Perceived Effectiveness of ‘No Turn on Red’ Traffic Signs. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(16), 2004–2007.
- Preusser, D. F., Leaf, W. A., DeBartolo, K. B., Blomberg, R. D., and Dunlap and Associates, Inc. (1981). The Effect of Right-Turn-on-Red on Pedestrian and Bicyclist Accidents (DOT-HS-806-182).
- Pulugurtha, S. S., Nambisan, S. S., Dangeti, M. R., and Vasudevan, V. (2010). Evaluation of Effectiveness of Traffic Signs to Enhance Pedestrian Safety (10–1016). Article 10–1016. *Transportation Research Board 89th Annual Meeting Transportation Research Board*. <https://trid.trb.org/view/909745>.
- Sadeghvaziri, E., Javid, R., and Jeihani, M. (n.d.). *Investigating Walking and Biking Activities Among Low-Income African Americans*.
- Swake, J., Jannat, M., Islam, M., and Hurwitz, D. (2013, January 1). *Driver Response to Phase Termination at Signalized Intersections: Are Driving Simulator Results Valid?*

- Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, 9(2), 314–318.
- Wagoner, W. (1992). Driver behavior at right-turn-on-red locations. Ite Journal-Institute of Transportation Engineers. <https://www.semanticscholar.org/paper/Driver-behavior-at-right-turn-on-red-locations-Wagoner/023650a178debba4a9dfaead7698803101770adc>.
- Zador, P. L. (1984). Right-turn-on-red laws and motor vehicle crashes: A review of the literature. *Accident Analysis & Prevention*, 16(4), 241–245.
- Zador, P., Moshman, J., and Marcus, L. (1982). Adoption of right turn on red: Effects on crashes at signalized intersections. *Accident Analysis & Prevention*, 14(3), 219–234.
- Zegeer, C. V., and Cynecki, M. J. (n.d.). Evaluation of Countermeasures Related to RTOR Accidents That Involve Pedestrians. *Transportation Research Record*.

Safety Benefits of Parcel Delivery Modes Using Geographically Weighted Negative Binomial Regression

Zhenyu Wang, Ph.D.¹; Pei-Sung Lin, Ph.D.²; and Yaye Mallon Keita, Ph.D.³

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: zwang9@usf.edu

²Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: lin@usf.edu

³Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: ykeita@usf.edu

ABSTRACT

Emerging urban parcel delivery (UPD) modes are anticipated to decrease surface UPD truck trips and stops, thus leading to less exposure of UPD trucks on surface roads and reduced UPD crashes. This paper evaluated the safety impacts of innovative last-mile delivery strategies in urban areas. The geographically weighted negative binomial regression (GWNBR) model was developed at zone levels based on the roadway, traffic, and demographic data collected in Hillsborough County, Florida. Future UPD scenarios were projected for coming years (2030, 2040, and 2050) with different replacement rates (10%, 30%, and 50%) of UPD truck stops by emerging UPD modes. The developed GWNBR model was used to predict UPD crashes for future scenarios. The results indicate that emerging UPD technologies cause a decrease in delivery truck stops and reduce UPD crashes by 3%, 11%, and 20% for 2030, 2040, and 2050, respectively.

INTRODUCTION

With the rapid evolution of e-commerce, urban parcel delivery (UPD) has emerged as a high-growth market undergoing rapid technological change, particularly in the business-to-consumer segment. Based on the National Household Travel Survey (NHTS) data, the frequency of household shopping in stores has decreased in the past 10-20 years. At the same time, deliveries by parcel delivery trucks and vans have escalated around cities every day (Stinson, Enam et al. 2019). The increased UPD activities imply boosted delivery truck travel miles and frequent on-street stops, which result in congestion, pollution, and crashes. Emerging technologies will change the UPD patterns on surface roads, not only traffic operations but also the safety performance of roadway networks. Automated Vehicles (AV), delivery robots, and Unmanned Aircraft Vehicles (UAVs) decrease the traffic exposure of surface UPD truck trips and stops; meanwhile, driverless operations eliminate human errors in the journeys of parcel delivery. When deployed in a large-scale urban area, the emerging UPD modes, especially for UAVs, help reduce road crash risks and mitigate traffic congestion.

With the potentially significant demand for emerging UPD delivery modes in the foreseeable future, the need for and importance of evaluating the impacts of the migration of UPD traffic caused by the emerging delivery technologies on roadway safety is arising. However, limited previous studies , and Jeong 2020) (Ibrahim, Ab Rashid, et al. 2018) (Chung, Song, and Yoon 2014) (Xie, Ozbay, investigated delivery crashes and the contributing factors. Due to the absence

of parcel delivery data, implementable Safety Performance Functions (SPFs) for predicting the delivery-related crash patterns for different UPD scenarios were not developed in previous studies.

To fill the research gaps, this study aims to develop an analytical framework to assess the safety impacts of future UPD technologies. Specifically, the research objectives include: (1) developing data collection methods and collecting needed data, (2) developing zone-based SPFs for UPD crashes, (3) designing future UPD scenarios with new delivery technologies, and (4) estimating the safety and mobility benefits of the proposed UPD scenarios.

DATA PREPARATION

This study adopted Hillsborough County, which is the fourth most populous county (1,459,762) in Florida, as the study site. The research team collected UPD crash data, roadway network data, and demographic and socioeconomic data from various data sources and matched the data at traffic analysis zones (TAZs).

UPD Crash Identification

A UPD Crash is a traffic collision event involving at least one UPD vehicle. Existing crash databases in Florida do not provide a data field to indicate if a crash is related to a UPD truck. This study used the Fuzzy algorithm (*Fuzzywuzzy*) to match pre-defined keywords of major delivery vendors (e.g., UPS, USPS, FedEx, DHL, Amazon, etc.) and their variants to match vehicle information fields and crash narratives in crash reports. If a crash was matched to the keyword, the crash was identified as a UPD crash. The research team applied the matching method to four-year (2016-2019) crash reports in Hillsborough County, Florida. A total of 214 UPD crashes were identified and matched to TAZs.

Geographic Matching

Traffic features, such as Annual Average Daily Traffic (AADT), truck percentage (TFCTR), and speed limit, are downloaded from the Florida Department of Transportation (FDOT) Roadway Characteristics Inventory (RCI) database. These data were spatially matched to each TAZ. Since one TAZ may contain one or more traffic features, the weighted average traffic features were calculated for each TAZ sing the following formulas:

$$F_j = \frac{\sum_i F_{ij} \cdot L_{ij}}{\sum_i L_{ij}} \quad (1)$$

where F_j is the weighted average of a feature for TAZ j ; L_{ij} is the length of segment i in TAZ j ; and F_{ij} is the feature value of segment i in TAZ j .

Roadway network data were retrieved from the Florida All Roadways, Intersections, and Streets (FLARIS) database and were spatially matched to TAZs. The total mileage of roadway routes and the number of intersections within each TAZ were calculated. Demographic data were retrieved from the Tampa Bay Regional Planning Model Networks (TBRPM) for each TAZ. The TBRPM has been developed for the planning activities under the Regional Transportation Analysis (RTA) for the Tampa Bay Area. The dwelling units and employment information have been extracted and matched to each TAZ. The collected data are summarized in Table 1.

Table 1. Descriptive Statistics of Collected Data (sample size: 884)

Feature	Description	Mean	SD	Source
CrashCount	The number of UPD crashes (4 years) in a TAZ	0.24	0.67	S4 ¹
WA_SpeedLi	Weighted average speed limit in a TAZ	42.21	8.48	RCI ²
WA_AADT	Weighted average AADT in a TAZ	25226.70	17103.17	RCI ²
TOTAL_AADT	Sum of AADT in a TAZ	179036.20	175198.7	RCI ²
WT_AADT	Weighted total AADT in a TAZ	75453.18	70548.85	RCI ²
WA_TFCTR	Weighted average TFCTR in a TAZ	7.07	1.83	RCI ²
TOTAL_TFCTR	The sum of TFCTR in a TAZ	49.36	32.85	RCI ²
WT_TFCTR	Weighted total TFCTR in a TAZ	24.63	28.28	RCI ²
TOTAL_Miles	The sum of the length of all routes in a TAZ	8.51	6.37	FLARIS ³
TRAFFIC_SIG	Total number of traffic signals in a TAZ	0.87	1.37	RCI ²
FLARIS_INT	Total number of FLARIS intersections in a TAZ	41.63	28.84	TBRPM ⁴
DU	Total dwelling units in a TAZ	635.76	597.93	TBRPM ⁴
SERV_REMP	Total number of service employees in a TAZ	620.26	925.62	TBRPM ⁴
COMM_REMP	Total number of commercial local employment in TAZ	158.48	262.44	TBRPM ⁴
TOTAL_EMP	Total number of employments in a TAZ	941.52	1062.48	TBRPM ⁴

¹Florida SignalFour Analytics, ²FDOT Roadway Characteristics Inventory, ³Florida All Roadways, Intersections and Streets, ⁴Tampa Bay Regional Planning Model Networks

METHODOLOGY

This study proposed an analytical framework, as shown in Figure 1, to predict the impacts of emerging UPD technologies on UPD crash patterns. A safety performance function (SPF) was developed to estimate the UPD crash frequency at TAZ levels based on TAZ's roadway and traffic characteristics, demographic features, and UPD demand/exposures on surface roads. Implementing emerging UPD technologies substitutes the UDP demand/exposures on surface roads. Future scenarios assume the migration of UPD patterns and the reduction of UPD demand/exposures on surface roads. The developed SPF was used to predict UPD crashes in future scenarios and estimate the safety benefits of emerging UPD technologies by comparing UPD crashes between the base scenario (all surface UPD delivery) and future scenarios.

Estimation of UPD Travel Demand/Exposures

UPD demand data and models are unavailable on the study site. This study adopted an existing travel demand model (Moore 2019) to estimate UPD demand/exposure on surface roads. The model was previously developed for the Columbus, Ohio, area. Using similar variables, the

research team estimated monthly truck stops per TAZ in Hillsborough County, Florida, to indicate UPD demand/exposure on surface roads. The model formula is given as follows:

$$\hat{Y} = 16.325 + 0.039x_1 + 0.041x_2 + 0.179x_3 + 0.033x_4 \quad (2)$$

where \hat{Y} is the monthly delivery truck stops per TAZ; x_1 is the total number of dwelling units per TAZ; x_2 is the total number of service employments per TAZ; x_3 is the total number of commercial employments per TAZ; and x_4 is the total number of other employments per TAZ. Notably, the observed truck stops for modeling were collected from a truck fleet (including 19 trucks) at a major parcel delivery company. The predicted truck stops (\hat{Y}) are not the total delivery truck stops in a TAZ but can be a scaled-down indicator of the total stops.

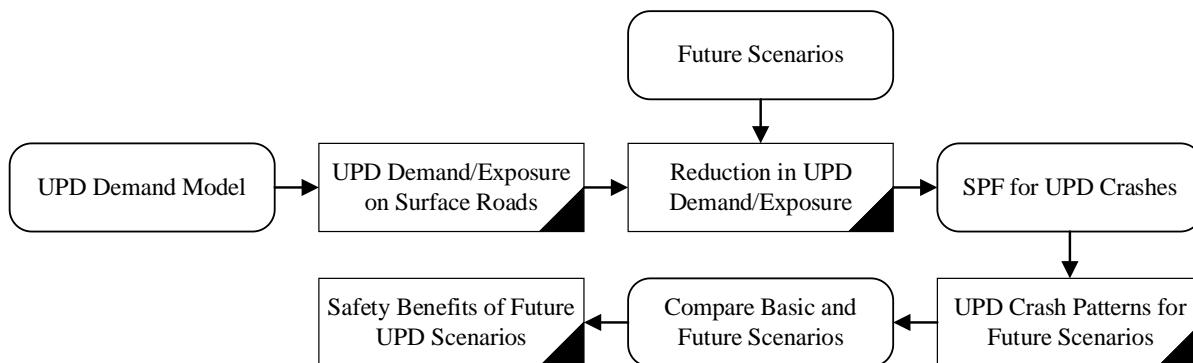


Figure 1. Analytical Framework to Predict the Safety Impacts of UPD Technologies

UPD Safety Performance Function

Researchers have developed various models to fit spatially aggregated crash data, such as geographically weighted regression (GWR) (Silva and Rodrigues 2016), macro-level CAR model, micro-level spatial joint model, random effects negative binomial model, and multivariate Poisson lognormal model with multivariate conditional auto-regressive prior (MVPLN-MCAR model) (Zhai, Huang et al. 2019, Obelheiro, da Silva et al. 2020, Yocum and Gayah 2022). Among these models, Geographically Weighted Poisson Regression (GWPR) and Geographically Weighted Negative Binomial Regression (GWNBR) are the more complex but effective in dealing with spatial autocorrelation and heterogeneity (Gomes, Cunto and da Silva 2017, Hezaveh, Arvin and Cherry 2019, Oluwajana, Park and Cavalho 2022). The concept of GWR is that the observed data next to zone i has more influence on the estimation of zone i 's coefficients than the data located further away from i , whose relationship is determined with a weighting function. GWR attempts to address spatial variation by adjusting a regression model to each point individually and using data subsamples while weighting with a distance function denominated by a kernel spatial function (Gomes, Cunto and da Silva 2017).

GWPR enables the modeling of discrete data at the zonal level using Poisson distribution. The limitation of GWPR is that it fails to consider overdispersion in crash data. The general form of the GWPR model is as follows (Gomes, Cunto and da Silva 2017):

$$y_j \sim \text{Poisson}[t_j \exp(\sum_k \beta_k(u_j, v_j)x_{jk})] \quad (3)$$

where y_j is the UPD crash frequency of Zone j , for $j = 1, \dots, J$; t_j is an offset variable for Zone j ; $\beta_k(u_j, v_j)$ is the coefficient related to the explanatory variable x_{jk} , for $k = 1, \dots, K$, with a spatial variation; (u_j, v_j) is location coordinates of zone centroids.

Similarly, GWNBR allows the modeling of counting data in a non-stationary way by including overdispersion of the data. This equation below shows the general form of the GWNBR model (Gomes, Cunto and da Silva 2017).

$$y_j \sim NB[t_j \exp(\sum_k \beta_k(u_j, v_j)x_{jk}), \alpha(u_j, v_j)] \quad (4)$$

where α is the parameter of overdispersion which is zone-specific.

To avoid difficulty in the estimation, Silva & Rodrigues (2014) also proposed the GWNBR with the global α (overdispersion parameter), called GWNBRg. In this model, the spatial variation is only allowed to $\beta_k(u_j, v_j)$. As there is no spatial variation for α , its contribution to the effective number of parameters of the model is unitary. The general form of the GWNBRg model is given in the following equation.

$$y_j \sim NB[t_j \exp(\sum_k \beta_k(u_j, v_j)x_{jk}), \alpha] \quad (5)$$

where α is the parameter of overdispersion keep constant over zones.

For GWR, it is necessary to determine the optimum bandwidth (smoothing parameter) that regulates the kernel size and controls the rate at which the weight of a given zone i decreases as it drifts away from the place of the regression point being analyzed (j). A way to estimate the optimum bandwidth is through the minimization of the AICc, which can be estimated using the following equation:

$$AIC_c = -2L(\beta, \alpha) + 2k + \frac{2k(k+1)}{n-k-1} \quad (6)$$

where, AICc is corrected Akaike Information Criterion; $L(\beta, \alpha)$ is log of maximum likelihood of the GWNBR; k is effective number of parameters – effective number of parameters of GWNBR can be written as $k = k_1 + k_2$, where k_1 and k_2 are the effective numbers of parameters due to β and α , respectively.

To date, it has been impossible to estimate k_2 —that is, the surface contribution of α in the effective number of parameters of the model. Thus, one way to overcome that challenge is to estimate the bandwidth by using the criterion of cross-validation (CV) given in Eq. 7.

$$CV = \sum_j [y_j - \widehat{y}_{\neq j}(b)]^2 \quad (7)$$

where CV is the cross-validation; y_j is the UPD crash frequency of Zone j ; $\widehat{y}_{\neq j}(b)$ is the estimated UPD crash frequency for zones other than j ; and b is the bandwidth.

To estimate the bandwidth, it is also important to select a kernel function. Two types of kernel functions can be considered as: Gaussian fixed kernel function and Biquadratic adaptive kernel function. The equations are as follows:

$$\text{Gaussian: } w_{ij} = \left\{ -0.5 \left(\frac{d_{ij}}{b} \right)^2 \right\} \quad (8)$$

$$\text{Biquadratic: } w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{G_i}\right)^2\right]^2 & \text{if } d_{ij} < G_i \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where w_{ij} is the weight value of an observation in j for the coefficient estimation in i; d_{ij} is the Euclidian distance between the points i and j, parameter b (bandwidth) regulates the kernel size and controls the rate at which the weight of a given point i decreases as it drifts away from the place of the regression point being analyzed (j). For the biquadratic function, G_i denotes the distance to the N_{th} nearest observation from regression point i obtained from the optimized bandwidth size (measured in neighbors).

The authors used all the possible methods to estimate the bandwidth. After the optimum bandwidth was selected for each method, a model was estimated using the optimum bandwidth and the model with the smallest AIC value was selected to be the best model. All the estimations were done using the SAS/IML macros by (Silva and Rodrigues 2016). The golden search was done to yield the optimum bandwidth (X_{min}) and corresponding golden score (AIC or CV). It is important to note that the AIC obtained during the golden search is different from the AIC obtained from the model.

Future Scenario Projection

Future UPD scenarios were projected for 2030, 2040, and 2050 with the replacement rate of 0%, 10%, 30%, and 50% for UPD truck stops by emerging UPD modes. In the scenario projection, the total miles of roadway, indicator of urban area, and percent dwelling units classified as working households with children and two cars, were kept constant while AADT and the variables for UPD truck stop estimation (dwelling units, service employments, commercial employments, and other employments) are projected based the data of the based year (2017) using the following formula:

$$\begin{cases} \text{Projected Data}_{\text{target year}} = \text{Data}_{2017} * [1 + GR] \\ GR = \frac{(\text{Population}_{\text{target year}} - \text{Population}_{2017})}{\text{Population}_{2017}} \end{cases} \quad (10)$$

The population data from 2020 to 2050 were retrieved from the TBRPM.

RESULTS

Model Estimation

The delivery truck stops were estimated from Eq.1 using the collected data (Table 1) for each TAZ in Hillsborough County, Florida. The estimated UPD stops are an independent variable of the GWR-based SPF. The estimation of GWRs requires searching for optimum bandwidth. The search results are illustrated in TABLE 2. The local GWNBR using the adaptive searching method and AIC type is the best model because the model has the lowest AICc among all estimated models. The X_{min} value, which represents the bandwidth, indicates that the best model (adaptive model) used 801 out of 875 TAZs to estimate values for TAZs, which means that it used 92 percent of all TAZs in Hillsborough County to estimate the local values. Although the

data is not completely global, it can be said that it is 92 percent global and 8 percent local. Nevertheless, the local model seems to be a slightly better model than the global model when considering the slight differences in AICc. The golden score or AIC score that yielded the optimum bandwidth was 1028.716. The AICc for that model (the best model) was 1029.014.

The 25th, 50th, and 75th percentiles of coefficient estimates (P5, P50, P75) and other statistics (mean, maximum, and minimum) for the best model (GWNBR) and the traditional NB model are provided in Table 3. The modeling efforts reveal that an increase in natural logarithm of number of delivery truck stops increased the frequency of UPD crashes. As the natural logarithm of AADT increases the frequency of UPD crashes also escalated. The results also convey that the total miles of roadway are positively related to UPD crashes. Urban areas are more likely to have UPD crashes than other area types. More parcel deliveries are likely in those areas than in other places. The model also suggests that as the percentage of dwelling units classified as working households with children and two cars increases, the frequency of UPD crashes decreases. This can be explained by the fact that households with two cars can use those vehicles to go to the stores and shopping places. Thus, they are more likely, for example, to drive to shopping places than households with no car who will probably do more online shopping.

Table 2. Golden Section Search Results

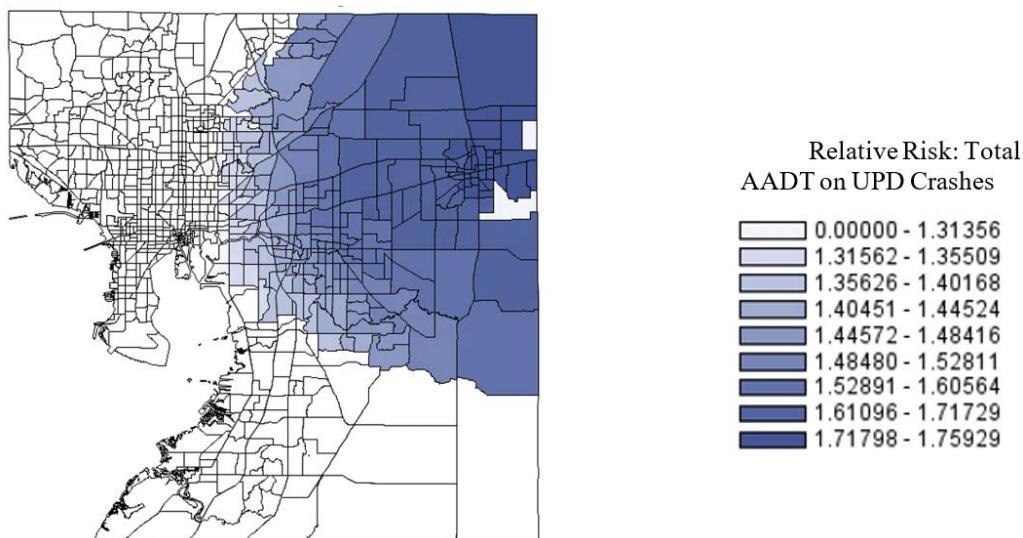
Model	Kennel Function	Searching Statistics	AIC or CV Score	Xmin (Bandwidth)	Model AICc
GWNBR	Fixed	AIC	1030.237	227741.83	1030.237
	Fixed	CV	379.2983	227741.83	1030.237
	Adaptive	AIC	1028.716	801	1029.014
	Adaptive	CV	7.23E+86	803	1716.707
GWNBRg	Fixed	AIC	1030.337	227741.83	1030.337
	Fixed	CV	379.2561	227741.83	1030.337
	Adaptive	AIC	1031.697	699	1031.697
	Adaptive	CV	378.4343	706	1031.722
GWPR	Fixed	AIC	1070.745	23989.587	1070.745
	Fixed	CV	380.993	227741.83	1078.682
	Adaptive	AIC	1067.721	689	1067.721
	Adaptive	CV	381.8985	875	1076.953
Negative Binomial	Traditional, Fixed	N/A	N/A	N/A	1030.155
Poisson	Traditional, Fixed	N/A	N/A	N/A	1078.849

Using the best model, maps of significant values and areas are produced. For AADT (FIGURE 2), the GWNBR model suggests that for the TAZs on the north-east of Hillsborough County (around Plant City), that an increase in AADT is (1.76 to 1.31) times likely associated with an increase in UPD crashes compared to other areas (see Figure 2). The areas in white are not significant; thus, for those areas AADT cannot be used to explain the increase in UPD Crashes. The reason for this finding may be that around the plant city area in Florida there are several freight hot spots. Also, since ADDT values are only available for major roadways, it may not be able to capture the traffic on Tampa local roadways.

Table 3. Estimated UPD Crash Models

	Intercept	Natural logarithm of delivery truck stops	Natural logarithm of AADT	Total miles of road routes	Indicator of urban area	Percent of dwelling units classified as working households with children and two cars
GWNBR (Eq. 4) (AICc = 1029.014)						
P25	-7.659	0.250	0.202	0.050	0.742	-0.052
P50	-6.765	0.322	0.238	0.060	0.879	-0.048
P75	-6.216	0.362	0.335	0.064	1.055	-0.046
Mean	-42.7539	11.576	-5.560	6.794	44.670	-1.466
Min	-86651.5	-4322	-3671	-1249	-45747	-780
Max	55414	14188	0.565	7144	84049	-0.039
Traditional Negative Binomial Model (AICc = 1030.155)						
Coef.	-7.111	0.368	0.267	0.046	0.860	-0.051
p-Value	<.0001	0.016	0.007	0.003	0.145	0.000

Figure 3 shows that TAZs in Tampa with higher total roadway miles are 1.07 times more likely to have UPD crashes compared to other areas. Total roadway miles cannot be used to explain an increase in UPD crashes in TAZs that are colored in white in Figure 3 because the values for those areas are insignificant. The results seem reasonable since Tampa has several freight hot spots and it has several miles of roadway, especially when considering all the local, regional, state, and interstate roadways that are available or pass through the city.

**Figure 2. Maps of relative UPD crash risk associated with AADT**

As the percentage of dwelling units classified as working households with children and two cars increases in the darkest blue TAZs, the risk of UPD crashes decreases by 0.96 compared to

0.94 in the lightest blue areas (see Figure 4). The values for TAZs in white are insignificant; thus, the percentage of dwelling units classified as working households with children and two cars cannot be used as an indicator of UPD crashes in those areas.

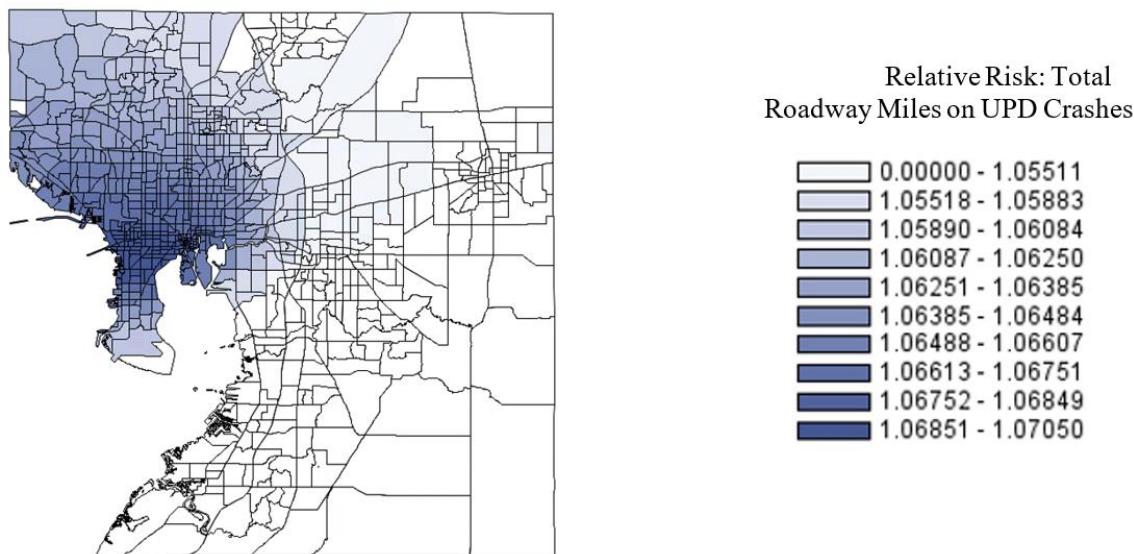


Figure 3. Maps of relative UPD crash risk associated with roadway miles

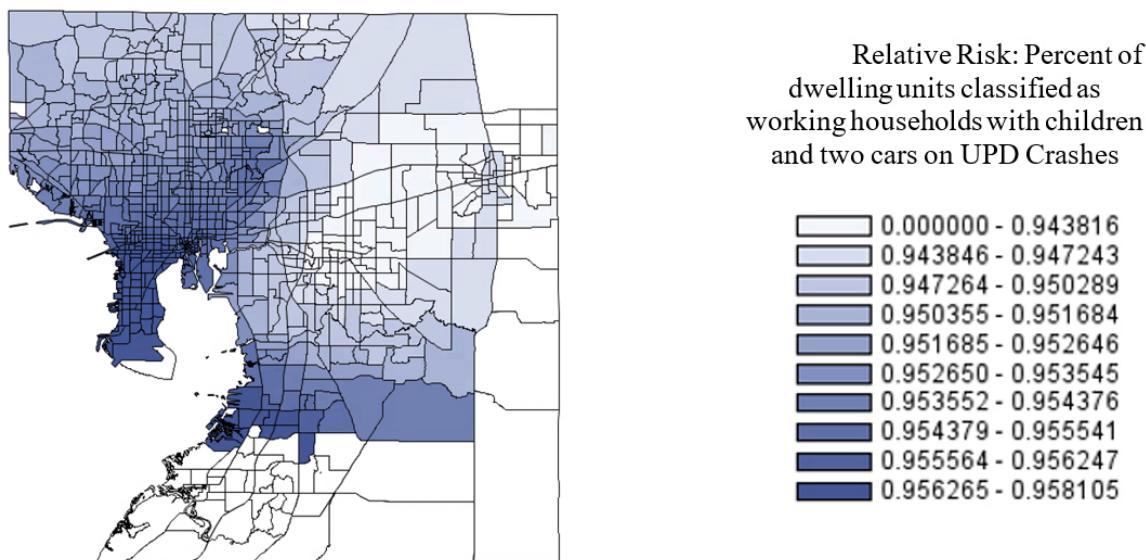


Figure 4. Maps of relative UPD crash risk associated with dwelling units

Future Scenarios

The results of the projected UPD crashes for future scenarios (2030, 2040, and 2050) with different UPD truck stop replacement rates (0%, 10%, 30%, and 50%) are presented in Figure 5. The results demonstrate that without the introduction of emerging UPD modes, UPD crashes will increase in the future. Replacing traditional UPD truck stops with emerging UPD mode stops

will decrease the frequency of UPD crashes in future scenarios. The crash reduction is even higher in the long term (e.g., 2050) than in the short term (e.g., 2030). Figure 5 shows that the reductions from base values are 3%, 11%, and 20% when 10%, 30%, and 50 % of UPD truck stops are replaced by emerging UPD modes. The results make sense because more UPD stops are expected in the future due to the growing number of online shopping; thus, replacing some with non-roadway modes (e.g., drones) could improve the safety of UPD in the future.

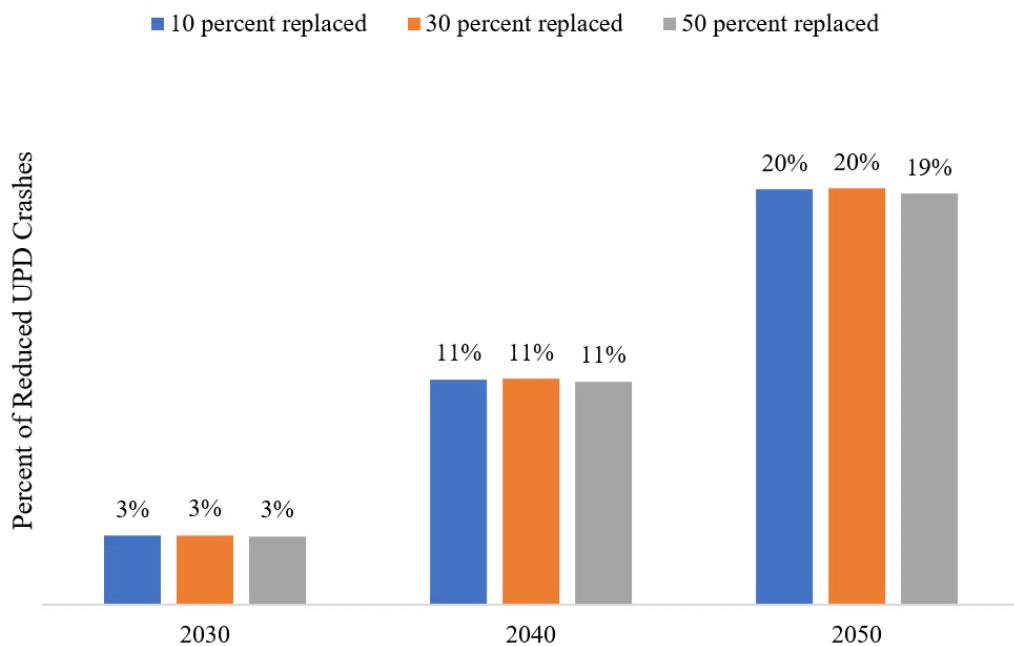


Figure 5. Relative Reduction of UPD crashes by Future Scenarios

SUMMARY AND CONCLUSIONS

This study developed a zone-based UPD crash model using the Geographically Weighted Negative Binomial regression model. The model was used to predict UPD crash frequencies for future UPD scenarios (substitution of surface delivery truck stops). By comparing the UPD crashes by different UPD scenarios for feature years, the safety benefits of new UPD modes were also estimated for future scenarios. Through the study, the following findings and conclusions are obtained:

- UPD crash frequencies at a TAZ level are associated with several factors, including delivery truck stops, total AADT, roadway mileage, area type, and household characteristics. The developed model implies that increases in delivery truck stops, total AADT, and roadway mileages are more likely to increase the UPD crashes. Urban areas also tend to experience a high UPD crash frequency. With an increase in the percentage of dwelling units classified as working households (with children and two cars), the UPD crashes are more likely to decrease.
- The comparisons between UPD scenarios (10%, 30%, and 50% substitution of delivery truck stops with new UPD modes) and baseline for future years show that the UPD crashes are more likely to be reduced by 3%, 11%, and 20% for the three scenarios, respectively.

Some limitations to be addressed in future studies include:

- Because parcel delivery demand data was not available in Florida, this study adopted a model developed for the Columbus area in Ohio (Moore 2019) to estimate the parcel delivery demand. Thus, the estimated delivery truck stops cannot directly measure parcel delivery demand in the study site. Although it is reasonable to assume the comparison of UPD scenarios and the baseline are valid based on the estimated parcel delivery stops, in the future it is important to develop local parcel delivery demand model to estimate the parcel delivery demand more accurately.
- The UPD scenarios were created by assumptions that 10%, 30%, and 50% substitution of delivery truck stops. The assumptions are simple and intuitive but may not capture the "true" parcel delivery demands in future UPD environments. A better estimation of the future parcel delivery patterns is needed to improve the estimation accuracy.
- This study used new data sources, such as roadway networks. However, some data are still incomplete. For example, AADT data were retrieved from the FDOT traffic database that just covered interstates and major roads. New data sources are needed to collect more detailed information, especially for minor roadway facilities.

ACKNOWLEDGMENTS

This work was funded by the National Institute of Congestion Reduction (NICR). The authors would also like to thank Professor Alan Ricardo da Silva for his support during the modeling process.

REFERENCES

- Gomes, M. J. T. L., F. Cunto, and A. R. da Silva. (2017). "Geographically weighted negative binomial regression applied to zonal level safety performance models." *Accident Analysis & Prevention* 106: 254-261.
- Hezaveh, A. M., R. Arvin, and C. R. Cherry. (2019). "A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level." *Accident Analysis & Prevention* 131: 15-24.
- Moore, A. M. (2019). "Innovative scenarios for modeling intra-city freight delivery." *Transportation Research Interdisciplinary Perspectives* 3: 100024.
- Obelheiro, M. R., A. R. da Silva, C. T. Nodari, H. B. B. Cybis, and L. A. Lindau. (2020). "A new zone system to analyze the spatial relationships between the built environment and traffic safety." *Journal of transport geography* 84: 102699.
- Oluwajana, S. D., P. Y. Park, and T. Cavalho. (2022). "Macro-level collision prediction using geographically weighted negative binomial regression." *Journal of Transportation Safety & Security* 14(7): 1085-1120.
- Silva, A. D., and T. Rodrigues. (2016). "A SAS® macro for geographically weighted negative binomial regression." <http://support.sas.com/resources/papers/proceedings16/8000-2016.pdf>. Data de acesso 1(06): 2016.
- Stinson, M., A. Enam, A. Moore, and J. Auld. (2019). Citywide impacts of E-commerce: Does parcel delivery travel outweigh household shopping travel reductions? *Proceedings of the 2nd ACM/EIGSCC Symposium on Smart Cities and Communities*.

- Yocum, R. L., and V. V. Gayah. (2022). "County-level crash prediction models for Pennsylvania accounting for income characteristics." *Transportation research interdisciplinary perspectives* 13: 100562.
- Zhai, X., H. Huang, P. Xu, and N. Sze. (2019). "The influence of zonal configurations on macro-level crash modeling." *Transportmetrica A: transport science* 15(2): 417-434.

Optimizing Speed Control Guidance at Urban Signalized Intersections: A Driving Simulator Study on Driver Behavior and Sociodemographic Factors

Parisa Masoumi¹; Eazaz Sadeghvaziri, Ph.D.²; and Mansoureh Jeihani, Ph.D.³

¹Ph.D. Student, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: pamas2@morgan.edu

²Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Macon, GA. Email: sadeghvaziri_e@mercer.edu

³Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: mansoureh.jeihani@morgan.edu

ABSTRACT

Urban intersections often suffer from high traffic congestion and emissions due to frequent braking and idling. This study proposes a Speed Control Guidance (SCG) system to optimize vehicle trajectories in mixed traffic, including both internal combustion engine vehicles (ICEVs) and battery electric vehicles (BEVs). A 3D driving simulator assesses driver responses to real-time SCG guidance at signalized intersections. Participants receive color-coded speed recommendations along their route in various scenarios, and their driving behavior is compared to scenarios without SCG. Sociodemographic factors like gender and age influence SCG effectiveness. Female drivers show lower compliance with speed guidance, and older drivers struggle to follow recommendations. To enhance SCG's effectiveness, it is crucial for researchers and vehicle manufacturers to develop strategies that address the specific needs and preferences of diverse driver groups.

Keywords: Eco-driving, Speed Control, Speed Optimization, Mixed Traffic, Driving Simulator.

INTRODUCTION

Urban areas are known for having high traffic volume and emissions at signalized intersections. These areas experience higher levels of braking and accelerating, as well as the longest periods of vehicle idling. Drivers frequently approach a green light at maximum speed and are forced to stop instantly when the light turns red. This lack of knowledge about the traffic signal's potential future condition leads to increased fuel consumption and longer travel times. Researchers have attempted to use connected vehicle and infrastructure technologies to develop eco-driving strategies aimed at optimizing vehicle speed. One of these strategies is the implementation of SCG, which aims to optimize the speed of connected vehicles by providing recommended trajectories from the beginning until the end of the route in the presence of signalized intersections.

The significant variability of fuel costs has led companies and governments to advise strategies that reduce energy consumption. Many researchers have examined the influence of different speed optimization methods on driving behaviors (Conlon & Lin, 2019; Khooban, 2019; Y. Lu et al., 2019; Wang et al., 2020). Several studies focused on developing algorithms and methods to improve fuel efficiency (Kramer et al., 2015; Shen et al., 2018; Z. Xu et al., 2021), insights into the effectiveness of speed limit facilities (F. Lu et al., 2016), and reduce travel time (Pourmehrab

et al., 2020) in the presence of signalized intersections and autonomous vehicles (G. Chen et al., 2019; Gamage & Lee, 2016; B. Xu et al., 2019). Speed profiles, especially on urban roads with many signal intersections, have a huge impact on fuel consumption (B. Chen et al., 2019; Liang et al., 2019; Nasri et al., 2018; Talati et al., 2021). Lim et al. proposed a distance-based eco-driving scheme that optimizes speed for an entire route before departure and adapts to real-time traffic conditions during the drive. The method focuses on nearby heavy traffic regions for adaptation while maintaining the effectiveness of the optimized speed profile elsewhere. The scheme aims to improve fuel efficiency and vehicle performance through real-time adaptation and long-term optimization (Lim et al., 2017).

This study employed a driving simulator with the SCG system to assess its effectiveness in optimizing speed control for diverse vehicles at signalized intersections. The research focused on color-coded speed information provided by the SCG system. Using the simulator, the study investigated drivers' responses to an in-vehicle SCG system offering real-time speed guidance throughout the entire route. The primary goal was to examine drivers' behavior when presented with color-coded speed guidance, specifically assessing compliance percentages and reactions, along with the ability to pass through intersections when signals are green.

METHODOLOGY AND DATA

Speed Control Guidance System

The Virginia Tech Transportation Institute (VTTI) developed a dynamic Speed Control Guidance (SCG) system (Almannaa et al., 2017; H. Chen et al., 2016) that calculates optimal vehicle trajectories near signalized intersections. This system incorporates SCG algorithms, allowing vehicles to identify suitable speed profiles based on information from nearby vehicles and upcoming intersections. Operating through an in-vehicle device, the system receives Signal Phase and Timing (SPaT) information via Vehicle-to-Infrastructure (V2I) communication. Utilizing this data, the system provides real-time speed recommendations to drivers, facilitating smooth passage through intersections when followed accurately. The SCG algorithm continually computes the optimal speed, ensuring vehicles approach intersections during the green or yellow phase to avoid unnecessary stops.

To present this information to drivers effectively, our algorithm displays it as either a green arrow up or a red arrow down. Through previous studies (Ahangari et al., 2019), we found this method of information provision to be the least confusing and most effective. It is worth noting that a previous version of this system was tested with the SCG functioning only near the intersections, and the intersections were separate entities. However, this time, we have successfully implemented the SCG to operate continuously throughout the entire network.

Driving Simulator

This study utilizes a 3D driving simulator with a Speed Control Guidance (SCG) system powered by VR-Design Studio software by Forum8 Company (*Forum8*, n.d.). The simulator closely mimics a real car's hardware, featuring a cockpit with essential components like pedals, a steering wheel, and three monitors for realistic visualizations. Safety features, including a seat belt and hazard button, are incorporated. The software enables the creation of lifelike driving scenarios with real-world elements, and the simulator records comprehensive data on driver and vehicle behavior, including speed, acceleration, and traffic signal information, every second.

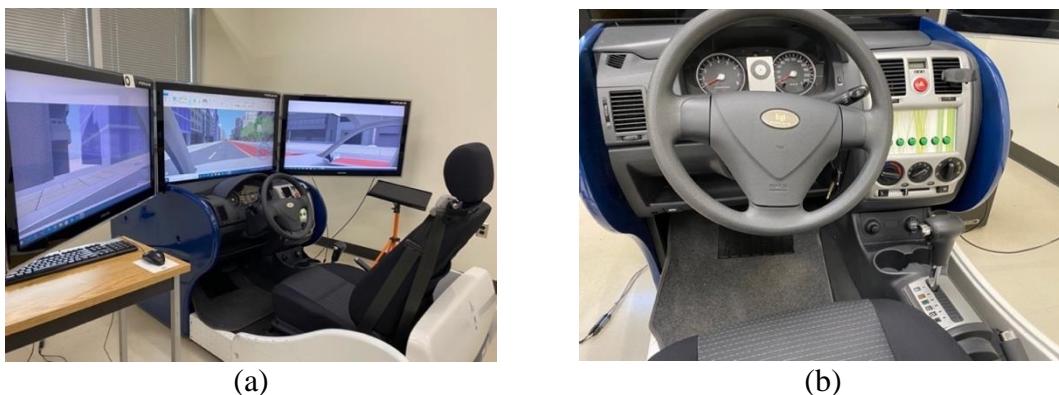


Figure 1. Views from (a) outside and (b) inside the driving simulator at MSU

Driving Scenarios

To study drivers' behavior, we created a road segment with eight signalized intersections, implementing four scenarios to simulate different weather conditions. Each scenario lasted three to four minutes. Scenario 1, the base scenario in sunny weather, assessed driving behavior without speed guidance. Scenarios 2-4 included recommended speeds for sunny, rainy, and nighttime conditions, aiming to facilitate smooth intersection passage without complete stops if participants adhered to the speeds.

To isolate the SCG system's effects and assess its impact on compliance, the road segment had a single lane in each direction. This setup allowed us to analyze the SCG system's direct influence on driver behavior, providing insights into its effects on compliance at signalized intersections. The study area consisted of a one-lane road network comprised of eight signalized intersections. The participants began driving in the base scenario, which served as a point of comparison for their driving behavior in relation to the other scenarios. Subsequently, participants experienced various scenarios with SCG provided throughout the entire network, including all eight intersections. In each scenario, participants received "Speed Guidance" through a color code system. In the Speed Guidance scenarios, a "Green Arrow" indicated the need to accelerate, while a "Red Arrow" signaled the need to decelerate (Figure 2). Participants were instructed to drive at a constant speed limit of 30 mph and adjust their speed according to the information provided via SCG to pass through the signalized intersections without stopping. The objective of the study was to assess the 'participants' ability to follow the SCG. Participants were instructed to follow an acceleration range of 1-3 mph in various driving scenarios. This guidance aimed to ensure a clear understanding of the designated speed range for acceleration. Additionally, the distances between intersections were strategically designed to prompt participants to encounter both acceleration and deceleration advice. This approach allowed for comprehensive observation of participants' responses to different speed adjustment recommendations throughout the experiments.

Data Collection

Before commencing the study, Institutional Review Board (IRB) approval was obtained to ensure ethical compliance. Recruitment methods included distributing flyers on the Morgan State University (MSU) campus and sending email invitations. Fifteen MSU participants were recruited, receiving detailed information about the study, eligibility criteria, and compensation. A valid

driver's license was a prerequisite, and participants received \$10 as compensation. Driving experiments encompassed four scenarios, including SCG and Base conditions, completed in a random order to eliminate the learning effect. Participants also completed pre-driving and post-driving surveys, providing demographic information in the former (Table 1) and feedback on the driving experience and simulator sickness in the latter (Figure 3).

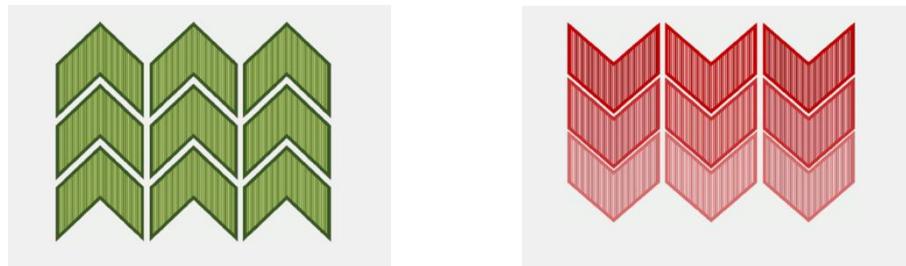


Figure 2. Color code type of speed control guidance

Table 1. Participants' Sociodemographic Characteristics

Variable	Categories	Percent
Gender	Female	46.7
	Male	53.3
Age	18-25	26.7
	26-35	26.7
	36-45	33.3
	46-55	13.3
Educational Status	Undergraduate	26.7
	Graduate	46.7
	Postgraduate	26.7

ANALYSIS AND RESULTS

Compliance Rate

Many studies use statistical analysis to develop policies to improve traffic safety, investigate and forecast travel behavior, and pinpoint deficiencies in transportation policy (Javid et al., 2023; Javid & Sadeghvaziri, 2023b, 2023a; Mehryar & Bandelt, 2022; Sadeghvaziri et al., 2023). A statistical analysis was conducted to evaluate the percentage of drivers who successfully **passed** intersections based on **following** the SCG, as well as to assess the effectiveness of the SCG under different weather conditions affecting visibility (Table 2). This suggests a preference for **following** the SCG when visibility is reduced during nighttime conditions, highlighting the effectiveness of the SCG in such circumstances. To take into consideration the inherent challenge of **following** the SCG accurately, the authors defined compliance as a participant's speed being within 5 mph of the recommended speed (i.e., recommended speed +/- 5 mph). If participants pass more than seven

intersections based on their adherence to the SCG, we consider that they have successfully passed the intersections for the purpose of this analysis. The criterion for **passing** intersections in this study is defined as successfully **passing** at least seven out of the eight intersections by **following** the SCG without stopping at red lights.

Table 2 summarizes the impact of SCG on participant behavior at intersections under various weather conditions. In Scenario 1 (Without SCG - Sunny Weather), participants had a low success rate, averaging 4.9% in **passing** more than seven intersections. In Scenario 2 (With SCG – Sunny Weather), 73% of participants successfully navigated intersections, resulting in an average of 6.9%. In Scenario 3 (With SCG – Rainy Weather), a 73% success rate led to an average of 6.8%. Scenario 4 (With SCG – Night Vision) saw an 80% adherence to SCG but a lower success rate of 40%, averaging 6.3%. Overall, Scenario 2 (With SCG - Sunny Weather) demonstrated the highest average success rate, highlighting SCG's effectiveness in promoting safer and more efficient traffic flow. Conversely, Scenario 1 (Without SCG - Sunny Weather) had the lowest average success rate, emphasizing the importance of SCG.

Table 2. Percentage of All Participants Who Follow SCG and Passed Intersections

Scenario	Follow SCG	Passing >= 7 Intersections	Avg Passing
Scenario 1 (Without SCG – Sunny Weather)	Without SCG	0%	4.9
Scenario 2 (With SCG – Sunny Weather)	73%	73%	6.9
Scenario 3 (With SCG – Rainy Weather)	73%	60%	6.8
Scenario 4 (With SCG – Night Vision)	80%	40%	6.3

ANOVA

Following Speed Control Guidance and Passing Intersections Analysis

An analysis of variance (ANOVA) was used to determine the proportion of drivers who **follow** the SCG and **passed** intersections based on different scenarios. Table 3 presents the findings of a study evaluating participants' adherence to SCG in three different scenarios. The "**Following SCG**" column shows the percentage of participants **following** SCG across scenarios for each category. In Scenario 2 (With SCG - Sunny Weather) and Scenario 3 (With SCG - Rainy Weather), younger participants aged 18-35 demonstrated higher **following** rate to SCG compared to those above 36. However, in Scenario 4 (With SCG - Night), participants aged 26-45 displayed a higher **following** rate of SCG among other age groups. Consistently across all three scenarios, male participants and participants with a graduate education level demonstrated higher **following** rate compared to others.

The "**Passing Intersection**" column shows the percentage of participants who **passed** more than seven intersections across scenarios for each category. In Scenario 2 (With SCG - Sunny Weather) and Scenario 3 (With SCG - Rainy Weather), younger participants aged 18-35 and participants with graduate level demonstrated higher **passing** rate compared to others. However, in Scenario 4 (With SCG - Night), participants aged 36-45 displayed a higher **passing** rate among other age groups. In all education subcategories, participants showed similar performance of

pasing rate within each category. Consistently across all three scenarios, female participants demonstrated higher **passing** rate compared to others. The ANOVA results show there is no signficancy different between **passing intersection** and **following SCG** among different scenarios.

Table 3. ANOVA Results for Following SCG and Passing Intersection

Scenario Name	Variable	Category	Following SCG			Passing Intersection		
			Percentage	F	Sig.	Percentage	F	Sig.
Scenario 2 (With SCG – Sunny Weather)	Age	18-25	36%	0.11	0.89	27%	1.75	0.19
		26-35	36%			27%		
		36-45	18%			18%		
		46-55	9%			18%		
	Gender	Male	55%			45%		
		Female	45%			55%		
	Education	Undergraduate	36%			27%		
		Graduate	55%			55%		
		Postgraduate	9%			18%		
Scenario 3 (With SCG – Rainy Weather)	Age	18-25	36%			33%		
		26-35	36%			33%		
		36-45	18%			22%		
		46-55	9%			11%		
	Gender	Male	55%			33%		
		Female	45%			67%		
	Education	Undergraduate	36%			33%		
		Graduate	55%			44%		
		Postgraduate	9%			22%		
Scenario 4 (With SCG – Night Vision)	Age	18-25	25%			33%		
		26-35	33%			0		
		36-45	33%			50%		
		46-55	8%			17%		
	Gender	Male	58%			33%		
		Female	42%			67%		
	Education	Undergraduate	25%			33%		
		Graduate	50%			33%		
		Postgraduate	25%			33%		

Regression Analysis

Following SCG and Passing Intersections for All Three Scenarios

To identify the relationship between the SCG **following** percentage and **passing** intersection as a dependent variables and the sociodemographic of participants as independent variables, binary

logistic regression analyses were performed on dataset. As shown in Table 4, the regression results for the **following** SCG and **passing** intersections as dependent variables with age, gender, and education as independent variables in all three scenarios, are as follows:

Table 4. Regression Results for Following SCG and Passing Intersections

Scenario	Variable	Coefficient	Standard Error	Sig.
Following SCG	age	-3.1254	1.475	0.034
	gender	2.2261	1.396	0.111
	education	-4.6858	1.804	0.009
Passing Intersections	age	0.5187	0.574	0.366
	gender	-1.6142	0.755	0.033
	education	0.2010	0.803	0.802

According to Table 4, there is a negative relationship between age and the **following** of SCG; younger drivers tend to **follow** SCG more than older drivers. Furthermore, there is a negative relationship between education and **following** SCG; drivers with an undergraduate degree are more likely to **follow** SCG. In this case, age and education have statistically significant effects on the likelihood of **following** the recommended speed. Moreover, binary logistic regression analysis was performed on the combined data to determine the relationship between the proportion of participants who pass at least seven intersections while the traffic light is green and the 'participants' sociodemographic characteristics as independent variables. According to the findings presented in Table 4, there is a negative relationship between gender and the likelihood of **passing** at least seven intersections while the traffic signal is green. Female drivers are more likely to surpass this threshold compared to male drivers.

Survey Analysis

Post Survey

The figures below present a descriptive analysis of post-driving surveys. In Figure 3 (a), participants' perceptions of SCG reveal that 13% found it useless, 7% considered it unhelpful, and about 33% had a neutral stance. In contrast, 20% found SCG helpful, while 27% found it extremely helpful. Figure 3 (b) shows opinions on the easiness of following SCG, with 20% strongly disagreeing, indicating challenges. No participants strongly agreed or disagreed, reflecting a lack of consensus. Approximately 53% had a neutral stance, while 27% agreed that following SCG was easy.

Moving to Figure 3 (c), 33% preferred the base scenario with no SCG for its autonomy, while 53% favored SCG in sunny weather for its benefits in speed consistency. Only 7% each preferred SCG in rainy weather or at night, indicating perceived benefits in challenging conditions. Figure 3 (d) shows that only 7% found the base scenario most challenging, while 60% reported SCG in rainy weather as the most challenging. In contrast, 33% found SCG at night challenging. These findings highlight challenges in rainy weather and nighttime driving, while SCG in sunny weather alleviates perceived difficulties.

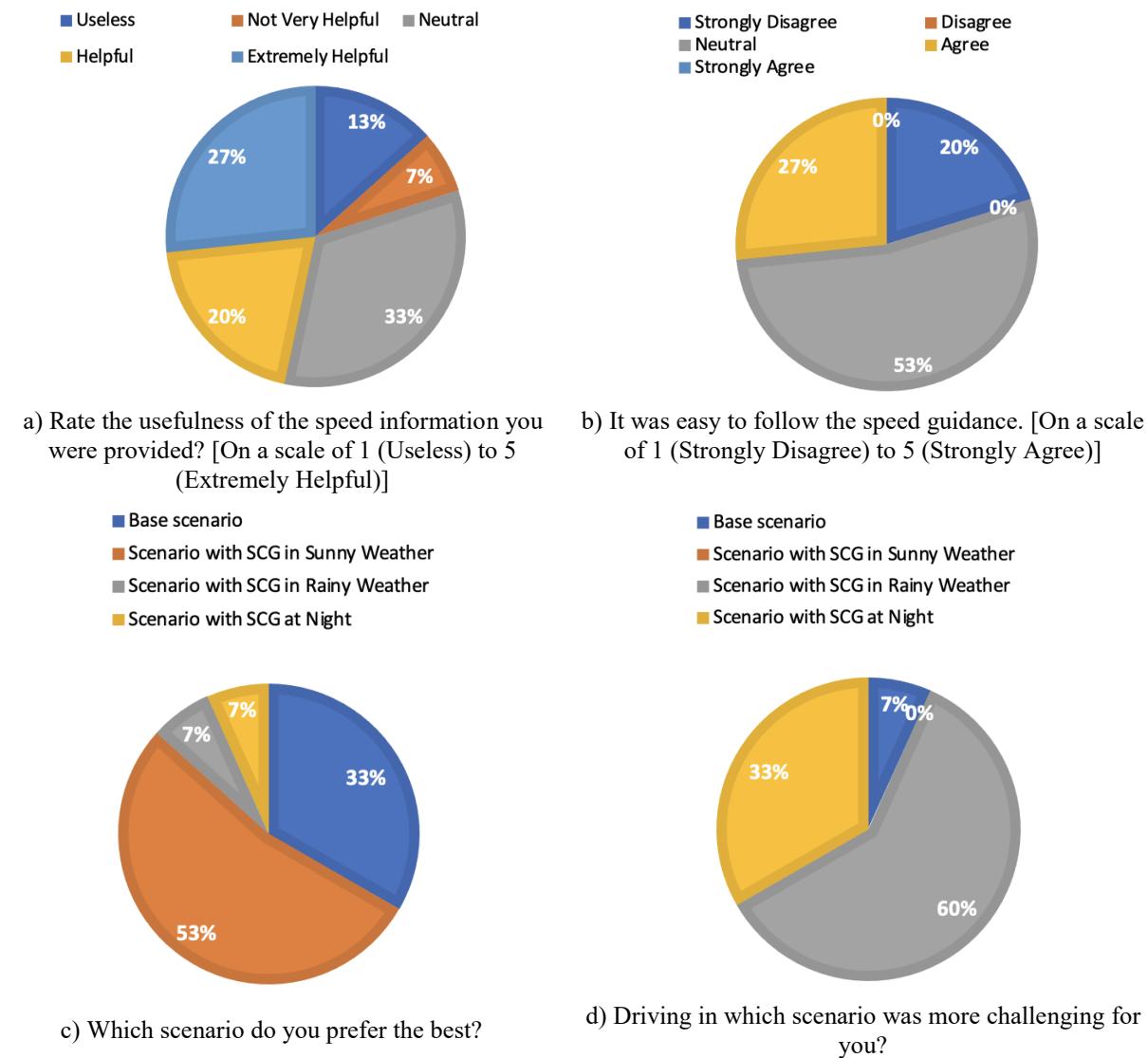


Figure 3. a) Usefulness of Speed Control Guidance b) Easiness of following Speed Control Guidance c) Preference of Type of Scenarios d) Comparison of Perceived Driving Challenge Between Scenarios

Taken together, the findings obtained from Figure 3 (c) and Figure 3 (d) highlight the significant influence of weather conditions, particularly visibility, on participants' preferences and perceived challenges in driving scenarios. Sunny weather with good visibility, combined with SCG, was generally preferred, and mitigated perceived difficulties. On the other hand, rainy weather, and nighttime driving, characterized by reduced visibility, posed challenges that even the presence of guidance systems could not fully overcome.

CONCLUSION AND DISCUSSION

This study employed a driving simulator to evaluate drivers' adherence to SCG under different weather and visibility conditions. The investigation aimed to assess SCG's impact on successfully

passing a minimum of seven intersections without stopping at red signals. In sunny conditions, male drivers tended to follow SCG more than females but had lower success in passing intersections. Younger participants displayed better adherence to follow SCG, while graduates and full-time workers exhibited higher compliance and success rates. Lower-income participants, especially those earning between \$20,000 and \$30,000, demonstrated a higher propensity to follow SCG. Under rainy conditions, male drivers again showed higher SCG adherence but lower success rates. Younger age groups, graduates, and full-time workers outperformed others. Notably, lower-income participants, particularly in the \$75,000 to \$100,000 range, excelled in passing intersections. In low visibility scenarios, male drivers demonstrated better SCG adherence but lower success rates. Middle-aged groups performed well, and graduates with full-time jobs showed superior compliance. Participants earning between \$20,000 and \$75,000 exhibited higher success rates. Overall, SCG positively influenced participants, increasing the likelihood of passing more than seven intersections. However, effectiveness varied with weather conditions, being more successful in sunny and rainy scenarios than at night.

A limitation is the divergence between simulator and real-world driving experiences. Implementing color-coded speed recommendations in actual driving scenarios is suggested for more reliable conclusions. The study focused solely on straight roads, neglecting turns. Future research should encompass varied turning scenarios to comprehensively evaluate the color-code SCG's effectiveness. The absence of actuated traffic lights in the simulator is another limitation. Incorporating these in future studies could enhance realism and applicability.

AUTHOR CONTRIBUTIONS

Study conception and design: P. Masoumi, E. Sadeghvaziri, M. Jeihani; data processing: P. Masoumi, E. Sadeghvaziri, M. Jeihani; analysis and interpretation of results; P. Masoumi, E. Sadeghvaziri, M. Jeihani; draft manuscript preparation: P. Masoumi, E. Sadeghvaziri, M. Jeihani. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Ahangari, S., Rashidi Moghaddam, Z., Jeihani, M., Chavis, C., Chen, H., Rakha, H., Kang, K., and Transportation Research Board. (2019). Investigating the Effectiveness of an Eco-Speed Control System in the Vicinity of Signalized Intersections Using a Driving Simulator (01698210). 20p. <https://trid.trb.org/view/1573249>.
- Almannaa, M. H., Chen, H., Rakha, H. A., Loulizi, A., and El-Shawarby, I. (2017). Reducing Vehicle Fuel Consumption and Delay at Signalized Intersections: Controlled-Field Evaluation of Effectiveness of Infrastructure-to-Vehicle Communication. *Transportation Research Record*, 2621(1), 10–20.
- Chen, B., Evangelou, S. A., and Lot, R. (2019). Series Hybrid Electric Vehicle Simultaneous Energy Management and Driving Speed Optimization. *IEEE/ASME Transactions on Mechatronics*, 24(6), 2756–2767. Scopus.
- Chen, G., Zhang, W., Li, X., and Yu, B. (2019). Adaptive speed control method for electromagnetic direct drive vehicle robot driver based on fuzzy logic. *Measurement and Control*, 52(9–10), 1344–1353.
- Chen, H., Rakha, H. A., Loulizi, A., El-Shawarby, I., and Almannaa, M. H. (2016). Development and preliminary field testing of an in-vehicle eco-speed control system in the vicinity of signalized intersections. *IFAC-PapersOnLine*, 49(3), 249–254.

- Conlon, J., and Lin, J. (2019). Greenhouse Gas Emission Impact of Autonomous Vehicle Introduction in an Urban Network. *Transportation Research Record*, 2673(5), 142–152.
- Forum8. (n.d.). Retrieved July 13, 2023, from <https://www.forum8.co.jp/>.
- Gamage, H. D., and Lee, J. (2016). A conceptual demonstration of self-learning eco-speed control at signalised intersection (01676251). 13p. <https://arrbknowledge.com>.
- Javid, R., and Sadeghvaziri, E. (2023a). Assessing Gender Gaps in Citi Bike Ridership in New York City. Available at SSRN 4358660.
- Javid, R., and Sadeghvaziri, E. (2023b). Equity Analysis of Bikeshare Access: A Case Study of New York City. *Findings*.
- Javid, R., Sadeghvaziri, E., and Jeihani, M. (2023). A Driving Simulator Study to Understand the Impact of Cell Phone Blocking Apps on Distraction. *Highlights of Vehicles*, 1(1), 17–28.
- Khooban, M. H. (2019). Hardware-in-the-loop simulation for the analyzing of smart speed control in highly nonlinear hybrid electric vehicle. *Transactions of the Institute of Measurement and Control*, 41(2), 458–467.
- Kramer, R., Maculan, N., Subramanian, A., and Vidal, T. (2015). A speed and departure time optimization algorithm for the pollution-routing problem. *European Journal of Operational Research*, 247(3), 782–787. Scopus.
- Liang, X. J., Guler, S. I., and Gayah, V. V. (2019). Joint Optimization of Signal Phasing and Timing and Vehicle Speed Guidance in a Connected and Autonomous Vehicle Environment. *Transportation Research Record*, 2673(4), 70–83.
- Lim, H., Su, W., and Mi, C. C. (2017). Distance-Based Ecological Driving Scheme Using a Two-Stage Hierarchy for Long-Term Optimization and Short-Term Adaptation. *IEEE Transactions on Vehicular Technology*, 66(3), pp 1940-1949.
- Lu, F., Tian, X., and ASCE. (2016). A Comparative Study on Speed Control Effects on an Expressway (01606951). pp 1485--1496.
- Lu, Y., Xu, X., Ding, C., and Lu, G. (2019). A Speed Control Method at Successive Signalized Intersections Under Connected Vehicles Environment. *IEEE Intelligent Transportation Systems Magazine*, 11(3), pp 117-128.
- Mehryaar, E., and Bandelt, M. J. (2022). Geological Modeling Along Tunnel Projects Using Machine Learning Techniques. *ARMA/DGS/SEG International Geomechanics Symposium*.
- Nasri, M. I., Bektaş, T., and Laporte, G. (2018). Route and speed optimization for autonomous trucks. *Computers and Operations Research*, 100, 89–101. Scopus.
- Pourmehrab, M., Elefteriadou, L., Ranka, S., and Martin-Gasulla, M. (2020). Optimizing Signalized Intersections Performance under Conventional and Automated Vehicles Traffic. *IEEE Transactions on Intelligent Transportation Systems*, 21(7), 2864–2873. Scopus.
- Sadeghvaziri, E., Javid, R., Jeihani, M., and Center, E. (2023). *Investigating Walking and Biking Activities Among Low-Income African Americans*.
- Shen, D., Karbowski, D., and Rousseau, A. (2018, April 3). Fuel Efficient Speed Optimization for Real-World Highway Cruising (01723487). SAE Technical Paper.
- Talati, S., Vekaria, D., Kumari, A., and Tanwar, S. (2021). An AI-driven object segmentation and speed control scheme for autonomous moving platforms. *Computer Networks*, 186, 107783.
- Wang, P., Jiang, Y., Xiao, L., Zhao, Y., and Li, Y. (2020). A joint control model for connected vehicle platoon and arterial signal coordination. *Journal of Intelligent Transportation Systems*, 24(1), 81–92.

- Xu, B., Chen, X., Li, K., Hu, M., Bian, Y., Yu, Q., and Wang, J. (2019). Double-layer speed optimization for reducing fuel consumption with vehicle-to-infrastructure communication. *Journal of Intelligent Transportation Systems*, 23(5), pp 513-524.
- Xu, Z., Pan, L., and Shen, T. (2021). Model-free reinforcement learning approach to optimal speed control of combustion engines in start-up mode. *Control Engineering Practice*, 111, 104791.

Exploring the Role of Human Behavior in Road Crash Occurrence through Analysis of Crash Variability in Highly Similar Road Environments

Nitesh Acharya¹ and Michael Henry, Ph.D.²

¹Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo, Japan; Ministry of Physical Infrastructure and Transportation, Government of Nepal.

Email: na21504@shibaura-it.ac.jp

²Dept. of Civil Engineering, Shibaura Institute of Technology, Tokyo, Japan.

Email: mwhenry@shibaura-it.ac.jp

ABSTRACT

Crashes occur due to the influence of the interaction between three factors: humans, vehicles, and the road environment. However, the human factor is often found to be the most dominant factor responsible for crash occurrences. The objective of this study is to assess the contribution of human behavior to the occurrence of road crashes by analyzing crash occurrence in highly similar road environments. Homogeneity in the road environment was achieved by clustering road segments with highly similar characteristics and by controlling the effect of vehicle-related factors, and the variation in crash occurrence between and within clusters was assessed to explore the role of human behavior. The approach developed in this study can be used to understand the variability of the human factor and its contribution to crash occurrence without knowing drivers' actual behavior. Furthermore, the results can be applied by road management authorities to mitigate road crashes and to prioritize and select appropriate interventions.

Keywords: road traffic crash, road environment, human behavior, cluster analysis

1. INTRODUCTION

1.1 Background

Road traffic crashes are increasing significantly in low- and middle-income countries while remaining a major problem for developed countries. Even though low-and-middle-income countries have approximately 60% of the world's total vehicle fleet, 93% of the world's total crashes occur in these countries (WHO 2022). Crashes may be occurring in these countries because of poor road infrastructure, weaker regulations, and ineffective management. With the increase in crash rates, the road management authorities in these countries are also facing challenges to mitigate the crashes. Considering the seriousness of this global issue, Sustainable Development Goal (SDG) target 3.6 was formulated to halve the global fatalities and serious injuries due to road traffic crashes by the year 2030.

Road traffic crashes occur due to the influence of humans, vehicles, and roadway factors. A study performed in Indiana, USA, by Treat et al. (1979) to study the causes of road crashes found that around 93% of road crashes occur due to factors involving humans. A summary of these results can be seen in Figure 1, and they are discussed in more detail in the Highway Safety Manual (AASHTO 2010). Vehicle factors, which include the power, size, design, maintenance, and features like braking system and navigation aids which enhance safety, are found to have the

least contribution among the three. While the road environment itself does not have much role in crash occurrence, its combined effect with human behavior cannot be ignored, as 27% of crashes occur due to the interaction between these two factors. Humans are easily and heavily influenced by road environments (FHWA 2017). The road environment comprises the natural conditions and man-made infrastructures on and around the road, and the roadway and its components form a part of driving conditions. Road design and geometry, surrounding infrastructures, traffic control devices, and existing safety measures are the major components that characterize the road environment. In addition, conditions like weather and topography can also impact drivers' ability to respond to external stimuli (Oppenheim & Shinar 2012).

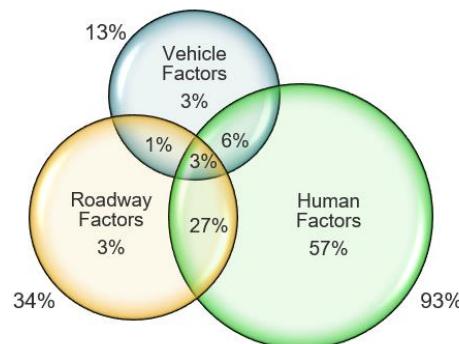


Figure 1: Venn diagram showing contributing factors to road crashes (adapted from the Highway Safety Manual published by AASHTO).

Humans perceive different information while driving and the ability to perceive, process and act depends on human factors like experience, habit, emotions, cultural norms, education, age, and fatigue, among many others (Petridou & Moustaki 2000). Humans are also prone to make mistakes, which can lead to fatal road crashes. The human factor comprises the driver's or road users' behavior, which is governed by two systems: deliberate and intuitive. In the deliberative system, once the road user receives information from the surrounding, the person processes the information and acts rationally, whereas the intuitive system is implicit, the person acts instantaneously, and the outcome of that act may or may not be favorable (FHWA 2017). Road crashes involving young drivers are found to be the result of factors like inexperience and over speeding, whereas those involving old drivers were due to physical and mental changes like a reduction in cognitive function (Bucsuhazy et al. 2020). Various conclusions have been drawn from previous research considering human and roadway environmental factors and their interrelationship. A study conducted in Alabama, USA, to study regional variability and driver characteristics contributing to fatal road crashes showed that 6.19% of the variability in fatal crash rates could be attributed to the city, 3.84% to the county, and the remaining 89.97% of variability to the driver's characteristics and other contributing factors (Adanu et al. 2021). Another study conducted in Oman to explore the human-road interaction for crash occurrence found that, among various road elements, drivers consider roads with wider carriageways to be safer, and factors like inappropriate speed, fatigue, inattentional (perceptual) blindness, and having a conversation while driving is related to road design elements (Plankermann 2013).

After a review of existing literature, it was found that most of the research on road traffic crashes has been focused on the prediction of crash frequency and severity using features like

age, sex, education, income, and other demographic features, as well as the impact of enforcement on human behavior. However, there is very limited literature that has specifically analyzed the occurrence of crashes by incorporating various aspects of road environment like geometry, safety features, traffic, and pavement conditions and their influence on humans while driving.

1.2 Research objectives

The primary objective of this study is to explore the role of humans in road crash occurrences. Since it is difficult to directly measure the extent of human behavior on crash occurrence, a novel approach is introduced in this paper to analyze the interrelationship between road environment and human behavior by assessing crash variability in highly similar driving conditions. It is hypothesized that if the road environments are the same and the impact of vehicle-related factors on road crashes is minimized, then the resulting crashes may be attributed to human behavior and, thus, can be studied through the variability in crash frequency between and within highly similar groups of road segments.

2. METHODOLOGY

2.1 Data collection

A mountainous, semi-urban highway named Dhulikhel-Sindhuli-Bardibas (H13) highway of Nepal was considered for the study. The 160-kilometer highway has huge socio-economic importance as it connects the capital Kathmandu to the plains and the eastern region of the country (Figure 2). Geographically, the highway passes through rolling, mountainous, and steep terrain, as well as through both river and hilly routes. Road sections following hilly routes are narrow, steep, and windy, whereas the sections following river routes are relatively straight and have flat gradients. Due to the difficult terrain in the mountains, the shoulder widths are restricted, making the roads even narrower, and the presence of multiple sharp and short curves makes traffic safety very challenging (Figure 3). Road crash records maintained by the Nepal traffic police show that road crashes are rife on this highway, with most of the crashes attributed to drivers' carelessness and mistakes. Eight-year crash statistics also reveal that there exist multiple crash blackspots, but there are also sections that have not recorded a single road crash to the date of data collection. Furthermore, some of the road sections have recorded road crashes even after adding road safety measures, which implies that the crash occurrence in those segments may be related to human behavior and not just the road environment.

To perform the analysis, it was necessary to divide the highway into smaller segments, hence the highway was divided using a fixed-length approach to segments of 250 meters. The length of 250 meters was considered appropriate for this study because it was judged to satisfactorily reflect the road characteristics, whereas larger segments would lead to a loss of homogeneity. The formation of finer segments, on the other hand, could result in a larger number of segments with zero crash records, which would make the analysis cumbersome. Therefore, taking these considerations into account, 639 road segments were formed and analyzed in this study.

For each road segment, necessary data was collected including information related to the road geometry, traffic volume, pavement condition, existing safety measures, and road traffic crashes. The data on the road geometry and features were extracted from the design and as-built

drawings of the highway and included information about carriageway width, shoulder type and width, length and radius of curves, and longitudinal gradient. The pavement condition of the highway segments was examined using the International Roughness Index (IRI), and Average Daily Traffic (ADT) was used as a parameter to assess traffic volume. The data on IRI and ADT were obtained from the annual survey conducted by the Department of Roads, Nepal, and the project office for the construction, operation, and maintenance of the study highway.

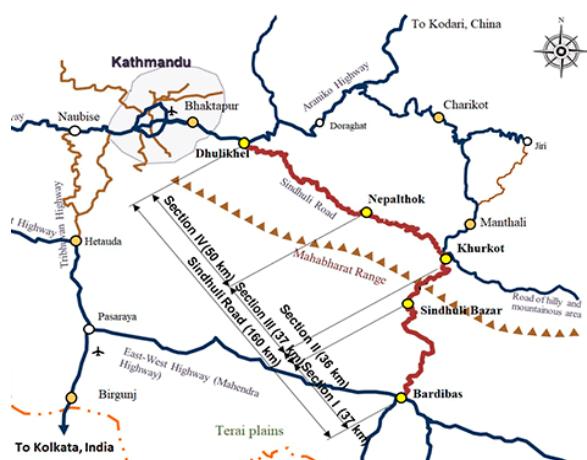


Figure 2: Map of Dhulikhel-Sindhuli-Bardibas Highway (Source: Kamei, 2019)



Figure 3: Photo of the study highway

Data on the existing road features were collected by field survey in July 2022 and include data on the presence and condition of various safety measures like crash barriers, signposts, road markings, delineation measures, pedestrian crossings, sidewalks, and street lightings. In addition, information about the surrounding environment, like the number of accesses in each segment and the level of ribbon development along the highway segments, was also gathered during the field survey. Road traffic crash data for the years 2021 and 2022 were obtained from the records maintained by the Nepal traffic police, which included information on the location of the crash, type of collision or crash, severity of the crash, vehicles involved in the crash, and probable cause of the crash.

2.2 Preparation of dataset

The preliminary data collected from the field survey and other sources were cleaned and processed to produce either numeric or categorical data. In addition, the subjective scores provided by safety inspectors for the existing road features, which were judged in reference to prevailing departmental and international guidelines during the field survey, were converted into three categorical levels (good, fair, and bad) during the data cleaning process. The resulting descriptive statistics of the road features and safety measures are shown in Tables 1 and 2.

The two-year crash database of the study highway comprised 225 crashes. However, 5 crashes were found to be attributed to vehicle factors, such as crashes occurring due to defects in the vehicle or mechanical faults and were removed from the analysis. It is hypothesized that the removal of vehicle-related crashes would result in scenarios involving only road environments and humans and would be easier to assess the role of humans while driving in highly similar road

environments. Among the 639 road segments, there were 482 segments without road crashes and 116 segments had recorded at least one crash in the period of two years. Furthermore, in this study, the number of crashes in each road segment was divided by Average daily traffic (ADT) and was expressed in the unit of crash number per 10,000 vehicles. Since some of the road segments had relatively lower traffic than others, the use of crash per traffic volume would better represent the crash frequency based on the exposure. The descriptive statistics of the ADT and crashes per 10,000 vehicles are also included in Table 1.

Table 1: Descriptive statistics for road features, traffic volume, and road crashes

Road features	Unit	Min.	Max.	Mean	Median	Mode	S.D.
Carriageway width	Meter	4.52	9.89	5.26	5.15	5.50	0.56
Shoulder width: Left	Meter	0.00	2.00	0.47	0.51	0.00	0.43
Shoulder width: Right	Meter	0.00	2.00	0.62	0.65	0.90	0.36
Grade	%	0.00	10.00	3.83	3.64	0.50	2.30
Curve radius	Meter	20.42	1025.00	134.42	86.24	500.00	128.72
Curve length	Meter	7.80	250.00	140.67	141.65	250.00	51.36
Number of accesses	Nos.	0.00	4.00	0.19	0.00	0.00	0.52
IRI	m/km	3.10	10.95	5.12	4.85	4.73	1.17
ADT	Veh./day	3430	12015	6030	5091	8988	2196
Crashes/ 10,000 vehicles	Nos.	0	8.80	0.61	0	0	1.29

Table 2: Frequency table for categorical variables

Road features	Frequency		Road features	Frequency		
<i>a) Shoulder type</i>	<u>Left</u>	<u>Right</u>	<i>d) Existing safety features</i>	<u>Poor</u>	<u>Fair</u>	<u>Good</u>
Composite	24	25	Crash barrier	439	55	145
Gravel	342	390	Centerline marking	620	13	6
Paved	69	167	Edge marking	134	168	337
None	204	57	Edge delineation	113	283	243
<i>b) Ribbon development</i>			Curve delineation	580	-	59
No	411		Pedestrian crossing	586	-	53
Yes	92		Footpath	631	-	8
Partial	136		Road signage	152	260	227
<i>c) Straight section</i>			Streetlight	622	14	3
No	612					
Yes	27					

2.3 Analytical approach

As previously introduced, the objective of this study is to examine the variation in the crash frequency under similar driving conditions. Homogenous driving environments were obtained by

performing cluster analysis to identify groups of road segments with similar road characteristics and existing safety features. Cluster analysis is an unsupervised machine-learning technique for identifying patterns within a dataset. This analytical tool groups objects or observations based on their characteristics such that there exists high intra-cluster similarity and low inter-cluster similarity and helps to gain valuable insight about the data by determining which group an observation falls into. Clustering is also useful when it is difficult to analyze individual observations, but the formation of clusters can provide more meaningful results.

There are various clustering techniques, the most common of which are distance-based clustering and density-based clustering. The selection of the clustering technique depends on the type of data and the objective of the clustering. In this study, the objective of the clustering was straightforward: to form clusters of road segments having similar characteristics. Hence, a distance-based clustering technique called partitioning clustering was used to form the k-number of clusters. The number and size of clusters are very important when analyzing any dataset using cluster analysis. As this study is focused on intra-cluster properties and variation, there was no restriction for the number of clusters. However, for the intra-cluster analysis, it was necessary to have at least two observations (road segments) in any cluster because it would be impossible to compare the crash variation within the segments if the cluster had only a single road segment.

In partitioning clustering, the observations are grouped into a specified number of clusters based on their similarities. There are two types of partitioning clustering; K-means clustering and K-medoid clustering (also called Partitioning Around Medoids, or PAM). In PAM clustering, the cluster is represented by one observation in the cluster, which is called the cluster medoid, and observations similar to the medoid are grouped into that cluster. As the data for this analysis is mixed-type data (numerical and categorical), Gower distance was considered as the suitable metric to measure the distance between observations (Gower 1971). Since PAM is less sensitive to outliers (Kassambara 2017) and is found to be effective with Gower distance (Wiper et al. 2022), it was used in this analysis. The open-source statistical analysis tool R (<https://cran.r-project.org/>) was used for analyzing the data. R package StatMatch (D’Orazio 2022) was used for finding the Gower distance, and package-cluster (Maechler et al. 2022) was used for clustering the data. After the formation of the clusters, road crash frequency in each road segment was assigned and the crash frequency was assessed for each formed cluster.

3. RESULTS AND DISCUSSION

3.1 Results of cluster analysis

As a result of cluster analysis, 45 clusters were identified to meet the criteria set for this analysis, which was to have no single-member clusters. The Dunn Index was used to assess the clustering tendency. It is calculated as the ratio of the lowest inter-cluster distance to the highest intra-cluster distance. The higher the Dunn Index, the better the clustering (Dunn 1974). The Dunn Index for this cluster analysis was 0.032, which means that the clusters are not very compact, and are not well-separated. However, for an analysis of this type with multiple features, it is very difficult to create highly compact and well-defined clusters. More compact clusters could be obtained by dividing them into even more similar clusters, but this would also increase the number of clusters and the likelihood of forming single-member clusters.

Figure 4 shows the number of road segments grouped in different clusters. The results show that the largest cluster was Cluster 32, with 39 road segments, and the smallest clusters were Clusters 1 and 41, with 4 road segments each.

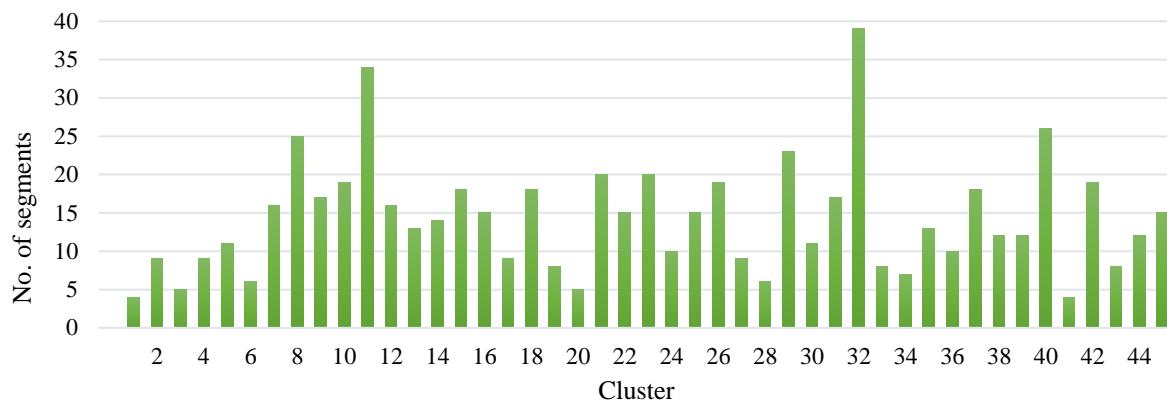


Figure 4: Histogram for the number of segments in each cluster

3.2 Variability of crash occurrence within clusters

While road segments within each of the formed clusters are likely to have similar characteristics, the number of crashes in the road segments within the clusters varies. Figure 5 shows the scatter plot of in-cluster means and standard deviations of crashes per 10,000 vehicles. It can be seen that there exists a somewhat linear relationship between the mean and standard deviation of the road crashes per 10,000 vehicles (Pearson's $r = 0.87$). This implies that the variation in the occurrence of crashes increases as the number of crashes increases, and this variation can be related to human behavior because the driving conditions are almost the same within the cluster.

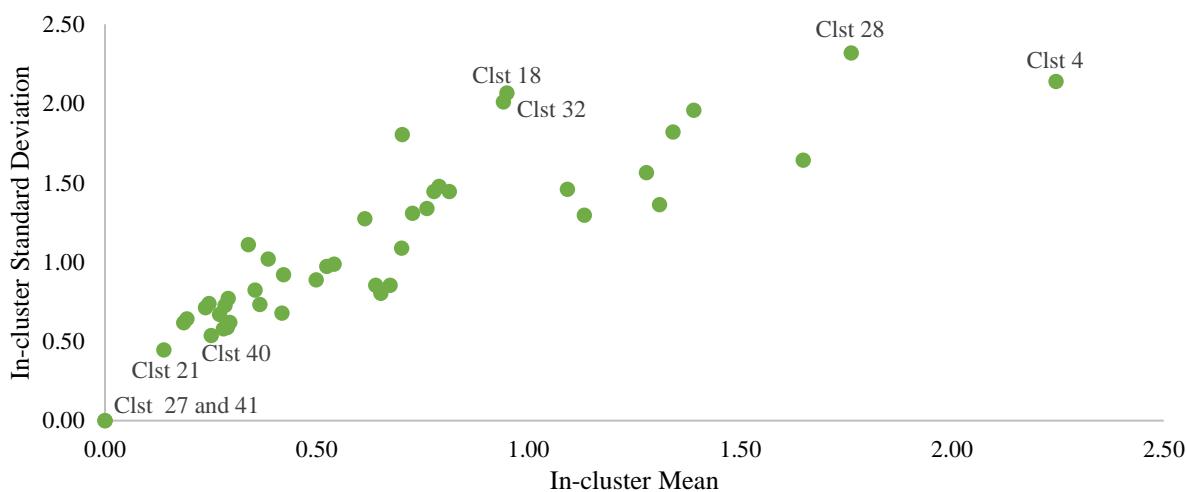


Figure 5: Scatter plot of in-cluster means and standard deviations of crashes per 10,000 vehicles

It would be cumbersome to analyze all the clusters and, since this study is more oriented toward analyzing the variability in crashes within highly similar road environments, a total of eight clusters (four with the highest variabilities and four with the lowest variabilities, as measured by in-cluster standard deviation in mean crashes per 10,000 vehicles) were selected for

further analysis. These clusters have been labeled in Figure 5 to distinguish them from other clusters. Tables 3 and 4 show more details of the clusters having the highest and least crash variability, respectively. The characteristics of the clusters having high and low crash variability can be compared to identify the road features that contribute to road crashes.

Table 3: Summary characteristics of clusters having high crash variability

Road features	Cluster 4		Cluster 18		Cluster 28		Cluster 32	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Carriageway width	6.24	1.15	5.45	0.19	5.82	0.35	5.08	0.22
Grade	2.57	0.84	2.07	1.27	6.89	2.00	4.95	1.74
Curve radius	264.92	191.92	306.95	145.62	47.33	32.93	93.10	106.96
Curve length	115.07	57.01	126.04	41.06	184.43	25.88	148.78	44.82
Access numbers	1.56	0.50	0.00	0.00	0.17	0.37	0.00	0.00
IRI	4.35	0.43	6.74	1.33	5.15	0.42	4.84	0.82
Left shoulder width	0.86	0.08	0.79	0.25	0.25	0.30	0.03	0.08
Right shoulder width	0.76	0.24	0.87	0.15	0.46	0.24	0.55	0.11
Left shoulder type	100% gravel		95% gravel		83% gravel		15% gravel	
Right shoulder type	100% gravel		94% gravel		67% paved		95% paved	
Ribbon development	89%		78% partial		0%		0%	
Straight section	0%		11%		0%		0%	
<i>Existing Safety features</i>								
Crash barrier	11%		6% fair		17%		87% good	
Centerline mark	33% good		0%		0%		0%	
Edge mark	89% good		83% good		83% good		100% good	
Edge delineation	78% fair		0%		83% good		98% good	
Curve delineation	0%		0%		100%		5% good	
Zebra marking	89%		11%		100%		0%	
Footpath	11%		0%		0%		0%	
Road signage	78% fair		16.67% fair		83% fair		87% good	
Streetlight	0%		0%		0%		0%	
Crash per 10,000 veh.	2.25	2.14	0.95	2.07	1.76	2.32	0.94	2.01

One distinct characteristic that can be identified by comparing the two sets of clusters is that the clusters having low crash variation (Table 4) have relatively smaller curve radius and carriageway width than the clusters having higher crash variability (Table 3). Furthermore, the clusters having large crash variations also have comparatively large variations in those two features. This shows that drivers tend to drive slowly and carefully when roads are narrower and windy (with sharp curves) with reduced chances of crash occurrence, whereas in the roads with wide carriageways and flatter curves, there is a higher probability of crash occurrence when there is larger variation in those two road features. Other road features like shoulder width and type, grade, IRI, ribbon development, number of accesses, and existing road safety features are found to be almost similar in clusters with both high and low crash variability. From this analysis alone, those features cannot be identified to have any influence on crash occurrence. Hence, to study the influence of road features on human behavior to cause road crashes, the intra-cluster analysis would be a more appropriate approach because it can provide clearer information on how a

change in certain road features, with other road features remaining the same, can impact driving behavior leading to road crash.

Table 4: Summary characteristics of clusters having low crash variability

Road features	Cluster 21		Cluster 27		Cluster 40		Cluster 41	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Carriageway width	4.94	0.18	5.21	0.35	4.98	0.30	5.04	0.20
Grade	3.70	1.39	5.32	2.28	2.95	1.43	3.12	1.15
Curve radius	66.56	48.87	73.00	57.56	68.05	45.05	82.03	29.15
Curve length	142.89	33.70	175.45	44.93	140.95	49.13	130.30	45.23
Access numbers	0.40	0.73	0.11	0.31	0.31	0.54	0.25	0.43
IRI	4.72	0.52	4.96	0.77	4.87	0.77	5.12	0.73
Left shoulder width	1.02	0.24	0.15	0.20	0.06	0.12	0.37	0.36
Right shoulder width	0.22	0.21	0.63	0.15	0.93	0.32	0.54	0.31
Left shoulder type	100% gravel		50% gravel		27% gravel		50% paved	
Right shoulder type	70% gravel		100% composite		92% gravel		50% paved	
Ribbon development	95% partial		0%		96% partial		75%	
Straight section	0%		0%		0%		0%	
<i>Existing safety features</i>								
Crash barrier	10%		22%		20%		0%	
Centerline mark	0%		0%		0%		25%	
Edge mark	80% fair		89% good		80% fair		100% fair	
Edge delineation	97% good		89% fair		62% fair		75% good	
Curve delineation	0%		0%		15%		0%	
Zebra marking	0%		0%		0%		100%	
Footpath	0%		0%		0%		25%	
Road signage	90% fair		89% good		85% fair		100% good	
Streetlight	0%		0%		0%		75% fair	
Crash per 10,000 veh.	0.14	0.45	0.00	0.00	0.25	0.54	0.00	0.00

3.3 Intra-cluster analysis of road crashes with an example

To assess the role of road features in human behavior to cause road crashes, an intra-cluster assessment of a cluster can be performed. As an example, Cluster 4 is discussed here, with its characteristics shown in Table 5. This cluster consists of nine road segments with highly similar features, but there is a huge variation in the number of crashes per 10,000 vehicles. Road segment 23 of this cluster has the highest number of crashes per 10,000 vehicles, which could be the result of certain road features or a group of road features which are different than other road segments. In this case, segment 23 has a relatively shorter curve radius (sharp curve) than other segments in this cluster while, at the same time, the grade is also much steeper. Segment 161 also has a short curve radius, but no crashes have occurred in this section because the grade here is less and the carriageway width is wider. Therefore, the combination of these features could have impacted the driver's behavior resulting in road crashes. Also, in segment 23 there is partial ribbon development together with crash barriers, which other segments in this cluster do not have. These features may psychologically influence drivers to drive at higher speeds because drivers feel safer when crash barriers are present and when there is less side friction (activities at

the side of the road) due to partial ribbon development. This analysis shows that, when humans drive in a similar road environment and if there is abrupt change in one or combination of any road features, there is chance that drivers get influenced and crash may occur if the driver is unable to cope with change in environment. Intra-cluster analysis of other clusters can be performed in a similar way to identify road features that relate to an increase in crash occurrence due to human behavior. This novel analytical method can be easily adapted by road safety practitioners to assess the role of humans in crash occurrences.

Table 5: Intra-cluster characteristics of Cluster-4

Segment number	5	9	23	24	77	154	160	161	376
Crash/ 10,000 veh.	2.12	3.21	6.70	4.48	0.0	1.84	1.87	0.0	0.0
Carriageway width	5.50	5.50	5.50	5.50	5.50	7.56	8.76	6.92	5.42
Grade	2.34	1.84	4.08	2.33	2.24	4.00	1.66	2.70	1.91
Curve radius	500.00	557.50	58.07	237.90	500.00	148.58	242.24	40.00	100.00
Curve length	204.80	130.61	103.64	151.97	15.55	126.32	166.30	33.39	103.09
Access numbers	2.00	1.00	1.00	2.00	1.00	2.00	2.00	1.00	2.00
IRI	3.70	4.51	4.69	4.69	4.19	5.05	4.33	4.33	3.68
LT shoulder width	0.90	0.90	0.90	0.90	0.90	0.96	0.75	0.75	0.75
RT shoulder width	0.90	0.90	0.90	0.90	0.90	0.11	0.75	0.75	0.75
LT shoulder type	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel
RT shoulder type	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel	Gravel
Ribbon-develop	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Yes
Straight section	No	No	No	No	No	No	No	No	No
Crash barrier	Poor	Poor	Good	Poor	Poor	Poor	Poor	Poor	Poor
Centerline mark	Poor	Poor	Poor	Poor	Poor	Fair	Good	Good	Good
Edge mark	Good	Good	Good	Good	Fair	Poor	Good	Good	Good
Edge delineation	Good	Poor	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Curve delineation	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
Zebra marking	Poor	Good	Good	Good	Good	Good	Good	Good	Good
Footpath	Poor	Poor	Poor	Poor	Good	Poor	Poor	Poor	Poor
Road signage	Fair	Fair	Good	Fair	Fair	Fair	Fair	Fair	Good
Streetlight	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor

4 CONCLUSIONS

The objective of this study was to explore the role of human behavior in crash occurrences by assessing highly similar driving conditions and to explore the interrelationship between road environment and human behavior. Previous studies have shown that humans are easily influenced by the road environment and its surroundings, which may lead to erroneous decisions and, subsequently, road crashes. Therefore, to mitigate road crashes, it is essential to study how the road environment can affect human behavior.

A semi-urban mountainous road in Nepal was selected and divided into finer segments to assess its road features and the number of road crashes occurring in those segments. K-medoid clustering was used to partition and cluster road segments with similar road environments, and a

total of 45 clusters were formed based on their road geometry, existing safety features, and pavement conditions. The road crash frequency per 10,000 vehicles recorded for two years in those segments (clusters) was used to assess the average crashes and their variation in each cluster. The clustering of road segments was conducted aiming for a highly similar road environment within each cluster, and hence it is very likely that any crash variability in those clusters may be due to differences in human behavior. Initially, inter-cluster assessment was performed by identifying and comparing the clusters with very high and low crash variations. The results showed that there are more crash occurrences in the roads having wider carriageways and larger curve radius, and these crashes could occur due to large variations in these two geometric properties of the road. Apart from these two features, there were no other features in either set of clusters having significantly distinct features that can relate to high or low crash variations. The preceding inter-cluster analysis was followed by an example of intra-cluster analysis, where it demonstrated how initially certain road segments having higher crashes can be identified within a cluster having highly similar road features and then its properties be assessed and related with its influence in human behavior. In this way, through this study, it was possible to analyze the role of humans in crash occurrences without having any information about the drivers and their actual behavior.

These results and the concept of this study can be helpful for road management authorities involved in mitigating road crashes. Road features may influence drivers to make mistakes like over speeding, neglecting pedestrians, reckless overtaking, etc., and these actions may lead to road crashes. Upon identifying such significant road features, road management authorities can easily manage them to reduce their effect on humans. For instance, road sections with sharp road curves can be treated with multiple interventions like hillside slope trimming to increase visibility, construction of on-lane rumble strips to reduce speed, and so on. Identification, prioritization, and selection of the most effective interventions can be conducted using this concept, which can serve as a robust reactive approach to treating road crashes blackspots. This study can be further investigated by creating more identical road environments and considering road crash data over a longer period. The results thereafter can provide much clearer and interpretable results. Furthermore, this analytical method can be used for other road types and in other countries where there are different road environments and driving behaviors. Moreover, the novel approach used in this study to assess the interaction between road environment and humans in crash occurrence demonstrates an effective way to analyze road crashes, and it is believed that this concept can provide other researchers with new directions to study the role of humans in road crash occurrences.

ACKNOWLEDGEMENT

This research was supported by a scholarship for road asset management from the Japan International Cooperation Agency (JICA). The authors are grateful to the Department of Roads, Nepal, for providing the necessary design and as-built drawings of the study highway.

REFERENCES

- AASHTO. (2010). *Highway Safety Manual*. AASHTO.
- Adanu, E. K., Penmetsa, P., Wood, D., and Jones, S. L. (2021). Incorporating systems thinking approach in a multilevel framework for human-centered crash analysis. *Transportation Research Interdisciplinary Perspectives* 9.

- Bucsuházy, K., Matuchova, E., Zúvala, R., Moravcová, P., Kostíková, M., and Mikulec, R. (2020). Human factors contributing to the road traffic accident occurrence. *Transportation Research Procedia* 45 (pp. 555-561). Rome: Elsevier B.V.
- Dunn, J. (1974). Well-Separated Clusters and Optimal Fuzzy Partitions. *Journal of Cybernetics*, 4:1, 95-104.
- FHWA (Federal Highway Administration), US DOT. (2017). Road Safety Fundamentals: Concepts, Strategies, and Practices that Reduce Fatalities and Injuries on the Road. In *Unit 2: Human Behavior and Road Safety* (pp. 4-5).
- Gower, J. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, 27(4), 857–871.
- Kamei, H. (2019). The Project for Construction of Sindhuli Road. Retrieved from JSCE The International Infrastructure Archives: <http://www.jsce.or.jp/e/archive/project/pj13.html>.
- Kassambara, A. (2017). *Machine Learning Essentials_Practical Guide in R*. STHDA.
- Marcello, D. (2022). *StatMatch: Statistical Matching or Data, R package version 1.4.1*.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2022). *cluster: Cluster Analysis Basics and Extensions. R package version 2.1.4*.
- Oppenheim, I., and Shinar, D. (2012). A context-sensitive model of driving behaviour and its implications for in-vehicle safety systems. *Cogn Tech Work* 14, 261–281.
- Petridou, E., and Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *European Journal of Epidemiology* 16, 819-826.
- Plankermann, K. (2013). *Human Factors as Causes for Road Traffic Accidents in the Sultanate of Oman under Consideration of Road Construction Designs* [Doctoral Thesis, University of Regensburg]. Burglengenfeld, Germany.
- Treat, J., Tumbas, N., McDonald, S., Mayer, R., Stansifer, R., and Castellan, N. (1979). *Tri-level study of the causes of traffic accidents: Executive summary*. Washington D.C.: US Department of Transportation, National Highway Traffic Safety Administration.
- WHO (World Health Organization). (2018). *Global Status Report on Road Safety 2018*. Geneva: World Health Organization.
- Wiper, O., Vuik, S., Cheatley, J., and Cecchini, M. (2022). Cluster analysis to assess the transferability of public health interventions, DELSA/HEA/WD/HWP (2022)1. Organization for Economic Co-operation and Development.

Investigating the Influence of Weather on Weigh-in-Motion Measurements Using In-Pavement Strain Sensors

Xinyi Yang¹; Xingyu Wang, Ph.D.²; Ying Huang, Ph.D.³; Pan Lu, Ph.D.⁴; Yihao Ren⁵; and Joseph Podolsky, Ph.D.⁶

¹Dept. of Civil and Environmental Engineering, North Dakota State Univ., Fargo, ND.

Email: xinyi.yang@ndsu.edu

²Dept. of Civil and Environmental Engineering, North Dakota State Univ., Fargo, ND.

Email: xingyu.wang@ndsu.edu

³Dept. of Civil and Environmental Engineering, North Dakota State Univ., Fargo, ND.

Email: ying.huang@ndsu.edu

⁴Dept. of Transportation, Logistics, and Finance, North Dakota State Univ., Fargo, ND.

Email: pan.lu@ndsu.edu

⁵Dept. of Transportation, Logistics, and Finance, North Dakota State Univ., Fargo, ND.

Email: yihao.ren@ndsu.edu

⁶Office of Materials and Road Research, Minnesota Dept. of Transportation, Maplewood, MN.

Email: joseph.podolsky@state.mn.us

ABSTRACT

Accurate vehicle weight monitoring is essential for efficient and effective traffic management and road maintenance, particularly in preventing potential damage from overweight vehicles to road infrastructure. Weigh-in-motion technology plays a pivotal role in traffic engineering due to its ability to rapidly collect data with minimal traffic disruption. However, the accuracy of weigh-in-motion stations need enhancement and may be affected by weather conditions, especially fluctuations in temperature. To address this concern, a field test was conducted, incorporating controlled vehicle weight, speed, and driving behavior. Additionally, electrical resistance strain transducers were configured for data collection, strategically positioned at the location of maximum bending, the bottom of the asphalt layer, 5 in. beneath the asphalt road surface. This field test, spanning 10 experiments over two years, was conducted to analyze the impact of weather factors on measured vehicle weights by strain sensor. The findings of this investigation will provide insights into the benefits of enhancing the accuracy and reliability of vehicle weight measurements, contributing to more effective traffic management and improved road maintenance practices.

INTRODUCTION

As the number of registered vehicles in the United States continues to rise, reaching approximately 282 million in 2021(*US Number of Registered Vehicles / Economic Indicators / CEIC*, n.d.), and with nearly 166 million private and commercial trucks on the road (*Statista Research Department*, n.d.), the importance of maintaining and ensuring the life-expectancy of pavements has become increasingly crucial. Accurate vehicle weight monitoring is a critical aspect of modern traffic management and road maintenance. Ensuring the adherence to weight limits is not only essential for preserving road infrastructure but is also instrumental in optimizing traffic flow and minimizing safety risks associated with overweight vehicles. Weigh-

in-motion (WIM) technology has emerged as a key tool in traffic engineering, offering the advantage of efficient data collection with no disruption to traffic. WIM stations are designed to capture real-time data on the weights of passing vehicles, providing valuable insights for traffic engineers and policymakers. The data collected aids in enforcing weight restrictions, optimizing infrastructure design, and identifying potential hazards on the road.

However, concerns persist regarding the precision of WIM stations, as variations in weather especially temperature can potentially compromise the accuracy of weight measurements. Various studies have investigated the impact of temperature variations on pavement stress-strain and strain distribution under dynamic loading (Ahmed et al., 2015; Cheng et al., 2021). Specifically, these studies utilized sensors embedded within an asphalt sample, subjecting it to dynamic loading in a lab to simulate a vehicle driving above it and to study the effectiveness of the WIM sensors. Nevertheless, the performance was within a laboratory setting, not replicating field scenarios that involve multiple layers and complex environmental conditions. Thompson et al. (2019) utilized field data from diverse 5-axle tractor and semi-trailer configurations passing through the WIM station. They calculated the average front wheel weight per month over a 21-month period to evaluate the impact of temperature on these vehicles' measurements. While this research encompasses a diverse range of vehicles and contributes valuable insights into WIM sensor performance and stability over time and temperature variations, it's essential to note that the inherent variability in the vehicles, including differences in driving behavior (Saifizul et al., 2013), speed (Gajda et al., 2007), and wheel location (X. Yang et al., 2023), introduces potential biases into the results. In order to thoroughly investigate the impact of weather conditions, especially temperature, on weight in motion, it becomes essential to establish controlled testing conditions that consider these influencing factors.

Hence, this study conducted a comprehensive field test on pavement road, controlling both vehicle speed and weight. Additionally, drivers were instructed to position the wheels directly above the wheel path. To address potential temperature- contributed influences on sensors (H. Yang et al., 2023), the research employed electrical resistance strain transducers configured as a full bridge. This configuration ensures uniform temperature sensitivity across sensors, a critical element for preserving the accuracy of weight measurements. The strain transducers were strategically placed five inches beneath the asphalt road surface, enabling direct monitoring of the dynamic loads imposed by passing vehicles.

The field test spanned ten experiments within two years and included forty-six passes of truck vehicles with identical weights, each traveling at similar low speeds. The collected data was subjected to thorough statistical analysis to unravel the intricate relationship between weather fluctuations and weight measurement precision. By exploring the statistical characteristics of the dataset, this study aims to shed light on the impact of temperature variations on the reliability of WIM stations. The significance of this investigation lies in its potential to enhance the accuracy and reliability of vehicle weight measurements, thereby contributing to effective traffic management and enhanced road maintenance practices.

FIELD TEST & DATA COLLECTION

The experimental dataset was collected through field tests conducted at MnROAD, an extensive pavement test track managed by the Minnesota Department of Transportation. MnROAD includes 3 miles of the primary mainline (I-94 interstate), 3.5 miles of the original westbound I-94, and a 2.5-mile low-volume road. Test sections are distributed across both the

mainline and the original westbound I-94, featuring over 80 distinct segments dedicated to various research investigations.

The experiments were carried out at the segments depicted in Figure 1, with an electrical resistance strain sensor strategically placed beneath the asphalt pavement, positioned 5 inches deep. The sensors operate in a full bridge configuration, ensuring that any temperature variations uniformly affect all sensors. This design maintains a constant resistance ratio, preventing changes in voltage output due to temperature fluctuations. A temperature sensor adjacent to the strain sensor records the mid-pavement temperature. Moreover, to investigate potential impacts of weather conditions on the outcomes, weather stations positioned on the north-west side of the test road section were utilized to gather essential weather data. A detailed summary of all collected weather factors, including descriptions and units, is presented in Table 1.

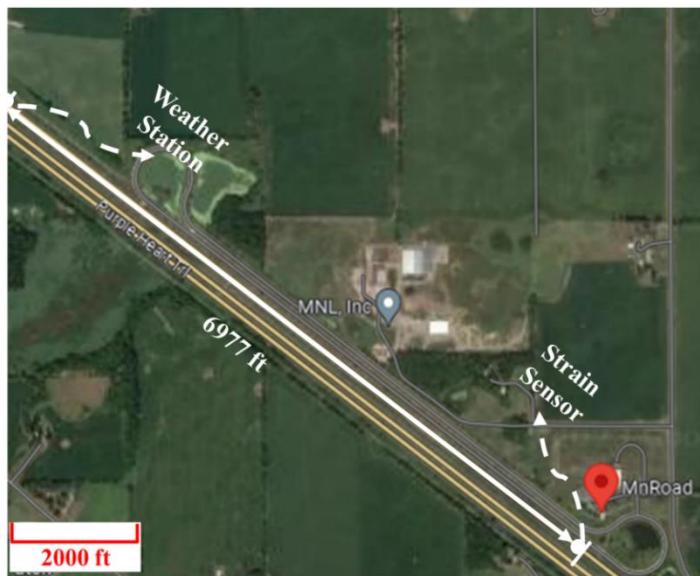


Figure 1. Satellite view of experimental area and weather station

Table 1. Weather factors collected by weather station.

Description	Unit
Air Temperature	Celsius
Atmospheric Pressure	Millibars
Heated Rain Gauge Precipitation	inches
Relative Humidity Percent	%
Net Radiation	Watts/Square Meter
Wind Direction from North	Degrees
Wind Average Speed	Meter/Second
Gust Maximum Speed	Meter/Second
Gust Direction from North	Degrees

To assess the impact of temperature on vehicle weight, this study employed a controlled-weight vehicle (5-axle semi-truck with a gross weight of approximately 8,000 lbs.) to traverse over the strain sensor at various temperatures. Consequently, 10 experiments were conducted

over two years, spanning temperatures from 7°C to 36°C. Considering the potential effects of sensor position and vehicle speed on weight measurements, the study focused on low-speed driving, maintaining an average speed of approximately 4 mph directly above the wheel path. Throughout the 10 experiments, the truck passed over the sensor a total of 46 times. Figure 2 illustrates the speed for each time.

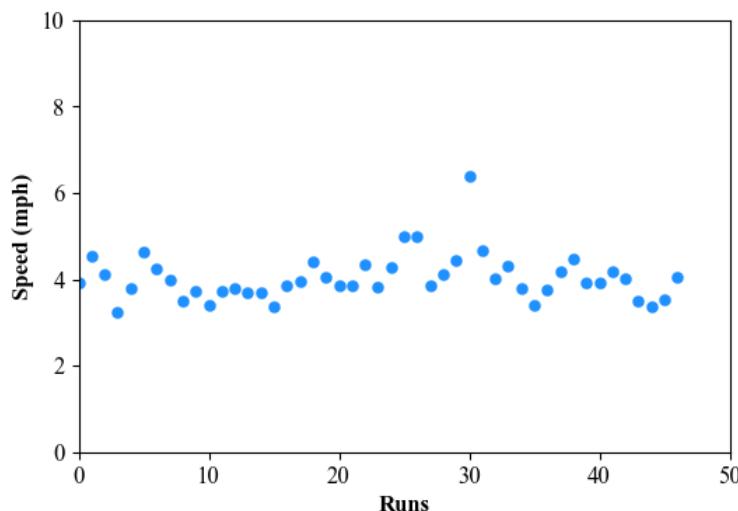


Figure 2. Speed distribution of controlled-weight vehicle runs over the strain sensor

DATA ANALYSIS APPROACH

The primary objective of this study is to evaluate the correlation between weather conditions, especially temperature, and strain data collected by the strain sensor, using real-world data. To accomplish this, a normality test will be conducted on both the strain data and the weather factor data. The outcome of these tests will guide the selection of the most suitable method for evaluating the relationship between the two sets of variables.

Two widely used methods for exploring relationships between variables are Pearson's and Spearman's correlation coefficients (Cho & Won, 2003; Xiao et al., 2016). Pearson's correlation coefficient requires normally distributed continuous variables and employs a parametric test, while Spearman's correlation coefficient is a non-parametric measure that does not assume normal distribution (Akoglu, 2018). Pearson's correlation coefficient is effective when dealing with linear relationships between variables. However, its reliance on normality assumptions can limit its applicability, especially when dealing with non-normally distributed data or ordinal variables. In contrast, Spearman's correlation coefficient is a robust alternative that assesses monotonic relationships, making it suitable for scenarios where linearity cannot be assumed. For two variables, X_1 and X_2 , each with n samples, the Pearson correlation coefficient (r_p) and Spearman's correlation coefficient (r_s) can be determined by:

$$r_p = \frac{\sum_{j=1}^n (X_{1,j} - \bar{X}_1)(X_{2,j} - \bar{X}_2)}{\sqrt{(\sum_{j=1}^n (X_{1,j} - \bar{X}_1)^2)(\sum_{i=1}^n (X_{2,i} - \bar{X}_2)^2)}}$$

where, $X_{1,j}$ and $X_{2,j}$ are the variable samples, and $j = 1, 2, 3, \dots n$. \bar{X}_1 and \bar{X}_2 are means of the values in X_1 and X_2 variable.

$$r_s = 1 - \frac{6 \sum_{j=1}^n d_j^2}{n(n^2 - 1)}$$

where, d_j^2 is the difference between the j^{th} pair of the ranked variables ($X_{1,j}$ and $X_{2,j}$). r_p and r_s both range in value from +1 to -1, where a value of +1 indicates a perfect positive correlation and a value of -1 indicates a perfect negative correlation between the two variables. The range 0.5 to 1 (or -1 to -0.5) indicates a strong correlation. The range 0.3 to 0.5 (or -0.5 to -0.3) indicates a moderate correlation, while 0.1 to 0.3 (or -0.3 to -0.1) suggests a weak relationship. A correlation close to zero implies either no relationship or a very weak one.

To determine the appropriateness of using either Pearson's correlation coefficient or Spearman's correlation coefficient, a normality test is imperative. In this investigation, the Shapiro-Wilk test, renowned for its widespread application (Alizadeh Noughabi, 2016), was selected for its demonstrated superior power in detecting non-normality compared to other normality tests (Mishra et al., 2019). For a variable O with ordered random samples, $O_1, O_2, O_3, \dots, O_l$ (l is the total number of the random samples), where $O_1 \leq O_2 \leq O_3 \leq \dots \leq O_l$, \bar{O} is the sample mean, Shapiro and Wilk (Shapiro & Wilk, 1965) proposed the W test based on the statistic is defined as follows:

$$W = \frac{(\sum_{l=1}^m a_l O_l)^2}{\sum_{l=1}^m (O_l - \bar{O})^2}$$

where,

$$a_l = (a_1, \dots, a_l) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^2}$$

where $m = (m_1, \dots, m_l)^T$ denote the expected values of the order statistics derived from samples of independent and identically distributed random variables, selected from the standard normal distribution. The covariance matrix, V , delineates the interdependence among these order statistics.(Razali & Wah, 2011).

RESULTS AND DISCUSSIONS

This section encompasses a comprehensive analysis to explore the intricate connections between temperature in the middle of the pavement layer, weather factors, and strain sensors. It includes scatterplots illustrating relationships among all variables, the Shapiro-Wilk Test for normality assessment on all variables to identify the most suitable method for evaluating the relationships, and the method to identify the relationship between the factors and strain (utilized as a metric for determining vehicle weight). The section further delves into showcasing the average and variance of the strains of the first wheel of the vehicle, emphasizing the most impactful factors among the features that influence strain dynamics.

Figure 3 presents an extensive grid of scatterplots, illustrating relationships among weather factors, middle pavement layer temperature, and the strain sensor. Diagonal histograms provide

insights into individual variable distributions, while scatterplots visually represent correlations, with a straight line indicating a linear relationship. An upward slope suggests a positive correlation, where both variables tend to increase together, while a downward slope indicates a negative correlation. It's important to note that the absence of a straight-line pattern doesn't exclude a relationship, which could be nonlinear or more complex. Examination of the figure reveals positive linear relationship between middle pavement layer temperature and air temperature with the strain.

After a preliminary exploration of the relationship between the factors and the strain collected by the sensor, the Shapiro-Wilk Test was conducted to assess the normality of the features. This step aimed to identify the most suitable correlation coefficient for analyzing the relationship between the collected strain and all factors. The p-values obtained from the Shapiro-Wilk Test for all features are presented in Table 2. The results indicate that the strain and all features do not follow a normal distribution, given that the p-values are less than 0.05. Consequently, Spearman's correlation coefficient is considered the most appropriate for the subsequent analysis in this study.

Table 2. Results of Shapiro-Wilk test for normality of middle pavement layer temperature, sensor and weather factors.

Features	p-Value
Strain	1.19E-03
Middle Pavement Layer Temperature	1.04E-05
Air Temperature	1.08E-02
Atmospheric Pressure	1.68E-02
Heated Rain Gauge Precipitation	3.61E-15
Relative Humidity Percent	1.47E-04
Net Radiation	7.20E-05
Wind Direction from North	1.94E-02
Wind Average Speed	2.00E-07
Gust Maximum Speed	1.90E-05
Gust Direction from North	5.61E-05

Then, Spearman's correlation coefficients are evaluated for all variables, the results are visually presented in Figure 4 through a heatmap, elucidating the interrelationships between different factors. Strong correlations are evident, particularly between middle pavement layer temperature and the strain, as well as air temperature and the strain, exhibiting Spearman's correlation coefficients of 0.65 and 0.63, respectively. Furthermore, moderate correlation coefficients are observed between the strain with wind direction, gust maximum speed, and gust direction, with Spearman's correlation coefficients of -0.47, -0.32, and -0.46, respectively.

Given that middle pavement layer temperature stands out as a significant factor, showing the highest Spearman's correlation coefficient among all features and a strong correlation with strain, a more in-depth analysis is conducted to explore the relationship between the average middle pavement layer temperature and strain. Figures 5 and 6 intricately illustrate the relationship between mean pavement temperature and strain dynamics across a series of ten tests. Each test, categorized by increasing mean pavement temperatures, provides insights into the relationship between rising temperature and the corresponding collected strain values.

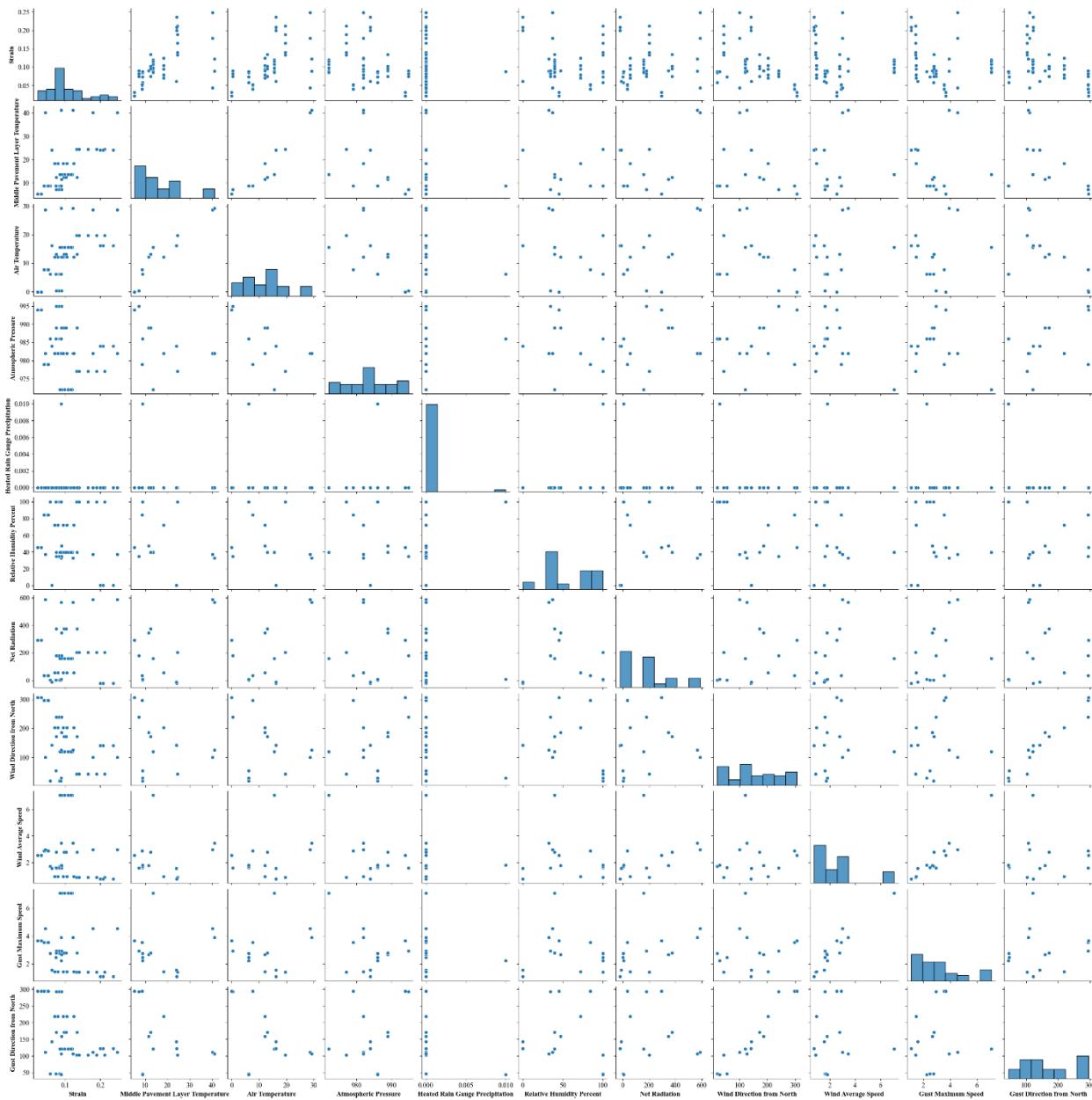


Figure 3. Exploration of correlations: scatterplots among middle pavement layer temperature, weather factors, and strain sensor data.

Figure 5 depicts the mean strain change with varying mean pavement temperatures, showcasing a clear upward trend that implies a positive correlation between higher pavement temperatures and increased mean strain. This observation suggests that as pavement temperature rises, strain tends to increase correspondingly. Figure 6 expands this exploration by introducing the variance in strain change. Simultaneously, the line plot depicting variance offers insights into the distribution and dispersion of strain data. Higher variances at elevated temperatures, particularly in tests 8, 9, and 10, suggest a broader range of responses, while at lower temperatures, variances exhibit less disparity.

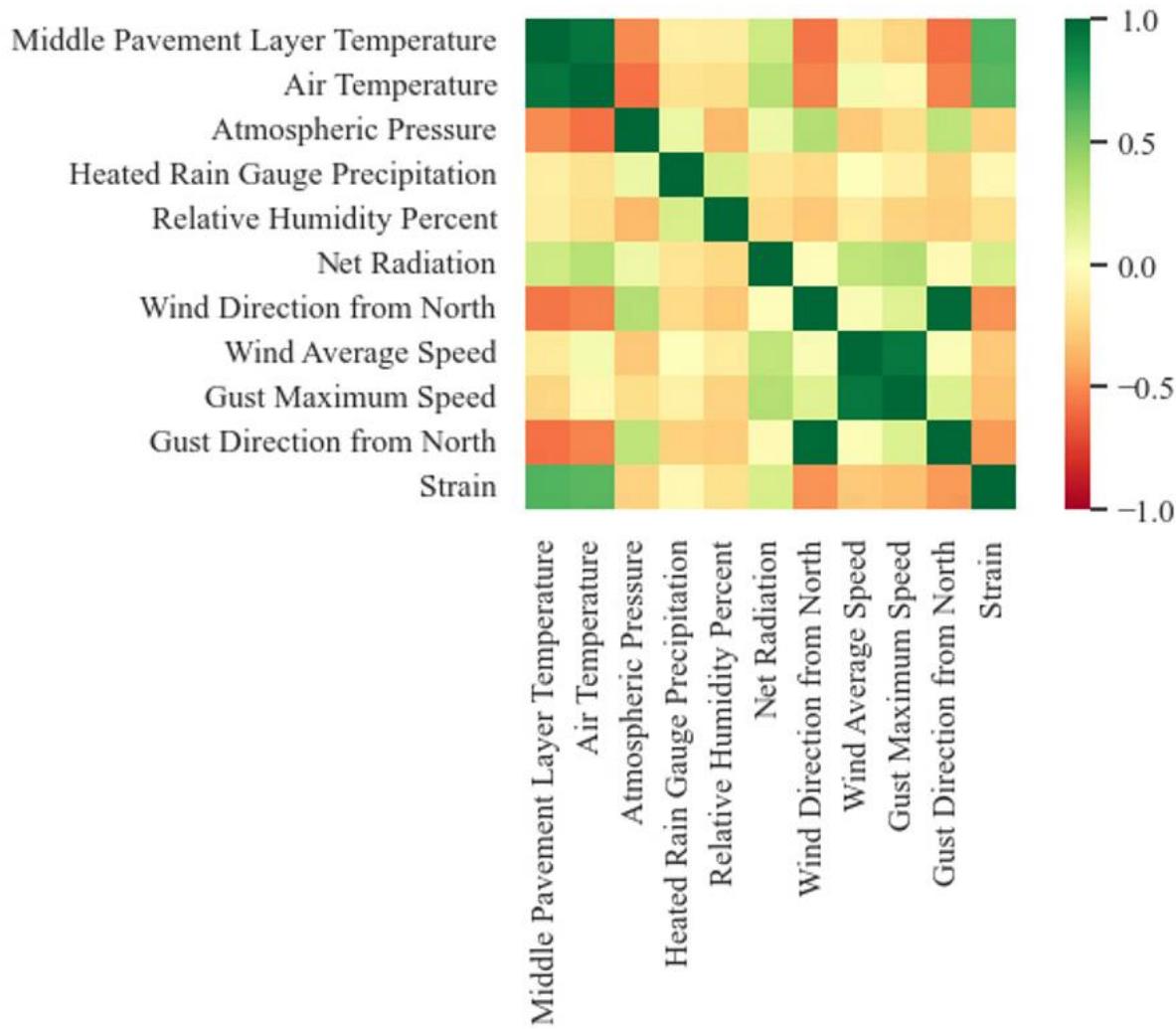


Figure 4. Correlation heatmap: relationships between variables

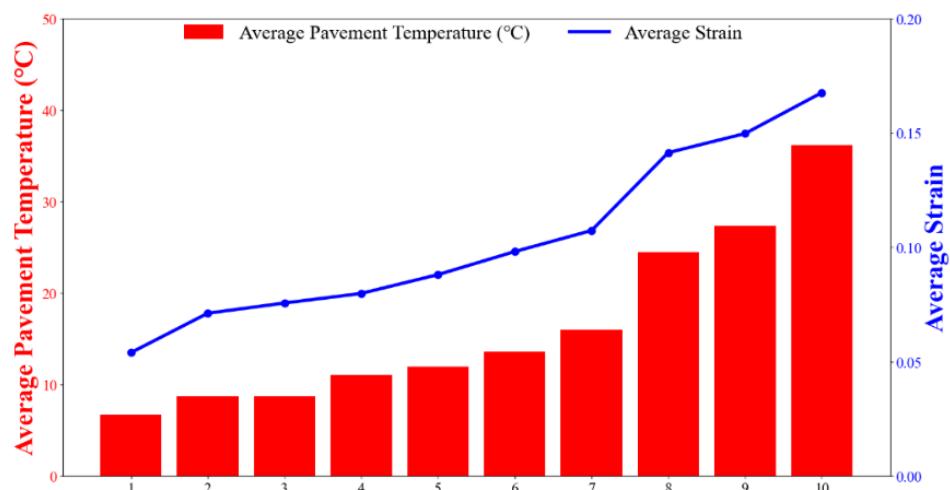


Figure 5. Average strain dynamics in response to varying pavement temperatures

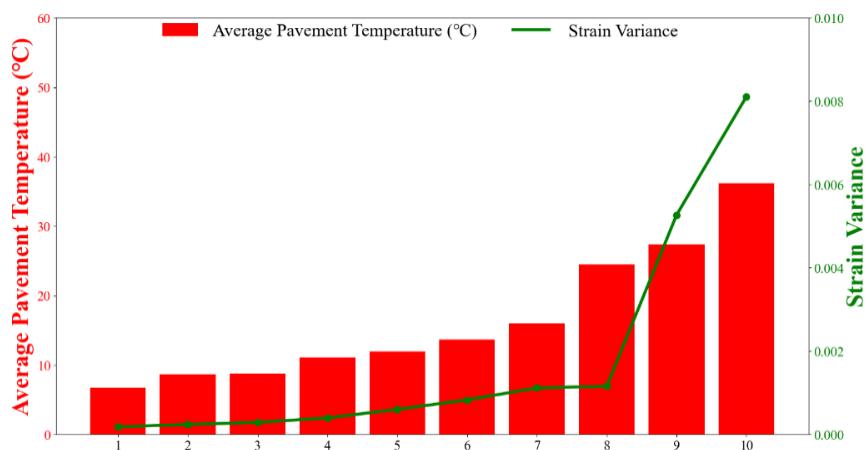


Figure 6. Strain variance dynamics in response to varying pavement temperatures

CONCLUSION

In conclusion, this study utilized electrical resistance strain transducers installed beneath asphalt, complemented by temperature sensors and weather data from a weather station. It aimed to uncover relationships between middle pavement layer temperature, weather factors, and strain (a metric for vehicle weight). Through scatterplots, normality tests, and correlation coefficient analysis, the research revealed robust connections among middle pavement layer temperature, weather factors, and strain sensors. Notably, mid-pavement temperature and air temperature exhibited strong correlations with strain, while wind direction, gust maximum speed, and gust direction showed moderate correlations. The subsequent examination of average strain and variance highlighted the intensified and nonlinear nature of the relationship between pavement temperature and strain under elevated conditions. Results indicated that average strain and strain variance undergo subtle changes at temperatures below 24.47 degrees but become more pronounced as temperatures exceed this threshold.

These findings highlight the impact of temperature variations on weigh-in-motion accuracy, offering valuable insights for optimizing WIM technology. In the future, these findings will be synthesized to establish parameters for assessing WIM, aiding in mitigating the impact of weather conditions on its accuracy.

ACKNOWLEDGEMENT

The authors express their gratitude to the funding provided to support this study from the U.S. Department of Transportation through Mountain-Plains Consortium (MPC), a university transportation center. The authors thank the support from the Minnesota Road Research Facility (MnROAD) to install sensors and collect data, especially thanks to the staff Benjamin Worel and Michael Vrtis.

REFERENCES

- Ahmed, M., Rahman, A., Islam, M., and Tarefder, R. (2015). Combined effect of asphalt concrete cross-anisotropy and temperature variation on pavement stress-strain under dynamic loading. *Construction and Building Materials*, 93, 685–694.

- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93.
- Alizadeh Noughabi, H. (2016). Two powerful tests for normality. *Annals of Data Science*, 3(2), 225–234.
- Cheng, H., Liu, J., Sun, L., and Liu, L. (2021). Critical position of fatigue damage within asphalt pavement considering temperature and strain distribution. *International Journal of Pavement Engineering*, 22(14), 1773–1784.
- Cho, S.-B., and Won, H.-H. (2003). *Machine learning in DNA microarray analysis for cancer classification*. 189–198.
- Gajda, J., Sroka, R., and Żegleń, T. (2007). Accuracy analysis of WIM systems calibrated using pre-weighed vehicles method. *Metrology and Measurement Systems*, 14(4), 517–527.
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., and Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67.
- Razali, N. M., and Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.
- Saifizul, A. A., Karim, M. R., Yamanaka, H., and Okushima, M. (2013). Empirical analysis on the effect of gross vehicle weight and vehicle size on speed in car following situation. *Asian Transport Studies*, 2(4), 351–362.
- Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- Statista Research Department. (n.d.). Number of motor vehicles in U.S. Statista. Retrieved May 24, 2023, from <https://www.statista.com/statistics/183505/number-of-vehicles-in-the-united-states-since-1990/>.
- Thompson, G., Kroll, K., Grangroth, C., and Arnold, J. (2019). WIM Sensor Performance and Stability Across Time Periods And Variations in Temperature. *ICWIM8*, 284.
- US Number of Registered Vehicles | Economic Indicators | CEIC. (n.d.). Retrieved August 3, 2023, from <https://www.ceicdata.com/en/indicator/united-states/number-of-registered-vehicles>.
- Xiao, C., Ye, J., Esteves, R. M., and Rong, C. (2016). Using Spearman's correlation coefficients for exploratory data analysis on big dataset. *Concurrency and Computation: Practice and Experience*, 28(14), 3866–3878.
- Yang, H., Yang, Y., Zhao, G., Guo, Y., and Wang, L. (2023). Development and Temperature Correction of Piezoelectric Ceramic Sensor for Traffic Weighing-In-Motion. *Sensors*, 23(9), 4312.
- Yang, X., Wang, X., Podolsky, J., Huang, Y., and Lu, P. (2023). Assessing Vehicle Wandering Effects on the Accuracy of Weigh-in-Motion Measurement Based on In-Pavement Fiber Bragg Sensors through a Hybrid Sensor-Camera System. *Sensors*, 23(21), 8707.

Navigating Road Safety and Equity: A GIS Analysis of Crash Data in Atlanta, GA

Eazaz Sadeghvaziri, Ph.D.¹; Ramina Javid²; and Lila Turbiville³

¹Assistant Professor, Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Macon, GA (corresponding author). Email: sadeghvaziri_e@mercer.edu

²Ph.D. Candidate, Dept. of Transportation and Urban Infrastructure, Morgan State Univ. Email: rajah1@morgan.edu

³Undergraduate Student, Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Pensacola, FL. Email: Lila.Evelyn.Turbiville@live.mercer.edu

ABSTRACT

Motor vehicle crashes, claiming over 100 lives daily in the US, pose a critical public health risk. This study aims to examine relationships between a five-year period of crash data and disadvantaged communities (DACs) in the city of Atlanta to assess the equity aspects of crash safety. The study used regression analysis to evaluate the association between transportation, health, economic, equity, and resilience indicators and crashes. Results indicate a positive relationship between DACs and crashes, with transportation disadvantage emerging as the most influential predictor. The findings emphasize the need for targeted policies addressing underlying factors contributing to elevated crash rates in DACs, such as poverty and limited access to transportation and healthcare. Furthermore, the study underscores the importance of considering community insights in traffic safety planning and informs the development of interventions to enhance road safety and equity.

INTRODUCTION

Motor vehicle crashes are a public health concern both in the United States and abroad. In the United States, motor vehicle crashes are a leading cause of death and kill over 100 people every day. In 2020, almost 41,000 people died in motor vehicle crashes in the United States. That is more than 110 people killed in crashes every day (CDC 2023). This social phenomenon particularly affects vulnerable road users of disadvantaged communities (DACs) and populations in low and middle-income communities (World Health Organization 2020). Previous research shows that underserved communities are more prone to observing traffic crashes (Li et al. 2022). Due to the increased awareness of these issues in recent years, state and federal transportation agencies have been prompted to address inequality and inequity in transportation. Nonetheless, there is a lack of studies looking into the social differences in crash risks, and there is no simple approach for evaluating the disparities in transportation. Therefore, how to effectively model the relationship between disadvantaged communities and their traffic safety conditions needs to be further explored.

The goal of this study is to analyze equity aspects of crash safety in Atlanta's transportation system to improve road safety and equity in Atlanta. Therefore, the overall objective of this study is to investigate the associations between DACs and five years of crashes in the city of Atlanta.

LITERATURE REVIEW

Existing road safety analysis primarily has focused on the exploration of relationships between traffic crashes and road geometry (Mohammed et al. 2023; Al-Aamri et al. 2021; Singh

and Katiyar 2021), and traffic flow (Wu et al. 2023). For instance, Mohammed et al. aimed to investigate Road Traffic Crashes (RTCs) and their causes by identifying crash hot spots and exploring whether they were primarily attributed to driver behavior or the geometrical design of roads and intersections. The results of this study showed that the majority of crashes occurred in the densely populated central-eastern section of the country (Mohammed et al. 2023). Al-Aamri et al. identified high-risk zones for crashes. They used a network analysis that used RTC and GIS data to analyze hotspots. There is statistical evidence that bridges and roundabouts have a higher risk of causing RTCs (Al-Aamri et al. 2021).

From the traffic flow perspective, Melodi Zhan-Moodie aimed to investigate the spatiotemporal traffic crashes and congestion patterns to gain a better understanding of why these traffic crashes may be occurring, and how they can be prevented. It was found that road segments with more complicated variables, including diverging, merging, and weaving lanes, were found to have more accidents. Like results found in the previous study, it was found that areas with a lower speed limit had a higher crash frequency. This study found that accidents are both a cause and a result of congestion, because during heavy traffic conditions, the safety of motorists and efficiency of traffic conditions are both at risk (Zhan-Moodie 2019). Similarly, a study aimed to propose a GIS-based crash risk identification model that uses machine learning techniques. This model recommended policy changes that could be implemented to enhance urban traffic safety in the future. The study found that spots with high crash risk are typically correlated with spots that have dense traffic flow (Wu et al. 2023). The findings of both studies could be used to make suggestions for road planning in urban areas to help aid traffic safety.

Several studies focused on the association between the built environment and crashes. The results of a study that aimed to identify physical characteristics of the environment (that are modifiable) that may increase the risk of pedestrian injury showed that pedestrian crashes were more likely to occur at intersections where a driver may have an abundance of visual stimuli (i.e., too many road signs). This study found that fair road conditions lead to a greater number of crashes – this could be attributed to the fact that drivers are likely to travel faster if a road is in better condition (Tauechel 2009). Ewing and Dumbaugh aimed to summarize and explore the relationship between traffic safety and the built environment. They concluded that the traffic environment in urban areas is safer than that of rural areas, despite congestion. Drivers seem to be more reckless on roads that are less densely populated, leading to more RTCs (Ewing and Dumbaugh 2009). Similarly, Dai and Jaworski aimed to investigate the influence of the built environment on pedestrian crashes. The results of this study found that these crashes were most likely to occur in areas with public transit stops (Dai and Jaworski 2016). Bavar et al. aimed to evaluate the effects of environmental factors on the frequency of crashes. Overall, the study found that the identification of crash environmental influencing factors will be incredibly useful in forming policies that can be used to reduce urban RTCs. For example, bus stop characteristics, traffic light patterns, and better control of land could all be modified to ensure better safety for drivers and passengers (Sedigh Bavar, Naderan, and Saffarzadeh 2023).

Moreover, the relationship between roadway safety and various socioeconomic and demographic factors, such as income, housing, population structure, investments in urban infrastructures, and living environments, at various scales, has been the subject of an increasing number of studies conducted from a social justice perspective (Chimba, Musinguzi, and Kidando 2018; Nantulya and Reich 2003; Patwary et al. 2024). For instance, Nantulya and Reich (2003) examined how traffic-related fatalities and serious injuries were distributed among various counties and different socioeconomic groups. According to this study, vulnerable and low-

income individuals are exposed to a higher risk of traffic injuries and fatalities (Nantulya and Reich 2003). Another study used disadvantaged communities (DACs) developed by USDOT and the results revealed that health, resilience, and transportation-disadvantaged tracts are associated with more fatal crashes (Patwary et al. 2024).

While existing literature examines the crash risks with some selected sociodemographic variables, there is still a lack of studies that systematically assess the spatial relationship between crashes and social vulnerability. Therefore, the goal of this study is to investigate the associations between DACs and crashes in the city of Atlanta.

DATA AND METHODOLOGY

This study focused on the city of Atlanta as a case study. Atlanta is the capital and most populous city in the state of Georgia. It is the seat of Fulton County, although a portion of the city extends into neighboring DeKalb County. With a population of 499,127 living within the city limits, Atlanta is the eighth most populous city in the Southeast, and the 38th most populous city in the United States according to the 2022 U.S. census. The three largest groups in Atlanta are Black or African American (48%), White (41%), and Asian (5%). In 2022, Atlanta had 18.5% of residents living below the poverty level (U.S. Census Bureau 2022).

Data

Two sources of datasets are used in this paper. The first dataset for the city of Atlanta was retrieved from the Georgia Department of Transportation (GDOT) Crash Data Dashboard (GDOT 2023). The Crash Data Dashboard contains crash-related information, including geographic location (latitude-longitude), date and time of the crash, number of fatalities, weather conditions, surface conditions, light conditions, etc. The crash information is retrieved from 2017 to 2022. According to the data, a total of 224,172 crashes occurred in Atlanta during these five years. Table 1 represents a summary of crashes that occurred from 2017 to 2022 in the city of Atlanta.

Table 1. Crash Statistics in Atlanta from 2017 to 2022

Description	Frequency
Total Crash Count	224,172
Fatal Crash Count	396
Number of Injuries	81,995
Number of Fatalities	435

The second dataset is associated with the United States Department of Transportation Disadvantaged Communities (USDOT DACs). The USDOT developed an interactive GIS-based mapping tool for public use, and it was launched as a dashboard to display six disadvantage theme indicators (Transportation, Health, Economy, Equity, Resilience, and Environmental), an overall disadvantage indicator, and underlying data used to create the indicators (USDOT, 2022). Underlying data that went into each theme was compiled at the census tract level from various sources, and was then normalized and averaged to generate overall indicators for each of the themes (Patwary et al. 2024). A brief overview of the six indicators is provided below:

1. Transportation Access Disadvantage
 - Communities and places that spend more and take longer to get to where they need to go (CDC Social Vulnerability Index, Census America Community Survey, EPA Smart Location Map, HUD Location Affordability Index).
 - Areas that:
 - Have a higher percentage of total population with a drive time to work of 30 minutes or longer.
 - Have a higher percentage of total population with no vehicle.
 - Are not supportive of walking based on economic and built-environment characteristics.
 - Have higher transportation costs as a percentage of income.
2. Health Disadvantage
 - Communities that have adverse health outcomes, disabilities, and environmental exposures. (CDC Social Vulnerability Index)
 - Areas that:
 - Have a higher percentage of population over the age of 64.
 - Have a higher percentage of population without health insurance.
 - Have a higher percentage of non-institutionalized population with any disability.
3. Environmental Disadvantage
 - Communities with disproportionately high levels of certain air pollutants and high potential presence of lead-based paint in housing units. (EPA EJ Screen)
 - Areas that:
 - Have a higher percentage of housing units built before 1960 (lead paint indicator).
 - Higher levels of diesel particulate matter in the air.
 - Higher risk for air toxic cancer.
 - Closer to areas with high vehicular traffic congestion.
 - Higher PM2.5 levels in the air.
 - Higher ozone levels in the air.
4. Economic Disadvantage
 - Areas and populations with high poverty, low wealth, lack of local jobs, low homeownership, low educational attainment, and high inequality (CDC Social Vulnerability Index, Census America Community Survey, FEMA Resilience Analysis & Planning Tool).
 - Areas that have:
 - A higher percentage of people age 25 and older with an education that is short of a high school diploma.
 - A higher percentage of renters.
 - Higher unemployment.
 - Higher income inequality.
 - Higher percentage of population at or below the area median income.
 - Higher percentage of people living below the Federal Poverty Level.
 - Higher Housing Costs as a percentage of income for a typical household in the local region.
5. Resilience Disadvantage
 - Communities are vulnerable to hazards caused by climate change. (FEMA National Risk Index)

- Areas that have a higher expected annual loss of life, building value, and agricultural value from 18 climate or natural hazards. These hazards include avalanche, coastal flooding, cold wave, drought, earthquake, hail, heat wave, hurricane, ice storm, landslide, lightning, riverine flooding, strong wind, tornado, tsunami, volcanic activity, wildfire, and winter weather.
6. Equity Disadvantage
- o Communities with a high percentile of persons (age 5+) who speak English "less than well" or are considered linguistically isolated. (CDC Social Vulnerability Index).

Table 2 shows the percentage of disadvantaged communities in each category in the city of Atlanta at the census tract level.

Table 2. Percentage of Disadvantaged Communities at the Census Tract Level.

Indicator	Percentage of census tracts
Transportation Disadvantage	13.53%
Health Disadvantage	19.55%
Economy Disadvantage	25.56%
Equity Disadvantage	6.77%
Resilience Disadvantage	0.75%

Methodology

The crash data was imported into the ArcGIS Pro using the longitudes and latitudes of the crashes. Moreover, the disadvantaged indicators were imported from ArcGIS Online at the census tract level. The "Summary Within" option in ArcGIS Pro was used to calculate the number of crashes at each census tract in the city of Atlanta. The crash-based dataset was then merged with the disadvantaged indicators dataset, but due to some crashes not being present in the second dataset, 0.4% of crash data was lost. The final cleaned dataset includes 223,262 crashes. ArcMap was used for spatial visualization of data.

A regression model was developed to study the relationship between DACs from Atlanta's census tracts and five years of crashes. Regression analysis is used for two purposes: to provide a simple outline to examine the relationship among a group of variables, and to predict the dependent variable and future outcome. The model can generally be expressed as Equation 1, where n represents the sample size, y denotes the dependent variable, X is the explanatory variable, β is the unknown regression coefficients, and ε is the error term.

$$y = X\beta + \varepsilon \quad (1)$$

$$r_i = y_i - \hat{y}_i \quad i=1, 2, \dots, n \quad (2)$$

$$\hat{\beta}_{OLS} = \arg \min_{\beta} \sum_{i=1}^n r_i^2(\beta) \quad (3)$$

Based on the estimated coefficients $\hat{\beta}$, the dependent variable can be estimated as \hat{y} . The residual r_i is then calculated for each observation, based on Equation 2. A typical regression analysis relies on the Ordinary Least Squares (OLS) method, which computes the model's

coefficients to minimize the sum of squared residuals, as indicated in Equation 3. The regression analysis provides a simple outline to examine the relationship among a group of variables. A regression model was conducted to find the best satisfactory result using the number of crashes as the dependent variable and Transportation Access Disadvantage, Health Disadvantage, Environmental Disadvantage, Economic Disadvantage, Resilience disadvantage, and Equity disadvantage as independent variables.

ANALYSIS RESULTS

A heatmap was generated using ArcGIS Pro and the longitude and latitude data of crashes retrieved from GDOT between 2017 and 2022 to illustrate areas with a higher concentration or density of crashes, utilizing a color gradient. This heatmap, as demonstrated in Figure 1, serves the purpose of identifying patterns and trends. Locations with a more intense color on the heatmap indicate a higher concentration of crashes. Figure 1 shows that the density of crashes over this five-year period was significantly higher in downtown Atlanta. Figure 2 shows the density of fatal crashes over this five-year period.

Figure 3 shows the disadvantaged communities at the census tract level in Atlanta. Census tracts with darker shades of red have higher scores on disadvantaged indicators. As shown, the south side, west side, and downtown of the city have the most disadvantaged Census Tracts.

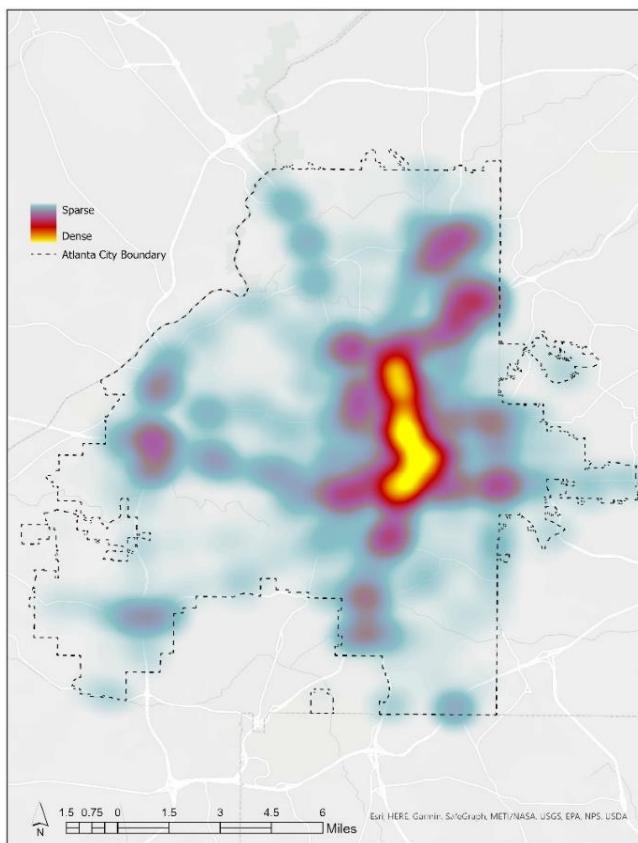


Figure 1. Heatmap of Density of Crashes in Atlanta

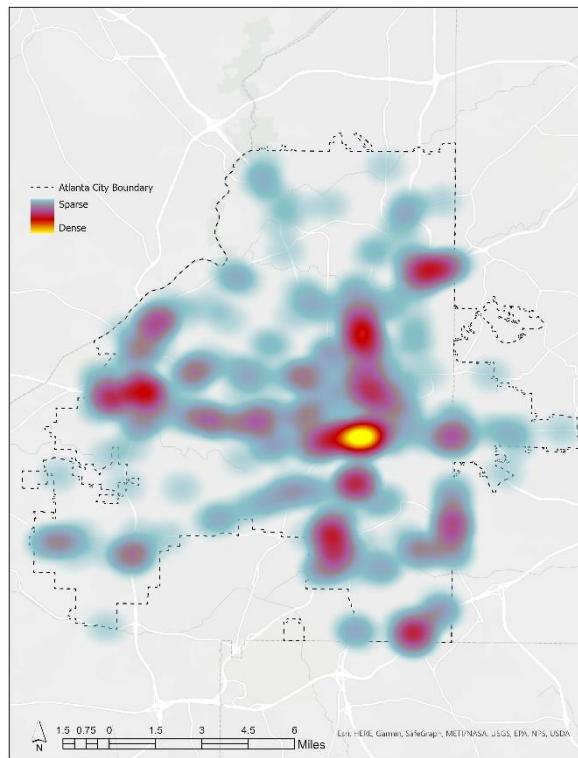


Figure 2. Heatmap of Density of Fatal Crashes in Atlanta

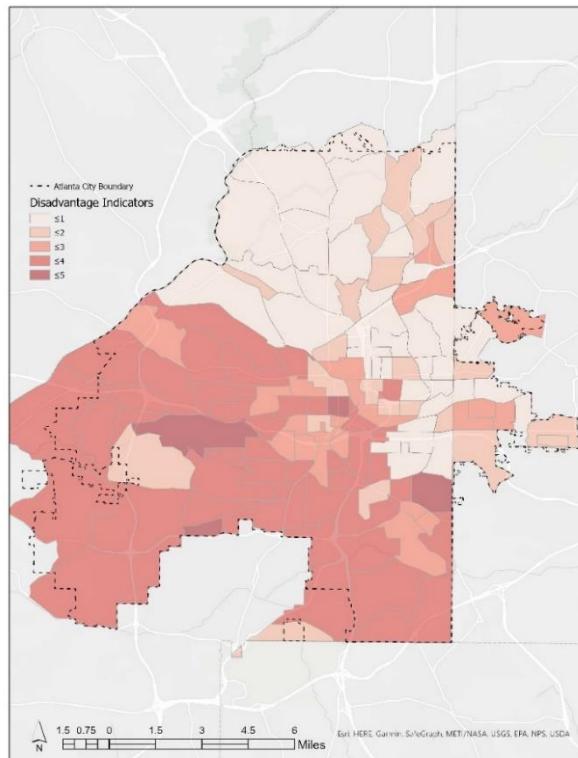


Figure 3. Disadvantaged Communities in Atlanta

Many studies have used statistical models to develop policies to improve traffic safety, investigate and forecast travel behavior, and pinpoint deficiencies in equity and transportation policies (Javid, Sadeghvaziri, and Jeihani 2023a; Sadeghvaziri, Javid, and Jeihani 2023; Jeihani, Javid, and Sadeghvaziri 2022; Javid, Sadeghvaziri, and Jeihani 2023b). Therefore, this study developed a linear regression model to investigate the relationship between the number of crashes and disadvantaged community indicators. Table 3 presents the results of the final models.

Table 3. Regression Model Results

	Estimate	Std. Error	z value	Pr(> z)	Signif. Codes
(Intercept)	-0.703	0.177	-3.974	0.000	***
Transportation Disadvantage	0.032	0.004	9.108	0.000	***
Health Disadvantage	0.021	0.003	8.066	0.000	***
Economy Disadvantage	0.012	0.003	4.080	0.000	***
Equity Disadvantage	0.011	0.002	5.595	0.000	***
Resilience Disadvantage	-0.002	0.003	-0.885	0.378	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
 Residual standard error: 0.5105 on 127 degrees of freedom
 Multiple R-squared: 0.8505, Adjusted R-squared: 0.8447
 F-statistic: 144.6 on 5 and 127 DF, p-value: < 2.2e-16

Table 3 shows the regression statistics for the regression model. The R-squared and adjusted R-squared (0.84) are both high, and the model explains 85% of the fitted data in the regression model, which shows acceptable goodness of fit for the model.

DISCUSSION AND SUMMARY

This study investigated the relationship between disadvantaged communities (DACs) and the number of crashes in Atlanta using crash data from 2017 to 2022 and disadvantaged indicators at the census tract level using a Linear Regression model. The findings suggest that there is a positive relationship between DACs and the number of crashes. Census tracts with higher scores on disadvantaged indicators tend to have a higher number of crashes. The regression model results show that all four of the disadvantaged indicators used in the study were statistically significant predictors of the number of crashes. This suggests that all of these factors are important in contributing to the higher crash rates in DACs. The Transportation Disadvantage indicator was the strongest predictor of the number of crashes, followed by the Health Disadvantage indicator, the Economy Disadvantage indicator, and the Equity Disadvantage indicator. The Resilience Disadvantage indicator was not a significant predictor of the number of crashes.

The heatmap generated in this study also shows that the density of crashes is higher in downtown Atlanta. This is consistent with the finding that DACs are often located in urban areas. The findings of this study have several implications for policy and practice. First, the findings underscore the need for policies that address the underlying factors that contribute to the higher crash rates in DACs. These factors include poverty, lack of access to transportation, and lack of access to healthcare. Second, the findings suggest that the location of transportation infrastructure and services should be considered when developing transportation plans. Third, the findings highlight the importance of community engagement in traffic safety planning.

Community members can provide valuable insights into the factors that contribute to crashes in their neighborhoods and can help develop and implement effective traffic safety interventions.

This study is a valuable contribution to the understanding of the relationship between DACs and the number of crashes. The findings of this study can be used to inform the development of policies and interventions that can help improve traffic safety in DACs.

REFERENCES

- Al-Aamri, A. K., G. Hornby, L.-C. Zhang, A. A. Al-Maniri, and S. S. Padmadas. 2021. "Mapping Road Traffic Crash Hotspots Using GIS-Based Methods: A Case Study of Muscat Governorate in the Sultanate of Oman." *Spatial Statistics, Towards Spatial Data Science*, 42 (April): 100458.
- CDC. 2023. "Transportation Safety | Transportation Safety | Injury Center | CDC." July 24, 2023. <https://www.cdc.gov/transportationsafety/index.html>.
- Chimba, D., A. Musinguzi, and E. Kidando. 2018. "Associating Pedestrian Crashes with Demographic and Socioeconomic Factors." *Case Studies on Transport Policy* 6 (1): 11–16.
- Dai, D., and D. Jaworski. 2016. "Influence of Built Environment on Pedestrian Crashes: A Network-Based GIS Analysis." *Applied Geography* 73 (August): 53–61.
- Ewing, R., and E. Dumbaugh. 2009. "The Built Environment and Traffic Safety: A Review of Empirical Evidence." *Journal of Planning Literature* 23 (4): 347–67.
- GDOT. 2023. "Crash Data & Reporting: Accident Reporting System - GDOT." 2023. <https://www.dot.ga.gov/GDOT/Pages/CrashReporting.aspx>.
- Javid, R., E. Sadeghvaziri, and M. Jeihani. 2023a. "A Driving Simulator Study to Understand the Impact of Cell Phone Blocking Apps on Distraction." *Highlights of Vehicles* 1 (1): 17–28.
- Javid, R., E. Sadeghvaziri, and M. Jeihani. 2023b. "Development and Evaluation of a Bayesian Network Model for Preventing Distracted Driving." *IATSS Research* 47 (4): 491–98.
- Jeihani, M., R. Javid, and E. Sadeghvaziri. 2022. "Educating the Public About Distracted Driving and Evaluating Distraction-Prevention Technologies."
- Li, X., S. Yu, X. Huang, B. Dadashova, W. Cui, and Z. Zhang. 2022. "Do Underserved and Socially Vulnerable Communities Observe More Crashes? A Spatial Examination of Social Vulnerability and Crash Risks in Texas." *Accident Analysis & Prevention* 173: 106721.
- Mohammed, S., A. H. Alkhhereibi, A. Abulibdeh, R. N. Jawarneh, and P. Balakrishnan. 2023. "GIS-Based Spatiotemporal Analysis for Road Traffic Crashes; in Support of Sustainable Transportation Planning." *Transportation Research Interdisciplinary Perspectives* 20 (July): 100836.
- Nantulya, V. M., and M. R. Reich. 2003. "Equity Dimensions of Road Traffic Injuries in Low- and Middle-Income Countries." *Injury Control and Safety Promotion* 10 (1–2): 13–20.
- Patwary, A. L., A. M. Haque, I. Mahdinia, and A. J. Khattak. 2024. "Investigating Transportation Safety in Disadvantaged Communities by Integrating Crash and Environmental Justice Data." *Accident Analysis & Prevention* 194: 107366.
- Sadeghvaziri, E., R. Javid, and M. Jeihani. 2023. "Active Transportation for Underrepresented Populations in the United States: A Systematic Review of Literature." *Transportation Research Record*, 03611981231197659.
- Sedigh Bavar, M., A. Naderan, and M. Saffarzadeh. 2023. "Evaluating the Spatial Effects of Environmental Influencing Factors on the Frequency of Urban Crashes Using the Spatial Bayes Method Based on Euclidean Distance and Contiguity." *Transportation Engineering* 12 (June): 100181.

- Singh, N., and S. K. Katiyar. 2021. "Application of Geographical Information System (GIS) in Reducing Accident Blackspots and in Planning of a Safer Urban Road Network: A Review." *Ecological Informatics* 66: 101436.
- Taquechel, E. P. 2009. "A Spatial Analysis of the Relationship between Pedestrian Crash Events and Features of the Built Environment in Downtown Atlanta." https://scholarworks.gsu.edu/iph_theses/117/.
- US Census Bureau. 2022. "U.S. Census Bureau QuickFacts: Atlanta City, Georgia; Georgia." 2022. <https://www.census.gov/quickfacts/fact/table/atlantacitygeorgia,GA/PST045222>.
- USDOT. 2022. "USDOT Transportation Disadvantaged Communities April 2022 - Overview." 2022. <https://msucivileng.maps.arcgis.com/home/item.html?id=44a94401adb3489cb19b08539bcb11ed>.
- World Health Organization. 2020. "Road Traffic Injuries." 2020. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- Wu, P., T. Chen, Y. D. Wong, X. Meng, X. Wang, and W. Liu. 2023. "Exploring Key Spatio-Temporal Features of Crash Risk Hot Spots on Urban Road Network: A Machine Learning Approach." *Transportation Research Part A: Policy and Practice* 173 (July): 103717.
- Zhan-Moodie, M. 2019. "Spatio-Temporal Analysis of Highway Congestion and Accidents: A Case Study of Interstate 285 in Georgia." https://digitalcommons.kennesaw.edu/msce_etd/3/.

Comparison of the Effect of In-Crosswalk Traffic Signs on Pedestrian Safety

Mulugeta D. Amare, S.M.ASCE¹; and Daba S. Gedafa, Ph.D., P.E., ENV SP, F.ASCE²

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of North Dakota, Grand Forks, ND.

Email: mulugeta.amare@und.edu

²Dept. of Civil Engineering, Univ. of North Dakota, Grand Forks, ND.

Email: daba.gedafa@und.edu

ABSTRACT

Pedestrian safety is a significant concern due to the severity of pedestrian-involved crashes. Pedestrian fatalities have increased by 77% over the past decade, constituting a 5% increase in pedestrian fatalities per the overall number of traffic-related fatalities. Prior studies have explored various regulatory traffic signs; however, there have been limited studies on in-crosswalk signs. This study compared the effectiveness of in-crosswalk "YIELD TO PEDESTRIANS" and "STOP FOR PEDESTRIANS" traffic signs. Yielding behavior and vehicle speed were used as safety indicators, and impacts were measured during different school sessions and times of day. The traffic signs' impact was evaluated through independent t-tests and two-proportion z-tests at a 95% confidence level. The results demonstrate that both traffic signs led to significant vehicle speed reductions and yield to pedestrian enhancements across all situations; however, the difference between the two traffic signs' impact was insignificant, suggesting they can be used interchangeably.

Keywords: Pedestrians, Safety, In-Crosswalk Traffic Signs, YIELD TO PEDESTRIANS, STOP FOR PEDESTRIANS

INTRODUCTION

The continuous flow of vehicles and people in modern urban settings creates complex interactions, resulting in significant traffic-related pedestrian crashes that profoundly impact individuals, families, and communities; therefore, prioritizing effective strategies to decrease the occurrence and severity of vehicle-pedestrian collisions remains essential in urban environments.

The 2022 Governors Highway Safety Association (GHSA) report states that pedestrian fatalities in the United States (US) are increasing each year (Kara and Elizabeth 2022). There was a projected 11.5% increase in pedestrian fatalities in 2021 compared to 2020. The US has seen the highest number of pedestrian deaths since 1981. Pedestrian fatalities have surged by 77% over the past decade, constituting a 5% increase in pedestrian fatalities per the overall number of traffic-related fatalities. The National Center for Statistics and Analysis (NCSA) reported that 3% of the total reported injuries in the past decade were pedestrian injuries. This percentage decreased to 2% in the last two years (NCSA 2023).

Traffic safety is a shared responsibility, and different sectors have been implementing strategies to address the issue. The US Department of Transportation employed a National Roadway Safety Strategy and selected five safe system approaches to traffic safety: safer people, safer roads, safer speed, safer vehicles, and post-crash care (USDOT 2022). These approaches could be achieved by implementing different strategies, such as Engineering, Enforcement,

Education, and Emergency Medical Services (Finkel et al. 2020). One innovative engineering strategy is implementing traffic signs, such as in-crosswalk traffic signs (NHTSA 2006). These signs help enhance crosswalk visibility and appear to be more intuitive since they are positioned directly in the traveler's path (Houten et al. 2012). In-street signs are also more impactful on low-speed, two-lane streets (Arhin et al. 2022; Redmon 2011).

Studies have revealed that implementing traffic signs within crosswalks can effectively reduce pedestrian accidents by enhancing speed regulations and promoting the yielding behavior of drivers (Ellis et al. 2007; Gedafa et al. 2014; Huang et al. 2000; Kannel et al. 2003; Strong and Ye 2010). Placing "YIELD TO PEDESTRIANS" signs within crosswalks notably reduced speed and enhanced the drivers' tendency to yield. A study by Gedafa et al. compared the effectiveness of the YIELD sign at five points from the crosswalk line and found that positioning the sign directly at the crosswalk line was the most effective (Gedafa et al. 2014). Additionally, a comparative study assessing the efficacy of these traffic signs positioned at distances of 0 ft, 20 ft, and 40 ft from the crosswalk found that the sign placed at 0 ft was equally or more effective than signs placed at other distances (Ellis et al. 2007).

Research conducted in Iowa analyzed the effects before and after implementing the State Law – Yield to Pedestrians at three locations and concluded that the sign positively affected driver behavior (Kannel et al. 2003). An observational study focused on the spillover effects of within-crosswalk signs reported that the signs positively impact and enhance motorist and pedestrian behaviors (Strong and Ye 2010). Another study comparing the single and gateway configurations of in-crosswalk signs discovered that all setups effectively increased the yielding percentage (Bennett et al. 2014).

Problem Statement

Past studies documented the significance of within-crosswalk traffic signs in reducing traffic speed and increasing the drivers' yielding behavior (Ellis et al. 2007; Gedafa et al. 2014; Huang et al. 2000; Pulugurtha et al. 2012). Studies have also discussed the effect of placing in-crosswalk traffic signs, such as State Law: Yield to Pedestrians Within Crosswalk and State Law: STOP to Pedestrians Within Crosswalk, at different locations along the road center line (Ellis et al. 2007; Gedafa et al. 2014); however, the authors recommended research on the repeatability of their results at other sites to increase the robustness of their findings. The impact of traffic signs on speeding and yielding may differ based on the type of within-crosswalk sign. A comparison of signage impacts in various time circumstances as well as during school and non-school sessions was not investigated; therefore, the primary objective of this research was to compare the effectiveness of in-crosswalk "YIELD TO PEDESTRIANS" and "STOP FOR PEDESTRIANS" signs at different times and locations. This research aims to contribute valuable insights for urban planners, transportation authorities, and safety advocates.

METHODOLOGY

Study Area and Materials

The traffic speed and yield data were collected at four locations in Grand Forks, North Dakota, USA. The main facilities in the city include business areas, residence areas, schools, and recreational parks. The city streets that are close to the recreational parks and schools experience

more pedestrians and bicyclists; therefore, those regions were selected for data collection. The streets selected for the study were 6th Ave N, S 25th St, Cherry St, and 11 Ave S. Figure 1 indicates the location of the study areas selected for speed and yield data collection.

The speed data were collected during in-school hours and times when schools were not in session at all locations; however, the yield data were collected at all locations for the in-school sessions only. Table 1 summarizes the main features and collected data types at each location.

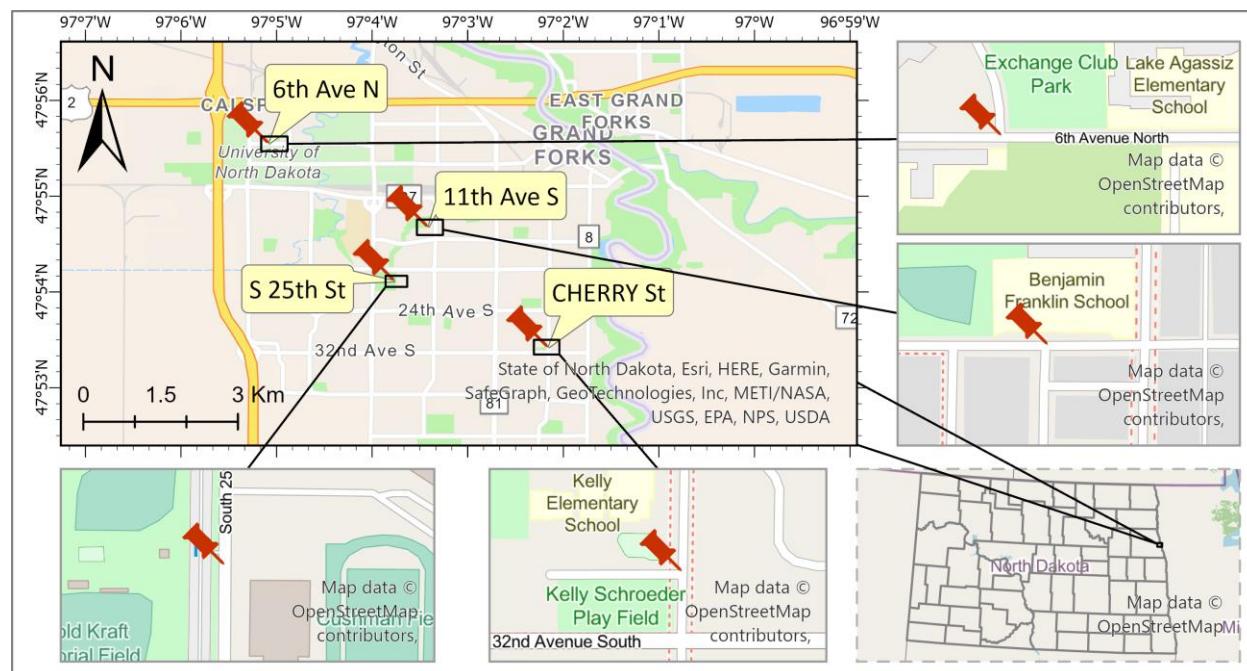


Figure 1. Study area

Table 1. Study location features (NDDOT 2021)

Location	Number of Lanes	AADT (2020)	Posted Speed Limit (mph)	Collected Data			
				School	No-School	School	No-School
6 th Ave N	Two-lane with turning-lane	3908	25	*	*	*	*
11 th Ave S	Two lane	2320	20	*	*	*	*
Cherry St	Two lane	3065	20	*	*	*	*
S 25th St	Two lane	1550	20	*	*	*	*

Data Collection

The regulatory in-street traffic signs described in Section 2B.12 of the FHWA Manual on Uniform Traffic Control Devices (FHWA 2009) were used. Figure 2 presents the two traffic signs placed at the edge of the crosswalk lines at 25th Ave S. Vehicle speed data were collected using a Scout Wireless Handheld Traffic Radar Gun by Decatur.

The speed and yield data were collected at the test streets with (W) and without (WO) the two within-crosswalk traffic signs. The data were collected twice a day from May 2023 to October 2023, during the morning (M) and afternoon (A) hours at 20-minute intervals. The speed and yield data were collected at free-flow traffic conditions and peak-hour conditions, respectively. These free-flow conditions are usually observed during off-peak hours (TRB 2000). The traffic signs were placed at the most effective location: the intersection of the road center line and crosswalk line (Ellis et al. 2007; Gedafa et al. 2014).

The minimum, average, 85th percentile, and maximum speeds were calculated. The 85th percentile speed is a basic element in setting speed limits (Forbes et al. 2012). The speed for turning vehicles was excluded from the analysis since the drivers reduced speed even without the presence of the traffic signs. The yield data were collected at peak hours and only during school sessions. The drivers were scored according to how they interacted with the pedestrians.

The leading vehicle's speed and yield score were considered when vehicles traveled closely. The stopping sight distance (SSD) determined vehicle proximity, and roads were marked at this distance from the pedestrian crossing line. The SSD was calculated based on posted speed limits at each site and consisted of brake reaction distance and braking distance (AASHTO 2011). Vehicles following another within a distance shorter than the SSD were excluded from the analysis. Drivers received a yielding score if they stopped or yielded for pedestrians. Drivers also received a yielding score if pedestrians appeared after drivers passed the SSD mark. A driver was marked as not yielding if the pedestrian reached the road crossing before the driver reached the SSD mark and did not yield. Any conflict between a driver and a pedestrian was considered as not yielding.

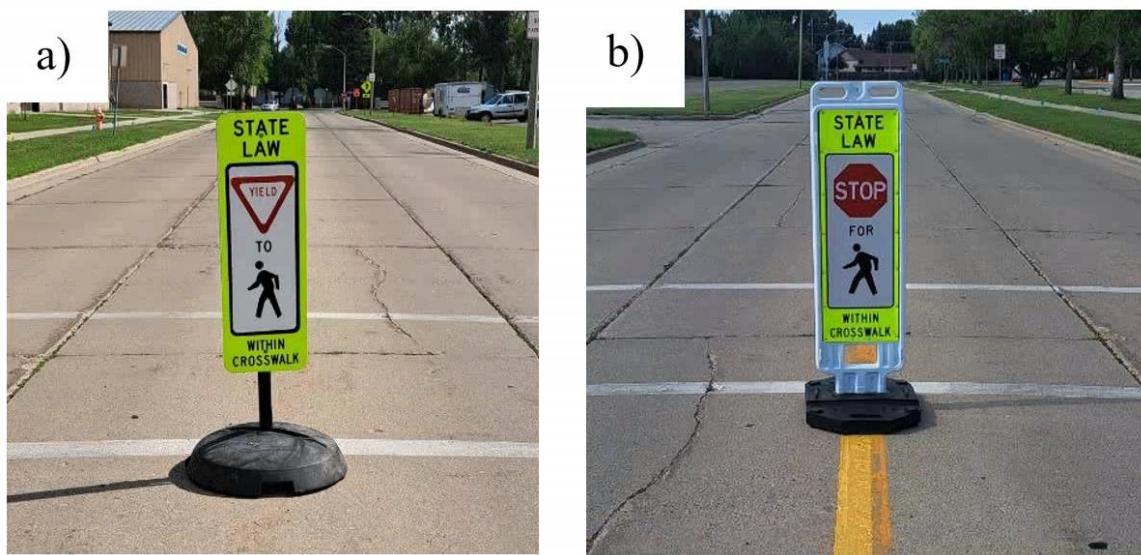


Figure 2. Within-crosswalk traffic signs at S 25th St: a) YIELD TO PEDESTRIANS: R1-6 and b) STOP TO PEDESTRIANS: R1-6a

Significance Difference Tests

Statistical tests were used to check the significant difference between the with and without traffic sign yield and speed data. A 95% confidence level was used for all statistical tests. An

independent t-test was used to test for the significant difference between the average speeds with and without traffic signs. This test can be used to make inferences about two independent means (Ott and Longnecker 2015). The null hypothesis for the t-test stated that the means of the two samples were not significantly different and could be rejected when the p-value was less than the selected significance level (Mendenhall et al. 2012).

Chi-squared and two-proportion tests were used to check the yielding proportion difference between the with and without conditions. The tests were used to test the significant difference between two categorical variable proportions, and the null hypothesis for these tests stated that there was no significant difference between the two sample proportions (Mendenhall et al. 2012; Ott and Longnecker 2015). Figure 3 summarizes the main steps followed while conducting this study.

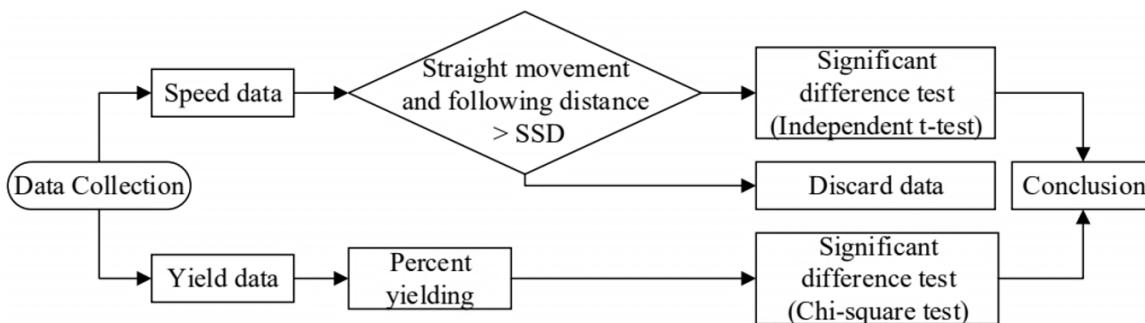


Figure 3. Study flowchart

RESULTS AND DISCUSSIONS

Effect of Traffic Signs on Speed

The minimum, average, 85th percentile, and maximum speeds at all locations were calculated from the collected data. The presence of the YIELD and STOP within-crosswalk signs resulted in a lower average speed for both in-school sessions and times when schools were not in session. The 85th percentile speed was also lower when the traffic signs were present on the road crosswalk. The minimum and maximum speeds observed were generally higher for the without conditions, and there were some exceptions where the drivers traveled at a higher speed regardless of the traffic signs. Figure 4 summarizes the speed data and standard deviation at all locations when schools were not in session. The speed reduction pattern was also similar for the in-school session data. Figure 5 summarizes the speed analysis results with standard deviation for the in-school session data.

An independent t-test with a significance level of 0.05 indicated the presence of significant differences in the average speeds at the two conditions. Table 2 presents the statistical test summary for both traffic signs during in-school sessions and times when schools were not in session. The results indicate that the presence of traffic signs resulted in a significant reduction in the average speed of drivers at all locations. A significant average speed reduction was observed in more than 93% and 87% of the total cases for the YIELD and STOP signs, respectively, when schools were not in session. Likewise, 81% and 75% of the cases attributed to YIELD and STOP signs, respectively, indicated a decrease in speed during in-school sessions. The standard deviation for more than 99% of the cases ranged from 3 mph to 5 mph.

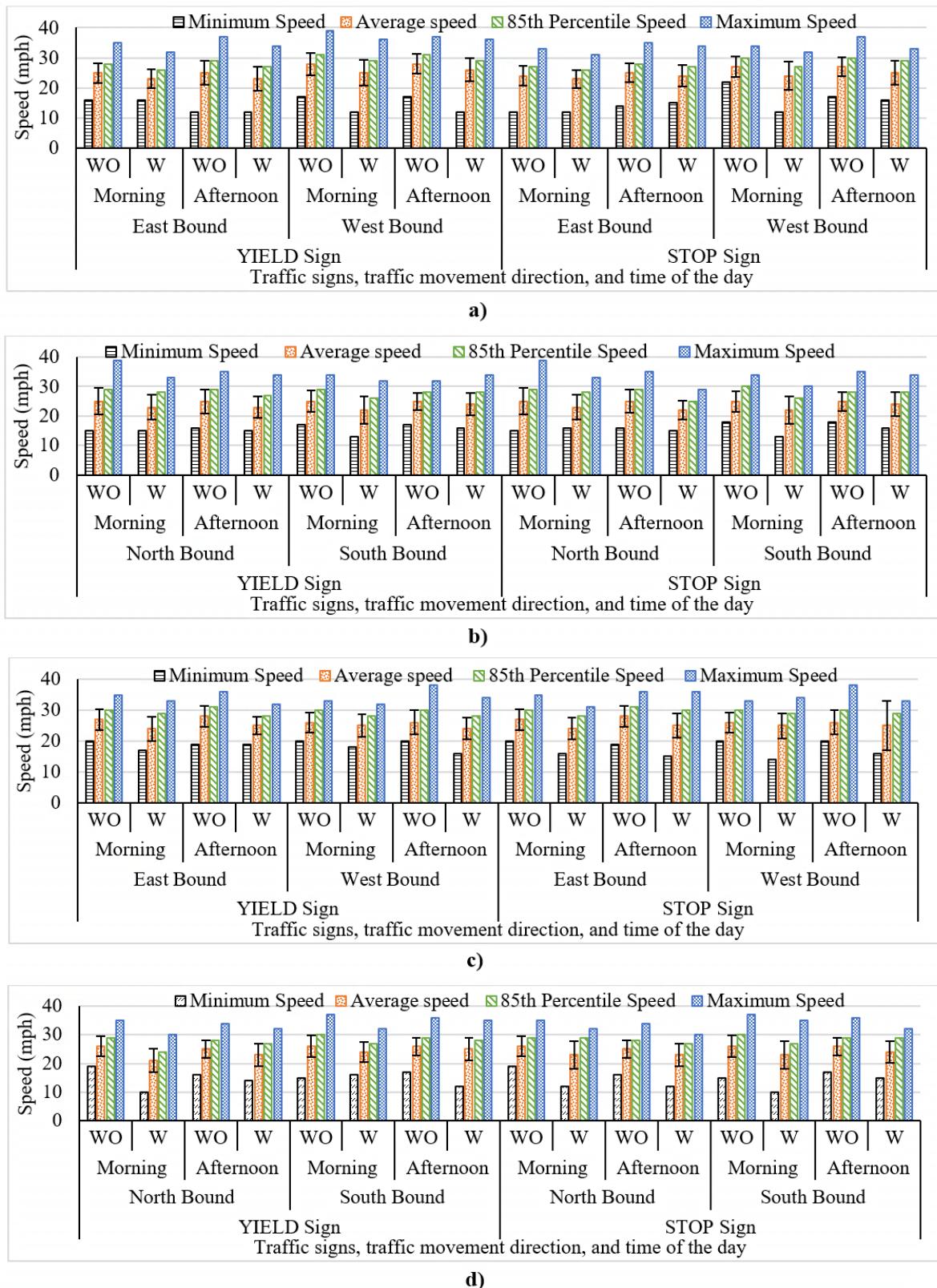


Figure 4. Speed data summary: no-school session a) 6th Ave N, b) 11th Ave S, c) 25th Ave S, and d) Cherry St

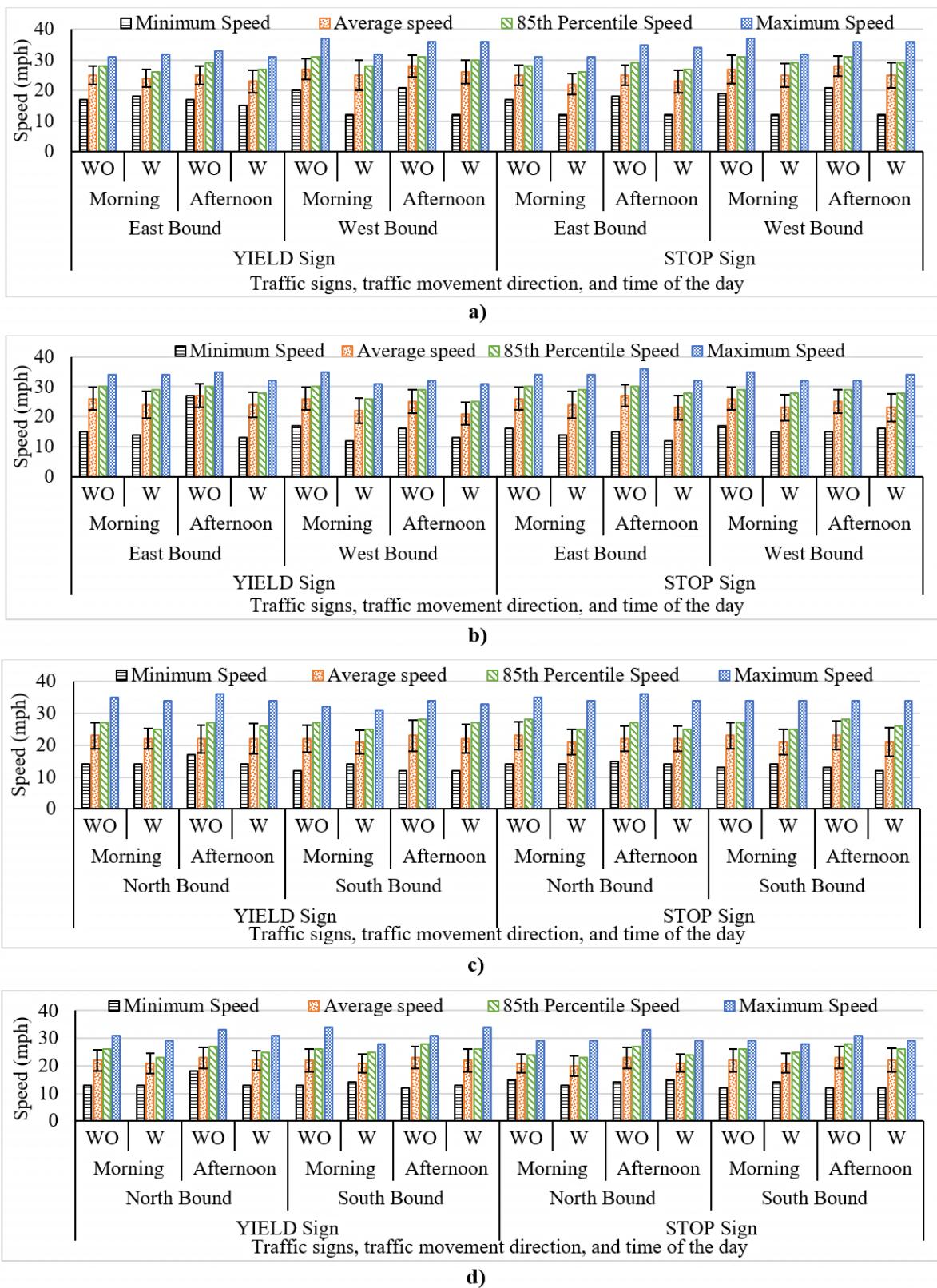


Figure 5 Speed data summary: in-school session a) 6th Ave N, b) 11th Ave S, c) 25th Ave S, and d) Cherry St

Table 2. Significant difference test for traffic speed

No- School session			YIELD sign		Sig. Diff.		STOP sign		Sig. Diff.					
Street name	Direction	Time	WO	W	p-value	(95% CI)	WO	W	p-value	(95% CI)				
			Avg Speed	n	Avg Speed	n			Avg Speed	n				
6 th Ave N	EB	M	25	193	23	168	<0.0001 S	24	168	23	153	0.0005 S		
		A	25	138	23	152	<0.0001 S	25	161	24	145	0.0016 S		
	WB	M	28	129	25	128	<0.0001 S	27	68	24	60	<0.0001 S		
		A	28	155	26	158	<0.0001 S	27	86	25	58	0.0158 S		
11 th Ave S	EB	M	27	40	25	52	0.0017 S	27	50	24	67	0.0001 S		
		A	28	63	25	86	<0.0001 S	28	56	25	76	<0.0001 S		
	WB	M	26	45	25	79	0.0372 S	26	52	25	73	0.0732 N		
		A	26	59	24	74	0.0005 S	26	62	25	67	0.0193 S		
Cherry St	NB	M	26	53	21	64	<0.0001 S	26	63	23	49	0.0017 S		
		A	25	100	23	94	<0.0001 S	25	88	23	82	0.0008 S		
	SB	M	26	70	24	89	0.0002 S	26	61	23	66	0.0005 S		
		A	26	99	25	84	0.0279 S	26	111	24	90	<0.0001 S		
S 25th St	NB	M	25	50	23	63	0.0095 S	25	49	23	52	0.0308 S		
		A	25	56	23	67	0.0044 S	25	54	22	56	<0.0001 S		
	SB	M	25	57	22	55	0.0006 S	25	44	22	50	0.0004 S		
		A	25	84	24	58	0.1942 N	25	68	24	58	0.2206 N		
Overall			M	24.1	637	22.3	698	<0.0001 S	23.9	555	22.4	570	0.0002 S	
			A	24.4	753	22.8	773	<0.0001 S	24.7	686	22.6	632	<0.0001 S	
School session														
6 th Ave N	EB	M	25	102	24	92	0.0015 S	25	96	22	89	<0.0001 S		
		A	25	94	23	88	<0.0001 S	25	85	23	94	<0.0001 S		
	WB	M	27	73	25	60	0.0004 S	27	89	25	99	0.0039 S		
		A	28	67	26	72	0.0067 S	28	80	25	74	0.0003 S		
11 th Ave S	EB	M	26	70	24	76	0.0023 S	26	60	24	67	0.0011 S		
		A	27	62	24	69	0.0035 S	27	73	23	55	<0.0001 S		
	WB	M	26	56	22	48	<0.0001 S	26	52	23	45	0.0011 S		
		A	25	44	21	54	<0.0001 S	25	43	23	62	0.0027 S		
Cherry St	NB	M	22	78	21	80	0.0362 S	21	94	20	83	0.2887 N		
		A	23	90	22	71	0.0063 S	23	55	21	67	0.0148 S		
	SB	M	22	81	21	62	0.0211 S	22	73	21	81	0.6718 N		
		A	23	98	22	75	0.2132 N	23	69	22	59	0.0487 S		
S 25 th St	NB	M	23	57	22	71	0.1697 N	23	59	21	78	0.0060 S		
		A	22	77	22	61	0.4975 N	22	53	22	66	0.3385 N		
	SB	M	22	67	20	54	0.0419 S	23	79	21	73	0.0122 S		
		A	23	70	21	62	0.0308 S	23	64	21	50	0.1740 N		
Overall			M	24.1	584	22.3	543	<0.0001 S	23.9	602	22.4	615	<0.0001 S	
			A	24.4	602	22.8	552	<0.0001 S	24.7	522	22.6	527	<0.0001 S	

S: Significant at a 0.05 significance level, N: Not significant at a 0.05 significance level.

Table 3. Significant difference test using Chi-square (χ^2) and Proportion test for Yielding

YIELD Sign		Time of the day	Yielding data (Proportion)			Significance test	Combined	
Street name	Direction		WO	W	χ^2 (p-value)			
6 th Ave	East	M	57 (66.7)	64 (89.1)	8.964 (0.003) S	-2.994 (0.003) S	-4.627 (<0.00001)	
		A	66 (69.7)	67 (91.0)	9.634 (0.002) S	-3.104 (0.002) S		
	West	M	63 (69.8)	56 (83.9)	3.270 (0.071) N	-1.808 (0.070) N	S	
		A	59 (67.8)	68 (79.4)	2.216 (0.137) N	-1.487 (0.136) N		
11 th Ave S	East	M	54 (72.2)	61 (90.2)	6.177 (0.012) S	-2.485 (0.013) S	-4.492 (<0.00001)	
		A	63 (76.2)	60 (91.7)	5.406 (0.020) S	-2.325 (0.020) S		
	West	M	53 (84.9)	47 (91.5)	1.023 (0.312) N	-1.011 (0.313) N	S	
		A	56 (76.8)	59 (96.6)	9.955 (0.002) S	-3.155 (0.002) S		
Cherry St	North	M	90 (68.9)	84 (83.8)	4.951 (0.026) S	-2.225 (0.026) S	-2.950 (0.0032)	
		A	83 (71.1)	81 (77.8)	0.964 (0.326) N	-0.982 (0.327) N		
	South	M	80 (68.8)	84 (81.0)	3.254 (0.071) N	-1.804 (0.072) N	S	
		A	70 (82.9)	76 (86.8)	0.452 (0.501) N	-0.672 (0.503) N		
S 25 th St	North	M	73 (74.0)	78 (92.3)	9.176 (0.002) S	-3.029 (0.002) S	-4.804 (<0.00001)	
		A	75 (76.0)	73 (86.3)	2.559 (0.109) N	-1.599 (0.109) N		
	South	M	83 (75.9)	87 (92.0)	8.191 (0.004) S	-2.862 (0.004) S	S	
		A	80 (73.8)	85 (87.1)	4.669 (0.031) S	-2.161 (0.031) S		
STOP Sign		6 th Ave	M	58 (60.7)	74 (93.2)	11.908 (0.001) S	-3.451 (0.001) S	
Cherry St	East		A	59 (76.3)	70 (80.0)	0.262 (0.609) N	-0.512 (0.610) N	
			M	63 (76.2)	64 (84.4)	1.345 (0.246) N	-1.159 (0.246) N	
	West		A	65 (75.4)	71 (91.5)	6.539 (0.011) S	-2.557 (0.011) S	
11 th Ave S	East	M	58 (72.4)	67 (91.0)	5.949 (0.015) S	-2.439 (0.015) S	-4.070 (<0.00001)	
			A	63 (73.0)	68 (92.6)	5.556 (0.018) S	-2.357 (0.018) S	
	West	M	56 (83.9)	49 (91.8)	1.507 (0.219) N	-1.227 (0.219) N	S	
		A	46 (73.9)	43 (90.7)	4.246 (0.039) S	-3.061 (0.039) S		
S 25 th St	North	M	81 (69.1)	78 (89.7)	10.26 (0.001) S	-3.203 (0.001) S	-4.273 (<0.00001)	
			A	74 (73.0)	77 (84.4)	2.958 (0.085) N	-1.720 (0.085) N	
	South	M	70 (72.9)	73 (83.6)	2.412 (0.120) N	-1.553 (0.121) N	S	
		A	73 (76.7)	75 (89.3)	4.198 (0.041) S	-2.049 (0.040) S		

S: Significant at a 0.05 significance level, N: Not significant at a 0.05 significance level.

The study areas have similar features, and the individual values can be added to check the overall significance of the differences (Gedafa et al. 2014). The overall tests revealed that the

speed reduction due to the traffic signs significantly reduced the average speed; therefore, the null hypothesis was rejected.

Effect of Traffic Signs on Yielding to Pedestrians

The proportion of drivers who yielded to pedestrians to the total number of scored drivers for each location was calculated and used for the statistical analysis. Table 3 presents the summary of the significant tests. The raw data illustrates that the YIELD and STOP signs both increased the proportion of drivers yielding to pedestrians; however, the yielding proportion was significant for only 56% and 68% of the individual cases for YIELD and STOP signs, respectively. The traffic sign conditions resulted in higher yielding proportions; however, sites such as 6th Ave N and Cherry St exhibited more cases where the results were insignificant. This discrepancy might be linked to higher driving speeds and relatively elevated instances of speeding violations at these locations.

The statistical tests demonstrated that the presence of traffic signs significantly increased the proportion of drivers yielding to pedestrians across all locations. Specifically, the STOP sign condition exhibited a higher number of significant cases. The null hypothesis can be rejected based on the calculated overall p-values, which were all below the significance level.

Comparison of the Effect of In-Crosswalk YIELD and STOP Signs

Table 4 summarizes the effectiveness comparison of the traffic signs on speeding. The results indicate that the overall effectiveness of the within-crosswalk STOP and YIELD signs was comparable. The effectiveness of the signs was significantly different at 6th Ave N at times when schools were not in session. The STOP signs resulted in a relatively lower average speed value than the YIELD signs; however, the average speeds for both cases had p-values higher than the confidence level at the other three locations. Furthermore, the differences in average speed values at all locations due to the traffic signs were insignificant. The null hypothesis cannot be rejected since the p-values for the overall cases were higher than 0.05.

Table 4. Significant difference test between YIELD and STOP signs: Speed data summary

Location	No-school Session				Sig. Diff (95% CI)	In-School Session				Sig. Diff (95% CI)		
	YIELD sign		STOP sign			YIELD sign		STOP sign				
	Avg Speed	n	Avg Speed	n		Avg Speed	n	Avg Speed	n			
6 th Ave N	24.1	606	23.5	416	0.0017 S	24.2	312	24.1	356	0.6599 N		
11 th Ave S	24.8	291	24.9	283	0.7064 N	23.0	247	23.2	229	0.5866 N		
Cherry St	23.2	331	23.4	287	0.5447 N	21.3	288	21.0	290	0.3122 N		
S 25 th St	23.1	243	23	216	0.7359 N	21.2	248	21.3	267	0.8949 N		
Overall	23.9	1471	23.7	1202	0.3410 N	22.5	1095	22.5	1142	0.8144 N		

S = significant at a 0.05 significance level, N = Not significant at a 0.05 significance level.

Another comparison between the effectiveness of the two signs was performed using the effect on yielding to pedestrians. Table 5 presents the yielding proportion differences between the two signs. The significance proportion test indicated that the yielding proportion differences between the two signs were insignificant at all locations; therefore, the signs had a comparable effect and can be used to reduce speed and increase yield to pedestrians on two or three-lane streets.

Table 5. Significant difference test between YIELD and STOP signs: Yield data summary

Location	YIELD	STOP	z-score (p-value)	Combined z-score (p-value)
6 th Ave N	255 (85.9)	279 (87.5)	-0.535 (0.596) N	-0.497
11 th Ave S	227 (92.5)	204 (91.2)	0.506 (0.610) N	(0.617)
Cherry St	325 (82.2)	303 (86.8)	-1.603 (0.110) N	N
S 25 th St	323 (89.5)	310 (87.1)	1.036 (0.298) N	

S Significant at a 0.05 significance level, N Not significant at a 0.05 significance level.

CONCLUSIONS

The subsequent conclusions can be drawn based on the results of the analysis:

- The introduction of crosswalk STOP and YIELD signs led to a decrease in both average and 85th percentile speeds, establishing significant reductions in speed.
- The changes in vehicle speed were significant across various times, including mornings, afternoons, and whether or not schools were in session. Implementing these regulatory signs could effectively lower the risk of speed-related traffic crashes.
- The presence of these traffic signs significantly enhanced yielding behavior toward pedestrians. Placing these signs at the crosswalk could potentially reduce traffic-related pedestrian crashes.
- There was no significant difference between the impact of the two types of traffic signs on speeding and yielding behaviors. This finding implies that transportation planners have the flexibility to use either sign to enhance pedestrian and overall road safety.

FUTURE WORK

A longitudinal study is recommended to assess the long-term effects of implementing crosswalk STOP and YIELD signs on speed reduction, yielding behavior, and the overall incidence of traffic accidents involving pedestrians. A study should also be performed considering the impact of contextual factors, such as urban versus rural settings or diverse traffic patterns, on the effectiveness of these traffic signs. These studies would provide a more comprehensive understanding of the sustained impact of these measures on different environments.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the Grand Forks-East Grand Forks Metropolitan Planning Organization and the North Dakota Department of Transportation

for providing crash data. Additionally, the authors thank the University of North Dakota's Civil Engineering Department for providing the necessary materials used in this research. The views and conclusions expressed in this manuscript belong solely to the authors and do not necessarily reflect the perspectives of the funding agencies.

REFERENCES

- AASHTO. (2011). *A Policy on Geometric Design of Highways and Streets*, American Association of State Highway and Transportation Officials, Washington, DC.
- Arhin, S. A., Gatiba, A., Anderson, M., Manandhar, B., Ribbisso, M., and Acheampong, E. (2022). "Effectiveness of Modified Pedestrian Crossing Signs in an Urban Area." *Journal of Traffic and Transportation Engineering (English edition)*, 9(1), 21-32.
- Bennett, M. K., Manal, H., and Van Houten, R. (2014). "A Comparison of Gateway In-street Sign Configuration to Other Driver Prompts to Increase Yielding to Pedestrians at Crosswalks." *Journal of Applied Behavior Analysis*, 47(1), 3-15.
- Ellis, R., Van Houten, R., and Kim, J.-L. (2007). "In-Roadway 'Yield To Pedestrians' Signs: Placement Distance and Motorist Yielding." *Transportation Research Record*, 2002(1), 84-89.
- FHWA (Federal Highway Administration). (2009). *Manual on Uniform Traffic Control Devices for Streets and Highways 2009 Edition*. Washington DC.
- Finkel, E., McCormick, C., Mitman, M., Abel, S., and Clark, J. (2020). *Integrating the Safe System Approach With the Highway Safety Improvement Program: An Informational Report*. United States. Federal Highway Administration. Office of Safety.
- Forbes, G., Gardner, T., McGee, H., and Srinivasan, R. (2012). "Methods and Practices for Setting Speed Limits: An Informational Report by the Institute of Transportation Engineers and Federal Highway Administration." *Institute of Transportation Engineers. ITE Journal*, 82(8), 20.
- Gedafa, D., Kaemingk, B., Mager, B., Pape, J., Tupa, M., and Bohan, T. (2014). "Impacts of Alternative Yield Sign Placement on Pedestrian Safety." *Transportation Research Record*, 2464(1), 11-19.
- Houten, R. V., LaPlante, J., and Gustafson, T. (2012). "Evaluating Pedestrian Safety Improvements: Final Report." Michigan Department of Transportation.
- Huang, H., Zegeer, C., and Nassi, R. (2000). "Effects of Innovative Pedestrian Signs at Unsignalized Locations: Three Treatments." *Transportation Research Record*, 1705(1), 43-52.
- Kannel, E. J., Souleyrette, R. R., and Tenges, R. (2003). "In-Street Yield to Pedestrian Sign Application in Cedar Rapids, Iowa."
- Kara, M., and Elizabeth, P. (2022). *Pedestrian Traffic Fatalities by State*. Governors Highway Safety Association, Washington, DC, USA.
- TRB (Transportation Research Board). (2000). *Highway Capacity Manual*. Transportation Research Board, Washington, DC.
- Mendenhall, W., Beaver, R. J., and Beaver, B. M. (2012). *Introduction to Probability and Statistics*, Cengage Learning.
- NCSA (National Center for Statistics and Analysis). (2023). *Pedestrians: 2021 Data*" Traffic Safety Facts, National Highway Traffic Safety Administration, Washington, DC, USA.

- NDDOT (North Dakota Department of Transportation). (2021). *2021 Traffic Volume Map: Grand Forks County*. North Dakota, USA.
- NHTSA (National Highway Traffic Safety Administration). (2006). "Uniform Guidelines for State Highways Safety Programs: Highway Safety Program Guideline No. 19." National Highway Traffic Safety Administration.
- Ott, R. L., and Longnecker, M. T. (2015). *An Introduction to Statistical Methods and Data Analysis*, Cengage Learning.
- Pulugurtha, S. S., Nambisan, S. S., Dangeti, M. R., and Vasudevan, V. (2012). "Evaluation of Effectiveness of Traffic Signs to Enhance Pedestrian Safety." *Transportation Research Board* 2299(1), 100-109.
- Redmon, T. (2011). "Evaluating Pedestrian Safety Countermeasures." *Public Roads*, 74(5).
- Strong, C., and Ye, Z. (2010). "Spillover Effects of Yield-to-Pedestrian Channelizing Devices." *Safety Science*, 48(3), 342-347.
- USDOT. (2022). *National Roadway Safety Strategy*. US Department of Transportation, Washington, DC.

Noteworthy Practices for Cross-Agency Traffic Safety Coordination

Taylor Dinehart, AICP¹; Tia Boyd, AICP²; Jeff Kramer, AICP³; Lama Alfaseeh, Ph.D.⁴;
and Sarah Caper, AICP⁵

¹Research Associate, Transportation Planning, Policy, and Processes, Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL. Email: tdinehart@usf.edu

²Research Associate, Transportation Planning, Policy, and Processes, Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL. Email: tiaboyd@usf.edu

³Program Director, Transportation Planning, Policy, and Processes, Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL. Email: kramer@usf.edu

⁴Planner, Community & Infrastructure Planning Dept., Hillsborough County, Tampa, FL. Email: alfaseehl@hillsboroughcounty.org

⁵Executive Planner, Community & Infrastructure Planning Dept., Hillsborough County, Tampa, FL. Email: capers@hillsboroughcounty.org

ABSTRACT

Ensuring the safety of roadways for travelers of all modes is a collaborative process that involves stakeholders across state and local government agencies, advocacy groups, research communities, and between the private and the public sectors. To better understand the policies and practices regarding data-driven safety coordination between counties and first responder agencies—especially coordination activities that leverage available data assets of local jurisdictions to benefit and maximize the resources of the first responder agencies—a survey was conducted of 24 of the most populous counties in Florida, representing the largest third of counties in the state. An in-depth literature review was completed to supplement information gathered from the survey and to assist in analyzing the results. From the literature review and survey findings, a set of suggestions to enhance coordination between counties and local first responder agencies on transportation safety were developed. Although this study was conducted in Florida, the findings and suggestions of this study are relevant to understanding agency coordination on transportation in both the US and internationally.

Key words: roadway safety; safety coordination; transportation planning coordination

INTRODUCTION

Traffic-related fatalities are increasing nationwide. In 2021, traffic crashes accounted for over 1.7 million injuries and 42,939 fatalities in the United States. This represents a 10% increase in traffic fatalities from 2020, the highest number of fatalities since 2005, and the largest ever annual percentage increase in the 48-year history of the Fatality Analysis Reporting System (FARS) (National Highway Traffic Safety Administration 2023). The state of Florida represented 14.6% (252,971) of those injuries (Florida Department of Highway Safety and Motor Vehicles 2023). In addition to the effects of injuries and fatalities on individuals and communities, these crashes create significant cost burdens, most of which are borne by individuals not involved in motor vehicle crashes. These cost burdens are experienced through insurance premiums, taxes, and congestion-related costs such as lost time, excess fuel consumption, and increased environmental impacts.

The profound influence of roadway safety on human societies has led to substantial investments in safety countermeasures, the passage and implementation of safety related traffic laws, and the enforcement of various regulations related to the production of safer vehicles for use on the country's highways. Although many of these initiatives have achieved some success in lowering the occurrence of roadway crashes and lessening their severity, it remains evident that more needs to be done. Collaboration among state and local government agencies, advocacy groups, research communities, and the private and public sectors is crucial for ensuring the safety of roadways for all modes of travelers. One vital point of collaboration and coordination among these groups, especially in the face of limited funds and personnel, is the use of shared data to identify patterns and prioritize efforts to improve roadway safety.

In an effort to make roadways safer and to relieve the cost burden that unsafe roadways place on Hillsborough County, Florida residents, the Hillsborough County Community & Infrastructure Planning Department asked researchers from the Center for Urban Transportation Research (CUTR) at the University of South Florida (USF) to work with the County to identify and summarize noteworthy and reproducible practices currently being used in other large Florida Counties regarding data-driven safety coordination between counties and first responder agencies. This research project explored methods county planners might employ to assist law enforcement and other first responders within their county in identifying the most dangerous areas for road users as part of a larger roadway safety strategy (i.e., the Safe System Approach and Vision Zero). This paper presents the findings and recommendations of that study and provides insight into current policies and practices that leverage both roadway safety data and collaboration between key stakeholders to further roadway safety initiatives.

STUDY METHODS

Step 1: Comprehensive review. This review covered a broad examination of various sources, including scholarly literature, reports, policies, and grey literature related to coordination between local jurisdictions and first responder agencies as it applies to roadway safety. Topics of this review included: roadway safety data and data sharing, the Safe System Approach, collaborative roadway safety strategies and programs, examples of collaboration between transportation agencies and first responder agencies, the role of enforcement in roadway safety, and examples of collaboration between research institutions and first responder agencies.

Step 2: Develop and administer survey. Information collected through the comprehensive review was used to inform the design of an online survey that targeted staff or leadership in county transportation planning/engineering departments. Using 2022 population data from the U.S. Census Bureau, Florida counties were sorted from largest to smallest by population, and the 24 most populous counties were chosen to be included in the study. Those 24 counties were:

- Miami-Dade County
- Broward County
- Palm Beach County
- Hillsborough County
- Orange County
- Duval County
- Pinellas County
- Lee County
- Polk County

- Brevard County
- Pasco County
- Volusia County
- Seminole County
- Sarasota County
- Manatee County
- Osceola County
- Lake County
- Collier County
- Marion County
- St. Lucie County
- Escambia County
- St John's County
- Leon County
- Alachua County

Qualtrics survey software was used to construct a 21-question electronic online survey. Questions were presented as multiple choice, checkboxes, or textboxes. Several of the questions included multiple parts using survey logic that displayed additional questions based on the response. The survey was organized into sections which included a screening section, coordination on transportation safety, data sharing and analysis practices, safety and data training, obstacles encountered, and lessons learned. The survey opened on September 5th, 2023, and closed on October 6th, 2023, and was distributed via email, with emails and phone calls for reminders and follow-up as needed.

Step 3: Draw conclusions and develop recommendations. The outcomes of this research effort are described in more detail later in this paper.

RESULTS AND FINDINGS

Comprehensive review.

The comprehensive review included scholarly journal articles, books, government reports, planning documents, policy documents, websites, webinars, and other similar sources looking at data-driven safety coordination practices between local jurisdictions and first responder agencies. The following represents a summary of the key findings from the literature review.

Roadway safety data and data sharing

Police crash report data is the most commonly used source of data for transportation safety research (Cherry et al. 2018; Hosseinzadeh et al. 2022; Watson et al. 2015). Florida-specific data sources related to roadway safety analysis and research include Signal Four Analytics, which uses Florida Highway Safety and Motor Vehicles (FLHSMV) data and is the official crash data repository and the sole source of crash information for Florida Department of Transportation (FDOT) safety analyses. FDOT's State Safety Office also provides analysis of FLHSMV data through their State Safety Office Geographic Information System (SSOGis), which is a valuable resource for planners and traffic safety analysts.

A number of post-crash datasets are available at the national level and can be used in conjunction with local crash data. The National Highway Traffic Safety Administration's

(NHTSA) analysis and research tools are the most commonly used national data source related to roadway safety. NHTSA provides national crash data through a number of tools including its core tool, the Fatality Analysis Reporting System (FARS). The NHTSA tools are useful for comparing Florida crash data to other states as well as to national averages.

There are numerous data quality issues associated with roadway safety data, particularly with crash data coming from police reports, including:

- Underreporting: Not all accidents or incidents get reported. Minor accidents or those resulting in no injuries might not be documented, leading to incomplete data and an inaccurate understanding of the true safety situation.
- Incomplete Data: Missing information in crash reports, such as details about road conditions, weather, or contributing factors, can limit the analysis and understanding of the causes behind accidents.
- Inconsistent Data Entry: Inaccuracies or inconsistencies in data entry, especially if multiple agencies are involved in reporting, can lead to discrepancies or errors in the recorded information.
- Data Bias: Certain factors might introduce bias in the reported data, such as overrepresentation or underrepresentation of specific types of accidents, locations, or demographics. This bias can skew the analysis and conclusions drawn from the data.
- Data Format and Compatibility: Roadway safety data might come from various sources or formats, making integration and analysis challenging. Incompatible data formats or lack of standardization can hinder comprehensive analysis.
- Lack of Timeliness: Delayed reporting or data updates might impact the real-time monitoring of safety trends or the immediate implementation of preventive measures.
- Data Source Reliability: Depending on the source of the data, its reliability might vary. Data from different jurisdictions or agencies may have different collection methods, accuracy levels, or reporting standards.

Noteworthy collaborative efforts for roadway safety

Cross agency, multidisciplinary collaborations to improve roadway safety are baked into modern roadway safety strategies at the national, state, and local levels and examples across the county abound. For instance, in the City of Dixon and Solano County, CA, a Safe Routes to School project combined education, enforcement, and engineering to increase traffic safety for children. This project involved collaboration between schools, local police, the county health department, and the county's transportation planners/engineers. Benefits to the community included better air quality, a reduction in traffic, and an increase in roadway safety for children and the larger community (Institute for Local Government 2013).

Another example involves the Iowa Traffic Safety Data Service (ITSDS) and local/state police who both participate in "Safety Data Service"—a data sharing effort in which ITSDS collects, stores, and processes police crash data for State and local agencies. Local governments can request crash histories for specific areas, roads, and intersections for the purpose of identifying high-crash locations, developing safety plans, and analyzing safety programs (Federal Highway Administration 2014).

A third example is the Missouri Coalition for Roadway Safety—a partnership of Missouri safety advocates that includes law enforcement agencies, health care providers, courts, government agencies of various levels, advocacy groups, planning organizations, and concerned

citizens—who created a Strategic Highway Safety Plan which they attribute to a 40% decrease in traffic fatalities from 2005 to 2013 (Federal Highway Administration 2014).

Additionally, local government and law enforcement agencies have both found success in partnering with local universities in collaborative efforts, especially those that require specialized expertise and capabilities in data collection, processing, and analyzing. The University of California—Berkely, Louisiana State University, Purdue, and University of Florida have each joined with state and local agencies, as well as other stakeholders in transportation safety, to develop software programs, interactive websites, and other products to aid in data-driven decision making for roadway safety (SafeTREC 2023; Louisiana State University Center for Analytics & Research in Transportation 2023; Tarko et al. 2020; University of Florida GeoPlan Center 2023).

From the comprehensive review, the following stood out as noteworthy and replicable practices across the United States:

- Establishing integrated data platforms that combine various sources of information, including crash data, traffic flow, emergency response times, and infrastructure details. These platforms enable comprehensive analysis and informed decision-making by both local government authorities and first responders.
- Implementing systems that allow for real-time data sharing between local governments and first responder agencies for immediate access to critical information during emergencies, facilitating quicker and more effective responses.
- Utilizing predictive analytics to forecast potential safety issues. By analyzing historical data trends, local jurisdictions can proactively identify high-risk areas or times, allowing for targeted interventions and resource allocation.
- Integrating geographic information system (GIS) technology with safety data allows local governments and first responders to visualize and analyze data spatially. This aids in identifying geographic patterns, hotspot areas, and specific locations requiring safety improvements.
- Using safety data to establish performance metrics and regularly evaluating the effectiveness of safety initiatives. Local governments and first responder agencies are using these metrics to assess the impact of interventions and adjust strategies accordingly.
- Creating collaborative task forces or committees with representatives from law enforcement, emergency services, transportation planners, and other relevant agencies to foster coordinated efforts. These groups work together to analyze data, share insights, and develop strategies for improving safety.
- Providing training programs and capacity-building initiatives for first responders on data utilization and analysis.
- Leveraging data to communicate safety information to the public. Utilizing data visualization techniques or mobile apps, local governments and first responder agencies disseminate safety tips, road closure information, and emergency alerts to improve community awareness and engagement.

National level task forces and working groups have endorsed data sharing and the use of open data, lending credibility to the idea of collaborative approaches to using shared data between local governments and first responder agencies to address roadway safety issues.

Barriers to effective collection, use, and transparency of policing data include absence of national data standards and guidance, lack of internal capacity and vendor barriers, fear of mischaracterization of data, lack of accessibility and culturally informed practices, and

inconsistencies of state data reporting mandates. Some strategies identified in the literature for overcoming identified barriers in data sharing include:

- Finding and using committed champions to overcome reluctance to collaborate between agencies.
- Gaining buy-in from agency leaders with authority to make policy and funding decisions.
- Identifying existing data through cleaning and cross-referencing practices.
- Ensuring that all stakeholders contribute their ideas to data collection processes.
- Creating opportunities to bridge divisions between disparate stakeholder groups.
- Partnering with local universities.
- Using deliverables to advocate for funding sustainability.

Survey Results.

Respondents from fifteen counties completed the survey, for a 62% response rate. The following counties completed the survey:

- Broward County
- Palm Beach County
- Hillsborough County
- Orange County
- Duval County
- Lee County
- Brevard County
- Pasco County
- Seminole County
- Sarasota County
- Manatee County
- Lake County
- Collier County
- Marion County
- Alachua County

The first three questions of the survey were screening questions which asked respondents if they coordinate with first responders on transportation safety, share transportation safety data with first responders, or use data provided by first responder agencies for transportation safety studies, analysis, or decision making. Of the fifteen respondents, thirteen answered affirmatively to at least one of the three screening questions and those respondents continued with the remaining survey questions.

Of the thirteen respondents that answered affirmatively to the screening questions, the responses are recapped as follows:

- The majority of the counties surveyed share data with and/or use data from law enforcement. This data is used differently in each county.
 - Counties that share data with and/or use data from fire & rescue, EMS, and other first responder agencies also do so with law enforcement.
- Although most agencies share data with and/or use data from first responder agencies, many do not analyze crash data for local first responder agencies and/or share trends and findings with these agencies.
- Coordination on transportation safety happens through multiple methods. Emails, phone calls, or working groups, were selected the most frequently, followed by meetings.

- Regardless of the mechanism used, most coordination is ad hoc, with monthly or other regularly scheduled coordination occurring in only a few instances.
- Data is typically shared between counties and first responder agencies as needed/upon request. A few agencies share data monthly.
- Most counties use crash data, including data from Signal Four Analytics, to analyze transportation safety. Other frequently mentioned data include traffic volume, speed, and public feedback.
- Most counties do not have formal data sharing agreements with local first responder agencies.
- The two most frequently selected challenges for coordinating with first responder agencies on transportation safety include:
 - Limited number of staff dedicated to safety and supporting functions, including coordination, and
 - Adequate time for coordination.

CONCLUSIONS AND RECOMMENDATIONS

Sharing crash data between local jurisdiction planning staff and first responder agencies is vital for several reasons. First, it is instrumental in informing infrastructure planning by offering insights into frequent crash zones. This knowledge aids planning staff in identifying hazardous areas, and strategizing infrastructure enhancements like road redesigns, traffic signal modifications, and pedestrian safety improvements to prevent future crashes. Next, analyzing crash data leads to a deeper comprehension of crash causality, allowing for targeted interventions aimed at addressing specific issues contributing to crashes. Third, the mutual sharing of crash data facilitates effective resource allocation. This allows both county staff and first responders to prioritize areas of concern. This way, first responders can concentrate resources in high-risk areas indicated by crash data, while planners can earmark budgets for essential infrastructure changes. Lastly, the evaluation of safety measures or infrastructure changes becomes possible through the analysis of crash data, establishing a feedback loop. This process is vital for refining future strategies taken by planners and first responders to continually enhance road safety.

Additionally, data sharing fosters collaboration, enabling joint strategies for crash prevention and road safety. Integrating crash data ensures a holistic safety approach, combining reactive and proactive measures to reduce crashes and their impact. Overall, sharing crash data between county planning staff and first responder agencies facilitates informed decision-making, targeted interventions, and collaborative efforts aimed at creating safer roadways and communities.

Based on the comprehensive review and survey results, the following recommendations were formed. Mention of ‘county staff’ below generally refers to persons within the county planning department, transportation planning department, public works, or other similar planning and/or transportation departments, but may also include first responder agency staff as first responders may be housed within a county.

Encourage open discussions/storytelling between local government staff and first responders. Storytelling can be used as a creative approach to demonstrate how safety data are applied and how transportation safety strategies (e.g., countermeasures) are used to improve safety. The time and effort that county staff puts into data analysis and trends reporting could be used by law enforcement to pinpoint areas in most need of traffic safety enforcement countermeasures. Similarly, the experiences and stories of law enforcement can also be used to supplement analysis and inform infrastructure improvements to better align strategies with the

Safe System Approach (SSA). As stated in the literature, law enforcement's knowledge and experience related to traffic behavior can be used to identify and address transportation safety issues.

Identify opportunities for county staff and first responders to observe and learn from each other. To overcome data sharing barriers, the literature suggests creating opportunities to bridge divisions through experience. One approach to accomplish this goal is by shadowing—inviting different groups to immerse themselves in each other's challenges can help close apparent divides and offer valuable perspectives on potential solutions. First responders can attend meetings or other events to see how the data they provide is used to evaluate transportation safety. These shadowing activities are effective when they are consistent and include well-documented observations that are shared and discussed with staff and leadership from both agencies.

Provide opportunities for feedback between agencies. Another suggestion in the literature to address barriers in data sharing is to allow county staff to view first responder data and provide feedback about the types of elements that county staff would like to see based on their experiences with data analysis. This feedback can also include information received by the public. This strategy can be applied in real-time or deferred in meetings or workshops or by using feedback forms. For example, Lopez and Hemenway (2018) recommended bridging a divide between a bicycle advocacy group and the police by allowing the group to view police data on bicycle crashes and provide feedback about the types of elements that they would like to see based on their experiences cycling in Boston. This example demonstrates an effective approach to providing feedback, and although it does not include the county, counties can adopt this approach.

Host peer exchanges and workshops to promote information sharing and the development of best practices for coordination between counties and first responders. Peer exchanges/workshops can take a variety of formats. They can be local, attended by county staff, with peers from the state or around the country invited to share their approaches to transportation safety and coordination for this purpose. It can be statewide, with agencies from around the state attending to learn from their peers around the country. The peer exchange/workshop may be designed for attendance by county staff only, or as a joint event with county staff and first responders. The latter provides a more integrated event and further supports coordination between these different agency types. A joint peer exchange/workshop encourages buy-in from all involved agencies on any notable practices shared during the event. A well-established, regularly occurring schedule for the peer exchange/workshop (e.g., annually, biannually, biennially, etc.) can encourage attendance and help hosting agencies appropriately allocate resources. The peer exchanges or workshops can also be added to existing forums or events to encourage attendance and effectively use existing resources.

Provide transportation safety and/or data training with the county and local first responders. Training can be local, regional, or statewide. The literature cites the national Task Force on 21st Century Policing, which emphasizes the need for law enforcement personnel to have skills and knowledge in subjects related to the collection of policing data such as bias awareness and analytical research and technology. According to the literature, training builds agency capacity and identifies resources to overcome barriers.

Partner with the local university and identify other local agencies that can support activities related to transportation safety. Partnering with local universities offers agencies numerous potential benefits. First, local universities offer access to a pool of researchers and

experts who can conduct studies, experiments, and analyses related to transportation challenges offering insights and solutions that agencies might not access otherwise. Second, universities often have specialized equipment, labs, and resources that transportation agencies may not possess. Next, partnering with universities allows transportation agencies to tap into the talent pool of students. Also, universities can serve as testing grounds for new technologies or initiatives, allowing agencies to pilot and refine concepts before wide-scale implementation. These partnerships also facilitate professional development opportunities, offering tailored training programs to upskill agency personnel. Lastly, collaborative projects can attract funding, supporting groundbreaking transportation initiatives through grants and external investments.

Community organizations, non-profit organizations, and other agencies that are well-connected with local communities can provide meaningful information about the community members, safety concerns that may not be as easily recognized in the data, and resources needed to support a more robust approach to transportation safety. For example, county staff can learn more about the community's transportation concerns by occasionally attending local community meetings (e.g., church groups, homeowner associations, advocacy groups, special interest groups, etc.). When attending these meetings, it is important to set expectations, ensuring that attendees are aware that the goal of county staff attendance is to hear their stories and learn from their experiences and the outcome may not include the immediate identification and development of a strategy to address all of the identified needs.

Develop and implement a formal agreement between the county and first responders if they are under different governing bodies. Formal agreements have many benefits: they document the roles and responsibilities of all involved parties, describe the activities and the anticipated outcomes of the activities conducted through the agreement, and foster collaboration and communication between agencies. The agreements also serve as a contract, providing a certain level of predictability for the agency by ensuring compliance with the terms outlined in the agreement. A well-written agreement enables each agency to benefit from the collaboration, providing an opportunity for the creation of mutually agreed upon terms that serve the interests of all parties.

Identify a liaison in each agency. A liaison facilitates coordination and helps maintain the relationship between agencies. The benefits of liaisons include consistent communication with other agencies, improved information sharing within their own agency, and reduced duplication of effort by providing a single contact person. The liaison's responsibilities may include answering general questions, sharing information, and/or attending meetings of other agencies. The liaison may also serve as the organizer for the peer exchanges/workshops, trainings, or shadowing listed in this document. They may also be responsible for ensuring that the data sharing agreement is implemented as outlined in the document. The potential responsibilities of a liaison are vast, and agency budget, staff time availability for these activities, and agency goals should be considered to focus the list of responsibilities.

Create a County-focused action plan with participation from first responders. The action plan can be used to 1) identify and provide updated visualizations of the most dangerous areas for road users, 2) develop and execute enforcement countermeasures in designated "high injury" or "high crash" areas, 3) monitor trends and patterns, and 4) distribute information about trends (positive and negative) and patterns. An action team would have the responsibility of executing the plan as well as identifying strategies to effectively communicate trends and patterns with county staff, first responders, and the public.

Evaluate available budget and staff time to quantify resources that are available for transportation safety in general, as well as for coordination with first responder agencies on transportation safety. For any agency and at any level, coordination uses resources, therefore, most of the suggestions offered in this research, including the peer exchange, shadowing, training, partnering with other local agencies, having a liaison, and developing an action plan require adequate funding and staff time for successful implementation. Assessing costs, evaluating time budgets, identifying funding sources, and other activities can help the county identify where resources are available, where resources can be shifted, and if there is a need to seek out additional funding for these purposes.

Other suggestions derived from the literature include:

Find and use a committed champion that can encourage relationship building and data sharing. Involve this champion in coordination activities and decision-making processes.

Gain buy-in from agency leaders with authority to make policy and funding decisions.

Approval from these leaders ensures that the agency is able to advance actions to improve transportation safety and coordination practices with local first responders.

REFERENCES

- National Highway Traffic Safety Administration. (2023). "Traffic Safety Facts: 2021 Data". Washington, DC: US Department of Transportation.
<https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813509>.
- Florida Department of Highway Safety and Motor Vehicles. (2023). Florida Department of Highway Safety and Motor Vehicles Traffic Crash Reports: Crash Dashboard. FLHSMV:
<https://www.flhsmv.gov/traffic-crash-reports/crash-dashboard/>.
- Cherry, C., Hezaveh, A. M., Noltenius, M., Khattak, A., Merlin, L., Dumbaugh, E., and Sandt, L. (2018). Completing the Picture of Traffic Injuries: Understanding Data Needs and Opportunities for Road Safety. Knoxville, TN: Collaborative Science Center for Road Safety.
https://www.roadsafety.unc.edu/wp-content/uploads/2018/11/CSCRS_R4_FinalReport.pdf.
- Hosseinzadeh, A., Karimpour, A., Kluger, R., and Orthober, R. (2022). Data linkage for crash outcome assessment: Linking police-reported crashes, emergency response data, and trauma registry records. *Journal of Safety Research*, 81, 21-35.
- Watson, A., Watson, B., and Vallmuur, K. (2015). Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accident Analysis and Prevention*, 83, 18-25.
- Institute for Local Government. (2013). Partnerships Make Roadways Safer. https://www.calig.org/sites/main/files/file-attachments/partnerships_make_roadways_safer_feb_2013_0_0.pdf?1380725678.
- Federal Highway Administration. (2014). Collaboration in the Name of Roadway Safety. Washington DC: FHWA. <https://www.txltap.org/media/Bfing/Collaboration%2010.2014.pdf>.
- Lopez, D., and Hemenway, D. (2018). Generating a city's first report on bicyclist safety; lessons from the field. *Injury Prevention*, 24, 312-318.
<https://injuryprevention.bmjjournals.com/content/24/4/312>.
- Louisiana State University Center for Analytics & Research in Transportation. (2023). LA eCrash. CARTS: <https://carts.lsu.edu/ecrash>.
- SafeTREC. (2023). Transportation Mapping System (TIMS). Transportation Injury Mapping System: <https://tims.berkeley.edu/>.

Tarko, A. P., Romero, M., Lizarazo, C., and Pineda, R. (2020). Statistical Analysis of Safety Improvements and Integration into Project Design Process. Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2020/09.

University of Florida GeoPlan Center. (2022). Signal Four Analytics. UF GeoPlan Center:
<https://www.geoplan.ufl.edu/portfolio/s4/>.

Predicting Bicycle-Involved Crashes in the SCAG Region: A Machine Learning Analysis Using HSIS Data from California State

Ramina Javid¹; Eazaz Sadeghvaziri, Ph.D.²; Seyedmirsajad Mokhtarimousavi, Ph.D.³;
and Hananeh Omidi⁴

¹Ph.D. Candidate, Dept. of Transportation and Urban Infrastructure, Morgan State Univ. (corresponding author). Email: rajah1@morgan.edu

²Assistant Professor, Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Macon, GA. Email: sadeghvaziri_e@mercer.edu

³Postdoctoral Research Associate, Dept. of Civil and Environmental Engineering, Florida International Univ., Miami, FL. Email: smokh005@fiu.edu

⁴Ph.D. Student, Graduate Researcher, and Teaching Assistant, Dept. of Geography and Environmental Sustainability, Sarkeys Energy Center, Univ. of Oklahoma, Norman, OK. Email: hana.omidi@ou.edu

ABSTRACT

This study investigates factors influencing the severity of bicycle-involved crashes in the US, particularly the Southern California Association of Governments (SCAG) region from 2013 to 2017, using Highway Safety Information System (HSIS) data. Employing a Bayesian network model with rigorous validation, the study yields a low error rate, emphasizing its effectiveness in analyzing crash data and enhancing rider safety insights. Two scenarios explore variables affecting the probability of fatal crashes, revealing the positive impact of proper lighting and surface on visibility. Additionally, the need for infrastructure capable of handling wet surfaces and providing adequate drainage is underscored. The results demonstrate the importance of effective infrastructure design, emphasizing proper lighting and visibility for cyclists to mitigate fatal crash risks. The study's implications extend to informing policymakers and transportation engineers on prioritizing safety measures, showcasing the Bayesian network model's efficacy in identifying critical factors in bicycle-involved crashes.

INTRODUCTION

Bicycle-involved crashes have become a major concern for public safety in many cities. Despite numerous efforts to increase bicycle riders' safety, bicycle crashes and fatalities continue to rise in many regions. Every year, road crashes in the United States result in over 130,000 injuries and close to 1,000 cyclist fatalities (CDC 2023; Noland and Quddus 2004). The Southern California Association of Governments (SCAG) region has experienced increased bicycle crashes, highlighting the need for further research into the factors affecting bicycle-involved crashes in this region.

This study aims to investigate the factors that affect bicycle-involved crashes in the SCAG region using the Highway Safety Information System (HSIS) data from the state of California and the SCAG region. This study uses a machine learning approach to analyze the HSIS data and identify the factors that contribute to bicycle-involved crashes. Machine learning algorithms have been used in various studies to identify the factors affecting traffic crashes, including bicycle-involved crashes (Lord and Mannering 2010; Lu et al. 2022). Previous studies have identified

various factors that contribute to bicycle-involved crashes, including human, environmental, and infrastructure factors. Human factors include the behavior and characteristics of bicyclists, motorists, and pedestrians. Environmental factors include weather, lighting, and road conditions. Infrastructure factors include roadway design, traffic control devices, and the presence of bicycle facilities (Lord and Mannering 2010; Shah et al. 2021; Chen 2015; Guo et al. 2017). Studies have shown that different regions may have unique factors contributing to bicycle-involved crashes.

Although an extensive literature review focuses on factors affecting road crashes, studying the factors affecting bicycle-involved crashes is still in its early stages. Therefore, this study aims to identify the specific factors contributing to bicycle-involved crashes in the SCAG region. A Bayesian network model was developed to analyze the influence of factors affecting bicycle-involved crashes using HSIS data to reach this goal. HSIS data is a comprehensive and reliable database that provides detailed information on traffic crashes. The HSIS data contains information on various factors, including crash characteristics, roadway characteristics, environmental conditions, and vehicle and driver characteristics.

LITERATURE REVIEW

Bicycling is a popular and environmentally friendly mode of transportation. Commuting on a bicycle is the third most utilized U.S. transportation mode, quickly gaining popularity as a commuting option. Commuters biking to work have increased by 65% nationwide from 2000 to 2019. However, crashes involving bicycles continue to be a significant global public safety issue. For effective strategies to be developed to reduce their occurrence, it is essential to comprehend the factors that contribute to these crashes because they have significant economic and social costs (Dash, Abkowitz, and Philip 2022; Javid and Sadeghvaziri 2023a).

Roadway design and infrastructure have been identified as significant factors in bicycle-involved crashes. The design of roads, including the number and width of lanes, the presence of bike lanes, and the types of intersections, can impact the likelihood of crashes. Research has shown that bicycle lanes and separate bike paths can reduce the risk of crashes by providing a safe and dedicated space for cyclists. Additionally, roundabouts and traffic calming measures, such as speed humps and chicanes, have reduced the incidence and severity of bicycle-involved crashes (Rothenberg, Goodman, and Sundstrom 2016; Wu et al. 2019). Another important factor in accidents involving bicycles is the presence of motor vehicles. The risk of collisions can be significantly increased by drivers' actions, such as speeding, distracted driving, and failing to yield to cyclists. According to previous studies, motor vehicles are usually involved in bicycle crashes, and the severity of the crash is frequently correlated with the vehicle's speed. A vehicle's size and type can also affect how severe a collision is, with bigger, heavier vehicles posing a greater risk to cyclists (Noland and Quddus 2004).

Environmental factors such as the weather and lighting can also play a role in bicycle-related crashes. Poor lighting and visibility can raise the risk of crashes, particularly during low-light conditions. By reducing traction and visibility on the roads, bad weather like rain, snow, and ice can also raise the risk of a crash (Retting, Ferguson, and McCartt 2003). Due to the kinetic energy generated during collisions between two masses with differing velocities and masses, cyclists are considered among the most vulnerable road users in mixed traffic. When a motor vehicle collides with a cyclist at speeds exceeding 20 miles per hour, the risk of severe injury or death is increased (Jacobsen and Rutter 2012; Jurewicz et al. 2016). This highlights the potential

danger that heavily trafficked urban corridors can pose to cyclists without adequate safety measures (Board, n.d.).

Researchers worldwide have conducted studies on bicycle crashes often focusing on specific areas or regions to pinpoint and address safety concerns in bicycle infrastructure and operations. These studies commonly utilize various modeling techniques such as the Poisson distribution, negative binomial models, linear regression models, logit models, ordered probit models, and multivariable logistic regression (Dash, Abkowitz, and Philip 2022).

Overall, bicycle-involved crashes are a significant public safety concern. Understanding the various factors that contribute to these crashes is crucial for developing effective strategies to reduce their occurrence. Roadway design and infrastructure, motor vehicle traffic, cyclist behavior, environmental factors, and demographic factors have all been found to contribute to bicycle-involved crashes. Addressing these factors through a combination of engineering, education, and enforcement strategies can help to improve the safety of cyclists on the roadways. While previous studies have made significant contributions to the study of factors impacting crash severity, there is still a need to investigate the factors impacting bicycle-involved crashes using HSIS data.

DATA AND METHODOLOGY

Data

The bicycle-involved crash data utilized in this analysis was obtained from the Highway Safety Information System (HSIS) from 2013 to 2017, covering the entire state of California. HSIS is a system that provides high-quality traffic data (i.e., crashes, and road characteristics). The HSIS dataset consists of multiple file groups, such as accident, curve, grade, roadway inventory, and vehicle files. In this dataset, a bicycle-involved crash is defined as whether any bicycle was involved in the crash. For the purpose of this study, the accident file was used.

Moreover, this study focuses on bicycle-involved crashes in the SCAG region. The SCAG is the nation's largest Metropolitan Planning Organization (MPO) and Council of Governments (COG), encompassing six counties, six county transportation commissions, 191 cities, and the region's Native American Tribes. It has a population of over 19 million in an area covering more than 38,000 square miles ("SCAG" 2020). Figure 1 shows the SCAG region within the boundary of the state of California. First, the "bike_flg" variable, which shows whether any bicycle was involved in the accident, was used to filter all bicycle-related crashes. Also, in order to investigate the bicycle-involved crashes in the SCAG region, ArcGIS Pro was utilized. Using HSIS shapefiles, all the bicycle-involved crashes were filtered based on the SCAG region. The shapefile for the SCAG region boundary was obtained from the SCAG Regional Data Platform ("RDP" 2022).

Figure 2 shows all the bicycle-involved crashes between 2013 to 2017 after the process mentioned above. During this period, 2,065 bicycle-involved crashes were recorded across the SCAG region. According to Table 1, more than 40% of bicycle-involved crashes in California were within the SCAG region.

Figure 3 shows the types of bicycle-involved crashes in the SCAG region. As demonstrated in this figure, 1.79% of all crashes were fatality crashes. To better reach the goals of this study, bicycle-involved crashes were aggregated into four categories: Fatal, Injury, and No Injury. Figure 4 shows the new aggregated data. It can be understood from the data that more than 90% of bicycle-involved crashes result in injury.



Figure 1. SCAG Region within the State of California

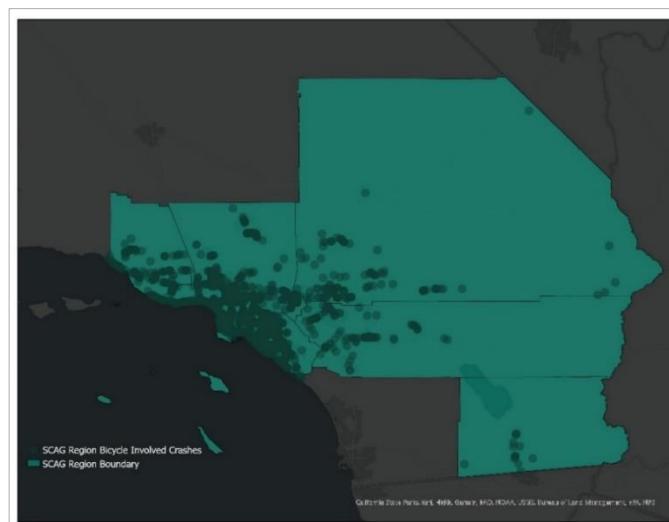


Figure 2. SCAG Region Bicycle Involved-Crashes from 2013-2017

Table 1. Bicycle-involved Crashes in California and SCAG Region

Year	Number of Bicycle-involved Crashes	
	California	SCAG Region
2013	1,188	494
2014	985	401
2015	1,010	399
2016	1,035	415
2017	886	356
Total	5,104	2,065

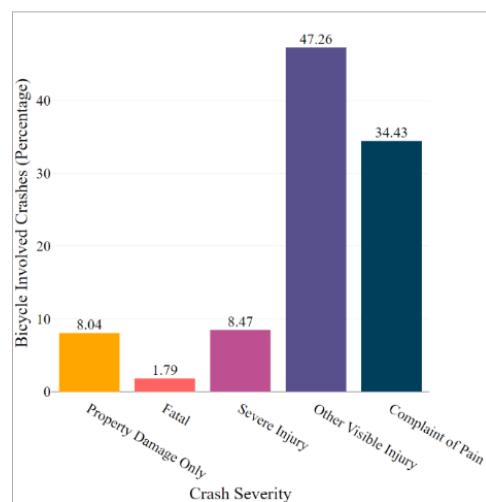


Figure 3. SCAG Bicycle-involved Crash Types in the SCAG Region;

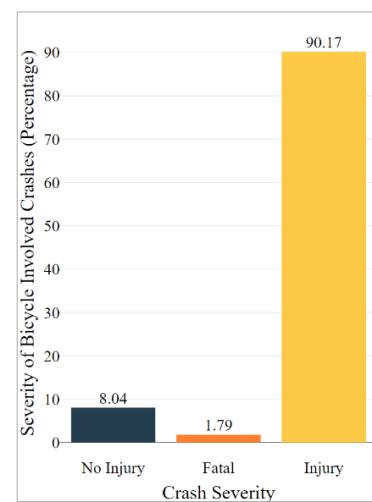


Figure 4. Severity of Bicycle-involved Crashes in the SCAG Region

Figure 5 shows the distribution of bike crashes by time of day, where times have been grouped into the following categories: (1) Early Morning (midnight-5:00 a.m.); (2) Morning (5:00 a.m.-9:00 a.m.); (3) Peak AM (9:00 a.m.-1:00 p.m.); (4) Afternoon (1:00 p.m.-5:00 p.m.); (5) Peak PM (5:00 p.m.-9:00 p.m.); and (6) Late Evening (9:00 p.m.-midnight). The frequency of bicycle crashes tends to increase as the day progresses, and it is at the highest point in the afternoon.

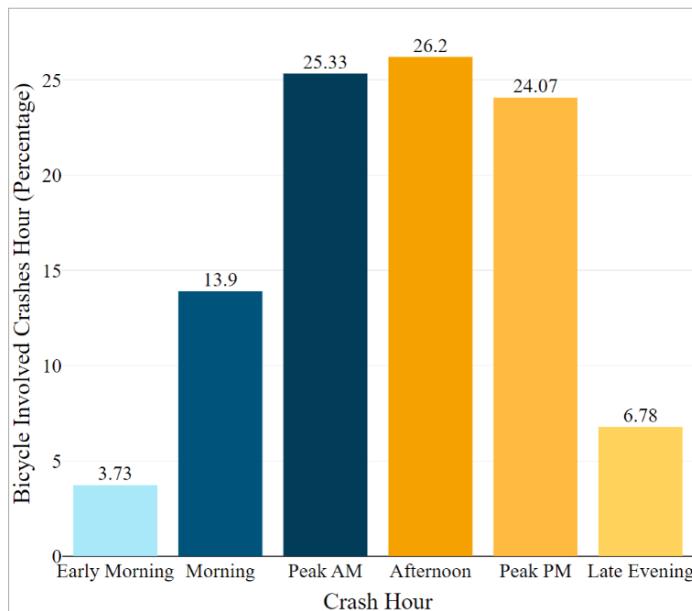


Figure 5. Bicycle-Involved Crashes Hours in SCAG Region

METHODOLOGY

Many studies use statistical models to develop policies to improve traffic safety, investigate and forecast travel behavior, and pinpoint deficiencies in transportation policy (Javid and Sadeghvaziri 2023b; Mokhtarimousavi et al. 2021; Sadeghvaziri et al. 2023; Baniasadi et al. 2022; Javid, Sadeghvaziri, and Jeihani 2023). Therefore, to reach the goal of the study, the methodology used in this study is based on the Bayesian network. The Artificial Intelligence Community introduced the Bayesian network during the 1980s. Despite its importance in modeling extremely unpredictable scenarios, the technique was not widely used until recently, aided by the emergence of new algorithms, software packages, and computer capabilities. Bayesian networks are now increasingly being used to model problems to support experts as they conduct their research. In the transportation field, Bayesian networks have been applied in predicting the probability of a road crash (HOSSAIN and MUROMACHI, n.d.), traffic management (Derbel and Boujelbene 2015), speed compliance (Ardeshtiri and Jeihani 2014), traffic flow forecasting (Sun et al. 2004), etc.

A Bayesian network usually reveals how a complicated system works by looking at the causal link between the influential variables. The variables and their relationships are represented by nodes and links in a Bayesian network graph. The influence of each variable on the final node and the correlations among explanatory variables are crucial elements (Ardeshtiri and Jeihani 2014). In this work, the initial Bayesian network was developed using the HSIS data. Hence, in

this study, a Bayesian network was developed to model the influence of different factors on bicycle-involved crashes. Several explanatory variables were used to build the Bayesian network construction. These variables represented hours when the crash occurred, the type/level of light that existed at the time of the crash, the location type of the crash, the condition of the road surface where the crash occurred, and the day of the week when the accident occurred. Additionally, the endpoint is having a severe injury (or fatality) and no injury in a bicycle-involved crash. Table 2 shows the variables applied in this research and their values.

Table 2. Variables Applied in this Research and Their Value

Variables	Symbols	Values
Severity type of the crash	SVR	- No Injury - Injury - Fatal
Hour when the crash occurred	HR	- Early Morning - Morning - Peak AM - Afternoon - Peak PM - Late Evening
The type/level of light that existed at the time of the crash	LGT	- Daylight - Dusk Dawn - Dark - Other
Location type of the crash	LOC	- Highway - Intersection - Ramp or Collector - Not State Highway
The condition of the road surface where the crash occurred	SRF	- Dry - Wet - Snowy, Icy - Other
Day of the week when the accident occurred	WKD	- Weekday - Weekend

In order to develop the Bayesian network graph, the authors used Netica software (“Norsys Software Corp. - Bayes Net Software” n.d.). The finished Bayesian network’s construction is shown in Figure 6, and each node’s conditional probability tables reflect the node’s states as belief bars.

ANALYSIS RESULTS

The Bayesian network model (Figure 6) demonstrates that the severity of bicycle-involved crashes depends on many variables. The model shows the percentage probability of each variable. Based on the data, it can be concluded that when a bicycle-involved crash occurs, there is an 11.1% probability of the death of the bicycle rider.

Bayesian Network Validation

Table 3 presents the results of the network validation. Performance evaluation ensures that a predictive model performs appropriately. The error rate indicates the percentage of false predictions. This is the proportion of cases where the prediction is wrong, which in our case is 9.5%. The logarithmic loss and quadratic loss vary from zero to infinity, with zero being the best fit. The result from our study is 0.35. Spherical payoff varies from zero to one, in which case one indicates the best fit (Ardeshtiri and Jeihani 2014). In this study, the spherical payoff is 0.9. Hence, the validation result shows that there is a strong consistency validation and no evidence of overfitting in the model. Additionally, the results assert that the Bayesian network model could be used for sensitivity analysis of the different variables.

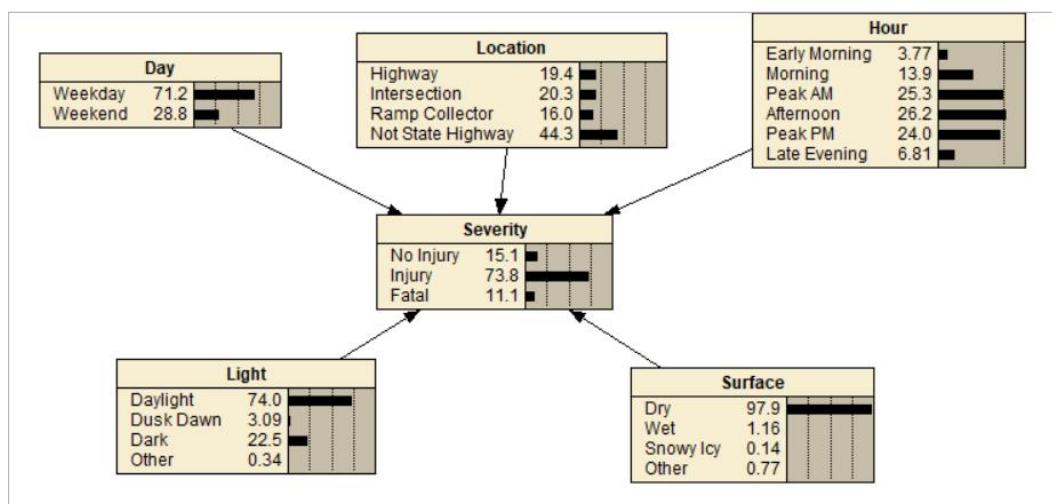


Figure 6. Bayesian Network Model of the Study

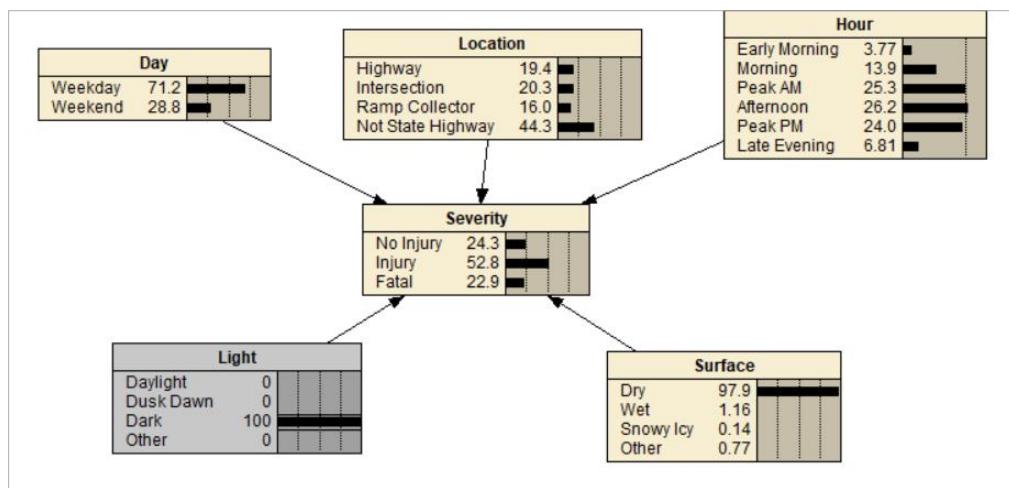
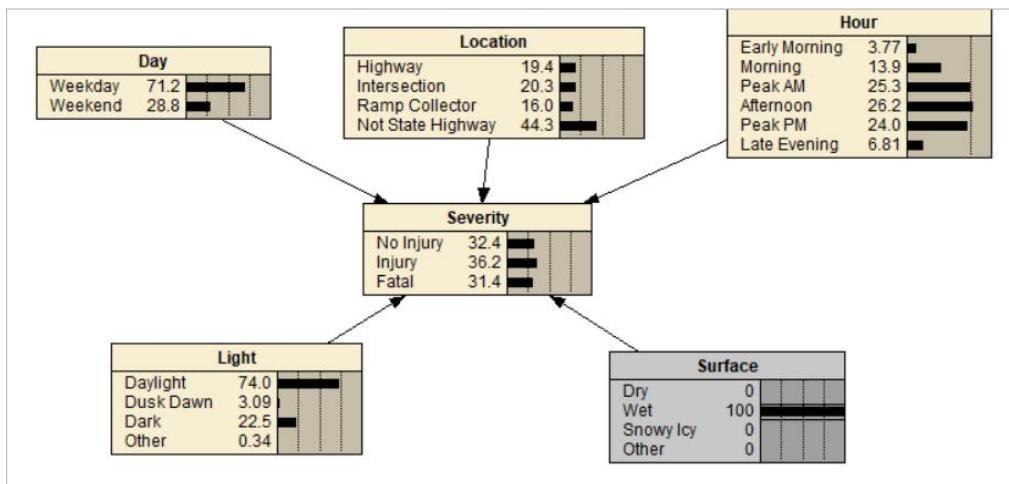
Table 3. Bayesian Network Validation

Error rate	9.54%
Logarithmic loss	0.3507
Spherical payoff	0.9066

Sensitivity Analysis

Sensitivity analysis is a technique that can help validate the probability parameters of a Bayesian network. This is done by investigating the effect of small changes in the model's numerical parameters ("BayesFusion" 2022). A Bayesian network can identify how different factors can impact bicycle-involved crashes, and how it can increase the safety of riders. For this study, two scenarios were developed. Figure 7, which shows the first scenario, indicates that if all bicycle-involved crashes occurred when it was dark outside, the probability of having a fatal crash would increase by 11%.

Figure 8, which shows the second scenario, indicates that if all bicycle-involved crashes were taking place on a wet surface, the probability of having a fatal crash would increase by 20%.

**Figure 7. Scenario One****Figure 8. Scenario Two**

DISCUSSION AND SUMMARY

This study used HSIS data to investigate the factors affecting the severity of bicycle-involved crashes in the SCAG region from 2013 to 2017. The Bayesian network model that was developed for the purpose of the study has a firm consistency validation. The model's error rate is 9.5%, indicating that the proportion of cases where the prediction is wrong is relatively low. The study also developed two scenarios to understand the impact of different variables on the probability of having a fatal crash. The first scenario shows that if all bicycle-involved crashes occurred when it was dark, the probability of having a fatal crash would increase by 11%. This highlights the importance of adequately lighting bicycles and wearing reflective clothing to increase visibility. The second scenario shows that if all bicycle-involved crashes occurred on a wet surface, the probability of having a fatal crash would increase by 20%. This emphasizes the need for designing roads and infrastructure that can handle wet conditions and provide adequate drainage.

The Bayesian network model developed in this study shows promising results in predicting the probability of bicycle-involved crashes and identifying the factors that contribute to them. The model's strong consistency validation and low error rate indicate its effectiveness in analyzing crash data and providing valuable insight into improving rider safety. The scenarios developed in the study also provide useful information on the impact of different variables on the probability of having a fatal crash. The results emphasize the importance of proper lighting and visibility on bicycles and the need for designing roads and infrastructure that can handle wet conditions. The sensitivity analysis conducted in the study further validates the effectiveness of Bayesian network models in identifying the most significant factors contributing to bicycle-involved crashes. The analysis results can help policymakers and transportation engineers prioritize measures to improve the safety of riders.

The findings indicate that improving infrastructure, enhancing visibility through proper lighting, and considering weather conditions in road design are crucial for reducing the severity of bicycle-involved crashes. The study's approach demonstrates the effectiveness of machine learning techniques, specifically Bayesian network models, in identifying and understanding the most significant factors contributing to bicycle-involved crashes. This knowledge can aid in the development of tailored strategies and safety measures to protect bicycle riders and improve overall road safety in the SCAG region and potentially in other areas with similar characteristics. Overall, the findings of this study provide valuable insight into improving the safety of riders and highlight the effectiveness of Bayesian network models in analyzing and predicting bicycle-involved crashes. Further research can focus on expanding the model to include more variables and validating the results in different geographical locations to develop tailored solutions for improving rider safety.

REFERENCES

- Ardeshiri, A., and M. Jeihani. 2014. "A Speed Limit Compliance Model for Dynamic Speed Display Sign." *Journal of Safety Research* 51: 33–40.
- Baniasadi, S., O. Rostami, D. Martín, and M. Kaveh. 2022. "A Novel Deep Supervised Learning-Based Approach for Intrusion Detection in IoT Systems." *Sensors* 22 (12): 4459.
- BayesFusion. 2022. Using GeNIe > Bayesian Networks > Sensitivity Analysis in Bayesian Networks. 2022. https://support.bayesfusion.com/docs/GeNIe/bn_sensitivitybn.html.
- Board, Safety. n.d. "Safety Research Report."
- CDC. 2023. "Bicycle Safety | Transportation Safety | Injury Center | CDC." 2023. <https://www.cdc.gov/transportationsafety/bicycle/index.html>.
- Chen, P. 2015. "Built Environment Factors in Explaining the Automobile-Involved Bicycle Crash Frequencies: A Spatial Statistic Approach." *Safety Science* 79: 336–43.
- Dash, I., M. Abkowitz, and C. Philip. 2022. "Factors Impacting Bike Crash Severity in Urban Areas." *Journal of Safety Research* 83: 128–38.
- Derbel, A., and Y. Boujelbene. 2015. "Bayesian Network for Traffic Management Application: Estimated the Travel Time." In *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, 1–6. IEEE.
- Guo, Y., J. Zhou, Y. Wu, and J. Chen. 2017. "Evaluation of Factors Affecting E-Bike Involved Crash and e-Bike License Plate Use in China Using a Bivariate Probit Model." *Journal of Advanced Transportation* 2017.

- Hossain, M., and Y. Muromachi. n.d. *Applicability of Bayesian network in transportation engineering for safety.*
- Jacobsen, P. L., and H. Rutter. 2012. *Cycling Safety*. Vol. 2007. MIT Press Cambridge, MA.
- Javid, R., and E. Sadeghvaziri. 2023a. "Assessing Gender Gaps in Citi Bike Ridership in New York City." Available at SSRN 4358660.
- Javid, R., and E. Sadeghvaziri. 2023b. "Equity Analysis of Bikeshare Access: A Case Study of New York City." *Findings*.
- Javid, R., E. Sadeghvaziri, and M. Jeihani. 2023. "Development and Evaluation of a Bayesian Network Model for Preventing Distracted Driving." *IATSS Research* 47 (4): 491–98.
- Jurewicz, C., A. Sobhani, J. Woolley, J. Dutschke, and B. Corben. 2016. "Exploration of Vehicle Impact Speed–Injury Severity Relationships for Application in Safer Road Design." *Transportation Research Procedia* 14: 4247–56.
- Lord, D., and F. Mannering. 2010. "The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives." *Transportation Research Part A: Policy and Practice* 44 (5): 291–305.
- Lu, W., J. Liu, X. Fu, J. Yang, and S. Jones. 2022. "Integrating Machine Learning into Path Analysis for Quantifying Behavioral Pathways in Bicycle-Motor Vehicle Crashes." *Accident Analysis & Prevention* 168: 106622.
- Mokhtarimousavi, S., M. Hadi, E. Sadeghvaziri, and A. Azizinamini. 2021. "Evolutionary Training Approaches for Machine Learning Models for Analyzing Classification Imbalanced Crash Datasets."
- Noland, R. B., and M. A. Quddus. 2004. "A Spatially Disaggregate Analysis of Road Casualties in England." *Accident Analysis & Prevention* 36 (6): 973–84.
- "Norsys Software Corp. - Bayes Net Software." n.d. Accessed May 17, 2022.
<https://www.norsys.com/>.
- "RDP." 2022. <Https://Hub.Scag.ca.Gov>. 2022. <https://hub.scag.ca.gov/>.
- Retting, R. A., S. A. Ferguson, and A. T. McCartt. 2003. "A Review of Evidence-Based Traffic Engineering Measures Designed to Reduce Pedestrian–Motor Vehicle Crashes." *American Journal of Public Health* 93 (9): 1456–63.
- Rothenberg, H., D. Goodman, and C. Sundstrom. 2016. *Separated Bike Lane Crash Analysis*.
- Sadeghvaziri, E., R. Javid, M. Jeihani, and Equity Center. 2023. *Investigating Walking and Biking Activities Among Low-Income African Americans*.
- SCAG. 2020. Southern California Association of Governments. 2020. <https://scag.ca.gov/ready-2020>.
- Shah, N. R., S. Aryal, Y. Wen, and C. R. Cherry. 2021. "Comparison of Motor Vehicle-Involved e-Scooter and Bicycle Crashes Using Standardized Crash Typology." *Journal of Safety Research* 77: 217–28.
- Sun, S., C. Zhang, G. Yu, N. Lu, and F. Xiao. 2004. "Bayesian Network Methods for Traffic Flow Forecasting with Incomplete Data." In *European Conference on Machine Learning*, 419–28. Springer.
- Wu, Y., M. Abdel-Aty, O. Zheng, Q. Cai, and L. Yue. 2019. "Developing a Crash Warning System for the Bike Lane Area at Intersections with Connected Vehicle Technology." *Transportation Research Record* 2673 (4): 47–58.

Evaluating Factors of Interstate Crashes in New Jersey

Fahmida Rahman, Ph.D.¹

¹Dept. of Civil and Environmental Engineering, Rowan Univ., Glassboro, NJ.
Email: rahmanf@rowan.edu

ABSTRACT

Interstates in New Jersey experience a high volume of traffic, and unfortunately, around 15,000 crashes occur on these roads every year. Statistics show that 21% of these crashes result in fatalities or severe injuries. Therefore, it is critical to understand the factors contributing to this high number of crashes. This study aims to identify these factors through the application of statistical techniques. Data was collected from the heavily traveled interstate corridors. Utilizing a negative binomial model, we analyzed the total number of crashes in relation to traffic, roadway geometry, and operational conditions. As expected, factors associated with exposure, such as traffic volume and segment length, were significant. As these factors increase, crashes tend to increase. Additionally, interstates with concrete barriers exhibited fewer crashes compared to those with curbed barriers, indicating the potential benefits of concrete barriers in reducing crashes. Interestingly, our study uncovered an aspect that is overlooked in the widely utilized safety assessment guidebook, the *Highway Safety Manual*—the impact of speed limits on crash frequency. Through a variable prioritization process, we identified speed limit as the second most important factor. Further exploration of the dataset revealed that crashes were more prevalent when the speed limit decreased from the previous segment on a specific interstate route. Such findings have practical implications for transportation agencies and policymakers. The insights gained can be instrumental in devising effective countermeasures, such as leveraging intelligent transportation systems (ITS) to provide advance signs for speed limit changes. Implementing such measures can help mitigate crashes, particularly those associated with speed limit alterations. Overall, this study identified the factors contributing to the high crashes in New Jersey's interstates. Understanding these factors can aid in formulating targeted strategies to improve the safety of these roads.

INTRODUCTION

The Highway Safety Manual (HSM) offers transferable Safety Performance Functions (SPFs) for predicting crashes, which can be adjusted to suit local conditions (HSM 2010). The New Jersey Department of Transportation (NJDOT) has already calibrated these SPFs for rural two-lane highways, rural multilane highways, and urban and suburban roads. They also developed New Jersey-specific prediction models for these facility types. Nonetheless, there is a requirement to develop such predictive models for New Jersey interstates while gaining an understanding of the contributing factors, for example, roadway geometry and operating conditions.

This paper aims to investigate how different roadway conditions affect crash frequency on New Jersey interstates and incorporate them in developing a crash prediction model specific to New Jersey interstates. Roadway conditions considered in this study included annual average daily traffic (AADT), section length, pavement width, number of lanes, shoulder width, median width, median type, and speed limit. After identifying the roadway conditions that significantly

affect the crash frequency, this study incorporated them into developing a crash prediction model for interstates of New Jersey. It is important to note that while the general approach of HSM does not include speed limit in the crash prediction model, this study considers speed limit along with other roadway conditions to evaluate its influence on interstate crashes.

LITERATURE REVIEW

HSM uses SPFs derived from observed crash data of similar facility types to predict average crash frequencies under base conditions (HSM 2010). To adapt these predictions to specific sites, one of the strategies is to utilize crash modification factors (CMFs). CMFs are adjusted for variations in geometric design and traffic control features, enabling a more accurate estimation of crash likelihood at a particular location. Another strategy is to develop jurisdiction-specific crash prediction models that may include variables outside of HSM, depending on the location.

Besides HSM, research has been conducted to develop crash prediction models for different segment types of the freeway facility. For example, NCHRP Project 17-45 developed a crash prediction method tailored to freeways, suitable for inclusion in the HSM (2021). It covers crash factors related to freeway segments and speed-change lanes, including factors like horizontal curvature, lane width, shoulder width, median width, barrier length, ramp-related lane changes, rumble strip presence, clear zone width, and recurring congestion extent. The NCHRP Project 17-89A developed a methodology to predict crash occurrence specifically for directional freeway segments featuring High Occupancy lanes, utilizing data gathered from California and Washington (2023).

In the past, NJDOT calibrated SPFs based on New Jersey's roadway features, traffic volumes, and crash data (Ozbay et al. 2019). They also developed new SPFs to better represent the state's conditions. Their research encompassed segments and intersections of various road types—rural two-lane two-way, rural multilane, and urban/suburban roads. During the analysis, they pinpointed areas in the datasets where improvements in data processing could improve data quality or efficiency. This effort aimed not only to create New Jersey-specific SPFs but also to enhance data usability for the stated purpose. However, one of the roadway types that was not considered in their analysis was interstates.

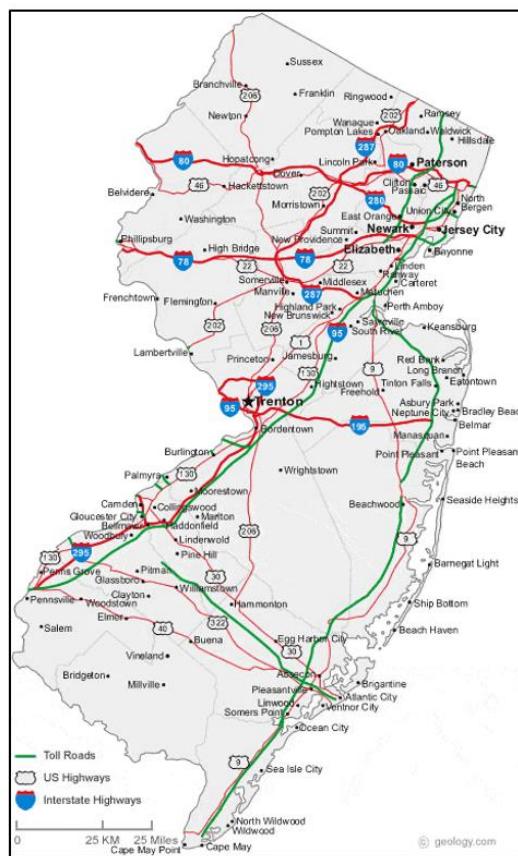
Several states have been practicing the calibration of HSM default SPFs based on the data specific to their jurisdictions. For example, Maryland and Missouri calibrated the HSM method for their interstate facilities (Shin et al. 2016, Sun et al. 2013). Considering the performance, the calibrated SPFs provided better accuracy compared to HSM default SPFs. Other states, such as Florida and Kansas, developed their own jurisdiction-specific SPFs for the interstates (Lu et al. 2014, Matarage and Dissanayake 2020). While comparing against the calibrated SPFs from HSM, the jurisdiction-specific SPFs fitted the local data better. These studies also established the fact that there is a need for developing jurisdiction-specific SPFs for interstate facilities.

In summary, existing studies show the importance of developing jurisdiction-specific SPFs for predicting crashes on interstates. In case of the interstates in New Jersey, such study still needs to be done. Therefore, this study attempts to develop crash prediction models specific to the interstates of New Jersey while considering different roadway conditions.

DATA COLLECTION AND PROCESSING

For this study, crash data were collected from the NJDOT between 2016-2020. The dataset included all the interstate highways in New Jersey, as shown in Figure 1. Roadway inventories

such as curb type, number of lanes, pavement width, median type, median width, shoulder width, AADT, section length, and speed limit were extracted from NJDOT's Straight Line Diagrams (SLD). These roadway attributes were conflated with the crash dataset. The conflated dataset was cleaned up by removing segments with no AADT values and a section length of less than 0.1 miles. All these processes resulted in a final dataset with 540 segments for model development. This sample size meets the HSM recommendation of using 30 to 50 sites for developing a crash prediction model.



crash frequency, especially when the counts are overdispersed. In a negative binomial regression, the expected crash frequency, μ_i , is determined as a function of the explanatory variables shown in Equation 1. An error term, ε_i , is also added to this equation to overcome the overdispersion issue.

$$\mu_i = \exp(\beta x_i + \varepsilon_i) \quad (1)$$

Where, x_i is the vector of explanatory variables presented in Table 1, and β is the vector of estimated coefficients.

Table 1. Descriptive Statistics of the Variables

	Mean	Standard Deviation	Minimum	Maximum
AADT	83040.8	33726.6	23229.0	166551.0
Section Length	0.7	0.6	0.1	4.3
Pavement Width	80.1	22.1	47.0	144.0
Number of Lanes	6.6	1.8	4.0	11.0
Shoulder Width	9.8	4.5	0.0	17.0
Median Width	26.9	20.4	2.0	70.0
Median Type*			0	2
Speed Limit	62.3	4.7	50.0	65.0
Crash Frequency	134.2	150.8	0.0	853.0

*Categorical variable

Correlation Coefficients

In a regression analysis, multicollinearity occurs when explanatory variables are highly correlated with each other. One common way to identify multicollinearity is by examining the correlation coefficients between the explanatory variables. Generally, a correlation coefficient of higher than 0.6 is considered an indicator of a high correlation between two explanatory variables (Li, Abdel-Aty and Yuan 2020). We followed this threshold to identify the multicollinearity during the analysis.

Goodness of Fit Measure

To assess the goodness of fit of a model, we employed several metrics: AICc, BIC, Root Mean Squared Error (RMSE), and R². AICc (Akaike Information Criterion corrected) and BIC (Bayesian Information Criterion) serve as selection criteria, gauging model complexity. Lower values indicate less complexity, with AICc typically preferred for smaller sample sizes due to its milder penalty on model complexity compared to BIC. RMSE is a statistical measure assessing model accuracy by quantifying the difference between predicted and actual values, where lower values indicate better accuracy. Additionally, R² was utilized to explain the proportion of variance in the response variable explained by the explanatory variables in the model. Ranging

between 0 and 1, an R^2 value closer to 1 signifies a more extensive explanation of variance in the response variable by the explanatory variables, implying a stronger fit of the model to the data.

To further assess the goodness of fit, we also utilized Cumulative Residual (CURE) plots. These plots illustrate the differences between predicted crash frequencies and actual data (Lyon, Persaud and Gross 2016). Analyzing residuals allows for an examination of model fitting. A model is deemed a good fit when its residuals are predominantly close to zero. Utilizing the CURE plot as a visualization tool aids in assessing SPF predictions based on the individual explanatory variables within the model. For an unbiased SPF, the CURE plot must fall within the confines of two standard deviations.

MODEL DEVELOPMENT AND RESULTS

Prior to integrating the explanatory variables into the model, a thorough investigation into multicollinearity was conducted. This analysis involved plotting a correlation coefficient matrix, presented in Figure 2, to analyze the relationships among the variables. Notably, the analysis revealed a strong correlation between pavement width and the number of lanes. Both of these factors also exhibited a noteworthy correlation, surpassing the 0.6 threshold, denoting a high correlation with AADT. Consequently, when experimenting with various models incorporating different sets of variables, we excluded those presenting multicollinearity issues.

	AADT	Total Crashes	Section Length	Pavement Width	Lanes	Shoulder Width	Median Width	Speed Limit
AADT	1	0.2123	-0.1778	0.6344	0.6387	0.2922	-0.3763	-0.1562
Total Crashes	0.2123	1	0.4638	0.1241	0.1255	0.0168	-0.0622	-0.2304
Section Length	-0.1778	0.4638	1	-0.207	-0.211	-0.124	0.2389	0.2046
Pavement Width	0.6344	0.1241	-0.207	1	0.9818	0.2856	-0.5106	-0.1869
Lanes	0.6387	0.1255	-0.2114	0.9818	1	0.2658	-0.5128	-0.1861
Shoulder Width	0.2922	0.0168	-0.124	0.2856	0.2658	1	-0.2866	-0.0783
Median Width	-0.3763	-0.0622	0.2389	-0.5106	-0.513	-0.2866	1	0.3662
Speed Limit	-0.1562	-0.2304	0.2046	-0.1869	-0.186	-0.0783	0.3662	1

Figure 2. Correlation Coefficient Matrix

After identifying the variables with multicollinearity issues, we finally came up with three separate model forms, as listed below:

- Model 1: with speed limit, median type, section length, and AADT
- Model 2: with speed limit, median type, section length, and number of lanes
- Model 2: with speed limit, median type, section length, and pavement width

Each model was developed following the negative binomial regression method. Table 2 presents all the three models tested in this study. Only the variables that were found significant at a 5% confidence interval are included in the table. The performance of these models was evaluated using the goodness of fit measures discussed previously. Table 3 shows the performance comparisons of these models.

Table 3 reveals that among the models evaluated, Model 3 exhibited the poorest performance across all measures of goodness of fit. Specifically considering AICc, BIC, and R² values, Model 1 appears to be better in comparison to Model 2. Moreover, to gain additional insights, CURE plots were generated for each model, plotting CURE against section length and AADT.

These plots, shown in Figure 3, provided valuable information. Notably, the cumulative residuals depicted in these plots for Model 1 align well within the range of two standard deviations. This highlights the robustness and favorable performance of Model 1, further confirming its superiority among the evaluated models.

Table 2. Models for New Jersey Interstates

	Model 1		Model 2		Model 3	
	coefficient	p-value	coefficient	p-value	coefficient	p-value
Intercept	1.1193739	0.2403	8.27904	<.0001	7.019414	<.0001
Speed Limit	-0.079452	<.0001	-0.07292	<.0001	-0.07227	<.0001
Median Type [Positive-Unprotected]	0.3636092	0.0011	0.570922	<.0001	0.609479	<.0001
Median Type [Curbed-Unprotected]	0.5455264	0.0027	0.651401	0.001	0.720354	0.0009
Log [Section Length]	0.941115	<.0001	0.935469	<.0001	1.121343	<.0001
Log[AADT]	0.7778348	<.0001	-	-	-	-
Number of Lanes	-	-	0.152329	<.0001	-	-
Pavement Width	-	-	-	-	0.010079	<.0001

Note: All the variables are significant at a 5% significance level

Table 3. Comparison of Model Performance

Models	AICc	BIC	RSME	R ²
Model 1	5885.74	5915.55	106.27	0.5
Model 2	5931.43	5961.25	106.27	0.46
Model 3	6056.99	6099.47	319.42	0.118

In this study, Model 1 was selected for detailed analysis, and its formulation is presented in Equation 2. Notably, both AADT and section length demonstrated a positive correlation with crash frequency, which aligns with their roles as exposure variables. Exploring further, the median type revealed an intriguing insight: segments with a positive barrier (like a concrete barrier) tended to experience fewer crashes compared to those with a curbed barrier. Interestingly, a negative relationship was observed between crash frequency and speed limit. Further examination of the dataset unveiled a noteworthy trend: a higher prevalence of crashes occurred when the speed limit decreased from the previous segment on a specific interstate route. This insight into the significance of speed limit was quantified through variable prioritization, as depicted in Figure 4. Surprisingly, the speed limit emerged as the second most influential variable, accounting for over 20% of the overall importance. This is an important finding in this study. Interestingly, existing research has also underscored the significance of speed-related factors (Dutta and Fontaine 2019, Pei, Wong and Sze 2012, Rahman 2022, Rahman, Zhang and Chen 2023). While the HSM approach tends to overlook speed-related variables in predicting interstate crashes, our study emphasized the critical role of such factors, especially highlighting the importance of the speed limit.

$$\mu_i = \exp(1.12 + 0.778 \log[AADT] + 0.941 \log[Section Length] + \beta Median Type - 0.079 Speed Limit) \quad (2)$$

[Note: if Median Type, Curbed: $\beta=0.546$; Positive: $\beta=0.364$; Unprotected: $\beta=0$]

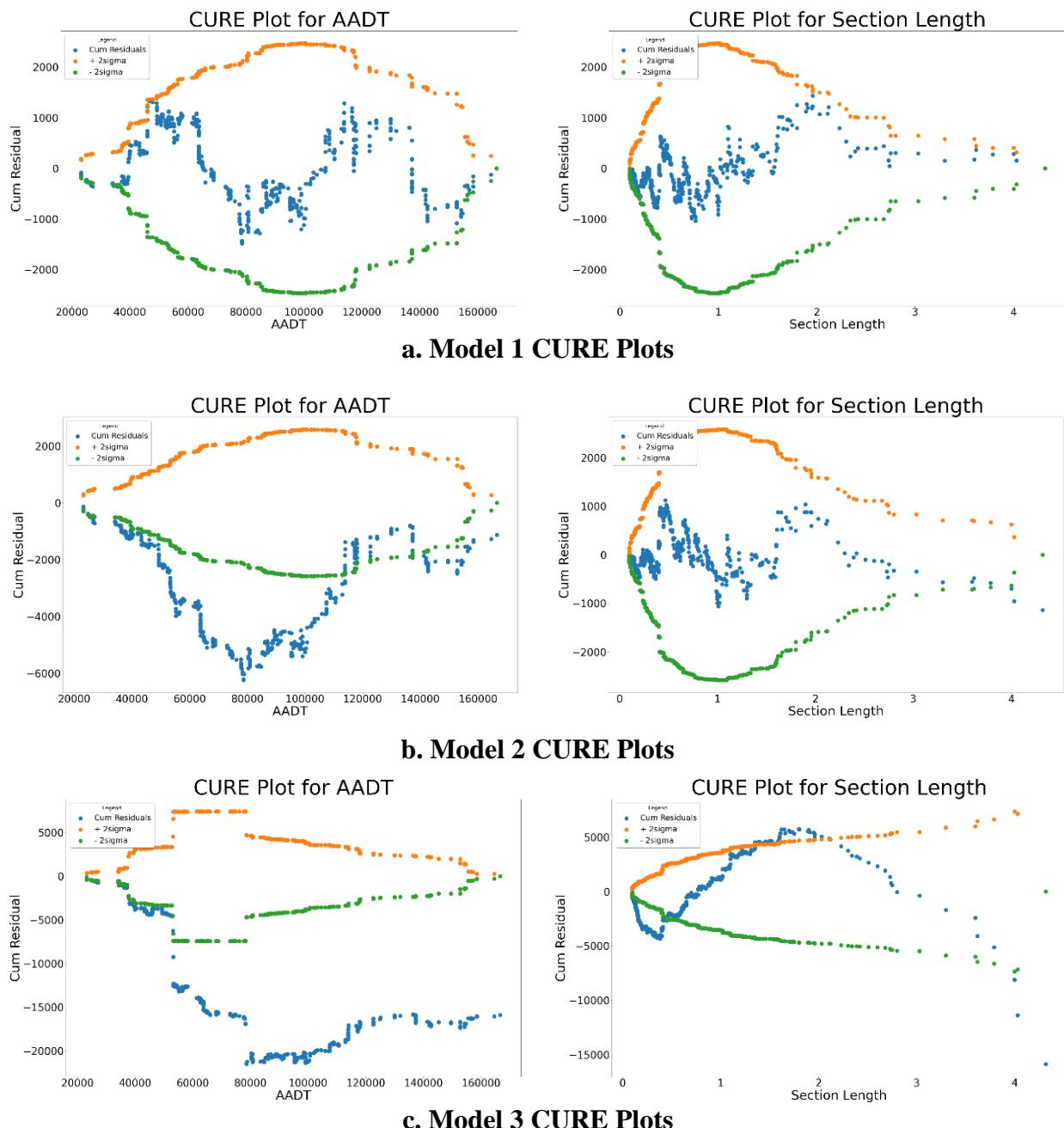


Figure 3. CURE plots for the Study Models

CONCLUSION

The primary goal of this study was to develop a model tailored specifically for New Jersey, intended to predict interstate crashes. To achieve this, we employed a negative binomial regression technique, incorporating a range of roadway attributes encompassing geometric variables, traffic volume, and speed limit considerations. After assessing various combinations of variables, our final model incorporated AADT, section length, median type, and speed limit as significant contributors. Notably, AADT and section length displayed a positive correlation with

crash frequency, which aligns with their role as exposure variables. Noteworthy was the revelation that interstates with concrete barriers experienced fewer crashes compared to those with curbed barriers, suggesting a potential advantage in leveraging concrete barriers for crash reduction.

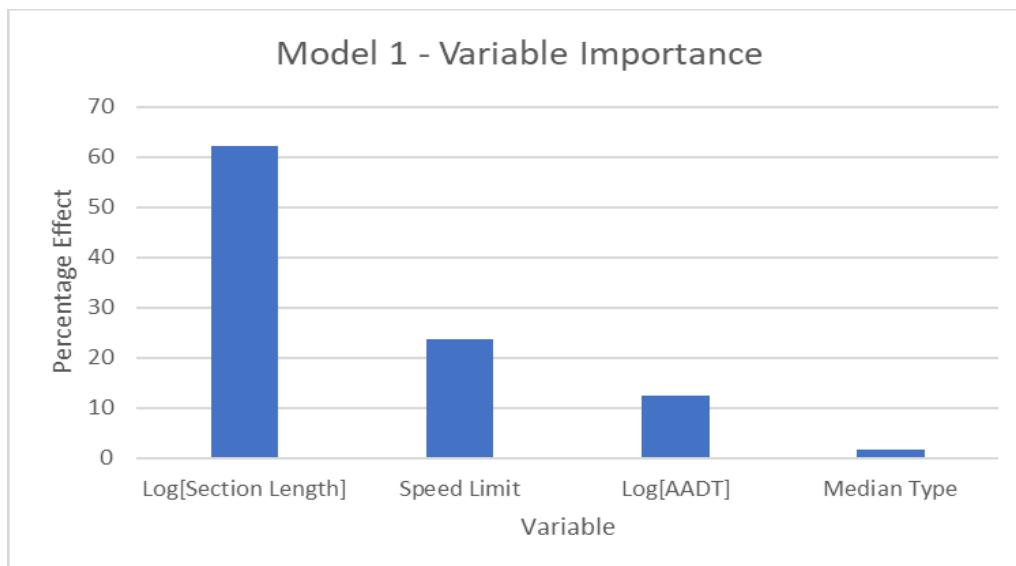


Figure 4. Importance of Variables in Model Predictions

Interestingly, our study uncovered an aspect that is overlooked in the widely utilized safety assessment guidebook, the Highway Safety Manual – the impact of speed limits on crash frequency. Through a variable prioritization process, we identified speed limit as the second most important factor. Further exploration of the dataset revealed that crashes were more prevalent when the speed limit decreased from the previous segment on a specific interstate route. These findings carry significant implications for transportation agencies and policymakers. Practical applications include the implementation of Intelligent Transportation Systems (ITS) to proactively signal upcoming speed limit changes. Such measures could effectively mitigate crashes, especially those associated with alterations in speed limits. Overall, this study identified the factors contributing to the high crashes in New Jersey's interstates. Understanding these factors can aid in formulating targeted strategies to improve the safety of these roads.

ACKNOWLEDGMENTS

The author extends sincere gratitude to James Danielwicz, John M McCleery, and Steven Lange at Rowan University for their invaluable assistance in the data collection process.

REFERENCES

- (2021). *Safety Prediction Methodology and Analysis Tool for Freeways and Interchanges*.
- (2023). *Safety Prediction Model for Freeway Facilities with High Occupancy Lanes*.
- Dutta, N., and M. D. Fontaine. (2019). Improving freeway segment crash prediction models by including disaggregate speed data from different sources. *Accident Analysis & Prevention*, 132, 105253.

- HSM. (2010). *Highway Safety Manual*. Washington DC: American Association of State Highway and Transportation Officials.
- Li, P., M. Abdel-Aty, and J. Yuan. (2020). Real-time crash risk prediction on arterials based on LSTM-CNN. *Accident Analysis & Prevention*, 135, 105371.
- Lu, J., K. Haleem, A. Gan, and P. Alluri. (2014). Safety performance functions for Florida's freeways. In *T&DI Congress 2014: Planes, Trains, and Automobiles*, 321-331.
- Lyon, C., B. N. Persaud, and F. B. Gross. (2016). *The Calibrator-An SPF Calibration and Assessment Tool User Guide*. United States. Federal Highway Administration. Office of Safety.
- Matarage, I. C., and S. Dissanayake. (2020). Quality assessment between calibrated highway safety manual safety performance functions and calibration functions for predicting crashes on freeway facilities. *Journal of traffic and transportation engineering (English edition)*, 7, 76-87.
- Ozbay, K., H. Nassif, B. Bartin, C. Xu, and A. Bhattacharyya. (2019). *Calibration/Development of Safety Performance Functions for New Jersey*. 171. New Jersey Department of Transportation Bureau of Research And US Department of Transportation Federal Highway Administration.
- Pei, X., S. Wong, and N.-N. Sze. (2012). The roles of exposure and speed in road safety analysis. *Accident analysis & prevention*, 48, 464-471.
- Rahman, F. (2022). *Incorporating Speed Into Crash Modeling for Rural Two-lane Highways*.
- Rahman, F., X. Zhang, and M. Chen. (2023). Evaluating Effect of Operating Speed on Crashes of Rural Two-Lane Highways. *Journal of Advanced Transportation*, 2023.
- Shin, H.-S., Y.-J. Lee, S. Dadvar, and S. Bharti. (2016). *The development of local calibration factors-phase II: Maryland freeways and ramps*. Maryland. State Highway Administration. Office of Policy & Research.
- Sun, C., H. Brown, P. K. Edara, B. Claros, and K. Nam. (2013). *Calibration of the highway safety manual for Missouri*. Mid-America Transportation Center.

Evaluation of Sinusoidal Rumble Strips on Noise Mitigation and Bicycle Accommodation

Pei-Sung Lin, Ph.D., P.E.¹; Cong Chen, Ph.D., P.E.²; Yaye Keita, Ph.D.³;
Shubhankar Shindgikar⁴; and Ela Białkowska-Jelińska⁵

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: lin@usf.edu

²Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: congchen1@usf.edu

³Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: ykeita@usf.edu

⁴Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: shindgikar@usf.edu

⁵Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: elzbieta@usf.edu

ABSTRACT

Lane departure crashes are among the most common crashes in Florida and in the United States. Many departments of transportation (DOTs) prioritize reducing both the frequency and severity of these accidents. Auditory vibratory treatments (AVTs), or rumble strips, are frequently used to prevent or reduce lane departure incidents. Traditional rumble strips produce exterior noises that become a burden to nearby residents. To address the noise issues, several DOTs considered a new type of AVT called sinusoidal rumble strips. The research team conducted an in-depth literature review, agency interviews, a focus group, and an assessment of studies conducted by the Florida Department of Transportation (FDOT) to evaluate the noise reduction and bicycle accommodation of sinusoidal rumble strips installed in Florida and across the nation. The research concludes that sinusoidal rumble strips can effectively mitigate external noise problems while effectively notifying drivers and occupants of lane departures.

INTRODUCTION

Lane departure crashes are among the most common crashes in the United States. The Federal Highway Administration (FHWA) defines a roadway departure (RwD) crash as a crash that occurs after a vehicle crosses an edgeline or a centerline or otherwise leaves the travel way. From 2016 to 2018, an average of 19,158 fatalities resulted from roadway departures, accounting for 51% of all traffic fatalities in the U.S. RwD is also a top contributing factor for traffic crashes in Florida. In 2015–2019, lane departures accounted for 34% of all crashes and 42% of traffic fatalities on Florida roadways. About one-third of lane departure crashes resulted in a collision with another moving vehicle, sometimes head-on, and two-thirds involved hitting a tree or another fixed object. Slightly more than half of fatal lane departure crashes occurred in rural areas where there were more two-lane roadways, narrow shoulders, and long stretches of relatively empty roadway.

Due to the high number of roadway departure crashes, prevention of those crashes has become the priority of many Departments of Transportation (DOTs) across the nation. Auditory Vibratory Treatments (AVTs), or rumble strips, are frequently used to prevent or reduce lane

departure incidents. Rumble strips, which have proven to be effective in reducing run-off-road and lane departure crashes by approximately 25% (Karkle, Rys, Russell, & Security, 2013; Wu, Geedipally, & Pike, 2018), typically are thermoplastic designs or consist of cylindrical design patterns milled into the pavement and placed perpendicular to the direction of traffic flow. They function by enabling tires to enter the groove, generating both noise and vibration. Raised rumble strips, installed as an alternative in states where they do now plow snow, encompass various forms such as side-by-side raised pavement markers or plastic inserts embedded within thermoplastic pavement markings. When a vehicle passes over a rumble strip, the driver receives an audible alert and feels a vibration, both of which serve as warnings of lane departure for drivers. These types of rumble strips are effective in alerting drivers, but generate external noise that may inconvenience nearby residents.

Certain transportation authorities, such as the Florida Department of Transportation (FDOT), have considered the use of milled sinusoidal rumble strips. These strips aim to mitigate external noise while effectively signaling drivers of potential roadway departures through adequate noise and vibration cues. This paper reviews the findings of the noise reduction effects of sinusoidal rumble strips implemented by selected state transportation agencies and conducted a detailed evaluation of sinusoidal rumble strips installed in Florida. It explores the potential for sinusoidal rumble strips to reduce exterior noise while appropriately alerting drivers when leaving the travel lane. A literature review, agency interviews on sinusoidal rumble strip noise control and bicycle accommodation, a focus group noise study, and FDOT field noise study evaluations of sinusoidal rumble strips installed in Florida were performed as part of the study.

The remainder of the paper is organized as follows: the literature review section elaborates on sinusoidal AVT design, implementation, and noise reduction, as well as the accommodation of bicyclists; the research approach section provides details on data collection and the analysis approach; the effects of sinusoidal rumble strips on interior and exterior noise are summarized in the results section; and the last section provides the conclusions.

LITERATURE REVIEW

Many states across the U.S. use rumble strips as a low-cost proven safety countermeasure to decrease or prevent lane departure collisions by providing a vibratory or audible warning, or both, to distracted and fatigued drivers. However, the exterior noises from traditional rumble strips, including cylindrical ground-in and profiled thermoplastic rumble strips, are the subject of many complaints from residents living near those strips. In response, some states are considering a different type of rumble strip, called sinusoidal rumble strips, to reduce noise pollution.

For example, in California the sinusoidal rumble strips were designed to ensure that when driving over the strips using standard vehicles, the inside noise and vibration are optimized, and the exterior noise is kept to a minimum. When compared to a conventional rumble strip design, the sinusoidal rumble strips in California lowered the external noise by 6 dBA for passenger vehicles and 3 dBA for a dump truck (Donavan, 2018; Staats, et al., 2020). Interior noise levels in passenger vehicles traveling on the California sinusoidal rumble and cylindrical rumble strips were 14.4 dBA and 13.9 dBA greater than the noise in vehicles traveling off the strips, respectively (Donavan, 2018; Staats, et al., 2020).

The Washington State Department of Transportation (WSDOT) evaluated new rumble strip designs to minimize roadside noise and increase safety (Laughlin et al., 2018). Four alternative

designs were analyzed by WSDOT, including a sinusoidal rumble strip design and three other cylindrical rumble strips (Staats et al., 2020). All four design types were evaluated using a mid-sized SUV moving at a speed of 60 mph. Exterior sound levels were recorded at 25 ft and 50 ft from the center of the travel lanes. Interior sound levels were measured from the passenger seat at ear level (Staats et al., 2020). The results suggested that the sinusoidal rumble strip design had the lowest sound levels outside the vehicle and produced sufficient interior noise compared to the other designs (Laughlin et al., 2018).

The Oregon Department of Transportation (ODOT) assessed the feasibility of replacing cylindrical rumble strips with sinusoidal rumble strips on Oregon roadways (Hurwitz et al., 2019; Kalathas et al., 2019). The results of the study confirmed that cylindrical rumble strips produced an exterior noise of 5 dBA higher than the baseline condition (no strips) for both a passenger car and a van, whereas the sinusoidal rumble strips generated exterior noise 3 dBA higher than the no-strips condition for the car and no noticeable difference compared to no strips for the van. The cylindrical rumble strips generated interior noise levels of 10 dBA and 12 dBA higher than the baseline for both the passenger car and the van, respectively, (Kalathas et al., 2019), and the sinusoidal rumble strips produced interior noise levels that were 4.6–5.8 dBA greater than the baseline (no-strips). According to FHWA, as little as a 3 dBA increase and ideally a 5 dBA increase in internal noise would be enough to alert drivers (Hurwitz et al., 2019; Staats et al., 2020).

The Minnesota Department of Transportation (MnDOT) funded a research study that led to the installation and evaluation of four types of sinusoidal rumble strip designs on a two-lane rural roadway in east central Minnesota. Three types of vehicles were used to assess the noise level—a passenger car (Ford Fusion), a small truck (Ford F-150), and a dump truck (Sterling Class 35). The four rumble strip designs performed differently in the study for all three vehicle categories (Terhaar et al., 2016).

Similarly, the Indiana Department of Transportation (INDOT) and Purdue University evaluated different sinusoidal rumble strip designs on Indiana's highway network, including sinusoidal rumble strips with 12-, 18-, and 24-in. wavelengths, and compared the interior and exterior noise produced from these rumble strips to Indiana's conventional rumble strip design. The three sinusoidal rumble strip designs produced 5–11 dBA less exterior noise and 9 dBA higher noise inside the vehicle when compared with the conventional rumble strips (Mathew et al., 2018; Staats et al., 2020).

According to the Federal Highway Administration (FHWA), rumble strips must be designed and installed to suit all roadway users. Rumble strips have been recognized as having a particularly negative impact on bicyclists. Often, bicyclists are forced to ride in travel lanes in circumstances where rumble strips are built on the shoulder without proper room for bicyclists (Ahmed et al., 2015). This condition exposes bicyclists to vehicle traffic threats that may result in crashes leading to severe injuries or fatalities. Thus, rumble strips should be implemented in a way that makes it easier for bicyclists to ride on the shoulders. Bicyclists should have at least 4 ft (1.2 m) of space between the rumble strips and the edge of the pavement, with more (5 ft) space if safety barriers are present (Ahmed et al., 2015).

Additionally, shoulder rumble strips are not recommended on routes designated as bicycle paths or in high bicycle-use zones. When constructing rumble strips in residential areas, many states consider bicyclists by either not installing rumble strips within city limits or, if necessary, evaluating collision data and adopting modified shallower depth rumble strips.

RESEARCH APPROACH

Noise assessment and evaluation of bicycle accommodation when implementing sinusoidal rumble strips in other states and in Florida were performed as part of this study. To accomplish this, in addition to the in-depth literature review and interviewing agency representatives from five states (California, Indiana, Kentucky, Minnesota, and Washington), the research team also conducted a focus group study involving 8 participants. The aim was to evaluate how effective three distinct designs of FDOT sinusoidal rumble strips are in alerting drivers about lane departures by considering internal noise levels and vibrations.

The research team further evaluated both external noise and cabin noise data provided by the FDOT Materials Office. Findings from the literature review were used to identify the agencies that would be suitable candidates for the interviews, notably the agencies that are pioneers in sinusoidal rumble strip implementation. The review also allowed us to identify gaps in the literature related to sinusoidal rumble strips implementation.

Agency interviews. Based on the literature review results, the research team conducted interviews with selected transportation agencies on their experience in sinusoidal AVT design and implementation to ascertain information that could benefit the sinusoidal AVT implementation in Florida and other states. As mentioned previously, representatives from five states that are pioneers in sinusoidal rumble strip implementations were interviewed, including: California, Indiana, Kentucky, Minnesota, and Washington.

The interview questionnaire was designed to collect more details on the following aspects:

- Sinusoidal AVT designs and implementations (design specifications, design variations and state guidelines, installation and maintenance experience and cost)
- Noise evaluation experience and results
- Consideration of bicycle accommodation in sinusoidal AVT implementation

Sample questions asked during the interview include:

- Has your agency or state used or adopted any sinusoidal rumble strip designs?
- Has your agency or state conducted noise studies of different sinusoidal rumble strips in your state?
- Were the implemented sinusoidal rumble strips in your state able to reduce the exterior noise while still providing enough interior noise to alert the driver?
- Does your agency or state consider bicyclists during shoulder sinusoidal rumble strip installation?
- What are specific bicycle accommodations of rumble strip implementation for your state?

Focus group study to examine inside vehicle noise levels to alert drivers. For the focus group, two different sites in Florida were studied—US-301 in Clay County and SR-100 in Putnam County. Each site had the three different FDOT sinusoidal rumble strip designs (Type 1, Type 2, and Type 3) as shown in Figure 1. They were installed on the shoulders and edgelines. Sample photos of FDOT sinusoidal rumble strip designs on shoulder and edgeline are shown in Figure 2. US-301 (a six-lane highway) southbound had six 1200-ft test sections with three different sinusoidal rumble strips on the shoulder and three different sinusoidal rumble strips at the edgeline. Similarly, SR-100 (a two-lane highway) eastbound and westbound had twelve 1200-ft test sections with three different sinusoidal rumble strips on the shoulder and three different sinusoidal rumble strips at the edgeline for each direction.

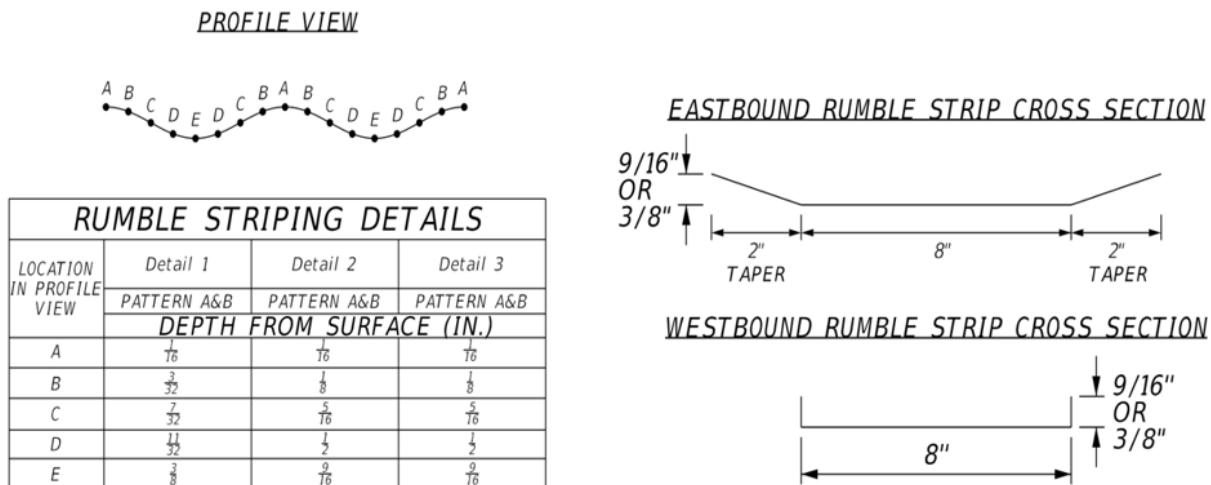


Figure 1. Details of three types of FDOT sinusoidal rumble strips

Source: FDOT, (n.d.)

Shoulder Rumble Strip



Edge Rumble Strip



Figure 2. Sample photos of FDOT sinusoidal rumble strip designs on shoulder and edgeline.

Four vehicles, as shown in Figure 3, were used to conduct the focus group study to test the inside noise levels of three FDOT sinusoidal rumble strip designs for the ability to alert drivers of potential roadway departures when driving over them (shoulder rumble strip and edgeline rumble strip) on US-301 and SR-100. Drivers and passengers took turns driving the vehicles at the designated speed limit and 5 mph above the speed limit. Some participants did two or three test runs when necessary to confirm their ratings. After each test, passengers and drivers completed the focus group questionnaire with their ratings and explanations for those ratings. When the testing at each site was complete, the team held a group discussion with participants about their ratings for that site and why they thought their selected type of rumble strip was better.

Analysis of FDOT Materials Office OBSI data on external noise reduction. The research team assessed the results from the noise studies performed by the FDOT Materials Office at the same two locations as the focus group study sites—US-301 and SR-100. FDOT used On-Board Sound Intensity (OBSI) testing to measure the outside noise at the tire/pavement interface using two microphones mounted vertically 4 in. from the outside tire sidewall of the rear passenger side tire (Figure 4). For each test section, the noise levels were measured three times for five seconds at 60 mph. FDOT also measured the cabin (inside) noise of the test truck by using a CESVA SC 310 sound level meter and by placing the sound level meter 5 in. above the center of the rear seat (Figure 4). Likewise, the noise levels were assessed three times for five seconds at 60 mph for each test section.



Figure 3. Vehicles used to conduct a focus group study.



Figure 4. FDOT OBSI testing set up (left) and cabin noise testing set up (right).
(Source: FDOT, 2022a)

RESULTS

This section covers the results from the different methods used to assess the effectiveness of sinusoidal rumble strips in reducing exterior noise from vehicular traffic on roadway departure crash prevention strips. The literature review reveals that sinusoidal rumble strips are more effective in addressing noise issues compared to other types of rumble strips, as they reduce exterior noise while still effectively alerting drivers when they depart from the travel lane. More details on the results from the agency interviews, focus group study, and FDOT materials office data are provided next.

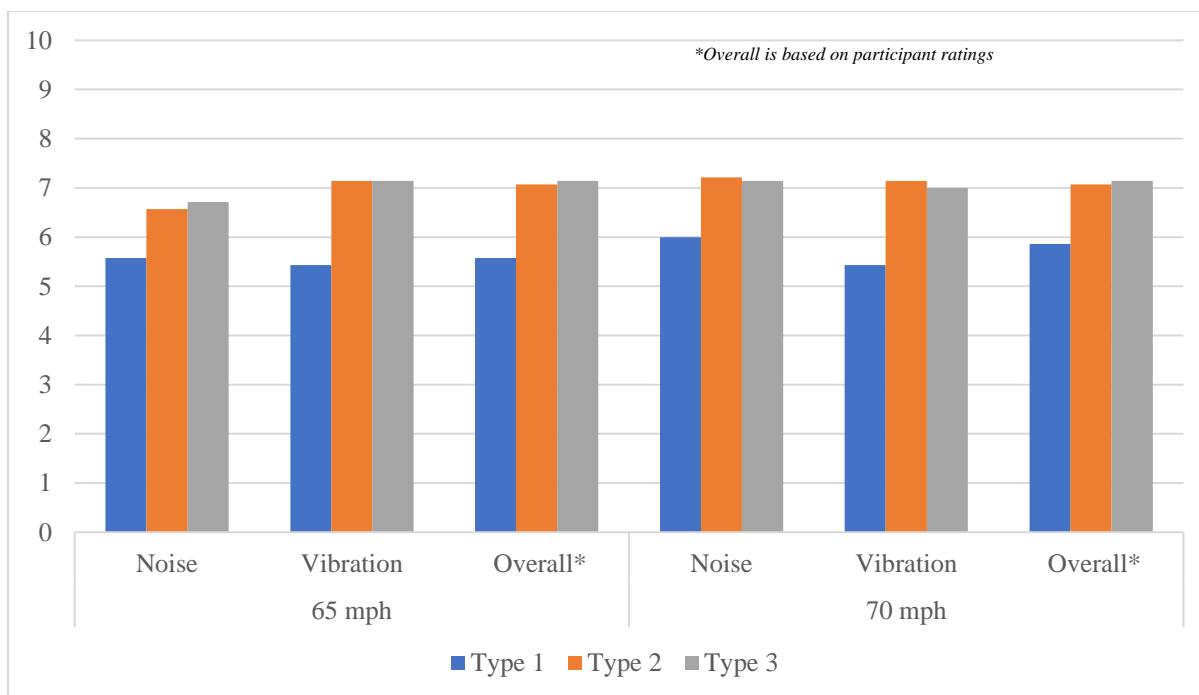
Noise analysis based on agency interviews. From the agency interviews with experts from California, Indiana, Kentucky, Minnesota, and Washington, it was found that sinusoidal rumble strips can alleviate the noise burdens on residents living near rumble strips. All the experts reached a consensus that using sinusoidal rumble strips is an efficient way to reduce the complaints of residents.

The interviews also revealed the types of design that can help accommodate bicyclists when implementing sinusoidal rumble strips. For example, it was highlighted that the negative impact of rumble strips on bicyclists can be mitigated by ensuring that it is easier for bicyclists to ride on roadway shoulders by having at least 4 ft (1.2 m) of space between the rumble strips and the edge of the pavement, with more (5 ft and over) if safety barriers are present. Paved shoulder width is a significant factor in deciding whether to use edgeline rumble strips or shoulder rumble strips. Edgeline rumble strips narrower than standard 12-inch can be used to accommodate pedestrians and bicyclists within existing road shoulders. Shoulder rumble strips should have a 10' gap every 50' if the shoulders are 3' or wider, or 12' gap every 48' as rumble strips gaps for bicyclists.

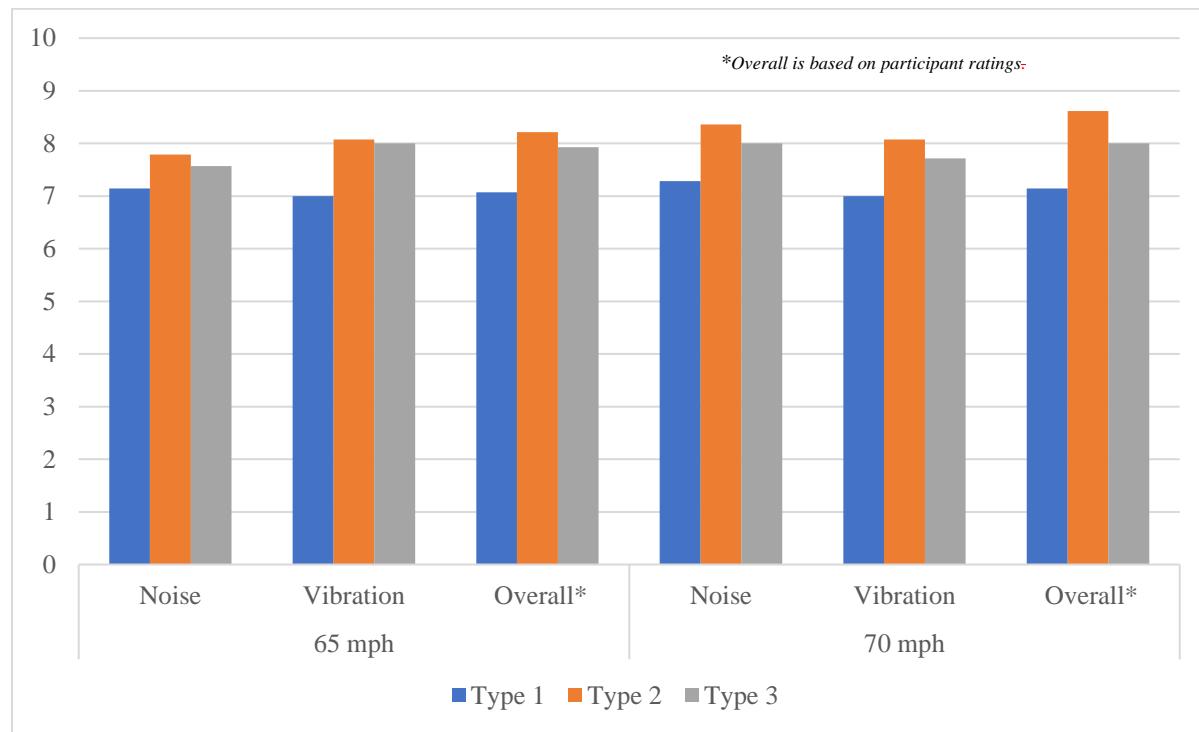
Noise analysis based on the focus group study. Figure 5 illustrates the findings of the focus group study, which examined drivers' experiences with three varieties of FDOT sinusoidal rumble strips along the six-lane US-301, where the posted speed limit is 65 mph. Figure 5(a) shows average driver ratings for US-301 shoulder sinusoidal rumble strips, and Figure 5(b) shows average driver ratings for US-301 edgeline sinusoidal rumble strips. The ratings for both Type 2 and Type 3 were higher than for Type 1 regardless of the speed or whether shoulder or edgeline was considered. Type 2 and Type 3 had similar ratings for shoulder and edgeline.

At higher speeds, for both Type 2 and Type 3, the noise from the rumble strips was more noticeable than the vibration. Vibration was the same for Type 2 at both speeds, but higher for Type 3 at lower speed. These results were the same for both shoulder and edgeline. For shoulders, overall, the performance of Type 2 and Type 3 were similar, and the difference was insignificant. Considering the edgeline, overall, Type 2 slightly outperformed Type 3. Comparing driver ratings for US-301, edgeline strips got higher ratings than shoulder strips at both 65 mph and 70 mph. Similar results were observed for passengers.

Another reason why Type 2 was found to be better during the focus group study is because of the pitch level. It is stated that human ears are generally more sensitive to sounds in the mid-frequency range, which corresponds to the frequencies of many human speech sounds. Among the three types of sinusoidal rumble strips evaluated during the noise studies, Type 1 was found to produce high-pitched sound while Type 3 was determined to yield low-pitched sound. Among the three types, the sound from Type 2 was intermediate, thus allowing drivers to be alerted effectively. The intermediate pitch level of the sound from Type 2 (not too high or too low) is effective and pleasant as an alert to people.



(a) Average driver ratings for US-301 shoulder sinusoidal rumble strips



(b) Average driver ratings for US-301 edgeline sinusoidal rumble strips

Figure 5. Focus group noise study results on US-301.

Unlike US-301, the winding two-lane section of SR-100 with a speed limit of 60 mph posed challenges for drivers in maintaining the required speed, potentially affecting their ratings.

Regardless, Type 1 underperformed compared to Type 2 and Type 3 for both shoulder and edgeline strips at 60 mph and 65 mph. For SR-100, Type 3 was rated slightly higher than Type 2 for shoulder and edgeline related to noise, vibration, and overall. At higher speed, for both Type 2 and Type 3, noise from the rumble strips was higher than at lower speed. Like US-301, edgeline strips had higher ratings than shoulder strips for SR-100 at 60 mph and 65 mph.

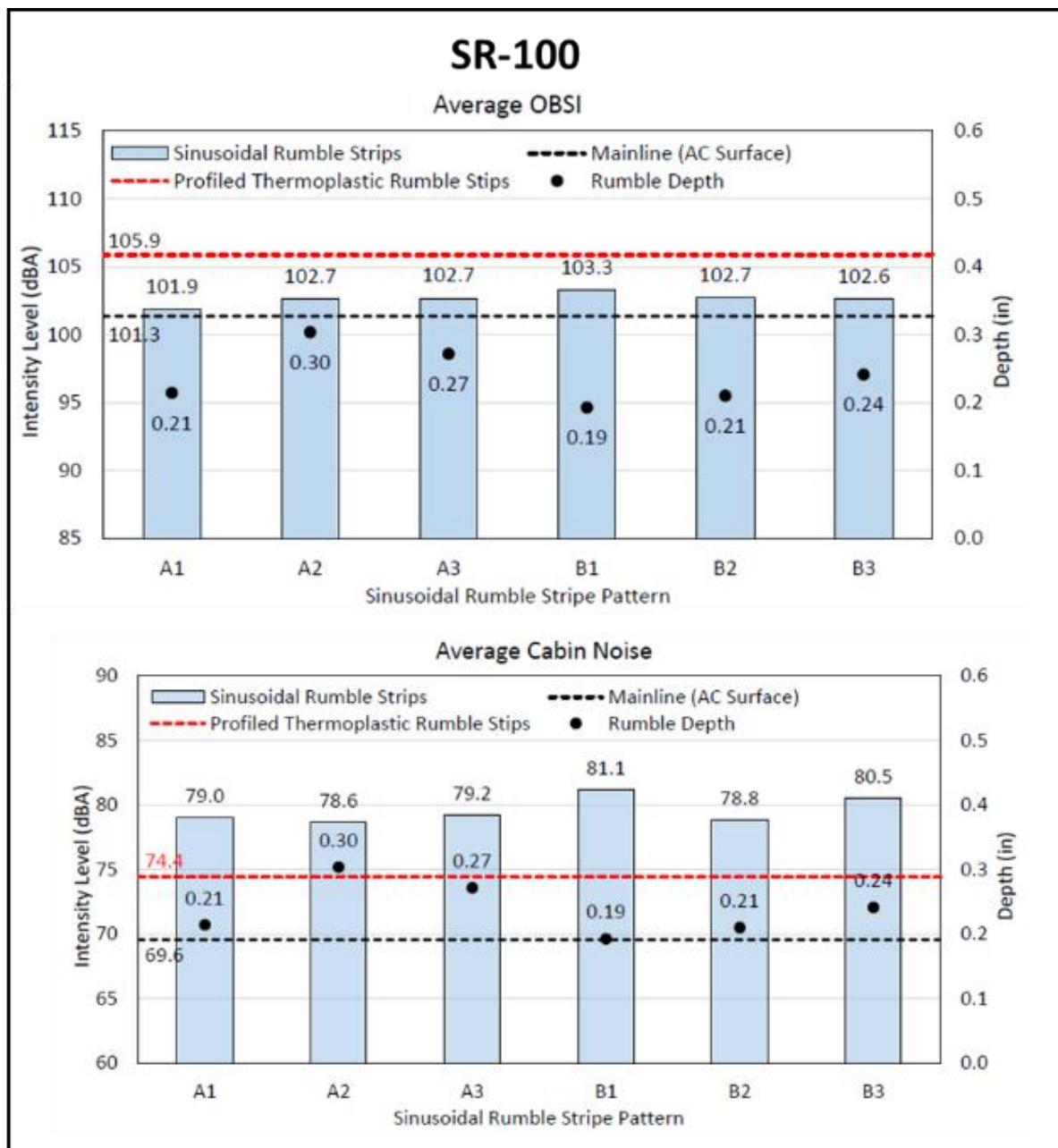


Figure 6. FDOT Materials Office testing results. Source: FDOT, (2022a and 2022b)

During the focus group study, the performance of Types 2 and 3 was better than Type 1. For the tests on US-301, the performance from Type 2 was slightly better than Type 3, but for SR-100, the performance from Type 3 was slightly better than that from Type 2. For the pitch of

sounds, Type 1 was the lowest, Type 3 was the highest, and Type 2 is high but lower than Type 3. Overall, the performance of Types 2 and 3 was similar. Participants from the focus group study preferred Type 2 slightly more than Type 3 considering drivers' experience, consistency of ratings, and the pitch of sounds.

Noise analysis based on the FDOT Material Office OBSI data. The analysis was conducted based on the FDOT Material Office OBSI data measuring external noise and cabin noise data measuring the inside vehicle noise to alert drivers. The data results were similar for both sites at US301 and SR-100. The FDOT Materials Office testing results on SR-100 are shown in Figure 6 as an example.

On the upper diagram of Figure 6, the intensity level of exterior noise without rumble strips was shown by a black dash line. The intensity level of exterior noise with traditional profiled thermoplastic rumble strips was shown by a red dash line. The exterior noise level produced by profiled thermoplastic rumble strips is significantly higher than that without the rumble strips. However, the intensity of exterior noise levels produced by the sinusoidal rumble strips regardless of any design or pattern is much lower than that from profiled thermoplastic rumble strips. It is only slightly higher than that without the rumble strips. The results clearly demonstrate the effectiveness of the sinusoidal rumble strips to reduce exterior noises and address the noise complaint issues.

On the lower diagram of Figure 6, the intensity level of cabin noise without rumble strips was shown by a black dash line. The intensity level of cabin noise with a common traditional profiled thermoplastic rumble strips was shown by a red dash line. The cabin noise level produced by profiled thermoplastic rumble strips is significantly lower than those produced by the sinusoidal rumble strips regardless of any design or pattern. The results clearly demonstrate that the implementation of sinusoidal rumble strips is more effective than that of profiled thermoplastic rumble strips to alert drivers about roadway departure risks.

CONCLUSIONS

This study assessed the noise reduction effects of sinusoidal rumble strips across the nation and in Florida. Based on the findings from the literature review, agency surveys, a focus group noise study, and an FDOT Materials Office noise data analysis, sinusoidal rumble strips are the most effective type of strips to address the noise issues while reducing roadway departure crashes.

The results from the focus group noise study and FDOT Materials Office noise data analysis revealed that FDOT's Type 2, or Type 3 sinusoidal rumble strip designs are more promising than Type 1 with relation to noise control. Researchers preferred the FDOT Type 2 design to the Type 3 design when considering drivers' experience, consistency of ratings, and the pitch of sounds. It is necessary to note that various other factors, such as speed, vehicle type, or vehicle model, can affect the noise levels both inside and outside of vehicles when driving on the sinusoidal rumble strips. The performance of the rumble strips may also depend on driving angle or whether driving in a straight line or weaving on the strips.

The analysis of the FDOT Materials Office noise data concluded that the intensity level of exterior noises produced by the sinusoidal rumble strips is much lower than that produced by common traditional profiled thermoplastic rumble strips. The implementation of sinusoidal rumble strips can significantly reduce exterior noise and address noise complaint issues from nearby residents. The implementation of sinusoidal rumble strips produces louder cabin noise or

inside vehicle noise, so it is more effective than the implementation of profiled thermoplastic rumble strips to alert drivers about roadway departure risks.

To mitigate the negative impact of rumble strips on bicyclists, rumble strips should be designed to make it easier for bicyclists to ride on roadway shoulders. Bicyclists should have at least 4 ft (1.2 m) of space between the rumble strips and the edge of the pavement, with more (5 ft and over) if safety barriers are present. This FDOT research project concludes that sinusoidal rumble strips can effectively address the exterior noise issues to address residents' complaints, and significantly alert drivers when leaving the travel lanes.

ACKNOWLEDGMENTS

This research project was sponsored by FDOT. Authors would like to express sincere appreciation to Gevin McDaniel, Brenda Young, David Amato, and Derwood Sheppard for their support, guidance, and assistance.

REFERENCES

- Ahmed, M., Sharif, M. A., and Ksaibati, K. (2015). Developing an effective shoulder and centerline rumble strips/stripes policy to accommodate all roadway users. Wyoming Department of Transportation. Retrieved from https://rosap.ntl.bts.gov/view/dot/29526/dot_29526_DS1.pdf.
- Donavan, P. (2009). Recommendation for initial sinusoidal rumble strip design. Technical Memorandum to Dennis McBride and Bruce Rymer at California Department of Transportation, August 4.
- FDOT (Florida DOT). (2022a). US-301 rumble strips noise test. Memorandum. Tallahassee, FL.
- FDOT (Florida DOT). (2022b). SR-100 rumble strips noise test in Putnam County. Memorandum.
- FDOT (Florida DOT). (n.d.). *US-301 Plan Sheet*.
- Hurwitz, D. S., Horne, D., Jashami, H., Monsere, C. M., and Kothuri, S. (2019). Quantifying the performance of low-noise rumble strips (No. FHWA-OR-RD-19-07). Retrieved from <https://www.oregon.gov/ODOT/Programs/ResearchDocuments/SPR800FinalReport.pdf>.
- Kalathas, P., Parrish, C., and Zhang, Y. (2019). Rumble strip design evaluation based on exterior noise using finite element analysis (No. OR-RD-19-10). Oregon Department of Transportation. Retrieved from <https://www.oregon.gov/ODOT/Programs/ResearchDocuments/RumbleStripsFinal.pdf>.
- Karkle, D. E., Rys, M. J., and Russell, E. R. J. J. o. T. S., & Security. (2013). Safety effectiveness of centerline rumble strips in Kansas. 5(1), 1-26.
- Laughlin, J., and Donahue, J. (2018). Evaluation of new rumble strip designs to reduce roadside noise and promote safety. Washington State Department of Transportation. Retrieved from <https://www.wsdot.wa.gov/research/reports/fullreports/881-1.pdf>.
- Mathew, J. K., Balmos, A. D., Plattner, D., Wells, T., Krogmeier, J. V., and Bullock, D. M. (2018). Assessment of alternative sinusoidal rumble stripe construction. Indiana Department of Transportation. Retrieved from <https://docs.lib.psu.edu/cgi/viewcontent.cgi?article=3191&context=jtrp>.

- Staats, W. N., Agent, K. R., and Howell, B. K. (2020). Evaluation of alternative rumble strip designs. Kentucky Transportation Center, Kentucky Transportation Cabinet. Retrieved from https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=2716&context=ktc_researchreports.
- Terhaar, E., Braslau, D., and Fleming, K. (2016). Sinusoidal rumble strip design optimization study (No. MN/RC 2016-23). Minnesota Department of Transportation. Retrieved from <https://www.dot.state.mn.us/trafficeng/safety/docs/sinusoidalrumblestripdesignoptimizationstudy.pdf>.
- Wu, L., Geedipally, S. R., and Pike, A. M. J. T. R. R. (2018). Safety evaluation of alternative audible lane departure warning treatments in reducing traffic crashes: an empirical Bayes observational before–after study. 2672(21), 30-40.

Effects of Law Enforcement Presence, Frequency, and Duration on Speeding

Pei-Sung Lin, Ph.D., P.E.¹; Yaye Keita, Ph.D.²; Shubhankar Shindgikar³; Emmeth Duran, P.E.⁴;
Ela Białkowska-Jelińska⁵; and Rakesh Rangaswamy⁶

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: lin@usf.edu

²Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: ykeita@usf.edu

³Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: shindgikar@usf.edu

⁴Florida Dept. of Transportation, Tampa, FL. Email: Emmeth.Duran@dot.state.fl.us

⁵Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: elzbiet@usf.edu

⁶Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: rrangaswamy@usf.edu

ABSTRACT

Speeding is the major cause of severe crashes in the United States. Speed management is essential to reduce fatalities and serious injuries for all road users in any crash, especially for vulnerable road users (VRUs), such as pedestrians and bicyclists. This study evaluated the effects of law enforcement presence on vehicle speeds for roadway segments with speeding issues. The project team used the citations and warning data collected through the Florida Department of Transportation (FDOT) District 7 Enhanced Law Enforcement Engagement (ELEE) project to assess the effects of law enforcement presence, frequency, and duration on vehicle speeds on study corridors. The research findings demonstrate that the effects of police presence and number of citations and warnings on vehicle speeds were instantaneous, but no lasting impact of police presence was observed. The effects of police presence on vehicle speed may be limited if the law enforcement officers are not visible to the public, or limited citations and warnings are issued by the police officers.

INTRODUCTION

Speeding, exceeding the posted speed limits or traveling too fast for conditions, was a contributing factor in almost 29 percent of all fatalities in 2021. Of the 42,939 fatalities that occurred on our Nation's roadways that year, 12,330 were speeding-related—an increase of 7.9 percent from 2020 (Stewart, 2023). Studies clearly show that higher speeds result in greater impact at the time of a crash, which leads to more severe injuries and fatalities (Elvik, 2005; WHO, 2008). Speed management is essential to reduce injury risk for all road users in any crash, especially for vulnerable road users (VRUs), such as pedestrians and bicyclists.

As one of five key elements of a Safe System Approach to reduce death and serious injuries, safe speed on our roadways is essential. Besides implementations of education and engineering strategies and associated countermeasures, law enforcement for roadway segments with speeding issues is important for speed management. When safe speed cameras cannot be installed at those critical locations, the frequent presence of law enforcement becomes crucial. Because law

enforcement officers cannot be at locations with speeding problems all of the time, investigating and providing recommendations on frequency and duration of law enforcement presence to achieve effective speed management is needed.

This study, sponsored by Florida Department of Transportation (FDOT) District 7, evaluated the effects of law enforcement activities on speeding. The research team used the citations and warning data collected through the FDOT District 7 Enhanced Law Enforcement Engagement (ELEE) project to assess the effects of law enforcement presence, frequency, and duration on vehicle speeds. An additional factor that is considered is the impact of the number of citations and warnings on vehicle speeds. It should be noted that when police officers want to catch speeders (drivers with excessive vehicle speeds exceeding a posted speed limit), they tend to be out of the sight of drivers. When they are not visible, the effect on the overall vehicle speed might be minimal. When they are visible to drivers, the effect on the overall vehicle speed should be significant. This study explored and analyzed how law enforcement activities affect vehicle speeds on roadways.

The remainder of the paper is organized as follows: the literature review section elaborates on the effects of speeding on safety and the effects of law enforcement on speeding; the research approach section provides details on data collection and the analysis methods; the effects of law enforcement activities on speeding are summarized in the results section; and the conclusions are provided in the last section.

LITERATURE REVIEW

Speeding is among the major causes of injuries and fatalities in the state of Florida and across the nation. According to the Federal Highway Administration (FHWA), speeding accounts for 29% of all fatalities in the U.S. (FHWA, 2023). This presents a significant concern for all roadway users, particularly those considered vulnerable road users.

Research conducted by Povey et al. (1996) in New Zealand aimed to investigate the relationship between enforcement activity, vehicle speeds, and injury crashes. The study revealed a correlation between increased enforcement efforts, lower vehicle speeds, and a subsequent decline in injury crashes. Another study evaluated the impact of police enforcement on vehicle speeds and the enduring effects of police presence on vehicular speed (Sisiopiku and Patel, 1999). They compared average speeds before and after police presence, as well as speeds at police enforcement locations versus downstream speeds. The results indicated no lasting effects of police presence on vehicular speed and that drivers reduced their speeds just before reaching the police car, but promptly resumed their previous speeds after passing that location.

Mountain et al. (2005) assessed the impact of different speed control strategies and enforcement measures on traffic speeds and the occurrence of traffic crashes. They concluded that various enforcement and speed management methods had different effects on reducing crashes. Dowling and Holloman (2008) examined the use of surprise tactics in law enforcement efforts. They discovered that when motorists were caught off guard by the presence of law enforcement, there was a significant increase in the effectiveness of lowering speeding habits, even among those who were not initially speeding. The element of surprise resulted in a prolonged decrease in speed downstream from the enforcement unit.

Elvik (2012) highlighted the crucial need to enforce speed limits for the well-being of the general population. As part of his work, the author showed the direct link between speed and road safety and emphasized the efficacy of law enforcement measures, such as speed cameras, in

diminishing instances of speeding and frequency of crashes. Stanojevic et al. (2013) conducted a study comparing regions with varying degrees of enforcement, illustrating the significant impact of enforcement on driver behaviors. They discovered that the lack of enforcement resulted in heightened instances of speeding, reduced utilization of seat belts, elevated occurrences of drunk driving, and more frequent occurrences of hazardous driving practices. In another study conducted by Simpson et al. (2023), the effects of saturation enforcement on speeding were evaluated by a group of researchers. It was found that the strategy could be effective at managing speed.

Automated speed enforcement is also a topic of discussion within the realm of speed management strategies. For example, the automated speed enforcement initiative in Maryland showed a significant 10% decrease in average speeds and a noteworthy decline in the number of cars exceeding the posted speed limit by more than 10 mph at camera locations (Hu and Mccartt, 2016). This investigation, conducted over many years, also revealed a significant decrease (39%) in the probability of crashes leading to severe or fatal injuries. The study confirmed the beneficial effect of speed cameras in reducing incidents caused by speeding.

The Safe System Approach for speed management (Kumfer et al., 2023) is instrumental in reducing speeding and enhancing road safety by encouraging a proactive and systemic approach to address speeding-related crashes. It was adopted by the U.S. Department of Transportation (DOT) and several states in recent years as the guiding paradigm to address roadway safety. For the Safe System approach, enforcement plays a crucial role in establishing initial compliance with target speeds.

RESEARCH APPROACH

The project team used the citations and warning data collected through the FDOT District 7 ELEE project to assess the effects of law enforcement presence, frequency, and duration on vehicle speeds on study segments. The ELEE data (1,449 records or activities at the time of this research) includes the date, time frame of enforcement, state road, and nearest intersection or mile marker, as well as the speeding related citations and warnings issued during each police activity. Figure 1 shows the distribution of the data collected from February 2022 to August 2022 for the roadway segments in the five counties in FDOT District 7. A police activity in this study is defined as a one-time presence of a police officer on a road segment for two hours or more.

Based on the warning and citation information, the team obtained the vehicle speed information from the Regional Integrated Transportation Information System (RITIS) on roadway segments before, during, and after law enforcement presence. To investigate the instantaneous and lasting effects of police presence on vehicle speeds, a combined number of warnings and citations issued for speeding was used from the ELEE data along with the law enforcement segment location details. Using geographic information systems (GIS) and the location information in the ELEE data, the authors generated traffic management center codes (TMC codes) for law enforcement activities, which was utilized to later link citation and warning data issued for speeding to the vehicle running speed data in RITIS.

To prepare the vehicle speed data on the corridors, among the various intervals (1, 5, 15, 30, and 60 minutes) of aggregated speed available in RITIS, the authors used the 15-minute average speeds. The 15-minute average speeds were chosen as they provided an effective interval to show speeding issues with proper samples on a roadway. The five-minute interval or less sample size is too small, and the 30-minute or 1-hour interval is difficult for researchers to exam the average speeds of drivers with excessive vehicle speeds.

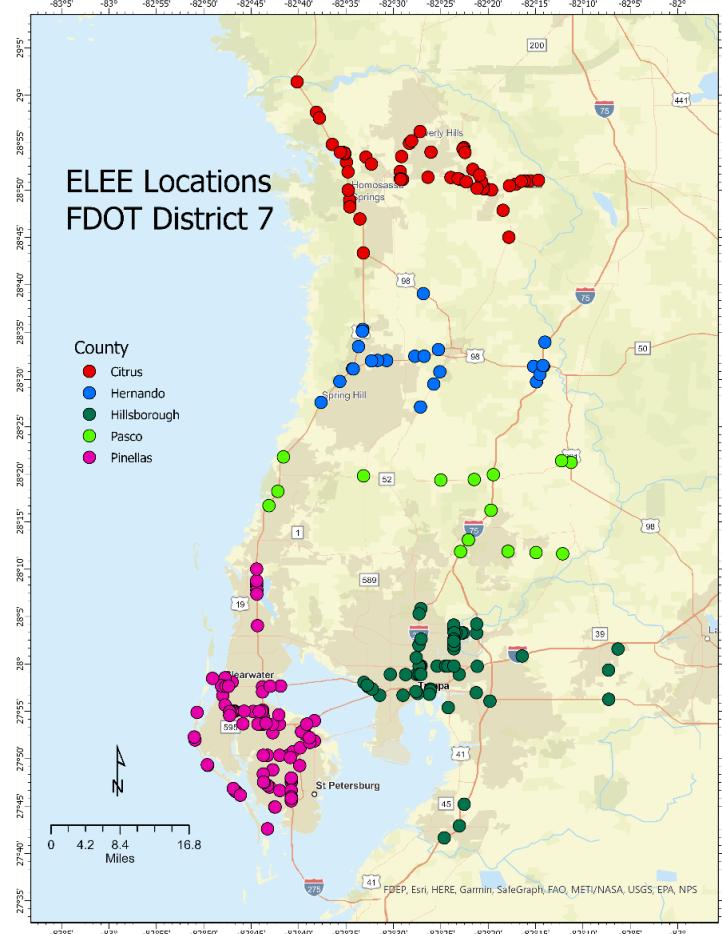


Figure 1. Law enforcement data distribution across counties in FDOT District 7.

Posted speed limits for all roadway segments were also incorporated into the speed dataset. Qualitative information such as travel direction, segment lengths, latitude, longitude, and TMC codes was also collected and merged with the speed data to create a comprehensive speed data profile. The speed data was then linked to the ELEE dataset using the unique TMC code for each record. The integrated dataset comprised the combined number of speeding related warnings and citations at each location during the law enforcement activity, along with the vehicle running speeds on that segment. In addition to those data preparation steps, data processing was essential for a more in-depth analysis. The initial format of the database was in a long format, resulting in multiple rows for different hours associated with a single record. This format posed challenges for analysis. Consequently, the data were transformed into a wide format, consolidating information into a single row for each record, facilitating a more straightforward analysis.

For the initial step, 11 roadway segments with the highest frequency of police activities (19 to 67 police activities per segment) were analysed. The segments generated a total of 124 activities out of which, a roadway segment of 20 activities experienced a more consistent reduction in speed during police presence. Second, the authors strategically selected 101 unique segments for analysis, which were selected based on both high frequency of police activities and high number of warnings and citations combined. After processing, those segments generated a total of 253 activities.

The data of the 253 activities was then divided into two distinct categories. The first category (Category 1) highlighted instances where the average speeds were noticeably higher than the posted speed limits in the absence of police presence, but decreased when police were present and issuing warning and citations. The second category (Category 2) comprised cases where speeds remained relatively consistent regardless of police presence. The outcome in the second category may be due to various conditions, including heavy traffic, poor weather with rain, two-lane roadways, many signals on a roadway segment, poor signal progression, work zone, etc., which limited the ability to speed. Among the 253 records, 49 belong to Category 1 (representing 10 roadway segments), and the other 204 records are in Category 2.

A focused analysis was conducted to investigate whether there was a substantial reduction in speed using the subset of 20 activities when just considering the roadway segments with the highest frequency of police activities and 49 activities when considering both high frequency of activities and high number of citations and warnings. This resulted in a total of 69 activities. For each of the 69 activities, the team evaluated the difference in speed between law enforcement presence and before and after period speed by not only considering the average vehicle speed, but also 5th, 25th, 75th, and 95th percentile speed. During the analysis, it was determined that the 95th percentile speeds served as a reliable representation of speeders. The 95th percentile was selected because it was the highest available in RITIS and it was difficult to detect roadways with speeding issues using the 75th percentile speeds.

After conducting basic statistical analysis based on the selected and available data, the average differences of 95th percentile speeds for (1) the three hours before and (2) during the police presence with active law enforcement activities of issuing warnings and citations shows speed reduction ranging from 2% to 7%. The other police activities with zero or minimal citations and warnings had limited effectiveness because of the lack of visibility of police presence during their operation. Three sites with 11 police activities were selected for detailed speed data analysis. This selection aimed to demonstrate various instances of how police activities affect vehicular speed. For the most part, many of the sites exhibited notable decreases in speed attributed to the police presence.

Several activities and locations were incorporated to illustrate the significance of police visibility (when a high number of citations and warnings are given) in influencing overall speed. Overall, the three sites that were selected proved to be excellent options for the detailed study due to their prominence as locations exhibiting either the highest levels of law enforcement activities or characterized by frequent police presence, or a notable frequency of warnings and citations issued. For the three study segments, the 95th percentile average speeds were examined three hours before and after police presence, as well as three days before and after police presence (during the same time period) to determine any changes in speed during police presence and activities on the corridors. During the evaluation, the corridor speed three hours before and three days before police presence is used as baseline speeds.

The aim for this research is to study the instantaneous and lasting effects of police activities for the three selected corridors. The lasting effects, for this study, mean sustained reductions in speeds beyond the immediate presence of police, which is still lower than the baseline speeds. The percentage reduction in speed was then calculated for each case. The corridors were also scrutinized for factors such as AM/PM peak hours, presence of signals, etc., to ensure that external variables did not influence speed variations. A detailed discussion of the findings for the three sites of 11 activities is in the next section.

RESULTS

The research team conducted a detailed analysis of speed data collected before, during, and after the presence of law enforcement. The detailed analysis results for each of the three roadway segments offer insights on the impact and lasting effects of law enforcement presence, frequency, and duration on vehicle speeds.

US-19

The first study segment of three activities lies on US-19 between Hexam Rd and Centralia Rd (see Figure 2). The image on Figure 2 only shows the section of the roadway segment considered in this study. There are no signals between the two intersections. The law enforcement officers were present on this two-mile long roadway segment on the following days and periods:

- July 22nd from 12:37 pm to 2:37 pm,
- July 27th from 2:15 pm to 4:15 pm, and
- July 29th from 2:00 pm to 4:00 pm.

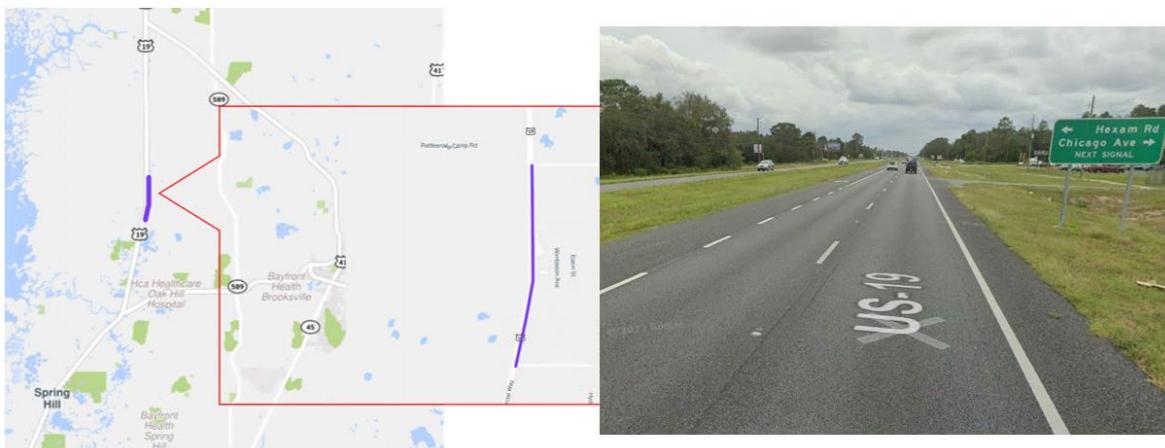


Figure 2. US-19 study location.

The results for the 95th percentile speeds before, during, and after the presence of law enforcement for this study segment are shown in Table 1 below.

As seen in Table 1, the average 95th percentile speed during police presence for all three days was lower than the speeds three days and three hours before and after police presence. A difference of 2 to 4 mph (3% to 7% reduction in speed) is observed when police officers are present. In general, the results show no lasting effect either a few hours or days after. Therefore, it can be determined that the police presence and activities did influence speeding on that roadway segment, but it was only an instantaneous effect. It should be noted that numerous citations and warnings were issued on that segment, which shows a consistent positive impact during police presence. However, the vehicle speeds after the law enforcement activities concluded immediately were back to levels similar to that of the baseline speed before the law enforcement presence, indicating no lasting effect. There was also no lasting effect for the same period one day, two days or three days after the law enforcement presence.

Table 1. Before-After Analysis based on 95th Percentile Speed for US-19

Speed Limit: 60 mph	22-Jul-22	27-Jul-22	29-Jul-22
Presence of Law Enforcement	12:37pm-2:37pm	2:15pm-4:15pm	2:00pm-4:00pm
Number of Citations and warnings	18	11	28
Average 3rd Hour Before	64.5	64.5	63.3
Average 2nd Hour Before	63.5	61.8	62.8
Average 1st Hour Before	62.0	63.0	62.3
3 Hour Average (Before)	63.3	63.1	62.8
Average During Presence	60.4	59.4	60.6
3 Hour Average (After)	63.7	63.1	64.6
Average 1st Hour After	62.5	61.3	65.0
Average 2nd Hour After	63.5	63.5	65.3
Average 3rd Hour After	65.0	64.5	63.5
Average 3 Days Before	63.8	65.6	63.3
Average 2 Day Before	63.3	62.1	60.8
Average 1 Day Before	62.0	63.6	63.4
3 Day Average (Before)*	63.0	63.7	62.5
Average During Presence	60.4	59.4	60.6
3 Day Average (After)*	63.8	63.1	63.9
Average 1 Day After	64.0	63.9	64.4
Average 2 Day After	65.0	60.6	64.0
Average 3 Day After	62.5	64.8	63.3

*Note: The 3 days before and 3 days after speeds were collected during the same hours as during the police presence. Speeds are in mph.

N 56th Street (FL-583)

The second study segment includes four activities, which is 1.5 miles long, lies on 56th Street (FL-583) between E Busch Blvd and Sligh Ave, as shown in Figure 3. The image illustrates the part of the roadway segment considered in the analysis. There are no signals between the two intersections. The police officers were present on this roadway segment during the following days and periods (they were present twice on April 23rd):

- April 22nd from 11:00 am – 1:30 pm,
- April 23rd from 1:00 am – 3:00 am, and 11:00 am – 1:10 pm, and
- April 24th from 1:30 am – 4:45 am.

The results for the 95th percentile speeds of before, during, and after the presence of law enforcement for this study are shown in Table 2.

For the N 56th Street segment, the results are mixed, as shown in Table 2. The effects of police presence are seen in some instances, but not for all. The mixed results may be partly due to the fact that less citations and warnings were given by police officers on that segment compared to the previous study segment (US-19). At daytime, the speeds before enforcement are occasionally lower than the speed limit, whereas during nighttime, the speeds before enforcement consistently exceed the speed limit. This may also be the reason limited citations and warnings were issued compared to first segment.

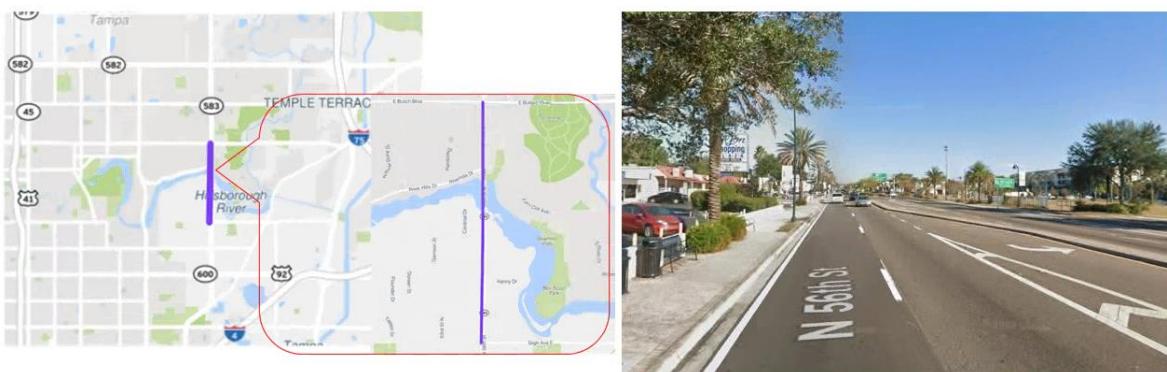


Figure 3. N 56th street study location.

Table 2. Before-After Analysis based on 95th Percentile Speed on N 56th Street

Speed Limit: 35 mph	April 22, 2022	April 23, 2022	April 23, 2022	April 24, 2022
Presence of Law Enforcement	11am-1:30pm	11am-1:10pm	1am-3am	1:30am-4:45 am
Citations and warnings	2	5	2	2
Average 3rd Hour Before	31.8	39.0	35.7	38.3
Average 2nd Hour Before	35.8	37.3	38.1	41.5
Average 1st Hour Before	35.5	38.0	36.3	43.3
3 Hour Average (Before)	34.3	38.1	36.7	41
Average During Presence	32.0	35.0	34.9	42.4
3 Hour Average (After)	30.9	32.8	35.2	44.6
Average 1st Hour After	33.0	32.5	34.8	43.5
Average 2nd Hour After	28.8	32.5	34.4	48.5
Average 3rd Hour After	31.0	33.3	36.5	41.8
Average 3 Days Before	34.1	33.6	44.4	43.9
Average 2 Day Before	33.6	33.4	42.3	45.4
Average 1 Day Before	33.3	32.1	33.0	43.1
3 Days Average (Before)*	33.7	33.0	39.9	44.1
Average During Presence	32.0	35.0	34.9	42.4
3 Days Average (After)*	35.5	33.8	33.1	44.8
Average 1 Day After	34.8	37.8	31.7	43.4
Average 2 Day After	37.8	31.7	33.7	46.1
Average 3 Day After	33.8	31.8	33.8	44.9

*Note: The 3 days before and 3 days after speeds were collected during the same hours as during the police presence.

From those results, it can be determined that police presence by itself may not be enough to influence the 95th percentile speed of a roadway segment if there are not enough citations and warnings issued. It may also be that for some occurrences, most groups of vehicles may have come at different times than the times the citations and warnings were given; thus, did not notice the police presence and continued to have higher speeds. A difference of up to 13% is observed

for the times where the 95th percentile speeds were lower during police presence than before (see 3-day average for before presence on April 23, 2022, in Table 2). In any case, no lasting effect of police presence was seen either a few hours or days after. Hence, the effect of police presence on this study segment was instantaneous.

SR-699

The third study segment of four activities lies on SR-699 between CR-694 and Park Blvd N, which is 3.8 miles long, as shown in Figure 4. The image on the right of Figure 4 is a screenshot of a section of the segment. There are no signals between the two intersections. This roadway was studied during police presence on March 1st, 2nd, and 4th. The police officers were present twice on March 1st. The specific days and periods are as follows:

- March 1st from 10:03 PM-12:45 AM and 12:12AM-2:21AM,
- March 2nd from 12:03AM-2:03AM, and
- March 4th from 8:02PM-10:26PM.

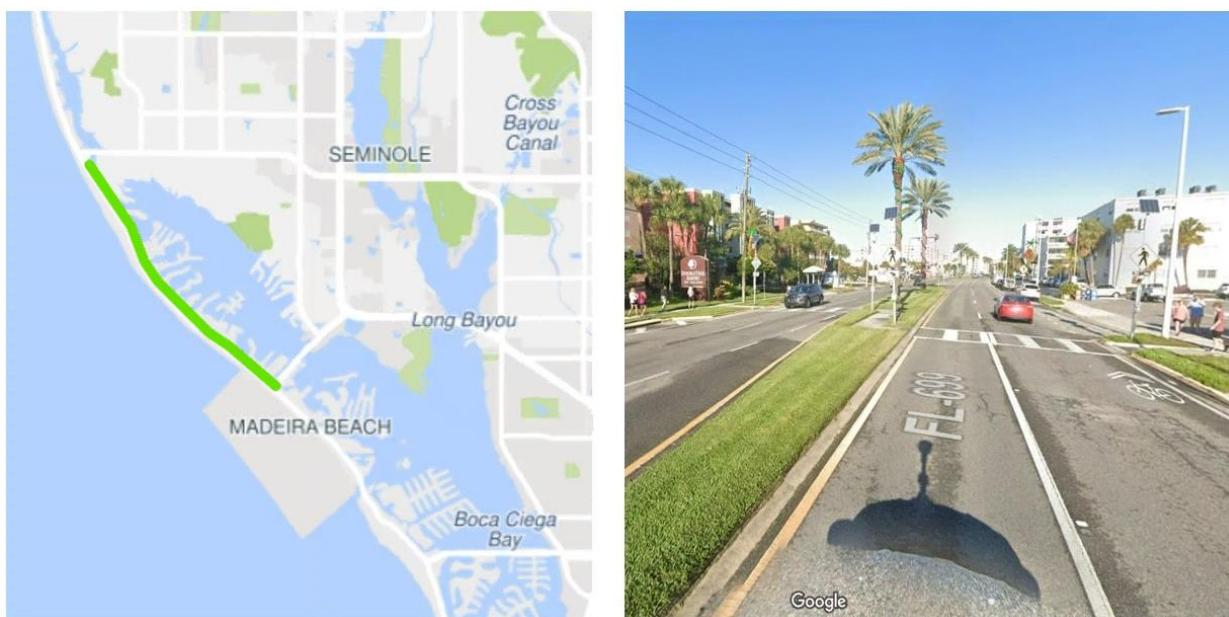


Figure 4. SR-699 study location.

The results for the 95th percentile speeds of before, during, and after the presence of law enforcement for this study segment are shown in Table 3.

The results for this nightly police presence are also mixed, as shown in Table 3. The speeds are mostly higher than the speed limit of 35 mph before and during enforcement, but not by much. Thus, the police presence had a slightly positive impact during those times and the differences in speed were up to 5% (see 3-day average before and average during presence on March 4th, 2022). However, it should be noted that one or no citations were issued for each of those times police officers were present. Thus, the effects of police presence are not consistently seen for those activities because police presence alone is not enough if citations and warnings are not given, or they are visible to drivers.

Multiple groups of vehicles could have traversed the roadway section without noticing the police presence because they remained largely concealed and were not as visible as when issuing warnings and citations. On the other hand, fewer citations and warnings would be issued if only a small number of drivers, driving at the 95th percentile speeds, were traveling at speeds significantly above the posted speed limit. The findings also reveal no lasting effects either a few days or hours after. Therefore, the impacts of police presence were also instantaneous on that segment.

Table 3. Before-After Analysis based on 95th Percentile Speed for SR-699

Speed Limit: 35 mph	March 1 st	March 1 st	March 2 nd	March 4 th
Presence of Law Enforcement	10:03pm-12:45am	12:12am-2:21am	12:03am-2:03am	8:02pm-10:26pm
Citations and warnings	1	0	0	0
Average 3rd Hour Before	31.5	35.0	35.5	31.0
Average 2nd Hour Before	34.8	37.0	37.8	30.8
Average 1st Hour Before	35.5	38.5	35.5	31.0
3 Hour Average (Before)	33.9	36.8	36.3	30.9
Average During Presence	36.1	35.7	36.0	33.4
3 Hour Average (After)	34.8	34.3	34.7	37.9
Average 1st Hour After	36.3	34.0	34.5	35.8
Average 2nd Hour After	34.0	34.0	34.0	39.0
Average 3rd Hour After	34.3	34.8	35.5	39.0
Average 3 Days Before	36.9	37.2	37.5	35.8
Average 2 Day Before	36.0	37.5	35.3	34.7
Average 1 Day Before	36.7	35.4	35.4	34.5
3 Day Average (Before)*	36.5	36.7	36.0	35.0
Average During Presence	36.1	35.7	36.0	33.4
3 Day Average (After)*	36.7	35.7	36.5	34.2
Average 1 Day After	36.5	35.7	35.8	32.8
Average 2 Day After	36.5	35.6	35.7	34.3
Average 3 Day After	37.2	36.0	38.0	35.6

*Note: The 3 days before and 3 days after speeds were collected during the same hours as during the police presence.

CONCLUSIONS

Speeding poses a significant safety challenge nationwide, including in Florida. Law enforcement represents one of the crucial techniques for controlling speed, thereby enhancing the safety of both motorists and non-motorists. Limited studies are available that evaluate the effects of repeated law enforcement activities on vehicle speeds. This study assessed the impacts of repeated police presence and the number of issued citations and warnings on vehicle speed. A case study approach was used to obtain the findings. An in-depth analysis was performed for three study segments on which police officers were either repeatedly present or issued the highest number of citations and warnings.

The research findings demonstrate the effectiveness of law enforcement in instantaneously reducing vehicle speeds during many instances of police presence. Except for the first study segment, which shows a consistent decrease in vehicle speeds when law enforcement was present and visible with many citations and warnings issued, the results for the last two cases are mixed. For those segments, the effectiveness of law enforcement activities is shown for some periods, but not for all. This may be because less citations and warnings are given by the law enforcement officers at those sites. This study's conclusions indicate that a visible law enforcement presence, characterized by issuing numerous citations and warnings, significantly influences motorists' awareness. Segment 1 demonstrates a substantial reduction in vehicle speeds due to this heightened visibility. Conversely, in segments 2 and 3, where law enforcement was present but not consistently visible to motorists, the observed impact on speed reduction was limited. It is necessary to note that the visibility of police presentation to drivers is important on vehicle speed reduction.

Overall, no lasting effects of law enforcement presence, frequency, and duration were observed in any of the cases. Vehicle speeds typically went quickly back to the speed similar as that of the base conditions (before periods) according to the before-after speed data analysis from this study. A similar finding was obtained in a previous study performed decades ago that concluded that as soon as drivers passed the location where police were present, they increased their speed (Sisiopiku and Patel, 1999). The results from the current study also reveal that police presence without being visible (not issuing citations and warnings) is not effective at reducing vehicular speed on corridors. Given the impracticality of continuous police presence across numerous locations requiring attention to speeding issues, the authors propose other strategies to supplement police presence and enforcement activities, including community education and awareness of speeding risks and the implementation of speed safety cameras where permitted, such as in school zones.

ACKNOWLEDGMENTS

This research project was sponsored and supported by the Florida Department of Transportation (FDOT) District 7. Authors would like to express sincere appreciation to Peter Hsu, District Safety Administrator, for providing the law enforcement citation and warning data for detailed analysis.

REFERENCES

- Dowling, K., and Holloman, E. (2008). The Effect of Conspicuous Traffic Enforcement on Speeding Behavior: A Study of Speed Reduction Response. *Int. Soc. Sci. Rev.* 83, 181–188.
- Elvik, R. (2005). Speed and Road Safety: Synthesis of Evidence from Evaluation Studies. *Transportation Research Record: Journal of the Transportation Research Board*, 1908(1), 59-69.
- Elvik, R. (2012). *Speed Limits, Enforcement, and Health Consequences*.
- FHWA (Federal Highway Administration). (2023). Speed Management Safety. Retrieved on 12/15/2023 from <https://highways.dot.gov/safety/speed-management>.
- Hu, W., and Mccartt, A. T. (2016). Effects of automated speed enforcement in Montgomery County, Maryland, on vehicle speeds, public opinion, and crashes on vehicle speeds, public opinion, and crashes. *Traffic Inj. Prev.* 17, 53–58.

- Kumfer, W., Martin, L., Turner, S., and Broshears, L. (2023). *Safe System Approach for Speed Management*. Federal Highway Administration (FHWA).
- Mountain, L. J., Hirst, W. M., and Maher, M. J. (2005). Are speed enforcement cameras more effective than other speed management measures ? The impact of speed management schemes on 30 mph roads 37, 742–754.
- Povey, L. J., Presenter, W. J. F., and Keall, M. D. (1996). An investigation of the relationship between speed enforcement, vehicle speeds, and injury crashes in New Zealand 1.
- Simpson, R., Frewing, Q., and Bayer, J. (2023). The Effects of Saturation Enforcement on Speed (ing) Along a Highway Corridor: Results from a Police-Directed Field Study. *Justice Evaluation Journal*, 6(1), 20-31.
- Sisiopiku, V. P., and Patel, H. (1999). Study of the Impact of Police Enforcement on Motorists' Speed. *Transportation Research Record: Journal of the Transportation Research Board*.
- Stanojevic, P., Jovanovic, D., and Lajunen, T. (2013). Influence of traffic enforcement on the attitudes and behavior of drivers. *Accid. Anal. Prev.* 52, 29–38.
- Stewart, T. (2023). Overview of Motor Vehicle Traffic Crashes in 2021. DOT HS 813 435. Washington, DC: NHTSA.
- WHO (World Health Organization). (2008). *Speed management: a road safety manual for decision-makers and practitioners*. Geneva: World Health Organization.

Causal Insights into Speeding Crashes

Subashish Das, Ph.D.¹; Ahmed Hossain²; Swastika Barua³; Shahrbanoo Kavianpour⁴;
and Abbas Sheykhfard, Ph.D.⁵

¹Assistant Professor, Ingram School of Engineering, Texas State Univ., San Marcos, TX.
ORCID: <https://orcid.org/0000-0002-1671-2753>. Email: subasish@txstate.edu

²Dept. of Civil Engineering, Univ. of Louisiana at Lafayette, Lafayette, LA.
ORCID: <https://orcid.org/0000-0003-1566-3993>. Email: ahmed.hossain1@louisiana.edu

³Ingram School of Engineering, Texas State Univ., San Marcos, TX. Email: qwx11@txstate.edu
⁴Dept. of Civil Engineering, Babol Noshirvani Univ. of Technology, Babol, Mazandaran, Iran.
ORCID: <https://orcid.org/0000-0003-1246-8535>. Email: Sh.kavianpour.civil@gmail.com

⁵Researcher, Dept. of Civil Engineering, Babol Noshirvani Univ. of Technology, Babol,
Mazandaran, Iran. ORCID: <https://orcid.org/0000-0002-9536-3108>.
Email: A.Sheykhfard@nit.ac.ir

ABSTRACT

Excessive speeding poses significant risks to road safety, impacting a driver's ability to maneuver safely around obstacles and leading to longer stopping distances and delayed reactions to hazardous situations. It is a major contributor to fatal and serious road trauma, accounting for over 20% of such incidents in the US. Eliminating speeding entirely could potentially substantially reduce fatal injuries by 20% or more. This study leveraged the comprehensive National Highway Traffic Safety Administration's (NHTSA) Fatality Analysis Reporting System (FARS) data, focusing on 50,081 speed-related motor vehicle traffic fatal crashes that occurred between 2016 and 2021, and a probabilistic graphical model to investigate the causal associations between key contributing factors involved in these speeding incidents. Ultimately, this high-impact research advanced our understanding of the risks posed by speeding and impaired driving, guiding the way for evidence-based interventions and transformative policies to build safer roads and protect all road users from preventable tragedies.

INTRODUCTION

The relationship between speed and crashes in road safety is a complex interplay of multiple factors. However, it is evident that as driving speeds increase, so does the crash rate. Moreover, individual vehicles traveling at higher speeds than other traffic on the same road are at a higher risk of crashes. These crashes also lead to more severe injuries for both the driver responsible for the crash and the other party involved. Speeding, encompassing exceeding speed limits or driving too fast for prevailing conditions, ranks among the most common factors contributing to motor vehicle crashes in the United States (U.S.). Various factors, such as road type, driver age, alcohol impairment, and roadway characteristics like curvature, grade, width, and adjacent land use influence the impact of speed on crash involvement.

In contrast, the link between speed and injury severity remains consistent and direct. Higher vehicle speeds result in more significant changes in velocity during a crash, directly correlating with the severity of injuries sustained. This correlation is particularly critical for vulnerable road users, such as pedestrians, who lack the protective structures of a vehicle.

The National Highway Traffic Safety Administration (NHTSA) categorizes drivers as speeding if their vehicle's speed falls into one of four categories: exceeding the speed limit, driving too fast for conditions, racing, or speeding with specific details unknown. Speed limits established by states for various road types define the first category. In contrast, the second category is based on the basic speed law, which requires drivers to operate at a reasonable and prudent speed given the prevailing environmental conditions. Between 2016 and 2021, speeding contributed to 50,081 traffic fatalities in the U.S., with a 26% increase in fatalities related to speeding observed from 2019 to 2021. The surge in speeding-related fatalities highlights the pressing need for a comprehensive analysis of speeding-related fatal crashes in the U.S. To address this, the current study utilized a probabilistic graphical method, Bayesian networks (BNs), to explore the causal relationships among key contributing factors. The findings of this research offer valuable insights for policymakers to implement more effective speed management strategies and interventions, reducing the prevalence and impact of speeding-related crashes.

By understanding the intricate relationships between speeding and crash outcomes, policymakers can devise evidence-based measures to promote safer driving behaviors and safeguard the well-being of all road users. This research is a powerful tool to inform the development of targeted policies and initiatives, fostering a culture of responsible and safe driving to protect lives and improve road safety across the nation.

LITERATURE REVIEW

Since higher speed is one of the major factors increasing crash risk and injury severities, there has been a lot of research in assessing the relationship between speeding and crash contributing factors related to driver, vehicle, roadway, and environmental characteristics. Some of the previous investigations explored speeding-related crashes by focusing on specific vehicle types. For instance, a Das et al. (2022) used crash data from Louisiana from 2010 to 2016 to identify the patterns associated with motorcycle crashes caused by speeding. Edwards et al. (2016) conducted interviews and ride-along observations to explore speeding cultures among drivers of heavy vehicles. The study found that truck drivers tended to consider speeding as generally safe, wanted to accelerate quickly, and did so to save time and earn more money.

Prior research has focused on driver demographics (age, gender) and behavior (seatbelt use, cellphone use, alcohol involvement) in crashes caused by speeding. For example, Se et al. (2023) conducted in Thailand explored the differences in crash-contributing factors affecting the injury severity of drivers (seatbelt-restrained and unrestrained) involved in speeding-related crashes. Using cellphone-involved single-vehicle crash datasets (2004 to 2019) from Pennsylvania, Wu et al. (2022) reported that combining cellphone usage with risky driving behaviors like speeding significantly increased the injury severity of drivers. The link between young drivers and speeding has also been explored by (Breen et al. 2020; Tay, 2005). For example, Ferguson, (2013) a national US-based report focused on speeding-related fatal crashes involving teen drivers aged 16-19. According to the investigation, speeding commonly results in single-vehicle and run-off-road crashes and is more common in males, at night, and when other teen passengers are present. (Bhalla et al. 2013; Høye, 2020; Romano et al. 2021; Ajay Kumar Yadav and Velaga, 2020) explored the role of drunk driving in speeding-related crashes. The common major finding from these studies was that an increase in Blood Alcohol Concentration (BAC) increases the likelihood of speeding-related crashes.

The crash risk associated with speeding can vary spatially, so some of the previous research addressed speeding-related crashes in specific locations. Yan et al. (2022) investigated speeding-related crashes on rural roadways focused on two distinct types of crashes – a) overturned and b) hitting fixed-object. The study identified some critical factors that present relative temporal stability, such as alcohol involvement, truck, aggressive driving, vehicle age (greater than 14 years old) and posted speed limit of less than 45 mph. Jiang et al. (2017) investigated overspeeding violation behaviors of drivers on expressways recorded by a global positioning system (GPS) enabled smartphone application. Kveladze and Agerholm (2020) identified speeding patterns on arterial roadways using Geovisual Analytics (GVA).

Prior studies have explored speeding-related crashes focusing on specific vehicle types, driver demographics, and behaviors. Compared to other models, BNs stand out for their ability to effectively capture complex relationships and dependencies among variables, making it a promising approach for understanding the intricate dynamics of speeding-related crashes more comprehensively and accurately.

METHODOLOGY

Data Collection

This study acquired fatal crash data for 6 years by the different types of speeding, including too fast for conditions, exceeded the speed limit, racing, and specifics unknown. Table 1 shows that over the six years from 2016 to 2021, the total number of recorded driving incidents increased from 8,334 cases in 2016 to 9,551 cases in 2021. Among the severity categories, incidents involving specifics unknown showed the most significant rise, increasing by 76.9%. Although the number of exceeded speed limit incidents remained relatively stable, there was a slight increase of approximately 0.8% by 2021. Incidents related to racing remained low and consistent throughout the years. However, incidents attributed to too-fast-for conditions saw a notable 9.6% increase.

Table 1. Fatalities by Year by Speeding Type

Year	Too Fast for Conditions	Exceeded Speed Limit	Racing	Specifics Unknown	Yearly Total
2016	3,742	3,496	58	1,038	8,334
2017	3,520	3,253	51	1,135	7,959
2018	3,473	3,056	51	1,054	7,634
2019	3,468	2,868	48	1,176	7,560
2020	3,685	3,628	104	1,626	9,043
2021	4,102	3,524	91	1,834	9,551
Grand Total	21,990	19,825	403	7,863	50,081

Table 2 shows the distribution of different variables considering different speeding types. The most common land use for all types was rural. When the specific functional system was known, other principal arterials were the most common for the too fast for conditions, exceeding the speed limit, and specifics unknown speeding types and minor arterials were the most common for racing. Crashes typically occurred on weekdays, during the daylight, and were

evenly split across all seasons. The most harmful event for too fast for conditions was a rollover or overturn, but for all other speeding types, it was motor vehicle in transport. Not collision with MVIT was the most common crash type. Drivers were usually 25 to 45 years old, not impaired by alcohol, had a valid driver's license, and were not ejected from the vehicle. Typically, no restraints were used. The roads usually had a straight alignment, a level road profile, and two lanes. Generally, the roadways were two-way undivided. The most common speed limit for too fast for conditions or specifics unknown was 50-55 MPH. For exceeding the speed limit and racing, however, the most common speed limit was 35-45 MPH. Cars were the most common body type, and usually, crashes involved single vehicles. Generally, no rollover occurred.

Table 2. Distribution of Variables Considering Different Speeding Types

Variable	Too Fast for Conditions N=21990	Exceeded Speed Limit N=19825	Racing N=403	Specifics Unknown N=7863	p-val.
rur_urb (Land Use)					.
Rural	11759 (53.5%)	8968 (45.2%)	75 (18.6%)	3155 (40.1%)	
Unknown	88 (0.40%)	62 (0.31%)	3 (0.74%)	38 (0.48%)	
Urban	10143 (46.1%)	10795 (54.5%)	325 (80.6%)	4670 (59.4%)	
func_sys (Functional System)					<0.001
Interstate	3305 (15.0%)	2085 (10.5%)	36 (8.93%)	1048 (13.3%)	
Major Collector	4028 (18.3%)	3562 (18.0%)	42 (10.4%)	1251 (15.9%)	
Minor Arterial	4138 (18.8%)	4423 (22.3%)	121 (30.0%)	1435 (18.3%)	
Other	5982 (27.2%)	5063 (25.5%)	86 (21.3%)	2233 (28.4%)	
Principal Arterial - Other	4537 (20.6%)	4692 (23.7%)	118 (29.3%)	1896 (24.1%)	
day_week (Day of the Week)					0.014
Weekday	13782 (62.7%)	12121 (61.1%)	246 (61.0%)	4865 (61.9%)	
Weekend	8208 (37.3%)	7704 (38.9%)	157 (39.0%)	2998 (38.1%)	
Season (Season)					<0.001
Autumn	5529 (25.1%)	5029 (25.4%)	97 (24.1%)	2201 (28.0%)	
Spring	5195 (23.6%)	5050 (25.5%)	118 (29.3%)	1887 (24.0%)	
Summer	5797 (26.4%)	5874 (29.6%)	118 (29.3%)	2252 (28.6%)	
Winter	5469 (24.9%)	3872 (19.5%)	70 (17.4%)	1523 (19.4%)	
lgt_cond (Lighting Condition)					<0.001
Dark - Lighted	3747 (17.0%)	4355 (22.0%)	148 (36.7%)	1980 (25.2%)	
Dark - Not Lighted	6542 (29.7%)	6237 (31.5%)	86 (21.3%)	2168 (27.6%)	
Dawn	402 (1.83%)	315 (1.59%)	1 (0.25%)	122 (1.55%)	
Daylight	10487 (47.7%)	8111 (40.9%)	145 (36.0%)	3263 (41.5%)	
Dusk	559 (2.54%)	550 (2.77%)	17 (4.22%)	210 (2.67%)	
Other	253 (1.15%)	257 (1.30%)	6 (1.49%)	120 (1.53%)	
m_harm (Most Harmful Event)					.

Fire/Explosion	417 (1.90%)	480 (2.42%)	9 (2.23%)	215 (2.73%)	
Guardrail Face	447 (2.03%)	361 (1.82%)	8 (1.99%)	181 (2.30%)	
Motor Vehicle In-Transport	6107 (27.8%)	5725 (28.9%)	155 (38.5%)	2340 (29.8%)	
Other	3573 (16.2%)	2880 (14.5%)	55 (13.6%)	1374 (17.5%)	
Parked Motor Vehicle	366 (1.66%)	249 (1.26%)	10 (2.48%)	164 (2.09%)	
Pedalcyclist	4 (0.02%)	3 (0.02%)	0 (0.00%)	1 (0.01%)	
Pedestrian	10 (0.05%)	13 (0.07%)	1 (0.25%)	29 (0.37%)	
Rollover/Overturn	6201 (28.2%)	4838 (24.4%)	66 (16.4%)	1681 (21.4%)	
Tree (Standing Only)	3906 (17.8%)	4173 (21.0%)	76 (18.9%)	1445 (18.4%)	
Utility Pole/Light Support	959 (4.36%)	1103 (5.56%)	23 (5.71%)	433 (5.51%)	
man_coll (Collision Type)					.
Angle	2512 (11.4%)	3039 (15.3%)	80 (19.9%)	1074 (13.7%)	
Front-to-Front/Rear	2977 (13.5%)	2435 (12.3%)	38 (9.43%)	1128 (14.3%)	
Not Collision with MVIT	15877 (72.2%)	13679 (69.0%)	218 (54.1%)	5360 (68.2%)	
Other	94 (0.43%)	85 (0.43%)	7 (1.74%)	49 (0.62%)	
Rear-end	49 (0.22%)	33 (0.17%)	1 (0.25%)	9 (0.11%)	
Sideswipe	481 (2.19%)	554 (2.79%)	59 (14.6%)	243 (3.09%)	
Age (Driver Age)					.
25-45 years	8819 (40.1%)	9334 (47.1%)	153 (38.0%)	3529 (44.9%)	
46-65 years	5001 (22.7%)	3238 (16.3%)	37 (9.18%)	1488 (18.9%)	
Children	370 (1.68%)	240 (1.21%)	5 (1.24%)	108 (1.37%)	
Infant	64 (0.29%)	43 (0.22%)	0 (0.00%)	16 (0.20%)	
Older than 65 years	1843 (8.38%)	723 (3.65%)	5 (1.24%)	542 (6.89%)	
Unknown	27 (0.12%)	29 (0.15%)	1 (0.25%)	18 (0.23%)	
Young	5866 (26.7%)	6218 (31.4%)	202 (50.1%)	2162 (27.5%)	
Drinking (Driver Impairment)					<0.001
No	6273 (28.5%)	5943 (30.0%)	113 (28.0%)	2020 (25.7%)	
Unknown	11133 (50.6%)	8619 (43.5%)	222 (55.1%)	4120 (52.4%)	
Yes	4584 (20.8%)	5263 (26.5%)	68 (16.9%)	1723 (21.9%)	
Ejection (Driver Ejection)					<0.001
Ejected	5464 (24.8%)	5697 (28.7%)	106 (26.3%)	1840 (23.4%)	
Not Ejected	11934 (54.3%)	9548 (48.2%)	207 (51.4%)	4186 (53.2%)	
Unknown	4592 (20.9%)	4580 (23.1%)	90 (22.3%)	1837 (23.4%)	
l_status (Driver Licensing Status)					.
Canceled or denied	62 (0.28%)	112 (0.56%)	2 (0.50%)	27 (0.34%)	
Expired	506 (2.30%)	321 (1.62%)	11 (2.73%)	193 (2.45%)	
Not licensed	1895 (8.62%)	1899 (9.58%)	73 (18.1%)	903 (11.5%)	
Revoked	401 (1.82%)	634 (3.20%)	3 (0.74%)	167 (2.12%)	
Suspended	2127 (9.67%)	2437 (12.3%)	35 (8.68%)	787 (10.0%)	

Unknown License Status	263 (1.20%)	221 (1.11%)	10 (2.48%)	137 (1.74%)	
Valid	16736 (76.1%)	14201 (71.6%)	269 (66.7%)	5649 (71.8%)	
rest_use (Restraint Usage)					.
Child Restraint	135 (0.61%)	62 (0.31%)	0 (0.00%)	21 (0.27%)	
Helmet	1505 (6.84%)	1415 (7.14%)	26 (6.45%)	447 (5.68%)	
None used	12160 (55.3%)	12438 (62.7%)	191 (47.4%)	4397 (55.9%)	
Restraint Used	6507 (29.6%)	4370 (22.0%)	109 (27.0%)	2001 (25.4%)	
Unknown	1683 (7.65%)	1540 (7.77%)	77 (19.1%)	997 (12.7%)	
Valign (Alignment)					.
Curve	10180 (46.3%)	7719 (38.9%)	108 (26.8%)	2633 (33.5%)	
Driveway Access	19 (0.09%)	7 (0.04%)	0 (0.00%)	6 (0.08%)	
Straight	11345 (51.6%)	12001 (60.5%)	286 (71.0%)	5101 (64.9%)	
Unknown	446 (2.03%)	98 (0.49%)	9 (2.23%)	123 (1.56%)	
vnum_lan (Number of Lanes)					<0.001
Five lanes	1253 (5.70%)	1045 (5.27%)	34 (8.44%)	404 (5.14%)	
Four lanes	1599 (7.27%)	1727 (8.71%)	77 (19.1%)	772 (9.82%)	
Other	1357 (6.17%)	842 (4.25%)	31 (7.69%)	415 (5.28%)	
Three lanes	1904 (8.66%)	1845 (9.31%)	69 (17.1%)	925 (11.8%)	
Two lanes	15877 (72.2%)	14366 (72.5%)	192 (47.6%)	5347 (68.0%)	
Vprofile (Road Profile)					0.000
Downhill	2190 (9.96%)	1783 (8.99%)	11 (2.73%)	299 (3.80%)	
Grade, Unknown Slope	2983 (13.6%)	2169 (10.9%)	49 (12.2%)	1128 (14.3%)	
Level	12734 (57.9%)	13499 (68.1%)	273 (67.7%)	5138 (65.3%)	
Other	1885 (8.57%)	1848 (9.32%)	25 (6.20%)	419 (5.33%)	
Unknown	2198 (10.00%)	526 (2.65%)	45 (11.2%)	879 (11.2%)	
vspd_lim (Speed Limit)					.
30 MPH and Lower	1943 (8.84%)	2840 (14.3%)	57 (14.1%)	1229 (15.6%)	
35-45 MPH	7029 (32.0%)	8414 (42.4%)	207 (51.4%)	2475 (31.5%)	
50-55 MPH	7920 (36.0%)	5918 (29.9%)	77 (19.1%)	2682 (34.1%)	
65-70 MPH	3717 (16.9%)	1931 (9.74%)	40 (9.93%)	860 (10.9%)	
75-85 MPH	732 (3.33%)	311 (1.57%)	5 (1.24%)	216 (2.75%)	
No Statutory Limit	92 (0.42%)	24 (0.12%)	2 (0.50%)	18 (0.23%)	
Unknown	557 (2.53%)	387 (1.95%)	15 (3.72%)	383 (4.87%)	
Vtrafway (Trafficway Type)					<0.001
Other	1636 (7.44%)	1181 (5.96%)	37 (9.18%)	640 (8.14%)	
Two-Way Divided Median Barrier	3005 (13.7%)	2252 (11.4%)	70 (17.4%)	1175 (14.9%)	
Two-Way Divided Unprotected Barrier	2717 (12.4%)	2888 (14.6%)	93 (23.1%)	1028 (13.1%)	
Two-Way Undivided	13724 (62.4%)	12600 (63.6%)	170 (42.2%)	4691 (59.7%)	

Two Way TWLCL	908 (4.13%)	904 (4.56%)	33 (8.19%)	329 (4.18%)	
body_typ (Vehicle Body Type)					.
ATV	369 (1.68%)	101 (0.51%)	6 (1.49%)	128 (1.63%)	
Bus	39 (0.18%)	7 (0.04%)	0 (0.00%)	3 (0.04%)	
Car	9000 (40.9%)	9344 (47.1%)	231 (57.3%)	3500 (44.5%)	
Light Truck	3347 (15.2%)	2467 (12.4%)	21 (5.21%)	877 (11.2%)	
Motorcycle	4497 (20.5%)	4481 (22.6%)	87 (21.6%)	1790 (22.8%)	
Motorhome	11 (0.05%)	2 (0.01%)	0 (0.00%)	2 (0.03%)	
Other	319 (1.45%)	114 (0.58%)	10 (2.48%)	110 (1.40%)	
Truck	665 (3.02%)	188 (0.95%)	2 (0.50%)	140 (1.78%)	
Utility	3225 (14.7%)	2778 (14.0%)	44 (10.9%)	1150 (14.6%)	
Van	518 (2.36%)	343 (1.73%)	2 (0.50%)	163 (2.07%)	
Rollover (Rollover Type)					<0.001
No Rollover	14507 (66.0%)	13034 (65.7%)	281 (69.7%)	5558 (70.7%)	
Other	646 (2.94%)	500 (2.52%)	11 (2.73%)	181 (2.30%)	
Rollover, Tripped by Object/Vehicle	5667 (25.8%)	5316 (26.8%)	99 (24.6%)	1763 (22.4%)	
Rollover	1170 (5.32%)	975 (4.92%)	12 (2.98%)	361 (4.59%)	
ve_total (Vehicles Involved)					<0.001
Multiple	7515 (34.2%)	7230 (36.5%)	214 (53.1%)	3039 (38.6%)	
Single	14475 (65.8%)	12595 (63.5%)	189 (46.9%)	4824 (61.4%)	

Bayesian Network

A Bayesian Network (BN) or belief network is a probabilistic graphical model used to represent and reason uncertain relationships between variables. It is based on Bayesian probability theory, which allows for incorporating prior knowledge and updating probabilities as new evidence is obtained. The network consists of nodes representing variables and directed edges representing probabilistic dependencies between the variables. There are several data-driven techniques available, including Naive Bayesian Networks (NBN), Augmented naive Bayesian Networks (ABN), and Tree-augmented naive Bayes Networks (TAN). TAN learning specifically constructs qualitative BNs that capture the interactive relationships among risk influential factors (RIFs). This approach proves valuable for extracting valuable insights into crucial human variables that play a role in different types of crashes, such as those related to speeding. Additionally, research by Friedman et al. (1997) highlighted that TAN outperforms NBN in terms of predictive accuracy while still maintaining computational simplicity and robustness. A study by (Howard and Matheson, 1984; Pearl, 1988) discussed an influence diagram, also known as a Bayesian belief net or decision diagram, depicts the factors that contribute to a particular conclusion or uncertainty within the context of speeding types in fatal crashes. Moreover, Shachter (2007) emphasized that the influence diagram allows for the incorporation of decisions, enabling decision-makers to comprehend how each alternative impacts the probability of a specific outcome.

RESULTS AND DISCUSSIONS

Several steps were involved in developing the BN model in this study. Firstly, Netica 6.04 (a user-friendly software specifically designed for BN applications) adopted by Norsys (2020) was used to construct a DAG representing the model. DAG is a causal network in which a node represents each variable's conditional probability distribution (CPD), and the edges show how those nodes are related to one another. To fit the network to the input-output combinations obtained from the Monte Carlo (MC) runs using the model, the Lauritzen (1995) expectation maximization (EM) technique was used. The BN model parameters (conditional probability distributions) were estimated through the learning process. These distributions provide information about the likelihood of different states within each node of the BN. Figure 1 illustrates the layout of the final BN network. The relationships between different nodes are depicted in this figure. Each node's conditional probability tables (CPTs) are represented as belief bars and reflect the probabilities associated with the various states of that node.

Figure 1 presents the initial data on different speeding types of vehicle crashes, modeled through a BN. The most frequently contributing speeding types to fatal crashes are too fast for conditions (43.9%) and exceeded speed limit (39.6%). Among fatal crashes, 51.8% occur in urban areas, with a majority on two-way undivided (62.3%) and two-lane (71.4%) roadways. The most common functional system is 'other' (26.7%) or an 'other' form of principle arterial (22.4%). Considering vehicle types and drivers' perspectives, it can be observed that drivers between 25 and 45 years old (44.1%), holding a valid driving license (73.6%), and driving a car (44.1%) tend to be involved in fatal crashes most often. Typically, no restraint is used (58.3%). Usually, drinking is either unknown (48.1%) or not involved (28.7%). These crashes predominantly occur when driving on straight (57.4%), leveled (63.2%), two-lane (71.4%), and two-way undivided (62.3%) roads in urban areas (51.8%). The most common speed limit is 35 to 45 MPH. The most common collision type is not a collision with an MVIT (70.2%). The most harmful event was often motor vehicle in transport. Usually, no rollover occurred, and the driver was generally not ejected. Usually, crashes involved a single vehicle (64.1%). These crashes often occur during the weekday (61.7%), in the daylight (43.9%), and are relatively evenly split across all seasons.

Counterfactual Scenarios

Figure 2 illustrates a counterfactual scenario considering all crashes to have been too fast for the conditions. There was an increase in crashes that occurred in rural areas (by 5.7%), making it the most dominant category. Siskind et al. (2011) study identified the over-representation of fatal crashes in rural areas where the drivers had been travelling at or below the posted speed limit, but too fast for the prevailing condition. Afghari et al. (2018) discussed the intuitive relationship between 'driving too fast for conditions' and 'curved alignment' reveals a critical safety concern in such settings. Curved roads often limit the driver's line of sight.

Figure 3 shows a counterfactual scenario considering all crashes occurred with a speeding type of exceeded the speed limit. Urban areas were again the most common location for these crashes, although it only increased by 2.7% compared to the original data. The relationship between exceeding the speed limit and urban area type is well documented in previous research (Heydari et al. 2014; Hu and Cicchino, 2020; Moradi et al. 2013; Pérez-Acebo et al. 2021; Yannis et al. 2013). Urban areas are often characterized by heavy traffic congestion, especially

during peak hours. Drivers may feel rushed, impatient, or stressed due to time constraints, leading them to speed to reach their destinations faster. A 35 to 45 MPH speed limit became the dominant category, increasing by 6.2%. There was also a slight increase in crashes that occurred on level roads of 4.9%. Interestingly, there was no notable change in any of the other variables.

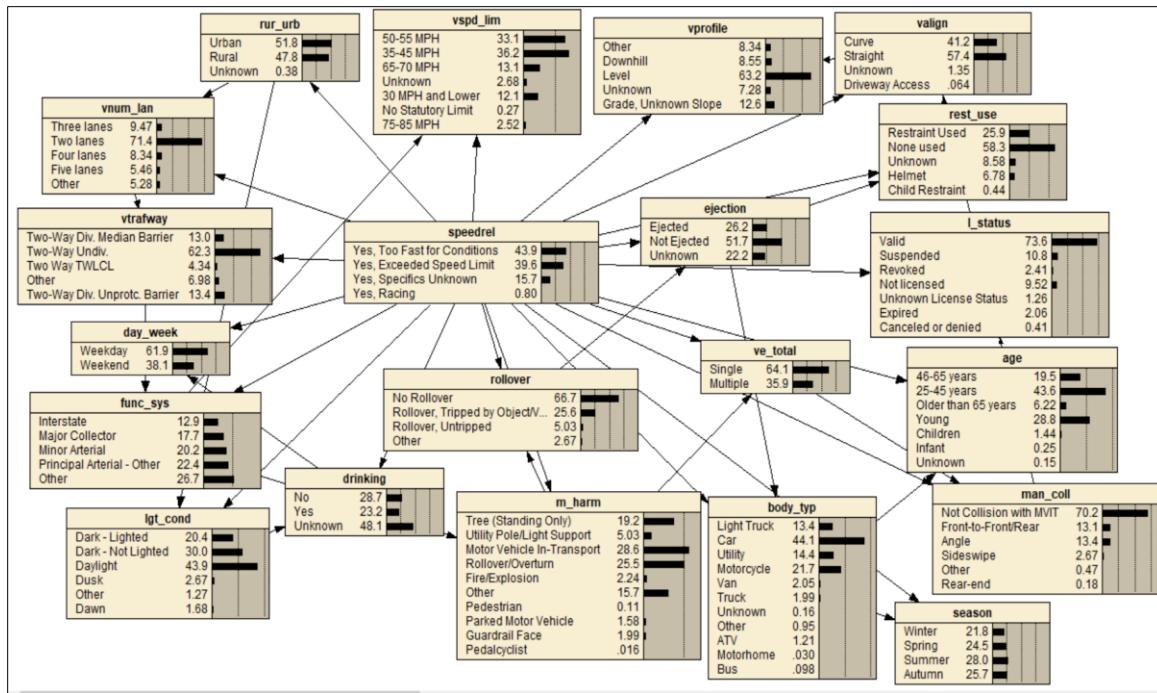


Figure 1. BN of full data

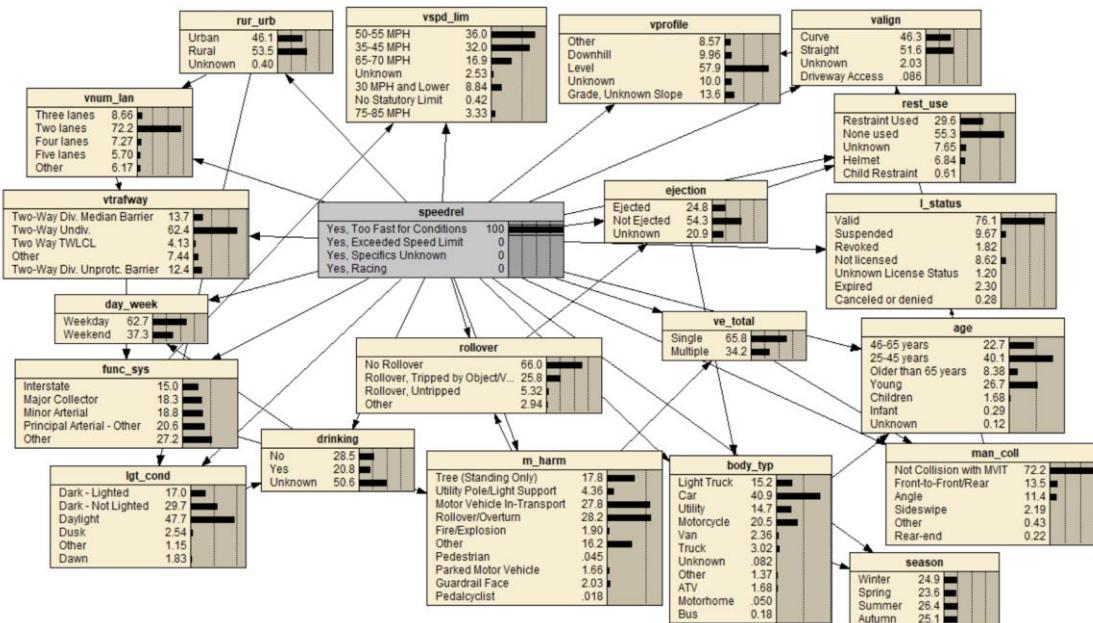


Figure 2. BN of counterfactual scenario considering all crashes occurred with the speeding type of too fast for conditions.

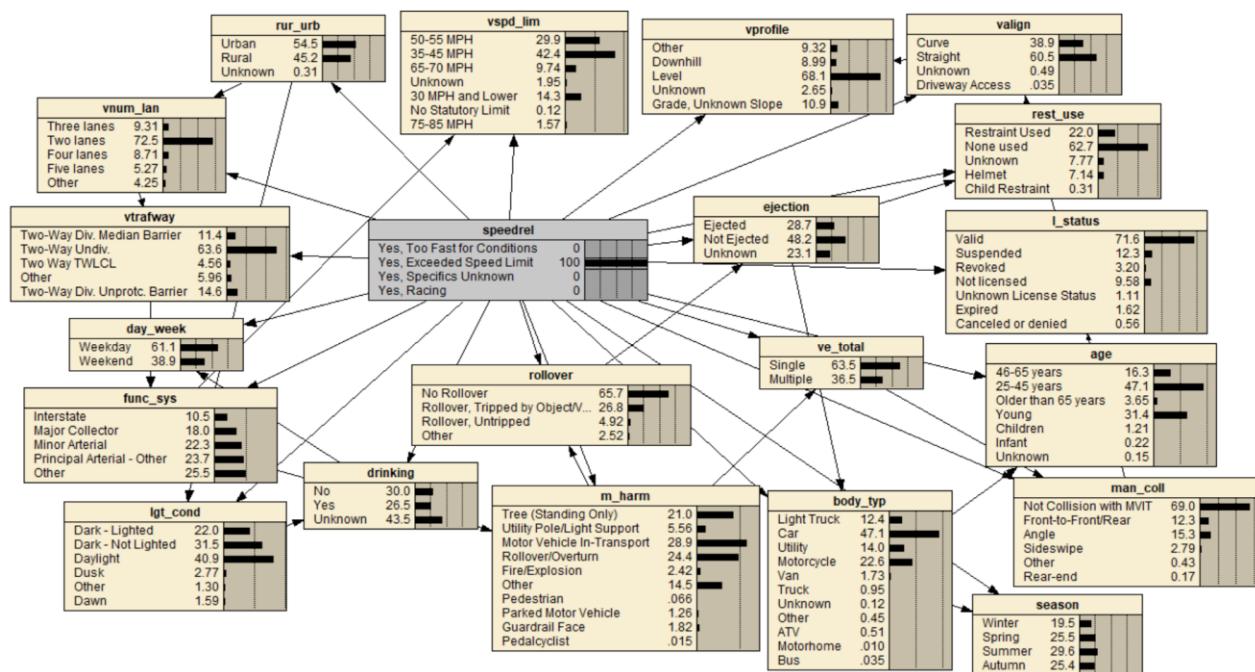


Figure 3. BN of counterfactual scenario considering all crashes occurred with a speeding type of exceeded speed limits

CONCLUSIONS

Speeding is one of the most prevalent factors in motor vehicle crashes in the United States. This study provides valuable insights into fatal crashes related to different types of speeding over six years from 2016 to 2021. A BN analysis uncovered the intricate relationships between speeding and crash outcomes. The initial data highlights that fatal crashes' most common speeding types were too fast for conditions and exceeded the speed limit. Urban areas accounted for slightly over half of fatal crashes, and drivers aged 25 to 45 with valid licenses were most involved. Counterfactual scenarios showed that altering specific variables could lead to notable shifts in crash patterns. For instance, increasing the speed limit to 35-45 MPH resulted in more urban crashes. In comparison, a speed limit of 75 to 85 MPH led to an increase in crashes involving driving too fast for conditions and a substantial increase in crashes on interstates.

The study's unique contributions involved using a BN analysis to understand the complex relationships between speeding and crash outcomes, offering insights into speeding types, road types, and driver characteristics related to fatal crashes. The research also employed counterfactual scenarios to assess the potential impact of interventions on crash patterns, aiding evidence-based policy decisions. The study findings also highlighted the need for speed management strategies, road infrastructure improvements, targeted enforcement and education, and comprehensive road safety initiatives to reduce speeding-related crashes and fatalities on U.S. roads.

The study's limitations include the focus on fatal crashes, potentially missing non-fatal incidents, and excluding combined speeding behaviors. Furthermore, improved data recording and reporting practices may influence the observed trends. BN modeling's accuracy may be impacted by changes in driving behavior or infrastructure over time. Future research should explore the effectiveness of intervention strategies and adopt a broader perspective on road

safety, considering injuries and property damage. Regular updates to the model with recent data are crucial. Addressing these limitations can enhance road safety measures and guide evidence-based policy decisions.

REFERENCES

- Afghari, A. P., Haque, M. M., and Washington, S. 2018. Applying fractional split model to examine the effects of roadway geometric and traffic characteristics on speeding behavior. *Traffic injury prevention* 19, 860–866.
- Bhalla, K., Li, Q., Duan, L., Wang, Y., Bishai, D., and Hyder, A. A. 2013. The prevalence of speeding and drunk driving in two cities in China: a mid project evaluation of ongoing road safety interventions. *Injury* 44, pp-S49-S56.
- Breen, J. M., Næss, P. A., Hansen, T. B., Gaarder, C., and Stray-Pedersen, A. 2020. Serious Motor Vehicle Collisions Involving Young Drivers on Norwegian Roads 2013–2016: Speeding and Driver-Related Errors Are the Main Challenge. *Traffic Injury Prevention* 21, pp-382-388.
- Das, S., Mousavi, S. M., and Shirinzad, M. 2022. Pattern recognition in speeding related motorcycle crashes. *Journal of Transportation Safety & Security* 14, pp-1121-1138.
- Edwards, J., Davey, J., and Armstrong, K. 2016. Safety culture and speeding in the Australian heavy vehicle industry. *Journal of the Australasian College of Road Safety* 27, 18–26.
- Ferguson, S. A. 2013. *Speeding-Related Fatal Crashes Among Teen Drivers and Opportunities for Reducing the Risks (Digital/other)*.
- Friedman, N., Geiger, D., and Goldszmidt, M. 1997. Bayesian Network Classifiers. *Machine Learning* 29, 131–163.
- Heydari, S., Miranda-Moreno, L. F., and Liping, F. 2014. Speed limit reduction in urban areas: A before–after study using Bayesian generalized mixed linear models. *Accident Analysis & Prevention* 73, pp 252-261.
- Howard, R., and Matheson, J. 1984. Influence diagrams. *the principles and applications of decision analysis*. Strategic Decisions Group.
- Høye, A. 2020. Speeding and Impaired Driving in Fatal Crashes—Results from In-Depth Investigations. *Traffic Injury Prevention* 21, pp-425-430.
- Hu, W., and Cicchino, J. B. 2020. Lowering the speed limit from 30 mph to 25 mph in Boston: effects on vehicle speeds. *Injury prevention* 26, 99–102.
- Jiang, Y., Zhang, J., Chikaraishi, M., Seya, H., and Fujiwara, A. 2017. Effects of a GPS-enabled Smart Phone App with Functions of Driving Safety Diagnosis and Warning Information Provision on Over-speeding Violation Behavior on Expressways. Presented at the *Transportation Research Procedia*, Elsevier, p. pp-1815-1823.
- Kveladze, I., and Agerholm, N. 2020. GeoVisual analytics for understanding the distribution of speeding patterns on arterial roads: assessing the traffic safety of vulnerable road users. *Journal of Location Based Services* 14, pp-201-230.
- Lauritzen, S. L. 1995. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis* 19, 191–201.
- Moradi, A., Motevalian, S. A., Mirkoohi, M., McKay, M. P., and Rahimi-Movaghhar, V. 2013. Exceeding the speed limit: prevalence and determinants in Iran. *International journal of injury control and safety promotion* 20, 307–312.

- Norsys. 2020. Netica 6.04 [WWW Document]. URL <https://www.norsys.com/netica.html> (accessed 7.18.23).
- Pearl, J. 1988. Morgan Kaufmann series in representation and reasoning. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann, San Mateo, CA, US.
- Pérez-Acebo, H., Otxoa-Muñoz, X., Marquina-Llaguno, M., and Gonzalo-Orden, H. 2021. Analysis of the efficiency of traffic lights turning red in case of exceeding speed limit. *Ingeniería e Investigación* 41.
- Romano, E., Fell, J. C., Li, K., Simons-Morton, B. G., and Vaca, F. E. 2021. Alcohol- and speeding-related fatal crashes among novice drivers age 18–20 not fully licensed at the time of the crash. *Drug and Alcohol Dependence* 218, 108417.
- Se, C., Champahom, T., Wisutwattanasak, P., Jomnonkwo, S., and Ratanavaraha, V. 2023. Temporal instability and differences in injury severity between restrained and unrestrained drivers in speeding-related crashes. *Sci Rep* 13, 9756.
- Shachter, R. 2007. Model building with belief networks and influence diagrams. *Adv Decis Anal Found Appl* 177–207.
- Siskind, V., Steinhardt, D., Sheehan, M., O'Connor, T., and Hanks, H. 2011. Risk factors for fatal crashes in rural Australia. *Accident Analysis & Prevention* 43, 1082–1088.
- Tay, R. 2005. The effectiveness of enforcement and publicity campaigns on serious crashes involving young male drivers: Are drink driving and speeding similar? *Accident Analysis & Prevention* 37, 922–929.
- Wu, P., Song, L., and Meng, X. 2022. Temporal analysis of cellphone-use-involved crash injury severities: Calling for preventing cellphone-use-involved distracted driving. *Accident Analysis & Prevention* 169, 106625.
- Yadav, A. K., and Velaga, N. R. 2020. Alcohol-Impaired Driving in Rural and Urban Road Environments: Effect on Speeding Behaviour and Crash Probabilities. *Accident Analysis & Prevention* 140, 105512.
- Yan, X., He, J., Wu, G., Zhang, C., Wang, C., and Ye, Y. 2022. Differences of overturned and hit-fixed-object crashes on rural roads accompanied by speeding driving: accommodating potential temporal shifts. *Analytic Methods in Accident Research* 35, 100220.
- Yannis, G., Louca, G., Vardaki, S., and Kanellaidis, G. 2013. Why do drivers exceed speed limits. *European Transport Research Review* 5, 165–177.

Extended Runtime at Signalized Intersection Using Backup Power after the Loss of Utility Power

Pei-Sung Lin, Ph.D., P.E.¹; Achilleas Kourtellis, Ph.D.²; and Zhenyu Wang, Ph.D.³

¹Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: lin@usf.edu

²Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: kourtellis@usf.edu

³Center for Urban Transportation Research, Univ. of South Florida, Tampa, FL.
Email: zwang9@usf.edu

ABSTRACT

Lessons learned after hurricanes revealed that although most traffic signal infrastructure withstood the impacts of the storms, signalized intersections were not operational due to lack of utility power. Current traffic cabinets employ a standard uninterrupted power supply (UPS) system, which is powered by lead acid batteries, an outdated technology that provides power for only up to 8 h of operation. For this study, a liquid propane generator was tested successfully to provide constant power without issues. A system that included a battery + solar backup was tested successfully to provide power for several days without issues. This system is not usually an issue after hurricanes, especially if adequate solar energy generation is provided to charge the battery during available daytime hours. The result of the testing provides options for transportation agencies for extended runtime at signalized intersections using backup power after the loss of utility power.

INTRODUCTION

Lessons learned from recent hurricanes revealed that although most traffic signal infrastructure withstood the impacts of the storms, signalized intersections were not operational due to lack of utility power. Because of this and problems associated with the use of generators to run intersections, such as purchase and replacement costs, storage issues, and costs of placement in the field after a storm, the subject needs further research for viable solutions.

Due to the large number of hurricanes and other storms experienced in Florida and their impact on Florida's transportation system, the Florida Department of Transportation (FDOT) conducted a research project (Lin et al., 2021) to investigate, test, and evaluate selected backup power systems. FDOT strives to be the nation's leader regarding extended or self-sustaining power at signalized intersections. The advantages of having a signalized intersection continue to operate during long hours of power outage during or after hurricanes could be substantial.

Implementation of backup power systems on a wide scale is needed, especially for traffic signals on major arterials or at major signalized intersections. Resources needed for implementation may be high for new signalized intersections and low for existing ones. Lack of a solution for this problem means continuing to have non-operational (dark) signalized intersections during normal daily traffic after loss of utility power during and after a storm. Currently, a dark intersection is treated as a four-way stop, which contributes to traffic congestion and safety concerns.

Study Objectives. The primary objective of this research was to improve safety and efficiency at signalized intersections during utility power loss during or after hurricanes. This study investigated the most appropriate methods for developing an extended runtime signalized intersection that would operate for a minimum of three days (72 hours) using backup power. The original goal was to provide power for five days (120 hours), but this was later modified due to the high costs of the available systems. Other areas of interest for the study were technologies that would lengthen battery backup runtime such as solar assist or other energy sources. Implementation is expected to be applied at new intersections and a retrofit solution that could be used with existing 120 VAC-powered intersections so benefits can be experienced immediately for many signalized intersections.

BACKGROUND

The objective of this study was to test backup power technologies to supply extended runtime for traffic signal systems at a typical signalized intersection during a power outage for a minimum of three days. The existing alternating current voltage level used for traffic signal power is 120 VAC. For example, the intersection of E Fowler Ave & Bruce B Downs Blvd in Tampa, Florida has 21 traffic and 8 pedestrian signal heads. The average LED traffic signal power is 9 W (Watts), and the pedestrian signal power is 6 W. This computation is based on specifications from actual traffic signals approved in Florida (Dialight 2019, FDOT 2023). The estimated power required is about 300 W or 0.3 kW (kilowatts), including powering all traffic signals, the communication system, vehicle detection in operation, etc. Table 1 shows the detailed computation.

Table 1. Traffic Intersection Calculated Power

Device Used for Power Calculation	Full Operation Power (Watts)	Flashing Mode Power (Watts)
Dialight LED signal heads (x21)	189	90
Dialight Pedestrian (x8)	18	18
McCain 170 Controller (1x1)	40	40
EMX Detection loops (x21)	30	30
Allied Vision Monitoring equipment (x4)	10	10
Total	295	188

This brings the required capacity to 21.3 kW for full operation and to 13.5 kW for flashing mode for 3 days.

Low Voltage Backup System Topology. Figure 1 shows the interconnection topology of a traffic signal backup power system. The traffic signal devices are connected and controlled by a controller cabinet and are powered by DC power from the uninterruptible power supply (UPS) in the controller cabinet. The AC power sources are connected to an AC/DC converter, and the DC sources (fuel cell or photovoltaic) can be imported to the controller cabinet directly or through a DC/DC converter.

During normal conditions, the system is powered by a grid network, which provides a 120 VAC power source. Once the grid is out of power, such as during a hurricane or storm, the switch will connect the system to the backup power source automatically. If the backup power is

an AC source, it will go through an AC/DC converter; if it is a DC source, it can be used to power the system directly.

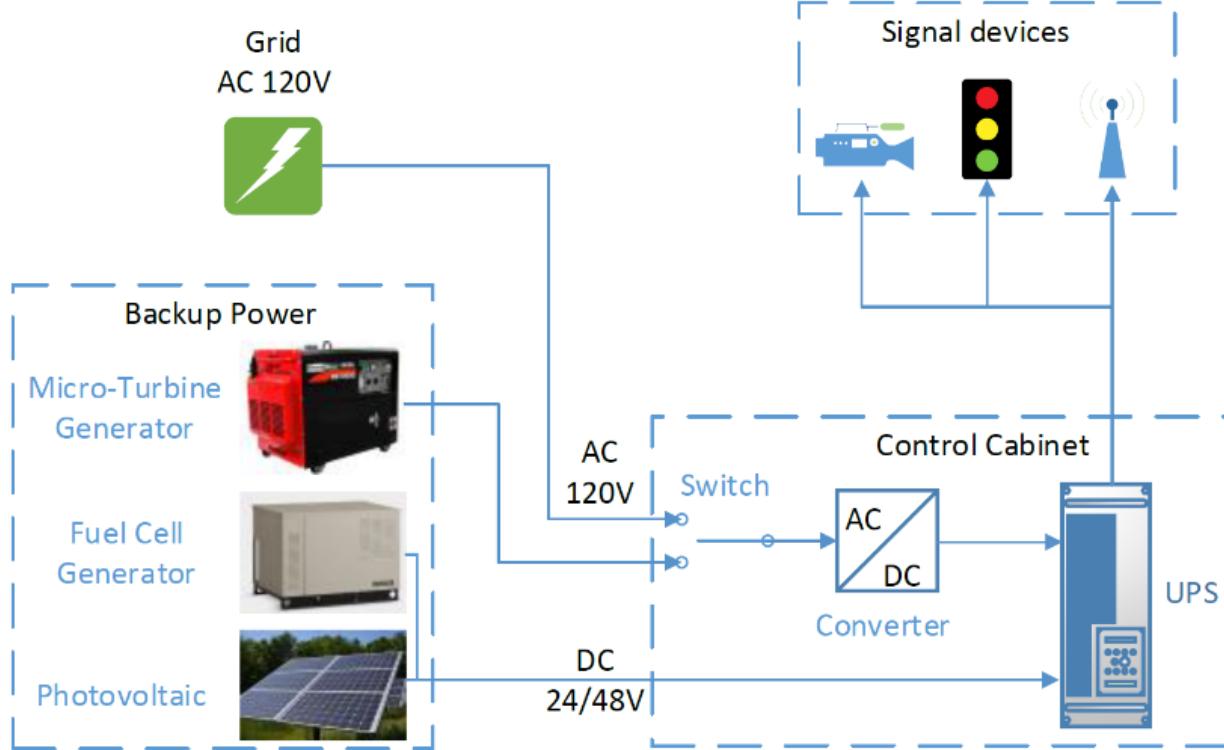


Figure 1. Concept of traffic signal backup system.

Backup Power Systems. Several backup power systems currently exist that would provide backup power to a traffic cabinet. These include battery backup systems (Schmidt-Rohr 2018), fuel cells (FCHEA 2023), natural gas/liquid propane generators (NaturalGas 2013), portable gasoline/diesel generators, solar power, wind generators and even supercapacitors (Goodwood 2014). There are pros and cons in using each of these systems, especially in the application of a traffic signal backup power. Many have been researched and tried in the past, but not all have been deemed appropriate for use. In addition, the topology to be applied is important since not all can be used in Florida. For example, wind turbines (generators) could not be used in Florida since the required wind is not widely available year-round. For Florida, major promising backup power technologies for meeting the needs of three days power include battery, fuel cell, natural gas/liquid propane, and solar power with battery.

Several promising backup power technologies have been commercialized and successfully applied to specific fields in several industries, but in the traffic signal field, some are in the early stages of testing or pilot implementation. A simple comparison of backup power technologies on their applications is shown in Table 2.

Based on the assessment of technologies and associated backup power systems, the research team recommended three candidate backup power technologies and associated systems to the FDOT Project Manager for consideration for testing: (1) battery backup system, (2) battery & solar backup system, and (3) natural gas or propane generator backup system.

Table 2. Backup Power Technology Applications

Backup Power Technology	Industrial Applications	Commercialized?	Applied in traffic signal systems?
Supercapacitor	Rapid and short-term charge/discharge, e.g., vehicles, elevators	Yes	Not used
Battery (lithium-ion)	Electric vehicles, emergency power, off-grid power systems, energy storage, smart house	Yes	Limited or in early stage
Battery (lead acid)	UPS, emergency lighting, hospital equipment	Yes	Often
Hydrogen fuel cell	Transit buses, backup power for facilities and homes	Yes	Limited or in early stage
Natural gas/liquid propane	Standby generators for facilities and homes, clean energy vehicles	Yes	Limited or in early stage
Portable generators (gasoline/diesel)	Backup power for small-size facilities and homes	Yes	Often
Solar power	Power generation in grid farm and home rooftop	Yes	Often; mostly used in flashing beacons or LED signs and lights
Solar-powered hydrogen fuel cell	Hydrogen-powered forklifts	No	Not used or not yet
Wind generator	Power generation in wind farm	Yes	Limited or in early stage

TESTING AND EVALUATION

A test site at the FDOT Traffic Engineering Research Laboratory (TERL) in Tallahassee FL, was used to evaluate three recommended backup power technologies for extending traffic signal operations after the loss of utility power supply in hurricane events. Before testing, data were collected at a live signalized intersection in Hillsborough County, Florida. Voltage, current, and power data for a traffic control system in full operation and in flashing mode was collected. Following a pretest, the three selected backup power technologies were installed and tested at the TERL, and data were recorded and fully analyzed to compare the selected systems.

Evaluation Procedure. Each backup system was tested under full operation as well as in flashing mode if the runtime in full operation mode was not satisfied. Full operation testing took priority, as the goal was to ensure that the traffic system was working normally as long as possible. Figure 2 presents the operational condition decision logic used, and Figure 3 shows the connection diagram for the system test.

Data Collection and Analysis. With voltage and current data from the tested system, the operational condition could be evaluated via data analysis. Figure 4 shows the logic flowchart used to make a judgment.

Based on the electric-required specification in FDOT Standard Specifications for Road and Bridge Construction, 685-2.2.4, the voltage should be in the range of 85–154 VAC. Therefore, the system will be in full operation if the voltage is within the normal range and the calculated

power is close to full power. The system will be in flashing operation if the voltage is within the normal range and calculated power is decreased to 50% level.

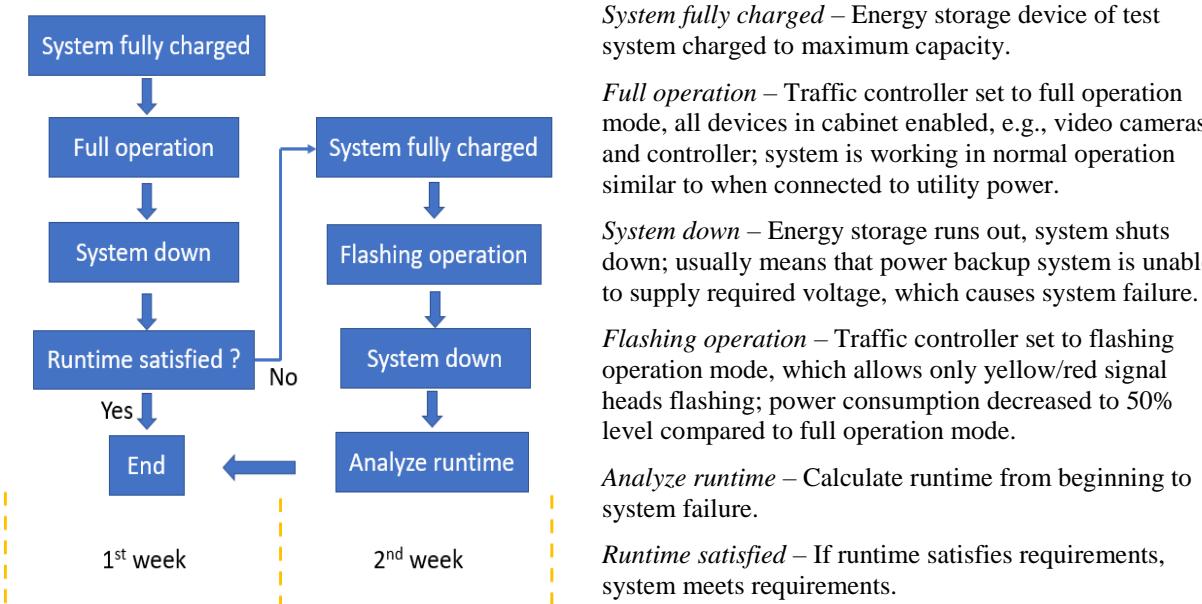


Figure 2. Operational condition decision.

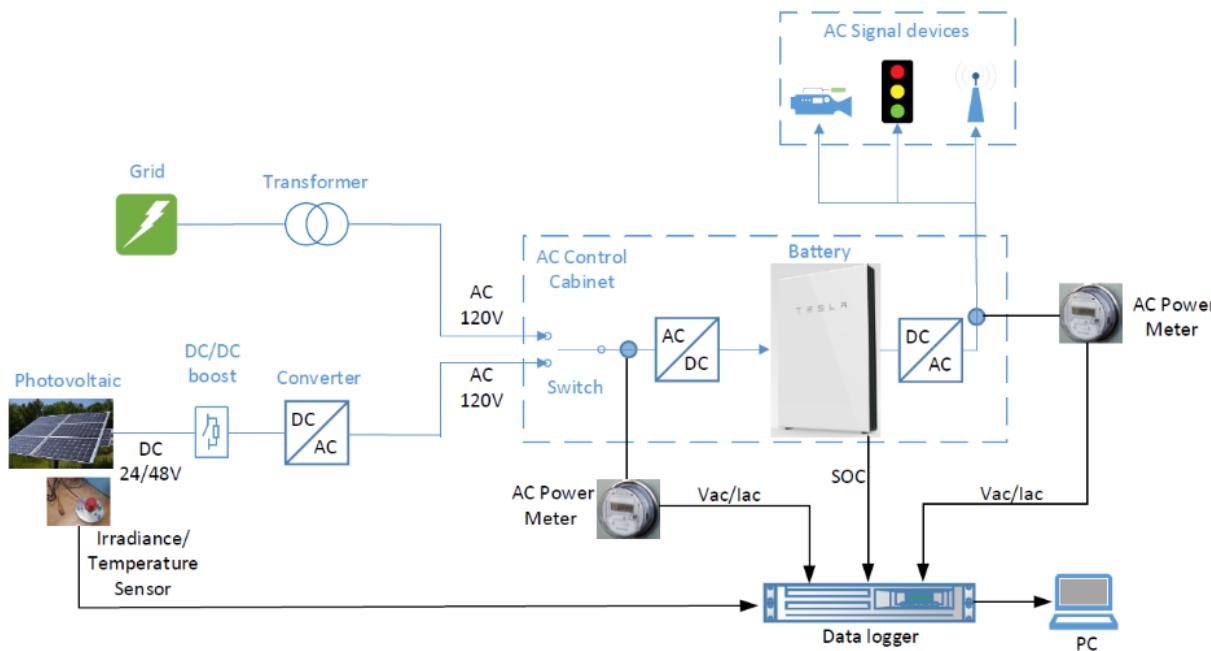


Figure 3. 120 VAC system assembly and data logger connection.

Battery & Solar Backup System. A 10 kWh capacity lithium battery and 8 solar panels with 2.8 kW maximum power were installed for testing as a backup power system. Figure 5 shows the system setup at the test intersection at the FDOT TERL facility in Tallahassee FL.

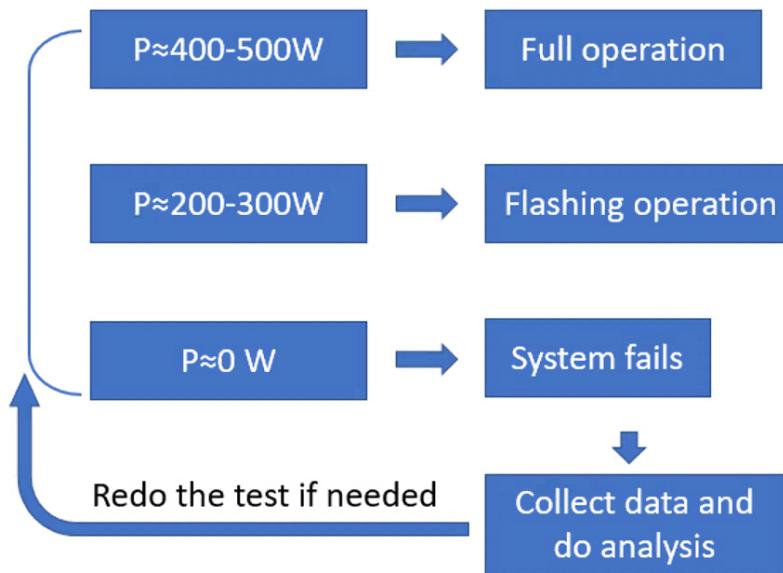


Figure 4. Operation condition evaluation flowchart.



Figure 5. Battery system and solar panels (left), inverter and battery (right)

The objective of using solar panels is to utilize a small battery (24 hrs capacity), but then use the solar panels to charge the battery during the day. Assuming there is adequate solar irradiance, the system can work indefinitely. This test was conducted after the battery was 100% fully charged and was conducted twice, with different outcomes. Test 1 included experiment data for a bad weather scenario, and Test 2 included experiment data for a good weather scenario, which affected how much solar power was available to charge the batteries and support the required runtime. The solar irradiance historical data were downloaded from the SOLCAST webpage (<https://solcast.com/>).

Figure 6 shows one week of data for reference. During the testing week, the first two days had cloudy and rainy weather, and the following five days had fewer clouds and more sunny weather. The two tests were run separately and are shown as a dashed line in Figure 6.

Battery & Solar System Data Analysis. The first test started at 10:04 AM. The battery's initial State of Charge (SOC) was 100%, and the traffic load was switched on. Eight 400 W solar panels were available to generate up to a rated 3.2 kWp. The test runtime was expected to be a minimum of 72 hours from start, which is the desired outcome of such systems. Solar power, load power, and battery SOC during this test are shown in Figure 7.

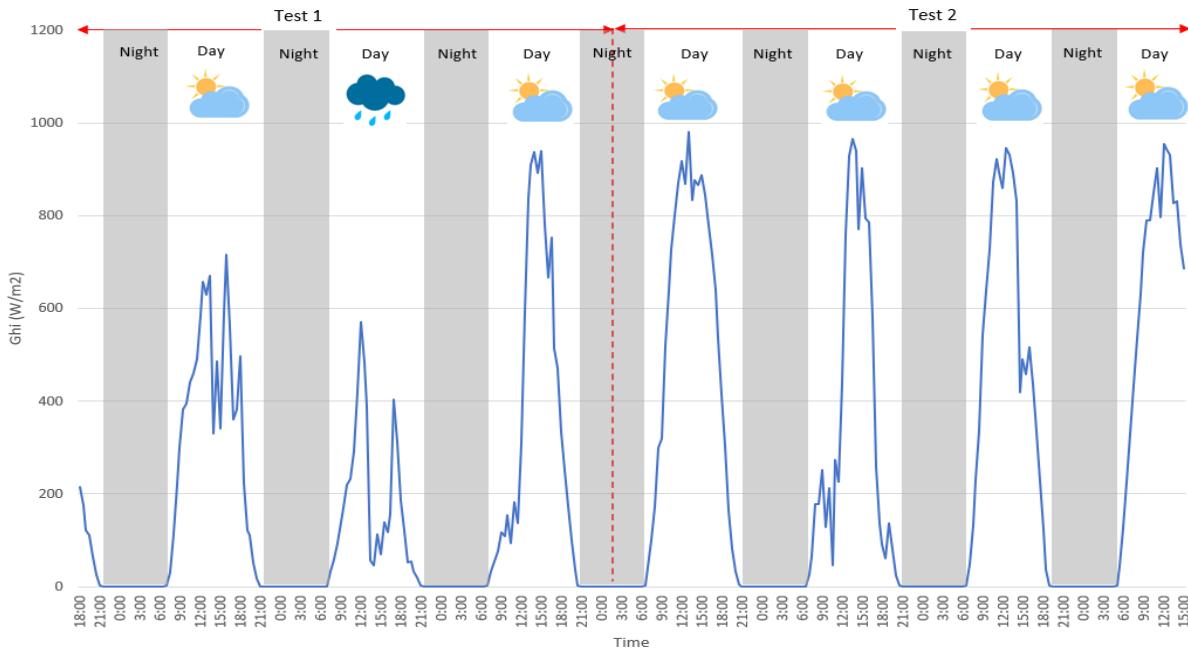


Figure 6. Solar irradiance data at TERL.

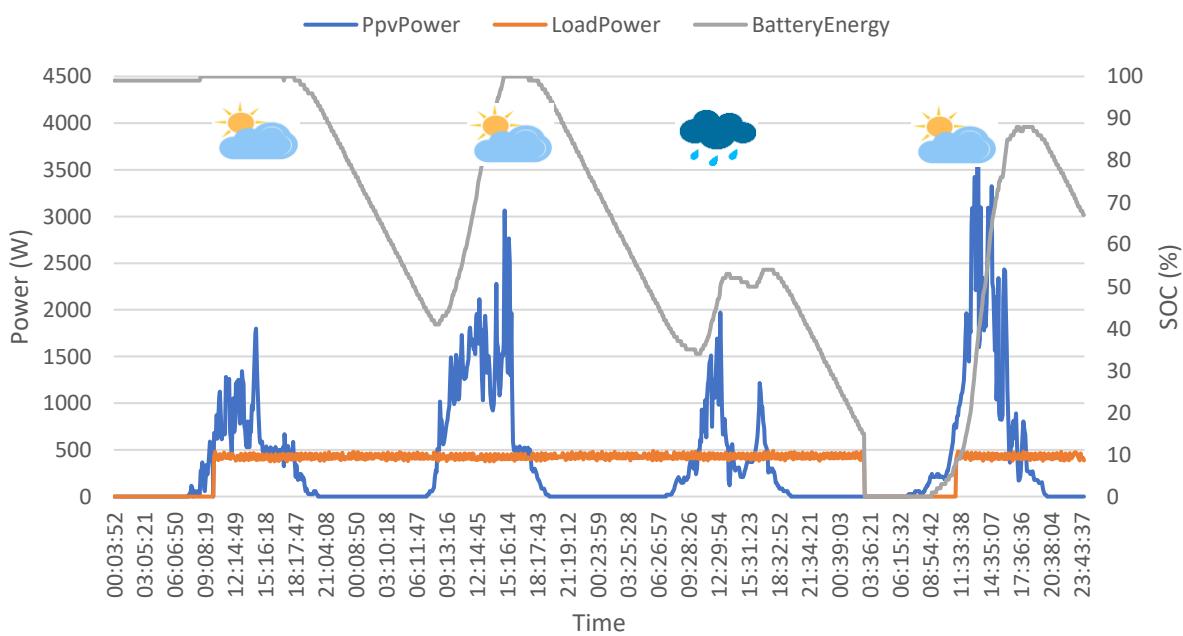


Figure 7. Battery & solar test 1 data.

The solar power shown in Figure 7 reflects the weather conditions during the first three days. Although there was not sufficient solar energy during the first day, the solar panels were able to keep the battery charged 100% until 6:15 PM. From 6:15 PM to 9:13AM the next day, the load was served by the battery with no solar power (nighttime). The next day, the solar panels were able to recharge the battery to 100% only briefly (3:21–5:42 PM). At 5:42 PM, the battery started to serve the load until the next morning at 8:53 AM when the solar panels started to produce energy. This time, however, the panels were not able to charge the battery, which reached a maximum of 63%. At 6:50 PM, the battery started to serve the load for the night and dropped to 15% at 3:00 AM, at which time the system shut down to protect the battery. The system remained shut down until 11:00 AM when the solar panels were able to charge the battery enough to switch it back on. The gap between 3:00 AM and 11:00 AM was due to the previous cloudy day when the battery was not charged more than 53%. The runtime requirement would have been met if the system had been operational for another seven hours (3:00–10:00 AM).

In this test, the battery and the eight solar panels supported the traffic system working in full operation for 2 days and 17 hours. The combination of battery capacity and amount of solar power provided can greatly affect the characteristics of the system related to how fast it discharges or charges and how long it can power the intersection without solar power. The test showed that the system used approximately 60–70% of the battery capacity overnight; however, this value was not accurate, as cloudy days meant that solar power was late to start charging in the morning or early to drop at night and a longer battery runtime was required.

Second Battery + Solar Test and Data Analysis. After the first test, the system continued to power the cabinet for the following days. A second test was recorded, in which solar irradiance at the TERL was relatively better and the battery was fully charged during daytime as a best-case scenario. The data collected during the next 3 days are presented in Figure 8.

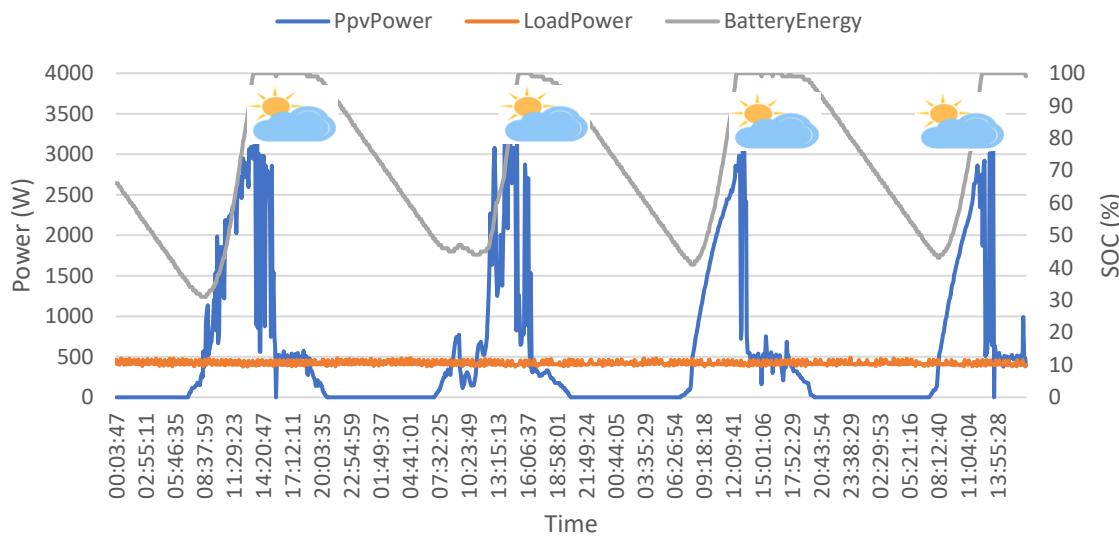


Figure 8. Battery & solar test 2 data.

The setup was the same as in Test 1 with the only difference being the weather during the three testing days. This time, the test did not start with a 100% battery charge; the initial battery SOC was about 30% before solar power started charging. Due to sufficient solar irradiance, the

battery was fully charged at 2:00 PM the next day while also powering the cabinet. The battery remained at 100% until the solar power could not provide enough power and dropped overnight from 100% to 44%. The second day, the solar panels charged the battery to 100% at 4:30 PM; subsequently, the battery took the load overnight and reached a low 41%. On the last day of the test, the solar panels charged the battery from 41% to 100% in five hours, the shortest time of all test days. The battery remained charged until nighttime operations. As shown in Figure 8, the load was always constant, with no loss of power (unlike Test 1). This test showed the expected pattern of charge and discharge cycles when the weather is clear enough such that the solar panels can provide enough energy to charge and maintain the system.

Propane Generator Test and Data Analysis. A liquid propane generator with a 57-gallon fuel tank was installed at the same intersection for testing purposes. The propane tank was fueled at 80% capacity for safety reasons, as propane tanks need space for the fuel to expand under high temperatures. As the total tank capacity was 57 gallons, the estimated initial propane fuel was approximately 45.6 gallons of fuel with a predicted runtime of 2–3 days under full operation conditions. Figure 9 shows the components installed at the site.



Figure 9. Propane generator backup power supply system at test site.

This test was straight forward as the liquid generator is used as a backup system for buildings, homes, and offices. The transfer switch is responsible for switching the power from utility to the generator once it senses that power is lost. The generator used was 12 kW rated, but smaller sized ones can be used. For the tested 400–500 W signal system, this generator was expected to work at 25% load operation. According to fuel conversion factors, the predicted runtime was approximately 51.3 hours with 45.6 gallons of LPG. Figure 10 shows the propane fuel level during the test. The fuel consumption is in a linear relationship with time, and the total generator running time was 51 hours. Thus, the propane generator used 0.87 gal/hr when supporting the traffic signal system. Actual fuel consumption was also compared with predicted fuel consumption.

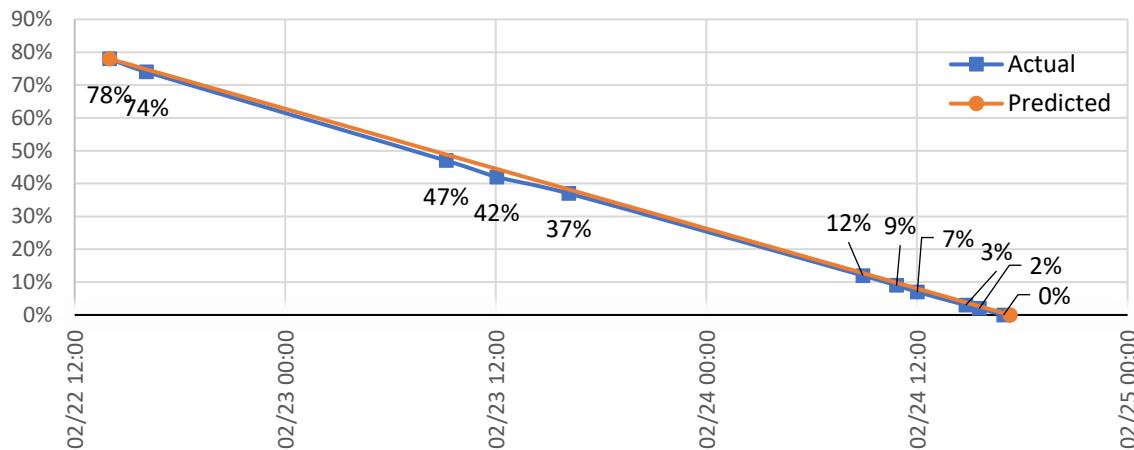


Figure 10. Propane fuel consumption during test.

Figure 11 shows that the traffic control cabinet was disconnected from utility power at 1:57 PM; the propane generator started to power the traffic system within 1 min. During the transition, the intersection was dark for 13 seconds, and the signals went to flashing mode for one cycle; within 30 seconds, the traffic signals were back to fully operating mode.

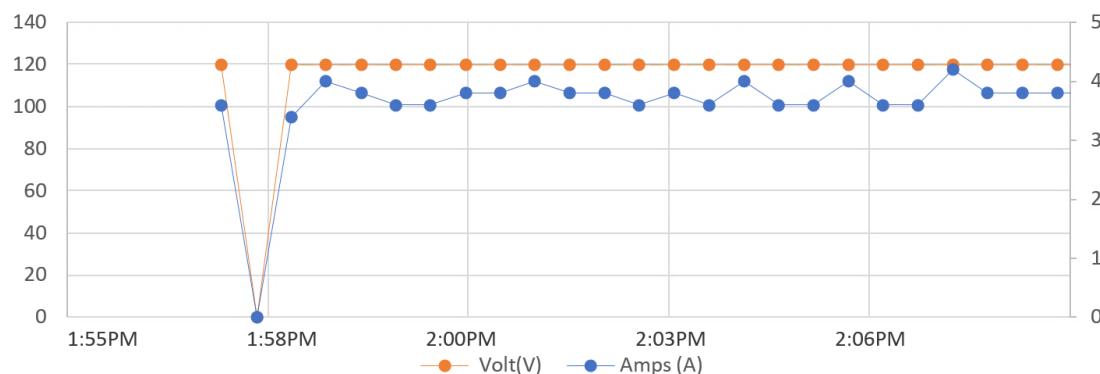


Figure 11. AC voltage/current data when switching from utility power to propane generator.

Test Summary. During testing, three different backup systems were acquired, installed, and tested at the TERL test intersection: a propane generator, a battery system, and a battery & solar system. All tests successfully showed the capability of the systems to extend traffic load runtime after losing power from the main grid. Comparison of the tested systems is summarized in Table 3.

Based on the comparison in Table 3, the following conclusions can be drawn:

- A **propane generator** is a low-cost backup power technology that will not be impacted by weather conditions; however, power quality is not as clean as utility power, and it might not be able to support sensitive electronic devices. A propane generator is loud and requires maintenance and refueling. In addition, the transfer time was not adequate for uninterrupted power supply, but it can be adjusted. A standard UPS can fill the gap between utility power loss and generator fully operational mode.

Table 3. Comparison of Tested Technologies

Attribute	Propane	Battery	Battery + Solar
Cost	\$9k	\$16k	\$17k
Power quality	Adequate	Good	Good
Runtime	2–3 days	20 hrs	Up to 5 days
Startup time	30 sec	Milliseconds	Milliseconds
Weather impact	No	No	Yes
Noise level	High	Low	Low
Air pollution	No	No	No
Maintenance	Quarterly	Yearly	Yearly
Refuel	Yes	No	No
Data monitor	Yes	Yes	Yes

- A **lithium battery system** (ion or Li-PO) provides reliable power quality and can be switched on within milliseconds without noticeable signal noise. It does not need refueling and requires less maintenance, and diagnostics can be monitored remotely. It is a more expensive system, especially when a larger-capacity battery is required. Battery life cycle issues must be addressed to ensure safe and reliable operation for years.
- A **battery & solar system** has all the advantages of a battery system and can extend runtime at a very low-cost rate, as solar panels are relatively inexpensive. Solar panels can charge the battery during the day, which discharges during the night, and this cycle can be sustained until the batteries need to be changed. The number of solar panels and battery capacity will vary depending on required load, available space, desired runtime, and available budget for this type of system.

CONCLUSION

The research recommended three specific technologies for consideration in testing extended backup power for traffic cabinets in case of power loss. Current traffic cabinets use a standard UPS system, which is powered by lead acid batteries, an outdated technology that provides power for up to eight hours of operation. The three backup power systems were approved and selected for testing including LPG generator system, Lithium battery system, and Lithium battery & solar backup system.

An LPG generator is a tested technology used in many applications, including telecommunications, rail, and other industries with remote facilities. The team installed and tested a generator that provided power to a traffic cabinet without issue for as long as it had fuel. During the switch, the generator took 13 seconds to transfer power, which would require further adjustments or the use of a small UPS to cover the gap between power loss and the generator fully supporting the load. In addition, the LPG generator requires a fuel tank, which can be hazardous if hit during a traffic crash; therefore, additional safety precautions must be taken (breakaway lines, automatic shutoff valve, or concrete barriers) to ensure safe storage close to a traffic intersection.

The lithium battery system is a relatively newer technology widely used in commercial and residential applications as energy storage and for off-grid applications. Recent advances in battery technology used in electric vehicles and energy storage allow for a compact form factor with high energy capacity. These batteries have a higher initial capital cost, but a longer life

cycle with up to a 10-year warranty compared to other batteries. The battery-only system transferred the load instantaneously and without delay and is an applicable solution to the problem of providing extended power to traffic cabinets.

To minimize the cost of the batteries, a recommended solution is to provide backup power using additional solar power. In Florida after a hurricane, typically there are sunny days that can provide enough solar power to recharge a smaller battery discharged during the night and cycle for the next day, a pattern that can go indefinitely. The team installed and tested this setup, providing enough solar power to charge a 10 kWh battery used to power an average 0.45 kW load for 14 hours overnight, from 6:00 PM to 8:00 AM. The solar power in the morning powered the load and charged the battery to be ready for the next night. In real-world applications, a specific design needs to be implemented for each intersection to assess power needs, available space, and budget to install a battery & solar battery system.

The systems tested showed a great promise in providing backup power in case of power loss due to a hurricane for an extended period. Selection and availability of each system will be the responsibility of each agency once the systems are approved and listed on the FDOT Approved Product List.

ACKNOWLEDGMENTS

This research project was sponsored and supported by the Florida Department of Transportation (FDOT). Authors would like to express sincere appreciation for the support, guidance, and assistance from Derek Vollmer and Matthew DeWitt with FDOT Traffic Engineering and Research Lab (*TERL*).

REFERENCES

- Lin, P.-S., Kourtellis, A., Miao, Z., Wang, Z., and Zhang, M. Development of Low Voltage/Extended Runtime Signalized Intersection Using Backup Power after the Loss of Utility Power Due to Hurricanes. Final Report. Florida Department of Transportation, 2021, 110p. <https://trid.trb.org/view/1984651>.
- Dialight. 12" ITE Compliant XL15 Long Life LED Traffic Signals Brochure. https://s3-us-west-2.amazonaws.com/catsy.557/Dialight_Traffic_LongLife_Aug2019.pdf.
- FDOT. FDOT Approved Product List (APL): 650 – Vehicular Traffic Signal Assemblies. 2023. <https://www.fdot.gov/traffic/traf-sys/product-specifications.shtm>.
- Schmidt-Rohr, K. How Batteries store and release energy: Explaining basic electrochemistry. *Journal of Chemical Education*, 2018. 95(10): p. 9.
- First one up the drive: A new sort of storage device gives lithium-ion batteries a run for their money. The Economist, Science and Technology Edition, The Economist.com, July 12th, 2014. <https://www.economist.com/science-and-technology/2014/07/10/first-one-up-the-drive>.
- FuelCells.org, Fuel Cell Basic Applications. 2023. <http://www.fuelcells.org/basics/apps.html>.
- Naturalgas.org, Natural gas and the environment. 2013. <https://www.elgas.com.au/blog/486-comparison-lpg-natural-gas-propane-butane-methane-lng-cng>.

Before-After Analysis of Highway Crashes on the Dhaka-Chattogram National Highway of Bangladesh

Afsara Taspia¹; Shadman Sakib²; Md. Moshiur Rahman³; Javed Bari, Ph.D., P.E., M.ASCE⁴;
Kazi Md. Shifun Newaz⁵; and Faria Tabassum, M.ASCE⁶

¹Dept. of Civil and Environmental Engineering, North South Univ., Bashundhara, Dhaka, Bangladesh. Email: afsara.taspia@gmail.com

²Dept. of Civil and Environmental Engineering, North South Univ., Bashundhara, Dhaka, Bangladesh. Email: shadmanananta102@gmail.com

³Dept. of Civil and Environmental Engineering, North South Univ., Bashundhara, Dhaka, Bangladesh. Email: souravrahman2017@gmail.com

⁴Dept. of Civil and Environmental Engineering, North South Univ., Bashundhara, Dhaka, Bangladesh. Email: javed.bari86@gmail.com

⁵Accident Research Institute, Bangladesh Univ. of Engineering and Technology, Dhaka, Bangladesh. Email: newazshifun@gmail.com

⁶Flood Control District, Maricopa County, Phoenix, AZ. Email: ftabassum.faria@gmail.com

ABSTRACT

This study critically assesses the impact of converting the Dhaka-Chattogram National Highway (NH-1) in Bangladesh from an undivided two-lane to a four-lane highway with a median on crash rates and patterns. Crash data pre- and post-conversion was sourced from the Accident Research Institute (ARI) and the Bangladesh police headquarters for detailed analysis. The transition resulted in a notable decline in head-on collision-related fatal crashes. However, concerning trends emerged post-conversion, revealing an increase in various other significant crash types and associated injuries and fatalities, particularly those involving pedestrians. Effective countermeasures are recommended based on the study findings to mitigate highway crashes and their severity. This investigation provides critical insights into the dynamics of highway safety post-expansion, highlighting the necessity for targeted interventions to address emerging risks and enhance overall safety on NH-1.

INTRODUCTION

Background. Globally, road crashes represent a significant concern, causing over 1.3 million fatalities and 20 to 50 million injuries annually, with low and middle-income countries accounting for 93% of these fatalities (WHO, 2022). Various factors contribute to highway crashes, including human error, vehicle condition, traffic, road geometry, and weather (Arun et al., 2022). Bangladesh, facing one of the highest rates of road fatalities among emerging nations, experiences highway crashes primarily due to reckless driving and excessive load (Islam et al., 2021). The Dhaka-Chattogram Highway (i.e., N1), a vital transit link connecting the capital Dhaka with Chattogram - the second largest and the main seaport city of the country, underwent significant expansion from two to four lanes in 2016, aiming to reduce transit time and bolster economic growth (Road Traffic Technology, n.d.). However, despite this expansion's benefits in enhancing connectivity and economic prospects, crashes persist due to factors such as reckless driving and increased traffic (Illius, 2016).

Literature Review. Several studies have addressed road safety concerns along the Dhaka-Chattogram Highway, identifying crash-prone areas and factors contributing to crashes. Previous investigations highlighted significant crashes post-expansion, indicating a rise in fatalities and injuries due to reckless driving (Chowdhury, 2020). The expansion aimed to reduce transit time and improve traffic flow, yet it resulted in increased crashes attributed to careless driving (Enan, 2022). However, there is a gap in comprehensive studies that bridge the pre-upgrade and post-upgrade crash scenarios. Studies pinpointed specific crash-prone zones and highlighted key factors like speeding and driver carelessness contributing to crashes such as (Alam et al., 2013; Saleh et al., 2015). Furthermore, Hamim et al. have utilized the AcciMap technique to analyze crash causes beyond individual driver responsibility (Hamim et al., 2019). Notably, existing research emphasized factors like high speeds, irresponsible driving, and vehicle overloading as primary contributors to crashes (Islam et al., 2021). A recent study delved into the impact of four-lane bridges on traffic flow, uncovering variations in speed and flow rates before and after the bridge expansion (Mamun et al., 2022). Additionally, a case study on the Dhaka-Faridpur highway (Rakib et al., 2022) highlighted the high-speed element and the predominance of buses in crashes. However, despite the existing body of research, a comprehensive study bridging the pre-upgrade and post-upgrade scenarios is lacking, motivating the need for this research.

Objective. This study aims to analyze crash rates and patterns on the Dhaka-Chattogram Highway (N1) before (2013-2015) and after (2018-2020) its upgrade. Additionally, the research endeavors to propose strategies to address road safety concerns based on the findings.

Study Area. The focus of this investigation spans the 233-kilometer stretch from Jatrabari ($23^{\circ}42'19.94''N$, $90^{\circ}26'40.33''E$) to Chattogram City Gate ($22^{\circ}22'25.58''N$, $91^{\circ}46'24.14''E$) along the Dhaka-Chattogram Highway (N1) of which 7.2 kilometers from Jatrabari to Kachpur Bridge have eight lanes and the remainder have four lanes.

METHODOLOGY

Before Study. To address the study objective, the MAAP-5 software was utilized at the Accident Research Institute (ARI) for compiling crash data on the N1 highway from 2010 to 2015 (Baguley, 1994). Subsequently, the data pertaining to the N1 highway, spanning from Jatrabari to Chattogram City Gate, was organized and processed.

After Study. For the "After" study on the N1 highway, crash data was collected from newspapers covering the time period between 2018 and 2021, sourced from the ARI. As the MAAP-5 software-based crash database of ARI extended only until 2015, the 2021 data was obtained from the Bangladesh Police Headquarters. Notably, the detailed crash data before 2021 was unavailable at the Police Headquarters, and the MAAP-5 database did not contain post-2015 data. Using the Microsoft Excel, the ARI (newspaper-based) and police headquarters (ARF-based) data were analyzed and plotted to identify variances. A predictive model (equation) was developed to convert the ARI's crash data into police data for 2018-2020. The pre-upgrade (2013-2015) MAAP-5 data was compared with the post-upgrade (2018-2020) predicted police data. For both datasets, the crash frequency, standard deviation, and pooled variance were estimated utilizing the following equations (Garbar et al., 2009):

$$\text{Crash Frequency} = \text{Number of Crashes} / \text{Period in years} \quad (1)$$

$$\text{Standard deviation, } S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N-1}} \quad (2)$$

where,

x_i = each value from the population

\bar{x} = population means

N = size of population

$$\text{Pooled variance, } S_p^2 = \frac{\sum_{I=1}^K (n_i - 1)s_i^2}{\sum_{I=1}^K (n_i - 1)} \quad (3)$$

where,

s_i^2 = sample variances

n_i = sample size

$$\text{T-test statistic, } T = \frac{\bar{x}_1 + \bar{x}_2}{S_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (4)$$

where,

\bar{x}_1 and \bar{x}_2 = sample means

n_1 and n_2 = sample sizes

S_p = square root of the pooled variance.

Additionally, a detailed field survey was conducted to get thorough insight of the traffic condition, road geometry, behavior of drivers and behaviors of pedestrian on N1.

ANALYSES AND FINDINGS

Comparison of crash statistics. The analysis presents a detailed comparison of highway crash statistics before and after the road expanded from two to four lanes. A summary is presented in Table 1. Initially, 306 incidents were recorded between 2013 and 2015, dropping to 282 post-expansion in 2018–2020. A 7.84% decrease in accidents occurred following the expansion.

Crash data collected from the Accident Research Institute (ARI) and Bangladesh Police Headquarters for 2021 prompted the development of an equation bridging differences in 2018–2020 data sources. Table 2 highlights the variations in the crash data of 2021 obtained from newspaper and police reports. Figure 1 provides the equation that established the relationship between the crash data obtained from accident reports and newspaper. Table 3 presents the crash data of 2013–2015 obtained from ARI and the comparable predicted crash data of 2018–2020.

Statistical analyses. Statistical analyses, including crash frequency calculations, T-tests, and variance assessments substantiate the observed trends. The overall crashes (irrespective of crash types), hit pedestrian-type, and rear-end-type crashes statistically increased, whereas the head-on collision and hit object type crashes statistically decreased post-expansion, as supported by the T-test results at a 5% level of significance.

Changes in crash statistics and trends. The crash frequency in crashes per year increased from 102 (before expansion) to 288 (after expansion). Each year's crash statistics from 2013 through 2020 were critically analyzed. Just as example, the crash statistics of 2015 and 2020 are

graphically shown in Figures 2 and 3. Notably, pedestrian accident before and after the expansion remains the highest, sustaining well-above 40%. Pedestrians' reluctance to use foot over bridges perpetuates the high pedestrian-related accident rates. Additionally, the highway's proximity to numerous food establishments contributes to these accidents.

Table 1. Comparison of highway crash statistics before and after the road expansion

Accident Types	"Before construction of four lane" Data (Source: ARI Map-5 Database)			"After construction of four lane" Data (Source: ARI Newspaper based)		
	2013	2014	2015	2018	2019	2020
Hit Object	1	4	5	0	1	1
Head On	25	28	26	4	5	4
Hit Pedestrian	43	24	55	45	38	29
Rear End	6	8	11	36	26	19
Side Swipe	9	3	5	2	0	0
Three Way Collision	0	0	0	6	5	3
Lost Control	0	0	0	15	16	0
Vehicle Roll-Over	0	0	0	0	0	0
Unknown	11	15	6	3	8	5
Fall from Vehicle	0	0	0	0	0	0
Over Turn	7	3	11	0	2	9
Total	102	85	119	111	101	70
Grand Total	306			282		

Table 2. Accident data of 2021 obtained from Bangladesh Police Headquarters and ARI

Accident Type	Accident data of the year 2021 collected from ARI (Source: Newspaper)	Accident data of the year 2021 collected from Bangladesh Police Headquarter
Axis	x	y
Hit Object	2	7
Head On	2	4
Hit Pedestrian	24	82
Rear End	26	74
Side Swipe	0	1
Three Way Collision	0	1
Lost Control	4	10
Vehicle Roll-Over	0	2
Unknown	1	0
Fall from Vehicle	1	0
Over Turn	1	0
Total	61	181

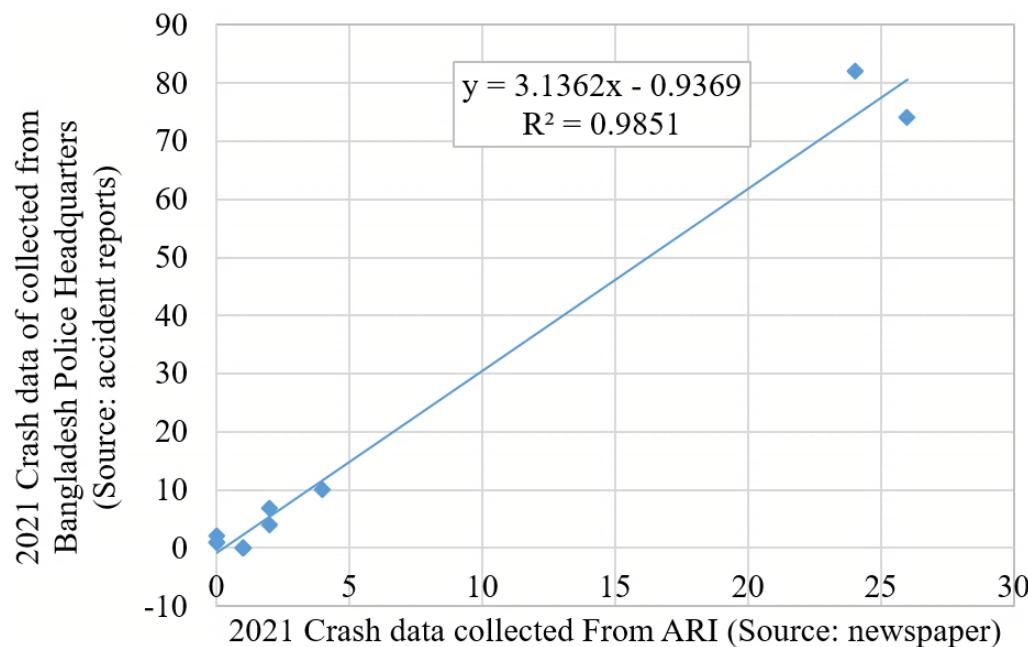


Figure 1. Year 2021 crash data obtained from accident reports and newspaper.

Table 3. Comparison of four-lane highway crash statistics before and after upgradation

Accident Types	"Before construction of four lane" Data (Souce: ARI Maap-5 Database)			"After construction of four lane" Data (Predicted)			Percent Reduction
	2013	2014	2015	2018	2019	2020	
Hit Object	1	4	5	0	2	2	-60%
Head On	25	28	26	12	15	12	-51%
Hit Pedestrian	43	24	55	140	118	90	185%
Rear End	6	8	11	112	81	59	908%
Side Swipe	9	3	5	5	0	0	-71%
Three Way Collision	0	0	0	18	15	8	0%
Lost Control	0	0	0	46	49	0	0%
Vehicle Roll-Over	0	0	0	0	0	0	0%
Unknown	11	15	6	8	24	15	47%
Fall from Vehicle	0	0	0	0	0	0	0%
Over Turn	7	3	11	0	5	27	52%
Total	102	85	119	341	309	213	182%
Grand Total	306			863			

Conversely, head-on collisions decreased significantly, from 22%-26% pre-expansion to around 5% post-expansion, though not reaching the anticipated 0%. The absence of service lanes for local vehicles like battery-operated rickshaws and auto-rickshaws leads to vehicles traveling in wrong directions, despite median construction on the Dhaka–Chattogram highway. While the presence of dividers significantly reduced head-on collisions, vehicles, especially during traffic congestion or accidents, still maneuver into oncoming traffic.

Rear-end collisions notably surged from 8%-10% pre-expansion to around 28%-29% after the lane expansion, possibly due to increased free-flow speeds, resulting in driver negligence and weather-related control loss. Similarly, incidents of loss of control rose from 0% to 11%, attributed to higher speeds despite smoother driving conditions. Contrarily, side-swipe collisions decreased from 6% to 1% following the highway's expansion. The expansion alleviated risks associated with overtaking slow vehicles but previously posed dangers due to limited space for maneuvering.

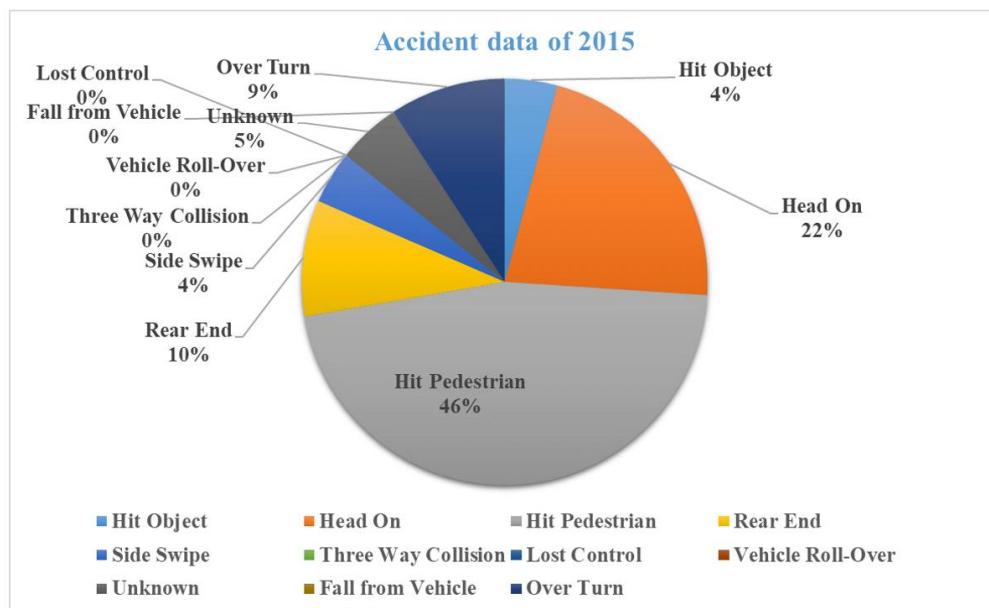


Figure 2. Crash statistics of 2015.

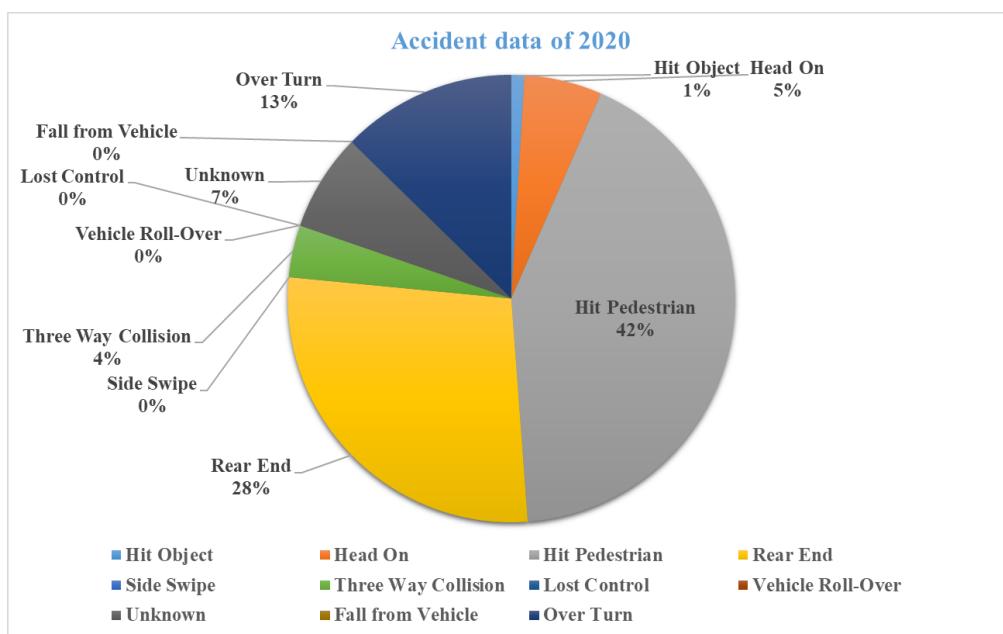


Figure 3. Crash statistics of 2020.

CONCLUSION

Comparing crash data between 2013–2015 and 2018–2020 highlighted significant shifts in numbers and types of incidents. Despite the road widening from undivided two-lane type to divided four-lane type, pedestrian-involved accidents persist above 40%. Head-on collisions reduced from 26% to 5%, though short of the targeted 0%. Meanwhile, rear-end crashes surged from 8% to 29%, while side-swipe incidents dropped from 6% to 1%. Previously, few control-loss type crashes resulted from slow free-flowing speeds, contrasting sharply with the current rise in control-loss incidents, now at 11%. This analysis stresses the ongoing necessity for robust road safety measures and proactive steps to mitigate future crashes. The study's outcomes reinforce the urgency of sustained initiatives to bolster road safety and implement strategies for minimizing future accident occurrences.

RECOMMENDATIONS TO ENHANCE ROAD SAFETY

The research underscores that pedestrian involvement dominates collisions due to insufficient crossing options. Despite the road expansion to four lanes, increased accidents, especially reckless driving and rear-end collisions, demand immediate attention. To mitigate these issues, aligning with the government's decision to construct service lanes, relocating roadside markets, erecting rest areas, foot-over bridges, and installing appropriate signboards and passenger shelters can effectively reduce accidents. Elevating medians strategically along the route will discourage random pedestrian crossings, reducing accidents significantly. Addressing drivers' conditions is crucial. Specifying working hours and instituting mandatory training sessions on overtaking, braking, and safety measures can notably decrease accidents, considering the long hours spent driving. Streamlining the driving license process, ensuring vehicle fitness, comprehensive knowledge of road signs, and driving expertise should drastically alter crash dynamics.

RECOMMENDATIONS FOR FUTURE STUDIES

Future research should gather comprehensive secondary data from diverse sources for accuracy. Comparative research requires data consistency for 18 to 24 months to accurately determine variance. Conducting extensive fieldwork will precisely delineate accident patterns and timings for a deeper understanding.

REFERENCES

- Alam, M., and Ahsan, H. M. (2013). Identification and Characterization of Hazardous Road Locations on Dhaka-Chittagong National Highway. *Global Journal of Researches in Engineering General Engineering*. Global Journals Inc., USA. Volume 13, Issue 4, Version 1.0. Retrieved from: <https://engineeringresearch.org/index.php/GJRE/article/view/955/887>.
- Arun, A., Haque, M. M. S., Washington, T. S., and Mannering, F. (2022). How many are enough?: Investigating the effectiveness of multiple conflict indicators for crash frequency-by-severity estimation by automated traffic conflict analysis. *Transportation Research Part C: Emerging Technologies*, 138 (2022), Article 103653, 10.1016/j.trc.2022.103653.

- Baguley, C. J. (1994). TRRL's Microcomputer Accident Analysis Package (MAAP): A Standard for SARTTO Countries. Transport Research Laboratory, TRID Accession No. 00646854, Washington, DC, USA.
- Chowdhury, Z. (2020). Reckless driving still rampant on Dhaka-Chittagong Highway. The Business Standard Newspaper. Published on October 22, 2020. Retrieved from: <https://www.tbsnews.net/bangladesh/reckless-driving-still-rampant-dhaka-chattogram--148279>.
- Enan, F. (2022). Dhaka Chittagong Highway, Present and Future. Growing Up with Bangladesh. Retrieved from: <https://www.growingupwithbd.com/2022/11/>.
- Garbar, N. J., and Hoel, L. A. (2009). *Traffic and Highway Engineering*. Cengage Learning.
- Hamim, O. F., Hoque, M. S., McIlroy, R. C., Plant, K. L., and Stanton, N. A. (2019). Applying the AcciMap methodology to investigate the tragic Mirsharai road accident in Bangladesh. *MATEC Web of Conferences*. Vol. 277, p. 02019. EDP Sciences.
- Illius, S. (2016). 4-lane Dhaka-Ctg highway ready for use. The Daily Independent Newspaper. Published on July 2, 2016. Retrieved from: <https://m.theindependentbd.com/printversion/details/50027>.
- Islam, M. A., and Dinar, Y. (2021). Evaluation and Spatial Analysis of Road Accidents in Bangladesh: An Emerging and Alarming Issue. *Transportation in Development Economics*. Retrieved from: <https://link.springer.com/article/10.1007/s40890-021-00118-3>.
- Mamun, G. M. A., Hasan, A. S., Rokhsana, K., and Pervin, A. (2022). Impact of four lane bridges on mobility of traffic flow in Dhaka-Chittagong Highway. *6th International Conference on Civil Engineering for Sustainable Development (ICCESD 2022)*. 10~12 February 2022, KUET, Khulna, Bangladesh (ISBN- 978-984-35-1972-6). Retrieved from: http://iccesd.com/proc_2022/Papers/TRE-4854.pdf.
- Rakib, R., Manik, M. H., Hoque, E., Tumpa, J. F., Ullah, I., Asadujjaman, M., Farjana, M., and Haque, R. (2022). Highway Accident Analysis: A Case Study Dhaka-Faridpur Highway. *Journal of Transportation Systems*. Volume-7, Issue-3 (September-December, 2022). Page 18-27.
- Road Traffic Technology. (n.d.). Dhaka-Chittagong National Highway Upgrade Project. Retrieved from: <https://www.roadtraffic-technology.com/projects/dhakachittagong/?fbclid=IwAR3Kg5qxjBJ29Rk4FKAWiAZVqZ482kZtIX9BZm8c5Spew9S4v4TDLCPoEqw>.
- Saleh, S., and Islam, S. (2015). Safety Investigation and Assessment of Dhaka- Chittagong National Highway in Bangladesh. *10th Conference on Global Engineering, Science and Technology Conference*. Volume: ISBN: 978-1-922069-69-6. BIAM Foundation, Dhaka, Bangladesh. Retrieved from: https://www.researchgate.net/publication/341105433_Safety_Investigation_and_Assessment_of_Dhaka-_Chittagong_National_Highway_in_Bangladesh.
- WHO (World Health Organization). (2022). *Road traffic injuries*. World Health Organization.

Machine Learning Based Multi-Modal Transportation Network Planner

Neeraj Menon Manghat¹; Vaishak Gopalakrishna²; Sai Bonthu³; Victor Hunt⁴; Arthur Helmicki⁵;
and Doug McClintock⁶

¹Research Assistant, Dept. of Computer Science, Univ. of Cincinnati, Cincinnati, OH.
Email: manghann@mail.uc.edu

²Research Assistant, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati,
Cincinnati, OH. Email: gopalavk@mail.uc.edu

³Research Assistant, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati,
Cincinnati, OH. Email: bonthusy@mail.uc.edu

⁴Professor, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, Cincinnati, OH.
Email: huntvj@ucmail.uc.edu

⁵Professor, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, Cincinnati, OH.
Email: helmicaj@ucmail.uc.edu

⁶Executive Director, Red Bike (Cincy Bike Share, Inc.), Cincinnati, OH.
Email: doug.mcclintock@cincyredbike.org

ABSTRACT

Multi-modal transportation relies on efficient network connectivity among active transportation services, such as biking and transit stops. In many cases, road users must walk or bike varying distances to reach a transit stop and then continue their journey to their final destinations. As shared mobility services like installations of shared bike docking stations increase in urban settings, planning additional docking stations that efficiently connect with existing bike network and transit stops becomes a significant challenge. This paper proposes a machine learning (ML)-based multi-modal transportation network planner to address the optimization challenge of expanding shared bike docking stations (Cincinnati Red Bike) that connect with the existing transit network (Cincinnati Go Metro) and amenities within neighborhoods of the city of Cincinnati. We identify relevant data sources for the planning problem, such as existing bike share ridership data, transit data from General Transit Feed Specification (GTFS), demographic data from the Census, and built environment factors such as amenities and land use data. The K-means clustering algorithm is used to model the region's features across the selected variables and identify potential areas to expand into. Network analysis is performed to observe the network effects of expansion and determine candidate locations where future bike share stations may be built.

INTRODUCTION

This paper deals with examining how to efficiently expand the shared bike network in the metro Cincinnati, OH and Northern Kentucky area. We source and integrate varied data sources that contribute to the planning process, such as demographics, employment, amenities and transit and apply modeling techniques to understand the characteristics of the region and find optimal areas to expand the bike network into. Data analysis could play a significant role in assisting urban planners and communities in selecting optimal locations for bike docking stations that are both data-driven and community-preferred. For example, Red Bike plans expansion specifically

into Walnut Hills, Avondale, and Evanston purely from a community development perspective. With intensive community engagement, the team is collecting data through surveys, where community members place pins on a map indicating where they would like to ride a bike. Survey results serve as data points to guide expansion into those community spaces. The team also considers the demographics of the bike stations' service areas. However, there is no hard and fast science that guarantees success or failure, as it depends on the community's adoption of bike riding. Combining data-driven insights for newly proposed expansion spots with other metrics, like community preference, would simplify and optimize the effectiveness of the expansion process of bike share systems. The approach used and results obtained may also be helpful for further mobility planning, such as choosing optimal locations for smart mobility hubs, park & ride, and other similar mobility services.

Federal level research studies (USDOT FHWA 2016, 2018) were initially explored to understand the features and data sources relevant for multi modal mobility problems. Broad categories that capture attributes related to both demand and equity were identified. The primary data sources identified are demographics (including population, household densities and motor vehicle access), employment characteristics and transit accessibility. The U.S. Environmental Protection Agency's (EPA) and U.S. General Services Administration (GSA) Smart Location Database (SLD), (U.S. EPA 2021), summarizes several demographics, employment, and built environment variables for every census block group (CBG) in the United States. This database was chosen as one of the primary data sources since relevant features across different sources such as Census, American Community Survey (ACS) and Longitudinal Employer-Household Dynamics (LEHD) have already been sourced and aggregated, enabling simpler integration into our model. The SLD dataset was then fused with the 2019 Census TIGER/Line Shapefiles that denote the boundaries of census block groups. Amenities and transit data from OSMnx (Boeing 2017) and GTFS respectively are used at a later stage to further narrow the search for profitable areas of expansion. As our analysis focuses on the expansion of the already existing Red Bike station network, we reduce the region scope based on the bounding latitude and longitude values of the bike network, plus a buffer (5 km) to accommodate areas that have not been ventured into. The reduced dataset is then input to a clustering model and network analysis is performed to identify suitable areas for bike network expansion.

CLUSTERING OF CENSUS BLOCKS

The U.S. EPA SLD dataset includes indicators of the commonly cited 'D' variables shown in the transportation research literature to be related to travel behavior (Cervero, R. & Kockelman. 1997). The 'D's include residential and employment density, land use diversity, design of the built environment, access to destinations, and distance to transit. Of the total features present in the SLD database, a subset of them that logically pertains to our problem are selected and shown in Table 1, along with the description for each variable.

The selected features fall into the following categories and a brief explanation of each, along with the original data source, is given below. These are detailed in the Smart Location Database Technical Documentation and User Guide (Chapman et al. 2021).

- a. **Demographics:** Demographic variables are from the 2018 Census American Community Survey (ACS) (5-Year Estimate) at the block group level. These include population and residential activity (dwelling units and households). Total population of all ages [TotPop] and housing units [CountHU] were tabulated. The percentage of working age

[P_WrkAge] population (between 18 years and 64 years of age) was identified. Auto ownership fields were derived by the EPA from ACS table B08201.

Table 1. Selected features from U.S. EPA Smart Location Database

Feature	Description
'TotPop'	Population, 2018
CountHU, HH	Housing units and occupied housing units, 2018
P_WrkAge	Percent of population that is working aged 18 to 64 years, 2018
AutoOwn0, AutoOwn1, AutoOwn2p	Number of households in CBG that own 0, 1 or 2+ automobiles, 2018'
'Workers'	Count of workers in CBG (home location), 2017
R_LowWageWk, R_MedWageWk, R_HiWageWk	Count of workers with monthly earnings in brackets of \$1250 or less, \$1250-\$3333, and \$3333 and above, 2017
'TotEmp'	Total employment (work location), 2017
D1A, D1B, D1C	Gross residential, population and employment density on unprotected land
D2A_JPHH	Jobs per household
D2A_WRKEMP	Household Workers per Job, by CBG
D3A, D3B	Street Connectivity Index and Intersection Density

- b. **Employment:** Employment information is based on LEHD LODES (LEHD Origin-Destination Employment Statistics) data. The employment variables report job activity and worker information for each CBG. Total employment [TotEmp] was summarized for each CBG from the LEHD Workplace Area Characteristics (WAC) tables. The number of workers [Workers] was summarized from LEHD Residential Area Characteristics (RAC) tables, which report employment based on worker residence. The LEHD RAC tables were also referenced to produce wage stratification variables based on worker residence. High wage workers earn more than \$3,333 per month, while low wage workers earn \$1,250 or less per month. Medium wage workers earned between \$1,251 and \$3,333 a month.
- c. **Densities & Diversity:** All density variables summarize population, housing, or employment within a CBG per unprotected CBG acreage. The primary density variables examine residential characteristics such as housing units [D1A], population [D1B] and employment [D1C]. Employment and housing diversity refer to the relative mix of employment and residential development within an analysis zone. These measures act as proxies for land use diversity by quantifying the relative blend of the number of jobs in different employment sectors and residential housing types. The most simplistic of the measures characterize the jobs to household balance [D2a_JpHH] and the workers' residential location to the employment location balance [D2a_WrkEmp] by CBG.
- d. **Urban Design:** Urban design variables measure connectivity or the ability to traverse distances in many directions along a street network. Areas with higher connectivity typically have a gridded street network with shorter block lengths than more disconnected

areas with fewer intersections and longer block lengths. The urban design (D3) variables measure connectivity in terms of street network density and street intersection density. The denominator used in street network density [D3a] and street intersection density [D3b] calculations was total land area. Additionally, street intersection density [D3b] also summarizes total intersection density weighted to emphasize pedestrian and bicycle travel connectivity. The EPA calculates network density (facility miles per square mile) and intersection density (intersections per square mile) for each CBG by dividing summary figures of facility miles by type (Auto-oriented, Multi Modal and Pedestrian-oriented) and intersection total by type and number of legs by the total land area for each CBG.

Once the variables were selected, the data frame was provided as input to a K-Means clustering model with cluster sizes of 5 and 8. A higher number of clusters partition the region more aggressively, so we begin with a low value such as 5 and increase if more granularity is required. The aim is to group together census block groups that have similar characteristics across the selected set of features. The K-means clustering algorithm (MacQueen 1967) partitions data into 'k' clusters by iteratively assigning points to the cluster with the closest centroid and updating centroids based on mean of points in each cluster. The output of the clustering for cluster sizes 5 and 8 are visualized using Kepler.gl (Kepler.gl 2019) and are shown in Figures 1 and 2. Figure 1 shows how the census blocks across the Cincinnati / Northern Kentucky region have been segregated into 5 different clusters based on the underlying attributes. Per-cluster statistics for Cluster 5 are shown in Table 2. This provides insights into the clusters themselves, with the mean values for the attributes in each cluster as well as the total number of constituent CBGs, the number of CBGs without current Red Bike presence and the number of Red Bike stations within each cluster. Although the K-Means algorithm is an unsupervised learning model, observation of the values in this table shed light on how the clusters were separated.

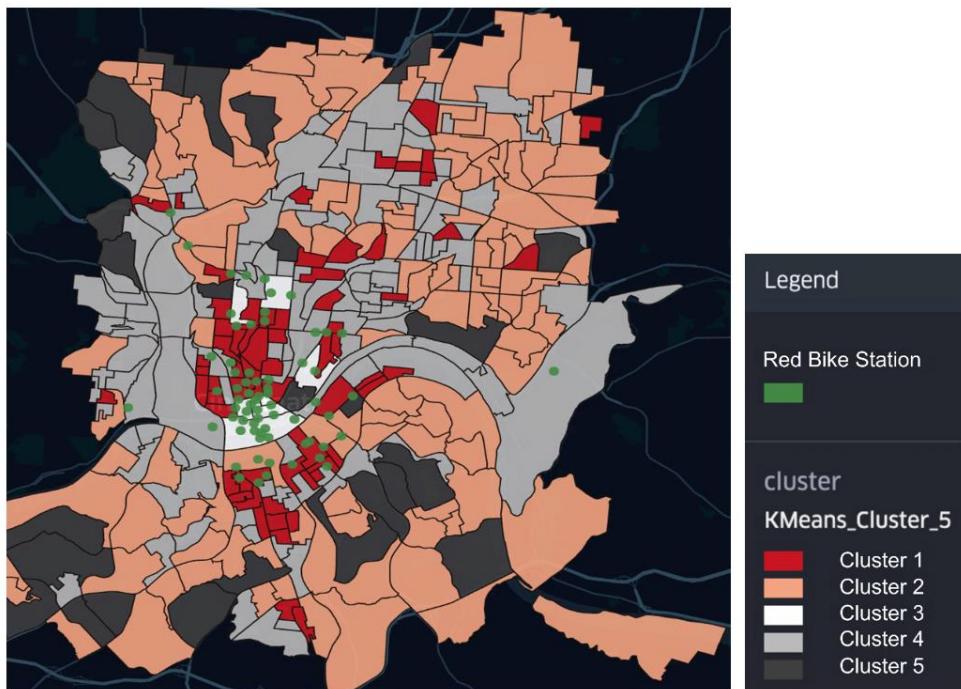


Figure 1. 5 Clusters

Table 2. Per-cluster statistics for 5 clusters

Cluster	No. Of CBGs	No. Of CBGs without Bike Stations	No. Of Bike Stations	D1A (HU/acre)	D1B (pop/a cre)	D1C (emp/acre)	D2A_JP HH	D2A_WRKEMP	D3A	D3B
1	68	50	28	10.36	18.81	7.93	1.0	10.99	33.92	284.62
2	123	120	6	4.24	8.15	2.39	0.9	7.49	16.66	75.56
3	7	1	21	6.5	13.26	121.82	32.44	0.06	39.78	292.59
4	121	110	12	4.05	7.04	4.45	8.36	9.28	21.83	116.55
5	26	25	1	4.03	8.27	2.86	0.75	7.66	15.15	60.7

From the above table, it can be seen that clusters 1 and 3 have the highest concentration of existing Red Bike stations, with 49 out of the total 68 stations localized in this region. On examining the attributes, we understand that CBGs in cluster 3 have a notably high value for attributes D1C, D2A_JPHH, D3A and D3B. This tells us that CBGs in cluster 3 have a high employment density and jobs to household ratio as well as high street connectivity and intersection density scores. Cluster 3 also has the lowest D2A_WRKEMP score which indicates that this area primarily caters to employment rather than housing. By looking at the map in Figure 1 we see that Cluster 3 (in white) corresponds to the Central Business District (CBD) area, which indeed has the characteristics observed above and makes up the most urban neighborhood in the city. Other blocks in cluster 3 that are not part of CBD are also seen towards the north, notably around the University of Cincinnati area, which again has a similar profile in terms of urban activity. Cluster 1 (in red) is then the second ‘best’ cluster, with CBGs that continue to maintain high D3A and D3B values indicative of dense street networks suitable for pedestrian and multimodal activities, but also including more housing units (due to a higher D2A_WRKEMP value of 10.99).

From a Red Bike expansion perspective, we see that Cluster 3 has already been well covered with only 1 out of 7 CBG’s left without any bike stations. Meanwhile, 50 out of 68 CBG’s from cluster 1 do not have current Red Bike presence, with existing stations concentrated only in the urban neighborhoods of Downtown and Over the Rhine. Thus, knowing that cluster 1 is the second best based on the variables examined above, it makes sense to choose CBGs that fall within this cluster when expanding to other neighborhoods. If more granularity is required, a higher number of clusters, such as 8 (Figure 2) can be chosen and a similar iterative process can be followed. In this fashion, one can easily examine and understand the qualities of census blocks across large areas and systematically plan the process of choosing new locations to branch out into.

BIKE NETWORK ANALYSIS

A network is created for the existing BSS, to maintain consistency and align with the absence of historical trip data for prospective new stations, an optimal radius is calculated and applied when creating connections between existing stations. This approach underscores the importance of spatial proximity and ensures a uniform spatial criterion for edge connections across the entire network, regardless of whether a station is existing or newly proposed.



Figure 2. 8 Clusters

Network metrics (Bloch 2023), a set of quantitative measures assessing various aspects of network connectivity and centrality, are then computed to establish a baseline for the existing system. Degree centrality quantifies the number of connections each station has, closeness centrality measures the proximity of a station to others, network efficiency gauges overall connectivity, and eigenvector centrality identifies stations with influential connections. These metrics collectively offer a comprehensive understanding of the current network dynamics and are treated as the baseline metrics of the network.

Census block groups in the Bike Share System (BSS) operation area are further filtered after clustering. The calculation of an optimal radius for bike share system expansion becomes a critical aspect. This radius is determined through a statistical approach, incorporating weighted mean and standard deviation calculations. The mean represents the average number of trips between bike stations and the distance between them, while the standard deviation provides a measure of variability. The chosen formula, $\text{radius} = \text{mean} + 2 * \text{standard deviation}$, ensures that the calculated radius extends beyond the mean, accounting for the variability in trip counts within the bike share system. This statistical framework ensures a nuanced consideration of the distribution of trips between stations, guiding the definition of a meaningful expansion radius for the bike share network.

Once the optimal radius is established, a spatial filtering process is implemented for the bike share system. This involves excluding census blocks falling outside the calculated radius, as well

as those already hosting bike stations. The spatial filtering is conducted based on the centroids of census blocks, employing OSMnx, a Python library for working with OpenStreetMap data, distance queries to determine their proximity to the existing bike share network. This filtering mechanism ensures that the expansion is confined to areas within a specified distance from existing bike infrastructure, mitigating the risk of redundancy and optimizing the use of resources within the bike share system as shown in Figure 3.

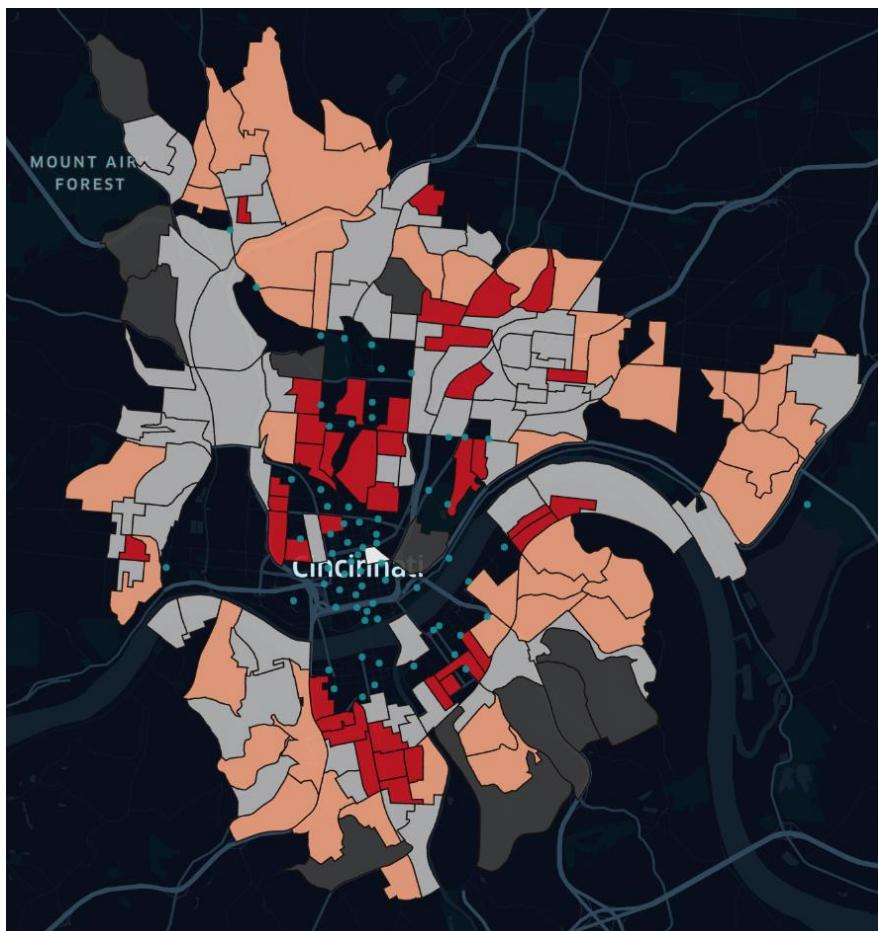


Figure 3. Filtered census blocks for cluster size 5

An iterative approach is adopted where new nodes, corresponding to the centroids of filtered census blocks, are introduced into the network. After each addition, network metrics are recalculated to assess the impact of the new node on the overall system. To gauge the cumulative impact of each new addition, a comparison is made between the metrics of the expanded network and the baseline metrics. This involves calculating the cumulative gain or loss in each network metric, revealing how the addition of new nodes influences the overall network structure. Positive values indicate improvement in the respective metric, while negative values signify a decline. Prospective areas for expansion are identified by considering the top-ranking nodes for each network metric after each iteration. These nodes as shown in Figure 4 are indicative of high centrality or efficiency, represent areas with the most potential for expansion based on diverse network characteristics.

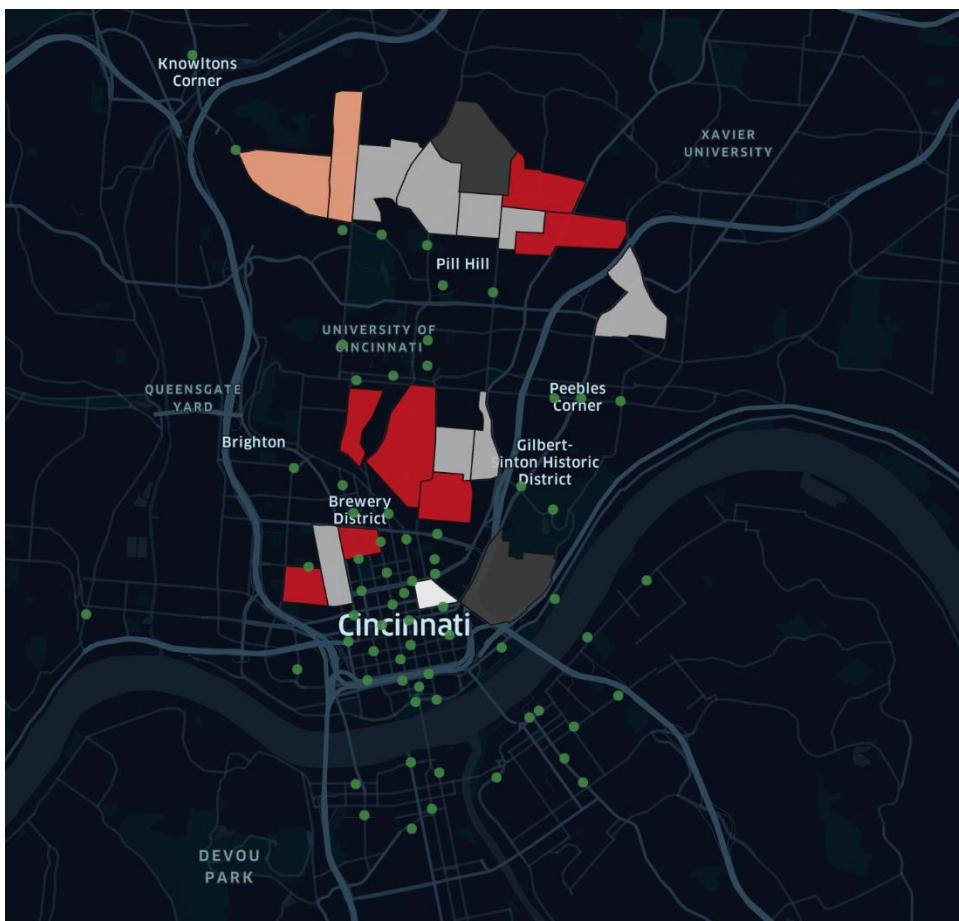


Figure 4. Top 10 census blocks based on eigenvector centrality and degree centrality, closeness centrality, network efficiency.

In our scenario, the top 10 census blocks for degree centrality, closeness centrality, network efficiency remained the same even though the ranking of the blocks varied within the top 10 for the metrics which are highlighted in the southern region as shown in Figure 4, while the northern region of Figure 4 depicts the eigenvector centrality. If the goal is to expand out of the network, we will use eigenvector centrality and pick a census block in the northern region, whereas if the goal is to increase the connectivity between existing stations and increase efficiency, we pick a census block in the southern region.

AMENITY AND TRANSIT CONNECTIVITY

The census blocks that were identified based on the cluster and network analysis are chosen for further examination from an amenities and transit perspective. The clusters that exhibit a higher density of existing bike stations are chosen for this analysis (cluster 1 and cluster 3). It is known that amenities and transit are key drivers in multi modal mobility and hence it makes sense to place bike stations in a way that provides strategic access to these resources.

Upon determining a specific census block, GTFS data is employed to ascertain the location of transit stops within the selected census block, providing valuable insights into the area's accessibility through public transportation. Simultaneously, OSMnx is utilized to retrieve

information on the density of amenities within the block, encompassing diverse urban features such as commercial establishments, public facilities, and recreational spaces. Proximity to transit stop and density of amenities within the census block contributes towards the final selection of the new station location (Rixey 2013).

CONCLUSIONS

This study explores the use of data-driven methods to support the expansion of bike share programs. K-means clustering was applied to model the characteristics of the underlying built environment and segregate census blocks into different clusters. By combining spatial considerations, statistical analyses, and network science principles, our approach further provides valuable insights into the complex interaction between spatial structures and urban mobility systems centered around the bike share system. Specifically, it shows how choosing specific census blocks would impact the current system highlighted by the network metrics, eigenvector centrality depicting the areas to expand outwards of the system and degree centrality, closeness centrality and network efficiency to increase the network connectivity and efficiency and filling gaps in the existing network. This approach enables a systematic process for the process of expansion and/or filling gaps in the network. The stepwise approach of filtering census blocks of interest based on multiple criterion such as socioeconomic, demographics, network efficiency, and transit and amenity connectivity ensure that the expansion process is fine-tuned to meet multiple sets of requirements and be of maximal benefit to the end-users of the bike share system.

ACKNOWLEDGEMENT

We gratefully acknowledge Red Bike for providing ridership data of the bike share system, the U.S. Environmental Protection Agency for making the Smart Location Database freely available to the public and the authors from UD4H and U.S. EPA for providing technical documentation for the Smart Location Database.

REFERENCES

- Bloch, F., Jackson, M., and Pietro, T. (2023). "Centrality Measures in Networks". *Soc. Choice Welf.* 2023, 61, 413–453.
- Cervero, R., and Kockelman, K. (1997). "Travel Demand and the 3Ds: Density, Diversity, and Design". *Transportation Research Part D.* 2 (3): 199-219.
- Chapman, J., Fox, E. H., Bachman, W., Frank, L. D., Thomas, J., and Reyes, A. R. (2021). *EPA smart location database technical documentation and user guide - Version 3.0.*
- Boeing, G. (2017). "OSMnx: New methods for acquiring, constructing, analysing, and visualising complex street networks" *Computers, Environment and Urban Systems*, vol 65, pp. 126-139.
- Kepler.gl. (2019). <<https://github.com/keplergl/kepler.gl>>.
- MacQueen, J. B. (1967). "Some methods for classification and analysis of multivariate observations". In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297). California: University of California Press.

- Rixey, R. A. (2013). Station-Level Forecasting of Bikesharing Ridership: Station Network Effects in Three U.S. Systems. *Transportation Research Record*, 2387(1), 46–55.
- USDOT FHWA (Federal Highway Administration). (2018). *Guidebook for Measuring Multimodal Network Connectivity*.
- USDOT FHWA (Federal Highway Administration). (2016). *Guidebook for Developing Pedestrian & Bicycle Performance Measures*.
- USEPA, Office of Policy, Office of Sustainable Communities (Publisher). (2021). “US Environmental Protection Agency Smart Location Database”,
[<https://catalog.data.gov/dataset/smart-location-database1>](https://catalog.data.gov/dataset/smart-location-database1).

Privacy-Preserved Federated Reinforcement Learning for Autonomy in Signalized Intersections

Negar Asadi¹; Seyed Hamid Hosseini²; Mahdi Imani³; Daniel P. Aldrich⁴;
and Seyedeh Fatemeh Ghoreishi⁵

¹Ph.D. Student, College of Engineering, Northeastern Univ. Email: asadi.n@northeastern.edu

²Ph.D. Student, College of Engineering, Northeastern Univ.

Email: hosseini.ha@northeastern.edu

³Assistant Professor, College of Engineering, Northeastern Univ.

Email: m.imani@northeastern.edu

⁴Professor, College of Social Sciences and Humanities, Northeastern Univ.

Email: d.aldrich@northeastern.edu

⁵Assistant Professor, College of Engineering and Khoury College of Computer Sciences,

Northeastern Univ. (corresponding author). Email: f.ghoreishi@northeastern.edu

ABSTRACT

The control of autonomous vehicles (AV) within complex signalized intersection scenarios poses a substantial challenge, requiring a significant amount of data for effective training of autonomous driving models. This challenge is compounded by limitations in relying solely on a single AV for data, which may fall short of capturing the diverse and dynamic nature of complex intersection settings. The advent of connected and autonomous vehicles (CAVs) introduces a transformative concept, enabling real-time traffic information exchange among interconnected vehicle networks. This shift empowers CAVs to exhibit vastly improved driving behaviors, leveraging shared information for enhanced accuracy, reliability, and efficiency. However, the imperative of preserving user privacy and ensuring data security in the context of AI-enabled CAVs remains a significant challenge. This paper proposes a federated reinforcement learning framework designed to establish a privacy-preserving knowledge-sharing strategy. The proposed framework efficiently aggregates and reuses knowledge learned by diverse CAVs, operating in varying intersection environments. This collective knowledge enhances decision-making precision and adaptability, contributing to safer and more efficient autonomous driving at signalized intersections.

KEYWORDS: Connected and Autonomous Vehicles (CAVs), Signalized Intersections, Reinforcement Learning, Federated Learning.

I. INTRODUCTION

In recent years, the field of autonomous vehicles (AV) has witnessed unprecedented advancements, with a significant focus on developing intelligent systems capable of navigating complex environments safely [1-4]. As the demand for more autonomous solutions has surged, the limitations of traditional rule-based approaches have become increasingly apparent, prompting a paradigm shift towards more adaptive and learning-based methods. Reinforcement Learning (RL) [5,6] has emerged as a powerful tool for addressing the intricate challenges associated with autonomous vehicle control. Unlike rule-based approaches that rely on pre-

defined instructions, RL empowers vehicles to learn optimal behaviors through interaction with their surroundings. This learning paradigm enables autonomous vehicles to adapt to diverse and dynamic scenarios, enhancing their ability to navigate through unpredictable situations and make informed decisions in real time.

Navigating complex environments successfully poses a significant challenge for autonomous driving, requiring a substantial amount of data for their training models. This challenge is exacerbated by the poor sample efficiency inherent in RL techniques, where a considerable volume of data or experiences is needed to determine optimal decision-making processes [7-10]. This issue becomes particularly more challenging in practical scenarios where resource-intensive, expensive, and time-consuming interactions are needed. Furthermore, the absence of a reliable computer simulation for real-world situations amplifies the challenge, requiring RL algorithms to rely on real data to learn. The demand for extensive real-world data, especially in critical domains like autonomous vehicle control, creates a major hurdle.

The advent of Connected and Autonomous Vehicles (CAVs) marks a paradigm shift, introducing a concept that revolves around the real-time exchange of traffic information among interconnected vehicle networks. This shift from traditional reliance on a driver's limited observation fundamentally alters how vehicles navigate their surroundings. By leveraging shared information, these vehicles can exhibit vastly improved driving behaviors characterized by enhanced accuracy, reliability, and efficiency. In particular, the integration of CAVs introduces a transformative concept that can significantly alleviate the data limitations in the training of RL approaches. CAVs facilitate the real-time exchange of knowledge among vehicle networks, allowing them to learn from shared experiences. By leveraging this collective knowledge, CAVs can enhance their decision-making performance, effectively reducing the need for large volumes of additional data.

Nevertheless, the imperative of preserving user privacy and ensuring data security presents a substantial challenge in the context of AI-enabled CAVs. As these vehicles leverage advanced artificial intelligence for real-time traffic information exchange and collaborative learning, the need to safeguard user data becomes paramount [11-16]. Striking a delicate balance between enabling information sharing for enhanced decision-making and preserving users' personal information poses a critical dilemma. Addressing the challenge of user privacy and data security in the realm of AI-enabled CAVs is pivotal for realizing the full potential of these transformative technologies in a manner that is both effective and responsible for individual privacy concerns.

Federated learning [17-20] has emerged as a promising paradigm for sharing knowledge among multiple agents without compromising individual privacy. In this decentralized learning approach, instead of sending raw data to a central server, the learning model is sent to the data sources, which locally compute updates based on their respective datasets. These updates are then aggregated to create a global model without exposing sensitive information. Federated learning thus enables collaborative learning across a network of autonomous vehicles while preserving the privacy of each contributing vehicle. This paper develops a federated reinforcement learning technique for knowledge sharing among heterogeneous autonomous vehicles with potentially different transition probabilities. The privacy-preserving aspect of the proposed federated reinforcement learning framework is aimed at safeguarding sensitive information while enabling collaborative learning among autonomous vehicles. One key aspect is the decentralized nature of the proposed learning process, where each autonomous vehicle retains control over its local data and computes updates independently. This approach shares only the learned policies among vehicles and ensures that raw or sensitive data, such as location

information or driving behavior, is not transmitted or stored centrally, reducing the risk of privacy breaches. The paper is concerned with the control of autonomous vehicles approaching signalized intersections, where vehicles are assumed to perform the same tasks with some discrepancies in their transitions in the environment.

The rest of the paper is organized as follows. In Section II, a detailed description of the proposed framework, including the problem statement, MDP modeling of the problem, and the proposed federated reinforcement learning action control policy, is provided. This is followed by Section III, which presents the numerical experiments demonstrating the performance of the proposed federated reinforcement learning compared to standard reinforcement learning policy. Finally, Section IV contains the concluding remarks.

II. PROPOSED FRAMEWORK

A. Problem Statement

In this study, we explore the scenario of an autonomous vehicle approaching a signalized intersection, aiming to regulate the series of actions undertaken by the vehicle to navigate the intersection safely amid diverse sources of uncertainty. Consider an autonomous vehicle situated at a distance d from the traffic light, moving towards the intersection with a velocity of v and an acceleration of a . The dynamic model describing the vehicle's behavior within a time interval of Δt can be expressed as [21]:

$$\begin{aligned} d' &= d - \frac{1}{2}a\Delta t^2 - v\Delta t + n_d, \\ v' &= v + a\Delta t + n_v, \end{aligned} \quad (1)$$

where $n_d \sim p_d$ and $n_v \sim p_v$ characterize the discrepancies of the distance and the velocity of the vehicle after executing acceleration a from the real values of distance and velocity. Several factors contribute to these deviations between the model and the actual system dynamics, encompassing approximations inherent in the model, distinctions among different vehicles, interactions involving road-tire friction, the number of passengers aboard, variations in road curvature, and various other relevant factors.

Let $\phi \in \{G, Y, R\}$ denote the current phase of the traffic light, corresponding to the green, yellow, and red colors. The duration of traffic phases often exhibits stochastic behavior, influenced by factors such as the time of day, weather, and traffic conditions. We denote the anticipated duration of the green, yellow, and red traffic signals as T_G , T_Y , and T_R , each accompanied by additive noise $n_\phi \sim p_\phi$ that models the uncertainty in the phase durations. Let t_ϕ represent the elapsed time since the last transition of the traffic light. The feasible values for t_ϕ are determined as follows:

$$\begin{aligned} \phi = G : t_\phi &\in [0 \ T_G + n_\phi] \\ \phi = Y : t_\phi &\in [0 \ T_Y + n_\phi] \\ \phi = R : t_\phi &\in [0 \ T_R + n_\phi] \end{aligned} \quad (2)$$

where, in a time interval of Δt , a transition occurs if $t_\phi + \Delta t$ exceeds the duration of the current phase. This transition follows the cyclic sequence of G → Y → R → G. If $t_\phi + \Delta t$ does not surpass the duration of the current phase, there is no change in the traffic light phase after the time interval Δt . Figure 1 illustrates the system at a specific moment when an autonomous vehicle, positioned at a distance d and moving with velocity v , approaches an intersection with a green traffic light, and the elapsed time is 15 s.

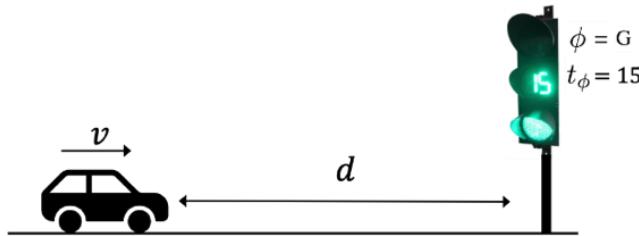


Figure 1. A schematic representation of the problem where an autonomous vehicle with distance d and velocity v approaches a green traffic light with the elapsed time of 15 s.

B. Markov Decision Process (MDP) Modeling

In this part, we formulate the problem as a Markov Decision Process (MDP), a mathematical framework essential for decision-making in complex systems. An MDP is characterized by the tuple $\langle \mathcal{X}, \mathcal{A}, \mathcal{T}, R, \gamma \rangle$, encompassing potential states $\mathbf{x} \in \mathcal{X}$ and feasible actions $\mathbf{a} \in \mathcal{A}$ that the agent can undertake. The transition probability, denoted as $\mathcal{T}(\mathbf{x}', \mathbf{x}, \mathbf{a}) = p(x' | x, a)$, outlines the likelihood of transitioning to state \mathbf{x}' when executing action \mathbf{a} in state \mathbf{x} , representing the system dynamics. The reward function $R(\mathbf{x}, \mathbf{a})$ quantifies the reward associated with taking action \mathbf{a} in state x . Finally, $\gamma \in (0,1)$ serves as the discounting factor, prioritizing immediate rewards over future ones. The agent's objective is to identify an optimal control policy $\pi^*: \mathcal{X} \rightarrow \mathcal{A}$ that maximizes the expected discounted accumulated reward, defined as:

$$\pi^*(\mathbf{x}) = \operatorname{argmax}_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k R(\mathbf{x}_k, \pi(\mathbf{x}_k), \mathbf{x}_{k+1}) | \mathbf{x}_0 = \mathbf{x} \right], \quad (3)$$

for $\mathbf{x} \in \mathcal{X}$. In the following parts, we formulate and construct each component of the MDP tuple for the control of an autonomous vehicle approaching a signalized intersection.

a) State Space: Let $\mathbf{x}_k \in \mathcal{X}$ denote the state vector at time step k ($k = 1, 2, \dots$), where each time step evolves in two sequential steps with the time interval of Δt . The state vector incorporates the current status of both the autonomous vehicle and the traffic light, expressed as:

$$\mathbf{x}_k = [d_k \ v_k \ \phi_k \ t_{\phi_k}]^T, \quad (4)$$

where d_k is the distance of the vehicle from the traffic light, v_k is the speed of the vehicle, ϕ_k is the phase of the traffic light, and t_{ϕ_k} is the time elapsed from the last transition of the traffic light, all at time step k .

b) Action Space: The action space consists of the acceleration a_k executed by the autonomous vehicle at time step k . The goal here is to sequentially determine the desired acceleration taken by the autonomous vehicle for traversing a signalized intersection safely.

c) Transition Model: Given the state \mathbf{x}_k and the vehicle's acceleration at time k , i.e., a_k , the state at time step $k + 1$ can be obtained as:

$$\mathbf{x}_{(k+1)} \sim p(\mathbf{x}_{k+1} | \mathbf{x}_k, a_k). \quad (5)$$

This represents the general state transition, illustrating the Markovian assumption wherein the state at time step $k + 1$ depends solely on the state and acceleration at time k . The state at time step $k + 1$ according to (1)-(2) can be obtained as [22]:

$$\begin{bmatrix} d_{k+1} \\ v_{k+1} \\ \phi_{k+1} \\ t_{\phi_{k+1}} \end{bmatrix} = \begin{bmatrix} d_k - \frac{1}{2} a_k \Delta t^2 - v_k \Delta t + n_{d,k} \\ v_k + a_k \Delta t + n_{v,k} \\ \phi_k \mathbf{1}_{[t_{\phi_k} + \Delta t \leq T_{\phi_k} + n_{\phi,k}]} + \phi_k^+ \mathbf{1}_{[t_{\phi_k} + \Delta t > T_{\phi_k} + n_{\phi,k}]} \\ (t_{\phi_k} + \Delta t) \mathbf{1}_{[t_{\phi_k} + \Delta t \leq T_{\phi_k} + n_{\phi,k}]} + t \mathbf{1}_{[t_{\phi_k} + \Delta t > T_{\phi_k} + n_{\phi,k}]} \end{bmatrix}, \quad (6)$$

where $\mathbf{1}$ is an indicator function, $n_{d,k} \sim p_d$, $n_{v,k} \sim p_v$, $n_{\phi,k} \sim p_\phi$, ϕ_k^+ represents the next phase from ϕ_k according to the phase transition G → Y → R → G, and $t \sim \mathcal{U}(0, \Delta t)$ to include the elapsed time in the duration of Δt where \mathcal{U} indicates a uniform distribution. Thus, the system dynamics can be represented as:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, a_k) + \mathbf{n}_k \quad (7)$$

where $\mathbf{f}(\cdot, \cdot)$ is the function reflecting the system dynamics and \mathbf{n}_k contains the noises associated with all state variables.

d) Reward Model: In this problem, our aim is to avoid the autonomous vehicle running a red light (RRL) and encourage it to pass the intersection as soon as possible while the phase of the traffic light is green or yellow. We assign a large positive reward, denoted as R_{term} , when the vehicle successfully passes the intersection—the terminal state—during a green or yellow traffic light phase. Furthermore, if the vehicle violates the red light, it incurs a large negative penalty, represented by R_{RRL} . Additionally, the vehicle is expected to minimize the number and value of actions required to cross the intersection. Consequently, a negative reward, denoted as R_{act} , is assigned for each action the vehicle takes. Another negative reward, R_{vel} , is implemented if the vehicle's velocity surpasses the allowed maximum speed limit. Furthermore, to ensure smooth acceleration changes for the autonomous vehicle, a penalty term R_{smooth} is considered for significant differences in consecutive actions executed by the agent. Combining these factors, the comprehensive reward function is expressed as:

$$\begin{aligned} R(\mathbf{x}_k, a_k, \mathbf{x}_{k+1}) = & R_{\text{term}} \mathbf{1}_{d_{k+1} \leq 0 \wedge \phi_{k+1} = \text{G} \vee \text{Y}} \\ & + R_{\text{RRL}} \mathbf{1}_{d_{k+1} \leq 0 \wedge \phi_{k+1} = \text{R}} \\ & + R_{\text{vel}} \mathbf{1}_{v_{k+1} < v_{\max}} \\ & + R_{\text{smooth}} \mathbf{1}_{|a_k - a_{k-1}| < \text{thr}_{\text{smooth}}} + R_{\text{act}} \times |a_k|, \end{aligned} \quad (8)$$

where \wedge and \vee represent logical AND and OR, v_{\max} is the maximum allowed speed, and $\text{thr}_{\text{smooth}}$ is the smoothness threshold for the difference of consecutive actions.

C. Reinforcement Learning (RL) Policy

Following the MDP modeling of the problem, we employ the Q-learning framework to approximate the stationary control policy outlined in (3). Q-learning is a model-free Reinforcement Learning technique that revolves around learning an action-value function, denoted as Q . The Q value represents the anticipated discounted reward for executing action a at state \mathbf{x} and subsequently following policy π [23]. The primary goal of Q-learning is to estimate the Q values for an optimal policy. In the context of the known state process described in (6), the algorithm iteratively calculates the reward function for a given control strategy by generating multiple sample trajectories of the system dynamics along with their associated rewards. The Q values in Q-learning are learned through learning based on simulated data generated from the state model. Assuming that in the n th episode of the Q-learning algorithm, the system is in state \mathbf{x}^j , and by executing a control input, i.e., acceleration a , the system transitions to state \mathbf{x}^i in episode $n + 1$. All Q values at time $n + 1$ remain consistent with those at time n , except for the Q value corresponding to state \mathbf{x}^j and acceleration a , which undergoes an update as follows:

$$Q_{n+1}(\mathbf{x}^j, a) = (1 - \alpha)Q_n(\mathbf{x}^j, a) + \alpha \left(R(\mathbf{x}^j, a, \mathbf{x}^i) + \gamma \max_{a' \in \mathcal{A}} Q_n(\mathbf{x}^i, a') \right), \quad (9)$$

where $0 < \alpha < 1$ is the learning rate, and $0 < \gamma < 1$ is the discounting factor. In [23], it has been shown that Q-learning converges to the optimum action-values with probability 1 as long as all actions are repeatedly sampled in all states and the action-values are represented discretely, i.e., $Q_n(\mathbf{x}, a) \rightarrow Q^*(\mathbf{x}, a)$, for all $\mathbf{x} \in \mathcal{X}$ and $a \in \mathcal{A}$ as $n \rightarrow \infty$. After obtaining the stationary action values $Q^*(\mathbf{x}, a)$, the autonomous vehicle is capable of decision-making for executing the acceleration at any state \mathbf{x}_k as:

$$a_k = \operatorname{argmax}_{a \in \mathcal{A}} Q^*(\mathbf{x}_k, a) \quad (10)$$

This action policy, referred to as the “Independent RL Policy” in the numerical experiments, is shown in Figure 2.

D. Proposed Federated Reinforcement Learning Policy

The RL techniques often exhibit a substantial need for large volumes of data to train and optimize models effectively. The learning process in RL relies on the exploration of various states, actions and their consequences, aiming to discover the most rewarding strategies. The demand for extensive data arises from the necessity to generalize across diverse and complex environments, ensuring that the learned policies are robust and adaptable [24]. This need for substantial data becomes particularly challenging in scenarios involving complex decision-making, such as the problem of autonomous vehicle control at intersections considered here, where the algorithms must learn optimal behaviors under a multitude of conditions. However, the acquisition of such datasets can be resource-intensive, time-consuming, and, in some cases, impractical.

Federated reinforcement learning addresses these challenges by allowing individual agents to collect and learn from their specific experiences locally without centralized data collection and the need to share raw data. In this framework, instead of sending data to a central server, the learning model is sent to individual devices, each of which computes updates based on its specific local dataset. These updates are then aggregated to create a global model. By leveraging the collective intelligence of diverse agents while preserving privacy, this decentralized approach not only mitigates the impracticality and resource intensity of central data collection but also promotes more efficient and adaptable learning.

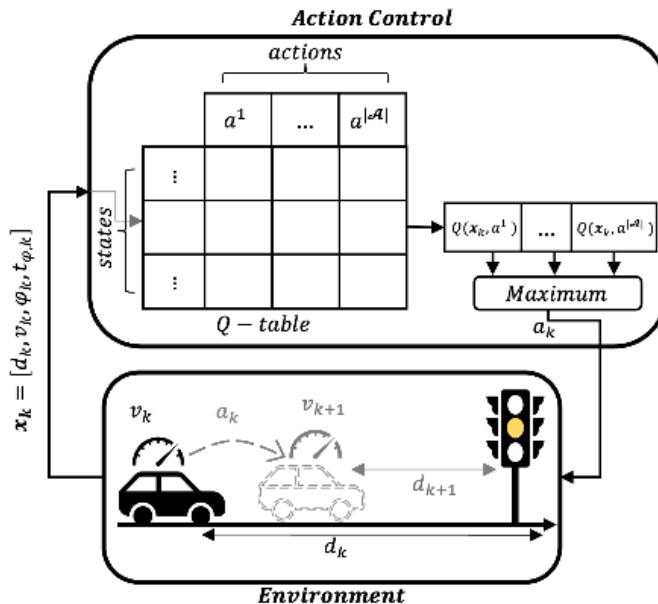


Figure 2. A depiction of the RL policy for the control of an autonomous vehicle in a traffic light environment.

Federated learning offers several advantages for the control of autonomous vehicles at signalized intersections. First, it ensures privacy preservation by allowing vehicles to learn locally without sharing sensitive raw data. This decentralized approach also facilitates adaptation to the diverse dynamics of signalized intersections, accommodating varying vehicle dynamics, traffic conditions, and signal timings. Moreover, federated learning efficiently addresses the heterogeneity across different intersections, enabling collaborative learning in heterogeneous environments. Real-time adaptation to changing traffic scenarios is another key benefit, allowing vehicles to continuously update their control policies based on the most recent local observations. By reducing dependence on centralized infrastructure, federated learning promotes efficiency and scalability, while the collective intelligence harnessed from vehicle collaboration enhances decision-making for more informed, efficient, and safe interactions at intersections. Therefore, the goal of this paper is to enable N autonomous vehicles to share their knowledge and experiences while preserving privacy through the federated reinforcement learning framework described below.

Consider N autonomous vehicles sharing the same state and action spaces (\mathcal{X} and \mathcal{A}), and the same reward functions (R), as described in II-B, while they might have different dynamics and transition probabilities. We indicate the i th vehicle's MDP model using the following tuple:

$\langle \mathcal{X}, \mathcal{A}, \mathcal{T}_i, R, \gamma \rangle$, where \mathcal{T}_i represents the transition probability of vehicle i . Let $Q_L^i(\mathbf{x}, a)$ be the latest local Q-values of vehicle i for $i = 1, \dots, N$. Then, if a new vehicle $N + 1$ starts learning within its environment, we use the global federated Q-values, which are the average of the local Q-values of N vehicles as the baseline for the initial behavior policy of the new vehicle as:

$$Q_G^{N+1} = \frac{1}{N} \sum_{i=1}^N Q_L^i. \quad (11)$$

This process ensures that the learning of the new vehicle is informed by the collective intelligence captured in the global Q-values established through federated learning. Hence, the actions of vehicle i are anticipated to carry more information compared to alternative policies, such as those relying solely on its local Q-values, denoted as an independent RL policy here. This leads to interactions that are more efficient in terms of sample utilization, facilitating accelerated learning. Regular updates to the global Q-values can be performed after each interaction, episode, or at scheduled intervals. This ensures that the most recent local Q-values contribute to the refinement of the global policy, guaranteeing that the system benefits from the latest knowledge acquired by the vehicles during their interactions. Using the ϵ -greedy as the behavior policy, the global Q-values will be used for action selection when the vehicle is at state \mathbf{x}_n as:

$$a_n \sim \begin{cases} \underset{a \in \mathcal{A}}{\operatorname{argmax}} Q_G^{N+1}(\mathbf{x}_n, a) & \text{with probability } (1 - \epsilon) \\ \text{Random}\{a^1, \dots, a^{|\mathcal{A}|}\} & \text{with probability } \epsilon \end{cases}, \quad (12)$$

where $0 \leq \epsilon \leq 1$ represents the parameter of the ϵ -greedy policy, controlling the level of exploration during the learning process. Given that ϵ is typically small, the vehicle tends to favor actions linked to the highest global Q-value. This action policy, referred to as the “Proposed Federated RL Policy” in the numerical experiments, is shown in Figure 3.

The proposed federated reinforcement learning model exhibits scalability and applicability to real-world traffic systems beyond simulations through several key features. Firstly, its decentralized nature allows for scalability as the number of autonomous vehicles increases. Since each vehicle learns independently and contributes to the collective intelligence through federated learning, the model can accommodate a growing fleet of vehicles without imposing significant computational or communication burdens. Additionally, the framework's modular design enables easy integration with existing traffic management systems and infrastructure. By leveraging real-time traffic data and communication networks, the model can adapt to dynamic traffic conditions and infrastructure changes in real-world scenarios.

III. NUMERICAL EXPERIMENTS

In this section, we examine the performance of the proposed federated RL method compared to the independent RL policy. We employ a simulation environment replicating a signalized intersection scenario with an approaching vehicle. As detailed in II-B, this simulation environment encompasses the distance of the vehicle from the intersection, its velocity, the current phase of the traffic light, and the elapsed time since the last phase transition as the system state. In all the experiments, the values associated with the variables in the model are $T_G = 10$ s, $T_Y = 5$ s, $T_R = 10$ s, $\Delta t = 1$ s and $a \in \mathcal{A} = \{-3, -2, -1, 0, 1, 2, 3\}$. The discretized values of

distance from the intersection and the vehicle's velocity are considered as $-8:8:120$ and $0:1:15$, respectively. The Q-learning parameters are set to $\alpha = 0.0025$, $\gamma = 0.99$, and $\epsilon = 0.1$. The reward function parameters are $R_{\text{term}} = 100$, $R_{\text{RRL}} = -200$, $R_{\text{vel}} = -100$, $R_{\text{act}} = -1$, $R_{\text{smooth}} = -10$, and $\text{thr}_{\text{smooth}} = 1$. We consider $N = 4$ vehicles with the following uncertainties associated with the state transition probabilities:

$$\begin{cases} p_d^1 = \mathcal{N}(0,1.2), & p_v^1 = \mathcal{N}(0,0.5), & p_\phi^1 = \mathcal{N}(0,0.05) \\ p_d^2 = \mathcal{N}(0,1.1), & p_v^2 = \mathcal{N}(0,0.6), & p_\phi^2 = \mathcal{N}(0,0.04) \\ p_d^3 = \mathcal{N}(0,1), & p_v^3 = \mathcal{N}(0,0.55), & p_\phi^3 = \mathcal{N}(0,0.06) \\ p_d^4 = \mathcal{N}(0,0.8), & p_v^4 = \mathcal{N}(0,0.4), & p_\phi^4 = \mathcal{N}(0,0.055) \end{cases},$$

and the 5th vehicle which uses the Q-values learned by these 4 vehicles, has the following state transition probabilities: $p_d^5 = \mathcal{N}(0,0.9)$, $p_v^5 = \mathcal{N}(0,0.45)$ and $p_\phi^5 = \mathcal{N}(0,0.045)$.

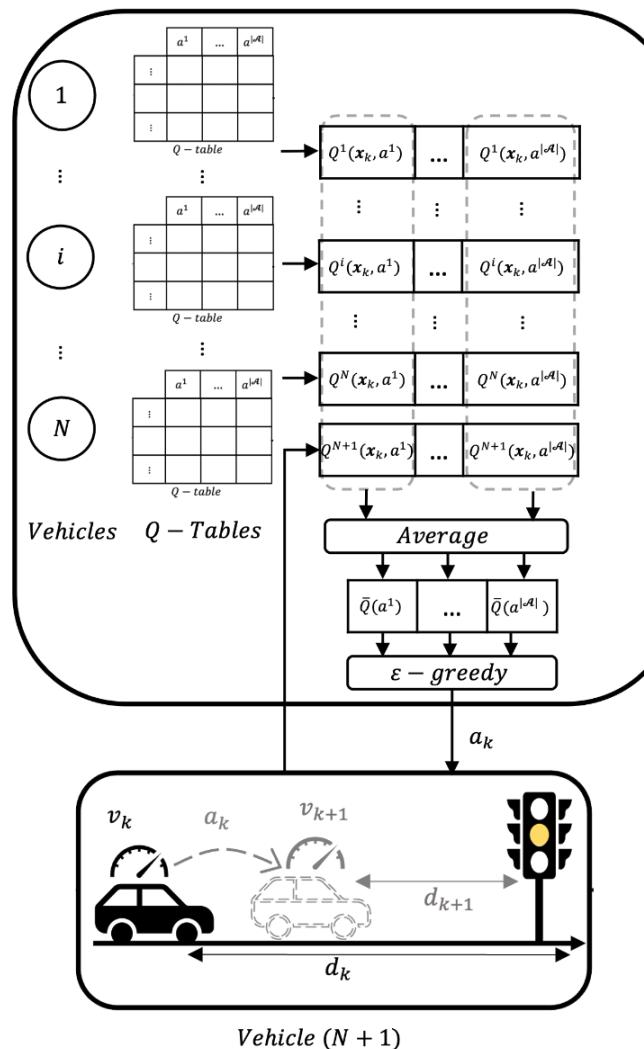


Figure 3. A schematic diagram of the proposed federated RL policy for the control of an autonomous vehicle in a signalized intersection.

Figure 4 represents the average accumulated reward per episode of learning in the proposed federated RL policy vs the independent RL policy. The results are averaged over 2000 runs starting from random initial states. The blue curve represents the result of the proposed federated Q-learning method, and the red curve indicates the independent Q-learning policy result. As can be seen, the proposed policy significantly outperforms the independent RL policy in the initial stage of the learning process, and as the number of episodes increases, the two policies perform similarly in terms of the average accumulated reward. This result shows the importance of the federated RL policy in the early stage of learning, especially in complex scenarios with large state and action spaces. The federated RL policy leverages collective information from all other vehicles, bypassing the need for initial exploration in the learning process which results in lower values of the average accumulated reward.

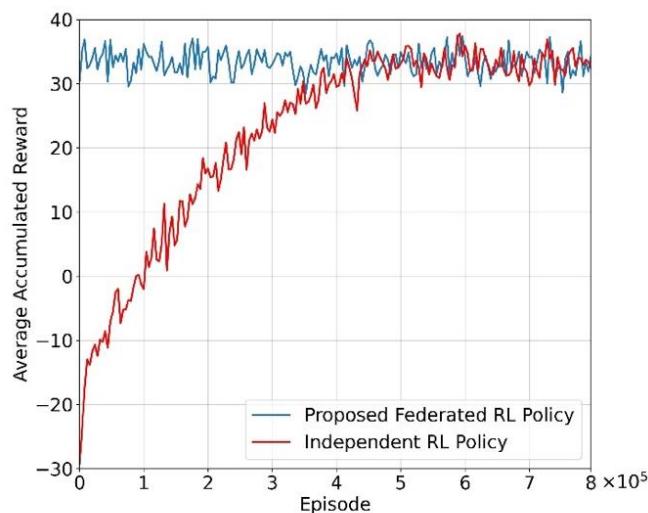


Figure 4. The average accumulated reward per episode of learning in the proposed federated RL policy vs the independent RL policy.

Figure 5 demonstrates the matching rate of the proposed federated RL policy and the independent RL policy with respect to the optimal policy over the learning episodes. It can be seen that the proposed federated RL policy has a high rate of matching with the optimal policy from the early stage. This is achieved by leveraging the knowledge acquired from other vehicles that establishes a solid foundation for the 5th vehicle to follow. It is important to highlight that this achievement is realized solely through the sharing of policies, specifically Q-values, from other vehicles, with no exchange of any additional information.

In this part of the numerical experiments, we aim to demonstrate the impact of the deviation of state transitions of the new vehicle from those of other vehicles whose policies have been incorporated into the federated RL policy. These deviations can arise from various factors. These may include differences in the vehicles' physical characteristics, environmental conditions, or individual driving behaviors. Additionally, variations in sensors, hardware, or software configurations among the vehicles can contribute to discrepancies in state transitions. Other external factors such as road conditions, traffic density, and unexpected events may also play a role in causing deviations in the state transitions of the new vehicle. In the proposed federated reinforcement learning model, the criterion of "Similar System Dynamics" is crucial in selecting

the participating vehicles. This criterion ensures that vehicles with comparable mechanical, environmental, and operational characteristics are chosen to collaborate in the learning process. Vehicles with similar system dynamics are more likely to encounter analogous situations and respond similarly in various states, thereby providing more relevant and actionable insights for the collective learning framework. By considering vehicles with similar system dynamics, the federated model can effectively leverage collective intelligence to develop robust and adaptable control policies that are applicable across a large fleet of autonomous vehicles.

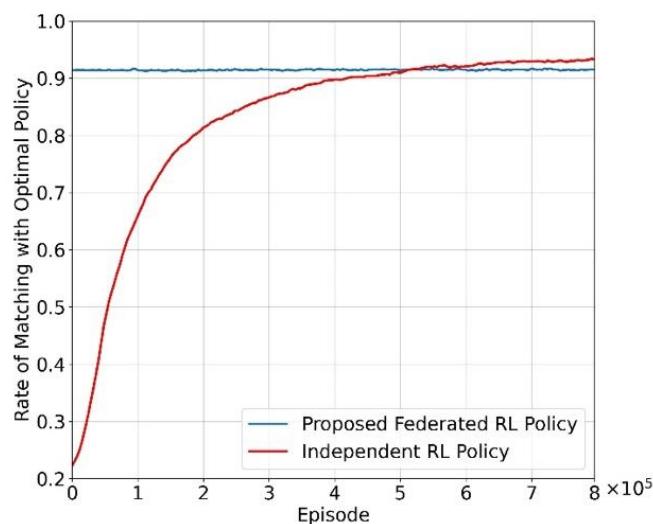


Figure 5. The matching rate of the proposed federated RL policy and the independent RL policy with the optimal policy over the learning episodes.

For this analysis, we consider an additive deviation factor δ in the distance state transition of the new vehicle. We analyze the results for four values of deviation factor, $\delta = 0, -1, -3, -5$. Figure 6 shows the average accumulated reward per episode of learning in the proposed federated RL policy vs the independent RL policy with different levels of deviation factor. As can be seen in the results, an increase in the level of deviation leads to a diminished performance of the proposed federated RL policy compared to the independent RL policy. This indicates the significance of considering vehicles with comparable system dynamics to leverage collective knowledge in the proposed federated RL policy. Otherwise, the knowledge from the group of vehicles integrated into the federated learning will dominate, preventing the new vehicle from recovering the Q-values as effectively as it could in the independent RL policy.

Figure 7 presents the average accumulated reward at episode 4×10^5 of learning in the proposed federated RL policy compared to the independent RL policy with different levels of deviation factor δ . In line with the earlier finding, it can be seen that as the deviation between the new vehicle and the other group of vehicles whose Q-values contribute to the federated learning increases, the average accumulated reward obtained by the proposed federated RL method decreases significantly compared to the independent RL policy. However, in the scenario where there is no discrepancy between the systems' dynamics, i.e., $\delta = 0$, the proposed federated RL policy outperforms the independent RL one in terms of average accumulated reward. As mentioned earlier, these findings emphasize the significance of integrating agents with

comparable system dynamics into the proposed federated RL policy. Identifying similarities among agents and integrating pertinent information will be the focus of our future research.

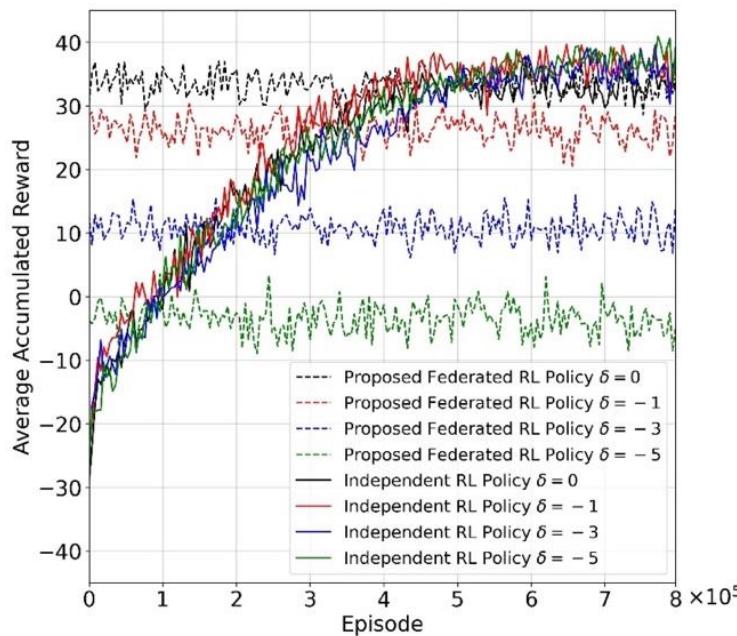


Figure 6. The average accumulated reward per episode of learning in the proposed federated RL policy vs the independent RL policy with different levels of deviation factor.

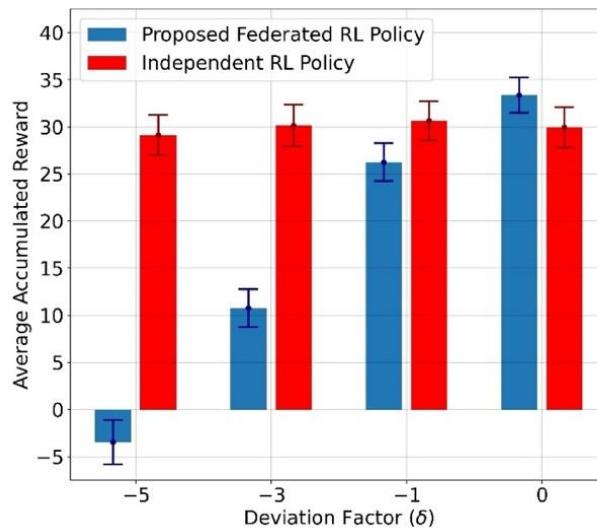


Figure 7. The average accumulated reward at episode 4×10^5 of learning in the proposed federated RL policy vs the independent RL policy with different levels of deviation factor.

IV. CONCLUSIONS

This paper proposed a federated reinforcement learning framework for autonomous vehicles navigating signalized intersections. The framework aimed at sharing knowledge and experiences

among vehicles while preserving privacy. The proposed framework enables the vehicles to leverage collective intelligence by sharing their learned policies instead of sharing sensitive raw data. The results showed that this approach proves beneficial in the early stages of learning, especially in complex scenarios with large state-action spaces. The proposed federated RL policy efficiently utilizes collective information, bypassing the need for initial exploration. However, it has been shown that the effectiveness of the federated learning approach depends on considering vehicles with similar system dynamics, emphasizing the need for future research in identifying and incorporating relevant information for optimal results.

V. ACKNOWLEDGEMENT

This research was supported by the ARMY Research Laboratory under award number W911NF2320179 and the TIER 1 Seed Grant/Proof of Concept Program at Northeastern University.

REFERENCES

- [1] S. Bijjahalli, and R. Sabatini. "A High-Integrity and Low-Cost Navigation System for Autonomous Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 356–369, 2021
- [2] W. Schwarting, J. Alonso-Mora, and D. Rus. "Planning and Decision-Making for Autonomous Vehicles," *Annu Rev Control Robot Auton Syst*, vol. 1, no. 1, pp. 187–210, 2018
- [3] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu. "A survey of deep learning techniques for autonomous driving," *J Field Robot*, vol. 37, no. 3, pp. 362–386, 2020.
- [4] Y. Ma, Z. Wang, H. Yang, and L. Yang. "Artificial intelligence applications in the development of autonomous vehicles: a survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 2, pp. 315–329, 2020
- [5] M. Zhu, Y. Wang, Z. Pu, J. Hu, X. Wang, and R. Ke. "Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving," *Transp Res Part C Emerg Technol*, vol. 117, p. 102662, 2020.
- [6] B. Osiński, et al. "Simulation-based reinforcement learning for real-world autonomous driving," in *2020 IEEE international conference on robotics and automation (ICRA)*, IEEE, 2020, pp. 6411–6418.
- [7] M. Imani, S. F. Ghoreishi, and U. M. Braga-Neto. "Bayesian control of large MDPs with unknown dynamics in data-poor environments," in *Advances in neural information processing systems*, 2018, pp. 8156–8166.
- [8] A. Ravari, S. F. Ghoreishi, and M. Imani. "Optimal Inference of Hidden Markov Models Through Expert-Acquired Data," *IEEE Transactions on Artificial Intelligence*, vol. PP, pp. 1–15, Feb. 2024
- [9] A. Ravari, S. F. Ghoreishi, and M. Imani. "Implicit Human Perception Learning in Complex and Unknown Environments," in *American Control Conference (ACC)*, 2024.
- [10] A. Ravari, S. F. Ghoreishi, and M. Imani. "Optimal Recursive Expert-Enabled Inference in Regulatory Networks," *IEEE Control Syst Lett*, vol. 7, pp. 1027–1032, 2022.
- [11] D. J. Glancy. "Privacy in autonomous vehicles," *Santa Clara L. Rev.*, vol. 52, p. 1171, 2012.
- [12] V. Dhar. "Equity, safety, and privacy in the autonomous vehicle era," *Computer (Long Beach Calif)*, vol. 49, no. 11, pp. 80–83, 2016.

- [13] J. M. Anderson, K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola. *Autonomous vehicle technology: A guide for policymakers*. Rand Corporation, 2014.
- [14] A. Nanda, D. Puthal, J. J. P. C. Rodrigues, and S. A. Kozlov. "Internet of Autonomous Vehicles Communications Security: Overview, Issues, and Directions," *IEEE Wirel Commun*, vol. 26, no. 4, pp. 60–65, 2019
- [15] A. Kazeminajafabadi, S. F. Ghoreishi, and M. Imani. "Optimal Detection for Bayesian Attack Graphs under Uncertainty in Monitoring and Reimaging," in *American Control Conference (ACC)*, 2024.
- [16] A. Kazeminajafabadi, and M. Imani. "Optimal Monitoring and Attack Detection of Networks Modeled by Bayesian Attack Graphs," *Cybersecurity*, Feb. 2023
- [17] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao. "A survey on federated learning," *Knowl Based Syst*, vol. 216, p. 106775, 2021
- [18] S. Niknam, H. S. Dhillon, and J. H. Reed. "Federated Learning for Wireless Communications: Motivation, Opportunities, and Challenges," *IEEE Communications Magazine*, vol. 58, no. 6, pp. 46–51, 2020
- [19] Y. Li, X. Tao, X. Zhang, J. Liu, and J. Xu. "Privacy-Preserved Federated Learning for Autonomous Driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8423–8434, 2022
- [20] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon. "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, 2016.
- [21] M. Emamifar, and S. F. Ghoreishi. "Modeling and State Estimation of Autonomous Vehicles in Signalized Intersections," Feb. 2023.
- [22] M. Emamifar, and S. F. Ghoreishi. "Uncertainty-Aware Reinforcement Learning for Safe Control of Autonomous Vehicles in Signalized Intersections," in *2023 IEEE Conference on Artificial Intelligence (CAI)*, 2023, pp. 81–82.
- [23] C. J. C. H. Watkins, and P. Dayan. "Q-learning," *Mach Learn*, vol. 8, no. 3–4, pp. 279–292, 1992.
- [24] M. Imani, and S. F. Ghoreishi. "Scalable inverse reinforcement learning through multifidelity Bayesian optimization," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 8, pp. 4125–4132, 2021.

Field Study on Unconfined Compressive Strength and Drilling Data of DSM Columns: A Machine Learning Approach

Seyed Meisam Alavi¹; Sajjad Shakeri Talarposhti²; Ahmad Ali Khodaei Ardabili³;
and Milad Aghamolaei⁴

¹Baspar Pey Iranian Company, Tehran, Iran (corresponding author).

Email: m.alavi@basparpey.com

²Baspar Pey Iranian Company, Tehran, Iran. Email: s.shakeri@basparpey.com

³Baspar Pey Iranian Company, Tehran, Iran. Email: a.khodaei@basparpey.com

⁴Baspar Pey Iranian Company, Tehran, Iran. Email: m.aghamolaie@basparpey.com

ABSTRACT

Numerous limitations of proper lands, especially in offshores, bring about outcomes to consider soil improvement methods. The deep soil mixing (DSM) method is an efficient ground improvement procedure using in situ soils for treatment plans. In order to ensure the quality of the treated soil, the DSM elements should be evaluated using unconfined compressive strength (UCS) tests on only 2%–4% of executed columns, based on the recommendations of the FHWA. Therefore, there are significant uncertainties regarding the UCS of the other 96% of the untested DSM elements. To reduce these deficiencies and bring an insight into the uncored DSM columns indirectly, artificial intelligence (AI) and machine learning (ML) methods (XGBoost, random forest, artificial neural network, and lasso regression) were implemented to predict a correlation algorithm between UCS results and several drilling rig machine parameters, including drilling, lifting, rotary speeds, rotary pressures, and the added water during pre-drilling phase. The extensive results of over 3,780 UCS tests and corresponding machine data of a DSM soil improvement project in Hormozgan, Iran, were considered to train and test the ML algorithms. In this case, instead of conducting a probability analysis to find out a DSM column UCS for using in design procedure, the predicted strength of each column by the well-trained ML method can be considered.

Keywords: Deep soil mixing (DSM) method; Field investigation; Unconfined compressive strength (UCS); Artificial intelligence (AI); Machine learning (ML); XGBoost; Random Forest (RF); Artificial neural network (ANN); Lasso Regression (LR).

1. INTRODUCTION

The demands for soil improvement have increased gradually due to the industrialization and demands for construction of structures on unfavorable soil layers, especially for transportation projects. Aiming to solve the problem, different soil improvement methods have been widely used for decreasing unallowed settlements, increasing the bearing capacity, controlling the liquefaction phenomena, or stabilizing weak soils. Several methods for improving soil quality, including high-pressure jet grouting, DSM elements, stone columns, and dynamic compaction are existed in the literature (Han 2015; Nicholson 2014), among them, deep soil mixing (DSM) is a method, by which the in situ soil mixes with the binder to form a homogenous soil cement column. High-speed implementation, efficiency, low cost in comparison with piles, and compatibility with the environment are the positive features of this method. A brief review of the

deep mixing literature shows that The previous contributions focused mostly on the effect soil fabric and mix designs on the strength of columns. (e.g., (Chen et al. 2013; Helson et al. 2017; Jamsawang et al. 2008; Jung et al. 2020; Rashid et al. 2018; Yao et al. 2016; Yi et al. 2019; Yin and Fang 2010)) using common quality control (QC) and quality assurance (QA) procedures.

Kitazume (2022) and Bruce et al. (2013) introduced different controlling methods that are necessary to be sure about the quality of executed columns. Many parameters including soil fabric, execution procedure, amount of injected grout, blade rotation number (BRN), etc., are involved in DSM projects and may affect the quality of the columns. It has been mentioned by Bruce et al. (2013) that about 4% of DSM columns should be cored to check the uniformity and unconfined compressive strength (UCS) tests and these tested columns are the indicator of whole of the site, but the rest of the columns (i.e., 96%) of columns may not be absolutely similar with the gathered information and it casts doubt on the quality of these elements.

To resolve the problem, indirect methods (statistical analyses, soft computing, machine learning, etc) can be employed to clarify the features and strength of columns, for instance by using machine data. Many newly developed rigs are highly instrumented that are capable of measuring data such as drilling pressure, rotary pressure, drilling, and withdrawal speed, etc, all of which may exhibit the soil layer condition during execution time. By dint of these accumulated information some quantitative measurements may be extracted if proper interpretation be done on this dataset. For instance, at a certain level of engine power, when the drilling speed reduces it means that the soil layer strength increases. The literature survey indicates that the number of contributions to this topic is rare, if not at all.

It is worth mentioning that in large projects that number of columns is notable, recording the rig's information provides a considerable amount of data to train statistical equations, the importance of which in the mega projects cannot be overstated. On the other hand, by using soft computing methods and machine learning procedures, it is achievable to correlate the rig data to the desired factors such as UCSs of the cored column and extrapolate them to predict the strength and quality of uncored columns. Using artificial intelligence (AI) methods, are of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science (Jordan and Mitchell 2015), and major breakthroughs can be achieved by using these models in geotechnics.

Based on the aforementioned statements, to fill the gap and investigate the contributions of rig data to assess the quality of executed columns, a comprehensive analysis has been performed in a mega soil improvement project (conducted by Baspar Pey Iranian Company), in which numerous numbers of columns have been cored for QA tests (more than 3780 UCS tests). Based on these results of cored samples and the recorded data of rigs, the correlation between the 15 input parameters regarding pre- and main-drilling phases and corresponding UCS values is investigated. In this regard, different machine learning algorithms have been employed in this study to find the most appropriate strategy that can propose the most fit prediction, including the Extreme Gradient Boosting (XGBoost), Random Forest (RF), Artificial Neural Network (ANN), and Lasso Regression (LR). Therefore, the well-trained algorithm with the lowest error metrics has been implemented in this project for the first time in the literature to predict the quality of all executed DSM elements without conducting UCS tests.

2. DESCRIPTION OF QUALITY CONTROL METHOD

During the execution of DSM columns, data logging systems is used to monitor the rig data. This helps in keeping track of various parameters and ensuring that the mixing operation is

functioning optimally. The collected data is influenced by the interaction of three key components - soil, auger, and slurry. Understanding these interactions is crucial to be insured that the rig is operating well. For instance, if the torque of the device increases, it could indicate that the mixing shear resistance is high. This could mean that the soil grains are still not properly disaggregated, and further action may be required to improve the quality and strength of mixing production.

At present, the FHWA guideline has suggested a complete QC/QA program for consideration in DSM soil improvement projects, which includes conducting UCS tests on 2-4% of executed columns (based on the project significance). Although this method extends these results for the entire improved site, there is still a huge uncertainty regarding the condition of the other 96% DSM elements. Besides, these extended results are only based on statistical concepts and do not represent untested columns. Hence, this study tries to develop an additional QC/QA approach using AI and drilling machine data to reduce these ambiguities and clarify the quality of all executed columns, even without conducting UCS tests. In the proposed QC/QA program, at first, the correlation between the drilling parameters of cored DSM elements and UCS results of these columns has been predicted using various ML methods. Next, the most fitted ML algorithm is used to determine the strength quality and UCS status of untested elements based on their drilling parameters.

The current study presents the results of a big DSM soil improvement project conducted in Hormozgan, Iran by Baspar Pey Iranian (BPI) company. Over 3780 UCS tests were conducted on DSM elements during the QC/QA programs. Furthermore, a Soilmech SR-95 single shaft drilling machine was used to execute DSM elements in this project (Figure 1). The Soilmech SR-95 drilling machine has the capability of accurately recording of various indexes every ten-centimeter steps during the execution of each element. Figure 1 presents a sample of the data log generated by the drilling rig device. By analyzing these data, parameters such as drilling and lifting speed, water consumption, and soil moisture can be effectively evaluated. Monitoring the drilling and lifting speed during the grouting phase is essential for obtaining an accurate binder factor. Based on the contractor previous experiences, the rotational speed and rotary pressure during the return phase are crucial indicators of the shear strength of the soil during mixing. High levels of applied shear energy during mixing indicate poor mixing quality. Another important parameter to consider is the spoil density, which is an effective indicator of the quality of mixing and resistance. Hence, through careful examination of these parameters, it is possible to predict the quality of the mixed soil and cement, as well as the final strength of it. Table 1 presents the descriptions of 15 input and one output (UCS) parameters considered in the current study for data analysis. The statistical values that have been extracted and calculated from the gathered data are also provided in Table 1.

3. AI AND ML TECHNIQUES

Various ML algorithms have been developed to evaluate the correlations between input (also called features) and output parameters for civil engineering problems (e.g., (Anitescu et al. 2019; Dang et al. 2023; Nghia-Nguyen et al. 2023; Nguyen and Wahab 2023; Samaniego et al. 2020; Tran et al. 2023; Wang et al. 2021)). In the current study, four well-known AI and ML algorithms in civil engineering have been considered to assess the relationship between drilling rig data and UCS of DSM elements, including the XGBoost, RF, ANN, and LR. The Python programming language version 3.10 was utilized to develop mentioned ML algorithms. The

dataset was divided into 85% for training and 15% for testing machine learning models. To mitigate overfitting in ANN and XGBoost algorithms, 15% of the training data was further set aside as a validation dataset, and the early stopping technique was implemented in both algorithms.

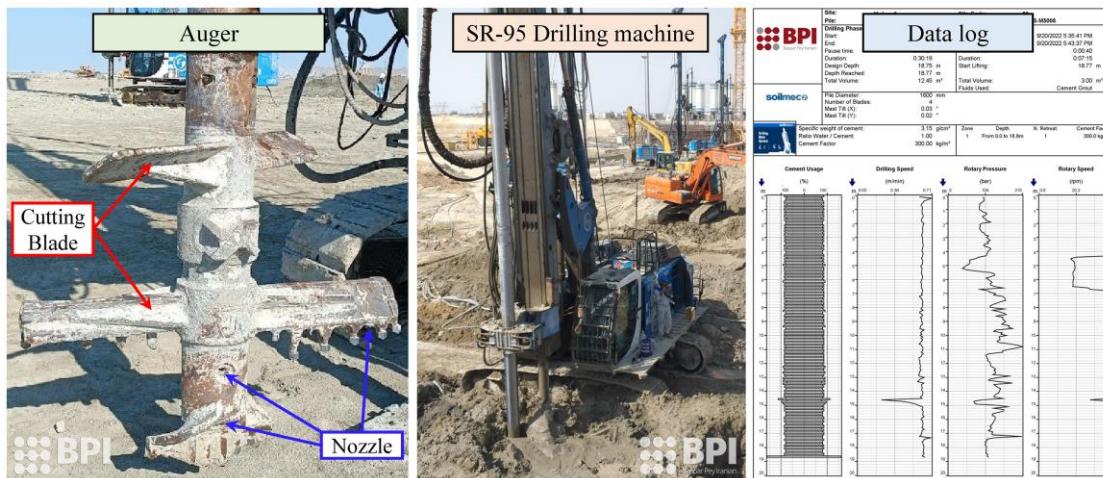


Figure 1. SR-95 drilling rig device, auger, and data log

Table 1. Descriptions of cosidered parameters regarding quality assessment of DSM elements

Abbreviation	Attribute	Unit	Count	Mean	Standard deviation	Range				
						Min	25%	50%	75%	Max
DSW	Drilling speed at water phase	m/min	3781	0.58	0.17	0.06	0.47	0.57	0.67	2.97
LSW	Lifting speed at water phase	m/min	3781	2.25	2.98	0.13	0.87	0.92	1.87	18.36
DSG	Drilling speed at grout phase	m/min	3781	0.56	0.13	0.34	0.55	0.57	0.58	5.82
LSG	Lifting speed at grout phase	m/min	3781	1.28	0.14	0.63	1.25	1.30	1.34	6.25
RSDW	Rotary speed at water drilling phase	rpm	3781	38.19	2.11	2.64	36.93	37.91	39.05	46.12
RSLW	Rotary speed at water lifting phase	rpm	3781	39.23	4.57	5.81	38.85	40.33	41.23	48.95
RSGD	Rotary speed at grout drilling phase	rpm	3781	40.72	2.29	16.51	39.97	40.73	41.42	48.19
RSGL	Rotary speed at grout lifting phase	rpm	3781	41.03	2.3	19.41	40.21	40.81	41.53	48.54
RPWD	Rotary pressure at water drilling phase	bar	3781	187.22	18.62	78.58	177.12	188.73	199.74	237.12
RPWL	Rotary pressure at water lifting phase	bar	3781	99.71	22.21	19.72	85.06	97.46	109.67	262.51
RPGD	Rotary pressure at grout drilling phase	bar	3781	85.19	15.41	30.23	75.34	83.24	92.42	220.99
RPGL	Rotary pressure at grout lifting phase	bar	3781	91.24	15.19	35.14	81.82	88.06	97.34	215.56
DSP	Spoil density	g/cm ³	3781	1.53	0.18	1.12	1.38	1.55	1.7	1.95
WD	Added water at drilling phase	ton	3781	0.68	0.20	0.05	0.54	0.65	0.78	2.14
WL	Added water at lifting phase	ton	3781	0.28	0.16	0.00	0.16	0.39	0.41	1.09
UCS*	Unconfined compressive strength	MPa	3781	2.84	1.31	0.39	1.91	2.56	3.5	7.78

* Output parameter. Others are input parameters.

3.1. Extreme Gradient Boosting (XGBoost)

The XGBoost is an scalable extension of gradient boosting-based algorithms that combines multiple decision trees to make precise predictions of output parameters (Han et al. 2022). This algorithm constructs its ensemble of decision trees using a greedy optimization algorithm (Chen and Guestrin 2016). At each iteration, the algorithm selects the tree that minimizes a predefined loss function, such as the mean squared error (Friedman 2001). The XGBoost algorithm considers a set of weak learners (i.e., decision trees that perform better than random guess) at the initial stage (Nguyen et al. 2021). Next, the gradients of the loss function with respect to the model predictions are calculated to train the next tree. Therefore, XGBoost develops a new a decision tree that minimizes the sum of weighted gradients of individual training examples. Finally, the model updates itself with gathering the predictions of the new and previous trees to enhance predictions in over and over iterations.

3.2. Random Forest (RF)

The RF is an ensemble ML algorithm that considers multiple decision trees and combines their predictions to get an overall result (Cakiroglu et al. 2022; Feng et al. 2021). It employs bootstrap aggregating technique for selecting of training data sets. In this method, multiple subsets of the training data are created by randomly sampling the original database with replacement. Hence, each subset might contain duplicates of the same data points, while other data points may not even be included in any training subset. Furthermore, the RF selects a random subset of input parameters (i.e., features) from the full set of features for each node of the every decision tree.

3.3. Artificial Neural Network (ANN)

The ANN is a machine learning mechanism that employs artificial intelligence to establish a nonlinear mapping between inputs and outputs. Its application in civil engineering has been documented in literature, such as studies by Gordan et al. (2019) and Pramanik and Babu (2022). If adequate databases are available, ANN can develop quick and practical estimation models (Dal et al. 2019). The ANN network consists of three layers with corresponding neurons, including the input layer (contains the independent variables), hidden layer (contains unobservable neurons which are functions of the input parameters), and output layer (contains the dependent variables which are functions of the hidden neurons). Besides, activation functions (i.e., mathematical functions) applies to the output of neurons to consider non-linearity into the model in hidden and output layers.

3.4. Lasso Regression (LR)

Lasso Regression (also known as Least Absolute Shrinkage and Selection Operator Regression) is a linear regression algorithm that penalizes the sum of the absolute values of the model coefficients. This penalty tends to shrink the coefficients towards zero, effectively reducing the complexity of the model. As a result, LR can be used to address issues like multicollinearity and overfitting, making it a valuable tool for data analysis tasks.

4. RESULTS AND DISCUSSION

To achieve a comprehensive understanding of the mutual relationship between input and output parameters in a dataset, it is imperative to analyze the correlations of features. This method is an essential step in the data analysis process and provides a reliable way to extract valuable information about the relationships between the all features. In the current study, the Pearson's approach is chosen to calculate the linear correlation coefficient (R). In this regard, the strength and direction of a relationship between two variables (i.e., features) can be measured by a number between -1 and 1. The formula for computing this coefficient is as follows:

$$R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (1)$$

where the x and y are individual feature points and n represents the sample size. Figure 2 shows the correlation matrix for all investigated parameters. The results well indicated the correlation between various features. For instance, the data on drilling speeds with water (DSW) is well-matched to the water consumed (WD) ($R = -0.82$), which is acceptable given the constant flow rate of the drilling. Hence, the correlation matrix reveals that WL, DSP, and RPGL are the most effective parameters on the UCS.

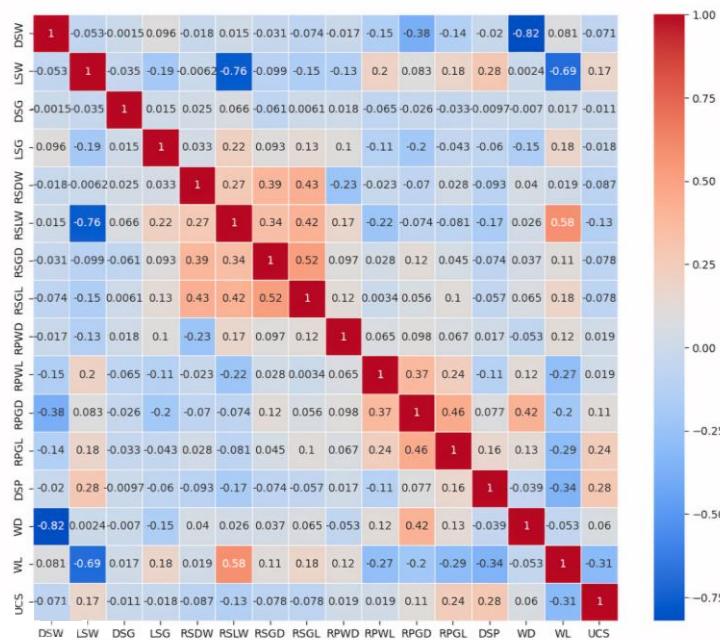


Figure 2. Correlation matrix of investigated features (abbreviations are denoted in Table 1)

The hyperparameters of ML models are configurations that are set before the model training process begins. Table 2 presents the considered hyperparameters for ML models, in which:

- The n_estimators indicates the number of generated gradient boosted trees in the model.
- The Learning_rate controls the learning speed, impacting both training time and generalization ability.

- The maximum depth that each tree can reach during any boosting round is controlled using the max_depth parameter.
- The L2 regularization is a technique used to prevent overfitting by adding a penalty term to the loss function. The reg_lambda controls the strength of L2 regularization.
- The min_samples_split idicates the minimum number of samples required to split an internal node.
- Number of neurons in ANN algorithm, strikes a balance between model complexity and learning capacity.
- The lambda value in the LR algorithm balances model complexity and overfitting.

Table 2. Charectristics and performance indicators of machine learning models

Abbreviation	Attribute	Hyperparameter							Performance indicator													
		n_estimators	learning_rate	max_depth	reg_lambda	min_samples_split	number of neurons	Training function	Lambda value	RMSE	MAE	MAPE (%)	R ²	RMSE	MAE	MAPE (%)	R ²	RMSE	MAE	MAPE (%)	R ²	
XGBoost	Extreme Gradient Boosting	5000	0.1	6	0.1	-	-	-	0.70	0.54	23.91	0.75	1.02	0.79	33.90	0.43	0.77	0.59	25.67	0.68		
RF	Random Forest	5000	-	5	-	10	-	-	0.36	0.28	11.93	0.95	0.94	0.71	29.05	0.45	0.53	0.36	15.36	0.86		
ANN	Artificial Neural Network	-	0.01	-	-	-	10	trainlm	-	0.99	0.76	32.16	0.43	1.09	0.83	34.32	0.35	1.02	0.78	32.81	0.41	
LR	Lasso Regression	-	-	-	-	-	-	-	0.005	1.12	0.87	38.06	0.29	1.07	0.80	34.12	0.29	1.11	0.85	37.27	0.29	

Four performance indicators were considered to evaluate the performance of the investigated ML algorithms, including the root mean square error (*RMSE*), mean absolute error (*MAE*), mean absolute percentage error (*MAPE*), and coefficient of determination (*R*²):

$$RMSE = \sqrt{\frac{1}{n} \sum_{m=1}^n (y_m - y_p)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{m=1}^n |y_m - y_p| \quad (3)$$

$$MAPE = \left[\frac{1}{n} \sum_{m=1}^n \left| \frac{y_m - y_p}{y_m} \right| \right] \times 100 \quad (4)$$

$$R^2 = 1 - \frac{\sum_{m=1}^n (y_m - y_p)^2}{\sum_{m=1}^n (y_m - \bar{y})^2} \quad (5)$$

where the y_m stands for measured data, y_p represents the predicted value, and y_{mm} is the mean of the measured data.

There are various challenges regarding the implementation of considered ML algorithms. To achieve the optimum performance of each machine learning model in the present study, hyperparameters were carefully fine-tuned through a trial-and-error approach. Despite considering 15 drilling input parameters, there are still more unconsidered effective parameters in every project, which can affect the UCS results. Therefore, several iterations for each ML method must be conducted to obtain the algorithm parameters with the lowest prediction error. In the tuning process of hyperparameters:

- Increasing the `n_estimators` generally improves model performance but comes at the cost of higher computational complexity and time.
- Considering a deeper tree in `max_depth` parameter can capture complex patterns but may lead to overfitting.
- The low values of `reg_lambda` (< 1) mean that the model is free to learn the training data as complexly as it wants, which includes the risk overfitting.
- A higher value of `min_samples_split` parameter discourages the formation of overly complex trees, thereby reducing overfitting.

Table 2 shows the error metrics corresponding to each ML algorithm. The RF and XGBoost algorithms had the most precise predictions for all train, test, and total database with R^2 values (for total database) more than 0.86 and 0.68, respectively. Figure 3 presents the results of ML predictions for UCS values. Three graphs are presented for each algorithm (i.e., Figures 3a-3c for XGBoost, Figures 3d-3f for RF, Figures 3g-3i for ANN, and Figures 3j-3l for LR), including the results of considering training dataset, test dataset, and total database. As is mentioned in Table 2, the RF algorithm well predicted UCS values, considering other ML models (Figure 3d-3f). Considering the R coefficient of 0.97 for training dataset in Figure 3d demonstrates superior outcomes of this algorithm for database considered in the current study.

The machine output data for a specific depth of a DSM element can be used to accurately predict the value of the output parameter (UCS) using any of the described algorithms. Theoretically, the most accurate algorithm is an algorithm where the difference between the predicted result and the actual value is zero. In reality, a tolerance can be granted to algorithms. By accounting for a margin of error, the effectiveness of an algorithm can be determined by analyzing the probability of accurate predictions within that margin. In this study, absolute prediction error which is defined as the absolute difference between the measured and the predicted values was used to compare the different algorithms.

Figure 4 depict the probability of correctly predicting the UCS values, as well as various absolute prediction errors for different algorithms. It should be noted that both the training and testing data were taken into account while generating these results. Figure 4a shows that the XGBoost algorithm can predict the strength of the samples with an error of ± 0.5 MPa with a probability of 51.67%. According to Figure 4b-4d, the probability of correct prediction with an error of 0.5 MPa for the other three methods, i.e., RF, ANN, and LR, are 82.7%, 42.11%, and 37.78%, respectively. Hence, The RF algorithm outperforms all the other ML methods in terms of prediction accuracy. Following that, XGBoost and ANN methods were ranked respectively, with the LR method having the weakest prediction. This is a clear results that the RF method is the most suitable for accurate predictions in this database. Besides, it is evident that there is a clear and distinct similarity in the trend observed while taking into account the error interval of ± 1 MPa, based on results reported in Figure 4.

It is crucial to determine a specific strength in soil improvement projects that use the deep mixing method. This determination is contingent upon the project specifications, soil properties, and other relevant factors. The specified unconfined 28- and 56-day UCSs for DSM projects in the United States are in the range of about 0.7 to 2.1 MPa (Bruce et al. 2013). According to the FHWA (Bruce et al. 2013), the strength acceptance criteria for treated soil is defined as the strength of 80% of the cored samples of each column being higher than the specified value. Furthermore, the strength of 90% of all samples should be higher than this specified value.

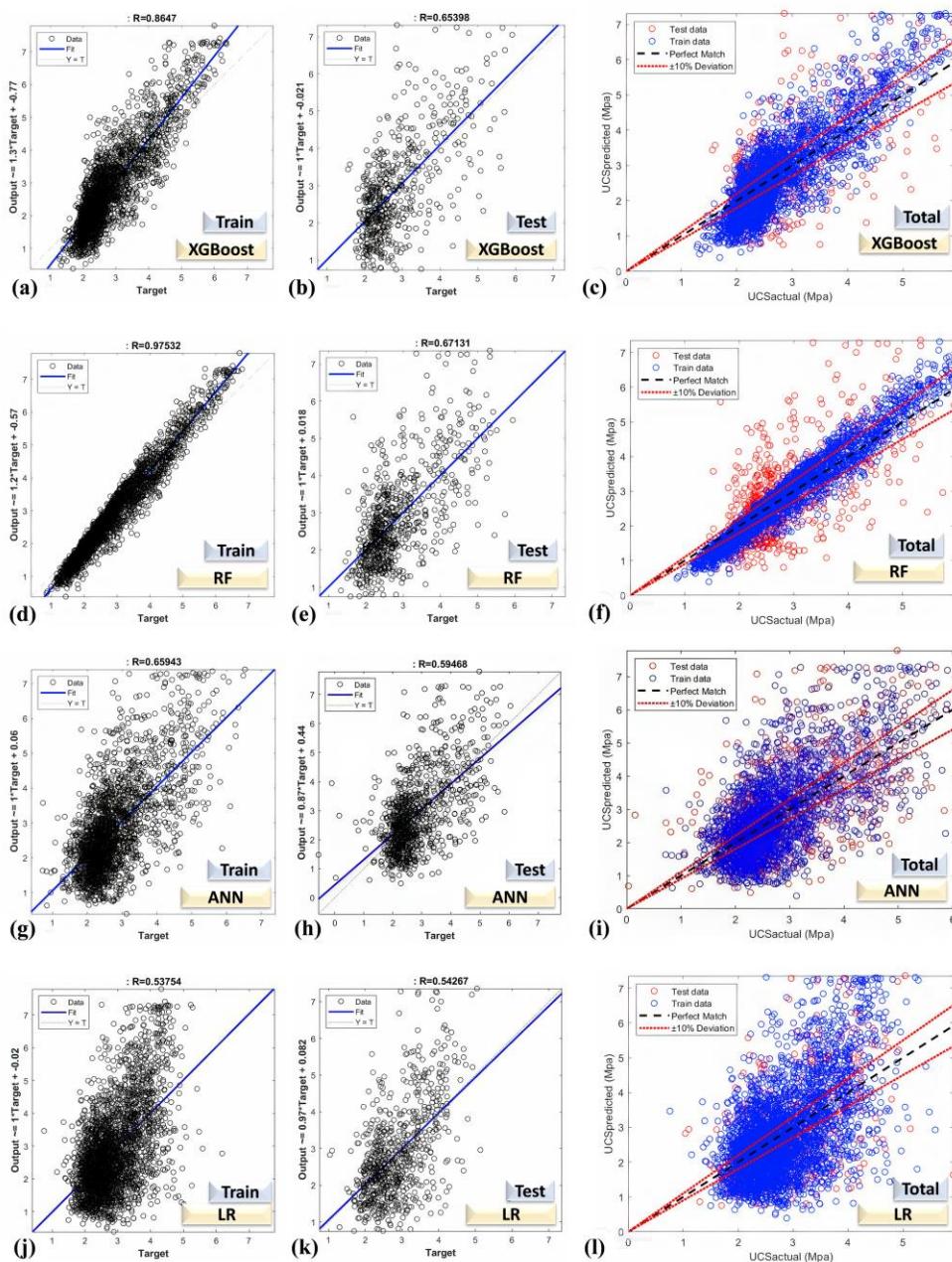


Figure 3. Results of UCS predictions by machine learning algorithms based on train, test, and total databases: a-c) XHBoost algorithm, d-f) RF algorithm, g-i) ANN algorithm, and j-l) LR algorithm

In the improvement project investigated in this study, if the specified strength is considered equal to 1.5 MPa, according to Figure 4c, with a probability of about 83%, the prediction of the trained RF model has an error of ± 0.5 MPa, and with considering a probability of about 97 %, the RF prediction has an error of ± 1 MPa. Besides, the RF prediction has a 100% accuracy, considering the error of about ± 1.5 MPa. Therefore, if at a specific depth for a DSM element, a strength of about 2 MPa is predicted using this well-trained RF algorithm, it is likely that the core in the UCS test should meet the strength criteria of 1.5 MPa with a confidence level of about 83%. Similarly, if the model prediction is approximately 2.5 MPa, the confidence level of acceptance of this element is about 97%. Moreover, if the prediction of UCS using this RF model exceeds 3 MPa, with 100% confidence level, the DSM element will meet the strength criteria. Implementing this approach in the field soil improvement projects, results in predicting the accept/reject status of an untested UCS element with various levels of confidence.

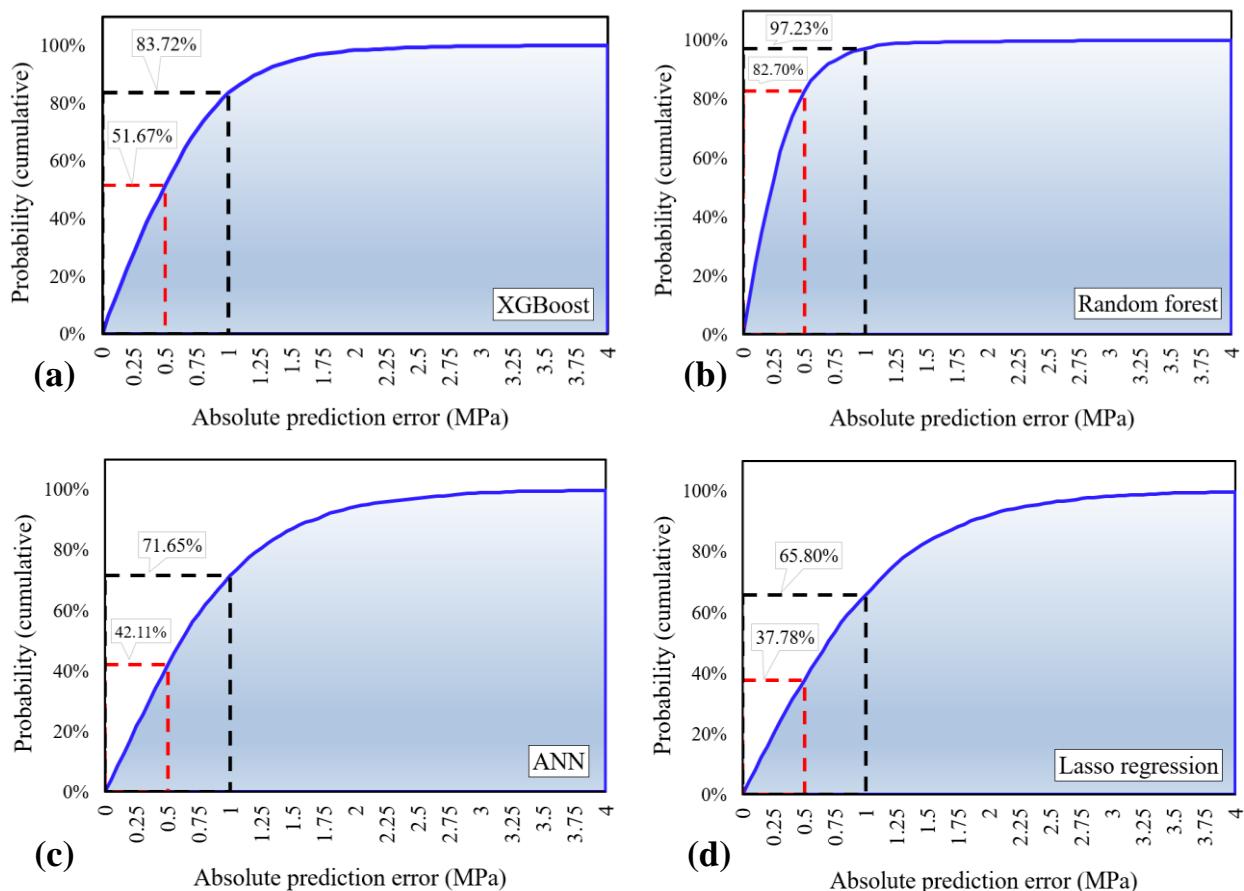


Figure 4. Cumulative absolute prediction error of UCS values by machine learning methods, including a) XGBoost, b) RF, c) ANN, and d) LR

5. CONCLUSION

Prediction of the quality of deep soil mixing column without conducting coring plan is a key that could bring about remarkable benefits. To address the correlation between the recorded data of rig and the strength of elements, comprehensive analyses have been carried out using

statsticial approaches and the ML algorithms, including XGBoost, RF, ANN, and LR. The main outcomes can be summarized as follows:

- 1- While QC/QA procedures are necessary in a soil improvement projects, but due to limited number of tests and cost considerations, indirect quality control methods are viable.
- 2- Based on the accuracy of the predictions, the RF algorithm provides the most fit prediction on the available data.
- 3- The accuracy of prediction is highly sensitive to the algorithms that the user uses, so it is recommended to different methods be considered to fit the most accuare method.
- 4- Using the statistical approach, it is achievable to predict the strength of DSM elements with an acceptable error and desired confidence level without conducting any UCS test, which is highly important in the projects.

6. ACKNOWLEDGEMENTS

This study was carried out in collaboration with the Baspar Pey Iranian (BPI) company. All of the data presented in the current study are provided by BPI company and the authors sincerely appreciate this company for all the support. The authors would also like to thank Mr. Amirmohammad Rabbani, an MS in earthquake engineering at the University of Tehran, for his constructive comments.

REFERENCES

- Anitescu, C., Atroshchenko, E., Alajlan, N., and Rabczuk, T. 2019. "Artificial Neural Network Methods for the Solution of Second Order Boundary Value Problems." 59 (1).
- Bruce, M. E. C., Berg, R. R., Filz, G. M., Terashi, M., Yang, D. S., Collin, J. G., and Geotechnica, S. 2013. *Federal Highway Administration design manual: Deep mixing for embankment and foundation support*. United States. Federal Highway Administation. Offices of Research & Development.
- Cakiroglu, C., Islam, K., Bekdaş, G., Kim, S., and Geem, Z. W. 2022. "Interpretable machine learning algorithms to predict the axial capacity of FRP-reinforced concrete columns." 15 (8), 2742.
- Chen, J.-J., Zhang, L., Zhang, J.-F., Zhu, Y.-F., and Wang, J.-H. 2013. "Field tests, modification, and application of deep soil mixing method in soft clay." 139 (1), 24-34.
- Chen, T., and Guestrin, C. 2016. "Xgboost: A scalable tree boosting system." In *Proc., Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 785-794.
- Dal, K., Cansiz, O., Ornek, M., and Turedi, Y. 2019. "Prediction of footing settlements with geogrid reinforcement and eccentricity." *Geosynth. Int.*, 26 (3), 297-308.
- Dang, B.-L., Nguyen-Xuan, H., and Wahab, M. A. 2023. "An effective approach for VARANS-VOF modelling interactions of wave and perforated breakwater using gradient boosting decision tree algorithm." *Ocean Eng.*, 268, 113398.
- Feng, D.-C., Wang, W.-J., Mangalathu, S., Hu, G., and Wu, T. 2021. "Implementing ensemble learning methods to predict the shear strength of RC deep beams with/without web reinforcements." 235, 111979.

- Friedman, J. H. 2001. *Greedy function approximation: a gradient boosting machine*. 1189-1232.
- Gordan, B., Koopalipoor, M., Clementking, A., Tootoonchi, H., and Tonnizam Mohamad, E. 2019. "Estimating and optimizing safety factors of retaining wall through neural network and bee colony techniques." *Eng. Comput.*, 35, 945-954.
- Han, J. 2015. *Principles and practice of ground improvement*. John Wiley & Sons.
- Han, J., Pei, J., and Tong, H. 2022. *Data mining: concepts and techniques*. Morgan kaufmann.
- Helson, O., Beaucour, A.-L., Eslami, J., Noumowe, A., and Gotteland, P. 2017. "Physical and mechanical properties of soilcrete mixtures: Soil clay content and formulation parameters." 131, 775-783.
- Jamsawang, P., Bergado, D., Bandari, A., and Voottipruek, P. 2008. "Investigation and simulation of behavior of stiffened deep cement mixing (SDCM) piles." 2 (3), 229-246.
- Jordan, M. I., and Mitchell, T. M. 2015. "Machine learning: Trends, perspectives, and prospects." 349 (6245), 255-260.
- Jung, C., Ceglarek, R., Clauvelin, T., Ayeldeen, M., and Kim, D. 2020. "Deep soil mixing in Sabkha soils for foundation support in United Arab Emirates." 6 (1), 3.
- Kitazume, M. 2022. *Quality Control and Assurance of the Deep Mixing Method*. CRC Press.
- Nghia-Nguyen, T., Kikumoto, M., Nguyen-Xuan, H., Khatir, S., Wahab, M. A., and Cuong-Le, T. 2023. "Optimization of artificial neural networks architecture for predicting compression parameters using piezocone penetration test." *Expert Syst. Appl.*, 223, 119832.
- Nguyen, D. H., and Wahab, M. A. 2023. "Damage detection in slab structures based on two-dimensional curvature mode shape method and Faster R-CNN." *Adv. Eng. Softw.*, 176, 103371.
- Nguyen, H., Vu, T., Vo, T. P., and Thai, H.-T. 2021. "Efficient machine learning models for prediction of concrete strengths." 266, 120950.
- Nicholson, P. G. 2014. *Soil improvement and ground modification methods*. Butterworth-Heinemann.
- Pramanik, R., and Babu, G. S. 2022. "Prediction of the Maximum Tensile Load in Reinforcement Layers of a MSE Wall Using ANN-Based Response Surface Method and Probabilistic Assessment of Internal Stability of the Wall." *Int. J. Geomech.*, 22 (8), 05022004.
- Rashid, A. S. A., Kueh, A. B. H., and Mohamad, H. 2018. "Behaviour of soft soil improved by floating soil-cement columns." 18 (2), 95-116.
- Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V. M., Guo, H., Hamdia, K., Zhuang, X., and Rabczuk, T. 2020. "An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications." *Comput. Methods Appl. Mech. Eng.*, 362, 112790.
- Tran, V.-T., Nguyen, T.-K., Nguyen-Xuan, H., and Wahab, M. A. 2023. "Vibration and buckling optimization of functionally graded porous microplates using BCMO-ANN algorithm." *Thin-Walled Struct.*, 182, 110267.
- Wang, S., Wang, H., Zhou, Y., Liu, J., Dai, P., Du, X., and Wahab, M. A. 2021. "Automatic laser profile recognition and fast tracking for structured light measurement using deep learning and template matching." *Measurement*, 169, 108362.
- Yao, K., Yao, Z., Song, X., Zhang, X., Hu, J., and Pan, X. 2016. "Settlement evaluation of soft ground reinforced by deep mixed columns." 9 (6), 460-465.

- Yi, Y., Liu, S., Puppala, A. J., and Jing, F. 2019. "Variable-diameter deep mixing column for multi-layered soft ground improvement: Laboratory modeling and field application." 59 (3), 633-643.
- Yin, J.-H., and Fang, Z. 2010. "Physical modeling of a footing on soft soil ground with deep cement mixed soil columns under vertical loading." 28 (2), 173-188.

Fast Machine Learning-Based High Fidelity Mesoscopic Modeling Tool for Traffic Simulation

Neeta A. Eapen, Ph.D.¹; Robert B. Heckendorf, Ph.D.²; and Ahmed Abdel-Rahim, Ph.D.³

¹Dept. of Computer Science, Univ. of Idaho, Moscow. Email: eape2084@vandals.uidaho.edu

²Dept. of Computer Science, Univ. of Idaho, Moscow. Email: heckendo@uidaho.edu

³Dept. of Civil Engineering, Univ. of Idaho, Moscow. Email: ahmed@uidaho.edu

ABSTRACT

Microscopic traffic simulators, such as VISSIM, use time-based step-by-step vehicle movements to predict outcome of proposed network configuration. This process is inherently slow and impractical for use in stochastic large network optimizations. Our solution is to use a mesoscopic traffic simulator with one-step movement approach between nodes in traffic network. Our model uses a predictor function for each road segment which predicts travel time distribution for traffic conditions using machine learning (AI). Our simulator chooses a random sample from predicted distribution as travel time. Travel time distributions may be sensitively dependent on parameters such as traffic and road conditions, and traffic behavior patterns which may be dependent on specific road. Experimental results show that our simulator's fidelity is like that of VISSIM for various traffic conditions. We demonstrate that our simulator is more than 100 times faster than VISSIM and provides network performance results that are comparable to these models.

Keywords: mesoscopic traffic simulator, VISSIM, machine learning, travel time

1. INTRODUCTION

Optimizing network traffic operations are complex process without reliable and situation specific closed-form formulas or models for optimization. This means simulation plays a critical role in any advancement in traffic network optimization. Microscopic models rely on simulating cars, pedestrians, and bicyclists behavior every time step, which could be as short as 100 milliseconds, leading to extremely long computing times that are impractical for any in-depth network-wide optimization. Furthermore, the connection, calibration, and validation with real traffic behavior is tenuous if the car behavior model is based on a few simple hard-coded abstractions that apply everywhere. This affects the reliability of predicting traffic flow across the diversity of street segments in a city. On the other hand, mesoscopic simulation is much faster in that it only requires the ability to estimate arrival times at the ends of each street segment link. Unfortunately, these estimations are difficult to make because they must be based on, not only engineering parameters, but also road conditions including congestion, pavement conditions, and subjective influencers such as roadway width, driver familiarity, and weather conditions, among other factors.

With the advent of specialized hardware and high-performance bus and memory architectures, machine learning has shown near-magical improvement in the last 10 years. Evidence of this is plainly visible on the cell phones we use every day such as speech recognition and classification of a picture library by subject. Machine learning has been applied to various

traffic problems, including methods of predicting congestion (Ata et al. 2019; Ide et al. 2015; Liu and Wu 2017) using techniques such as random forests and neural networks. Other examples are intersection traffic prediction (Alajali et al. 2018), traffic density prediction using LSTMs (Abbas et al. 2018), traffic forecasting with Bayesian networks (Zhou et al. 2020), and accident prediction by distinguishing normal from abnormal traffic behaviors (Dogru and Subasi 2012). We will apply this technology to predict the behavior of traffic on individual road segments allowing us to quickly assess the arrival times.

1.1. Our Approach

We propose to produce a fast and reliable mesoscopic traffic simulation model with a key component being a predictor function for each road segment that predicts the travel time distribution for each car across that segment. Let $d_s(c, \vec{\theta})$ be a predictor function for a segment s where c is the congestion on the segment and $\vec{\theta}$ is a vector of auxiliary parameters such as road conditions and weather conditions. Let p_s denote the travel time distribution predicted by the predictor function $d_s(c, \vec{\theta})$, on road segment s . During the running of the simulation, random samples from the distribution p_s are used.

To get these predictor functions, we propose learning the functions using machine learning techniques based on monitoring real traffic behavior. We aim to test out our ideas by using machine learning to learn predictor functions for the streets in various model cities run with the popular VISSIM micro simulator. That is, VISSIM will stand-in for real world traffic observations. To measure our success, we will compare the timings seen in VISSIM with those in our mesoscopic simulator. Furthermore, because the mesoscopic simulator does not have to simulate every small step along a road, it will run much faster. Perturbation of VISSIM traffic data controlled by auxiliary parameters will allow us to test more difficult traffic simulation conditions.

An advantage of our approach is the development of functions that can be street and time specific. This could include behaviors that standard simulators might find difficult to handle like: optically narrow streets, 5 o'clock traffic that is driving facing a setting sun, intersections with complex characteristics based on awkward sized turn lanes, etc. This is possible by monitoring actual traffic behavior and using machine learning methods to build a predictor for that case. The overall result would be simulations that run at greatly increased speed and fidelity.

2. BACKGROUND

Traffic simulation is a fundamental tool for understanding and predicting the behavior of vehicular and pedestrian traffic. There are three main types of traffic simulators based on the level of detail and abstraction: microscopic, mesoscopic, and macroscopic. Microscopic traffic simulators model individual vehicles, pedestrians, and bicycles as discrete objects and simulate their interactions with each other and the environment. These models often use car-following and lane-changing rules to mimic the behavior of real-world drivers. Some of the most popular microscopic traffic simulators include SUMO (Simulation of Urban MObility) (Krajzewicz et al. 2012; Lopez et al. 2018), VISSIM (PTV Group 2020), and AIMSUN (Yunex Traffic Group 2023).

Macroscopic traffic simulators model traffic flow as a continuous fluid-like system and use partial differential equations to describe the evolution of traffic density, speed, and other

macroscopic variables. These models are the least computationally intensive, but provide the most abstract representation of traffic behavior. Examples of macroscopic traffic simulators include LWR (Lighthill-Whitham-Richards) model (Lighthill and Whitham 1955) and CTM (Cell Transmission Model) (Daganzo 1994).

Mesoscopic traffic simulators model traffic at an intermediate level of detail. They typically predict travel time and traffic patterns between points in the road network by considering traffic characteristics such as traffic density, flow rates, and capacity of road segments. These models are less computationally intensive than microscopic models but provide more detail than macroscopic models. Examples of mesoscopic traffic simulators include DYNASMART (DYnamic Network Assignment-Simulation Model for Advanced Road Telematics) (Mahmassani 1992; Mahmassani and Abdelghany 2002) and Dynasmart-P (Mahmassani et al. 2001).

Overall, the choice of traffic simulator depends on the specific research question and the level of detail required. Microscopic simulators provide the highest level of detail, mesoscopic simulators provide a balance between detail and computational efficiency, and macroscopic simulators provide the most abstract representation of traffic flow.

Of the available traffic models, we choose a mesoscopic traffic simulator. Our simulator predicts the travel time for each car along its given path by predicting the distribution of travel time for each road segment on the path, using machine learning.

3. METHOD

In this section, we describe our mesoscopic traffic simulator. Our simulator predicts the travel time of each car, for each road segment, during the simulation. Our simulator contains a predictor function for each road segment. We begin by demonstrating that the travel time of a vehicle is a distribution of values rather than a constant. We use machine learning to create a predictor function for each road segment.

For our proof of concept, we use VISSIM road traffic simulator as real world (PTV Group 2020). VISSIM is a widely used microscopic traffic simulator which is accurate, methodical, and adequately reproduces the traffic patterns in the real world. Collecting real world data to learn from, was difficult for this work. So, we use VISSIM as a reasonable stand-in for real world. Since VISSIM is powerful enough to duplicate real world traffic behavior, we are using data collected from VISSIM to learn the traffic patterns and behavior in our simulator. We then show our high fidelity mesoscopic traffic simulator reproduces the travel time distributions observed in VISSIM, which stands in for the real world. We demonstrated high fidelity of our simulator (shown in Experimental Results section) on the following tests:

- Testing on different topologies of networks of road segments and lanes
- Testing on different congestions/vehicle input rates
- Testing on different speed limits
- Testing on road segments which are uphills or downhills

3.1. Travel time is a distribution rather than a constant

For a particular set of traffic conditions, existing travel time prediction algorithms (Rice and Van Zwet 2004; Lin et al. 2019; Qiu and Fan 2021) usually predict a single value. In this section, we show that travel time for the same set of traffic conditions on a road segment can result in a

range of values. The driving patterns of different people are different, even for the same parameters. Therefore, the travel time is not the same even for the same traffic conditions on a road segment. Therefore, the travel time should be considered as a distribution, rather than a constant.

To verify that VISSIM adheres to this view, we collected the traffic data for a single road segment using the VISSIM traffic simulation tool (PTV Group 2020). We kept all the traffic parameters as constant, with only congestion on the road changing. In Figure 1, the x-axis is the number of cars on road (congestion), and the y-axis is the travel time of vehicles in seconds. We can see that for even a particular value of ‘Number of Cars on Road’, the travel time is a range of values rather than a constant. Therefore, we need a machine learning model which predicts a distribution of travel time. We propose a mixture density neural network model (Bishop 1994) for predicting the distribution of travel time of vehicles.

3.2. Our Distribution Prediction Model For A Road Segment

A mixture density neural network (MDN) (Bishop 1994) is a type of neural network that is designed to model complex probability distributions. It is used in situations where a simple parametric distribution, such as a Gaussian distribution, cannot accurately model the data. In an MDN, the neural network outputs parameters that define a mixture of Gaussian distributions. Each Gaussian component in the mixture is associated with a weight that determines its contribution to the overall distribution. The network is trained to optimize the parameters of the mixture model to maximize the likelihood of the observed data.

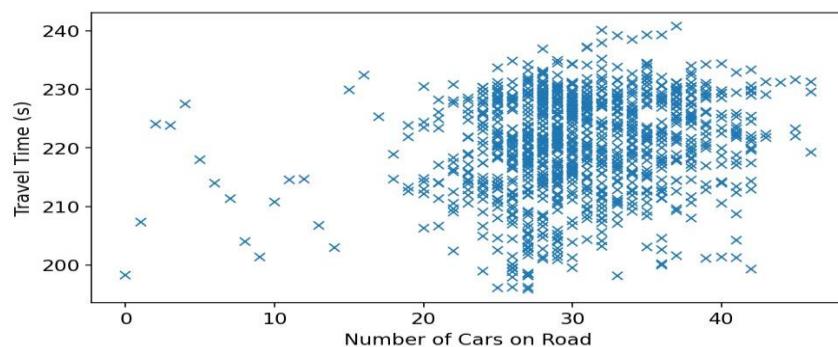


Figure 1: Travel time VS number of cars on road: travel time is a range of values rather than a constant.

We use a mixture density neural network for predicting travel time distributions in our model. Our model is a combination of a neural network and a Gaussian mixture model as shown in Figure 2. Any distribution can be considered as a combination of multiple Gaussian distributions. The parameters of a Gaussian distribution are mean and standard deviation. MDN fits the mean and standard deviation of multiple Gaussian distributions. It also finds the mixing coefficients for each Gaussian distribution, which is used for understanding how each Gaussian distribution is combined to obtain the travel time distribution of a vehicle.

Figure 2 shows the architecture of the proposed model. The neural network used in our work is composed of an input layer with 9 neurons, two hidden layers with 30 neurons each, and an output layer with 30 neurons. The input features of the neural network are: congestion on road

segment and auxiliary parameters such as length of road segment, gradient of road segment, speed limit of road segment, vehicle type, average length of vehicles, vehicle input rate of the road network, maximum number of cars on one lane, and number of lanes on road segment. The output layer consists of parameters for 10 Gaussian distributions used in a Gaussian mixture model: mixing coefficients (α_i) which is the probability of selecting from Gaussian distribution i , means(μ_i), and standard deviations(σ_i) for $i = 1$ to 10. Note: $\sum_{i=1}^{10} \alpha_i = 1$. The 10 Gaussian distributions p_1, p_2, \dots, p_{10} are combined to obtain the distribution p as shown in Equation 2. The number of Gaussian distributions is chosen as 10 since this value gave the best travel time distribution prediction in our experiments. Note that x denotes the input features given to the model. Our model chooses a random sample t from the distribution $p(x)$ as the travel time.

$$p(x) = \sum_{i=1}^{10} \alpha_i * p_i(x | \mu_i, \sigma_i) \quad (1)$$

Travel time is selected from $p(x)$ by:

$$t = \text{rand}(p_i(x | \mu_i, \sigma_i)) \quad (2)$$

where $i =$ one of $\{1, 2, \dots, 10\}$ and rand function chooses a random sample.

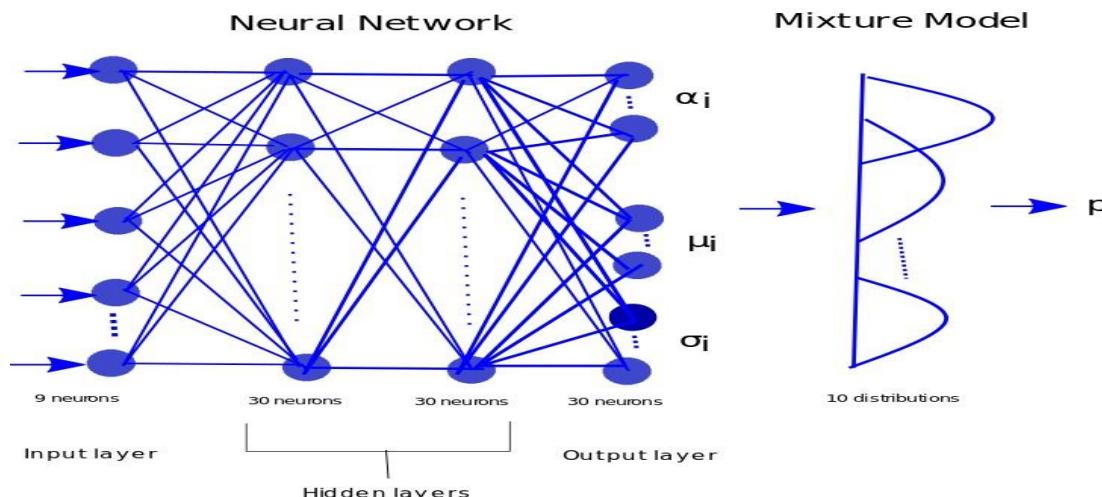


Figure 2: Model Architecture: A mixture density neural network with nine neurons in the input layer, 30 neurons in each of the two hidden layers, and 30 neurons in the output layer. The 10 Gaussian distributions in output layer are combined to obtain the distribution p .

3.3. Mesoscopic simulator

Algorithm 1 shows the steps performed by our mesoscopic traffic simulator. The input to the simulator is the road network, predictors for each road segment, vehicle input rate of road network R (number of vehicles/hour), and simulation time T (seconds). The output is the total travel time of each of the vehicles. We assume for this proof of concept that the road network has

no traffic signals. First, the algorithm calculates the number of vehicles to simulate n . Note that $\lfloor \cdot \rfloor$ denotes the floor function. Then, it finds the inter-start times S for the vehicles from an exponential distribution.

Algorithm 1 Mesoscopic Traffic Simulation

Require: Road network, Predictor functions for each road segment, Vehicle input rate of road network R (number of vehicles/hour), Simulation time T (seconds). **Ensure:** Travel time of vehicles (seconds).

- 1: Calculate number of vehicles to simulate $n = \lfloor (R / (60 \times 60)) \times T \rfloor$.
 - 2: Initialize a priority queue π for storing the vehicles in the simulation. Priority queue is sorted based on time to start simulating a vehicle.
 - 3: Calculate mean of inter-start times for vehicles $m = T/n$
 - 4: Draw n samples from an exponential distribution with rate parameter = m . Let S be the set of samples drawn.
 - 5: Calculate the start time for each vehicle using the inter-start times in S and insert the vehicles into π .
 - 6: **while** π is not empty **do**
 - 7: Pop a vehicle from the priority queue π . Let $t_{current}$ be the start time of the vehicle.
 - 8: Let s be the road segment through which the vehicle needs to travel next.
 - 9: Let d_s be the predictor for the road segment s .
 - 10: Choose a sample travel time t from the travel time distribution predicted by d_s .
 - 11: $t_{new} = t_{current} + t$
 - 12: **if** the vehicle has not reached its destination and $t_{new} < T$ **then**
 - 13: Insert the vehicle into π with updated start time t_{new}
 - 14: **end if**
 - 15: **end while**
-

An exponential distribution is used since it can estimate the vehicle start times, given the average time between vehicle start times. It assumes that vehicle start times are independent and random, and that the rate at which vehicles enter a road network is constant over time. Moreover, the exponential distribution has a memoryless property, which means that the time elapsed since the start time of the last vehicle does not affect the probability of the start time of the next vehicle. This property makes it a suitable choice for modeling interstart times (that is the time between vehicle start times) in vehicle simulation, where the probability of a vehicle start time at any given time depends only on the average time between vehicle start times, and not on how much time has elapsed since the last vehicle started.

Using the set of inter-start time values in S , the simulation start time for each vehicle is calculated, and the vehicle is inserted into a priority queue π . The priority queue π is used for storing the vehicles in simulation and it is sorted based on the start time of next action of each vehicle. Algorithm 1 shows a while loop which runs until π is empty. In the while loop, each vehicle is popped and its travel time along the next road segment (the road segment it needs to travel next, say s) is calculated by using a predictor d_s . The predictor d_s predicts the travel time distribution. A sample of the distribution is chosen as the travel time t of the vehicle, through the road segment s . Implementation note: The sample t is repeatedly chosen between t_{min} and t_{max} where t_{min} and t_{max} are the minimum and maximum values of travel time observed on this road segment during training. This is to ensure that practical values are chosen as travel time since

Gaussians have infinitely long tails. If the vehicle has not yet reached its destination, the vehicle is added back to the priority queue with the updated start time $t_{current} + t$ where $t_{current}$ is the start time of the vehicle on road segment s .

4. EXPERIMENTAL RESULTS

The experiments are written as a prototype using Python 3.8.5 programming language. We collected data from VISSIM 6 (PTV Group 2020) to get the data comparable to real-life scenarios.

4.1. Statistical Test

The distribution of total travel time of vehicles in our simulator is tested against the distribution of total travel time of vehicles in VISSIM, for various road networks. Our simulator reproduces the travel time distributions observed in VISSIM with high fidelity. For our tests, the null hypothesis is that, between all points A and B, the travel time distribution of our simulator is the same as the travel time distribution of VISSIM. We used Mann-Whitney U Test (Mann and Whitney 1947) to support our hypothesis, since this test can compare whether the two distributions are similar or different from each other. We obtain a p-value, for total travel time of each pair of source and destination points, using the Mann-Whitney U Test. If p-value is less than 0.05, we reject the null hypothesis that the two distributions are the same. Else, we fail to reject the null hypothesis that the two distributions are the same. We denote ‘Reject’ to reject the null hypothesis and ‘Accept’ to fail to reject the null hypothesis.

4.2. Road networks Tested

We collected data from VISSIM for the road networks R_1 , R_2 , and R_3 as shown in Figures 3, 4, and 5. Figure 3 shows a road network (R_1) with a single road segment/edge and two nodes. The route of vehicles has node 1 as origin and node 2 as destination, with a single path 1-2. Since there is only a single path, this path is selected with 100% probability. Figure 4 shows a road network (R_2) with two road segments/edges and three nodes. The route of vehicles has node 1 as origin and node 3 as destination, with a single path 1-2-3. Since there is only a single path, this path is selected with 100% probability. The road network in Figure 5 is a road network (R_3) with multiple road segments and has 6 nodes and 7 edges. The route of vehicles has node 1 as origin and node 5 as destination. Three paths exist between nodes 1 and 5. They are 1-2-3-4-5, 1-2-3-6-4-5, and 1-2-6-4-5. The path 1-2-3-4-5 is selected with 40% probability, the path 1-2-3-6-4-5 is selected with 20% probability, and the path 1-2-6-4-5 is selected with 40% probability. Both our simulator and VISSIM used the same probabilities for routes. We tested our simulator against VISSIM for different parameters on these road networks. This is shown in the following sections. These road networks were chosen to test the fidelity of our simulator for road networks with single, two, and multiple segments. We believe these road networks provide a good basic test. 80% of the collected VISSIM data was used for training and 20% of the data was used for testing.

Our simulator reproduces the travel time distribution of each pair of points with high fidelity. For example, Figure 6 shows the total travel time distributions of our simulator and VISSIM, for road segment 6-4 in the road network R_3 , in one of our experiments. The histogram of travel time

distribution is shown with the x-axis as the travel time and the y-axis as the frequency/count of the travel time. VISSIM travel time distribution (observed travel time) is shown in blue and the travel time distribution of our simulator (predicted travel time) is shown in orange. The shaded (gray) portion is the shared/overlapping portion of the travel time distributions of our simulator and VISSIM. We can observe that the travel time distribution of our simulator is similar to that of VISSIM. We obtained the p-value for the Mann-Whitney U Test as 0.94 which statistically shows the fidelity of our simulator.



Figure 3: Road network with a single road segment (R_1)



Figure 4: Road network with two road segments (R_2)

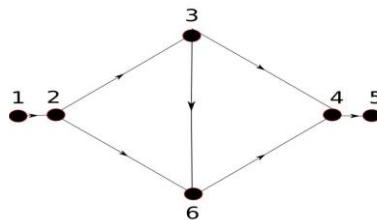


Figure 5: Road network with multiple road segments (R_3)

An example for distribution comparison of total travel time of our simulator and VISSIM between two non-adjacent points (points 2 and 4 in the road network R_3) is shown in Figure 7. We can observe that the travel time distribution of our simulator is similar to that of VISSIM. We obtained the p-value for the Mann-Whitney U Test as 0.73 which statistically shows the fidelity of our simulator.

4.3. Testing on different number of road segments and lanes

We tested our simulator against VISSIM for road networks R_1 , R_2 , and R_3 which have 1, 2, and 7 segments/edges. We set vehicle input rate to 100 vehicles/hr. We tested for number of lanes of each segment = 1, 2, and 3. The distribution of travel time in our simulator is similar to that of VISSIM for different number of road segments and lanes. This can be observed in Table 1. In all the cases, we got ‘Accept’ which indicates that our simulator works the same way as VISSIM for single, two, and multiple road segments and for different number of lanes.

4.4. Testing on different congestions/vehicle input rates

Driving behavior of an individual varies with congestion on the road. The congestion during the daytime is different from early morning and night. Congestion might be higher in cases where people are leaving from any major events, or if they are going to or from work. At other times, say late night or noon, congestion might be lower. We tested our simulator against

VISSIM for different congestions (vehicle input rates). Testing was performed on the road network R_3 (Figure. 5) with 3 lanes. The distribution of travel times in our simulator is tested against the distribution of travel times in VISSIM using the Mann-Whitney U Test (Mann and Whitney 1947). We show the p-values and Reject/Accept values in the Table 2 for different scenarios. We compared the distributions of total travel time for the vehicles in our simulator and VISSIM. We can see that in all the cases, we got ‘Accept’ indicating that our simulator can work the same way as VISSIM for different congestions.

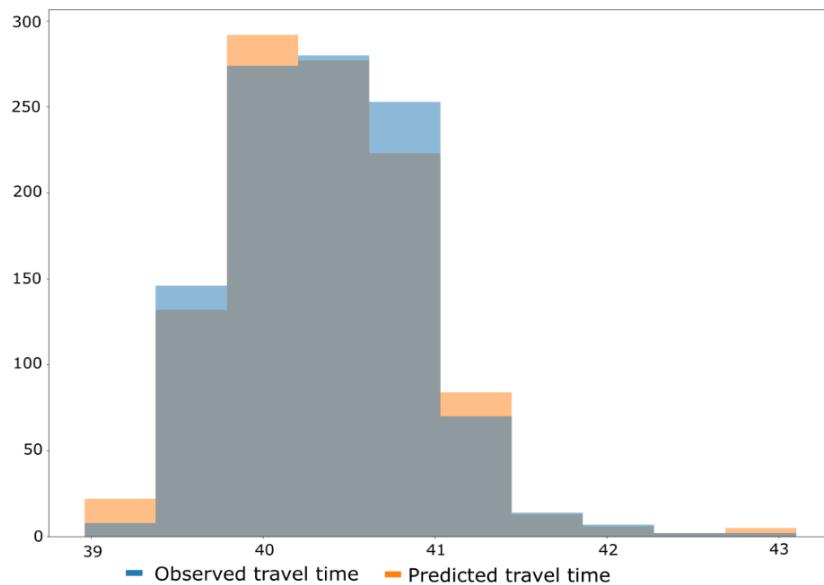


Figure 6: Comparison of distribution of travel time in our simulator (predicted travel time: shown in blue) with VISSIM (observed travel time: shown in orange) for road segment 6-4 in the road network R_3 . The shaded (gray) portion is the shared/overlapping portion of the travel time distributions of our simulator and VISSIM. We can observe that the travel time distribution of our simulator is similar to that of VISSIM. We obtained the p-value for the Mann-Whitney U Test as 0.94 which statistically shows the fidelity of our simulator.

Table 1: Testing our simulator against VISSIM for different number of road segments and lanes.

Road network	Num. of segments	Num. of Lanes	P-value	Reject/Accept
R_1	1	1	0.95	Accept
R_1	1	2	0.93	Accept
R_1	1	3	0.97	Accept
R_2	2	1	0.75	Accept
R_2	2	2	0.89	Accept
R_2	2	3	0.89	Accept
R_3	7	1	0.77	Accept
R_3	7	2	0.73	Accept
R_3	7	3	0.73	Accept

4.5. Testing on different speed limits

Driving behavior of an individual varies with speed limits of roads. We set vehicle input rate to 100 vehicles/hr. We tested our simulator against VISSIM for different speed limits. The distribution of travel times in our simulator is tested against the distribution of travel times in VISSIM using the Mann-Whitney U Test (Mann and Whitney 1947). We show the p-values and Reject/Accept values in the Table 3 for different scenarios. Testing was done on the road network R_3 (Figure. 5) with single lane. We compared the distributions of total travel time for the vehicles in our simulator and VISSIM. We can see that in all the cases, we got ‘Accept’ indicating that our simulator can work the same way as VISSIM for different speed limits.

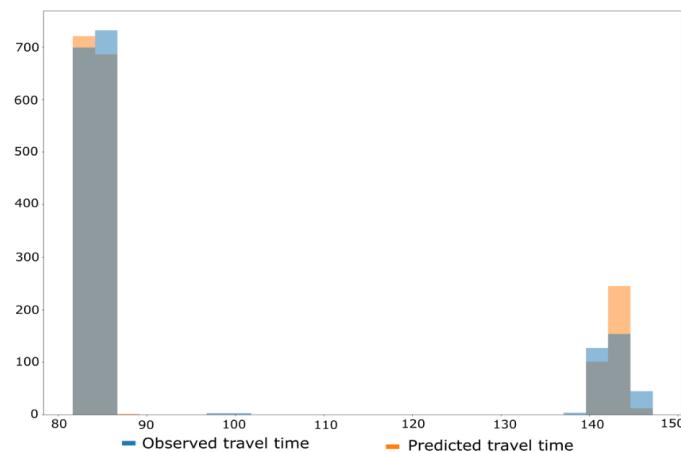


Figure 7: Comparison of distribution of total travel time in our simulator (predicted travel time: shown in blue) with VISSIM (observed travel time: shown in orange) between points 2 and 4 in the road network R_3 . Note that we picked two non-adjacent points. The shaded (gray) portion is the shared/overlapping portion of the travel time distributions of our simulator and VISSIM. We can observe that the travel time distribution of our simulator is similar to that of VISSIM. We obtained the p-value for the Mann-Whitney U Test as 0.73 which statistically shows the fidelity of our simulator. Note that the distribution (shown in this figure) is bimodal. This is because, of the three paths between points 2 and 4 in the road network R_3 (2-3-4, 2-6-4, and 2-3-6-4), the path 2-3-6-4 takes a longer travel time than the other two paths since it has a longer distance.

Table 2: Testing our simulator against VISSIM for different congestions.

Congestion (Vehicles/hr)	P-value	Reject/Accept
10	0.93	Accept
50	0.59	Accept
100	0.73	Accept
500	0.87	Accept
1500	0.68	Accept
1800	0.52	Accept
2100	0.61	Accept

4.6. Testing on road segments which are uphills or downhills

Driving behavior of an individual on an uphill is different from the driving behavior on a downhill. We performed testing on uphill and downhill road segments. The uphill road segments have a positive slope/gradient and the downhill road segments have a negative slope/gradient. We set vehicle input rate to 100 vehicles/hr. We tested our simulator against VISSIM for different gradient/slope values as shown in Table 4. We believe these gradient values are good basic tests. 100%, -100%, 200%, and -200% are used for testing hypothetical large slope values. 10%, -10%, 20%, -20%, and 0% are used for testing practical slope values. The distribution of travel times in our simulator is tested against the distribution of travel times in VISSIM using the Mann-Whitney U Test (Mann and Whitney 1947).

We show the p-values and Reject/Accept values in the Table 4 for different scenarios. Testing was done on the road network R_3 (Figure. 5) with single lane. We compared the distributions of total travel time for the vehicles in our simulator and VISSIM. We can see that in all the cases, we got ‘Accept’ indicating that our simulator can work the same way as VISSIM for different gradient/slope values.

Table 3: Testing our simulator against VISSIM for different speed limits.

Speed limit (km/hr)	P-value	Reject/Accept
20	0.79	Accept
30	0.78	Accept
40	0.84	Accept
50	0.77	Accept
60	0.76	Accept
70	0.79	Accept
80	0.84	Accept
90	0.80	Accept
100	0.81	Accept

Table 4: Testing our simulator against VISSIM for different gradient/slope values.

Gradient/Slope values on segments	P-value	Reject/Accept
0% on all segments	0.73	Accept
10% on 2-3, 2-6; -10% on 3-4, 6-4; 0% on 1-2, 4-5, 3-6	0.94	Accept
20% on 2-3, 2-6; -20% on 3-4, 6-4; 0% on 1-2, 4-5, 3-6	0.73	Accept
100% on 2-3, 2-6; -100% on 3-4, 6-4; 0% on 1-2, 4-5, 3-6	0.89	Accept
100% on 2-3; -100% on 3-4; 200% on 2-6; -200% on 6-4; 0% on 1-2, 4-5, 3-6	0.78	Accept
10% on 2-3; -10% on 3-4; 20% on 2-6; -20% on 6-4; 0% on 1-2, 4-5, 3-6	0.86	Accept

4.7. Running Time Comparison

Figure 8 shows the box and whisker plot of running time (seconds) of simulations performed on our simulator and VISSIM. The plot shows running time of 72 simulations. The box and whisker plot shows that our simulator has a significantly lower running time than VISSIM. The median, minimum, and maximum running time in VISSIM is 492.74 s, 452.47 s, and 593.58 s. The median, minimum, and maximum running time in our simulator is 2.92 s, 2.69 s, and 5.73 s. The ratio of running time of simulations in VISSIM and our simulator has the range 100.21 to 214.52. In other words, our simulator runs 100.21 to 214.52 times faster than VISSIM, even though we used Python programming language. Python is around 100 times slower than C (Lion et al. 2022). Our simulator is still at least 100.21 times faster.

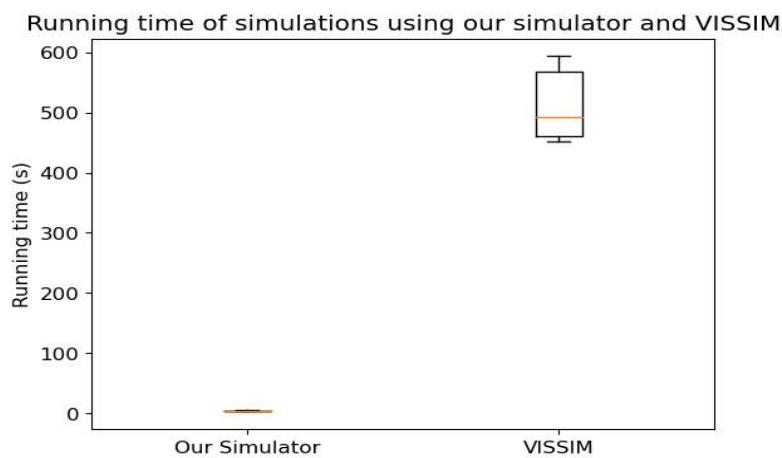


Figure 8: Box and Whisker plot of running time (s) of simulations using our simulator and VISSIM. The plot shows that our simulator has a significantly lower running time than VISSIM. Our simulator runs 100.21 to 214.52 times faster than VISSIM, even though we used Python programming language.

5. CONCLUSION

We proposed a mesoscopic simulation model which has a machine-learning based travel time predictor for each road segment. The predictor of a road segment predicts the travel time distribution for the particular traffic conditions. The mesoscopic simulator selects a random sample of the predicted distribution as travel time. We compared the fidelity of our simulator with VISSIM (PTV Group 2020) which stands in for real time traffic data. We demonstrated through our experiments that the travel time distributions predicted by our simulator is similar to that of VISSIM for various scenarios. We also conducted experiments to test the running time of simulations using our simulator and VISSIM, and found that our simulator runs at least 100 times faster than VISSIM, even though we used Python programming language. This makes our simulator much more appropriate for optimizations. Our work provides transportation researchers and professionals with a fast and reliable network modeling an optimization tool. But like all machine learning studies, in order for the simulator to work best, it needs to learn from a wide range of possible data sets representing a wide range of input conditions. Otherwise, the simulator will produce approximated results. Future work will include signals and extreme

values of congestion. The effect on fidelity of variations in the amount of data collected will be explored.

6. ACKNOWLEDGEMENTS

We thank Pacific Northwest Transportation Consortium (PacTrans) for funding this research. We also thank National Institute for Advanced Transportation Technology (NIATT) and University of Idaho for the support.

REFERENCES

- Abbas, Z., Al-Shishtawy, A., Girdzijauskas, S., and Vlassov, V. (2018). Short-term traffic prediction using long short-term memory neural networks. In *2018 IEEE International Congress on Big Data (Big Data Congress)*, pages 57–65. IEEE.
- Alajali, W., Zhou, W., Wen, S., and Wang, Y. (2018). Intersection traffic prediction using decision tree models. *Symmetry*, 10(9):386.
- Ata, A., Khan, M. A., Abbas, S., Ahmad, G., and Fatima, A. (2019). Modelling smart road traffic congestion control system using machine learning techniques. *Neural Network World*, 29(2):99–110.
- Bishop, C. (1994). Mixture density networks. Working paper, Aston University.
- Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation research part B: methodological*, 28(4):269–287.
- Dogru, N., and Subasi, A. (2012). Traffic accident detection by using machine learning methods. In *Third International Symposium on Sustainable Development (ISSD'12)*, page 467.
- Ide, C., Hadjii, F., Habel, L., Molina, A., Zaksek, T., Schreckenberg, M., Kersting, K., and Wietfeld, C. (2015). Lte connectivity and vehicular traffic prediction based on machine learning approaches. In *2015 IEEE 82nd Vehicular Technology Conference (VTC2015Fall)*, pages 1–5. IEEE.
- Krajzewicz, D., Erdmann, J., Behrisch, M., and Bieker, L. (2012). Recent development and applications of sumo-simulation of urban mobility. *International journal on advances in systems and measurements*, 5(3&4).
- Lighthill, M. J., and Whitham, G. B. (1955). On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345.
- Lin, D.-J., Tsao, W.-L., Yu, C.-H., Liu, H.-H., and Chang, Y.-C. (2019). The travel time prediction by machine learning methods with traffic data in chiayi city, taiwan. In *2019 4th International Conference on Electromechanical Control Technology and Transportation (ICECTT)*, pages 257–260. IEEE.
- Lion, D., Chiu, A., Stumm, M., and Yuan, D. (2022). Investigating managed language runtime performance: Why JavaScript and python are 8x and 29x slower than c++, yet java and go can be faster? In *2022 USENIX Annual Technical Conference (USENIX ATC 22)*, pages 835–852.
- Liu, Y., and Wu, H. (2017). Prediction of road traffic congestion based on random forest. In *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, volume 2, pages 361–364. IEEE.

- Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Luček, L., Rummel, J., Wagner, P., and Wießner, E. (2018). Microscopic traffic simulation using sumo. In *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE.
- Mahmassani, H. (1992). Dynamic traffic assignment and simulation for advanced network informatics (dynasmart). In *the 2nd International Seminar on Urban Traffic Networks*, 1992.
- Mahmassani, H., Abdelghany, A., Abdelghany, K., Chiu, Y.-C., Huynh, N., and Zhou, X. (2001). DYNASMART-P: Intelligent Transportation Network Planning Tool – DYNASMART-P Version 0.92 User's Guide. report number: ST067-85-P.
- Mahmassani, H. S., and Abdelghany, K. F. (2002). Dynasmart-ip: Dynamic traffic assignment meso-simulator for intermodal networks. In *Advanced modeling for transit operations and service planning*. Emerald Group Publishing Limited.
- Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60.
- PTV Group. (2020). PTV VISSIM. <https://www.ptvgroup.com/en/solutions/products/ptv-vissim/>.
- Qiu, B., and Fan, W. D. (2021). Machine learning based short-term travel time prediction: Numerical results and comparative analyses. *Sustainability*, 13(13):7454.
- Rice, J., and Van Zwet, E. (2004). A simple and effective method for predicting travel times on freeways. *IEEE Transactions on Intelligent Transportation Systems*, 5(3):200–207.
- Yunex Traffic Group. (2023). AIMSUN. <https://www.aimsun.com/>.
- Zhou, F., Yang, Q., Zhong, T., Chen, D., and Zhang, N. (2020). Variational graph neural networks for road traffic prediction in intelligent transportation systems. *IEEE Transactions on Industrial Informatics*, 17(4):2802–2812.

Improved Prediction of Soil Thermal Properties Using Recurrent Neural Networks

Scott Michael Slone¹; Zachary Zody²; Robert Ibey³; and Wade A. Lein, P.E.⁴

¹Cold Regions Research and Engineering Laboratory, Engineer Research and Development Center, US Army Corps of Engineers, Hanover, NH. Email: scott.m.slone@usace.army.mil

²Cold Regions Research and Engineering Laboratory, Engineer Research and Development Center, US Army Corps of Engineers, Hanover, NH

³Cold Regions Research and Engineering Laboratory, Engineer Research and Development Center, US Army Corps of Engineers, Hanover, NH

⁴Cold Regions Research and Engineering Laboratory, Engineer Research and Development Center, US Army Corps of Engineers, Hanover, NH

ABSTRACT

Frost effects, such as thaw weakening and frost heave, can significantly degrade pavements. Effective mitigation relies on accurate prediction of soil thermal properties. Current prediction methods use empirical equations or finite element analysis, with recent progress using machine learning. Recurrent neural networks, a type of machine learning model designed for prediction of sequential data, may be appropriate. Further, the model's output can be fed back into itself recursively to make long-term predictions. Using this method and training data from Hill Air Force Base, Utah, and Air Force Academy, Colorado, we were able to forecast soil parameters including temperature, thermal conductivity, and moisture content for frost susceptible soils. The model had a deviation from experimental values of no more than 10%, with the most significant contributions to accuracy being the use of gated recurrent unit neurons and the incorporation of multiple soil parameters.

INTRODUCTION

Frost effects, such as frost heave and thaw weakening, can significantly degrade pavements, and there have already been several decades of research into predicting or forecasting these effects, commonly finding difficulty due to the changing thermal properties of freezing soil (Pollard, 2017). A common solution has been to estimate or assume properties as negligible, such as in the saturated state, which has been applied both to analytical and numerical techniques (Bianchini & Gonzalez, 2012).

The Modified Berggren equation is a common analytical method of predicting frost penetration depth, a common frost effect, and is effectively a modified form of the Stefan equation, with the equation being:

$$P = \lambda \sqrt{\frac{48k * n * CFI}{L}}$$

where P is the frost depth penetration in feet, λ is a correction factor, k is the average thermal conductivity of frozen and unfrozen soil in BTU/(ft*h*°F), n is the air temperature to surface temperature ratio, CFI is cumulative freezing degree-days, and L is the volumetric latent heat of

fusion in BTU/ft³ (Bianchini & Gonzalez, 2012). Theory assumes that the correction factor corrects for assumptions in soil thermal properties, but in practice overestimates frost penetration depth. Often, site-specific empirical versions are used with custom correcting factors, but this limits their wide-spread utility (Rajaei & Baladi, 2015). Thankfully, this overestimation is well known, and most current models have focused on numerical methods, such as with finite element modeling.

Finite Element Modeling

Due to the complex nature of soil freezing, numerical modeling and simulation have shown to be more appropriate than previous methods. The most common is finite element modeling, where a complex system is approximated by a large number of small 2- or 3-dimensional elements, with interactions between elements represented by differential equations. This method has proven effective for a wide variety of phenomena, such as heat transfer, electromagnetics, and mechanics (Dayarathne, Hawlader, Phillips, & Robert, 2023), but has also been proven to be effective at modeling soil in unfrozen, frozen, and transition phases. However, these models often require calibration and site-specific information, along with large computational resources to simulate multi-season or region-sized models (Lein, Slone, Smith Jr., Bernier, & Oren, 2019).

Neural Network Models

An alternative to the previous methods that may have improved results while also requiring fewer computational resources are neural networks. Neural networks are a form of machine learning where complex operations are approximated using non-linear functions known as neurons, which connect to each other similarly to neurons in a brain. These models have shown to be appropriate for a variety of applications, with recent success in weather modeling using recurrent neural networks, which have identical inputs and outputs into their functions to allow for reading in sequential data (Alzubaidi, et al., 2021).

In 2016, Zaytar et al demonstrated that a recurrent neural network using long short-term memory neurons (LSTMs) was capable of effective forecasting of air temperature, humidity, and wind speed for 72 hours with a relatively short training period of 7 hours. The model's accuracy was attributed to the LSTMs, which allow the model to better estimate long-term trends by storing previous results in sequential data, making them highly appropriate for weather forecasting (Zaytar & Amrani, 2016). Recently, Gated Recurrent Units, which are similar to LSTMs but include a forget gate to remove irrelevant data, have been shown to have similar results to LSTMs, but have reportedly faster processing times (Barooni, Ziarati, & Barooni, 2023). Recently, neural networks models have been used to forecast soil properties, as the problem is largely similar to weather forecasting (Choi, Kim, Kim, & Won, 2022). Recent successes have produced mean absolute errors for prediction of soil temperature for a 24 hour forecast of 2 to 5 °F (Talsma, Solander, Mudunuru, Crawford, & Powell, 2023).

However, one limitation of these models has been their input data, with several parameters, such as temperature, thermal conductivity, and soil moisture content likely having some co-dependence, as in prediction of one property may depend on knowledge of the others. Incorporating these parameters into a joint hydraulic-thermal neural network may improve the accuracy of each parameter further, while also allowing for prediction of soil properties that are important to understanding frost effects but are not directly calculable from soil temperature. Based on the above literature results, there is evidence supporting use of both LSTMS and

GRUs. Due to there being fewer literature examples for frost depth prediction yet with notable claims of their benefits in other fields, it was recommended GRUs were likely the best model choice for further exploration.

RESULTS

Test Plan

Accurate characterization of soil properties is crucial for numerous applications, spanning geotechnical engineering, agriculture, and environmental science. A custom sensor platform, referred to as the frost probe and shown in Fig 1, was designed for efficient and precise in-situ measurements during freezing and thawing conditions. The frost probe system comprised a 5-foot-long tine array, with sensor tines distributed with a higher resolution in the upper soil as compared to lower regions. The upper section of the probe had 13 closely spaced tines in 1-inch intervals, while the lower section had 5 pairs of 1-inch separated tines, with the pairs separated by 6 inches. Soil temperature, thermal conductivity, and moisture content were recorded from two field sites: Hill Air Force Base in Utah from 2020 to 2022 and the Air Force Academy in Colorado from 2021 to 2022. It should be noted that both sites have similar soil types, being classified as high plasticity clay.

The probe system's tines integrated three distinct technologies to facilitate in-situ measurement of soil properties, including temperature and moisture content. High-fidelity temperature measurements and an embedded heating element allowed for low energy thermal conductivity measurements, which was inferred by controlled heating and measuring temperature decay. Time Domain Reflectometry (TDR) was utilized to determine volumetric moisture content, which would allow for accurate measurements in unfrozen soils but could be inaccurate in frozen conditions. Temperature and moisture content readings were recorded at 15-minute intervals, while thermal conductivity was assessed once daily due to a 1-hour extended duration after heating to measure decay.



Figure 1. Frost Depth Sensor Platform. Left, conceptual design. Right, installation at Hill AFB, 2020.

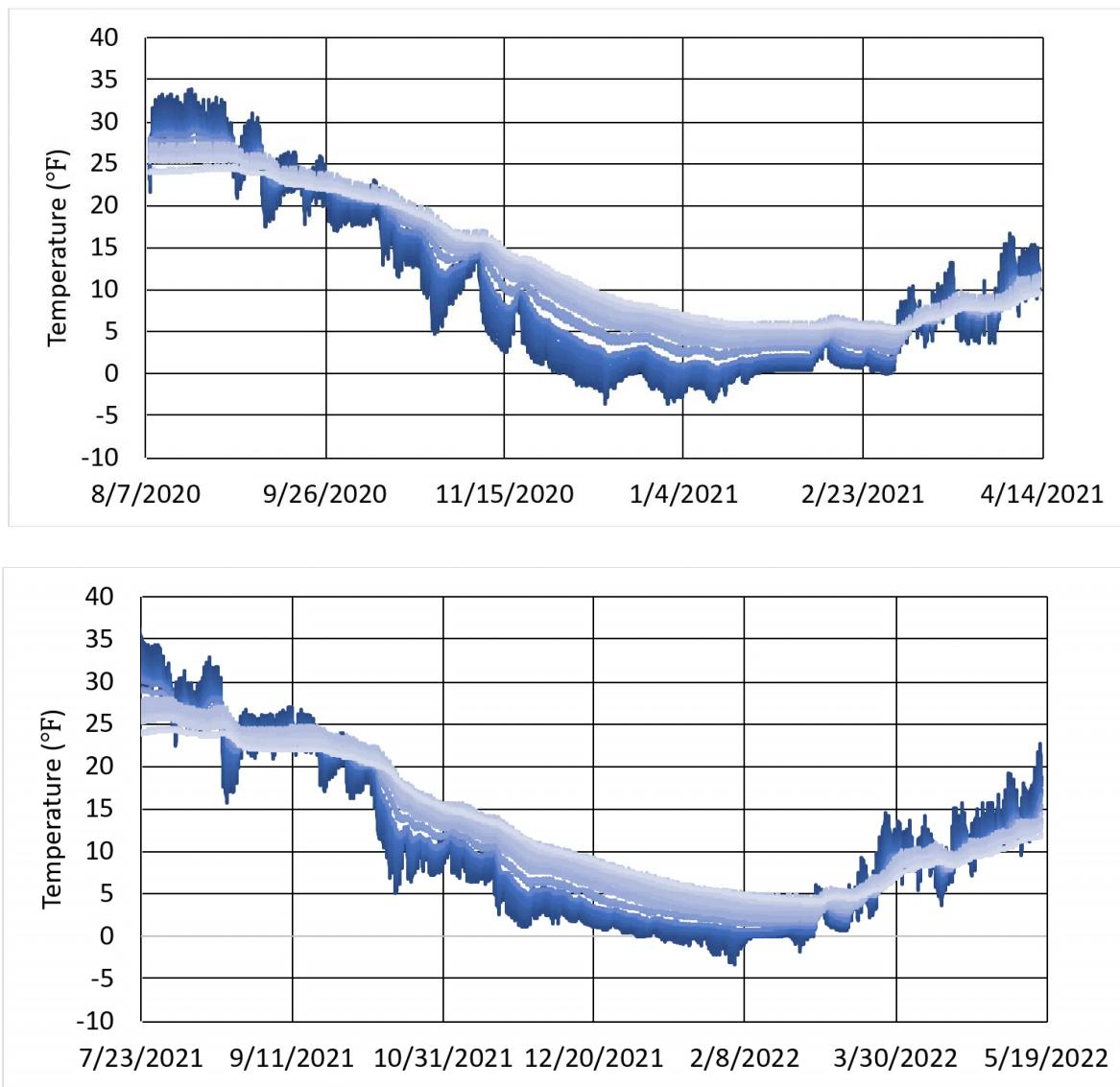


Figure 2. Soil temperature profile for Hill Air Force Base, Utah. Top, Winter 20/21. Bottom, Winter 21/22. Darkest blue trend is at 1 in depth, lightest blue is at 60 in depth.

Model Design

Based on literature, a recurrent neural network using GRUs was selected. Based on an initial evaluation of Hill AFB Winter 20/21 results, an architecture of 32 recurrent neurons and 32 GRUs was selected, using 48 hours of previous data with a 1 hour resolution as input and 1 hour as output, with the model developed in Python and Tensorflow. The model was trained for 120 epochs, with a batch size of 32, and with a staggered learning rate of 1e-3, 5e-4, 2.5e-5, and 1.25e-5 for every 30 epochs, respectively.

Due to there being multiple types of input data, four versions of the model were trained, and labeled based on if they included Temperature (T), Volumetric Moisture Content (M), or Thermal Conductivity (K). The trained models losses and validation losses are shown below in Fig 3.

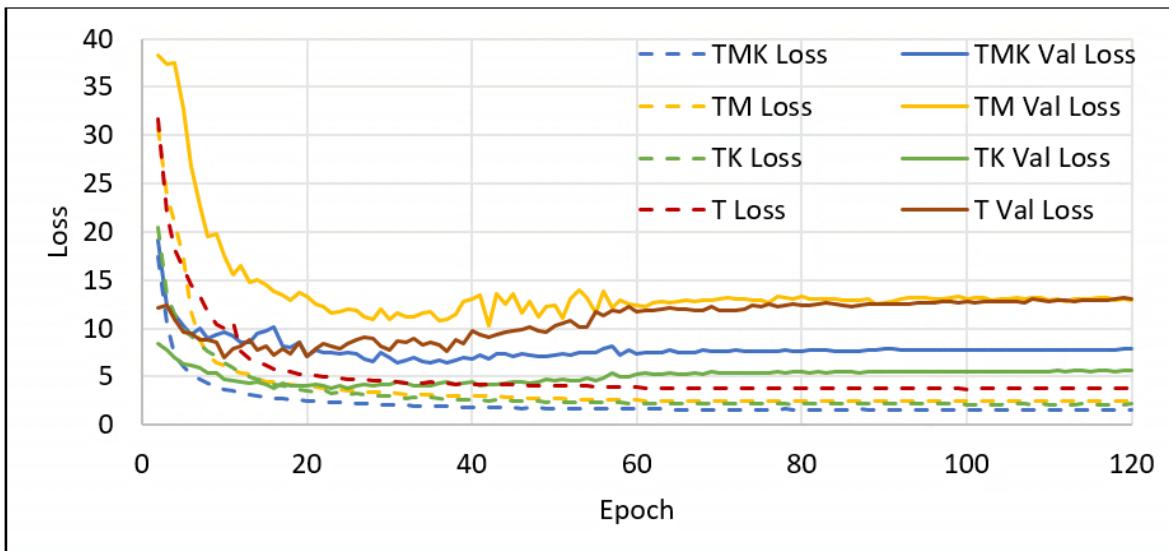


Figure 3. Training and Validation Losses for Frost Penetration Depth Recurrent Neural Network. Labels define which input parameters were used in training, with T, M, and K, meaning Temperature, Volumetric Moisture Content, and Thermal Conductivity, respectively.

Of the four models tested, Model T had the highest training loss and a similar validation loss as Model TM, having a final value of 13.06, implying it would be the worst predictor of the four in practice, with Model TM likely having similar results. The inclusion of thermal conductivity in Models TMK and TK lowered the validation losses to 7.95 and 5.60, respectively. Models that did not include moisture content tended to have lower losses than those that did, implying moisture content did not contribute greatly to accuracy. To verify which model is the most accurate, comparison to experimental results will be necessary.

DISCUSSION

When comparing to experimental results, all three outputs were analyzed individually. For brevity and ease of reading, only the properties at the shallowest testing depth (1 in.) will be shown, along with the mean absolute error across all results. Discussion of temperature will also include forecasting accuracy.

Temperature

All four models included a temperature output, which is a crucial metric to determine other soil properties, and thus benefits from a high accuracy. The measured temperature at the 1 in. depth and four model outputs for the Hill AFB 21/22 Winter Cycle, as well as the mean absolute error across all measured depths are shown in Fig. 4.

The models have largely similar results throughout the year, usually underestimating temperature in the summer and spring periods. The greatest accuracy was during the November 2021 to Early February 2022 period, when temperatures would be their lowest. Overall, Model TK, using only temperature and thermal conductivity, appeared to have the lowest mean absolute error.

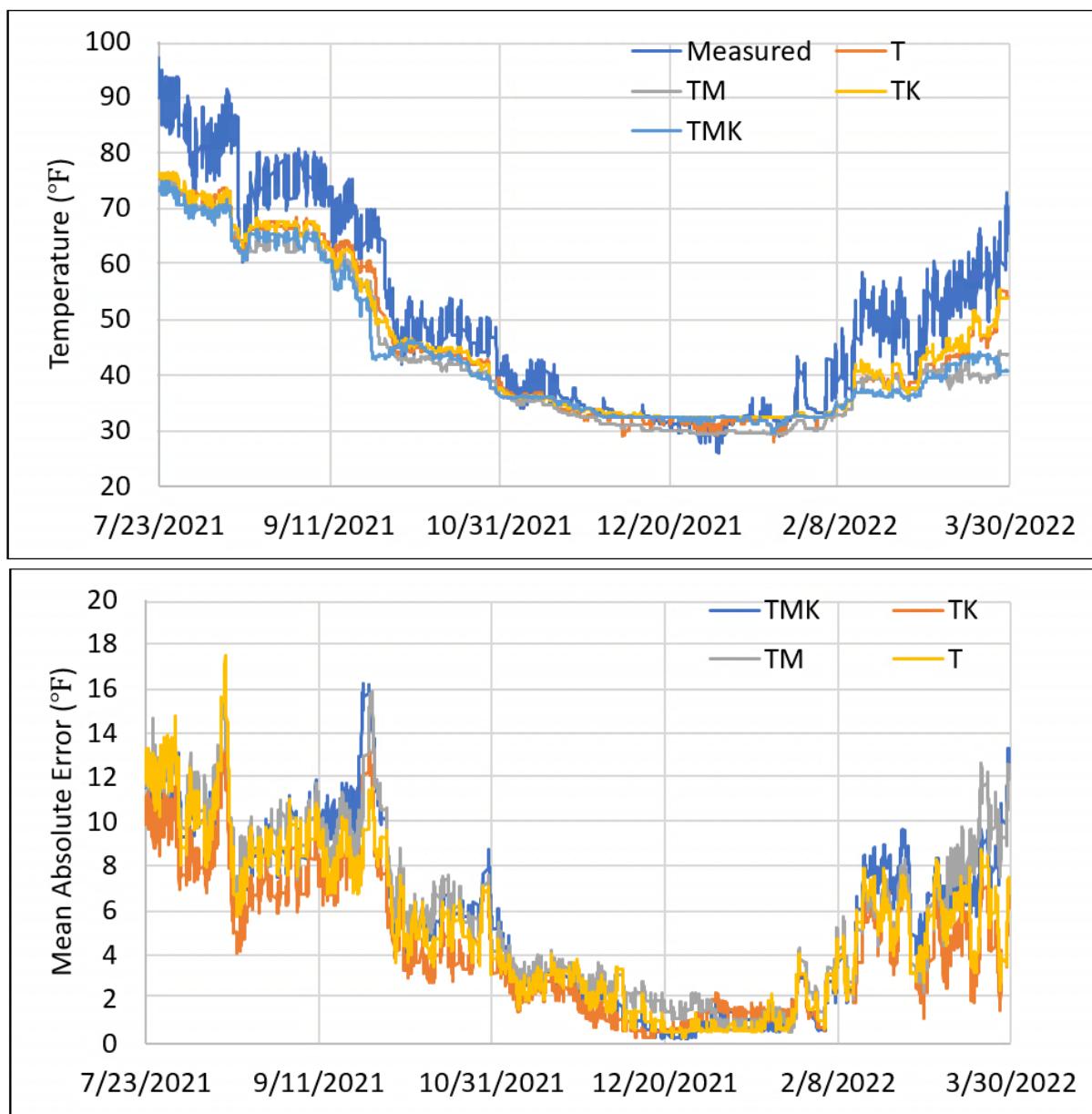


Figure 4. Temperature model results for Hill AFB, Winter 21/22. Top, Temperature for 1 in. depth. Bottom, Temperature Mean Absolute Error across all depths. Models are labeled if they used Temperature (T), Volumetric Moisture Content (M), and/or Thermal Conductivity (K) as input.

To further understand the models' utility, their capabilities for predicting temperature were also investigated. Rather than using historical data, a randomly selected 48 hour segment was selected from the Hill AFB 21/22 winter period and used as input for the four models, which then forecasted the next 120 hours of soil temperatures, with the accuracies of these shown below in Fig 5.

In Fig 5., we see the accuracies for the models are largely dependent on which parameters are selected for input. Using only temperature as an input was initially similar to using all

parameters, but its error steadily rises, surpassing Model TM at approximately 54 hours. At 120 hours, Model TM had an error of 7.69 °F, which is more than twice that for a traditional weather prediction model at is 3.45 °F (Novak, et al., 2014). Model TMK, which used all parameters, had a final error after 120 hours of 5.46 °F, a notable improvement over Models TM and T. Overall, Model TK was the optimal model with the lowest mean absolute error, with an error after 24 hours of 3.90 °F and after 120 hours of 4.76 °F. Further improvement to the architecture are recommended to further lower the prediction error. Due to the benefits of adding thermal conductivity, adding additional parameters or improving amount of training data using the above parameters is recommended.

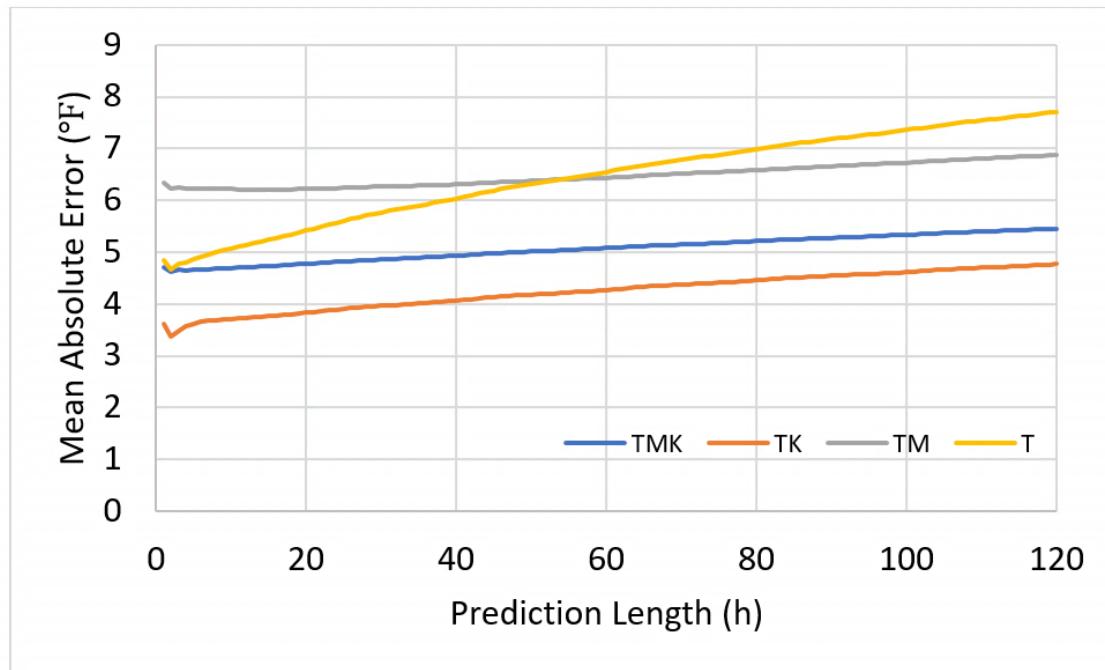


Figure 5. Mean Absolute Error for Temperature Forecasting. Prediction length is number of hours from the end of input data. Models are labeled if they used Temperature (T), Volumetric Moisture Content (M), and/or Thermal Conductivity (K) as input.

Volumetric Moisture Content

For Volumetric Moisture Content, as seen in Fig 6, there were notable discrepancies between the model results and measured values. The biggest difference being the inability of the models to recognize clear saturated soil conditions, such as seen from November 2021. While Model TMK had a lower mean absolute error than Model TM, neither would be reasonable for moisture content predictions.

Thermal Conductivity

For thermal conductivity, the dedicated Model TK performed better than the general Model TMK, generally following the measured result and having a lower mean absolute error over all.

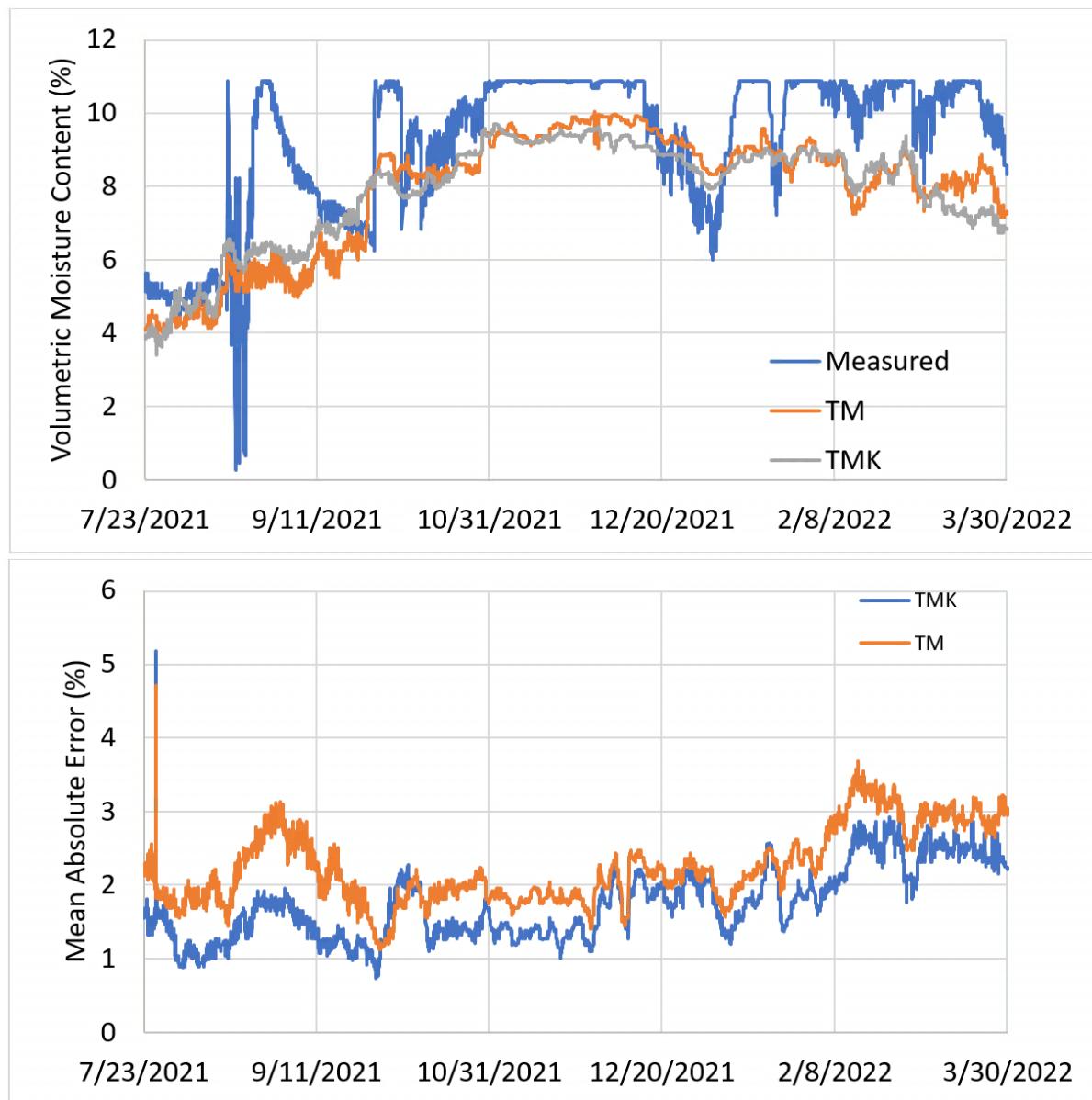


Figure 6. Volumetric Moisture Content model results for Hill AFB, Winter 21/22. Top, Volumetric Moisture Content for 1 in. depth. Bottom, Volumetric Moisture Content Mean Absolute Error across all depths. Models are labeled if they used Temperature (T), Volumetric Moisture Content (M), and/or Thermal Conductivity (K) as input.

Final Discussion

Based on the validation loss results and the results for temperature and thermal conductivity, the optimum model for prediction of the soil temperature profile was determined to be Model TK, with its full temperature profile prediction compared to measurement shown below in Fig 8. The model predicted a maximum freezing front of 14.4 in. compared to the experimental result of 16.7 in, while the Modified Berggren predicted 24 in. Model TK's accuracy was higher away from the freezing front, predicting a maximum depth for 35 °F at 17.0 in. compared to the

experimental result of 27.9 in., showing the model is currently more accurate for above-freezing temperature prediction. Model TMK performed similarly in many respects, and allowed for prediction of volumetric moisture content, but had a higher error for forecasting, which lowers its potential utility. When calculating moisture content, however, Model TMK was the superior option. For prediction and forecasting of soil temperatures, Model TK is able to do so with reasonable accuracy compared to existing weather models. However, there are limitations, shown in its errors for near freezing temperatures and due to being trained only on high plasticity clays. The model is not recommended for predicting in other soil types until it has been trained in a larger variety of soil conditions. The model is recommended for predicting in soil types under airfields, specifically when they are high plasticity, frost-susceptible soils in cold regions.

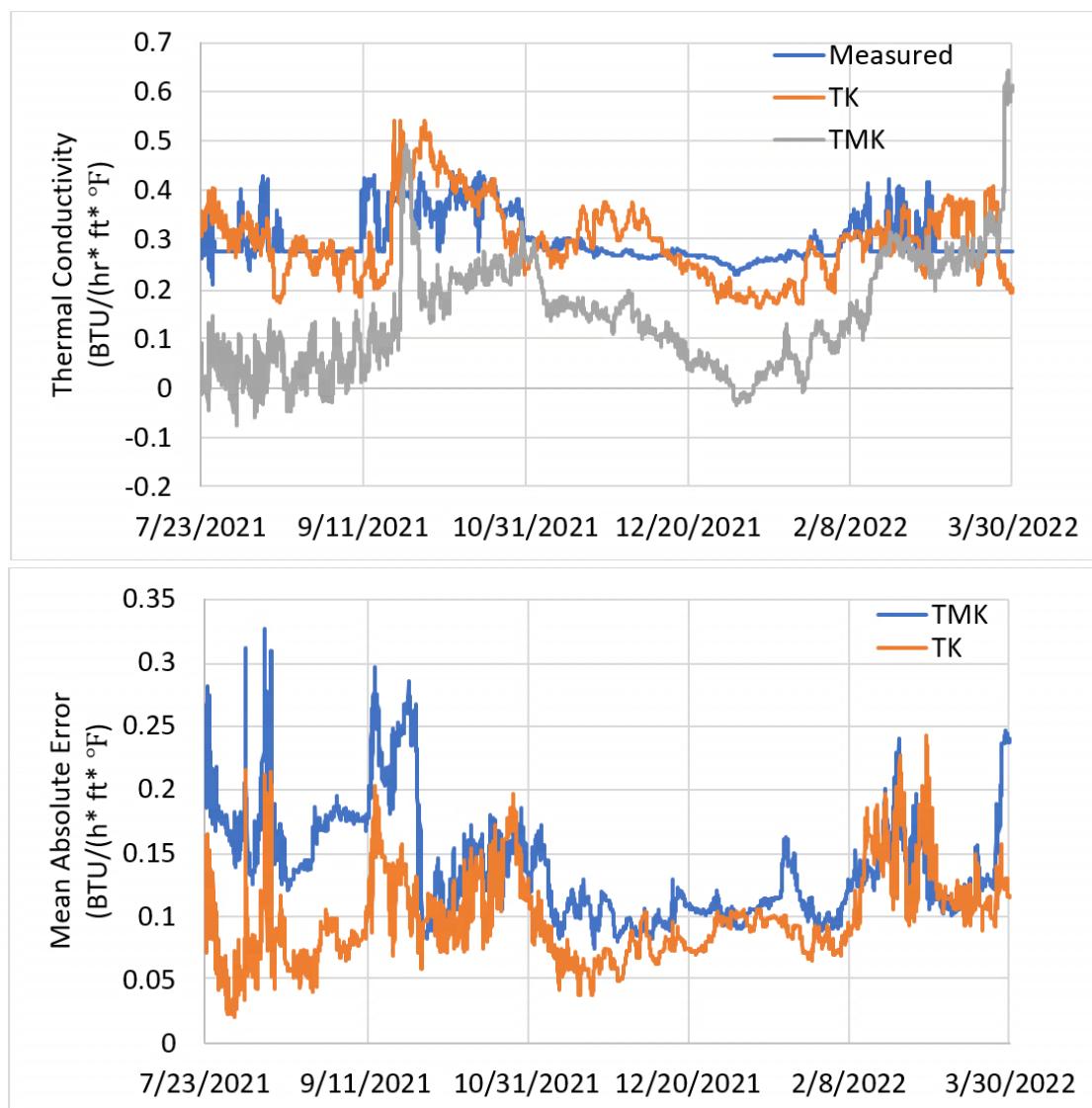


Figure 7. Thermal Conductivity model results for Hill AFB, Winter 21/22. Top, Thermal Conductivity for 1 in. depth. Bottom, Thermal Conductivity Mean Absolute Error across all depths. Models are labeled if they used Temperature (T), Volumetric Moisture Content (M), and/or Thermal Conductivity (K) as input.

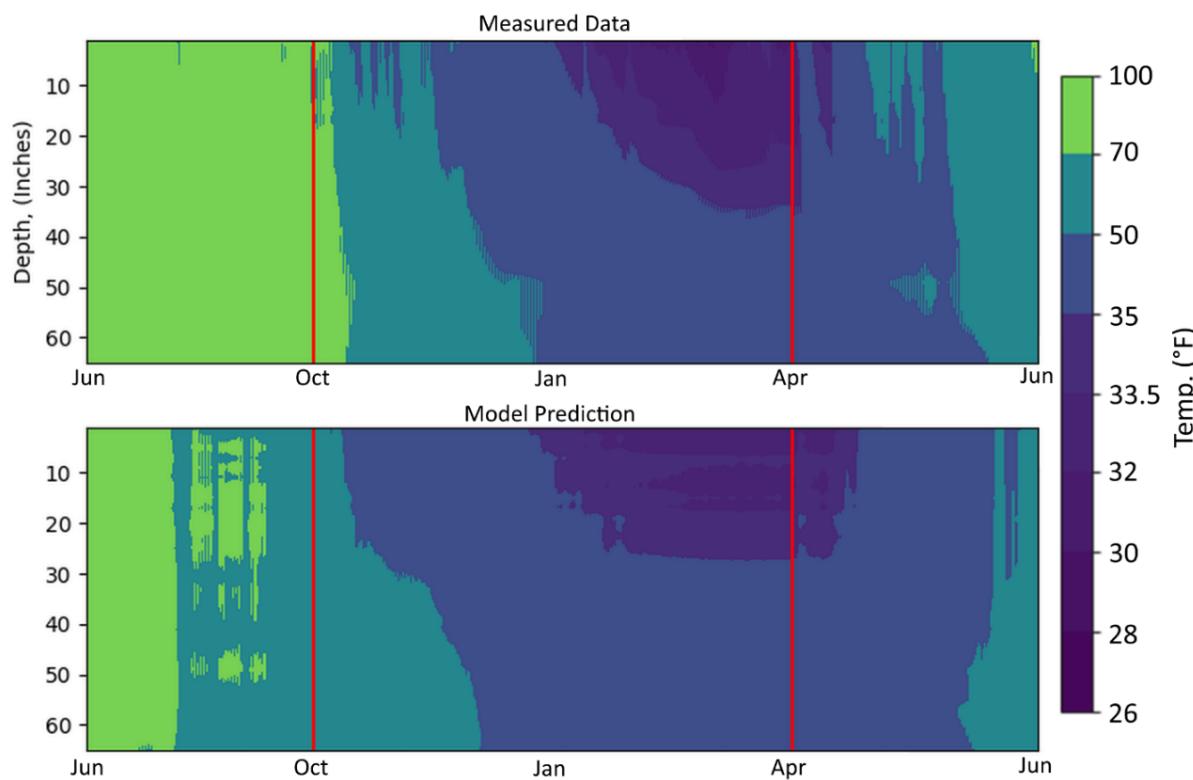


Figure 8. Soil Temperature Profile for Hill AFB, Winter 21/22, Top, compared to Model TK prediction, Bottom. Red lines correspond with October 1st 2021 and April 1st 2022 Model is labeled by input parameters, with Temperature, T and Thermal Conductivity, K.

CONCLUSION

Recurrent neural networks, especially those using Gated Recurrent Units, have shown to be effective at modeling soil properties relevant to frost effects, including temperature, volumetric moisture content, and thermal conductivity. Of the candidate models, the temperature-only model performed the worst while the joint temperature-thermal conductivity model performed the best. The general model using all three inputs also performed well, implying that soil properties should not be analyzed individually, and can be used to infer each other. In terms of forecasting, the temperature-thermal conductivity model had an error for a 120 hour period of 4.77 °F, which could be of use for traffic forecasting for airfields and runways in frost affected regions. Of the properties besides temperature, thermal conductivity provided the most significant benefits, being linked to increases in forecasting accuracy and a lower temperature mean absolute error. Volumetric moisture content, however, did not seem to consistently improve prediction accuracy, potentially due to inaccuracies in TDR measurements in frozen soils. It is recommended in the future to use alternate methods of recording volumetric moisture content, and to train this architecture on more soil conditions, as Hill AFB and Air Force Academy are in largely the same region of the United States and are both high plasticity clay soils. Training data from additional sites would be beneficial and further improve the model's accuracy and forecasting capabilities, leading to a new capability for modeling soil frost effects.

REFERENCES

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., and Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8.
- Barooni, M., Ziarati, K., and Barooni, A. (2023). Frost Prediction Using Machine Learning Methods in Fars Province. *28th International Computer Conference*. Tehran: Computer Society of Iran.
- Bianchini, A., and Gonzalez, C. R. (2012). *Pavement-Transportation Computer Assisted Structural Engineering (PCASE) Implementation of the Modified Berggren (ModBerg) Equation for Computing the Frost Penetration Depth within Pavement Structures*. Vicksburg: US Army Corps of Engineers.
- Choi, H.-J., Kim, S., Kim, Y., and Won, J. (2022). Predicting Frost Depth of Soils in South Korea Using Machine Learning Techniques. *Sustainability*(14), 9767.
- Dayarathne, R., Hawlader, B., Phillips, R., and Robert, D. (2023). Two-dimensional finite element modeling of long-term frost heave beneath chilled gas pipelines. *Cold Regions Science and Technology*, 208.
- Lein, A. W., Sloane, S. L., Smith, C. E., Jr., Bernier, A. P., and Oren, J. I. (2019). *Frost-Depth Penetration and Frost Heave in Frost-Susceptible Soils*. Hanover: US Army Corps of Engineers.
- Novak, D. R., Bailey, C., Brill, K. F., Burke, P., Hogsett, W. A., Rausch, R., and Schichtel, M. (2014). Precipitation and Temperature Forecast Performance at the Weather Prediction Center. *Weather and Forecasting*, 29(3), 489-504.
- Pollard, W. H. (2017). Chapter 15 - Periglacial Processes in Glacial Environments. In J. Menzies, & J. J. van der Meer, *Past Glacial Environments* (pp. 537-564). Elsevier.
- Rajaei, P., and Baladi, G. Y. (2015). Frost Depth: General Prediction Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2510, 74-80.
- Talsma, C. J., Solander, K. C., Mudunuru, M. K., Crawford, B., and Powell, M. R. (2023). Frost prediction using machine learning and deep neural network models. *Frontiers in Artificial Intelligence*.
- Zaytar, M. A., and Amrani, C. E. (2016). Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks. *International Journal of Computer Applications*, 143(11), 0975-8887.

Exploring the Determinants of Pedestrian Crash Severity Using an AutoML Approach

Amir Rafe¹ and Patrick A. Singleton²

¹Dept. of Civil and Environmental Engineering, Utah State Univ., Logan, UT
(corresponding author). Email: amir.rafe@usu.edu

²Dept. of Civil and Environmental Engineering, Utah State Univ., Logan, UT.
Email: patrick.singleton@usu.edu

ABSTRACT

This study investigates pedestrian crash severity through automated machine learning (AutoML), offering a streamlined and accessible method for analyzing critical factors. Utilizing a detailed dataset from Utah spanning 2010–2021, the research employs AutoML to assess the effects of various explanatory variables on crash outcomes. The study incorporates SHAP (SHapley Additive exPlanations) to interpret the contributions of individual features in the predictive model, enhancing the understanding of influential factors such as lighting conditions, road type, and weather on pedestrian crash severity. Emphasizing the efficiency and democratization of data-driven methodologies, the paper discusses the benefits of using AutoML in traffic safety analysis. This integration of AutoML with SHAP analysis not only bolsters predictive accuracy but also improves interpretability, offering critical insights into effective pedestrian safety measures. The findings highlight the potential of this approach in advancing the analysis of pedestrian crash severity.

Keywords: Pedestrian Crash Severity, AutoML, Random Forest, SHAP

INTRODUCTION

Pedestrian safety is a critical concern in urban transportation networks, as pedestrians represent a highly vulnerable group in traffic-related incidents. The urgency to enhance their protection and mitigate the severity of pedestrian-involved crashes cannot be overstated. Recent statistics from the National Highway Traffic Safety Administration (NHTSA) (2022) highlight this concern: in 2021, there were 7,388 pedestrian fatalities in the United States due to traffic accidents. The nature and severity of pedestrian crashes are influenced by a multitude of factors. Analysis of NHTSA data from 2021 reveals that pedestrian fatalities are predominantly occurring in urban areas (83%), at intersections (23%), and in conditions of low visibility, such as under dark lighting (39%). Furthermore, a report from the World Health Organization (WHO, 2022) underscores a startling fact: the likelihood of fatality for a pedestrian struck by a vehicle increases exponentially with the vehicle's speed, rising by 4.5 times with every 10mph increase in speed. Understanding these contributing elements is vital in devising strategies to decrease the number of pedestrian casualties and enhance overall traffic safety for pedestrians.

To effectively discern the determinants of pedestrian injury severity, it is common practice to examine historical crash data and employ predictive modeling techniques. The field of pedestrian safety research extensively utilizes a variety of analytical methods, including statistical analysis, machine learning, and deep learning algorithms, to assess the impact of different factors on the severity of pedestrian-involved accidents. Enhancing the precision and efficiency of these

predictive models is crucial for policymakers and traffic safety engineers. A thorough understanding of the key variables that contribute to reducing pedestrian injuries and fatalities enables these stakeholders to focus on the most effective safety interventions in their planning. Furthermore, the efficacy of these interventions can be assessed through comparative studies conducted before and after their implementation. In summary, refining these predictive models is instrumental in developing focused and efficacious strategies to protect pedestrians and decrease the incidence of injuries and fatalities in traffic crashes.

The advancement of artificial intelligence and the integration of machine learning (ML) in various fields have led to this study's objective: **introducing an Automated Machine Learning (AutoML) methodology for predicting the severity of pedestrian crashes.** This AutoML approach is designed to simplify the process of evaluating different ML methods on safety data, selecting the most effective one. It particularly focuses on providing an accessible, low-code solution for transportation engineers and decision-makers who may not be well-versed in computer science yet seek high-accuracy predictions to inform reliable decision-making. Furthermore, this study aims to pinpoint key factors influencing pedestrian injury severities. To accomplish this, we utilized pedestrian crash data from Utah spanning the years 2010 to 2021. We applied various ML methods and developed a web application centered around AutoML, making it adaptable to any crash severity dataset. Additionally, to interpret the results from the chosen ML method, we utilized SHapley Additive exPlanations (SHAP), a method grounded in game theory.

In the sections that follow, we will conduct a literature review to examine prior studies that have implemented innovative techniques in crash severity prediction. This will be followed by a detailed presentation of our data and methodologies, the results we obtained, and a discussion focusing on the principal findings, interpretation of the model, and our concluding insights.

LITERATURE REVIEW

Extensive research has been conducted to understand the multifaceted nature of pedestrian crash severity, with studies examining a range of influential factors. A thorough examination of the current literature, as summarized by (Shrinivas et al., 2023), identifies several key determinants. These include demographic aspects of the pedestrian such as age and gender, characteristics of the vehicle involved including its speed and type, environmental factors like the location of the collision and the time it occurred, and behavioral elements such as the involvement of alcohol or drugs in either the pedestrian or the driver. Additionally, the use of safety gear, such as helmets and reflective clothing, has been noted as an important factor.

Various studies have implemented ML approaches to assess pedestrian crash severity. Al-Mistarehi et al. (2022) explored several techniques, including decision trees, k-nearest neighbors (KNN), naive Bayes, and AdaBoost, and found that the random forest model outperformed others in accurately predicting different injury types, boasting the lowest error rate. In a different study, Goswamy et al. (2023) employed XGBoost and Random Parameters Discrete Outcome Models (RPDOM), uncovering that Rectangular Rapid Flashing Beacons (RRFB) notably reduce the severity of nighttime pedestrian crashes. Their study highlighted XGBoost's superior predictive accuracy of 97%, in contrast to RPDOM's 73.8%. Rahman et al.'s (2023) research on Utah state highways utilized boosted decision trees to identify risk factors affecting pedestrian crash frequencies. Their analysis revealed a correlation between higher motor vehicle and pedestrian volumes, transit stops, and certain demographic areas with increased crashes, while also noting a 'safety-in-numbers' effect where higher pedestrian volumes corresponded to lower crash rates.

Effati & Vahedi Saheli (2022) compared logistic regression and classification and regression trees (CART) for estimating pedestrian crash occurrences in rural settings. Their findings suggested a marginally better performance of logistic regression over CART in identifying risky segments for pedestrians. Lastly, Yang et al. (2022) investigated the effectiveness of three ML methods—Support Vector Machines (SVM), Ensemble Decision Trees (EDT), and k-Nearest Neighbors (KNN)—each optimized using a Bayesian algorithm for predicting pedestrian fatalities in road crashes. The study found that while all models improved in performance due to optimization, the KNN model showed the most significant accuracy enhancement. However, the SVM and EDT models still exhibited higher overall accuracy than the KNN model.

Following the exploration of various machine learning models in crash severity analysis, it is pertinent to discuss the benefits and challenges these models present. One significant advantage is their higher predictive accuracy. Machine learning methods can outperform traditional statistical models, especially in situations involving complex, non-linear relationships among variables or in the analysis of large and intricate datasets (Komol et al., 2021). Additionally, these models excel at identifying and understanding the importance of different features (explanatory variables), elucidating intricate relationships with crash severity. This capability often yields insights that might be less apparent or more difficult to extract using statistical models (Komol et al., 2021). However, machine learning models are not without their limitations. A notable challenge is their interpretability; these models can be more complex and less intuitive than their statistical counterparts, potentially complicating the explanation of how various factors relate to crash severity (Infante et al., 2022). Moreover, there is the risk of overfitting—where a model becomes too tailored to the training data, losing its ability to generalize and perform accurately on new, unseen data. This issue is particularly prevalent when models are trained with a vast number of features or a limited dataset (Komol et al., 2021).

Addressing the limitations of traditional machine learning approaches, AutoML stands as a promising solution in crash severity analysis. AutoML enhances efficiency by simplifying the model selection process, reducing the need for in-depth machine learning expertise and computational resources (Angarita-Zapata et al., 2021). This is particularly beneficial in complex scenarios like crash severity prediction, where choosing an appropriate model can be challenging. Moreover, AutoML can test a broader range of ML methods on the data, unlike traditional studies that often explore a limited selection. Empirical results indicate that AutoML's capabilities go beyond mere convenience; it can match or even exceed the performance of conventional methods with reduced human intervention (Angarita-Zapata et al., 2021). This reflects AutoML's sophisticated algorithmic prowess, adept at sifting through various machine learning models to identify the most effective one for a specific dataset. The implications for research in road safety systems are profound. AutoML holds the potential to refine model selection and enhance prediction accuracy in this vital field. Recognizing these encouraging outcomes, this paper suggests a broader investigation into the utility of AutoML for pedestrian crash severity analysis. Such exploration could yield deeper insights into its efficacy and versatility, paving the way for more sophisticated and precise road safety systems.

DATA AND METHOD

Data and Variables

For our research, we employed crash data reported by the Utah Department of Public Safety (UDPS, 2023), focusing on pedestrian crash severities in Utah from 2010 to 2021. The severity

of pedestrian crashes in this study was categorized into three levels: fatal, serious injury, and minor injury. Figure 1 illustrates the geographical distribution of these crashes, including a heatmap that specifically emphasizes the locations of fatal crashes within the dataset. This figure highlights that the density of pedestrian crashes is predominantly concentrated in Salt Lake City, followed by Ogden and Provo.

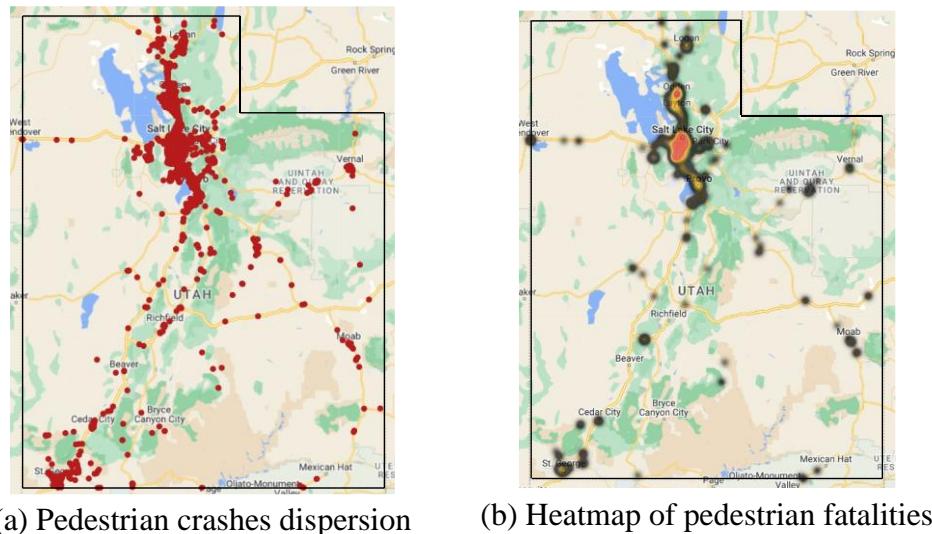


Figure 1. The spatial configuration of pedestrian crashes

In this study, we analyzed 8,319 pedestrian crashes and examined the influence of 17 explanatory variables on crash severity. Table 1 presents the descriptive statistics for these variables. For the scope of this study, non-injury pedestrian crashes (Property Damage Only, PDO) were not included, as the primary focus was on analyzing factors influencing the severity of injuries in pedestrian crashes. Regarding crash severity categorization, the 'minor injury' level in this research encompasses both 'minor injury' and 'possible injury' classifications according to the KABCO scale. According to the data presented in Table 1, males are more frequently involved in pedestrian crashes, representing 59% of all incidents, and they account for a higher proportion of fatal (65%) and serious injuries (62%). The age distribution presents a noteworthy finding: individuals aged 30 to 59 years are involved in 34% of crashes and constitute 42% of fatalities, while the younger cohort of 10 to 29 years old exhibits a higher incidence of serious injuries, at 41%. Substance influence is also a critical variable; crashes with both alcohol and drugs detected are less common but have a 100% fatality rate, highlighting the exacerbated risk of pedestrian fatalities when both substances are involved. Most pedestrian crashes happened under conditions where DUI was not reported (97%) and without distracted driving (92%).

The involvement of teenage drivers appears in 10% of fatal cases, suggesting a potential area for intervention. Intersections emerged as hotspots for pedestrian crashes, comprising 61% of all incidents. Concerningly, close to 3% of these intersection-related crashes had fatal outcomes for pedestrians. As for the nature of the crashes, incidents involving both left-turn and right-turn movements showed comparable rates of occurrence. Crashes predominantly occur in urban settings (97%) and on arterial roads (52%), pinpointing these as high-risk environments. The data also reveals that a significant number of fatal (37%) occur in dark-not-lighted conditions, which may signal a need for improved lighting to enhance pedestrian safety.

Table 1. Descriptive statistics of the variables

Characteristics	Class	Total	Fatal	Serious injury	Minor injury
Pedestrian crashes		8319	476 (6%)	1363 (16%)	6480 (78%)
Sex	Male	4933 (59%)	309 (65%)	849 (62%)	3775 (58%)
	Female	3386 (41%)	167 (35%)	514 (38%)	2705 (42%)
Age group	0 to 9	782 (9%)	33 (7%)	118 (9%)	631 (10%)
	10 to 29	3795 (46%)	119 (25%)	556 (41%)	3120 (48%)
	30 to 59	2842 (34%)	202 (42%)	509 (37%)	2131 (33%)
	> 59	900 (11%)	122 (26%)	180 (13%)	598 (9%)
Alcohol-drug test result	Both-Positive	11 (0%)	11 (2%)	0 (0%)	0 (0%)
	Drug-Positive	34 (0%)	34 (7%)	0 (0%)	0 (0%)
	Alcohol-Positive	15 (0%)	13 (3%)	2 (0%)	0 (0%)
	Negative	9 (0%)	9 (2%)	0 (0%)	0 (0%)
	Not related	8250 (99%)	409 (86%)	1361 (100%)	6480 (100%)
DUI	No	8102 (97%)	413 (87%)	1305 (96%)	6384 (99%)
	Yes	217 (3%)	63 (13%)	58 (4%)	96 (1%)
Distracted driving	No	7651 (92%)	430 (90%)	1218 (89%)	6003 (93%)
	Yes	668 (8%)	46 (10%)	145 (11%)	477 (7%)
Teenage driver involved	No	7523 (90%)	432 (91%)	1205 (88%)	5886 (91%)
	Yes	796 (10%)	44 (9%)	158 (12%)	594 (9%)
Holiday	No	7308 (88%)	397 (83%)	1185 (87%)	5726 (88%)
	Yes	1011 (12%)	79 (17%)	178 (13%)	754 (12%)
Right-turn involved	No	6743 (81%)	464 (97%)	1254 (92%)	5025 (78%)
	Yes	1576 (19%)	12 (3%)	109 (8%)	1455 (22%)
Intersection involved	Yes	5091 (61%)	136 (29%)	718 (53%)	4237 (65%)
	No	3228 (39%)	340 (71%)	645 (47%)	2243 (35%)
Left-turn involved	No	6652 (80%)	441 (93%)	1144 (84%)	5067 (78%)
	Yes	1667 (20%)	35 (7%)	219 (16%)	1413 (22%)
Work zone involved	No	7953 (96%)	447 (94%)	1298 (95%)	6208 (96%)
	Yes	366 (4%)	29 (6%)	65 (5%)	272 (4%)
Road type	Urban	8082 (97%)	419 (88%)	1296 (95%)	6367 (98%)
	Rural	237 (3%)	57 (12%)	67 (5%)	113 (2%)
Functional class	Local	2526 (30%)	71 (15%)	352 (26%)	2103 (32%)
	Collector	1493 (18%)	71 (15%)	232 (17%)	1190 (18%)
	Arterial	4300 (52%)	334 (70%)	779 (57%)	3187 (49%)
Roadway surface is dry	Yes	7175 (86%)	409 (86%)	1181 (87%)	5585 (86%)
	No	1144 (14%)	67 (14%)	182 (13%)	895 (14%)
Lighting condition	Dark-Not lighted	1117 (13%)	176 (37%)	285 (21%)	656 (10%)
	Dark-Lighted	1836 (22%)	138 (29%)	332 (24%)	1366 (21%)
	Daylight	4953 (60%)	141 (30%)	678 (50%)	4134 (64%)
	Dusk	224 (3%)	10 (2%)	40 (3%)	174 (3%)
	Dawn	189 (2%)	11 (2%)	28 (2%)	150 (2%)
Weather condition	Clear	6378 (77%)	355 (75%)	1068 (78%)	4955 (76%)
	Cloudy	1141 (14%)	69 (14%)	176 (13%)	896 (14%)
	Rain	484 (6%)	31 (7%)	80 (6%)	373 (6%)
	Fog, Smog	24 (0%)	3 (1%)	4 (0%)	17 (0%)
	Snowing	201 (2%)	10 (2%)	26 (2%)	165 (3%)
	Others	91 (1%)	8 (2%)	9 (1%)	74 (1%)
Vertical alignment	Level	6515 (78%)	360 (76%)	1108 (81%)	5047 (78%)
	Uphill	57 (1%)	3 (1%)	7 (1%)	47 (1%)
	Downhill	50 (1%)	2 (0%)	12 (1%)	36 (1%)
	Others	1697 (20%)	111 (23%)	236 (17%)	1350 (21%)

For uniform encoding of the dataset, numerical assignments were utilized such as “Yes” being coded as 1, “No” as 0; “Male” as 1, “Female” as 0; “Rural” as 1, “Urban” as 0. Other categories were numerically encoded in ascending order, beginning with 1 as outlined in Table 1. Moreover, age was treated as a continuous variable, rather than categorizing it into discrete age groups.

METHOD

As previously mentioned, this study utilized an AutoML methodology to investigate the impact of various explanatory variables (EV), or features, on the outcomes of pedestrian crash injuries, based on the pipeline depicted in Figure 2. This methodology encompasses several critical steps, each employing specific techniques to ensure the robustness and accuracy of the predictive models.

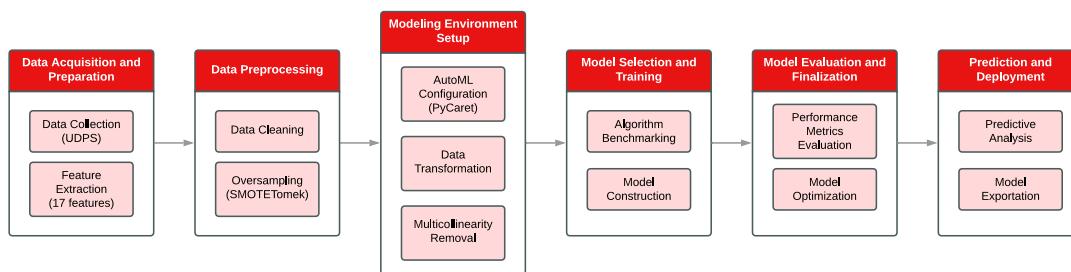


Figure 2. The AutoML pipeline for pedestrian crash severity analysis in this study

Initially, the study embarked on data acquisition, sourcing information from the UDPS and conducting feature extraction to identify the most relevant variables for the analysis. Subsequent data preprocessing involved cleaning and addressing class imbalance through oversampling methods to prepare a reliable dataset for model training. To address class imbalance in the dataset, the Synthetic Minority Over-sampling Technique coupled with Tomek link removal (SMOTETomek) (Batista et al., 2003) was employed to balance the severity classes. This method not only augments the dataset with additional synthetic minority class examples but also refines the class boundaries, thereby enhancing the model's predictive accuracy for less represented classes.

In setting up the AutoML environment, PyCaret (Moez Ali, 2020) was instrumental in configuring the necessary parameters, including data normalization and transformation processes. Additionally, features exhibiting high multicollinearity were identified and removed to prevent potential overfitting and improve model interpretability. The core of the AutoML approach was the model selection and training phase. Here, a variety of machine learning algorithms were evaluated and benchmarked to identify the model that best fits the data, considering the complex nature of crash severity outcomes. This was followed by a comprehensive evaluation of the chosen model, utilizing a suite of performance metrics such as F1 score, accuracy, precision, and recall assessing its predictive capabilities. Accuracy is the metric that measures the percentage of correctly classified observations across all categories. Precision quantifies the proportion of true positives among all positive predictions, indicating the model's exactness. Recall measures the proportion of actual positives that were correctly identified, reflecting the model's completeness. The F1 score, a calculated harmonic mean of

precision and recall, provides a balanced measure of the model's precision and robustness in identifying positive classes. The mathematical formulations for these metrics are delineated by Equations 1 to 4.

$$\text{accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} \quad (1)$$

$$\text{precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

$$\text{recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

$$\text{F1 score} = \frac{2 \times (\text{precision} \times \text{recall})}{\text{precision} + \text{recall}} \quad (4)$$

Finally, the selected model underwent optimization and finalization, which involved fine-tuning to enhance its performance. The robustness of the model was validated, and upon satisfactory evaluation, it was deployed for predictive analysis.

In our study, we employed SHAP (Lundberg et al., 2018; Lundberg & Lee, 2017) to interpret the outcomes of selected ML model. SHAP provides a detailed measure of feature importance, assigning an importance value to each EV for a given prediction. This approach enhances the transparency of model predictions by translating them in terms of input EVs. Originating from cooperative game theory, SHAP values distribute the "payout" (in this case, the predicted crash severity) among the "players" (the EVs), based on their individual contributions to the prediction. To understand this mathematically, if $f(x)$ represents the model's prediction for a specific instance x , and $E[f(X)]$ represents the expected model prediction (calculated as the mean prediction over the training set), then the contribution of each EV can be quantitatively expressed using an additive attribution model.

$$f(x) - E[f(X)] = \sum_{i=1}^N \varphi_i \quad (5)$$

Additionally, the significance of each EV, also known as the Shapley value for the i -th EV, denoted as φ_i is computed as follows:

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \left[\frac{|S|!(|N|-|S|-1)!}{|N|!} \right] (f_i(S \cup \{i\}) - f_i(S)) \quad (6)$$

In this context, N is the set of all EVs, S is a subset of N that includes the i -th EV, $|S|$ is the size of S , and f_i is a version of f where only the EVs in S and i (if it's included) are used.

MODEL RESULTS

The evaluation of various machine learning models yielded a range of performance outcomes, as summarized in Table 2. This table provides a detailed comparison of the models based on key performance metrics. The Random Forest model stood out in our comparative analysis of machine learning algorithms for predicting pedestrian crash severity. To assess its predictive performance, we employed several metrics visualized in Figures 3. The Receiver

Operating Characteristic (ROC) curve and its corresponding Area Under the Curve (AUC) value in Figure 3 (d) measure the model's ability to correctly classify the severity levels, with a higher AUC reflecting a better overall performance. The confusion matrix in Figure 3 (c) provides a detailed breakdown of the model's predictions, showing the proportions of true positives, false positives, true negatives, and false negatives, which are critical for understanding the model's classification accuracy. In Figure 3 (b), the Precision-Recall Curve offers a view of the model's precision, or the accuracy of predicting positive classes, against its recall, the model's ability to capture all actual positives, which is particularly important in the context of imbalanced classes. Finally, Figure 3 (a), through SHAP values, ranks the features by their importance, showing how each one influences the model's predictions and providing insights into the decision-making process of the Random Forest model. These metrics collectively paint a comprehensive picture of the model's performance and its interpretive power regarding the factors that affect pedestrian crash severity.

Table 2. Performance metrics of ML models for pedestrian crash severity prediction

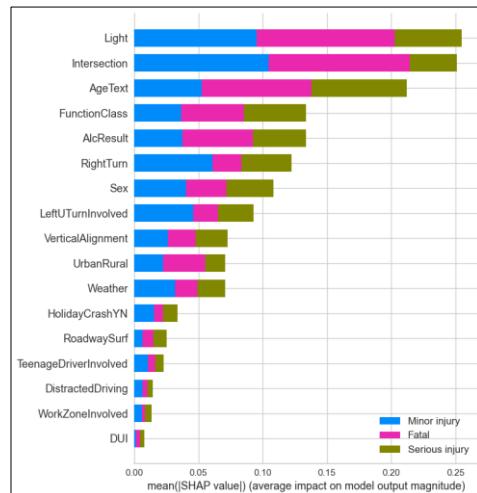
Model	Accuracy	AUC	Recall	Precision	F1 score
Random Forest Classifier	0.83	0.94	0.83	0.83	0.83
Extra Trees Classifier	0.82	0.93	0.82	0.82	0.82
Extreme Gradient Boosting	0.81	0.93	0.81	0.81	0.81
Decision Tree Classifier	0.80	0.87	0.80	0.80	0.80
CatBoost Classifier	0.80	0.93	0.80	0.80	0.80
Light Gradient Boosting Machine	0.78	0.92	0.78	0.78	0.78
K Neighbors Classifier	0.75	0.89	0.75	0.75	0.75
Gradient Boosting Classifier	0.69	0.87	0.69	0.70	0.69
Ada Boost Classifier	0.66	0.80	0.66	0.66	0.66
Logistic Regression	0.60	0.79	0.60	0.60	0.60
Linear Discriminant Analysis	0.59	0.79	0.59	0.61	0.60
Ridge Classifier	0.60	0.00	0.60	0.60	0.60
SVM - Linear Kernel	0.58	0.00	0.58	0.57	0.55
Quadratic Discriminant Analysis	0.57	0.77	0.57	0.57	0.55
Naive Bayes	0.57	0.77	0.57	0.58	0.55
Dummy Classifier	0.33	0.50	0.33	0.11	0.16

MODEL INTERPRETATION AND DISCUSSION

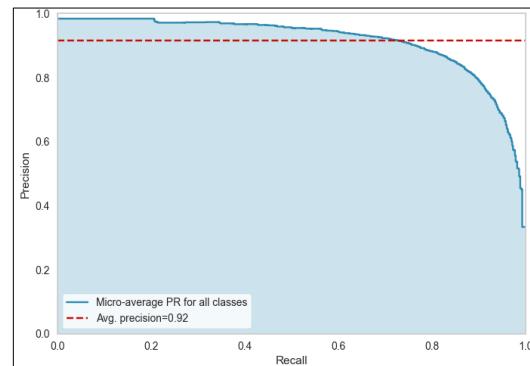
In this section, we delve into the results yielded by the AutoML process, with a particular focus on the Random Forest classifier, which was identified as the best-performing model. The Random Forest model demonstrated superior performance with an AUC of 0.94, indicating a high capability to distinguish between the different severity levels of pedestrian crashes. The ROC curve, which plots the true positive rate against the false positive rate, further confirmed the model's robustness in classification across varying thresholds. The model's precision and recall were balanced, as evidenced by the Precision-Recall Curve, with an average precision of 0.92. This balance is crucial in practice, as it ensures that the model is not only accurate in its predictions but also consistent in identifying the true instances of each class.

The confusion matrix provided a more granular view of the model's predictive accuracy, revealing that the model successfully classified 78% of the class 0 (minor injury) instances, 81% of class 1 (serious injury), and 94% of class 2 (Fatal) instances correctly. This distribution

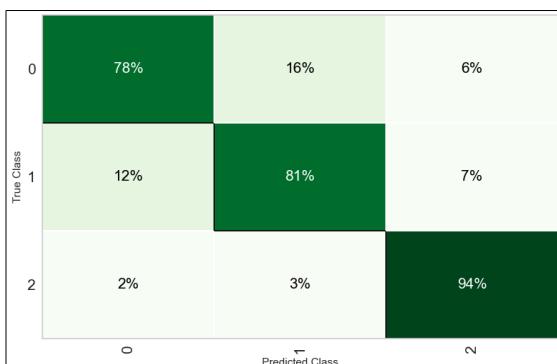
indicates a strong predictive power, particularly for the severe class, which is often the most challenging to predict due to the typically lower number of instances.



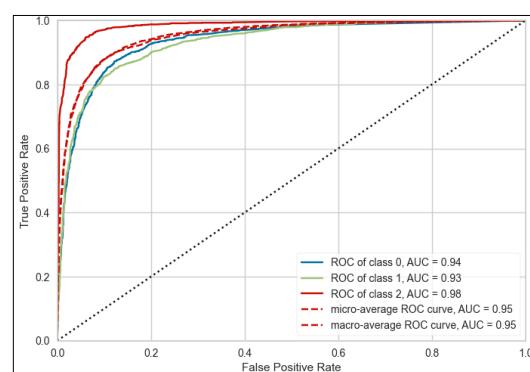
(a) Feature importance Derived from SHAP Values



(b) Precision-Recall curve



(c) Confusion matrix (0: Minor injury, 1: Serious injury, 2: Fatal)



(d) ROC curves

Figure 3. Comprehensive evaluation results of the Random Forest model

The SHAP analysis, which offers a game-theoretic approach to feature importance, shed light on the impact of various explanatory variables on the model's predictions. Features like lighting condition, whether the crash occurred at the intersection, and age were shown to have a substantial influence on the model's output, indicating that these factors play a significant role in the severity of pedestrian crashes. The SHAP values allow us to quantify the contribution of each feature to the prediction of each class, providing a clear picture of the underlying factors that the model considers when making predictions.

In assessing the Random Forest model's interpretation for each injury severity class, the SHAP value plots reveal the varied influence of explanatory variables. Figure 4 displays a SHAP summary plot that depicts how EV features influence the likelihood of each crash severity class. Each feature is represented by a row in the plot, with dots colored according to the feature's value—red indicating higher values and blue for lower ones. The placement of the dots horizontally indicates the direction of the feature's impact on the model's prediction: right for an

increase and left for a decrease in the likelihood of the positive class. A feature's consistent influence on prediction is suggested by dots clustering on one side, while the spread of dots across the horizontal axis indicates the variability of the feature's impact. The dispersion of dots within a row suggests the presence of interactions with other features that affect the predictive outcome. The ordering of the rows reflects the aggregate strength of the SHAP values across all data points, thus providing a ranked overview of feature importance in the model's predictions.

In the context of minor injuries, our analysis reveals that 'Light' and 'Intersection' significantly influence model predictions, as indicated by their notable SHAP values (Figure 4a). This observation corroborates findings from previous research, which underscores the critical role of lighting conditions and intersection presence in pedestrian crash severity (Harris et al., 2023). Specifically, well-lit areas and intersections, by enhancing visibility and imposing traffic controls, are associated with reduced injury severity following crashes. Such environments provide cues for safer pedestrian and vehicle interaction, potentially mitigating the impact of accidents. Furthermore, the 'VerticalAlignment' factor's prominence in our model aligns with broader safety studies, albeit less directly examined in pedestrian contexts. Insights from research on vehicle dynamics on mountainous freeways suggest that level road geometries contribute to lower crash severities, a principle extendable to pedestrian safety scenarios. Our findings suggest that flat terrains, enhancing pedestrian visibility and influencing vehicle speed, can similarly reduce the severity of pedestrian injuries.

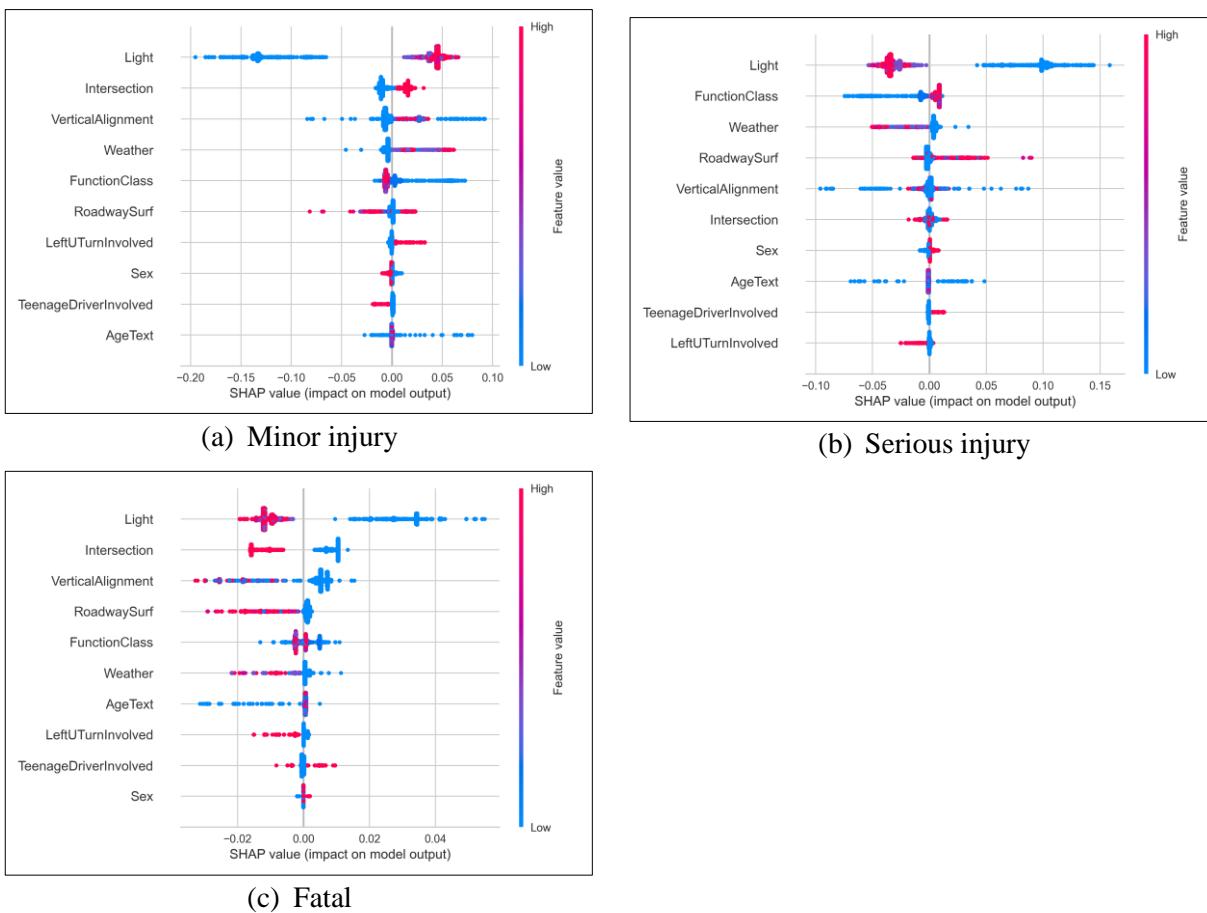


Figure 4. The SHAP summary plot for each crash severity class in Random Forest model

In our analysis of serious injuries, as depicted in Figure 4b, the 'RoadwaySurf' feature underscores the heightened risk associated with non-dry road conditions, corroborating findings from recent studies that emphasize the adverse effects of wet or icy surfaces on crash severity. These conditions notably compromise vehicle control and extend stopping distances, thereby elevating the likelihood of severe pedestrian injuries. This aligns with research highlighting the critical role of road surface conditions in influencing crash outcomes, where slippery roads are identified as a significant contributor to crash incidence and severity (Lee et al., 2023).

Furthermore, the analysis reveals a critical relationship between lighting conditions and injury severity, with 'Light' indicating that diminished visibility in lower light conditions amplifies the risk of serious injuries. This observation is supported by studies identifying darkness as a key factor in exacerbating pedestrian crash severity, reflecting the challenges drivers face in detecting pedestrians under such conditions (Lee et al., 2023).

Also, the 'FunctionClass' feature demonstrates a differentiated impact on the likelihood of serious injuries in pedestrian crashes. The SHAP values suggest that roads classified as 'Arterial' have a greater association with serious injury outcomes compared to 'Local' or 'Collector' roads. This could be due to arterial roads typically having higher traffic volumes, speeds, and more complex traffic patterns, which may increase the risk and potential severity of pedestrian accidents. The 'FunctionClass' variable offers additional insights into the risk associated with different road types. Arterial roads, characterized by higher traffic volumes and speeds, are shown to have a stronger association with serious injury outcomes, a finding that echoes research on the impact of road classification on pedestrian safety. Studies from the UK (Salehian et al., 2023) suggest that major roads, analogous to the 'Arterial' classification in our study, pose increased risks due to their traffic characteristics, reinforcing the need for targeted safety measures on such thoroughfares. The convergence of these insights underscores the critical need for enhanced safety protocols and infrastructure improvements on arterial roads to mitigate the higher risks of serious injuries in pedestrian crashes.

In Figure 4c, pertaining to the fatal injury class, the 'Light' feature reveals a profound impact, with darker conditions, particularly 'Dark-Not lighted', significantly increasing the probability of a fatal outcome. This aligns with the expectation that poor lighting can severely impair a driver's visibility and reaction time, leading to more severe accidents (Ferenchak et al., 2022). 'Intersection' also stands out as a critical factor, indicating that incidents occurring at intersections are more likely to result in fatal injuries. This may be due to the complex traffic dynamics and potential for higher-speed collisions at these locations (Siddiqui et al., 2006). Lastly, 'AgeText' appears to have a noticeable effect on the model's predictions for fatal injuries. This suggests that age, as a proxy for factors like physical vulnerability and reaction time, is a significant predictor of the severity of outcomes when pedestrians are involved in traffic incidents, with certain age groups possibly being at a higher risk for fatal outcomes (Kitali et al., 2017).

To navigate the intricacies of the dot summary plot and deepen our understanding of how each EV contributes to the model's predictions, we employed SHAP force plots. Figure 5 showcases such a plot for a specific case in our dataset, observation #313. This visualization (Figure 5a) elucidates the Random Forest model's inclination to categorize the observation as a minor injury, where the prediction score $f(x)$ is 0.82 against a base value of approximately 0.775. Notably, the lighting condition emerges as the most influential factor in this classification.

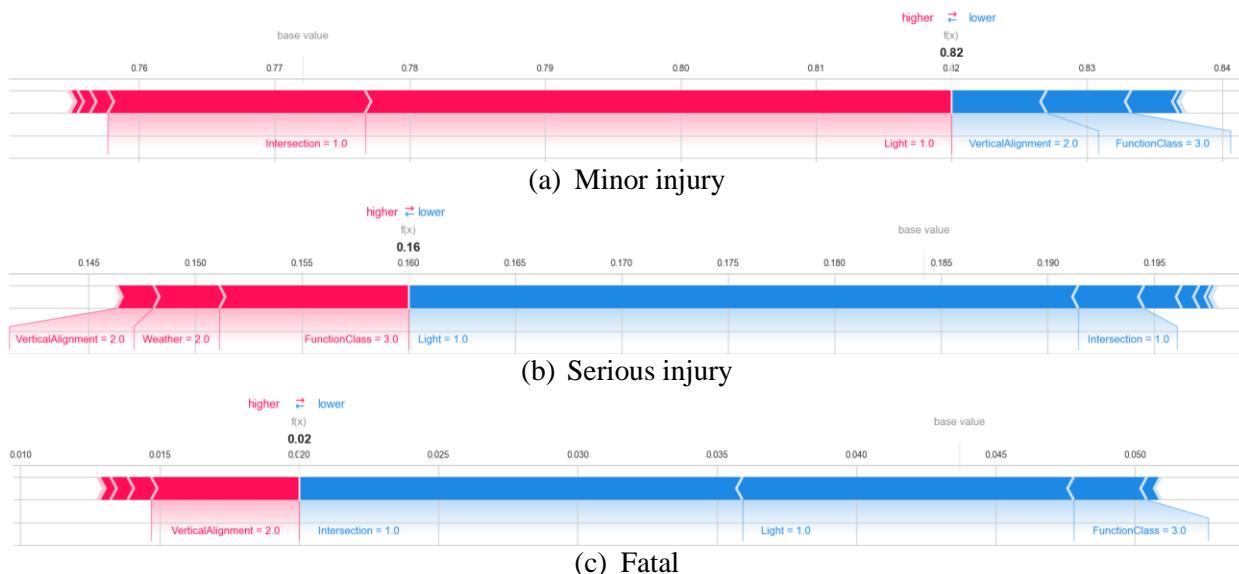


Figure 5. The SHAP values, explaining the contribution of EVs to the raw Random Forest model output for a specific observation.

CONCLUSION

In summary, this paper has highlighted the viability of AutoML for pedestrian crash severity analysis, offering a practical alternative to more complex deep learning models such as TabNet, which, although previously applied to this data achieving around 95% accuracy (Rafe & Singleton, 2023b), require greater computational resources and technical knowledge. The relative simplicity and reduced resource needs of AutoML position it as an accessible and efficient solution for diverse applications. We've also developed a user-friendly web application that enables users to build crash severity prediction models via a straightforward data upload process, with AutoML at its core. This tool, along with its source code, has been made available on GitHub (Rafe & Singleton, 2023a), facilitating easy access and collaborative enhancement. Through this research and its outcomes, we encourage further exploration into the integration of user-friendly tools like AutoML in various domains, promoting the adoption of data-driven decision-making in fields where expertise is traditionally siloed.

REFERENCES

- Al-Mistarehi, B. W., Alomari, A. H., Imam, R., and Mashaqba, M. (2022). Using Machine Learning Models to Forecast Severity Level of Traffic Crashes by R Studio and ArcGIS. *Frontiers in Built Environment*, 8, 860805.
- Angarita-Zapata, J. S., Maestre-Gongora, G., and Calderín, J. F. (2021). A bibliometric analysis and benchmark of machine learning and automl in crash severity prediction: The case study of three colombian cities. *Sensors*, 21(24).
- Batista, G. E. A. P. A., Bazzan, A. L. C., and Monard, M. C. (2003). Balancing Training Data for Automated Annotation of Keywords: a Case Study. In *Proceedings of the Second Brazilian Workshop on Bioinformatics*, January.

- Effati, M., and Vahedi Saheli, M. (2022). Examining the influence of rural land uses and accessibility-related factors to estimate pedestrian safety: The use of GIS and machine learning techniques. *International Journal of Transportation Science and Technology*, 11(1), 144–157.
- Ferenczak, N. N., Gutierrez, R. E., and Singleton, P. A. (2022). Shedding light on the pedestrian safety crisis: An analysis across the injury severity spectrum by lighting condition. *Traffic Injury Prevention*, 23(7).
- Goswamy, A., Abdel-Aty, M., and Islam, Z. (2023). Factors affecting injury severity at pedestrian crossing locations with Rectangular RAPID Flashing Beacons (RRFB) using XGBoost and random parameters discrete outcome models. *Accident Analysis & Prevention*, 181, 106937.
- Harris, L., Ahmad, N., Khattak, A., and Chakraborty, S. (2023). Exploring the Effect of Visibility Factors on Vehicle–Pedestrian Crash Injury Severity.
- Infante, P., et al. (2022). Comparison of Statistical and Machine-Learning Models on Road Traffic Accident Severity Classification. *Computers*, 11(5), 80.
- Kang, Y., and Khattak, A. J. (2022). Deep Learning Model for Crash Injury Severity Analysis Using Shapley Additive Explanation Values. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(12), 242–254.
- Kitali, A. E., Kidando, E., Sando, T., Moses, R., and Ozguven, E. E. (2017). Evaluating Aging Pedestrian Crash Severity with Bayesian Complementary Log–Log Model for Improved Prediction Accuracy. *Transportation Research Record*, 2659(1).
- Komol, M. M. R., Hasan, M. M., Elhenawy, M., Yasmin, S., Masoud, M., and Rakotonirainy, A. (2021). Crash severity analysis of vulnerable road users using machine learning. *PLOS ONE*, 16(8), e0255828.
- Lee, D., Guldmann, J. M., and von Rabenau, B. (2023). Impact of Driver's Age and Gender, Built Environment, and Road Conditions on Crash Severity: A Logit Modeling Approach. *International Journal of Environmental Research and Public Health*, 20(3).
- Lundberg, S. M., Erion, G. G., and Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. ArXiv Preprint ArXiv:1802.03888.
- Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 2017–December.
- Ali, M. (2020). PyCaret: An open source, low-code machine learning library in Python. <https://pycaret.org/>.
- NHTSA. (2022). Pedestrian Safety: Prevent Pedestrian Crashes. <https://www.nhtsa.gov/road-safety/pedestrian-safety>.
- Rafe, A., and Singleton, P. A. (2023a). CrashAutoML. <https://github.com/pozapas/CrashAutoML>.
- Rafe, A., and Singleton, P. A. (2023b). Exploring Factors Affecting Pedestrian Crash Severity Using TabNet: A Deep Learning Approach. <http://arxiv.org/abs/2312.00066>.
- Rahman, R., Vahedi Saheli, M., and Singleton, P. A. (2023). Risk Factors for Pedestrian Crashes on Utah State Highway Segments: Results from Parametric and Non-Parametric Approaches Controlling for Pedestrian Exposure. *Transportation Research Board (TRB) 102nd Annual Meeting*.
- Salehian, A., Aghabayk, K., Seyfi, M. A., and Shiawakoti, N. (2023). Comparative analysis of pedestrian crash severity at United Kingdom rural road intersections and Non-Intersections using latent class clustering and ordered probit model. *Accident Analysis and Prevention*, 192.

- Shrinivas, V., Bastien, C., Davies, H., Daneshkhah, A., and Hardwicke, J. (2023). Parameters influencing pedestrian injury and severity – A systematic review and meta-analysis. *Transportation Engineering*, 11, 100158.
- Siddiqui, N. A., Chu, X., and Guttenplan, M. (2006). Crossing Locations, Light Conditions, and Pedestrian Injury Severity. *Transportation Research Record: Journal of the Transportation Research Board*, 1982(1).
- UDPS. (2023). Utah Crash Summary. <https://udps.numetric.net/utah-crash-summary#/>.
- WHO. (2022). Road traffic injuries. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- Yang, L., Aghaabbasi, M., Ali, M., Jan, A., Bouallegue, B., Javed, M. F., and Salem, N. M. (2022). Comparative Analysis of the Optimized KNN, SVM, and Ensemble DT Models Using Bayesian Optimization for Predicting Pedestrian Fatalities: An Advance towards Realizing the Sustainable Safety of Pedestrians. *Sustainability*, 14(17), 10467.

Automated Traffic Safety Assessment Tool Utilizing Monocular 3D Convolutional Neural Network Based Detection Algorithm at Signalized Intersections

Ahmed Mohamed¹; Lizhe Li, Ph.D.²; and Mohamed M. Ahmed, Ph.D., P.E.³

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamea9@mail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: li3le@ucmail.uc.edu

³Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamed.ahmed@uc.edu

ABSTRACT

Traffic conflict analysis has gained significant attention over the last few years. Recent research has made significant progress in the development of detection and tracking systems by utilizing various technologies such as closed-circuit televisions (CCTVs), unmanned aerial vehicles (UAV), and sensor fusion exploiting light detection and ranging (LiDAR). Challenges impeding widespread adoption of artificial intelligence-based traffic conflict analysis include issues with monitoring devices, detection algorithms, and output processing. To address these, this study employs a convolutional neural network (CNN)-based algorithm designed for autonomous vehicles, enhancing detection and tracking of road users at signalized intersections. The algorithm's strength lies in depicting road users with three-dimensional bounding cuboids, streamlining data processing. Post-processing algorithms refine tracking precision, enabling seamless trajectory reconstruction and accurate vehicle state estimation. Tested on video footage, the algorithm excels in high-traffic scenarios, achieving 95.08% precision and 92.81% recall. It identifies diverse conflict instances using four unique traffic conflict indicators.

KEYWORDS: Artificial Intelligence, Computer Vision, Vehicle Trajectory, Vehicle Detection, Signalized Intersections, Conflict Analysis, Traffic Safety

INTRODUCTION

Intersection crashes contribute significantly to road fatalities and injuries, comprising more than 50% of total crashes on different U.S. roadway types (Federal Highway Administration 2023). The complexity of interactions among road users at intersections is a key factor (Kim et al. 2006). Traditionally, crash reports spanning at least three years are used for intersection assessment (Mohanty et al. 2022; Diaz-Corro et al. 2021). Alternative approaches, like analyzing traffic conflicts, expedite assessment processes (Sayed et al. 1994; Sayed and Zein 1999). Coined in 1967 by Perkins and Harris, traffic conflicts represent situations where road users' movements pose collision risks (Perkins et al. 1967). Key conflict indicators include Time-to-Collision (TTC), Post Encroachment Time (PET), and deceleration, with additional indicators derived from them (Johnsson et al. 2018). Although less common, indicators related to TTC were developed, such as time-to-line crossing or time-to-zebra, they remain valuable (Anarkooli et al. 2021). Additionally, conflicting speeds are crucial for understanding conflict-crash relationships, categorized by risk level: critical, average, and minimum severity conflicts (Muhlrud 1993).

This study focuses on evaluating intersection safety using implemented surveillance cameras. It introduces a detection method employing convolutional neural network (CenterNet) to enhance the accuracy of road users' trajectories for safety analysis. The study further develops a spatial analysis method to identify traffic conflicts at signalized intersections. Various post-processing steps are implemented to improve the precision of vehicle tracking, involving error analysis, data reduction, and the reconstruction of vehicle trajectories, including position, speed, and acceleration. The enhanced tracking results are utilized to estimate various traffic indicators, showcasing the effectiveness of the proposed tracking process.

The study is structured as follows: The next section provides a literature review on traffic conflict detection techniques. Section 3 outlines the employed methodology. In section 4, the video data collection, observation process, and algorithm formulations are demonstrated. The extracted outputs are validated and assessed in section 5. Section 6 introduces post-processing algorithms, where outputs undergo analysis and processing to eliminate errors. In section 7, the analysis of traffic conflict instances is presented. Finally, the study concludes with Section 8, presenting a summary and conclusions.

BACKGROUND

Early traffic conflict detection studies relied on manual techniques, where human observers were trained on various conflict types and predefined thresholds (Parker et al. 1989). These manual procedures were categorized as subjective and objective methods (Kraay et al., 2013). Subjective methods involved trained observers investigating conflicts and determining their severity levels. Objective methods, on the other hand, utilized time measurements, with Time-to-Collision (TTC) being a fundamental measure in these approaches. However, manual techniques faced limitations in coverage, cost, and judgment consistency among observation teams (Glauz et al. 1985; Ismail et al. 2009).

Upon the rapid development in computer vision, tracking techniques for vehicles using diverse cameras along roadway sections have emerged (Ismail 2010). Computational algorithms leverage these tracking methodologies to detect traffic conflicts (Hou et al. 2014; Essa and Sayed 2015).

A hybrid tracking method integrating Kalman filtering and the Kanade-Lucas-Tomasi feature tracker was initially developed for feature-based tracking on highway sections (Beymer et al. 1989), allowing the computation of traffic parameters like flow rate, average speed, and average spatial headway for individual lanes (Gazis and Edie 1968). This approach was extended by (Saunier and Sayed 2006) to include vehicle tracking at intersections, with adjustments in the transformation matrix to accommodate intersection-specific geometric features. The system achieved a vehicle detection accuracy of 88.4%, displaying efficiency in handling pedestrians and two-wheeled vehicles. However, challenges persisted, including errors in feature grouping at far distances, issues with camera jitter, over-segmentation of trucks and buses, and instances of over-grouping. To tackle these challenges, (Cavallaro et al. 2005) proposed a hybrid strategy addressing appearance, disappearance, splitting, and partial occlusion problems through the interaction between objects and regions' characteristics. While exhibiting good accuracy, the technique allowed real-time execution of all algorithm components on a standard PC, excluding the region segmentation stage.

In a study by (Veeraraghavan et al. 2003), traffic conflicts at intersections were investigated using video monitoring, employing low-level image-based tracking and a high-level Kalman

filter for position and shape estimation. The tracking process involved determining bounding boxes for moving objects based on timestamps, labels, velocities, and other features. A conflict detection module predicted potential collisions between vehicles by comparing distances between bounding boxes.

(Wu et al. 2020) introduced the "Automated Roadway Conflict Identification System" (ARCIS), utilizing Mask R-CNN for vehicle detection in UAV videos. The system employed a channel and spatial reliability tracking algorithm to track vehicles and generate trajectories, focusing on calculating PET at the pixel level for conflict identification. (Abdel-Aty et al. 2022) used CCTV videos to detect 3D vehicle key points with Mask-RCNN and occlusion-Net algorithms. An integrated system, including ARCIS, enhanced vehicle detection and tracking using UAV images, significantly improving localization precision from the 2D plan view. PET values were effectively used to identify traffic conflicts between vehicles.

(Anisha et al. 2022) proposed an integrated detection system combining CCTV cameras, LiDAR, sensor fusion, trajectory extraction, and safety evaluations. The system demonstrated real-time detection and tracking capabilities, with precision rates of 90.32% for cameras and 97% for LiDAR, while the fusion technique detected 97.38% of vehicles. However, limitations included lower LiDAR recall and exclusion of pedestrians and bicyclists, impacting the assessment of conflicts involving vulnerable road users.

Recent research emphasizes the continuous effort to combine technologies like CCTV cameras, drones, LiDAR, and sensor fusion for enhanced traffic monitoring, vehicle detection, tracking, and safety analysis. The goal is to optimize the accuracy and efficiency of current systems, tackling issues like occlusions and over-segmentation with advanced algorithms and integrated data. However, challenges such as drone altitude restrictions, real-time surveillance limitations, adverse weather conditions, and intricate data alignment impede the widespread adoption of these techniques in safety assessments. This study introduces an innovative tracking approach solely based on surveillance cameras, integrating computer vision and neural network algorithms to mitigate these challenges.

This study presents two primary contributions: (a) the detailed CNN-based algorithm, CenterTrack, for detecting road users, pedestrians, and vehicles using surveillance cameras at signalized intersections. CenterTrack's monocular 3D detection capability enhances detections in traffic congestions by estimating occluded edges, facilitating the transformation process to a top-down view; (b) the paper introduces a comprehensive post-processing approach involving data reduction, trajectory smoothing, and the estimation of vehicle states and traffic indicators. The results demonstrate that implementing this tracking algorithm and post-processing procedure significantly enhances the accuracy of traffic tracking. These accomplishments contribute to a proposed detection framework for traffic conflicts at signalized intersections by employing four conflict indicators.

METHODOLOGY

A comprehensive approach is employed, involving multiple steps. Initially, video footage at a signalized intersection was recorded, scrutinized for instances of traffic conflicts, and meticulously reviewed for integration into the detection and analysis framework. The CenterTrack algorithm was deployed for tracking road users in the specified video sequences. To assess CenterTrack's detection accuracy, YOLOv7 algorithm was utilized. Following this, post-processing techniques were employed to validate the extracted coordinates, convert them from

the image plane to a top-down perspective, and enhance the trajectories of road users. The framework's performance was evaluated through testing on a collection of extracted traffic conflict incidents. Traffic conflict indicators were then computed to pinpoint instances of traffic conflicts. presented framework for traffic conflict detection and analysis.

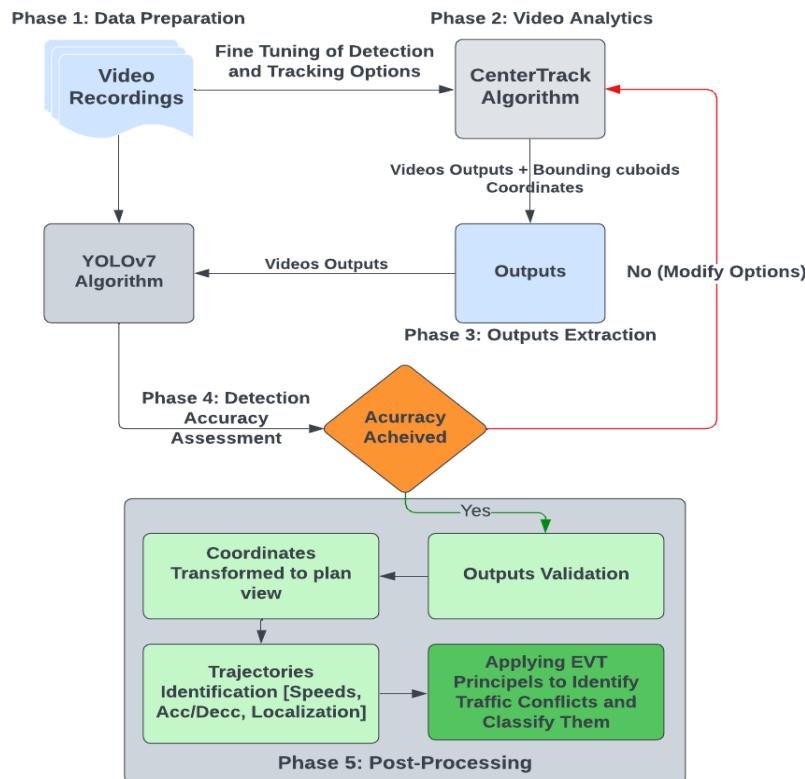


Figure 1: Traffic Conflict Detection and Analysis Framework

DATA PREPARATION

A cumulative duration of 40 hours, encompassing both live and pre-recorded video content, was collected from a CCTV camera positioned at the Town Square intersection in Jackson Hole, Wyoming, a rural signalized intersection. This footage was subjected to meticulous examination by two trained observers, aiming to identify instances of traffic conflicts and violations. To ensure the reliability and precision of the extracted data, two researchers rigorously scrutinized the observations. The CenterTrack algorithm was employed to track road users within the designated video sequences, which were deliberately chosen to include both weekdays and weekends.

The identification of traffic conflicts strictly followed the methodology stated in "Traffic Conflict Techniques for Safety and Operations" by the Federal Highway Administration (Parker and Zegeer 1989). A consolidated list of the identified traffic conflicts was generated, including associated Time-to-Collision thresholds (refer to Table 1). To conduct a comprehensive analysis, a subset of the compiled conflict list was selected, and various traffic conflict indicators were calculated. The video analytics process was applied to a total of 3865 frames, encompassing diverse camera resolutions and capturing all identified types of traffic conflicts.

Table 1: Observed Conflicts Distribution with TTC Thresholds

Conflict type	TTC (Sec.)							Total
	1	1.5	2	2.5	3	3.5	4	
Rear-end	4	9	71	25	49	8	21	187
Angle/side swipe	2	1	22	--	5	1	2	33
Right angle	--	--	1	--	--	--	--	1
Head-on	--	--	1	--	--	--	--	1
Pedestrian involved	2	4	13	4	21	4	10	58
Total	8	14	108	29	75	13	33	280

ALGORITHM FORMULATIONS

The algorithm "CenterTrack" has become a notable focus in research, originating from the paper "Objects as Points" authored by (Zhou et al. 2019). CenterTrack presents an end-to-end approach for monocular 3D object detection and tracking, capitalizing on the synergies between 2D object detection and 3D object localization. By framing 3D object detection as a problem of key point estimation, CenterTrack attains enhanced efficiency and accuracy in localizing and estimating the orientation of 3D objects. This proves especially effective in challenging scenarios involving occlusion and intricate scenes.

The analysis of the video footage from the case study intersection employed the CenterTrack detection and tracking algorithm. In preliminary trials, a 20-second video was analyzed to estimate the processing time, considering the computational power of the device used (Intel(R) Xeon(R) CPU E3-1240 v3 @ 3.40 GHz, 16.0 GB installed RAM, 1 GB NVIDIA Quadro K620). The analysis took 650 seconds using the CPU and significantly less, only 205 seconds, when utilizing the GPU. Given this, there was a need to enhance the computational power by either upgrading the PC or utilizing an external GPU. Consequently, a Google external GPU was employed through the Google Collaboratory environment, a cloud-based platform for machine learning. The analysis time for the 20-second video in Google Collaboratory was 41 seconds, leading to its selection for running the analysis. The video analytics for a single frame using the CenterTrack algorithm is depicted in Figure 2. To verify the detection accuracy of CenterTrack, the YOLOv7 algorithm, another state-of-the-art method, was chosen for comparison (Wang et al., 2022).

DETECTION ASSESSMENT

During this phase, the evaluation of CenterTrack's detection accuracy was essential. Therefore, a selected video set comprising 1,000 frames was subjected to analysis using both the CenterTrack and YOLOv7 algorithms. The chosen scenes encompassed scenarios with traffic congestions and varying weather and surface conditions during both day and night. A comparative analysis was conducted between the two algorithms, and performance indices were calculated, as detailed in Table 2.

The assessment process revealed that CenterTrack achieved an overall precision of 95.08% and a recall of 92.81%. One notable limitation of CenterTrack is its inability to detect road users until their entire form is visible within the video frame. This specific challenge is not present in YOLOv7, making it a more adaptable tool in scenarios where detecting partial visibility of road users is crucial.



Figure 2: Detection and Tracking for Road Users Using CenterTrack

Table 2: Detection Assessment of CenterTrack

Algorithm	Status	No. of Frames	fps	Intersection Over Union (IOU)	Mean average precision (mAP)	Recall	False Positive Rate (FPR)
YOLOv7	(Congested, Clear, Daytime)	350	30	88.06%	91.48%	95.89%	5.10%
	(Normal, Rainy, Nighttime)	300	25	87.45%	88.50%	98.58%	11.73%
	(Congested, Snowy, Daytime)	350	30	95.29%	96.61%	98.57%	3.19%
CenterTrack	(Congested, Clear, Daytime)	350	30	88.73%	94.62%	93.38%	8.27%
	(Normal, Rainy, Nighttime)	300	25	86.95%	97.24%	88.19%	2.76%
	(Congested, Snowy, Daytime)	350	30	89.71%	93.38%	96.87%	7.45%

POST PROCESSING

The extraction of accurate trajectories and orientation for road users is crucial for precise data processing, particularly in determining speed, acceleration, and capturing reliable movement. Despite the tracking process, challenges persist: (a) dependency on raw camera video introduces coordinate system misalignment, hindering accurate determination of vehicle position, speed, and acceleration, and (b) tracked data often contains errors, such as miss-tracking and outliers. Addressing these challenges and obtaining precise vehicle information necessitates careful post-processing techniques to correct errors arising from raw camera coordinates and tracking inaccuracies.

Transformation Homography Calculation

Given that the original camera's viewpoint, located on a first-floor rooftop, is unacceptable for research use, it becomes necessary to establish a common coordinate system. To achieve this, the

image of the same intersection area obtained from Google Earth is cited. In projective geometry a homography transformation is a bijection that maps lines to lines (Baer 2005), therefore by utilizing the Google Earth image as a reference, a consistent coordinate system can be established by matching the same feature points from two images. It is able to calculate the homography between the raw and projected road (see Figure 3).



Figure 3: Feature Points Matching from Raw Camera View (Left, Red Points) to High-Vantage View (Right, Green Points).

Trajectories Reconstruction and Smoothing

The challenges posed by poor video quality and congested traffic conditions give rise to various tracking errors, encompassing (a) duplicated detection due to multiple tracking instances, (b) missed tracking results when vehicles are not tracked in specific frames, (c) incorrect vehicle ID causing segmentation of a single trajectory, and (d) untracked vehicles leading to trajectory gaps when obscured by other objects. Addressing these issues in the data post-processing involves an initial step of associating vehicle positions among tracked points. This entails implementing a straightforward automated process that matches each tracked position to construct a unified trajectory (Coifman et al. 2018).

In the data post-processing stage, the following equation is utilized to minimize the error between current and subsequent frame data, specifically targeting the reduction of discrepancies in the road user's position across frames. This phase addresses tracking errors by automating the association of positions among tracked points, minimizing inter-frame errors, and introducing a threshold for accuracy. The outcome is an improved reliability of trajectory information with the elimination of outliers and gaps in tracking data.

$$\text{Min} \begin{cases} E_x(t) = \sum_{i=1}^4 (x_i(t) - x_i(t-1))^2 \\ E_y(t) = \sum_{i=1}^4 (y_i(t) - y_i(t-1))^2 \end{cases}$$

After addressing errors in the raw trajectories, this research employs a robust locally weighted regression and smoothing procedure (Cleveland 1979). Utilizing a bi-square function

defined by the following equation, this robust regression guards against distortions from deviant points, contributing to the refinement of trajectories.

$$w_i = \begin{cases} (1 - (\frac{r_i}{6 \cdot MAD})^2)^2, & |r_i| < 6 \cdot MAD \\ 0, & |r_i| \geq 6 \cdot MAD \end{cases}$$

Where r_i is the residual of the i^{th} data point produced by the regression smoothing procedure. While, MAD is the median absolute deviation of the residuals.

After the smoothing process, a Savitzky-Golay filter is performed to eliminate residual errors in vehicle trajectories.

TRAFFIC CONFLICTS DETECTION

After the trajectories were extracted and smoothed, four traffic conflict indicators were calculated to identify different types of traffic conflicts. The assessment of the video footage utilized three time-based indicators; TTC, Modified Time-to-Collision (MTTC), and PET and one acceleration-based indicator; Deceleration Rate to Avoid a Crash (DRAC).

$$TTC_t = \frac{x_{L,t} - x_{F,t} - D_L}{v_{F,t} - v_{L,t}}; (v_{F,t} - v_{L,t}) > 0$$

$$MTTC_t = \frac{\Delta v_t \pm \sqrt{\Delta v_t^2 + 2\Delta a_t(x_{L,t} - x_{F,t} - D_L)}}{\Delta a_t}$$

$$PET_t = t_{F,t} - t_{L,t}$$

$$DRAC_{Ft} = \frac{(v_{F,t} - v_{L,t})^2}{2(x_{L,t} - x_{F,t} - D_L)}$$

Where $x_{L,t}$ and $x_{F,t}$ are the positions of leading and following road users at time t , D_L is the length of the leading road user, $v_{L,t}$ and $v_{F,t}$ are the speeds of leading and following road users, respectively. Δv_t , and Δa_t are the difference in speeds and accelerations between the two colliding road users, correspondingly. $t_{L,t}$ and $t_{F,t}$ are the departure time of the offending vehicle, and the arrival time of the conflicting vehicle.

Time-to-Collision (TTC) measures the duration for the non-leading vehicle to reach the conflict area (Sayed and Zein 1999). while Modified Time-to-Collision (MTTC) addresses speed constraints assumed by TTC, allowing for variable speeds (Ozbay et al. 2008). Post-Encroachment Time (PET) signifies the temporal difference between two vehicles approaching the projected conflict zone (Cooper 1984). Deceleration Rate to Avoid a Crash (DRAC) calculates speed and distance extrapolations to assess conflict severity, particularly when hard brakes are not employed (Gettman and Head 2003). The study conducted an analysis of conflicts, excluding cases involving an unidentifiable three-wheeled truck. Traffic conflict indicators were computed using a code for road users, and the extreme value method detected sudden changes indicating conflict occurrences (Wang et al. 2019; Borsos et al. 2020; Zheng and Sayed 2020).

Figure 4 and Figure 5 show the extracted trajectories and code outputs for a side swipe conflict exemplify the video analytics process, showcasing instances where predicted collision points did not align with road users' actions, resulting in unnoticeable PET during detected conflicts.

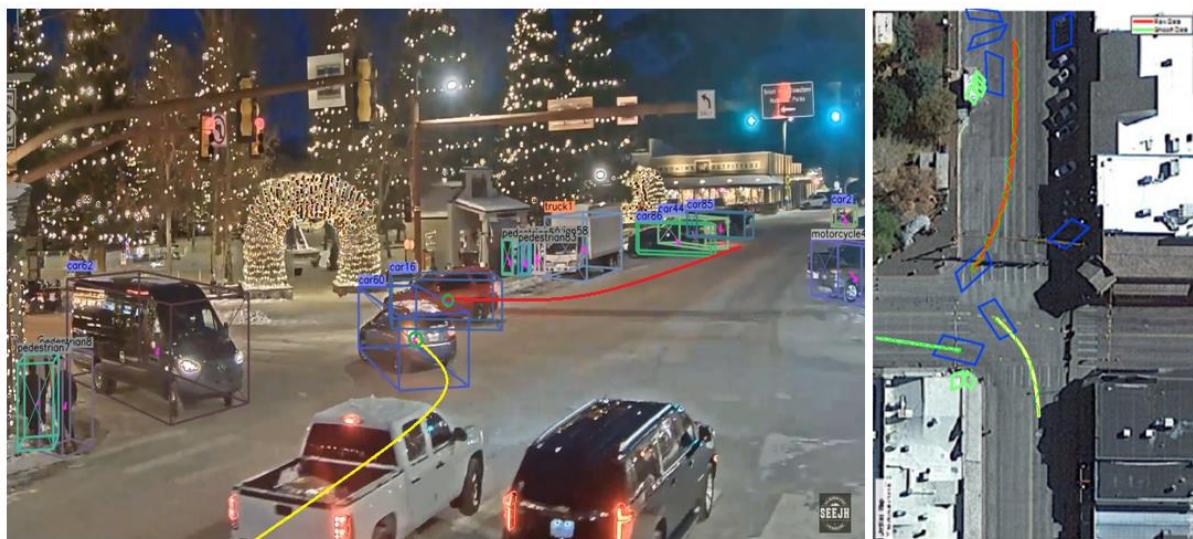


Figure 4: Video Analysis Output for a Conflict utilizing CenterTrack Algorithm

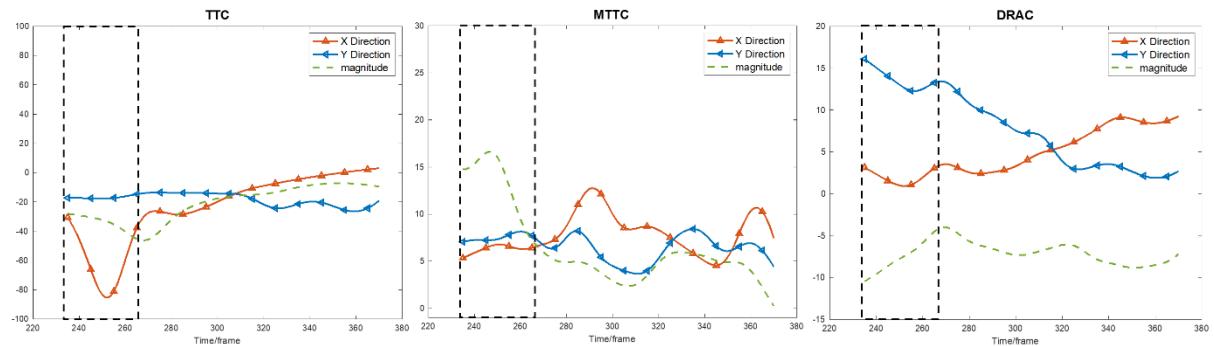


Figure 5: Traffic Conflict Indicators Extracted Values Versus Time Frame.

CONCLUSIONS AND DISCUSSIONS

In conclusion, this research effectively employed the CenterTrack algorithm to analyze traffic conflicts at intersections, achieving an impressive overall precision rate of 95.08% and providing valuable contributions to the field. The comparison with YOLOv7 and the implementation of a transformation homography for plan view visualization underscored the algorithm's efficacy in accurately identifying road users and understanding their spatial distribution. The introduced smoothening algorithm significantly enhanced trajectory precision, allowing for a more accurate analysis and prediction of traffic conflicts by reducing noise and irregularities in the data. The computation of four conflict indicators, coupled with the application of extreme value theory, successfully detected traffic conflict occurrences, contributing to intersection safety evaluations. In summary, this study emphasizes the

CenterTrack algorithm's significance in intersection analysis, offering practical implications for traffic management, intersection design, and overall road safety improvement. However, acknowledging CenterTrack's limitations, including height constraints and category limitations, opens avenues for future research to advance methods for preventing and analyzing traffic conflicts.

ACKNOWLEDGMENTS

This work was sponsored by the Wyoming Department of Transportation (WYDOT). Grant Number: RS04221. The writings were enhanced by using ChatGPT which helped in checking for grammatical errors and paraphrasing sentences in a more concise manner.

AUTHOR CONTRIBUTIONS

The authors confirm the contribution to the paper as follows: study conception and design: Ahmed Mohamed, and Lizhe Li; data preparation and reduction: Ahmed Mohamed and Lizhe Li; analysis and interpretation of results: Ahmed Mohamed and Lizhe Li; draft manuscript preparation: Ahmed Mohamed, Lizhe Li and Mohamed M. Ahmed. All authors reviewed the results and approved the final version of the manuscript. The authors have no conflict of interest to declare.

REFERENCES

- Abdel-Aty, M., Wu, Y., Zheng, O., and Yuan, J. 2022. Using Closed-Circuit Television Cameras to Analyze Traffic Safety at Intersections Based on Vehicle Key Points Detection. *Accident Analysis and Prevention*, Vol. 176.
- Anarkooli, A. J., Persaud, B., Milligan, C., Penner, J., and Saleem, T. 2021. Incorporating Speed in a Traffic Conflict Severity Index to Estimate Left Turn Opposed Crashes at Signalized Intersections. In *Transportation Research Record*, SAGE Publications Ltd, pp. 214–225.
- Anisha, A. M., Abdel-Aty, M., Abdelraouf, A., Islam, Z., and Zheng, O. 2022. Automated Vehicle to Vehicle Conflict Analysis at Signalized Intersections by Camera and LiDAR Sensor Fusion. *Transportation Research Record: Journal of the Transportation Research Board*, p. 036119812211288.
- Baer, R. 2005. *Linear algebra and projective geometry*. Courier Corporation.
- Beymer, D., McLauchlan, P., Coifman, B., and Malik, J. 1989. *A Real-Time Computer Vision System for Measuring Traffic Parameters*.
- Borsos, A., Farah, H., Laureshyn, A., and Hagenzieker, M. 2020. Are Collision and Crossing Course Surrogate Safety Indicators Transferable? A Probability Based Approach Using Extreme Value Theory. *Accident Analysis and Prevention*, Vol. 143.
- Cavallaro, A., Steiger, O., and Ebrahimi, T. 2005. Tracking Video Objects in Cluttered Background. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 4, pp. 575–584.
- Cleveland, W. S. 1979. *Robust Locally Weighted Regression and Smoothing Scatterplots*.
- Coifman, B., Li, L., and Xiao, W. 2018. Resurrecting the Lost Vehicle Trajectories of Treiterer and Myers with New Insights into a Controversial Hysteresis. *Transportation Research Record*, Vol. 2672, No. 20, pp. 25–38.

- Cooper, P. J. 1984. Experience with traffic conflicts in Canada with emphasis on “post encroachment time” techniques. In *International calibration study of traffic conflict techniques* (pp. 75-96). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Diaz-Corro, K. J., Moreno, L. C., Mitra, S., and Hernandez, S. 2021. Assessment of Crash Occurrence Using Historical Crash Data and a Random Effect Negative Binomial Model: A Case Study for a Rural State. In *Transportation Research Record*, SAGE Publications Ltd, pp. 38–52.
- Essa, M., and Sayed, T. 2015. Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts? *Transportation Research Record*, Vol. 2514, pp. 48–57.
- FHWA (Federal Highway Administration), US DOT. 2023. Traffic Safety Facts. URL <https://highways.dot.gov/research/research-programs/safety/intersection-safety>.
- Gazis, D. C., and Edie, L. C. 1968. Traffic Flow Theory. *Proceedings of the IEEE*, Vol. 56, No. 4, pp. 458–471.
- Gettman, D., and Head, L. 2003. Surrogate safety measures from traffic simulation models. *Transportation Research Record*, 1840(1), 104-115.
- Glauz, W. D., Bauer, K. M., and Migletz, D. J. 1985. Expected traffic conflict rates and their use in predicting accidents. *Transportation Research Record*, 1026, 1-12.
- Hou, J., List, G. F., and Guo, X. 2014. New Algorithms for Computing the Time-to-Collision in Freeway Traffic Simulation Models. *Computational Intelligence and Neuroscience*, Vol. 2014,
- Ismail, K., Sayed, T., Saunier, N., and Lim, C. 2009. Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. *Transportation Research Record*, No. 2140, pp. 44–54.
- Ismail, K. A. 2010. *Application of computer vision techniques for automated road safety analysis and traffic data collection* (Doctoral dissertation, University of British Columbia).
- Johnsson, C., Laureshyn, A., and De Ceunynck, T. 2018. In Search of Surrogate Safety Indicators for Vulnerable Road Users: A Review of Surrogate Safety Indicators. *Transport Reviews*, Vol. 38, No. 6, pp. 765–785.
- Kim, D.-G., Washington, S., and Oh, J. 2006. Modeling Crash Types: New Insights into the Effects of Covariates on Crashes at Rural Intersections.
- Kraay, J. H., Van Der Horst, A. R. A., and Oppe, S. 2013. Manual conflict observation technique DOCTOR (Dutch Objective Conflict Technique for Operation and Research).
- Mohanty, M., Panda, R., Gandupalli, S. R., Arya, R. R., and Lenka, S. K. 2022. Factors Propelling Fatalities during Road Crashes: A Detailed Investigation and Modelling of Historical Crash Data with Field Studies. *Heliyon*, Vol. 8, No. 11, p. e11531.
- Muhlrad, N. 1993. Traffic conflict techniques and other forms of behavioural analysis: Application to safety diagnoses. *6th ICTCT workshop*.
- Ozbay, K., Yang, H., Bartin, B., and Mudigonda, S. 2008. Derivation and Validation of New Simulation-Based Surrogate Safety Measure. *Transportation Research Record*, No. 2083, pp. 105–113.
- Parker, M. R., Jr., and Zegeer, C. V. 1989. Traffic conflict techniques for safety and operations: Observers manual. No. FHWA-IP-88-027, NCP 3A9C0093. United States. Federal Highway Administration.
- Perkins, S. R., and Harris, J. I. 1967. *Criteria for traffic conflict characteristics, signalized intersections*. Research Laboratories, General Motors Corporation.
- Saunier, N., and Sayed, T. 2006. A Feature-Based Tracking Algorithm for Vehicles in Intersections. No. 2006, pp. 59.

- Sayed, T., Brown, G., and Navin, F. 1994. "Simulation of traffic conflicts at unsignalized intersections with TSC-Sim." *Accident Analysis & Prevention* 26.5: 593-607.
- Sayed, T., and Zein, S. 1999. Traffic Conflict Standards for Intersections. *Transportation Planning and Technology*, Vol. 22, No. 4, pp. 309–323.
- Veeraraghavan, H., Masoud, O., and Papanikolopoulos, N. P. 2003. Computer Vision Algorithms for Intersection Monitoring. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, pp. 78–89.
- Wang, C., Xu, C., and Dai, Y. 2019. A Crash Prediction Method Based on Bivariate Extreme Value Theory and Video-Based Vehicle Trajectory Data. *Accident Analysis and Prevention*, Vol. 123, pp. 365–373.
- Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. M. 2022. *YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors*.
- Wu, Y., Abdel-Aty, M., Zheng, O., Cai, Q., and Zhang, S. 2020. Automated Safety Diagnosis Based on Unmanned Aerial Vehicle Video and Deep Learning Algorithm. *Transportation Research Record*, Vol. 2674, No. 8, pp. 350–359.
- Zheng, L., and Sayed, T. 2020. A Bivariate Bayesian Hierarchical Extreme Value Model for Traffic Conflict-Based Crash Estimation. *Analytic Methods in Accident Research*, Vol. 25.
- Zhou, X., Wang, D., and Krähenbühl, P. 2019. *Objects as points*.

A Deep Learning Approach for Multiple Weather Conditions Detection Based on In-Vehicle Dash Camera Video

Lizhe Li, Ph.D.¹; and Mohamed M. Ahmed, P.E.²

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: li3le@cumail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: ahmed2mh@ucmail.uc.edu

ABSTRACT

Weather conditions significantly influence traffic safety, with adverse conditions often leading to hazardous driving environments. This study presents the development of a weather detection model based on convolutional neural networks (CNNs), specifically trained on in-vehicle dash camera videos. Using thousands of images sourced from the SHRP2 trajectories level dataset, seven distinct weather categories were identified: clear, light snow, heavy snow, light rain, heavy rain, distant fog, and near fog. For enhanced training efficiency, images were strategically segmented. An innovative approach was employed where, on average, nine images were selected from every minute of the video; these images underwent processing using various pre-trained CNN architecture, such as "AlexNet," "GoogleNet," "ResNet," among others. The predominant detection results then informed the overall weather condition for the given video segment. The model exhibited promising performance, underscoring its viability for real-time in-vehicle weather detection and consequent enhancement of road safety.

INTRODUCTION

Ensuring driver safety is a paramount concern in the development and advancement of automotive technologies, with weather detection playing a crucial role. The Federal Highway Administration (FHWA) reports startling statistics: on average, there are 1,235,145 weather-related crashes each year, accounting for 21% of vehicle crashes. Among these, 5,376 are fatal crashes, representing 16% of crash fatalities. Notably, wet pavement is the most common condition in these incidents, contributing to 860,286 crashes, or 15% of vehicle crashes. Following closely is rain, resulting in 556,151 crashes, which is 10% of vehicle crashes. These figures underscore the significant impact of weather conditions on road safety, influencing visibility, road surface conditions, and vehicle performance. In light of these statistics, the development of effective weather detection systems becomes imperative for enhancing driver safety and making informed decisions under adverse conditions.

The development of weather detection methods has transitioned from reliance on traditional meteorological data to the incorporation of machine learning algorithms. Initially, machine learning techniques, such as decision trees, support vector machines, and simple neural networks, were employed. These methods offered improvements over conventional approaches by enabling the analysis of complex patterns in data. As development of deep learning in recent years, particularly Convolutional Neural Networks (CNNs), are capable of automatically extracting and learning features from raw data for image classification. This is particularly advantageous for weather detection, as CNNs excel in analyzing visual weather imagery. Their ability to learn

hierarchical feature representations makes them exceptionally well-suited for interpreting the nuanced variations in weather conditions captured by in-vehicle cameras and other camera related sensors.

LITERATURE REVIEW

Weather detection has evolved significantly over the years, transitioning from rudimentary meteorological tools to sophisticated computational models. Traditional methods, such as radar and satellite imagery (*Johnson et al., 2005*), have been pivotal in weather forecasting but often fall short in providing real-time, localized data necessary for driver safety.

In the United States, numerous weather detection systems are employed by state Departments of Transportation (DOTs) to alert drivers about abrupt changes in weather conditions. Local weather stations gather meteorological data, which is then disseminated to drivers via radio broadcasts. However, the cost of installation and maintenance is high (*U.S. DOT, 2020*). The advancement of image processing technology in recent decades has spurred researchers to develop more sophisticated models for weather detection. This involves training classifiers with these models and then employing machine learning for accurate detection. Machine learning leverages a wide variety of data sources and utilizes numerous image processing techniques to enhance accuracy and efficiency. *Hautiere et al. (2006)* proposed a method for detecting daytime fog and estimating visibility distances. *Roser and Moosmann (2008)* presented an approach to distinguish multiple weather situations based on the classification of single monocular color images. *Zhang et al. (2015)* extract multiple weather features and combine these features into high dimensional vectors to learn an adaptive classifier. The classifier could detect multiple weather in single images. *Chu et al. (2017)* analyzed the various correlations between weather properties and metadata based on large-scale and rich online image dataset and built a mode to estimate weather information for any given image. *Qian et al. (2016)* developed a detection system using static images from a dashboard camera. The detection process uses the distribution of road pixel locations, then fuses normalized luminance and texture features probabilistically to categorize the segmented road surface. *Khan et al. (2021)* proposed several machine learning classifiers by utilizing histogram of oriented gradient (HOG) and local binary pattern (LBP) as classification parameters to train the weather detection models.

In recent years, deep learning has gained popularity in various research fields, including weather detection. Many researchers now utilize deep learning techniques to train models, owing to their superior efficiency and effectiveness (*Nielsen, 2015*). *Liu et al. (2016)* designed deep learning-based weather forecasting using Deep Neural Network. *Fane et al. (2021)* proposed extreme weather prediction using deep learning. *Essa et al. (2022)* utilized LSTM neural network model for thunderstorm severity prediction. *Kang et al. (2018)* introduced extraction the ROIs of a weather image to increase the efficiency of deep learning models. *Toğuçar et al. (2021)* proposed weather images classification using Spiking Neural Network with GoogLeNet and VGG-16 models. *Khan et al. (2020)* introduced a deep learning architecture featuring five convolutional layers specifically designed for distinguishing between clear and foggy weather conditions. Their comparison with various machine learning algorithms demonstrated the superior efficacy of this deep learning approach.

This study particularly uses Convolutional Neural Networks (CNNs), which are uniquely suited for image processing tasks. Unlike previous discussions that focused on using deep learning to classify one or two special weather categories, this research expands the weather

scope into 7 categories, including clear, frog, rain and snow. The paper first proposed an image segment method to analysis image feature in different weather scenario. Then a deep learning algorithm is employed to train on a vast array of images, creating a model capable of classifying seven common weather categories. CNNs' specialized structure makes them highly effective in interpreting and analyzing image data, significantly enhancing model performance in tasks like weather detection. This model is specifically designed for use with dash camera videos located in vehicles, providing real-time weather detection results to drivers. The aim is to increase safe driving by offering timely and accurate weather updates directly from a vehicle's perspective.

METHDOLOGY

Data Preparation

The image data used in this study were collected from the Second Strategic Highway Research Program Naturalistic Driving Study, which often referred to simply as the " SHRP2 NDS.". It is a research project aimed at collecting real-world driving data for driver behavior in both ordinary and extraordinary situations. Thousands of vehicles were equipped with sensors and cameras to record continuous data on driver actions, vehicle movements, and external conditions, resulting in millions of recorded vehicle miles traveled across the U.S. This unique dataset has been used in many studies to improve the safety of roadways. In this paper, the driver data from 2021 ~ 2022 is utilized for the research.

The primary objective of this paper is to develop a method for detecting the external weather conditions from within a vehicle. Based on the data gathered from the SHRP2 NDS, 7 weather categories have been manually determined: clear, distance frog, near frog, heavy rain, light rain, heavy snow, and light snow. Figure 1 shows example images of all 7 categories.



Figure 1: The weather is manually classified as 7 categories.

In Vehicle Camera Images Processing

Images captured by in-vehicle dash cameras are pivotal for weather classification and should exhibit distinctive features pertinent to various weather conditions. To enhance the accuracy of

our classification, it's advantageous to segment these images into specific parts, each representing a unique aspect relevant to a particular weather category. For instance, a clear weather scenario displayed in Figure 2 (a), the image is divided into sections such as 'sky', 'roadside', and 'road surface'. The 'sky' section, being intrinsically linked to weather, should display a clear, blue sky for sunny conditions. The 'roadside' areas, encompassing the left and right road margins, reflect the environmental context, ideally showing a clean, snow-free landscape. Lastly, the 'road surface' segment focuses on the condition of the road, which, in clear weather, should appear dry. This methodical segmentation ensures a comprehensive analysis of each weather-specific element within the image. Figure 2 (b), shows the case of light snow conditions, the segmented approach is equally informative. The 'sky' displays a grey background, indicative of non-clear weather. The 'roadside' shows snow patches, while the 'road surface' often appears wet, reflecting the quick melting of light snow.

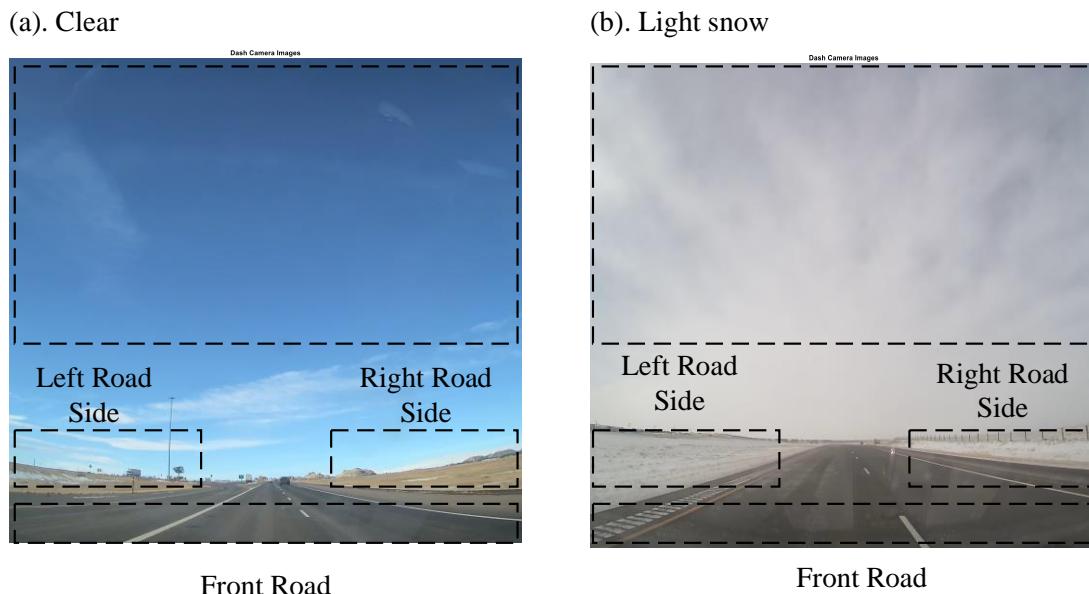


Figure 2: Images features for (a) clear weather image; (b) light snow weather image.

CNN NETWORK PROCESS

Convolutional Neural Network Introduction

Deep learning has pioneered advances in image classification, and central to this advancement is the Convolutional Neural Network (CNN). CNN is inspired by the visual cortex of the human brain, which contains specialized neurons responsive to overlapping regions in the visual field. The architecture of a CNN is comprised of several key components: (1), Convolutional layers, where spatial features are detected, using filters that slide over the input image. (2), Pooling layers, which reduce the spatial dimensions, concentrating on the most salient features. (3), Fully Connected Layers, which compiles the features to produce the final classification outcome. Through these layers, the network learns to detect patterns, starting from the simplest visual cues and gradually progressing to more intricate details.

CNN for image analysis is to learn features directly from image data without human intervention. By autonomously determining the best features for classification, CNN is able to

provide a decent image recognition solution. The self-learning ability ensures that the networks can adapt and refine their ability of determining the best images feature, leading to highly accurate and efficient recognition performances for image classification.

Pre-Trained Convolution Neural Network

In recent decades, the architecture of Convolutional Neural Networks (CNNs) for image classification has undergone significant evolution. Contemporary models exhibit increased depth and complexity in their design. This study employs six prominent CNN architectures for training on road-surface images. These include AlexNet, SqueezeNet, GoogleNet, ResNet18, ResNet50, and ShuffleNet. Each network is characterized by its unique layer configurations and specific requirements for training image size. Details of each network model are displayed in Table 1

Table 1: Network Requirements

Network	Depth	Image Input Size (Square, Pixels)
AlexNet	8	227
SqueezeNet	18	227
GoogleNet	22	224
ResNet18	18	224
ResNet50	50	224
ShuffleNet	50	224

Network Training Processing

This section outlines the training procedures for six CNN models. Each model was trained using a dataset comprising 2,500 images per category. The dataset was split such that 80% was allocated for training and model validation, while the remaining 20% was reserved for validation purposes. Key training parameters included: a maximum of 3 epochs, a batch size of 300, and the use of Stochastic Gradient Descent (SGD) as the optimizer. The initial learning rate was set at 0.001, with a learning rate drop factor of 0.1 and a drop period of 8 epochs. The subsequent figures provide a detailed illustration of the training outcomes for each network, specifically focusing on road surface condition detection.

AlexNet

The training progression of AlexNet is depicted in Figure 3. The overall training process is efficient, with the training completing in just 1 minute and 25 seconds on a single Nvidia GPU setup. The training culminates with a validation accuracy of 97.32%, indicating that the models are highly accurate in classifying road surface conditions.

SqueezeNet

Figure 4 illustrates the training trajectory of another SqueezeNet. The final accuracy is closed to 91.64%. This is slightly lower than AlexNet. Notably, the training time for this model is

remarkably short, with the process finishing in just 44 seconds, suggesting that SqueezeNet is particularly efficient in terms of computational resource utilization.

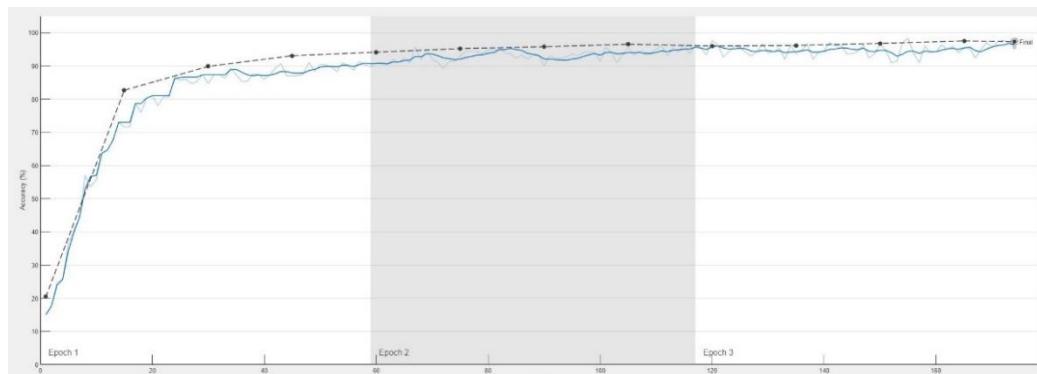


Figure 3: Training Process for AlexNet

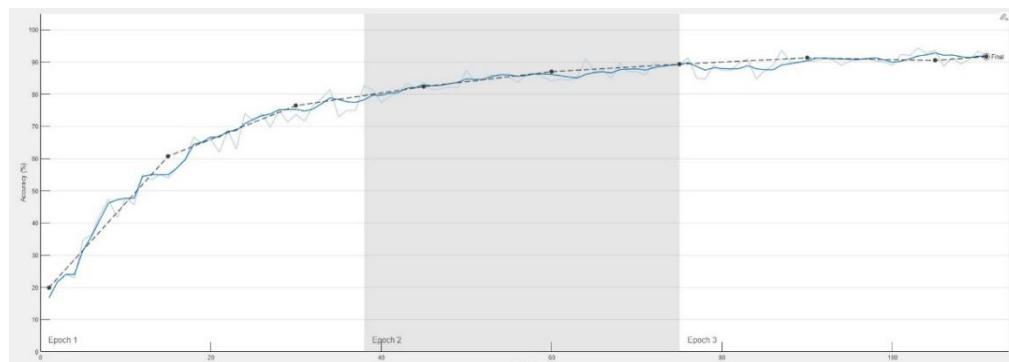


Figure 4: Training Process for SqueezeNet

GoogleNet

Figure 5 illustrates the training progression of GoogleNet. Its final accuracy remains at 92.57%, a bit lower than AlexNet. The training process is completed in 1 minute and 48 seconds, proving it is an efficiency model.

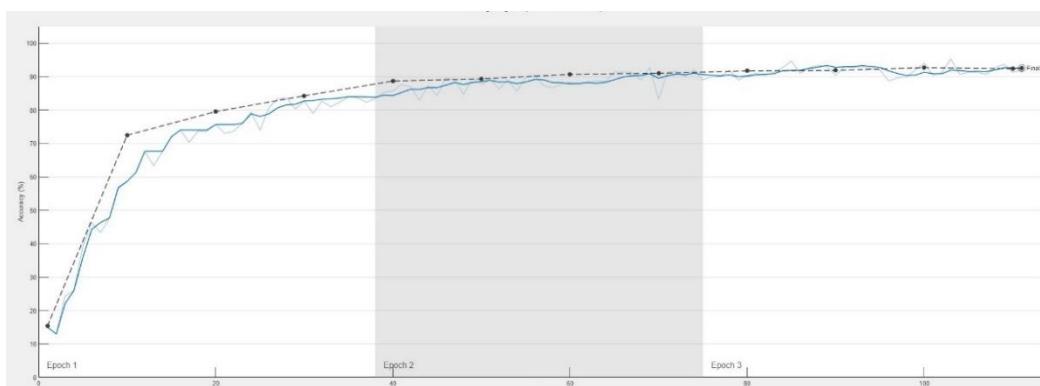


Figure 5: Training Process for GoogleNet

ResNet-18

The training progression for the ResNet18 model is presented in Figure 6. The final validation accuracy achieved is 98.1%, which is the highest of all. The training process completed its training in 1 minute and 54 seconds, demonstrating a balance between computational efficiency and high-performance metrics.

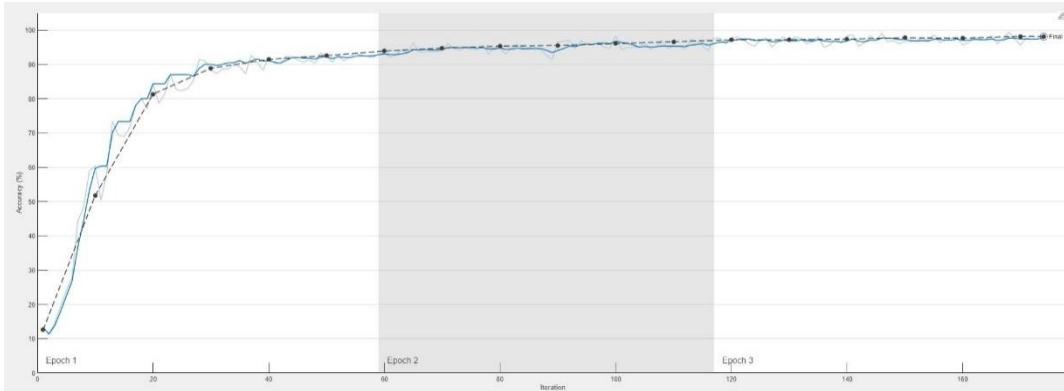


Figure 6: Training Process for ResNet-18

ResNet-50

Figure 7 shows the training progress graph for the ResNet50. The final accuracy follows a similar trajectory with 95.04%, a bit lower than ResNet-18. However, A significant observation from the training session is the duration, which, at 21 minutes and 50 seconds, is longer than the previous models. This extended training time can be attributed to the more complex architecture of ResNet50.

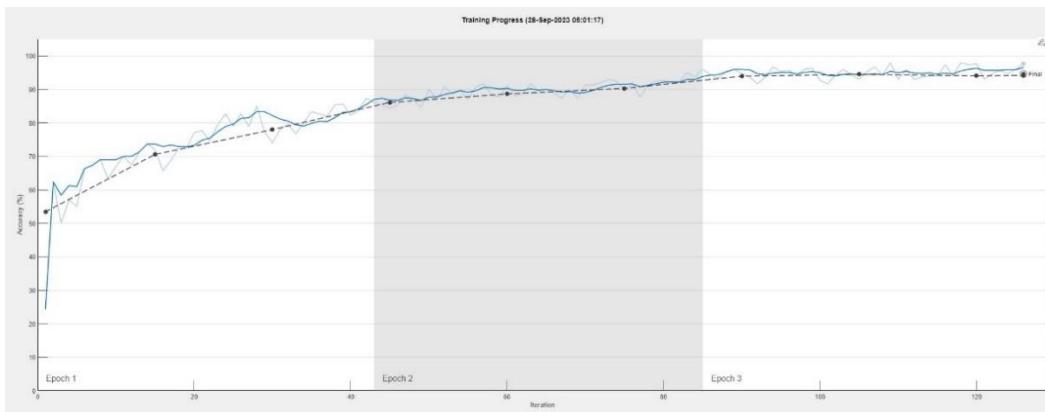


Figure 7: Training Process for ResNet-50

ShuffleNet

Figure 7 presents the training process of the ShuffleNet. The accuracy is impressive to 97.61%. The training duration is only 1 minute and 56 seconds. The swift training, combined

with the high validation accuracy makes ShuffleNet an ideal choice for real-time weather analysis.

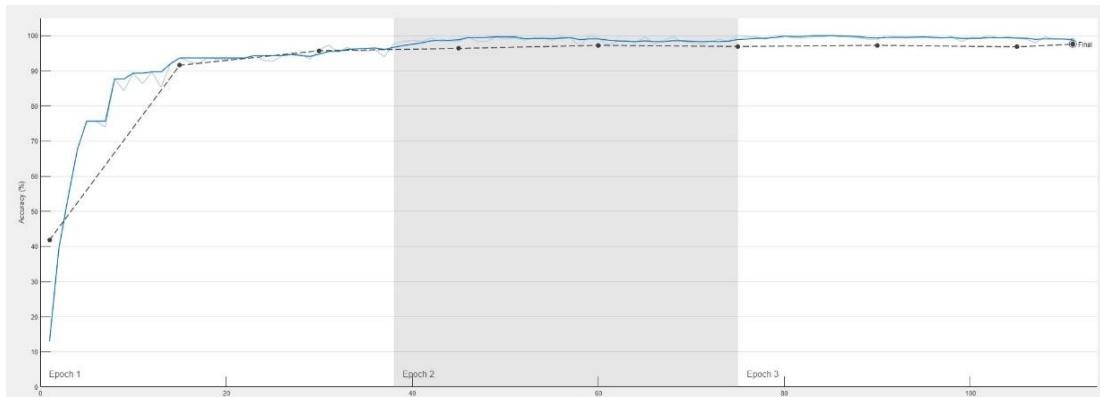


Figure 8: Training Process for ShuffleNet

VALIDATION RESULT

Table 2 displays the performance of six CNN models—AlexNet, SqueezeNet, GoogleNet, ResNet18, ShuffleNet, and ResNet50—in diverse weather conditions, revealing a balance between classification accuracy and training efficiency. AlexNet excels in Heavy Snow with 99.40% accuracy, boasting an overall 99.20% accuracy with a rapid training time of 1 minute 25 seconds. SqueezeNet, despite its efficient 44-second training, shows variable performance, with the lowest accuracy in Distance Fog at 85.60%. GoogleNet, consistently accurate across conditions, has a notable performance in Heavy Rain and Heavy Snow at 96.00%, and an overall 94.20% accuracy, trained in 1 minute 48 seconds. ResNet18 and ShuffleNet perform exceptionally in Clear and Heavy Snow, with ResNet18 achieving 100% in Light Rain; both have training times just under 2 minutes and accuracies above 97%. In contrast, ResNet50, with a deeper architecture, takes longer to train (21 minutes 50 seconds) but maintains a high 94.80% accuracy. The table highlights that while models like AlexNet, ResNet18, and ShuffleNet combine high accuracy with quick training, SqueezeNet sacrifices some accuracy for speed, and ResNet50's depth results in longer training without compromising performance, though its extended training time limits its suitability for weather image detection training.

Table 2: Details of detection Accuracy for each Weather Category

Network	Image Category							Test Accuracy (%)	Training time
	Clear	Distance Fog	Heavy Rain	Heavy Snow	Light Rain	Light Snow	Near Fog		
AlexNet	98.6	94.30%	93.80%	99.40%	98.30%	97.60%	99.2%	99.20%	1min 25s
SqueezeNet	93.60%	85.60%	88.00%	97.60%	95.50%	89.90%	85.2%	85.20%	44s
GoogleNet	96.80%	92.60%	96.00%	96.30%	92.30%	92.30%	94.2%	94.20%	1min 48s
ResNet18	98.80%	94.80%	96.50%	98.80%	100%	98.20%	97.8%	97.80%	1min 54s
ResNet50	97.80%	93.10%	95.20%	97.20%	94.10%	95.40%	94.8%	94.80%	21min 50s
ShuffleNet	98.80%	95.80%	96.10%	99.40%	96.60%	99.00%	97.6%	97.60%	1min 56s

WEATHER DETECTION MODEL TESTING FOR REAL VIDEO

The SHRP2 videos utilized in this research were captured using dash cameras inside vehicles, recording the front view. Each video is recorded in 4K resolution at 30 fps and has a duration of 1 minute. Given the belief that weather conditions remain consistent within a 1-minute span, select an average of 9 frames from each video as representative samples. Figure 9 illustrates examples of the averaged 9 frames chosen to represent these 1-minute videos. In this figure, we can clearly observe the vehicle's driving behavior.

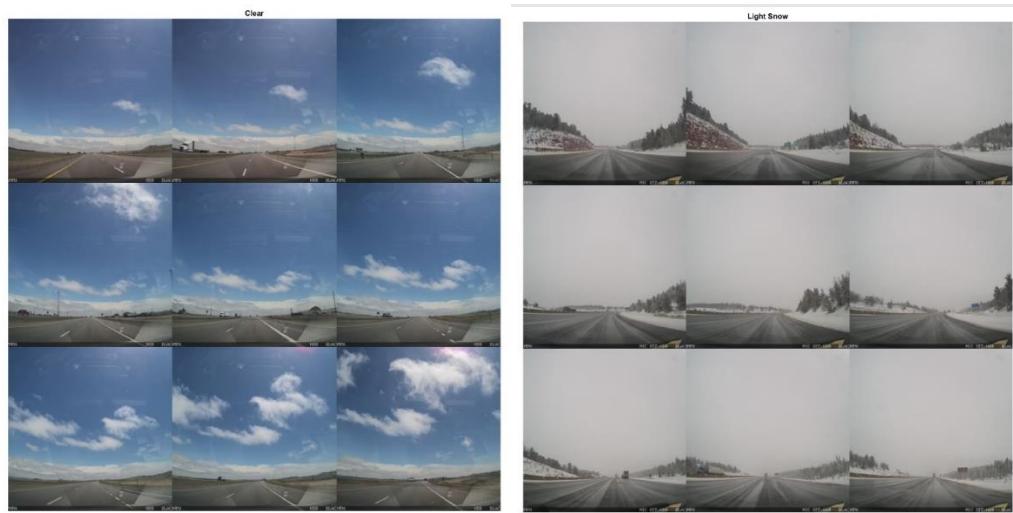


Figure 9: Example of averaged 9 frames for weather detection, from 1 min video recorded from dash camera. Left detection result is clear, right detection result is light snow.

CONCLUSION

This paper presents a thorough analysis of CNNs in weather classification tasks related to road-surface analysis under diverse weather conditions. It reveals that CNNs not only achieve high accuracy but also maintain consistency across various environmental scenarios. Evaluated architectures—AlexNet, SqueezeNet, GoogleNet, ResNet18, ResNet50, and ShuffleNet—demonstrate the adaptability and efficiency of CNNs over traditional machine learning methods, which rely on hand-engineered features. CNNs, through deep learning, identify intricate patterns, resulting in more robust, generalizable models. The study shows high-test accuracies consistently above 85%, with several models exceeding 97%. Despite their complexity, these CNNs exhibit relatively short training times, with SqueezeNet being exceptionally efficient, while ResNet50, although requiring a longer training period, is recommended for complex classifications due to its high accuracy and depth of learning.

These findings support the superiority of CNNs over traditional methods for weather condition classification in road-surface images. CNNs' ability to autonomously learn from images, coupled with their training agility and high-dimensional data handling, sets new benchmarks in computer vision. The study strongly recommends CNNs for training models in complex weather detection, offering weather predictions based on dash camera videos to enhance driver safety.

REFERENCES

- FHWA (Federal Highway Administration). (2020). How Do Weather Events Impact Roads? - FHWA Road Weather Management. <http://www.ops.fhwa.dot.gov/weather/q1_roadimpact.htm> (Sep. 20, 2019).
- Khan, M. N., and Ahmed, M. M. (2020). Trajectory-level fog detection based on in-vehicle video camera with TensorFlow deep learning utilizing SHRP2 naturalistic driving data. *Accident Analysis & Prevention*, 142, 105521.
- US DOT. (2020). RITA | ITS | Costs: Road Weather Information System (RWIS) Environmental Sensor Station (ESS) <<https://www.itskrs.its.dot.gov/its/benecost.nsf/>> (Apr. 21, 2020).
- Hautiere, N., Tarel, J. P., Lavenant, J., and Aubert, D. (2006). Automatic fog detection and estimation of visibility distance through use of an onboard camera. *Machine vision and applications*, 17(1), 8-20.
- Roser, M., and Moosmann, F. (2008, June). Classification of weather situations on single color images. In *2008 IEEE intelligent vehicles symposium* (pp. 798-803). IEEE.
- Zhang, Z., and Ma, H. (2015, September). Multi-class weather classification on single images. In *2015 IEEE International Conference on Image Processing (ICIP)* (pp. 4396-4400). IEEE.
- Chu, W. T., Zheng, X. Y., and Ding, D. S. (2017). Camera as weather sensor: Estimating weather information from single images. *Journal of Visual Communication and Image Representation*, 46, 233-249.
- Qian, Y., Almazan, E. J., and Elder, J. H. (2016, September). Evaluating features and classifiers for road weather condition analysis. In *2016 IEEE International Conference on Image Processing (ICIP)* (pp. 4403-4407). IEEE.
- Khan, M. N., Das, A., Ahmed, M. M., and Wulff, S. S. (2021). Multilevel weather detection based on images: A machine learning approach with histogram of oriented gradient and local binary pattern-based features. *Journal of Intelligent Transportation Systems*, 25(5), 513-532.
- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25, pp. 15-24). San Francisco, CA, USA: Determination press.
- Liu, Y., Racah, E., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., and Collins, W. (2016). Application of deep convolutional neural networks for detecting extreme weather in climate datasets. arXiv preprint arXiv:1605.01156.
- Fang, W., Xue, Q., Shen, L., and Sheng, V. S. (2021). Survey on the application of deep learning in extreme weather prediction. *Atmosphere*, 12(6), 661.
- Essa, Y., Hunt, H. G., Gijben, M., and Ajoodha, R. (2022). Deep Learning Prediction of Thunderstorm Severity Using Remote Sensing Weather Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 4004-4013.
- Kang, L. W., Chou, K. L., and Fu, R. H. (2018, December). Deep learning-based weather image recognition. In *2018 International Symposium on Computer, Consumer and Control (IS3C)* (pp. 384-387). IEEE.
- Toğaçar, M., Ergen, B., and Cömert, Z. (2021). Detection of weather images by using spiking neural networks of deep learning models. *Neural Computing and Applications*, 33(11), 6147-6159.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., and Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint arXiv:1602.07360.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Zhang, X., Zhou, X., Lin, M., and Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6848-6856).

Vehicle-to-Vehicle Conflicts Detection at Signalized Intersections: Leveraging Multi-Cameras Utilizing Convolutional Neural Network-Based Pose Estimation Algorithm

Ahmed Mohamed¹; Lizhe Li, Ph.D.²; and Mohamed M. Ahmed, Ph.D., P.E.³

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamea9@mail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: li3le@ucmail.uc.edu

³Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamed.ahmed@uc.edu

ABSTRACT

Over the past decade, researchers have shown significant interest in traffic monitoring using various devices such as surveillance cameras, unmanned aerial vehicles (UAV), loop detectors, and sensor fusion exploiting light detection and ranging (LiDAR). These techniques, widely applied for vehicle detection and tracking, have established robust safety assessment methodologies. However, nationwide deployment is hindered by constraints in monitoring devices, detection algorithms, and output processing. This study addresses these limitations by introducing an innovative approach that integrates three surveillance cameras with meticulously designed framework. Leveraging two distinct convolutional neural network (CNN)-based detection algorithms, this method accurately detects and tracks vehicles, providing reliable outputs that are further analyzed, smoothed, and integrated for a comprehensive reconstruction of traffic scenes at signalized intersections. The proposed framework enables the extraction of crucial traffic flow features, facilitating the identification of traffic conflicts instances. Through the synergistic integration of surveillance cameras and CNN-based algorithms, this research contributes significantly to advancing traffic safety practices, offering an efficient and accurate tool for safety assessment.

KEYWORDS: Vehicle Detection, Pose Estimation, Computer Vision, CNN, Signalized Intersections, Conflict Analysis, Traffic Safety

INTRODUCTION

One of the most crucial topics in the traffic safety field is the assessment of safety at intersections. Intersection crashes in the United States (US) are contributing with more than 50% of total crashes on different roadway types (*Federal Highway Administration 2023*). This higher crash risk at intersections could be interpreted by the complexity of movements between different types of road users (*Kim et al. 2006*). In addition to the complex nature of road users' movements, the contributing factors to crash occurrence encompass inadequate geometric design of intersection approaches, improper road user perceptions during traffic encounters, and deficient signal phase plans that fail to accommodate traffic flow demands through the day. Moreover, the prevalence of crash types can vary depending on whether the road environment is urban or rural, as the contributing factors differ in their nature. Accordingly, providing an assessment tool that could accurately monitor intersections, and provides an estimation of the safety performance within intersection areas is one of the active research topics.

This study is devoted to analyzing and assessing the safety at signalized intersections by integrating the field of views (FOVs) of the implemented surveillance cameras at intersections regardless of their types or output resolutions. Two types of CNN-based algorithms with two different forms of detection, bounding boxes and key points, are used to detect vehicles out of each camera output. Consequently, a detection framework is proposed based on the algorithms' results from the three cameras to detect the vehicles that fall within the integrated FOVs for better safety assessment.

The study is organized as follows: The next section provides a literature review on vehicle detection techniques. Section 3 outlines the followed methodology throughout the study, data preparation, algorithm formulations, and outputs analysis and data assessment. Section 4 demonstrates the traffic conflicts detection process. Finally, Section 5 presents the summary and conclusions.

VEHICLE DETECTION

The field of computer vision has seen rapid and continuous evolution, leading to the development of tracking techniques for vehicles using various devices (Ismail 2010). Computational algorithms have been utilized to track road users across different roadway sections (Hou et al 2014; Essa and Sayed 2015). Commonly developed object tracking techniques include region-based tracking (Stauffer et al. 2000; Veeraraghavan et al. 2003; Magee 2004; Maurin et al 2005), 3D model-based tracking (Dahlkamp et al. 2004), contour-based tracking (Koller et al. 1994), and feature-based tracking (Cavallaro et al. 2005; Saunier et al. 2006).

An early study conducted by (Beymer et al. 1997) implemented a hybrid method, merging Kalman filtering and Kanade-Lucas-Tomasi feature tracking, to monitor vehicle characteristics on freeways. This method, as detailed by (Gazis and Edie 1968), computed traffic parameters such as flow rate and average speed. (Saunier et al. 2006) later extended this methodology to include vehicle tracking at intersections, addressing challenges like feature grouping errors, camera jitter, and over-segmentation through a hybrid strategy proposed by (Cavallaro et al. 2005). Another approach, explored by (Veeraraghavan et al. 2003), utilized low-level image-based tracking and high-level Kalman filtering for traffic conflict analysis at intersections, building upon the detection module described by (Stauffer and Grimson 1999).

In a different vein, (Zheng et al. 2023) employed drones and the CitySim dataset for vehicle detection and safety analysis. Building on this groundwork, (Wu et al. 2020) introduced "Automated Roadway Conflict Identification System" (ARCIS), which utilized Mask R-CNN to enhance vehicle detection precision. (Abdel-Aty et al. 2022) utilized CCTV footage to detect vehicle-to-vehicle conflicts, integrating Mask R-CNN and the Occlusion-Net algorithm for improved detection accuracy.

(Anisha et al. 2022) proposed an integrated detection system combining CCTV and LiDAR technology, achieving high precision rates. Despite these advancements, challenges such as restricted drone height and real-time surveillance duration persist. Additionally, issues related to data fusion processes and high implementation costs hinder the widespread adoption of these techniques in safety assessment processes.

Our research work aims to fill the gaps by presenting a detailed framework that integrates the output video footage from the mounted cameras at signalized intersections, then employing CNN-based algorithms to detect vehicles key points. Consequently, a comprehensive post

processing technique is established to project the outputs coordinates to the top-down view, integrates the extracted key points from each camera simultaneously to reconstruct vehicle polygons on plan view, extract their trajectories, and analyze traffic conflict instances accurately.

METHODOLOGY

A detailed framework is organized, starting with recording video footage from mounted cameras at the site. Three cameras' recordings were synchronized, covering weekdays and weekends evenly across morning and evening peak hours. An intersection between cameras' FOVs was established to observe coverage. Two CNN-based algorithms, YOLOv7 (You Only Look Once) and OpenPifPaf (Open Pose In Full Pose Articulation Framework), were used for vehicle detection, providing bounding boxes and detecting 24 key points per vehicle, respectively. Refinement of regions of interest through multiple trials ensured accurate detection results. Post-processing transformed extracted coordinates to the plan view and integrated them across cameras. Intersected FOVs restored missing vehicle key points from single camera, employing Time-To-Collision (TTC) for identifying traffic conflict instances. The proposed framework for multi-camera traffic conflict identification is depicted in Figure 1.

Data Preparation

A cumulative duration of 72 hours, including real-time and archived data, was collected from surveillance cameras at the Town Square intersection in Jackson Hole, Wyoming, a rural signalized intersection. Trained observers meticulously examined the footage to identify traffic conflicts and violations, with two researchers ensuring data accuracy and consistency. The selection process aimed to include both weekdays and weekends, following the methodology outlined in (Parker and Zegeer 1989). Details such as road user movements, fault attribution, and prevailing weather were documented for each conflict, as per (Parker and Zegeer 1989). A comprehensive list of observed traffic conflicts, excluding pedestrian-related incidents, was compiled, along with associated Time-To-Collision values, detailed in Table 1. Subsequently, a subset of this list underwent analysis to compute comprehensive traffic conflict indicators, spanning 3605 frames from the video footage and accommodating various camera resolutions.

Table 1: Observed Conflict Distributions with TTC Thresholds

Conflict type	TTC (Sec.)							Total
	1	1.5	2	2.5	3	3.5	4	
Rear-end	4	8	68	21	44	7	20	172
Angle/side swipe	3	2	22	--	5	2	2	36
Right angle	--	--	1	--	--	--	--	1
Head-on	--	--	1	--	--	--	--	1
Total	7	10	92	21	49	9	22	210

Figure 2 depicts the arrangement of the three cameras, outlining their respective coverage areas and identified obstructed areas. The upper left image in the image sequences provides the positions of the three cameras and the FOV for each camera. While, the lower image

demonstrates the intersected FOVs. Noteworthy, the highlighted green areas are the covered areas by two or more cameras and the other regions covered by only one camera. The image columns on the right-hand side demonstrate the perspective image of each camera.

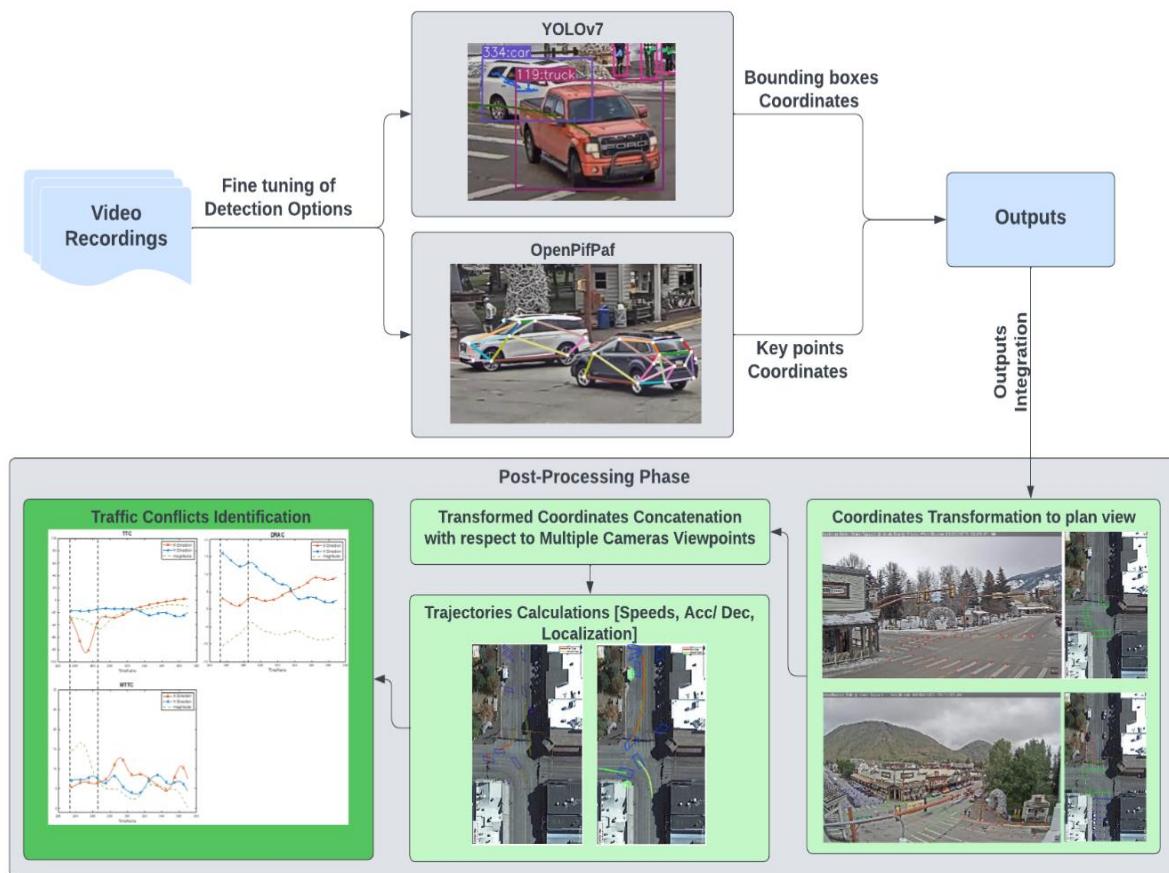


Figure 1: Traffic Conflicts Identification Framework

Detection Algorithms

This study integrates two distinct algorithms for vehicle detection, YOLOv7 and OpenPifPaf, to enhance accuracy and capture detailed information. YOLOv7 generates a bounding box for each identified vehicle. It has been trained with the COCO dataset, covering 80 object categories, utilizing 8 NVIDIA V100 GPUs over a training period of approximately 36 hours (Wang et al. 2022). Notably, its real-time performance evaluation demonstrated the capability to achieve speeds of up to 60 frames per second on a single GPU. A comparison between different YOLO versions verified that YOLOv7 was the most convenient algorithm to be utilized in the study of traffic conflicts analysis. Figure 3 shows the comparison between YOLOv7 performance and other YOLO versions used in real-time detection algorithms.

OpenPifPaf originally introduced by Sven Kreiss et al. (Kreiss et al. 2019) to present a new bottom-up methodology for multiple detections of human poses. Accordingly, the study was followed by several plugins' implementation for the algorithm on different classes (e.g., vehicles, and animals) (Kreiss et al. 2021). For the vehicles key points detection, ApolloCar3D dataset

was utilized for the training process including 5,277 driving images and involving more than 60,000 cars in their captions with a resolution of 4K.

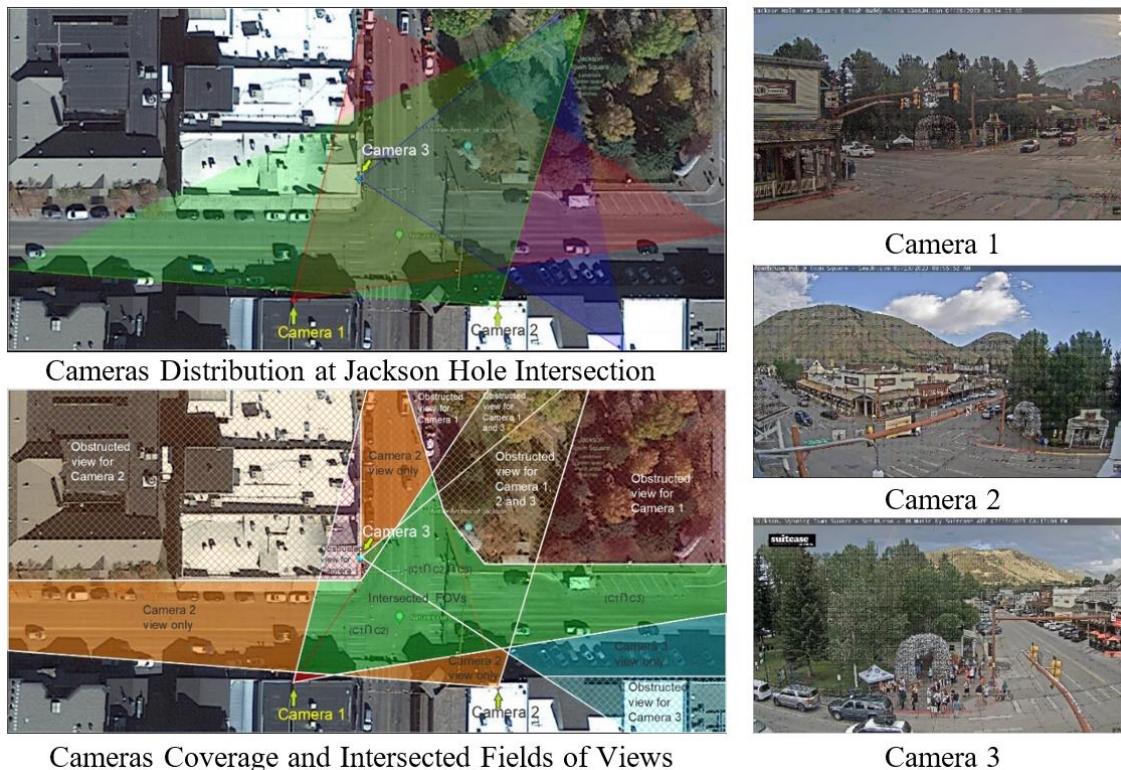


Figure 2: Cameras Positions and Views at the Site of Interest

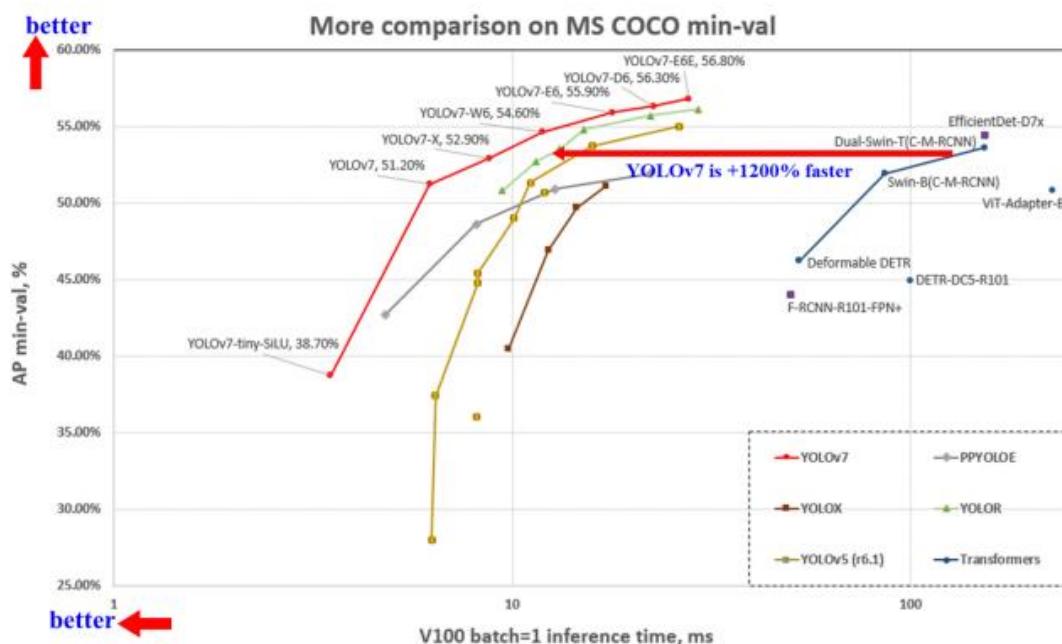


Figure 3: Comparison YOLOv7 to Other Real-time Object Detectors (Wang et al. 2022)

In this study, vehicle key points detection is accomplished using the OpenPifPaf algorithm with the "predict" command and a specific flag ("--checkpoint=shufflenetv2k16-apollo-24") representing vehicles with 24 key points. The research focuses on integrating multiple cameras using a top-down view, emphasizing key points located at or near the ground level for effective representation. Figure 4 illustrates the vehicle's structure and the distribution of 24 key points with corresponding IDs, providing a visual understanding of the spatial arrangement crucial to the study's methodology.

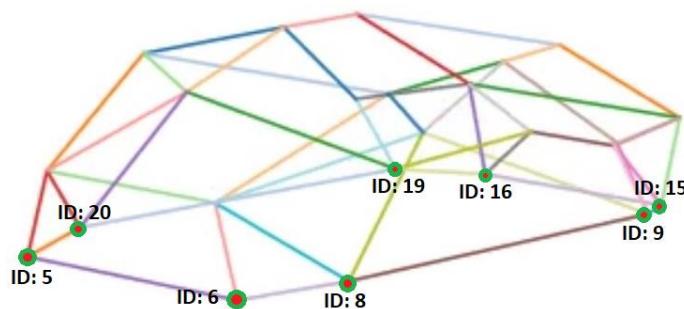


Figure 4: Vehicle Key Points Identification At/Near to Ground Level

Data Analysis and Detection Assessment

Detection Framework

The proposed detection framework is illustrated in Figure 5. The black SUV is detected from the three cameras. Consequently, the video analytics process was performed utilizing YOLOv7 and OpenPifPaf algorithms. The outputs were extracted, calibrated, and transformed to the plan view. The common area between the projected bounding boxes is determined and the projected key points are integrated. The undetected points are assumed based on the average vehicle dimension with respect to the detected points.

Detection Accuracy Assessment

A video set of 1,050 frames was analyzed using each camera solely and with the integrated framework. The selected set included traffic congestions, and hazardous weather and surface conditions at day and night times. The assessment process yielded to find that the proposed framework achieved 98.2% in overall precision with a false detection rate of 2.35% and average Intersect Over Union (IOU) of 0.955, Table 2.

Table 2: Detection Assessment of Proposed Framework

	(Intersection Over Union (IOU)) / No. of Detected Vehicles		
	(Normal, Clear, Daytime)	(Congested, Rainy, Nighttime)	(Congested, Snowy, Daytime)
Camera 1	(0.374)	5845	(0.447)
Camera 2	(0.519)	8155	(0.460)
Camera 3	(0.299)	4690	(0.303)
Framework	(0.939)	14735	(0.967)
Ground Truth	(1.000)	15680	(1.000)
		8785	(0.427)
		9030	(0.474)
		5950	(0.302)
		19005	(0.960)
		19635	(1.000)
			18235
			19005

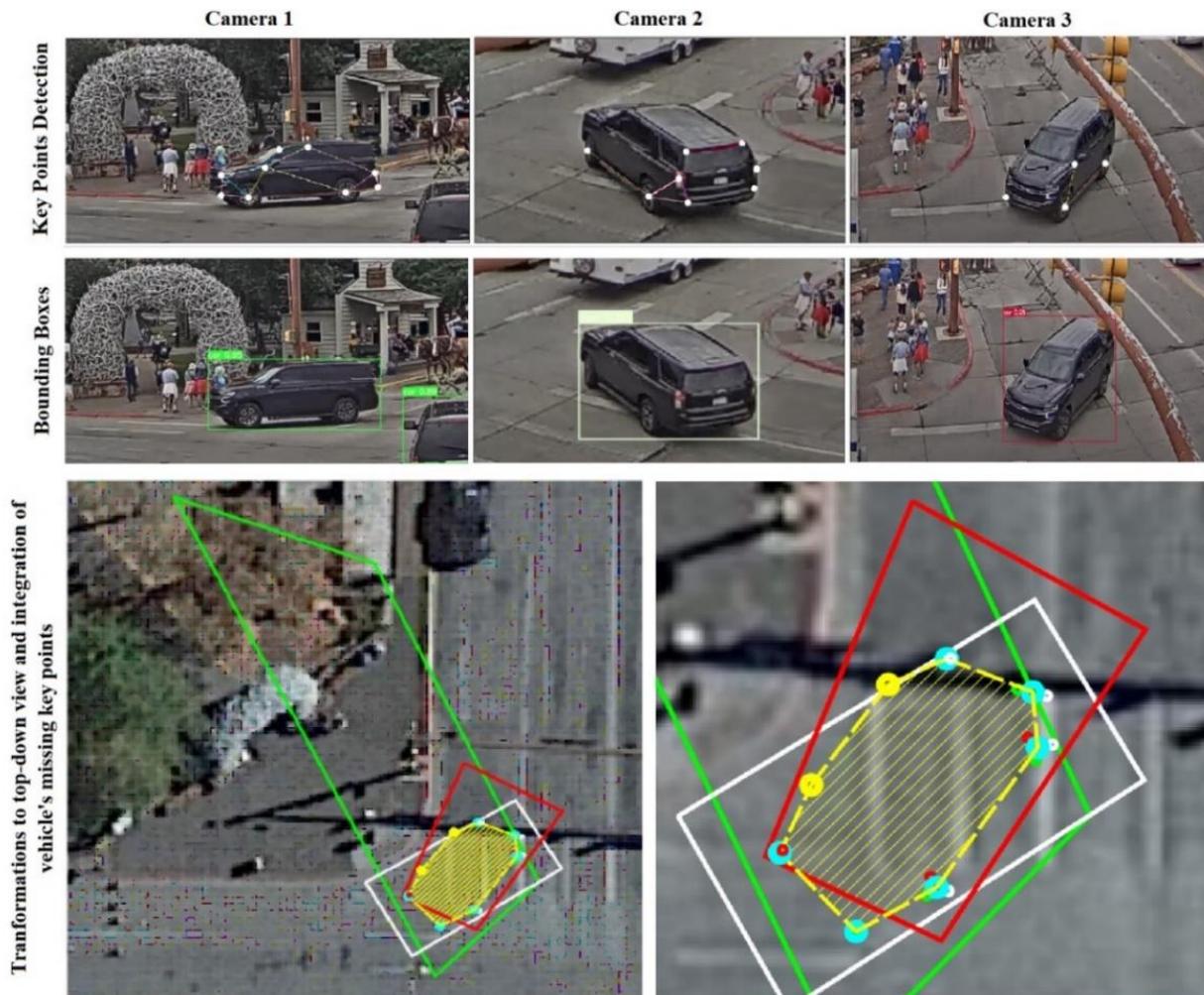


Figure 5: Detection Framework

Vehicle Key Points Projection and Trajectories Construction

The key points were projected onto a new reference road surface, which is sourced from Google Earth. Figure 6 illustrates the projection of key points of a single frame onto the new reference plane. By identifying the front (red points) and rear points (green points), the positions of the majority of the vehicles are determined.

The key points were projected, then, an identification process for each vehicle through multiple frames was performed by setting a minimum threshold between the same key point through each two successive frames.

$$\text{Min Distance}(t) = \sum_i \sqrt{(x_i(t) - x_i(t-1))^2 + (y_i(t) - y_i(t-1))^2}$$

Afterward, the raw trajectories containing errors are constructed. To further refine the trajectories, a smoothing filter, Savitzky-Golay, was utilized to overcome mis-tracking that produces data gaps and incorrect tracking that results in outliers.

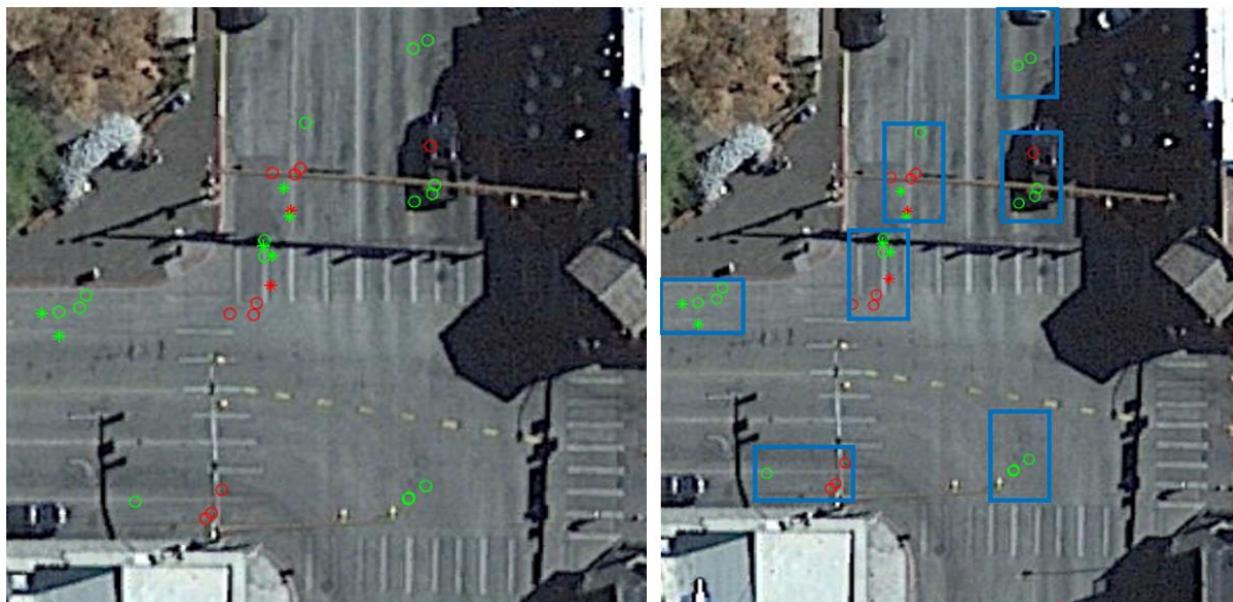


Figure 6: Transformed Video Image and The Projected Vehicles Coordinates in Google Earth Image. The Clustered Points Within the Blue Square Signify That They Belong to A Single Vehicle.

TRAFFIC CONFLICTS DETECTION

After the trajectories were extracted and smoothed, TTC was computed as the conflict indicator to identify the occurrence of traffic conflict instances. TTC denotes the duration it takes for the non-leading vehicle to reach the conflict area (Sayed and Zein 1999).

The analysis of the proposed conflicts was conducted. Figure 7 shows the extracted trajectories from both camera and top-down views as an example for the video analytics process of a side swipe conflict. While the drawn curves depict the code outputs for the same conflict showcasing the identification of traffic conflict occurrence. The analysis of a rear-end conflict is illustrated in Figure 7. The conflict occurred between two left turning vehicles when the motion of the leading one (Vehicle 1) was obstructed by the congestion and the following vehicle (vehicle 10) applied a sudden brake to avoid a rear end collision. The TTC curve showed a sudden drop under 2 second value while the speeds of both vehicles were drafted and the trajectories were extracted.

CONCLUSIONS AND DISCUSSIONS

In conclusion, this research paper successfully developed vehicle detection framework by leveraging the strengths of integrating multiple surveillance cameras that are mounted at signalized intersections and employing CNN-based algorithm to detect vehicles key points, presenting valuable insights and contributions to the field. While, the restored vehicles key points' coordinates from multiple cameras were clustered, integrated, and trajectories calculations were held. The smoothing algorithm played a vital role in improving the precision and dependability of the identified trajectories, speeds, and accelerations. Through noise reduction and elimination of irregularities in the data, this algorithm allowed for a more precise

depiction of road user behavior, resulting in more accurate analysis and prediction of traffic conflicts. By employing TTC with a predefined threshold the occurrence of traffic conflicts was successfully detected, providing valuable data for evaluating intersection safety and identifying potential areas of concern.

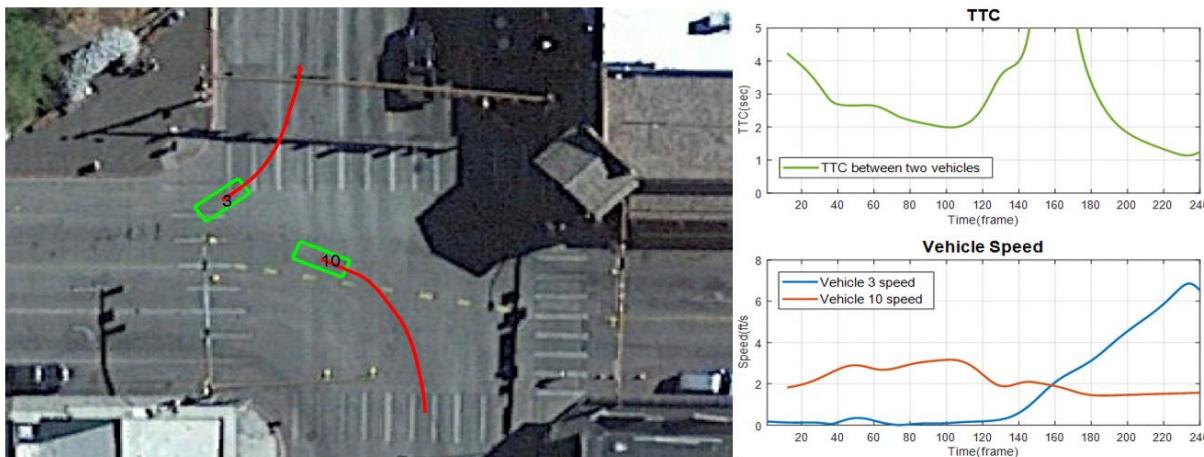


Figure 7: Example of a Sideswipe/Angle Traffic Conflict with TTC Identifying Its Occurrence

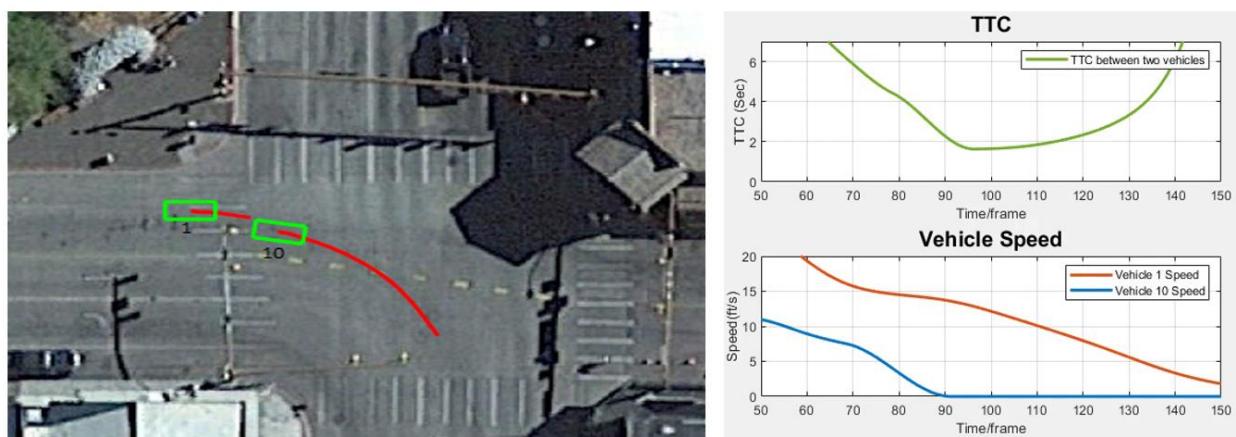


Figure 8: Example of a Rear-End Traffic Conflict with TTC Identifying Its Occurrence

Through the application of the proposed framework for vehicle detection and a subsequent comparison with the outputs from each camera solely, an overall precision rate of 98.2% was achieved, highlighting the effectiveness of the approach in accurately identifying vehicles. Furthermore, the visualization technique provided a comprehensive understanding of the spatial distribution of vehicles at the intersection, enabling a more insightful analysis of potential conflicts and their occurrence patterns.

In summary, this study highlights the importance of employing multiple cameras to detect vehicles at signalized intersections as an accurate and cost-effective mean for safety assessment. Through the integration of two CNN-based detection algorithms and the application of a smoothening technique for the extracted vehicles trajectories, our understanding of traffic

patterns and intersection safety is broadened. The results of this research hold practical implications for enhancing traffic management strategies, improving intersection design, and promoting road safety as a whole. Future research can leverage these findings to advance the field and create more advanced methods for analyzing and preventing traffic conflicts.

ACKNOWLEDGMENTS

This work was sponsored by the Wyoming Department of Transportation (WYDOT). Grant Number: RS04221. The writings were enhanced by using ChatGPT which helped in checking for grammatical errors and paraphrasing sentences in a more concise manner.

AUTHOR CONTRIBUTIONS

The authors confirm the contribution to the paper as follows: study conception and design: Ahmed Mohamed, and Lizhe Li; data preparation and reduction: Ahmed Mohamed and Lizhe Li; analysis and interpretation of results: Ahmed Mohamed and Lizhe Li; draft manuscript preparation: Ahmed Mohamed, Lizhe Li and Mohamed M. Ahmed. All authors reviewed the results and approved the final version of the manuscript. The authors have no conflict of interest to declare.

REFERENCES

- Abdel-Aty, M., Y. Wu, O. Zheng, and J. Yuan. (2022). "Using Closed-Circuit Television Cameras to Analyze Traffic Safety at Intersections Based on Vehicle Key Points Detection". *Accident Analysis and Prevention*, Vol. 176.
- Anisha, A. M., M. Abdel-Aty, A. Abdelraouf, Z. Islam, and O. Zheng. (2022). "Automated Vehicle to Vehicle Conflict Analysis at Signalized Intersections by Camera and LiDAR Sensor Fusion". *Transportation Research Record: Journal of the Transportation Research Board*.
- Beymer, D., P. McLauchlan, B. Coifman, and J. Malik. (1997). "A real-time computer vision system for measuring traffic parameters," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, PR, USA, pp. 495-501.
- Cavallaro, A., O. Steiger, and T. Ebrahimi. (2005). "Tracking Video Objects in Cluttered Background". *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 4, pp. 575–584.
- Dahlkamp, H., A. E. C. Pece, A. Ottlik, and H.-H. Nagel. (2004). "Differential Analysis of Two Model-Based Vehicle Tracking Approaches". *DAGM-Symposium*.
- Essa, M., and T. Sayed. (2015). "Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured 3 Conflicts?". *Transportation Research Record*, Vol. 2514, pp. 48–57. 4.
- Federal Highway Administration, US DOT. (2023). Traffic Safety Facts. Retrieved from [https://highways.dot.gov/research/research-programs/safety/intersection-safety].
- Gazis, D. C., and L. C. Edie. (1968). "Traffic Flow Theory". *Proceedings of the IEEE*, Vol. 56, No. 4, pp. 458–471.
- Hou, J., G. F. List, and X. Guo. (2014). "New Algorithms for Computing the Time-to-Collision in Freeway 6 Traffic Simulation Models". *Computational Intelligence and Neuroscience*, Vol. 2014, 7.

- Ismail, K. A. (2010). *Application of computer vision techniques for automated road safety analysis 1 and traffic data collection* (Doctoral dissertation, University of British Columbia).
- Kim, D.-G., S. Washington, and J. Oh. (2006). "Modeling Crash Types: New Insights into the Effects of 4 Covariates on Crashes at Rural Intersections". *Journal Of Transportation Engineering*, pp. 282-292.
- Koller, D., J. Weber, and J. Malik. (1994). *Multiple Car Tracking with Occlusion Reasoning*. Lecture Notes in Computer Science, vol 800. Springer, Berlin, Heidelberg.
- Kreiss, S., L. Bertoni, and A. Alahi. (2019). "PifPaf: Composite Fields for Human Pose Estimation". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 11977-11986.
- Kreiss, S., L. Bertoni, and A. Alahi. (2021). "OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association". *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 13498-13511.
- Magee, D. R. (2004). "Tracking Multiple Vehicles Using Foreground, Background and Motion Models". *Image and Vision Computing*, Vol. 22, No. 2, pp. 143–155.
- Maurin, B., O. Masoud, and N. P. Papanikolopoulos. (2004). "Tracking All Traffic: Computer Vision 13 Algorithms for Monitoring Vehicles Individuals, and Crowds". *IEEE Robotics and Automation Magazine*, Vol. 12, No. 1, pp. 29–36.
- Parker, M. R., Jr., and C. V. Zegeer. (1989). "Traffic conflict techniques for safety and operations: Observer's manual". No. FHWA-IP-88-027, NCP 3A9C0093. US Federal Highway 33 Administration.
- Saunier, N., and T. Sayed. (2006). "A Feature-Based Tracking Algorithm for Vehicles in Intersections". *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, Quebec, Canada, 2006, pp. 59-59.
- Sayed, T., and S. Zein. (1999). "Traffic Conflict Standards for Intersections". *Transportation Planning and Technology*, Vol. 22, No. 4, pp. 309–323.
- Stauffer, C., and W. E. L. Grimson. (1999). "Adaptive Background Mixture Models for Real-Time Tracking". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 246–252.
- Stauffer, C., and W. E. L. Grimson. (2000). "Learning Patterns of Activity Using Real-Time Tracking". *IEEE 16 Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 747–757. 17.
- Veeraraghavan, H., O. Masoud, and N. P. Papanikolopoulos. (2003). "Computer Vision Algorithms for 19 Intersection Monitoring". *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, pp. 78–89.
- Wang, C.-Y., A. Bochkovskiy, and H.-Y. M. Liao. (2022). "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors".
- Wu, Y., M. Abdel-Aty, O. Zheng, Q. Cai, and S. Zhang. (2020). "Automated Safety Diagnosis Based on Unmanned Aerial Vehicle Video and Deep Learning Algorithm". *Transportation Research Record*, Vol. 2674, No. 8, pp. 350–359.
- Zheng, O., M. Abdel-Aty, L. Yue, A. Abdelraouf, Z. Wang, and N. Mahmoud. (2023). "CitySim: A Drone- Based Vehicle Trajectory Dataset for Safety Oriented Research and Digital Twins". *Transportation Research Record*, 0(0).

Pedestrian Tracking at Signalized Intersections Leveraging Multi-Camera Field of Views Using Covolutional Neural Network-Based Pose Estimation Algorithm

Ahmed Mohamed¹ and Mohamed M. Ahmed, Ph.D., P.E.²

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamea9@mail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamed.ahmed@uc.edu

ABSTRACT

Pedestrian detection poses significant challenges for traffic safety researchers, given diverse motion patterns, clothing colors, partial occlusions, and varying positions concerning detection devices. Surveillance cameras, light detection and ranging (LiDAR), and microwaves beams have been utilized in pedestrians' detection, leading to robust safety assessment methodologies. Despite their efficacy, widespread deployment encounters constraints. This study introduces an optimized approach by combining surveillance cameras with a distinctive convolutional neural network (CNN)-based pose estimation algorithm for precise pedestrian detection at signalized intersections. A geometrical spatial proximity method, grounded in linear and curvilinear perspectives, restores pedestrian joint coordinates from the image plane to a top-down view. These coordinates are clustered and integrated to construct pedestrian trajectories. The proposed framework serves as an efficient and accurate tool for assessing pedestrian motions, addressing challenges in widespread deployment and contributing significantly to advancing traffic safety practices at signalized intersections.

KEYWORDS: Pedestrian Detection, Pose Estimation, Computer Vision, CNN, Signalized Intersections, Traffic Safety

INTRODUCTION

The assessment of pedestrian safety, particularly at intersections, is a critical focus in traffic safety research. In the United States (US), intersection-related crashes accounting for more than half of all recorded traffic incidents across diverse roadway types (Federal Highway Administration 2023). The elevated risk at intersections is attributed to complex interactions among various road user categories, with pedestrians playing a central role (Kim et al. 2006). Contributing factors to these crashes include inadequate geometric designs for intersection approaches, challenges faced by pedestrians in perceiving and navigating traffic, and signal plans struggling to adapt to dynamic pedestrian flow. Furthermore, crash patterns vary significantly between urban and rural settings due to differences in contributing factors nature. A detailed analysis of recent crash data highlights a notable increase in incidents involving vulnerable road users, placing pedestrians at the forefront of intersection-related incidents (Cicchino 2022).

In light of pedestrians involved crashes statistics, there is an increasing urgency to develop assessment tools for seamless intersection monitoring, efficient pedestrian detection, and comprehensive safety evaluation. This persistent research frontier stands as a critical and ongoing endeavor, driven by traffic safety researchers, with the ultimate aim of enhancing pedestrian safety at intersections.

This study is devoted to detect, analyze, and monitors the pedestrians' trajectories at signalized intersections by integrating the field of views (FOVs) of surveillance cameras regardless of their types. A CNN-based key points detection algorithm is employed to detect pedestrians. Also, a spatial analysis method is developed to transform the detected key points coordinates to the top-down view by utilizing linear and curvilinear perspective principles. Subsequently, a post processing technique was developed to integrate and cluster the restored coordinates. As a result, pedestrians' trajectories are extracted, smoothed, and their features are calculated.

The study is structured as follows: the next section provides a literature review on pedestrian detection techniques. Section 3 outlines the employed methodology through the study. Sections 4 and 5 demonstrate video data collection and algorithm formulations, respectively. In Section 6, an outputs spatial modification technique is presented. Section 7 presents the assessment process of the proposed framework. The representation of pedestrians' trajectories based on the transformed coordinates is demonstrated in section 8. Lastly, Section 9 presents the summary and conclusions of the study.

PEDESTRIANS DETECTION

Based on the ongoing progression in computer vision, robotics and autonomous vehicles fields, the development in detection of vulnerable road users is gaining a huge momentum. Various tracking techniques for road users by utilizing different types of devices have been developed (Ismail 2010). Moreover, several computational algorithms have been employed to track road users at different roadway sections (Hou et al 2014; Essa and Sayed 2015), region-based tracking (Stauffer et al. 2000; Veeraraghavan et al. 2003; Magee 2004; Maurin et al 2005), 3D model-based tracking (Dahlkamp et. al. 2004), contour-based tracking (Koller et al. 1994), and feature-based tracking (Cavallaro et al. 2005; Saunier et al. 2006) are considered the common objects tracking techniques that have been developed.

(Enzweiler and Gavrila 2010) provided a comprehensive overview of both methodological and experimental aspects in pedestrian detection. The study included a survey covering essential elements of pedestrian detection systems and their underlying models, along with experimental analysis sections. The utilized dataset encompassed numerous training samples and a 27-minute test sequence featuring over 20,000 images with annotated pedestrian positions. Evaluation considered both general and vehicle-specific settings. Results revealed the superiority of histogram-oriented gradient (HOG) and linear support vector machine (linSVM) models in high-resolution, slow processing scenarios, while AdaBoost cascades performed better in low-resolutions and near real-time contexts. (Dollár et al. 2012) evaluated sixteen pedestrian detectors on the Caltech pedestrian dataset, comprising 350,000 labeled bounding boxes and 250,000 frames. Evaluation metrics covered features, learning rates, and detection details. General observations highlighted that 80% of pedestrians with a size between 30 to 80 pixels were missed, and nearly all pedestrians went undetected when their size was under 30 pixels and partial occlusion reached 35%. A vision-based system was developed by (Alahi et al. 2014) to identify pedestrians in front of vehicles using low-resolution cameras, showcasing real-time operation with minimal memory usage. The system utilized a cascade of compact binary strings for appearance modeling and matched pedestrians across cameras. While (Zhang et al. 2017) introduced a dataset built on Cityscapes, CityPersons, that enabled the evaluation of Faster R-CNN. The study demonstrated the model's effectiveness on both Caltech and CityPersons

datasets, showcasing improved performance in challenging scenarios like heavy occlusion and small scale. Extensive literature reviews in (Antonio and Romero 2018; Brunetti et al. 2018) examined algorithms for pedestrian detection using images from video surveillance or standard cameras. The studies covered a spectrum of algorithms, revealing the efficacy of deep learning-based architectures in enhancing system performance. A recent study by (Alfred Daniel et al. 2023) introduced a sensor fusion method for LiDAR-camera fusion, improving pedestrian detection accuracy within a range of 10 to 30 meters. The proposed model demonstrated effectiveness in precise detection and identification.

These ongoing studies aim to leverage surveillance cameras, LiDARs, microwaves, and sensor fusion for pedestrian monitoring. While advancements address challenges, limitations in accuracy, real-time surveillance, adverse weather conditions, and high implementation costs hinder widespread adoption in safety assessments.

Our research work aims to fill the gaps by presenting a detailed framework that integrates the output video footages from the mounted cameras at signalized intersections, then, employing CNN-based algorithm to detect pedestrians. Consequently, a comprehensive technique is established to project the outputs coordinates to the top-down view by employing linear and curvilinear perspectives principles. Then, integrates the extracted key points from each camera simultaneously to calculate pedestrians' trajectories, and reconstruct the traffic scene.

METHODOLOGY

A comprehensive methodology was employed, comprising distinct steps to enhance the accuracy of pedestrians' detection at signalized intersections. Initially, simultaneous video footages were recorded from strategically positioned cameras, ensuring uniform frames for analysis. Recordings spanned weekdays and weekends, capturing morning and evening peak hours. Subsequently, a meticulous analysis phase utilized the YOLOv7-human-pose-estimation (You Only Look Once) algorithm, detecting 17 key points for each person. Trials on camera-specific regions of interest refined settings for accurate detection. Simultaneously, a comprehensive technique based on linear and curvilinear perspectives established a unified grid for each crosswalk. This grid facilitated the transformation of pedestrians' coordinates from each camera to an accurate plan view. Coordinated projection, clustering, and integration of extracted coordinates enabled the determination of each moving pedestrian's center point on the crosswalks. Finally, a tracking process extracted pedestrians' trajectories and calculated walking speeds, presenting a robust framework for multiple-camera-based pedestrian detection, as depicted in Figure 1.

DATA PREPARATION

A total of 40 hours was compiled, comprising real-time and pre-recorded video footage from three surveillance cameras strategically placed at the Town Square intersection in Jackson Hole, Wyoming. The dataset aimed to capture diverse scenarios, including various weather conditions, daytime and nighttime, and morning and evening peak hours. For focused analysis, a subset of videos was thoughtfully chosen, adhering to specific guidelines: a) ensuring matching start and end timestamps for uniform frame counts across all three cameras, b) including scenarios with typical and congested pedestrian traffic on crosswalks, and c) encompassing different weather conditions to assess detection accuracy across varied scenarios.

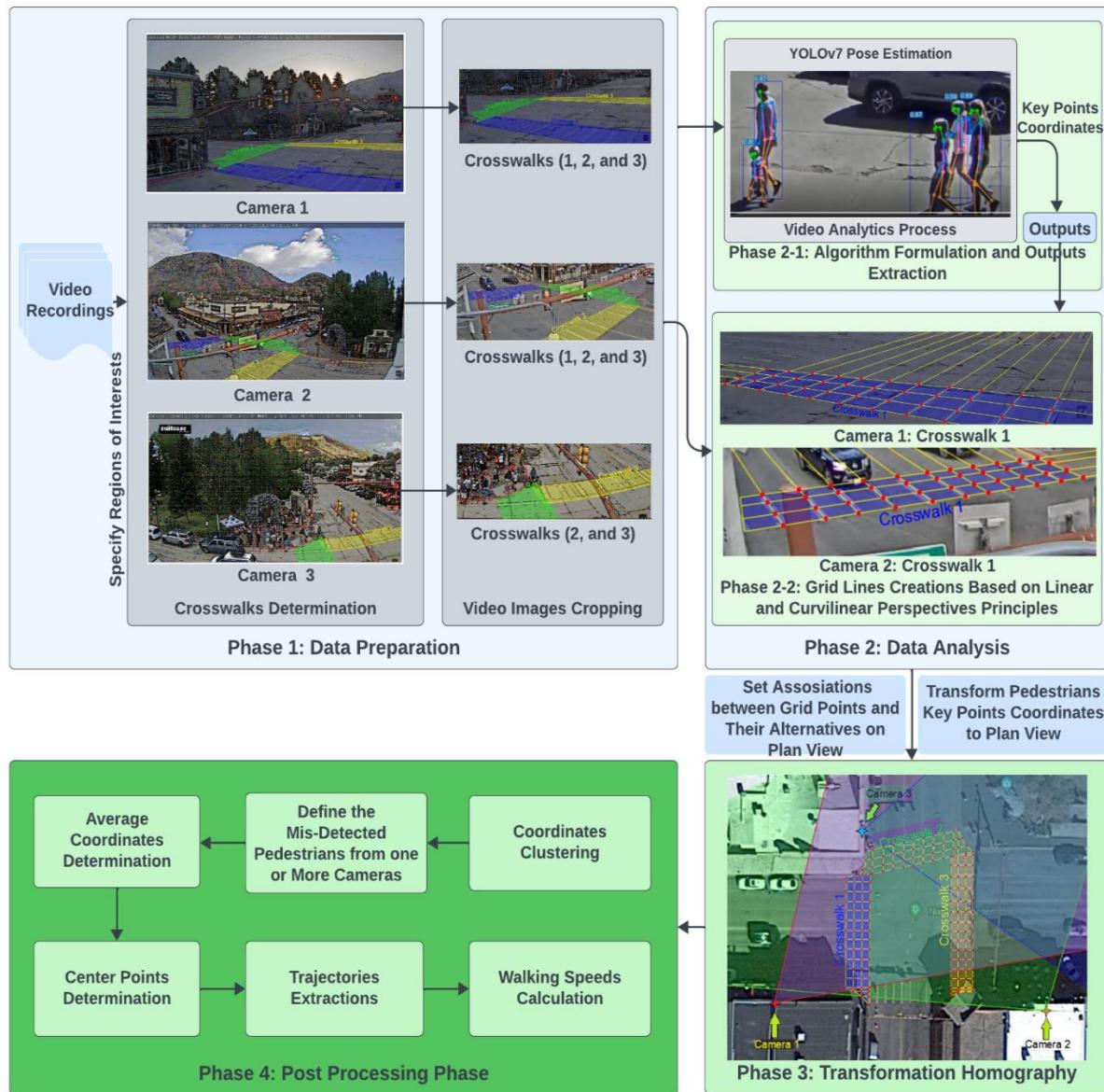


Figure 1: Pedestrian Detection Framework

Afterward, the regions of interest were identified, crosswalks, from the three cameras viewpoints and the integration between FOVs were set to create a unified grid for each crosswalk. Simultaneously, the videos were analyzed using YOLOv7-pose estimation algorithm to extract the pedestrians' key points. Figure 2 demonstrates the distribution of the three cameras and their coverage. Additionally, the figure illustrates the intersected fields of views between cameras, and the positions of each crosswalk with respect to each camera viewpoint.

ALGORITHM FORMULATIONS

YOLOv7, a powerful object detection model, is trained using the COCO dataset with 80 object categories, utilizing 8 NVIDIA V100 GPUs over a training period of approximately 36 hours (Wang et al. 2022). It demonstrated remarkable real-time performance, achieving speeds

of up to 60 frames per second on a single GPU, as evaluated on diverse benchmark datasets such as COCO, Waymo Open, and Udacity Autonomous Driving.

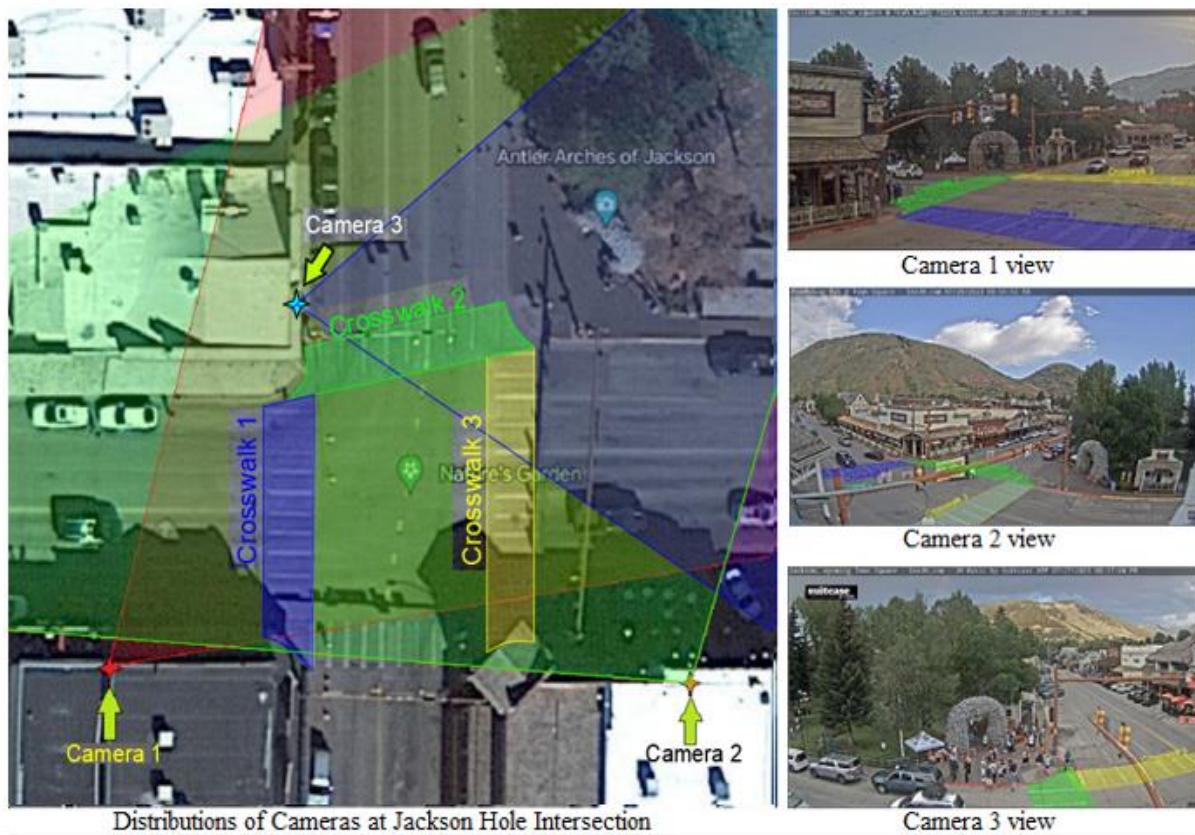


Figure 2: Cameras Positions and Views at The Site of Interest

For pedestrian detection, the YOLOv7 Human-Pose-Estimation Algorithm, available on Rizwan Munawar's GitHub repository, is employed. This algorithm provides bounding boxes for detected persons, along with 17 key points representing human joints and facial features as illustrated in Figure 3. The pretrained model "yolov7-w6-pose.pth" is utilized in the detection process.

Initially, sample videos were recorded from each camera viewpoint to assess the algorithm's detection accuracy and define regions of interest focusing on pedestrian-vehicle interactions, specifically crosswalks. However, the original video from each camera did not yield optimal detection results, prompting the cropping of video images to concentrate on the crosswalks.

Despite accurate pedestrian detection in the analyzed video from a single camera within the regions of interest, a small number of pedestrians remained undetected at the entrances and exits of each camera's defined regions. To address this, integration of detected pedestrians' coordinates was performed on the transformed outputs to plan view. This integration aimed to validate and enhance accuracy, detecting missing pedestrians from one camera by leveraging detections from others. For undetected pedestrians in a single camera's field of view, post-processing phase localization was conducted, restoring their positions in the missing frames based on their relative positions in frames where they were detected.

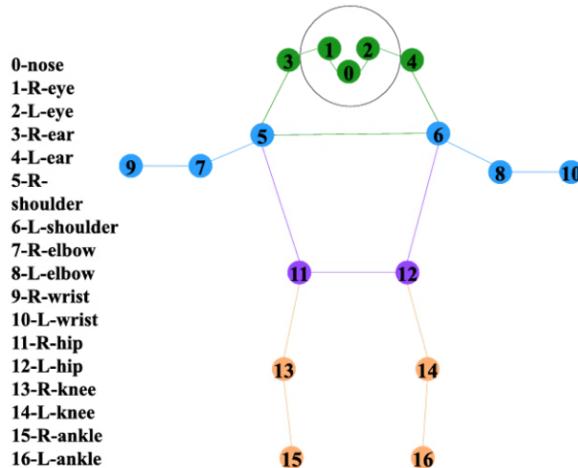


Figure 3: Human Key Points Distribution (Fu et al. 2023)

CROSSWALKS GRIDLINES CREATION

Upon assessing images from three cameras, a feature point extraction deficiency was evident for precise matching with counterparts on the plan view. Cameras 2 and 3, being semi-sphere cameras, displayed curvilinear distortions, making reliance on raw images for transformation inaccurate. To overcome this, a correction technique utilizing linear and curvilinear perspective properties was introduced. This correction enhances feature points, establishes a unified grid for each crosswalk, streamlines coordinate transformation, and improves overall accuracy.

Linear Perspective

For camera 1, it was observed that the extracted images suffer from linear distortion. Consequently, the properties of linear perspective, such as the horizon line and vanishing points, were restored from the image. Adhering to the principles of linear perspective, where parallel lines converge at vanishing points in the image perspective, and all vanishing points lie on the horizon line (Saunders and Backus 2006; Wu et al. 2007). The determination of grid points on the plan-view followed the procedure outlined in (Brownson 1981). Subsequently, these points were aligned with their counterparts in the image perspective. The performed grid in image perspective for camera 1 is demonstrated in Figure 4.

In Figure 4, the horizon line (H.L.) is defined as the line where all vanishing points are located (e.g., ground floor rooftop). V1-CW1, CW3 represents the first vanishing point for crosswalks 1 and 3, while V1-CW2 is the first vanishing point for crosswalk 3. A vanishing point is where parallel lines in the real world converge in the image perspective. Notably, the lines of paints for crosswalks 1 and 3 are parallel, leading them to share vanishing points.

Curvilinear Perspective

The video outputs from cameras 2 and 3 exhibited curvilinear distortions, requiring a modification procedure based on (Santoyo and Santoyo 2021) to establish gridlines on their images. In the curvilinear perspective, lines are represented as arc sections, and parallel lines

intersect at two vanishing points. For horizontal lines, vanishing points lie on the horizon line at the camera's fixing position, while for vertical lines, vanishing points are on a reference vertical line at the middle of the camera's lens. Reference lines are determined using two or more parallel lines of significant length, preferably located farther from the camera lens. Figure 5 illustrates the creation of crosswalk gridlines by applying curvilinear perspective properties to camera 2's output image.

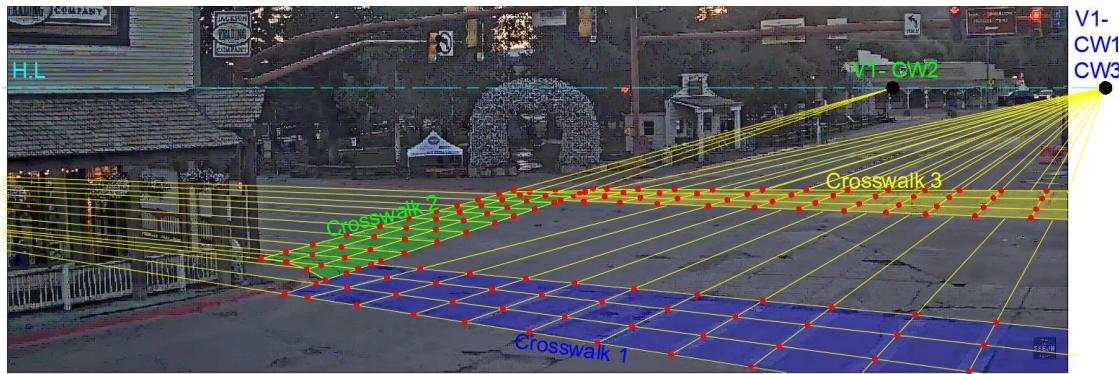


Figure 4: Gridlines Formation on Camera 1 Image Perspective

A control point selection tool was employed to establish associations between grid points in camera views and their counterparts on plan view. Additionally, the tool was used to validate the proposed procedure by selecting a set of points for testing the effectiveness of the transformation.

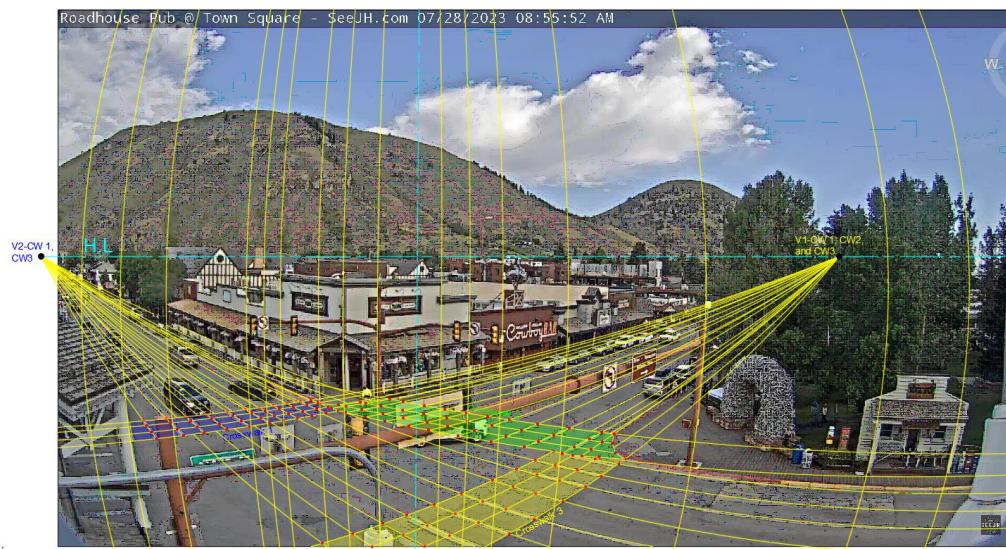


Figure 5: Gridlines Formation on Camera 2 Curvilinear Perspective

DETECTION ASSESSMENT AND DATA ANALYSIS

In this phase, the detection accuracy of the proposed framework using the three cameras was assessed. Referring to Figure 2, camera 1 covers all three crosswalks, while cameras 2 and 3

cover specific portions. A combination of video analysis outputs from cameras 1 and 2 was chosen to assess crosswalk 1, while all three cameras were employed for crosswalks 2 and 3. A selected set of 1,200 frames was analyzed for different conditions, Table 1. The assessment process utilized Intersection over Union (IOU) and the detected number of pedestrians to measure accuracy.

$$IOU = \frac{\text{Number of Detected Pedestrians} \cap \text{Ground Truth}}{\text{Number of Detected Pedestrians} \cup \text{Ground Truth}}$$

Table 1: Detection Assessment of Proposed Framework

		No. of Detected Pedestrians (Intersect of Union (IOU))								
		(Un-congested, Clear, Daytime)			(Un-congested, Rainy, Nighttime)			(Congested, Snowy, Daytime)		
		CW1	CW2	CW3	CW1	CW2	CW3	CW1	CW2	CW3
Camera 1	1804	1285	1480	1322	767	924	2163	1711	1944	
	(0.869)	(0.712)	(0.703)	(0.744)	(0.777)	(0.488)	(0.879)	(0.804)	(0.753)	
Camera 2	1683	836	1720	1164	422	1240	1986	1032	2096	
	(0.811)	(0.463)	(0.817)	(0.655)	(0.428)	(0.655)	(0.807)	(0.485)	(0.812)	
Camera 3	---	632	742	---	413	535	---	737	1025	
	(0.000)	(0.350)	(0.352)	(0.000)	(0.418)	(0.283)	(0.000)	(0.346)	(0.397)	
Combined	1983	1668	1997	1575	857	1664	2280	1908	2377	
	(0.956)	(0.924)	(0.948)	(0.887)	(0.868)	(0.879)	(0.927)	(0.896)	(0.921)	
Ground	2075	1805	2106	1776	987	1892	2460	2129	2582	
Truth	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

The assessment process revealed that the proposed framework achieved an overall precision of 97.73%, a false detection rate of 1.76%, and an average Intersection over Union (IOU) of 0.912. Nighttime detection showed lower IOU due to inadequate street lighting, impacting the algorithm's ability to detect pedestrians wearing colored clothes that blended with pavement color pattern. Camera 3's placement was deemed ineffective for clear crosswalk coverage but could be useful for monitoring eastbound vehicle traffic. The integration of cameras 1 and 2 significantly enhanced detection accuracy, with each camera serving specific roles in covering and improving detection at different parts of the crosswalks.

TRAJECTORIES EXTRACTION

Accurate trajectory extraction and orientation are crucial for post-processing phases to obtain precise walking speeds and movement information. Challenges such as miss-tracking, wrong-tracking, and outliers in video frames necessitate meticulous post-processing techniques. A comprehensive code is developed to calculate transformed coordinates from each camera, averages them, calculates pedestrian feet key points' average coordinates, sets a pixel threshold to relate coordinates to the nearest object, integrates missing frames, and calculates walking speeds. A pixel threshold of 10 pixels and a conversion factor of 1 foot to 3.25 pixels were

applied. Walking speed calculations consider frame rate and calculation intervals. The following equations illustrate the walking speed calculation process.

$$T_d = i / fr$$

$$S_{(i,P_j)} = \frac{fp * \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{T_d}$$

Walking speed (S) calculations involve frame number (i), time duration (T_d), frame rate (fr), transformation factor (fp), and pedestrian id (P_j). Figure 6 displays the code output for pedestrians crossing crosswalk 1.

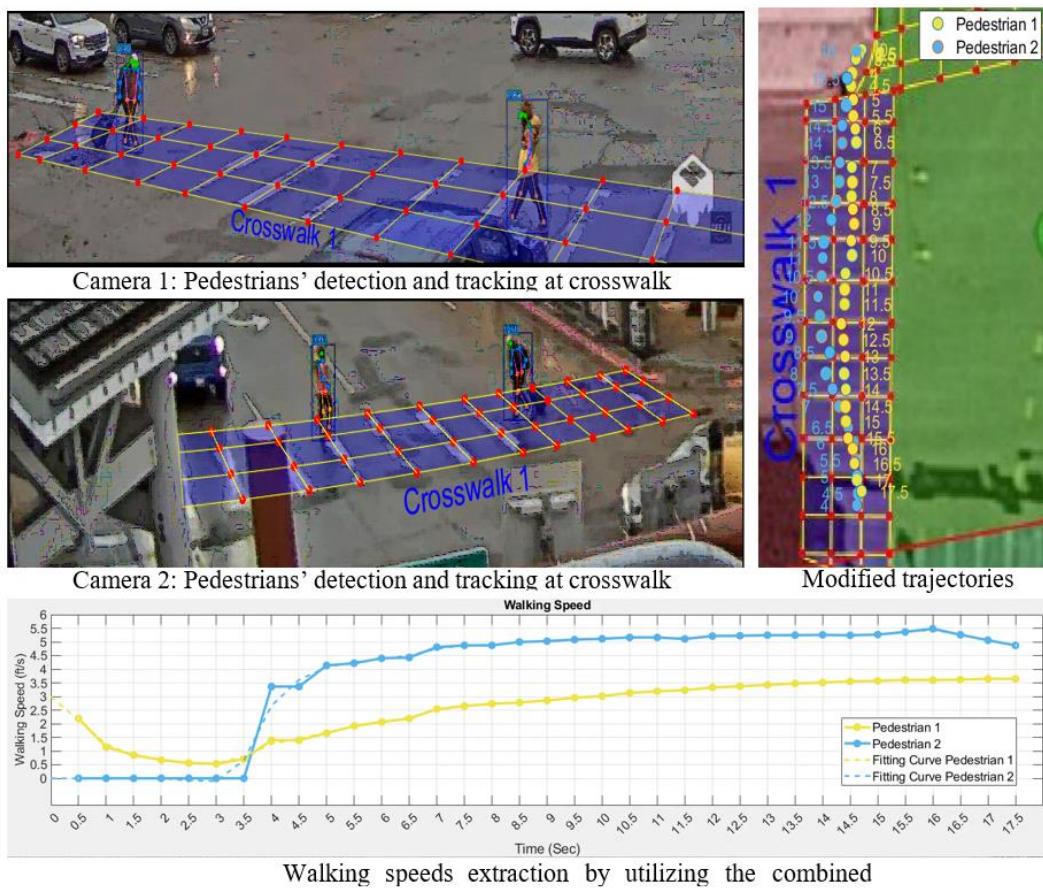


Figure 6: Detection and Tracking of Pedestrians Using the Proposed Framework

Initially, the code outputs reveal trajectories blending. A correction loop, with a selected pixel threshold and multiple trials, accurately relocates coordinates, and the associated time is plotted for trajectory representation. Speed curves show distinctions between an elderly man and a young man crossing the intersection, Figure 6. While the code provides accurate results, improvements are needed for better detection, classification, and speed calculations, especially in congested scenarios, to enhance efficiency and provide valuable insights for traffic conflict analysis frameworks.

CONCLUSIONS AND DISCUSSIONS

In conclusion, this research paper successfully employed the YOLOv7_human_pose_estimation algorithm to analyze pedestrian movements at intersections, making significant contributions to the field. The proposed framework integrates multiple camera fields of view to detect, track, and reconstruct pedestrian trajectories, enhancing accuracy through a spatial methodology. Post-processing phases, facilitated by a developed code, further improved trajectory precision and walking speed calculations. Achieving an overall precision of 97.73%, the study underscores the framework's effectiveness in accurately identifying pedestrians. The integration of surveillance cameras for pedestrian detection holds practical implications for traffic management and intersection design. Future research could expand on this approach to include various road users, advancing the analysis of traffic encounters and creating advanced assessment tools for potential crash analysis.

ACKNOWLEDGMENTS

This work was sponsored by the Wyoming Department of Transportation (WYDOT). Grant Number: RS04221. The writings were enhanced by using ChatGPT which helped in checking for grammatical errors and paraphrasing sentences in a more concise manner.

AUTHOR CONTRIBUTIONS

The authors confirm the contribution to the paper as follows: study conception and design: Ahmed Mohamed and Mohamed M. Ahmed; data preparation and reduction: Ahmed Mohamed; analysis and interpretation of results: Ahmed Mohamed and Mohamed M. Ahmed; draft manuscript preparation: Ahmed Mohamed and Mohamed M. Ahmed. All authors reviewed the results and approved the final version of the manuscript. The authors have no conflict of interest to declare.

REFERENCES

- Alahi, A., M. Bierlaire, and P. Vandergheynst. (2014). "Robust Real-Time Pedestrians Detection in Urban Environments with Low-Resolution Cameras". *Transportation Research Part C: Emerging Technologies*, Vol. 39, pp. 113–128.
- Alfred Daniel, J., C. Chandru Vignesh, B. A. Muthu, R. Senthil Kumar, C. Sivaparthipan, and C. E. M. Marin. (2023). "Fully Convolutional Neural Networks for LIDAR–Camera Fusion for Pedestrian Detection in Autonomous Vehicle". *Multimedia Tools and Applications*, Vol. 82, No. 16, pp. 25107–25130.
- Antonio, J., and M. Romero. (2018). "Pedestrians' Detection Methods in Video Images: A Literature Review". *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 354-360.
- Brownson, C. D. (1981). "Euclid's Optics and Its Compatibility with Linear Perspective. Vol. 24, No. 3 pp. 165-194, 1981. <https://www.jstor.org/stable/41133617>.
- Brunetti, A., D. Buongiorno, G. F. Trotta, and V. Bevilacqua. (2018). "Computer vision and deep learning techniques for pedestrian detection and tracking: A survey". *Neurocomputing*, 300, 17-33.

- Cavallaro, A., O. Steiger, and T. Ebrahimi. (2005). "Tracking Video Objects in Cluttered Background". *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 4, pp. 575–584.
- Cicchino, J. B. (2022). "Effects of Automatic Emergency Braking Systems on Pedestrian Crash Risk". *Accident Analysis and Prevention*, Vol. 172.
- Dahlkamp, H., A. E. C. Pece, A. Ottlik, and H.-H. Nagel. (2004). "Differential Analysis of Two Model-Based Vehicle Tracking Approaches". *DAGM-Symposium*.
- Dollár, P., C. Wojek, B. Schiele, and P. Perona. (2012). "Pedestrian Detection: An Evaluation of the State of the Art". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 4, pp. 743–761.
- Enzweiler, M., and D. M. Gavrila. (2010). "Monocular pedestrian detection: Survey and experiments". *IEEE Trans. on PAMI* 31(12), 2179–2195. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 31. 2179-2195.
- Essa, M., and T. Sayed. (2015). "Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured 3 Conflicts?". *Transportation Research Record*, Vol. 2514, pp. 48–57. 4.
- Federal Highway Administration, US DOT. (2023). Traffic Safety Facts. Retrieved from [https://highways.dot.gov/research/research-programs/safety/intersection-safety].
- Fu, H., J. Gao, and H. Liu. (2023). "Human Pose Estimation and Action Recognition for Fitness Movements". *Computers & Graphics*.
- Hou, J., G. F. List, and X. Guo. (2014). "New Algorithms for Computing the Time-to-Collision in Freeway 6 Traffic Simulation Models". *Computational Intelligence and Neuroscience*, Vol. 2014, 7.
- Ismail, K. A. (2010). *Application of computer vision techniques for automated road safety analysis and traffic data collection*. (Doctoral dissertation, University of British Columbia).
- Kim, D.-G., S. Washington, and J. Oh. (2006). "Modeling Crash Types: New Insights into the Effects of 4 Covariates on Crashes at Rural Intersections". *Journal Of Transportation Engineering*, pp. 282-292.
- Koller, D., J. Weber, and J. Malik. (1994). *Multiple Car Tracking with Occlusion Reasoning*. Lecture Notes in Computer Science, vol 800. Springer, Berlin, Heidelberg.
- Magee, D. R. (2004). "Tracking Multiple Vehicles Using Foreground, Background and Motion Models". *Image and Vision Computing*, Vol. 22, No. 2, pp. 143–155.
- Maurin, B., O. Masoud, and N. P. Papanikolopoulos. (2004). "Tracking All Traffic: Computer Vision 13 Algorithms for Monitoring Vehicles Individuals, and Crowds". *IEEE Robotics and Automation Magazine*, Vol. 12, No. 1, pp. 29–36.
- Santoyo, J., and M. A. Santoyo. (2021). "The 360° Curvilinear Perspective: A Hybrid Hypercubic Angular Space Grid Based on the 1968 Barre and Flocon Proposal". *Nexus Network Journal*, Vol. 23, No. 3, pp. 717–735.
- Saunders, J. A., and B. T. Backus. (2006). "The Accuracy and Reliability of Perceived Depth from Linear Perspective as a Function of Image Size". *Journal of Vision*, Vol. 6, No. 9.
- Sauzier, N., and T. Sayed. (2006). "A Feature-Based Tracking Algorithm for Vehicles in Intersections". *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, Quebec, Canada, 2006, pp. 59-59.
- Stauffer, C., and W. E. L. Grimson. (2000). "Learning Patterns of Activity Using Real-Time Tracking". *IEEE 16 Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 747–757.

- Veeraraghavan, H., O. Masoud, and N. P. Papanikolopoulos. (2003). "Computer Vision Algorithms for 19 Intersection Monitoring". *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, pp. 78–89.
- Wang, C.-Y., A. Bochkovskiy, and H.-Y. M. Liao. (2022). "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors".
- Wu, B., Z. J. He, and T. L. Ooi. (2007). "The linear perspective information in ground surface representation and distance judgment". *Perception & Psychophysics* 69, 654–672.
- Zhang, S., R. Benenson, and B. Schiele. (2017). "CityPersons: A Diverse Dataset for Pedestrian Detection". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017): 4457-4465.

Clarifying Rear-End Crash Propensity with Spatial Spillover Effect: Artificial Bayes-GLM Method

Wei Lin¹; Heng Wei, Ph.D., P.E., F.ASCE²; Zhixia Li, Ph.D.³; and Dong Nian⁴

¹ART-Engines Transportation Research Laboratory, Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH.

ORCID: <https://orcid.org/0000-0001-7231-8542>. Email: lin2wi@mail.uc.edu

²Professor and Director, ART-Engines Transportation Research Laboratory, Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH.

ORCID: <https://orcid.org/0000-0001-5308-8593>. Email: heng.wei@uc.edu

³Associate Professor, Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH.

ORCID: <https://orcid.org/0000-0002-7942-4660>. Email: lizx@ucmail.uc.edu

⁴Ph.D. Candidate, ART-Engines Transportation Research Laboratory, Dept. of Civil and

Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH.

ORCID: <https://orcid.org/0000-0002-1395-4933>. Email: niandg@mail.uc.edu

ABSTRACT

Numerous studies have highlighted the significant influence of adjacent geographic units on rear-end crashes, termed as the spatial spillover effect. This paper introduces an innovative systematic model based on Bayes-GLM to illustrate the propensity for rear-end crashes for improving safety countermeasures on a broader scale. The approach integrates hotspot analysis with K-means clustering into the crash causation model, creating a multi-layered strategy, seeking to improve functionality and scalability by systematically screening potential factors contributing to rear-end crashes. It further identifies relationships between the spillover effect and microscopic impacting factors. The primary advantage of employing such an integrated macro- and microscopic approach lies in the swift identification of critical areas. Moreover, a heuristic analysis of vehicle-to-vehicle interactions, reflecting pre-crash behaviors, is systematically connected to the spillover effect. This linkage enables the integration of a larger scale of crash-influencing factors, thereby enhancing the understanding of the rear-end crash occurrence mechanism.

1. INTRODUCTION

According to the National Highway Traffic Safety Administration (NHTSA), 2023, about 31,785 people were killed by traffic crashes during the first nine months of 2022 (1). Over 30% of crashes are at the rear-end, including 7% of fatalities and 31.1% of injuries (Sharafeldin et al., 2022; Bieber et al., 2022). Rear-end crashes occur when the following vehicle fails to maintain safe following distance and strikes the leading vehicle. This type of crash results in a range of injuries from “possible injury” to “severe injury”, which is determined by speed and other impacting factors such as actions of driver prior to and during crashes. Various factors from vehicles (i.e., breakdown), drivers (i.e., speeding, distractions, tailgating) and environmental conditions (i.e., long curve, bad weather) contribute to crashes. Therefore, it is always the top prioritized strategy to clarify crash mechanisms to better support remedial countermeasures.

Numerous comprehensive studies have been carried out to study rear-end crashes, following into different focuses, analysis of crash frequency and severity (Chen et al., 2016; Das et al., 2011; Jo et

al., 2019; Moussa et al., 2022), effectiveness of remedial countermeasures (Cicchino, et al., 2017; Wu et al., 2017). The rear-end crash analysis is conventionally conducted at microscopic level to study vehicle-to-vehicle interactions by using micro-simulation models (Dimitriou et al., 2018; Wang et al., 2022). The purpose is to identify contributing factors directly associated with pre-crash behaviors (e.g., speed, gap, headway) (Peng et al., 2018) and impacted infrastructures (Arvin et al., 2019; Wang et al., 2019). In this context, those factors reflecting spatial effects at large scale (e.g., land use, demographics, and socioeconomic factors) are not considered. However, some studies found out that the rear-end crashes may not only be affected by the traffic and infrastructural features of the crash location, but also be correlated with the geometric, traffic, or environmental conditions of the adjacent road facilities (e.g., immediately connected segments or intersections). Such an effect is termed as “spatial spillover effect” that measures the impact of exogenous variables beyond an exact crash location. The spillover effect is identified at macroscopic level (i.e., highly aggregated crash datasets over a larger spatial scale such as county level) (Lee et al., 2018; Lin et al., 2022). Therefore, the spill effect has a great potential to be used a safety impact factor for evaluating a roadway corridor or network improvement alternative.

Abundant studies have confirmed that a crash occurred at a location which is close to the boundary of an adjacent geographic unit (e.g., traffic analysis zone, corridor, county) that is likely correlated strongly with the socioeconomic and/or land use factors of that unit (Cai et al., 2016; Huang et al., 2016). Both the Bayes inference theorem and generalized linear models (GLMs) have been widely used to develop crash analysis approaches to identify crash contributing factors and interpret cause-and-effect mechanisms. For example, Mitra (2017) and Kumar (2013) proposed Bayesian models for crash analysis. Aguero-Valverde (2014) further developed models for identifying spatial effects. Park (2016) and Lee (2017) adopted the GLMs to identify the relationship between contributing variables and crashes.

To further explore its applicability, this paper proposes an innovative systematic Bayes-GLM based data-driven model to enhance the capability of rear-end crash propensity at integrated micro and macro level via systematically screen potential factors contributing to rear-end crashes and identify potential relationships among factors and rear-end crashes. Unlike traditional Bayes inference and GLMs, hotspot analysis with K-means clustering and crash causation model are incorporated into the proposed method to enhance functionality and scalability of the Bayes-GLM based data-driven model. The crash causation model consists of Swiss cheese model, Ishikawa diagram and Fault Tree Analysis (FTA) that enables a multi-layered approach to safety through systematic perspective. A rear-end crash occurrence does not result from a singular factor or action, it is normally the outcome of a complex interplay of multiple factors or consequence of sequential actions. The proposed method aims to identify vulnerabilities of safety defense and root causes through drill down into each potential factors and explore reasons behind rear-end crashes, near crashes and other safety concerns from the perspective of drivers, vehicles and environmental conditions. It is not only degrading complexity of systems but also making rear-end crash problems interpretable and visually presented. A proof-of-concept study in the state of Ohio is conducted to test effectiveness of the developed method. In this study, hotspot counties are first filtered and then based to perform an in-depth analysis of the crash mechanism via vehicle-driver-traffic environment interactions.

2. METHODOLOGY

Figure 1 illustrates the methodological framework with the data context for the rear-end crashes analysis. The dependent variable is rear-end crash. The combined independent

variables/datasets are inputs to the Bayes-GLM integrated data-driven model to measure spillover effectives while identifying contributing factors to rear-end crashes. Then, the K-means clustering technique is used to generate Z value in an ArcGIS platform, which is used as inputs to the Bayes algorithm to determine the hotspots. Fault tree analysis is conducted to clarify pre-crash behaviors, which are used to rebuild the vehicle-driver-traffic environment interactions that can be reflective of an inferred crashing mechanism within the hotspots. Additionally, safety assessment is provided for recommending countermeasures for both planning and engineering and management.

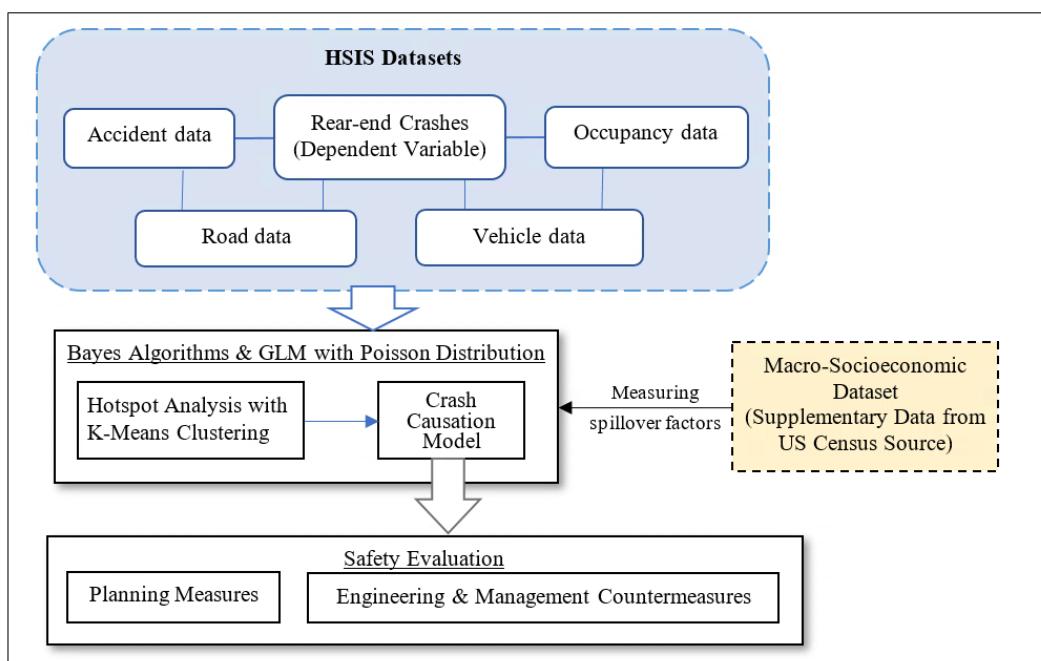


Figure 1. Modeling for Rear-end Crash Analysis

2.1 GLM with Poisson Distribution

While some cases adopt the GLM with a hyper-Poisson distribution for crash data analysis, this approach overlooks the hierarchy and spatial correlation of adjacent areas. Neglecting these factors can lead to erroneous statistical inferences. Integrating GLM with Poisson regression models offers a promising solution to enhance flexibility and extendibility. This is particularly relevant for crash count data, which often exhibits characteristics of either over- or under-dispersion. Over-dispersed data implies that the variance exceeds the mean of crash counts, whereas under-dispersed data indicates that the variance is lower than the mean of crash counts. The GLM is composed of three basic components, namely, random component (i.e., random vector attributes), systematic component that interprets relationships among the dependent variable and independent variables, and the link function that connects the first two components (58). Random vector attributes are expressed by Eq. (1). The distribution of the rear-end crash frequency (r_i) measures the probability of the rear-end crashes occurrence by combining r_i and crash rate (λ) into the Poisson distribution model, as expressed by Eq. (2). The link function $\ln(P_r)$ is expressed by Eq. (3).

$$\mathbf{y} = [y_1, y_2, \dots, y_n] \quad (1)$$

$$P_r = \frac{\lambda^{r_i} \exp(-\lambda)}{r_i!} \quad (2)$$

$$\ln(P_r) = \alpha + \beta_1 y_1 + \beta_2 y_2 + \dots + \beta_n y_n = \alpha + \sum_1^n \beta_n y_n \quad (3)$$

Where, α is a constant coefficient. y_n is nth variable and β_n is the corresponding regression coefficients. The positive (or negative) value of β_n indicates increasing (or decreasing) magnitude of the rear-end crash frequency.

2.1.1 Hotspot Analysis with K-means Clustering

The hotspot analysis is a spatial autocorrelation technique that is widely used to identify whether an event causes spatial clusters, as expressed by Eq. (4). It is represented by a heuristic-supervised learning algorithm to classify factors into K groups based on distance characteristics. The objectives of K-means are to minimize the intra-cluster distances and maximize the inter-cluster distances, which are usually measured by p-norm, as expressed by Eq. (5).

$$G_i = \frac{\sum_1^n w_n^i y_n - \frac{\sum_1^n y_n \sum_1^n w_n^i}{n}}{\sqrt{\frac{\sum_1^n (y_n)^2}{n} - (\frac{\sum_1^n y_n}{n})^2} \sqrt{\frac{n \sum_1^n (w_n^i)^2 - (\sum_1^n w_n^i)^2}{n-1}}} \quad (4)$$

$$|d_{tr} - d_{te}|_p = (\sum_1^d |d_{tr}^i - d_{te}^i|^p)^{\frac{1}{p}} \quad (5)$$

Where, G_i is Z value that provides a criterion for identifying hotspots (if Z value is equal to or more than 1.5). w_n^i is spatial weight between variables (or features) i and n . n is total number of features. d_{tr} and d_{te} is training dataset and testing dataset, and d_{tr}^i and d_{te}^i are their corresponding data, respectively. d is dimensional space, p is the power. Other variables are predefined as aforementioned in the paper.

2.1.2 Bayes Theorem

Bayes inference theorem originated from the description of joint probability of two dependent events, as shown by Eq. (6), or Eq. (7). If normalizing constant σ is replaced with $\frac{1}{P(Y)}$, the equation can be simplified as Eq. (8).

$$P(X \& Y) = P(Y)P(X|Y) = P(X)P(Y|X) \quad (6)$$

$$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)} \quad (7)$$

Where, $P(X \& Y)$ is probability of both events X and Y occur. $P(X)$ or $P(Y)$ is the probability of occurrence of event X or Y . $P(X|Y)$ is the probability of event X occurrence, given Y has occurred. $P(Y|X)$ is the opposite meanings.

$$P(X|Y) = \sigma P(X)P(Y|X) \quad (8)$$

Eq. (8) and Eq. (9) illustrate distinct dimensional interpretations of the Bayes inference theorem. If vector attributes are used, the Bayes inference theorem can be expressed by Eq. (9) to describe multi-dimensional relationships.

$$P(\vec{X}|\vec{Y}) = P(\vec{X})^T P(\vec{Y}|\vec{X}) = \sigma P(x_i)P(y_1|x_i)P(y_2|x_i) \dots P(y_n|x_i) \quad (9)$$

Where, x_i denotes the case whether an event occurs or not. y_i is a contributing factor influencing an event's occurrence.

2.2 Crash Causation Model

Swiss cheese model, Ishikawa diagram and FTA are three sub models of the crash causation model. They are interacting with each other. The Swiss cheese mode was proposed by a famous British psychologist, James Reason from University of Manchester, in the 1990s. It is a widely acceptable model making significant contributions to human error management (Reason, 1990; Ishikawa, 1987). In its application in end crash analysis, Swiss cheese model needs to be used with Ishikawa diagram and FTA to enable functionality of crash causation model.

The Ishikawa diagram is also called fishbone diagram, a causal diagram created by Kaoru Ishikawa to identify potential causes of an event. Figure 2 illustrates the concept of Ishikawa diagram, shown as a fish bone. The fish's head faces to the right side. There are many ribs branching off the backbone. Each rib means one category of main causes to the unsafety event occurring. Normally four categories are used, including culture, people, equipment, and environment. Each rib contains many sub-branches. They are potential reasons contributing to the unsafety event. These ribs and sub-branches are customized to different areas depending on needs. In traffic safety analysis, it is classified into three aspects, i.e., drivers, vehicles, infrastructure, and environment in corresponding to people, equipment, culture, culture, and environment.

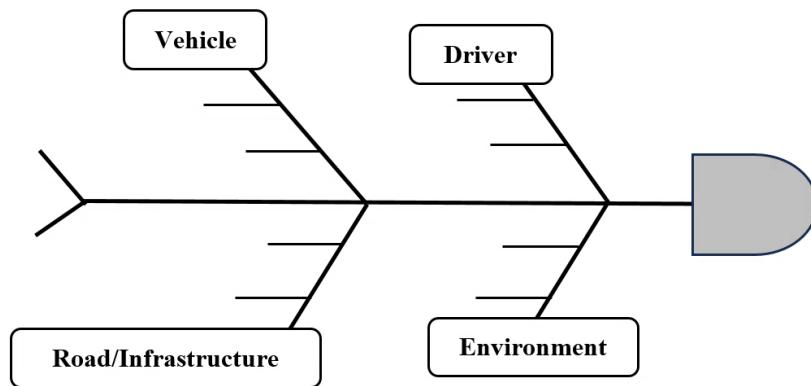


Figure 2. Architecture of Ishikawa Diagram

The FTA is a failure analysis technique used for identifying ways to reduce risks through finding root causes leading to a top incident. The FTA is a graphical tool for analyzing rear-end crashes occurrence from systematic considerations of involved driver-vehicle-infrastructure-environment

interactions. Boolean logic forms the basis of the FTA to combine lower level of events (i.e., indirect events or behaviors that could cause top event occurs) through a top-down-to-bottom searching approach to identify all indirect events that might cause a top event occurrence (i.e., crashes in this study). Two elements, events and logic gates (i.e., and gate, or gate, priority and gate, exclusive gate, inhibit gate) are embedded into the FTA to describe the relationship of different indirect events. Top event is expressed by E , intermediate events are shown by E_1, E_2 . Probability of and gate and or gate is expressed by Eqs. (10) and (11).

$$P_{AND} = E_1 E_2 \dots E_n = \prod_{i=1}^{i=n} E_i \quad (10)$$

$$P_{OR} = E_1 + E_2 + \dots + E_n = \sum_{i=1}^{i=n} E_i \quad (11)$$

3. PRACTICAL EXEMPLARY STUDY WITH HSIS DATA AND CENSUS DATA

Data used in this study consists of HSIS data and macro-socioeconomic datasets obtained from U.S. Census Bureau (USCB, 2020). The HSIS is used as the main data source, and the macro-socioeconomic datasets are supplementary to meet the study objectives. Two types of crash files (accident and road) in HSIS databases were merged into one file based on cntytrte, begmp and endmp variables in road file and cntytrte and milepost variables in accident file. Additionally, the combined file was integrated into occupancy file and vehicle file into one file based on caseno variable that defines unique accident case number. Meanwhile, ytomodel variable is calculated by accyr variable (accident year, originally from accident file) minus vehyr variable (vehicle model year, originally from vehicle file). A total of 199,416 rear-end crashes datasets were filtered out from the HSIS data source and used in the study.

The results from the GLM analysis with Poisson distribution approach are summarized in Table 1. The Sig value indicates the significance of variables.

An independent variable is significant as its Sig value is lower than 0.05. The sign of B means the relationship with a significant variable and rear-end crashes frequency, where plus denotes positive relationship and minus denotes negative relationship. The value of B describes corresponding magnitude of rear-end crash frequency. Table 1 shows the significant variables with Sig values of less than 0.05. Within the macro-socioeconomic datasets, variables of FE_TPO, MA, PPL_TPO, TP_TPO, TB_TPO, TM_TPO have negative relationships with rear-end crashes frequency – a higher value indicates lower probability of rear-end crash frequency. However, other variables like CGR, TV_TPO, TW_TPO, TT_TPO, TH_TPO, ND_TPO, DP_TPO have opposite results. Those people such as females, old people, poverty people, minority people are viewed as “disadvantaged group of population” that need more protections from society. The counties with a greater number of disadvantaged populations usually have fewer rear-end crashes because of more households with no vehicles. Especially for poverty people, they may have financial difficulties in recovering from accidents. It’s critical for them to enhance their self-defense awareness of safe driving, which may build on the corresponding society culture that unintentionally forces them drive more carefully.

In those counties, more commuting travels made by public transportation or bicycles, rather than private cars, are supposed to help reduce the chance of rear-end crash occurrences. On the contrary, increased vehicular trips build the circumstance leading to a higher rear-end crashes propensity. Meanwhile, increased walking trips may increase the chance of rear-end crashes at intersections or

urban streets. Even more, the increased opportunities of working from home may provide those parents more flexibility to participate in more volunteering work at their children's school or other community services. This may be another notable factor attributing to more rear-end crashes in urban areas. This phenomenon pointed out by Bibby has been further confirmed with some other studies (Bibby, 2017).

Table 1. Results of GLM Analysis

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald	Chi-Square	df		Lower	Upper
Macro-socioeconomic dataset										
CGR	+0.05	0.02	0.00	0.10	4.47	1.00	0.03	1.05	1.00	1.10
FE_TPO	-0.05	0.00	-0.06	-0.04	120.79	1.00	0.00	0.95	0.94	0.96
MA	-0.01	0.00	-0.02	-0.01	15.56	1.00	0.00	0.99	0.98	0.99
PPL_TPO	-0.01	0.00	-0.01	0.00	6.72	1.00	0.01	1.00	0.99	1.00
TV_TPO	+0.05	0.00	0.05	0.06	283.83	1.00	0.00	1.05	1.05	1.06
TP_TPO	-0.20	0.02	-0.23	-0.16	121.39	1.00	0.00	0.82	0.79	0.85
TW_TPO	+0.09	0.01	0.07	0.11	80.13	1.00	0.00	1.10	1.07	1.12
TB_TPO	-0.42	0.02	-0.46	-0.37	342.91	1.00	0.00	0.66	0.63	0.69
TT_TPO	+0.16	0.01	0.14	0.19	176.70	1.00	0.00	1.18	1.15	1.20
TH_TPO	+0.21	0.01	0.20	0.23	735.93	1.00	0.00	1.24	1.22	1.26
ND_TPO	+0.07	0.00	0.06	0.07	341.35	1.00	0.00	1.07	1.06	1.08
DP_TPO	+0.06	0.00	0.05	0.06	388.54	1.00	0.00	1.06	1.05	1.06
TM_TPO	-0.01	0.00	-0.01	0.00	32.45	1.00	0.00	0.99	0.99	1.00
Accident dataset										
hour	+6.41	0.40	5.63	7.19	259.22	1.00	0.00	608.64	278.87	1.33E3
light	-3.43	0.22	-3.87	-2.99	233.40	1.00	0.00	0.03	0.02	0.05
loc_type	+8.95	0.85	7.28	10.63	110.08	1.00	0.00	7.73E3	1.45E3	4.12E4
rd_char1	+2.47	0.24	1.99	2.94	102.16	1.00	0.00	11.77	7.30	18.98
rdsurf	+1.84	0.24	1.37	2.31	58.29	1.00	0.00	6.30	3.93	10.10
weather	-6.11	0.28	-6.67	-5.56	465.41	1.00	0.00	0.00	0.00	0.00
weekday	-3.59	0.21	-4.00	-3.19	301.74	1.00	0.00	0.03	0.02	0.04
Occupancy dataset										
physcond	+1.61	0.10	1.40	1.81	241.11	1.00	0.00	4.99	4.07	6.11
rest1	-1.65	0.08	-1.81	-1.49	398.38	1.00	0.00	0.19	0.16	0.23
sob_tst	-1.75	0.11	-1.96	-1.54	270.94	1.00	0.00	0.17	0.14	0.21
Vehicle dataset										
contrib1	-11.60	0.81	-13.19	-10.00	202.77	1.00	0.00	0.00	0.00	0.00
drv_age	-0.06	0.01	-0.07	-0.05	91.66	1.00	0.00	0.94	0.93	0.96
drv_sex	+8.13	0.18	7.78	8.48	2081.74	1.00	0.00	3.40E3	2.40E3	4.82E3
num_occ	-1.37	0.13	-1.62	-1.13	119.13	1.00	0.00	0.25	0.20	0.32
trf_cntl	+0.30	0.06	0.18	0.42	23.11	1.00	0.00	1.35	1.20	1.53
vehtype	-1.93	0.27	-2.45	-1.41	52.28	1.00	0.00	0.14	0.09	0.24
yтомodel	-0.36	0.01	-0.38	-0.35	2458.50	1.00	0.00	0.69	0.68	0.70
Road dataset										
func_cls	-0.56	0.04	-0.65	-0.48	170.72	1.00	0.00	0.57	0.52	0.62
no_lanes	+0.96	0.01	0.93	0.99	4246.84	1.00	0.00	2.62	2.55	2.70

As a result of analyzing the accident datasets, the variables of light, weather and weekday have negative relationships with rear-end crash frequency, and the variables of hour, loc_type, rd_char1, rdsurf have positive relationships. According to the definitions of variables shown in Table 1, daylight and good weather along non-curve geometry during weekdays are a notable factor correlated with high rear-end crashes propensity. More travels are often made under good weather with good visibility conditions, resulting in more traffic and more chance of crashes. Meanwhile, carelessly driving in high speed could easily happen in non-curve roads under good driving conditions as less driving tasks could catch drivers' attention. In addition, rear-end crashes may highly occur during nighttime or at non-intersection related facilities or on bad road surface. The first two phenomenon might happen due to light traffic or free-flow condition, driving speed may opt to be higher than speed limit. Under bad road surface conditions such as wet or ice, it is hard for drivers, especially truck drivers, to easily control braking distance.

Variables of rest1 and sob_tst within the occupancy datasets have a negative relationship with rear-end crashes and the variable of physcond has a positive relationship. It is noticed that bad physical conditions of occupants or safety equipment unused by occupants or under influence of alcohol or drug led to more rear-end crashes. That confirms to the patterns that most of crashes caused by driving errors or improper driving behaviors.

Vehicle variables of contrib1, drv_age, num_occs, vehtype, ytomodel have negatively correlated with rear-end crashes, and variables of drv_sex and trf_cntl have positively correlated. The data analysis shows that speeding, female driver, truck driver and traffic conditions, as well as new vehicles are factors highly correlated with rear-end crashes frequency. However, older drivers or high lane changing frequency has a lower correlation with rear-end crashes, because of the low speed associated with those factors. In addition, a greater number of occupants within a car result in lower rear-end crashes, because more people pay attention to driving behaviors. Road variables of func_cls and no_lanes have negative and positive relationships with rear-end crashes, respectively. It can be noticed that the freeway is an infrastructure that is highly correlated with rear-end crash propensity.

Figure 3 shows the results from the hotspot analysis. The county is identified as a hotspot if the output Z value is equal to or greater than 1.5. The identified hotspot counties include Cuyahoga, Franklin, Hamilton, Lucas, and Summit counties.

Crash causation model outputs the possible risk categories and corresponding potential factors, as shown in Table 2. There are two events contributing to rear-end crashes, including unacceptable gap occurrence and unacceptable speeding. Behind unacceptable gap occurrence, there are four intermediate events might cause near crashes under gate 1, namely failure to yield, improper turn, followed too closely, swerving to avoid. Meanwhile, some basic events under end gate could directly result in unacceptable gap occurrence, including left of center, ran red light or stop sign, improper backing, improper lane changing. Behind unacceptable speeding, there are two intermediate events that might cause near crashes, namely aggressive drivers and unsafe speed. There are some basic events under sub gate 1 contributing to intermediate events.

Microscopic variables impacting rear-end crashes are further identified through cause effect mechanisms analysis, as shown by Figure 4. The FTA analysis process consists of three categorized events: top event (rear-end crashes), intermediate events (i.e., contributing factors represented by rectangles, except the top one), and basic events (circles). The “+” sign within gate means OR; “.” means AND. They are gate symbols to determine relationships among events working for top event. On the basis of the top event, the logic of the FTA is to find out what intermediate events (contributing factors) contribute to which and what basic events contribute to which via intermediate events. In Ohio, eight intermediate events and nineteen basic events are identified and provide

indications of potential risks to rear-end crashes. The probability of each basic event is shown in the fault free diagram. On top of that, the probability of rear-end crashes, P_T is estimated by Equation (12). The cut set of fault free diagram means that a set of basic events have high risks to rear-end crashes. The cut set S_i is determined as a combination of one item from subset S_1 , S_2 and S_3 , respectively, by using Equations (13) through (15). There are seventy-two possible combinations, five of them have high rear-end crash propensity as illustrated in Figure 3:

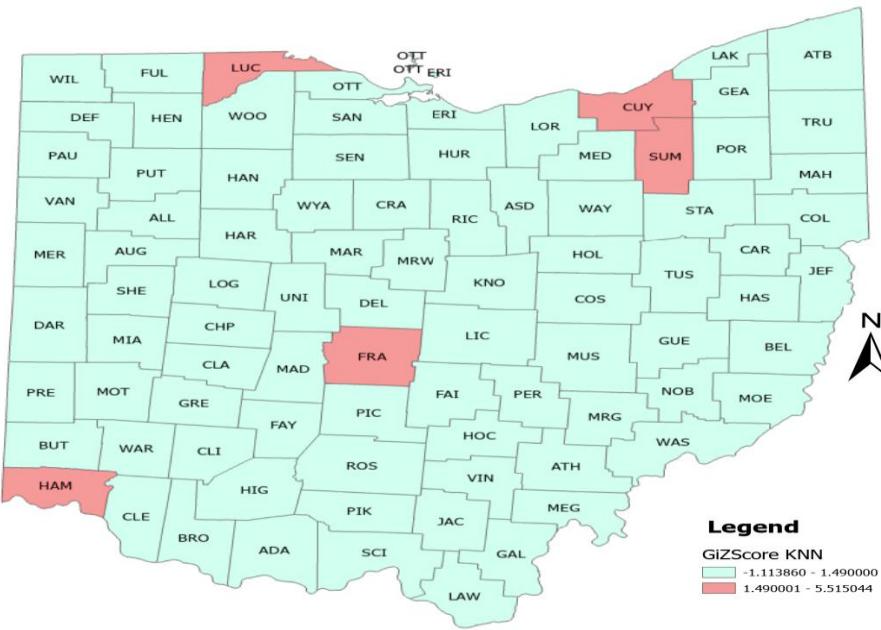


Figure 3. Hotspots Analysis in Ohio

Table 2. Possible Risk Categories and Potential Factors

Risk Category	Risk			
	Gate Level		Sub gate Level	
Unacceptable gap occurrence E_1	Gate 1	Failure to yield E_3	Sub gate 1	No traffic control device x_1 Old drivers x_2
		Improper turn E_4	Sub gate 1	Making U turn x_3 Turning right x_4 Turning left x_5
		Followed too closely E_5	Sub gate 1	Bad light conditions x_6 Severe weather x_7 Slowing/Stopped in traffic x_8
		Swerving to avoid E_6	Sub gate 1	Bad road surface x_9 Curve-graded x_{10}
			End gate	Left of center x_{11} Ran red light or stop sign x_{12} Improper backing x_{13} Improper lane changing x_{14}
Unaccepted speeding E_2	Gate 1	Aggressive drivers E_7	Sub gate 1	Emotional drivers x_{15} Drunk drug driving x_{16}
	Gate 1	Unsafe speed E_8	Sub gate 1	Passing x_{17} Driverless x_{18} Fatigued x_{19}

- Combination 1 includes X8, X15, and X18, the case of emotional drivers using automatic driving functions while following slow traffic stream or a stopped vehicle.
- Combination 2 includes X9, X15, and X18, the case of emotional drivers using automatic driving functions driving through bad conditions of the road surface.
- Combination 3 includes X6, X15, and X18, the case of emotional drivers using automatic driving functions under bad light conditions.
- Combination 4 includes X7, X15, and X18, the case of emotional drivers using automatic driving functions under severe weather.
- Combination 5 includes X8, X16, and X17, case of driving under the influence of drunk or drug when it is overtaking slow or stopped vehicles ahead in traffic flow.

$$\begin{aligned}
 P_T &= E_1 E_2 = (E_3 E_4 E_5 E_6)(E_7 E_8) \\
 &= (x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_{10} + x_{11} + x_{12})(x_{15} + x_{16})(x_{17} + x_{18} + x_{19}) \\
 &= (\sum_{i=1}^{12} x_i)(x_{15}x_{17} + x_{15}x_{18} + x_{15}x_{19} + x_{16}x_{17} + x_{16}x_{18} + x_{16}x_{19}) \\
 &= 1.4337 * (2.19203E - 06) = 0.000003
 \end{aligned} \tag{12}$$

$$S_1 = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{10}, x_{11}, x_{12}\} \tag{13}$$

$$S_2 = \{x_{15}, x_{16}\} \tag{14}$$

$$S_3 = \{x_{17}, x_{18}, x_{19}\} \tag{15}$$

$$\forall s_1 \in S_1, \forall s_2 \in S_2, \forall s_3 \in S_3, S_i = s_1 \cup s_2 \cup s_3 \tag{16}$$

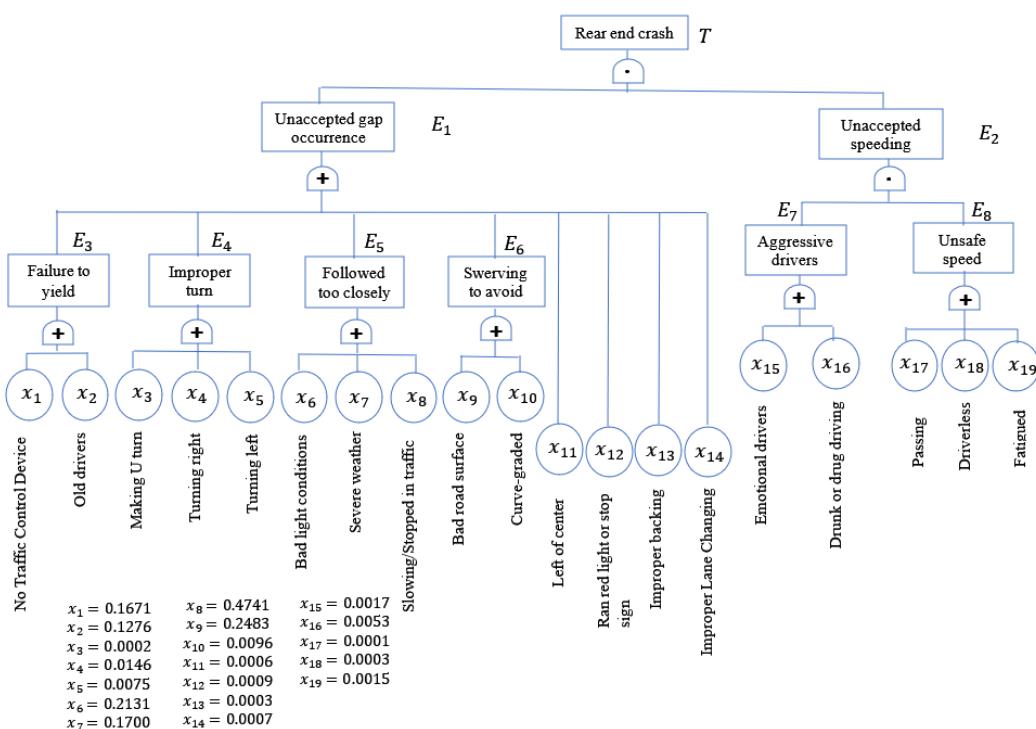


Figure 4. Exemplary Result of FTA

4. CONCLUSIONS

This paper presents an innovative Bayes-GLM based data-driven model by integrating macroscopic variables and microscopic variables into the rear-end crashes analysis. The HSIS data source is greatly helpful to make the proposed methods become testable. In general, the benefits from the integrated macro- and micro-scope approach are two-fold: 1) to enable a quick identification of the critical areas (i.e., hotspot counties in the case study) through employing the Bayes algorithms and GLM with Poisson distribution, alongside spillover effect analysis; 2) the outputs of the hotspot analysis with K-means clustering facilitate to generate inputs to crash causation model which provides a heuristic vehicle-to-vehicle interactions reflecting pre-crash behaviors, so that the rear-end crash occurrence mechanism could be quickly clarified with the supportive HSIS data. As a result, the analysis outcomes will be valuable to enhance the safety evaluation of recommended countermeasures from planning and engineering perspectives. Meanwhile, the contributions of this paper flow into those aspects:

- The Bayes-GLM-Based method contributes to pinpointing the underlying reasons for clarifying rear-end crash propensity through analysis of contributing factors from vehicles, drivers and environmental conditions, not just focusing on immediate causes.
- The FTA diagram visualizes relationships among rear-end crashes and causes reducing complexity of systems as well as making it easier to be understandable.
- A systematic analysis is conducted to identify how contributing factors interact and lead to rear-end crashes via comprehensive analysis of all possible failure paths.
- Uncertainty is quantified in traffic safety analysis for more realistic and reliable estimates by incorporating Bayes inference into the proposed method.
- Overcoming limitations of traditional linear regression models, the proposed method is capable of handling with non-normality and heteroscedasticity of traffic data resulting in better-fitting models.

Future studies are planned to extend the developed method into different types of crashes, such as commercial vehicles involved at work zone safety.

ACKNOWLEDGEMENT

The author would like to acknowledge U.S. Department of Transportation Federal Highway Administration for providing HSIS crash data used in this study and sponsorship from the CYBER-CARE, USDOT Tier I University Transportation Center. Any opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Agency; therefore, no official endorsement should be inferred.

REFERENCES

- NHTSA. NHTSA Estimates for First Nine Months of 2022 Suggest Roadway Fatalities Beginning to Level Off After Two Years of Dramatic Increases. 2023. Retrieved February 2023, from <https://www.nhtsa.gov/press-releases/nhtsa-estimates-traffic-deaths-2022-third-quarter#:~:text=The%20National%20Highway%20Traffic%20Safety,the%20same%20time%20in%202021.>

- Sharafeldin, M., Farid, A., and Ksaibati, K. Injury Severity Analysis of Rear-End Crashes at Signalized Intersections. *Sustainability*, 2022, 14, 13858.
- Bieber, C., and Ramirez, A. Rear-End Collisions: Fault & Compensation. 2022. Retrieved February 2023, from <https://www.forbes.com/advisor/legal/auto-accident/rear-end-collision/>.
- Chen, C., Zhang, G., Yang, J., Milton, J. C., and Alcántara, A. D. 2016. An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accident analysis and prevention*, 90, 95-107.
- Das, A., and Abdel-Aty, M. A. 2011. A combined frequency-severity approach for the analysis of rear-end crashes on urban arterials. *Safety science*, 49, 1156-1163.
- Jo, Y., Oh, C., and Kim, S. 2019. Estimation of heavy vehicle-involved rear-end crash potential using WIM data. *Accident analysis and prevention*, 128, 103-113.
- Moussa, G. S., Owais, M., and Dabbour, E. 2022. Variance-based global sensitivity analysis for rear-end crash investigation using deep learning. *Accident analysis and prevention*, 165, 106514-106514.
- Cicchino, J. B. Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. *Accident Analysis & Prevention*, 2017, 99, 142-152.
- Wu, Y., Abdel-Aty, M., Park, J., and Zhu, J. 2018. Effects of crash warning systems on rear-end crash avoidance behavior under fog conditions. *Transportation research. Part C, Emerging technologies*, 95, 481-492.
- Dimitriou, L., Stylianou, K., and Abdel-Aty, M. A. Assessing rear-end crash potential in urban locations based on vehicle-by-vehicle interactions, geometric characteristics and operational conditions. *Accident Analysis & Prevention*, 2018, 118, 221-235.
- Wang, X., Zhang, X., Guo, F., Gu, Y., and Zhu, X. Effect of daily car-following behaviors on urban roadway rear-end crashes and near-crashes: A naturalistic driving study. *Accident Analysis & Prevention*, 2022, 164, 106502.
- Peng, Y., Wang, X., Peng, S., Huang, H., Tian, G., and Jia, H. Investigation on the injuries of drivers and copilots in rear-end crashes between trucks based on real world accident data in China. *Future Generation Computer Systems*, 2018, 86, 1251-1258.
- Arvin, R., Kamrani, M., and Khattak, A. J. How instantaneous driving behavior contributes to crashes at intersections: Extracting useful information from connected vehicle message data. *Accident Analysis & Prevention*, 2019, 127, 118-133.
- Wang, C., Xu, C., and Dai, Y. A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. *Accident Analysis & Prevention*, 2019, 123, 365-373.
- Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M., and Cai, Q. Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects. *Accident; Analysis and Prevention*, 2018, 111, 12-22.
- Lin, W., Wei, H., and Ash, J. E. Modeling spatial spillover effect on intersection crash propensity: a case study at the county level in Ohio, *Journal of Transportation Safety & Security*, 2022. 10.1080/19439962.2022.2129892.
- Cai, Q., Lee, J., Eluru, N., and Abdel-Aty, M. Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. *Accident; Analysis and Prevention*, 2016, 93, 14-22.

- Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., and Abdel-Aty, M. Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of Transport Geography*, 2016, 54, 248–256.
- Mitra, S., and Washington, S. On the nature of over-dispersion in motor vehicle crash prediction models. *Accident Analysis and Prevention*, 2007, 39(3), 459–468.
- Kumar, C. N., Parida, M., and Jain, S. S. Poisson family regression techniques for prediction of crash counts using Bayesian inference. *Procedia - Social and Behavioral Sciences*, 2013, 104, 982–991.
- Aguero-Valverde, J. Direct spatial correlation in crash frequency models: Estimation of the effective range. *Journal of Transportation Safety & Security*, 2014, 6(1), 21–33.
- Park, M., Lee, D., and Jeon, J. Random parameter negative binomial model of signalized intersections. *Mathematical Problems in Engineering*, 2016, 1–8.
- Lee, J., Abdel-Aty, M., and Cai, Q. Intersection crash prediction modeling with macro-level data from various geographic units. *Accident; Analysis and Prevention*, 2017, 102, 213–226.
- Reason, J. 1990. *Human Error*, Cambridge, Cambridge University Press.
- Dekker, S., and Hollnagel, E. 2023. *Resilience Engineering: New directions for measuring and maintaining safety in complex systems*.
- USCB. Advanced search. 2020. Retrieved June 2021, from <https://data.census.gov/cedsci/>.
- Bibby, A. Why work from home? Here are 9 big reasons. 2017. Retrieved June 2021, from <https://www.flexjobs.com/blog/post/why-work-from-home/>.

Highway Abandoned Objects Recognition Based on Open Vocabulary Object Detection Approach

Siwen Liu¹; Yifan Zhuang²; Nanfang Zheng³; Yufei Ji⁴; Pei Liu⁵; and Ziyuan Pu⁶

¹School of Transportation, Southeast Univ. Email: 213202087@seu.edu.cn

²School of Transportation, Southeast Univ. Email: zhuangyf2012@gmail.com

³School of Transportation, Southeast Univ. Email: 213202292@seu.edu.cn

⁴School of Transportation, Southeast Univ. Email: 213201789@seu.edu.cn

⁵School of Transportation, Southeast Univ. Email: liupeiluna@foxmail.com

⁶School of Transportation, Southeast Univ. Email: ziyuan.pu@monash.edu

ABSTRACT

Highway abandoned objects have seriously threatened traffic safety, which may lead to vehicle crashes and block roadways. It is significant for transportation agencies to recognize and clear them as soon as possible. However, it is challenging to detect these abandoned objects because they have various sizes/categories and appear at different locations. Although computer-vision-based object detection algorithms have succeeded in traffic monitoring and autonomous driving, most SOTA algorithms cannot satisfy the demand for highway abandoned object detection due to undetermined target categories. Inspired by impressive progress in contrastive learning and large language models, this paper proposes a two-stage open-vocabulary object detection algorithm for highway abandoned object detection. The first stage is the region proposal, with the pre-trained image encoder presenting all region candidates on multiple scales. Then, the second stage adopts CLIP and computes similarities of embeddings from proposed regions. In contrast to conventional object detection algorithms, this algorithm is not constrained by limited training data categories. The proposed algorithm is validated using real-world highway surveillance videos collected in Nanjing, China, containing varying light conditions and multiple video angles. Compared with SOTA object detection models, the proposed one has shown extraordinary performance and generalization abilities for abandoned object detection, with higher precision and running efficiency, that is, in low-light conditions. This approach is helpful for transportation agencies to provide early warning in time and find accident perpetrators.

1. INTRODUCTION

Highway abandoned objects have become one of the leading causes of severe highway accidents, with over 10% (Traffic Management Bureau, Ministry of Public Security, China) caused by abandoned objects in China. Recognizing highway abandoned objects will help transportation agencies obtain real-time road status and avoid potential traffic accidents. Considering the “infinite” monitoring spots, it is almost impossible to recognize them by agency staff. Although computer vision technologies have been broadly implemented in transportation applications, it is still challenging for most solutions to handle this task. Unlike previous object detection tasks, highway abandoned objects have unconstrained object categories, scales, and positions. Most existing computer vision algorithms work on a fixed range of object categories and cannot meet the demand. Meanwhile, this application needs 7x24 traffic monitoring, i.e.,

low-light scenarios, which requires the algorithm to be robust at various light and detection distance conditions.

In order to liberate human resources as well as improving traffic efficiency and safety, non-deep-learning algorithms have already been used in highway abandoned object detection and classification. However, those conventional methods are usually inefficient and have low precision, especially when detecting diverse object scales i.e., small objects, and suffering from complex lighting conditions. Deep-learning methods are now introduced to this problem and zero-shot object detection methods such as open-vocabulary object detection (OVD) models (Zareian et al., 2021) have shown extraordinary detection accuracy, without further fine-tuning on large labeled data.

Thus, this paper proposes a two-stage open-vocabulary object detection method to complete highway abandoned object recognition and classification. This work has two primary contributions: 1) constructing a vocabulary list containing all kinds of potential abandoned object classes, and 2) designing a multi-modal detection model used to accomplish precisely detection and classification without further fine-tuning on target datasets. The rest of paper is organized as follows. Section II reviews related works about vision-based object detection in intelligent transportation applications. Section III describes the proposed method for abandoned object detection. Section IV provides our experimental details. The last section summarizes this work and puts refinement plans.

2. LITERATURE REVIEW

Traditionally, aiming at abandoned object detection, background subtraction models are widely used for foreground object extraction (Bhinge et al., 2015; Cuevas et al., 2017, 2017; Tripathi et al., 2013), while edge-based object recognition method (Dahi et al., 2017), gaussian mixture models (Saluky et al., 2020) and some deep learning methods (Wang et al., 2023) are also popular to solve this problem. As for classification tasks, approaches, e.g., CNN (Saluky et al., 2020), SVM classifier (Ketaki Shet-Talathi & S.V.Khobragade, 2017), independent component analysis (Bhinge et al., 2015), and clustering methods (Dahi et al., 2017), play a crucial role in predicting correct categories for abandoned objects. Moreover, some SOTA methods, including Faster RCNN (Ren et al., 2016) and YOLO series (Redmon et al., 2016) are also broadly used for object detection. However, since those methods rely on training on a large amount of manually labeled data and are restricted to limited fore-defined object categories, they do not have inference capability when facing strange classes.

Recently, open-vocabulary object detection algorithms and multi-modal models are proposed for object detection and classification beyond annotated label space. The attractive feature of these algorithms is their promising transferring abilities and robustness. Open-vocabulary object detection methods use visual-related language vocabulary data, e.g., image captions, as auxiliary supervision in open vocabulary settings (Wu et al., 2023) and do not request strict independence between training and testing data, which overcomes the shortcoming of previous zero-shot learning approaches yielding unsatisfying results in novel classes. In image classification tasks, combining simple model structures with visual-language encoders pretrained on large-scale datasets, e.g., CLIP, can lead to a better understanding of semantics on given visual inputs (Ding et al., 2021; Liu et al., 2023) and strong backbone models have been produced by simple pretraining tasks (Radford et al., 2021).

OVD extends its understanding ability to open-world objects by cross-modal aligning image region features and descriptive text of the object to be detected. Thus, it is possible to detect a flexible set of object classes that are only defined at test time via a text query (Chen et al., 2023; Minderer et al., n.d.). In this way, a fixed object category list can be replaced by an open-vocabulary word list. Moreover, without training by a large amount of manually labeled data, introducing multi-modal models like CLIP into object detection tasks, precisely detecting objects and recognizing new classes are realizable.

3. METHODOLOGY

This paper proposes a two-stage open-vocabulary object detection method for highway abandoned object recognition and classification. Meanwhile, a potential abandoned object vocabulary list is built instead of using the public open vocabulary list to reduce negative impacts of unrelated categories and improve model inference efficiency.

3.1 Architecture

As shown in Figure 1, the left side shows the first stage of region proposal, and the right side presents the proposal classification. Firstly, we use Segment Anything (SAM) to get all potential target regions. The region proposal network from Faster R-CNN is also tested but has suboptimal accuracy. Another advantage of using SAM is that these proposed region boundaries are pretty accurate, which do not require further refinement. Then these regions patches cropped from original images are sent into CLIP encoder to generate visual embeddings, while the text encoder converts the pre-defined abandoned object vocabularies to a list of text embeddings. After calculating and ranking similarities, this model predicts target regions as well as their categories.

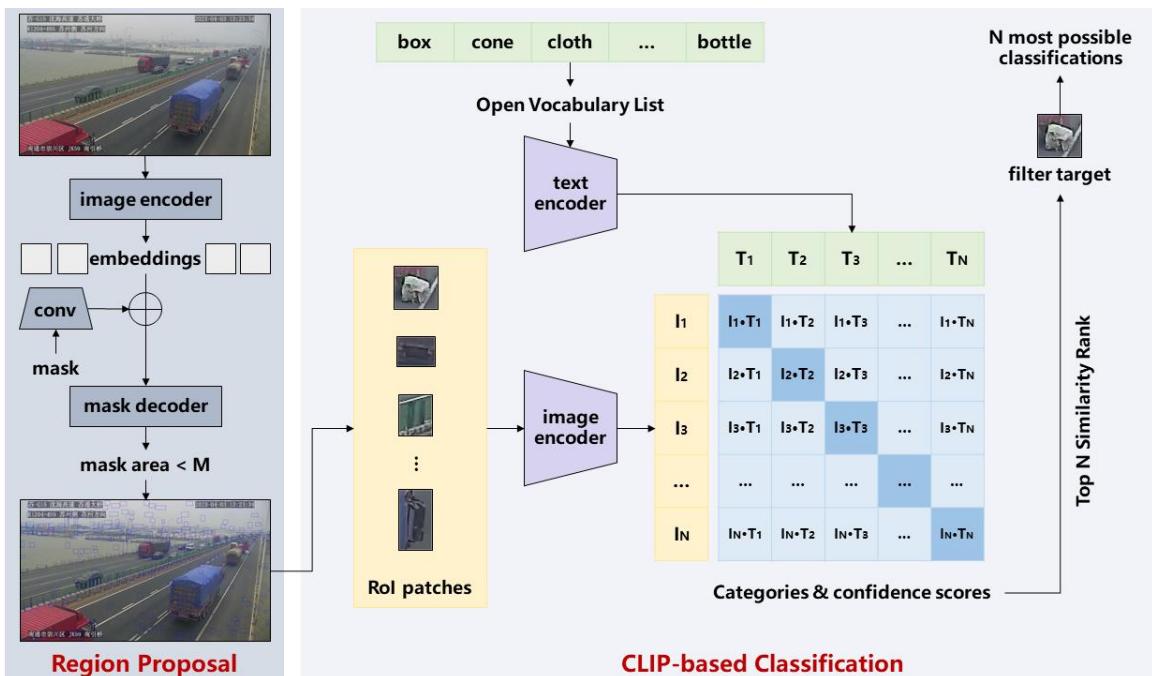


Figure 1: Framework of two-stage open-vocabulary object detection method.

3.2 Two-stage Open-vocabulary Object Detection

3.2.1 Open Vocabulary List

We build up an open vocabulary list for highway abandoned objects based on an existing open image class named *Open Images Dataset V7*, a dataset of ~9M images annotated with image-level labels (61,404,966 image-level labels on 20,638 classes), object bounding boxes (15,851,536 boxes on 600 classes), object segmentation masks (2,785,498 instance segmentations on 350 classes), visual relationships (3,284,280 relationship annotations on 1,466 relationships), and localized narratives (675,155 localized narratives). This vocabulary list determines the relationship of each keyword through super-subordinate relation, part-whole relation, antisense relation, etc., and use the characters of abandoned objects, such as color, shape, material, and scale, to establish a semantic relationship network among words. Moreover, an initial 1423 object categories list is first used for abandoned object recognition, and new classes can be added at test time. These 1423 classes include litter (paper, plastic bags, bottles, cans and other discarded items that people throw away casually), dirt or dust (soil or dust particles produced by road construction or other reasons), building materials (wood, bricks, steel bars and other waste materials), etc.

3.2.2 SAM-based Region Proposal

SAM is a large foundation model for image segmentation (Kirillov et al., 2023), which supports flexible prompts and computes masks in real-time. Trained on 11 million images (Yang et al., 2023), SAM can produce high-quality masks/bounding boxes and do zero-shot segmentation in generic scenarios. Considering its excellent ability in segmentation, SAM can easily find micro-details through low-pixel images, which is helpful to extract all potential abandoned object regions. There are three steps to generate potential regions – SAM model initialization, Intersection over Union (IoU) calculation, and RoI patch proposal.

- 1) **SAM initialization.** This paper uses the pre-trained SAM to automatically search whole images and have to initialize the model by setting a suitable number of search points per side, reserved minimum mask area and other related parameters. All these parameters are manually set according to image pixels and sharpness. Generally, more search points mean more meticulous search as well as more time consumed, and the same is true for other parameters. Thus, choosing proper parameters thoughtfully is important for abandoned object recognition.
- 2) **IoU calculation.** Considering major targets are small objects, it is important to customize a resealable maximum mask area to filter out predicted thousands of masks. Set those mask areas that are smaller than the maximum one as potential target regions, and record these areas' bounding boxes, coordinates, and IoU value, which is a popular metric to measure localization accuracy.
- 3) **RoI patches proposal.** This paper uses Non-Maximum Suppression (NMS) to remove preserved regions and only keep those with higher IoU values than the presupposed threshold that decides the quality of detection results. Those regions are finally proposed as RoI patches, whose bounding boxes and coordinates will be used for object recognition and localization.

3.2.3 CLIP-based Classification

CLIP (Contrastive Language–Image Pre-training) is a pre-trained multi-modal model containing visual and text encoders. Its pretrained encoders are widely used for downstream tasks, including zero-shot learning, natural language supervision, and multi-modal learning (Radford et al., 2021). Using the self-supervised training paradigm to extract features without label information, CLIP is a typical contrastive learning pattern, which is widely applied in image-text matching and classification areas. In this paper, CLIP encoders are used for abandoned object classification tasks.

- 1) **RoI patches recognition and Abandoned object localization.** Put those RoI patches into the pre-trained CLIP encoder, and then, together with the coordinates of patches, extract the most probable abandoned object regions by calculating cosine similarities, and add newly detected words into the open vocabulary list.
- 2) **Abandoned object classification.** Mapping the final regions to the initial object categories list, calculate and sort the possibilities of each category. Take the Top K as the final abandoned object categories. In this paper, K is 10. According to auxiliary information, transportation agencies could make a feasible decision in time, e.g., cleaning the road surface when finding abandoned objects blocking traffic.

4. EXPERIMENTAL DESIGNS

4.1 Dataset

Video Surveillance Dataset. We have extracted 28805 images whose scales are 1920*1080 pixels from highway surveillance videos captured in Nanjing, China. Within these videos, there are 35 traffic incidents related to abandoned objects and 4007 frames containing at least one abandoned object. As shown in Figure 2, the dataset not only contains various light conditions (day and night time), but also has different surveillance angles, including extreme challenging scenarios. For example, Figure 2 (a1) and (b1) contain tiny and ambiguous abandoned objects. Figure 2 (a3) shows objects in dust conditions with low contrast and poor texture. Figure 2 (b2) shows target objects are partially occluded by vehicles.

4.2 Evaluation Metrics

mAP (mean Average Precision). Mean Average Precision is simply the average of AP (Average Precision) of all classes and is calculated by taking the mean AP over all classes and/or over all IoU thresholds (Medewar). It helps to assess the balance between Precision and Recall for an object detection model across various probability thresholds, while a higher mAP value indicates better performance of the object detection model. The formulation of mAP is shown below.

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k \quad (1)$$

In which n means the number of classes we have proposed, and AP_k means the average precision of class k and in practice AP is the precision averaged across all recall values between 0 and 1.

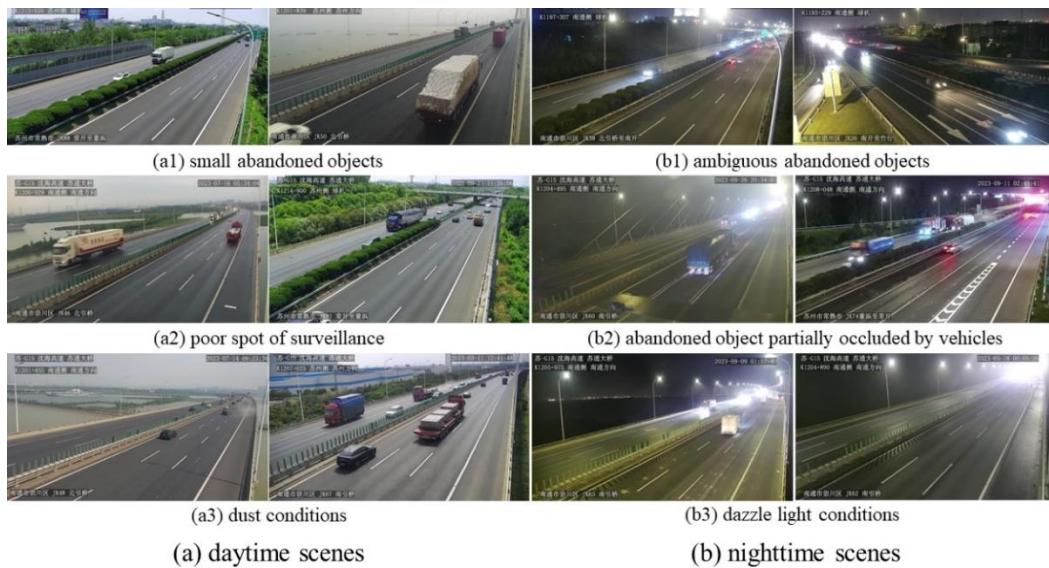


Figure 2: Different incidents and scenes of video surveillance dataset.

Detected rate (DR). Considering we use pre-trained multi-modal models without fine-tuning, counting the rate of abandoned objects being detected is an essential metric to evaluate the performance of our method. And the calculation formula is as follows.

$$DR = \frac{N_{\text{abandoned objects have been detected}}}{N_{\text{abandoned objects}}} \quad (2)$$

In which $N_{\text{abandoned objects}}$ means the number of all abandoned objects, while $N_{\text{abandoned objects have been detected}}$ means the number of abandoned objects that have been detected out.

4.3 Models and Parameters

For region proposal, we employ an automatic SAM, setting 64 search points for each side of images. And the maximum mask area in our tasks is set at 1000, with 0.86 as an IoU threshold. As for the CLIP encoder, we directly use *ViT-B/32* model without fine-tuning and just replace the text-token to our open vocabulary list and input RoI patches to the image encoder. Finally, we take the top ten possible object categories as prediction results. The following experiments are conducted on Linux system with an Intel Xeon Silver 4310 CPU, 125GB RAM, and NVIDIA GeForce RTX 4090 GPU.

To validate the generalization ability of the proposed algorithm, this research compares three representative detection models: Mask R-CNN, OWL-ViT and YOLO. All methods are applied directly without fine-tuning on the target dataset.

Mask R-CNN: a conceptually simple, flexible, and general model for object instance segmentation (He et al., 2018). As an instance segmentation algorithm, Mask R-CNN is a multi-tasking network that is widely used for object detection, object instance segmentation, and object key point detection. The Mask R-CNN model is pre-trained on the COCO dataset (Lin et al., 2015).

OWL-ViT: a recipe for transferring image-text models to open-vocabulary object detection (Minderer et al., 2022). OWL-ViT transfers image-text pre-trained model to OVD tasks and has achieved well-behaved results on both zero-shot text query detection and few-shot image query detection.

YOLO: using a single convolutional network simultaneously to predict multiple bounding boxes and the class probabilities of these boxes(Redmon et al., 2016). YOLO is widely used in object detection area and has been used for real-time detection.

5. EXPERIMENT RESULTS

5.1 Results of Two-stage Open-vocabulary Object Detection

After evaluating the proposed open-vocabulary object detection method on our video surveillance dataset, 200 potential ROI patches would be predicted on every frame under settings in Section 4.3. The prediction results are shown in Figure 3. What's more, 77.8% of small objects can be detected, with 92.9% being classed correctly. Samples of classification results are presented in Figure 4, where top 10 similarities rank are considered. It can be observed that most predicted categories are related based on the object appearance. Another advantage of this method is finding unknown objects in historical videos and data.



Figure 3: Results of SAM-based ROI patches proposal.

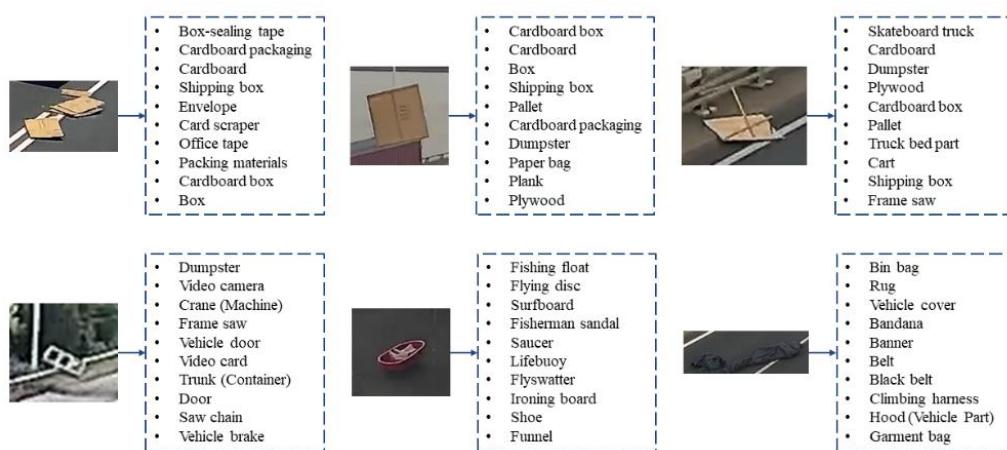


Figure 4: Sample Results of CLIP-based object classification.

5.2 Comparison with SOTA

The comparison results of the three methods are shown in Table 1, while the visual results of Mask RCNN, OWL-ViT and YOLO are shown in Figure 5, Figure 6 and Figure 7 respectively. It illustrates that the two methods are inefficient in handling low-light scenes and can barely recognize small abandoned objects. This may relate to the fact that Mask R-CNN, OWL-ViT and YOLO require on fine-tuning on specific datasets in order to get reasonable results.

Table 1: Comparison results of metrics for different methods.

	DR (%)	mAP (%)
Mask RCNN *	27.8	34.3
OWL-ViT (v1)	11.1	75.0
YOLO	13.9	20.0
Two-stage OVD (this paper)	77.8	92.9

* Since Mask RCNN asked image scales less than 1333 pixels in width and less than 800 pixels in height, we adjust original images (1920*1080) to 960*540 pixels.



Figure 5: Results of Mask RCNN.



Figure 6: Results of OWL-ViT.



Figure 7: Results of YOLO.

6. CONCLUSION

This paper proposes a two-stage open-vocabulary object detection method for highway abandoned object recognition. To avoid shortages of conventional object detection with a fixed range of detection categories, a specific vocabulary list containing all potential categories and multi-modal models, i.e., SAM and CLIP, are combined to precisely recognize abandoned objects even though they do not appear in the historical videos or pre-trained datasets. Compared to SOTA object detection methods, e.g., Mask R-CNN, OWL-ViT and YOLO, this method shows higher detection accuracy, especially in extreme scenes such as small objects and low-light conditions. With good generalization ability and robustness, it is easy to transfer this model to other similar downstream tasks like pedestrian detection. In the future, this method could act as the teacher model to distill conventional models for super-efficient object detection. It will help transportation agencies make road alerts in time to improve traffic safety and find perpetrators.

REFERENCES

- Bhinge, S., Levin-Schwartz, Y., Fu, G.-S., Pesquet-Popescu, B., and Adali, T. (2015). A data-driven solution for abandoned object detection: Advantages of multiple types of diversity. *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 1347–1351.
- Chen, K., Jiang, X., Hu, Y., Tang, X., Gao, Y., Chen, J., and Xie, W. (2023). OvarNet: Towards Open-vocabulary Object Attribute Recognition (arXiv:2301.09506). arXiv. <http://arxiv.org/abs/2301.09506>.
- Cuevas, C., Martinez, R., Berjon, D., and Garcia, N. (2017). Detection of Stationary Foreground Objects Using Multiple Nonparametric Background-Foreground Models on a Finite State Machine. *IEEE Transactions on Image Processing*, 26(3), 1127–1142.
- Dahi, I., Chikr El Mezouar, M., Taleb, N., and Elbahri, M. (2017). An edge-based method for effective abandoned luggage detection in complex surveillance videos. *Computer Vision and Image Understanding*, 158, 141–151.
- Ding, H., Jiang, X., Shuai, B., Liu, A. Q., and Wang, G. (2018). Context Contrasted Feature and Gated Multi-scale Aggregation for Scene Segmentation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2393–2402.
- Ding, H., Liu, C., Wang, S., and Jiang, X. (2021). Vision-Language Transformer and Query Generation for Referring Segmentation (arXiv:2108.05565). arXiv. <http://arxiv.org/abs/2108.05565>.
- He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2018). Mask R-CNN (arXiv:1703.06870). arXiv. <http://arxiv.org/abs/1703.06870>.
- Shet-Talathi, K., and Khobragade, S. V. (2017). A Real Time Abandoned Object Detection Hardware Approach. *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India.
- Kirillov, A., et al. (2023). Segment Anything (arXiv:2304.02643). arXiv. <http://arxiv.org/abs/2304.02643>.
- Kuo, W., Cui, Y., Gu, X., Piergiovanni, A. J., and Angelova, A. (2023). F-VLM: Open-Vocabulary Object Detection upon Frozen Vision and Language Models (arXiv:2209.15639). arXiv. <http://arxiv.org/abs/2209.15639>.

- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollár, P. (2015). Microsoft COCO: Common Objects in Context (arXiv:1405.0312). arXiv. <http://arxiv.org/abs/1405.0312>.
- Liu, C., Ding, H., and Jiang, X. (2023). GRES: Generalized Referring Expression Segmentation. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 23592–23601.
- Minderer, M., et al. (n.d.). *Simple Open-Vocabulary Object Detection with Vision Transformers*.
- Radford, A., et al. (2021). Learning Transferable Visual Models From Natural Language Supervision (arXiv:2103.00020). arXiv. <http://arxiv.org/abs/2103.00020>.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection (arXiv:1506.02640). arXiv. <http://arxiv.org/abs/1506.02640>.
- Ren, S., He, K., Girshick, R., and Sun, J. (2016). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks (arXiv:1506.01497). arXiv. <http://arxiv.org/abs/1506.01497>.
- Saluky, S., Supangkat, S. H., and Nugraha, I. B. (2020). Abandoned Object Detection Method Using Convolutional Neural Network. *2020 International Conference on ICT for Smart Society (ICISS)*, 1–4.
- Tripathi, R. K., Jalal, A. S., and Bhatnagar, C. (2013). A framework for abandoned object detection from video surveillance. *2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, 1–4.
- Wang, G., Zhou, M., Wei, X., and Yang, G. (2023). Vehicular Abandoned Object Detection Based on VANET and Edge AI in Road Scenes. *IEEE Transactions on Intelligent Transportation Systems*, 24(12), 14254–14266.
- Wu, J., et al. (2023). Towards Open Vocabulary Learning: A Survey (arXiv:2306.15880). arXiv. <http://arxiv.org/abs/2306.15880>.
- Yang, J., Gao, M., Li, Z., Gao, S., Wang, F., and Zheng, F. (2023). Track Anything: Segment Anything Meets Videos (arXiv:2304.11968). arXiv. <http://arxiv.org/abs/2304.11968>.
- Zareian, A., Rosa, K. D., Hu, D. H., and Chang, S.-F. (2021). Open-Vocabulary Object Detection Using Captions (arXiv:2011.10678). arXiv. <http://arxiv.org/abs/2011.10678>.

Is ChatGPT a Reliable Source of Transportation Equity Information for Scientific Writing?

Boniphace Kutela, Ph.D., P.E.¹; Shoujia Li²; Subasish Das, Ph.D.³; and Jinli Liu⁴

¹Roadway Safety Program, Texas A&M Transportation Institute, Houston, TX.
Email: b-kutela@tti.tamu.edu

²Texas A&M Univ., College Station, TX. Email: shoujiali@tamu.edu

³Texas State Univ., San Marcos, TX. Email: subasish@txstate.edu

⁴Texas State Univ., San Marcos, TX. Email: jinli.liu@txstate.edu

ABSTRACT

Transportation equity is an interdisciplinary agenda that requires both transportation and social inputs. Traditionally, transportation equity information is sourced from public libraries, conferences, television, and social media, among others. Artificial intelligence (AI) tools, including advanced language models such as ChatGPT, are becoming favorite information sources. However, their credibility has not been well explored. This study explored the content and usefulness of ChatGPT-generated information related to transportation equity. It utilized 152 papers retrieved through the Web of Science (WoS) repository. The prompt was crafted for ChatGPT to provide an abstract given the paper's title. The ChatGPT's abstracts were then compared to human-written abstracts using statistical tools and unsupervised text mining. The results indicate a weak similarity between ChatGPT and human-written abstracts. On average, the human-written and ChatGPT-generated abstracts were about 58% similar, with a maximum and minimum of 97% and 1.4%, respectively. The keywords from the abstracts of papers with over the mean similarity score were more likely to be similar, whereas those below the average score were less likely to be similar. Themes with high similarity scores include access, public transit, and policy, among others. Contrarily, the findings from collocated keywords were inconclusive. The study findings suggest that ChatGPT has the potential to be a source of transportation equity information. However, currently, a great amount of attention is needed before a user can utilize materials from ChatGPT.

Keywords: Transportation equity; Artificial intelligence; ChatGPT

INTRODUCTION

The use of artificial intelligence (AI) is radically altering how individuals live their lives and carry out the many activities they do daily. People's ways of living and working have been revolutionized because of the widespread adoption of cutting-edge technologies such as smartphones and smart watches. In addition, the use of voice command devices like SIRI and Alexa has transformed the day-to-day activities that people participate in.

In November 2022, OpenAI released ChatGPT, an advanced language model, that interacts with users by providing a set of written directives and produces the written text according to the instruction given (Noever & Ciolino, 2022). This tool has gained the attention of experts in different fields, such as academicians, economists, social scientists, engineers, and computer scientists (ChatGPT & Perlman, 2022; Gao et al., 2022). Most of the concerns surrounding this

tool have been about whether ChatGPT will replace some human-generated activities, such as writing codes/algorithms, preparing poems, movie transcripts, etc. (Qadir, 2022). Some scholars even argue that certain jobs will be replaced by ChatGPT (Qadir, 2022), whereas others disagree by indicating that the tool is not capable of taking over most human-generated jobs and thus will have minimal impact (Aydin & Karaarslan, 2022; Frye, 2022). However, ChatGPT is based on the advanced language model, which uses reinforcement learning, so it is expected to get better when new observations are included in retraining the model.

There has been discussion from researchers and the public about the replacement of Google with ChatGPT (Bindra, 2023; Greene, 2023). Google searches have been used by various researchers as the starting point for searching for various information. A key question among researchers has been whether the future of Google searches is on the brink of collapsing. Due to its ability to create informative texts given a certain prompt, researchers and the public have been using ChatGPT as a source of information. Users have prompted topics related to politics, social science, comedy, business, art, healthcare, games, coding, computer science, and transportation, among others (Kim & Lee, 2023; Yalalov, 2022). In the transportation field, public transportation accessibility and affordability have been among a few areas where users have been searching for information using ChatGPT (Mobility-Innovators, 2023). Public transportation accessibility and affordability form a part of transportation equity, which is among the hot topics both in the United States and globally.

Transportation equity is among the emerging topics of interest in the transportation field. The concept of transportation equity pertains to guaranteeing fair access to transportation alternatives for all individuals, regardless of their socioeconomic position or other individual circumstances (di Ciommo & Shiftan, 2017). With the global trend of urbanization leading to a surge in population and associated problems such as traffic gridlock and environmental pollution, transportation equity is gaining prominence as a pressing issue in various regions. Transportation equity is crucial as it has a pivotal role in advancing social and economic inclusivity (di Ciommo & Shiftan, 2017; Martens et al., 2019). Access to transportation is critical for individuals to access essential services like education, healthcare, and employment opportunities. Inadequate transportation infrastructure or unaffordable transportation options can impede people's ability to access such opportunities (Martens et al., 2019). By enhancing transportation choices and ensuring equal access, transportation equity can contribute to reducing poverty, improving social mobility, and boosting economic development.

Considering the importance of transportation equity, various outlets have been used as sources of information for this topic. Such sources include the National Association of City Transportation Officials (NACTO), the National Center for Mobility Management (NCMM), the American Public Transportation Association (APTA), and the U.S. Department of Transportation (USDOT). Other sources of information for transportation equity issues include the Web of Science, Google Scholar, and Transportation research in Transport Research International Documentation (TRID) (Clarivate, 2023; TRID, 2023). The information included in these sources include policy papers, design guides, toolkits, webinars, policy briefs, and case studies that address transportation equity issues (APTA, 2023; NACTO, 2023; NCMM, 2023; USDOT, 2023). Information from these sources has been used by researchers and practitioners to address transportation equity issues. Further, researchers utilize such sources for manuscript preparation for publication.

Recently, ChatGPT has been acquired by Microsoft and added to the Bing website, whereby users can chat with it and obtain text-based answers for whatever they are looking for. Like other

search engines, ChatGPT can be biased, error-prone, or provide misinformation in its content. There have been some complaints about biased information from ChatGPT in various fields, such as political science, earth science, and social science (Alba, 2023; Bass, 2023; Biddle, 2023). The key difference between ChatGPT to other sources is that other search engines, such as Google search, provide the user with multiple sources of the queried information, whereas ChatGPT gives the user a single document with all the information in it. Thus, a user has no opportunity to filter misinforming content. In addition to researchers, the general community relies on search engines to obtain various information. With how easy it is to use ChatGPT, several people might be influenced to use it for day-to-day searches. Thus, the information from this source should be error-free and consistent with other trusted sources to maintain a clear understanding of various issues within the community.

Despite the debate of the biases, misinformation, and errors from ChatGPT, efforts to evaluate the extent of similarity and the content of similar and dissimilar information have still not been done to a great extent. A few studies have attempted to understand the difference between published articles and ChatGPT-generated materials. (Gao et al., 2022; Kutela, Msechu, et al., 2023). These studies concluded that ChatGPT is a good tool for academic writing but needs more human inputs to make the content more human-like. However, these studies did not provide comparable statistical analysis for the two sources of information. In this study, the abstracts of manuscripts were used to explore the difference between human-generated texts in peer-reviewed journals and ChatGPT-generated texts. The intention was to evaluate whether the information from the two sources, human and AI, was similar, and if so, to what extent and what specific content was more similar than others. Further, this study added to the body of literature regarding the methodological approach needed to explore the key difference between human-generated scientific content and ChatGPT texts. It is found that ChatGPT is still not well adept in developing correct citations. In many of the lengthy scientific writing generated by ChatGPT generates either non-existing or fake citations. This study thus explores only abstracts (abstracts are citation free) to investigate the authenticity of the contents generated by ChatGPT.

The remaining sections are presented as follows. The next section presents the study methodology and discusses the data description and analytical approaches. The results and discussion section follows, then the conclusion and future studies are presented last.

METHODOLOGY

As described earlier, this study intends to explore whether ChatGPT can produce publication-ready materials comparable to human-generated text in published journals. This section presents the methodological approach used to attain the study objectives. The section is divided into two main sections, data description and analytical methods.

Data Description

To assess the similarities of the information retrieved from different sources, two types of data are necessary, human-written text and ChatGPT-generated text. In this study, authors utilized the abstract section of published papers as the human-written text.

Authors extracted transportation equity papers from the Web of Science database in which “transport equity” and “transportation equity” keywords were used to obtain the transportation equity-related manuscripts. A total of 251 manuscripts containing these keywords in their

abstracts were extracted. Upon further pre-processing and duplicate checking, a total of 152 manuscripts were retained for further analysis. The abstract sections, titles, author keywords, and year of publication are the few variables of interest available in the downloaded data. Most of these papers were published in the Transportation Research Record, Transportation Research Part D-Transport and Environment, the Journal of Transport Geography, Transport Policy, and Transportation Research Part A-Policy and Practice.

To obtain the corresponding abstracts generated by ChatGPT, a prompt, which contains the directives of the actions to be taken by ChatGPT, is necessary. The following prompt was used for retrieval:

"I want you to develop write an unstructured abstract with minimum of 300 words and maximum of 500 words for publication in a scientific journal that focuses on transport equity. I will give you several titles then I want you to give me the unstructured abstract. You need to adopt a persona of high-profile researcher in transport equity, who has exceptional writing skills and has published over 100 manuscripts from various parts of the world. The abstract should have the details for at least introduction, objectives, methodology, key findings, and study implications. The first title is "Title of the paper."

The authors supplied all 152 titles to ChatGPT. The corresponding abstracts generated by ChatGPT were extracted and matched to the human-written abstracts in an Excel sheet for further analysis.

Analytical Methods

Two analytical methods, document similarity analysis and text network analysis were applied to the text data to explore the similarities and content of the transportation equity information from two sources. The document similarity analysis shows the similarity scores between the abstracts generated by ChatGPT and those written by a human. However, it does not describe the content of such abstracts. To explore the similarities and differences in the content the text network analysis was applied.

Document Similarity Analysis

Document similarity analysis is used to determine the similarities between different documents. Normally, researchers and the public consider documents to be similar if they are semantically close and describe similar concepts/themes (Selivanov, 2018; Zach, 2020).

In this study, the document is defined as the abstract. The comparison is between the abstracts written by humans to the corresponding abstracts written by ChatGPT. For each given title of the study, the human-written abstract and ChatGPT-generated abstracts are extracted and stored in the spreadsheet.

Various approaches, such as Jaccard distance, Cosine distance, Euclidean distance, and Relaxed Word Mover's Distance, can be applied to determine document similarity (Selivanov, 2018; Zach, 2020). This study used Cosine similarity with the LSA approach to determine similarities between two corresponding abstracts.

First, the documents were transformed into bag-of-words to compute the similarity between the two documents, so each document will be a sparse vector. Thus, the similarity between two documents (abstracts) can be computed as:

$$\text{Similarity } (doc_1, doc_2) = \text{Cos}\theta = \frac{doc_1 \cdot doc_2}{|doc_1||doc_2|}$$

Whereby doc_1 is the human-written abstract of the paper and doc_2 is the ChatGPT-generated abstract of the same paper. The similarity score varies between 0 and 1, where a score of 1 implies that the two documents are duplicates and a score of zero implies that the two documents are not similar at all (Zach, 2020).

Text Network Analysis

Text Network Analysis (TNA) has been utilized in various fields such as literature and linguistics (Hunter, 2014), traffic safety and operations (Kutela, Das, et al., 2021; Kutela & Teng, 2021; Kwayu et al., 2021), and bibliometrics in transportation studies (Jiang et al., 2020). TNA uses nodes and edges to establish relationships between keywords within a corpus (see Figure 1), and its strength lies in its ability to visualize keywords and establish connections among them (Jiang et al., 2020; B. Kutela et al., 2021; Boniphace Kutela et al., 2021; Paranyushkin, 2011). The frequency and co-occurrence of keywords in the network are represented by the size of the nodes and the edges, respectively.

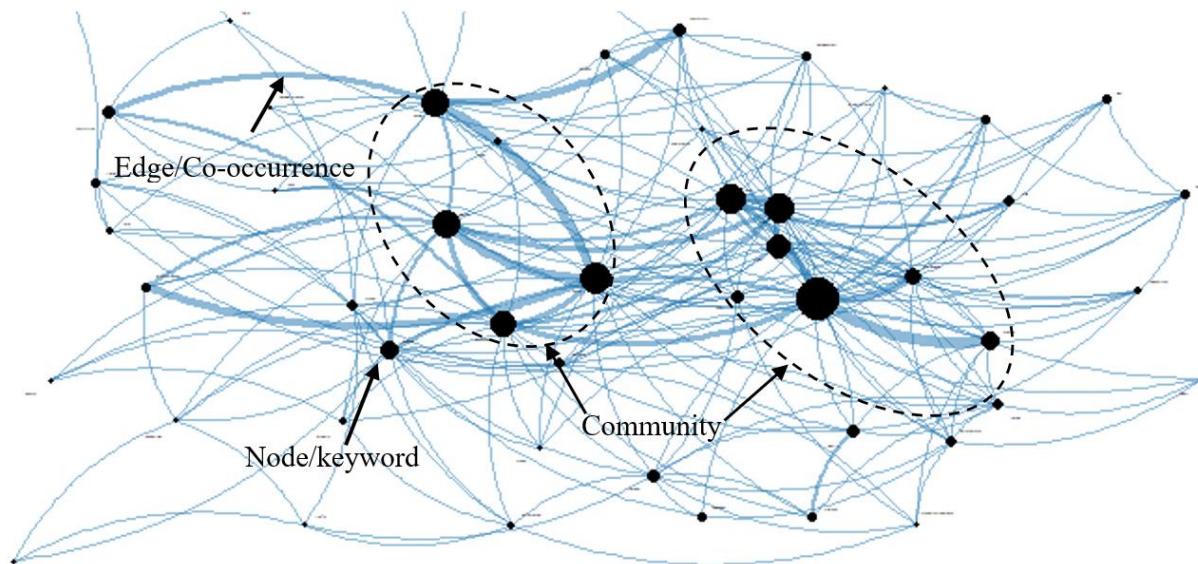


Figure 1. A Skeleton of the Text Network

Various processes are performed on the data when conducting TNA analysis. Normalization is the first process, whereby unstructured data is converted to structured data, all symbols are removed, and all texts are converted to lowercase then the output is used to create a matrix of keywords with their frequencies of occurrence (Das & Dutta, 2020; Kutela, Magehema, et al., 2022; Kutela, Novat, et al., 2022). The constructed matrix is then visualized with keywords represented as variously sized nodes based on their recorded frequencies. Various metrics can be utilized for comparative analysis, but document and collocated frequency (Kutela, Combs, et al.,

2022; Kutela, Kitali, et al., 2023) are the two metrics used in this study to compare human-generated and ChatGPT-generated introductions. Document frequency is the number of documents that contain the keyword of interest. Keyword frequency, on the other hand, focuses on the number of times the keyword appears in the document. Prior to producing keyword frequency, text stemming was necessary to reduce keywords to their roots/base form. Text stemming allows for better comparison and matching of words with the same root, even if they have different suffixes (Das & Dutta, 2020). For example, "run," "runs," "running," and "runner" can all be stemmed to "run," making it easier to identify all occurrences of the root word. Collocation frequency assesses the number of times the keywords are located next to each other and offers greater insights than individual keywords do since it focuses more on the closeness between two keywords in the corpus. The collocation of keywords in a text network plays a great role in the formation of text clusters, typically referred to as a community of keywords. A community of keywords represents a group collectively clustered in the text network; there can be two or more in a text network (see Figure 1).

RESULTS AND DISCUSSION

This section presents the results and discussion. It covers the document similarity results and the text network of the abstracts.

Document Similarity Results

Table 1 presents the similarity scores between human-written and ChatGPT-generated abstracts. It is observed that the abstract is divided into six different metrics used to measure quantitative assessments. The similarity score is measured by how much they are similar on a scale of zero to one. However, some algorithms may produce negative scores, indicating dissimilarity. The similarity score increases from the minimum to the maximum category as more abstracts are developed. For example, there is a notable change between the minimum and the first quartile. On average, the content from human-written and ChatGPT-generated abstracts is 58% similar. The table also shows that the minimum similarity score is 1.4% and the maximum similarity score is nearly 97%. In the statistic, the median is 62%, in which the first quartile is 41% and the third quartile is 80%.

Table 1. Similarity Scores between Human-Written and ChatGPT-generated Abstracts

Metrics	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
	Similarity Score	0.01402	0.41	0.61618	0.58428	0.79738
						0.96602

Although the document similarity scores present the relationship between the two documents being compared, it does not show the content of the text. Thus, text network analysis was applied to explore the pattern of the key themes of the abstracts. To perform content analysis, the documents were divided into two parts, the high and low similarity scores. The high similarity score document included all documents with similarity scores above the mean, while the low similarity score document included all documents with mean or lower similarity scores. The next section presents the content analysis results.

Text Network Results of the Abstracts

The content of the abstracts written by humans and by ChatGPT was evaluated. Figure 2 presents the text network of the human-written abstract section and ChatGPT-generated abstract section with high and low similarity scores, while Table 2 and Table 3 present the performance metrics of the text networks. The text network of human-written abstracts with high similarity scores (Figure 2 (a)) is heavily centered on the keywords *accessibility*, *public*, *transit*, *travel*, and *planning*. This is because the studies used in this paper are transportation-related. Further, the human-written abstract sections constitute keywords such as *plan*, *safety*, *environmental*, and *travel*, indicating the abstract also covers travel concerns in various studies. Some other keywords have a relatively low representation, including *time*, *regional*, *commuting*, *mobility*, and *services*, among others.

Figure 2 (b) presents the text network for ChatGPT-generated abstracts with high similarity scores. The text network is heavily centered on the keywords *accessibility*, *urban*, *transit*, *plan*, and *public*, similar to the human-written text network. However, the human-written text network nodes appear to be larger than those presented on the ChatGPT network. This outcome implies that these keywords appear more frequently in human-written abstract sections than in the ChatGPT-generated ones.

Table 2. Topmost Keywords from Abstracts with High Similarity Scores

Keywords				Collocations							
GPT			Human			GPT			Human		
Feature	Freq	Docfreq	Feature	Freq	Docfreq	Collocation	Count	Collocation	Count	Collocation	Count
access	431	113	access	225	67	key findings	139	public transit	39		
use	288	114	use	212	105	highlights importance	55	21st century	25		
find	274	151	public	211	69	public transit	66	act 21st	23		
need	258	117	transit	187	51	provides valuable	59	planning process	15		
includ	249	129	plan	149	64	literature review	49	flexible efficient	15		
provid	245	124	travel	145	51	valuable insights	49	new york	15		
polici	233	106	servic	118	46	highlights need	39	accountable flexible	15		
data	224	112	polici	110	63	case studies	28	safe accountable	15		
transit	216	38	urban	107	52	importance considering	51	efficient act	14		
implic	215	148	model	106	44	qualitative quantitative	27	legacy users	14		
plan	203	67	develop	106	66	current state	26	act legacy	14		
public	202	63	analysi	106	60	conclusion provides	32	land use	14		
key	196	141	research	100	58	findings indicate	66	job accessibility	13		
urban	192	64	system	100	54	important implications	35	disadvantaged			
import	189	117	area	94	51	implications suggest	34	groups	12		
communiti	188	61	result	88	70	methodology includes	26	travel demand	12		
improv	187	81	group	85	48	inform development	24	public services	12		
research	186	89	studi	85	37	population groups	24	environmental			
servic	186	54	social	84	44	policies programs	25	review	11		
analysi	178	96	mobil	83	38	review existing	24	york city	10		
								autonomous			
								vehicles	10		
								united states	10		

Observing the right side of the ChatGPT-generated text network with high similarity scores (Figure 2 (b)), the linked keywords include *accessibility* and *key*, *low* and *income*, *communities* and *income*, and *public* and *transit*. These linked keywords can be related to transportation accessibility and the impact of income on transportation. On the other hand, the left side of the

text network contains linked keywords such as *key* and *finding*, *survey* and *data*, *data* and *analysis*, *literature* and *review*, *data* and *used*, and *survey* and *methods*, which are related to analysis and recommendations/findings.

Table 3. Topmost Keywords from Abstracts with Low Similarity Scores

Keywords				Collocations					
GPT		Human		GPT		Human			
Feature	Fre q	Docfre q	Feature	Fre q	Docfre q	Collocation	Coun t	Collocation	Coun t
find	98	55	use	76	37	key findings	50	low income	24
implic	77	54	polici	45	30	public transit	33	public transit	22
object	67	53	public	86	28	low income	33	e scooters	9
key	74	52	result	35	28	findings indicate	27	land use	9
includ	88	48	transit develo	110	27	united states	25	subway accessibility	8
need	89	47	p	42	27	provides valuable	23	public services	8
provid	92	46	citi	41	27	literature review	21	95 ci	8
data	82	43	analysi	45	26	best practices	20	irr 95	8
import	66	43	access	63	24	highlights importance	19	rural areas	7
use	100	42	servic	56	22	highlights need	19	lower income	7
highlight	57	42	incom	39	21	valuable insights	19	planning processes	7
access	163	40	improv	32	21	importance considering	18	bike sharing	7
polici	106	40	provid	24	21	ride hail	18	new york	7
signific	56	38	plan	62	20	road safety	17	transit pass	7
analysi	72	37	mobil	47	20	marginalized communities	16	last mile	7
methodolo	43	37	data	26	20	current state	15	act 21st	7
g	59	36	system	39	19	conclusion provides	15	21st century	7
develop	48	36	paper	23	19	mass transit	15	shared mobility	6
literatur	73	34	urban	41	18	mixed methods	14	york city	6
research	40	34	area	39	18	public services	14	first last	6

Note: Docfreq = is the document frequency, which is the number of documents

Figure 2 (c) presents the text network for human-written abstracts with low similarity scores. The text network is heavily centered on the keywords *transit*, *public*, *income*, *mobility*, and *access*, similar to the text network with high similarity scores. However, the text network nodes appear to be larger than those presented on the human-written network with high similarity scores. This outcome implies that these keywords appear more frequently in human-written sections with higher similarity scores than in the sections with low similarity scores. The methodology keywords, *research*, *analysis*, and *model*, are also presented in this section.

Figure 2 (d) presents the text network for ChatGPT-generated abstracts with low similarity scores. The text network is heavily centered on the keywords *safety*, *transit*, *urban*, *planning* and *public*, similar to the human-written text network. Observing the left side of the ChatGPT-generated text network with high similarity scores (Figure 2 (d)), the linked keywords include *low* and *income*, *communities* and *income*, and *public* and *transit*. These linked keywords are similar to the text network for ChatGPT-generated abstracts with high similarity scores. On the other hand, the right side of the text network also contains linked keywords such as *key* and *finding*, *identify* and *data*, *data* and *analysis*, *literature* and *review*, *approach* and *used*, and *data* and *used*, which are related to terms of analysis and recommendations/findings.

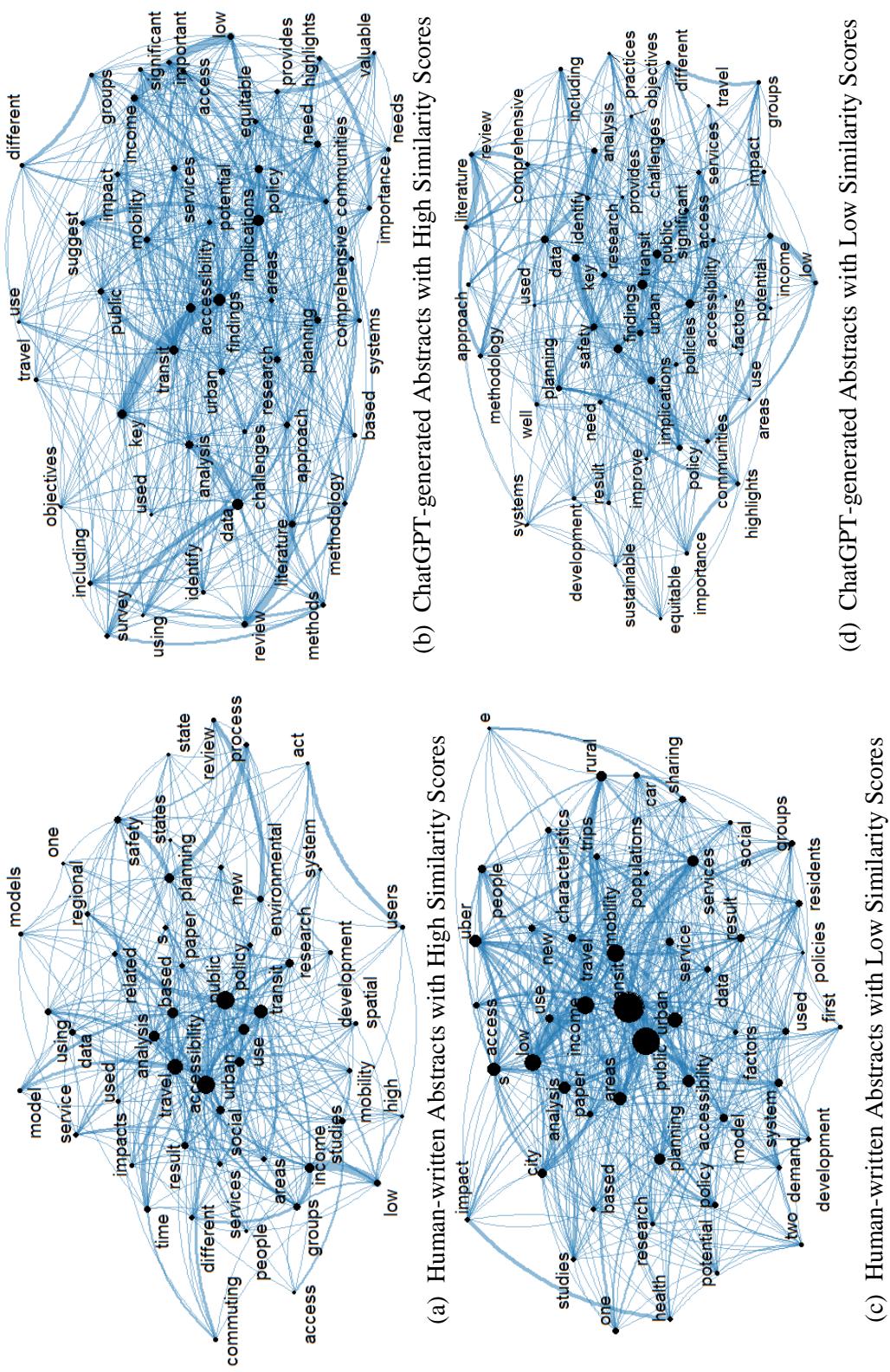


Figure 2. Text Networks of the Content of the Abstracts

Although Figure 2 shows that the four networks share some similarities and portray some differences, a comparative analysis of the four networks can be performed using the keyword and collocation frequencies. According to the results in Table 2 in the group with high similarity scores, among the top 20 keywords, ten are common for both sides. Even for the ten common keywords, the ranking varies significantly. For instance, the keyword *transit* in ChatGPT-generated metrics is ranked ninth, appearing in 216 abstracts, while it is ranked fourth in the human-written metrics, appearing 187 times. Keywords such as *travel, model, system, group*, etc., appear only in the human-written metrics. On the other hand, the keywords *find, need, data, key*, and *community*, among others, appear only in the ChatGPT-generated metrics.

Table 3 shows similar results to the metrics generated by ChatGPT and humans with low similarity scores. There are only seven keywords common for both sides with significantly different ranks, like *polici* which is ranked third in human-written abstracts, appearing 86 times, while only ranked 13th in ChatGPT-generated abstracts, appearing 106 times. Keywords such as *find, object, key* and *include*, among others, appear only in ChatGPT-generated metrics. Keywords *result, income, transit*, and *improve*, among others, appear only in human-written metrics. This observation indicates that ChatGPT is capable of replicating some keywords.

In addition to the individual keywords, the collocated keywords results can be used to distinguish ChatGPT-generated abstracts from human-written abstracts. The results in Table 2 and Table 4 show that the two approaches differ significantly. In Table 2, which shows the metrics with high similarity scores, only one pair of collocated keywords, *public transit*, is common for both approaches. According to the results in Table 3, there are three pairs of collocated keywords that are common for both approaches with low similarity scores, *public transit, low income* and *public services*. The keywords associated with the use of the study can be observed in the collocated keywords. Such keywords include *key findings, highlight importance, conclusion provides, findings indicate*, etc. These findings indicate that the ChatGPT algorithm tends to summarize the conclusions and the possible use of the research, while that is not the case for most human-written scientific writeups. The text network and associated metrics provide the details of the content of the abstracts, which facilitates the comparison of the two.

CONCLUSIONS AND FUTURE STUDIES

This study presents a comparative analysis of transportation equity-related themes by considering the source of information. The two sources compared in this study are the traditional source, which is human-written materials, and AI-generated materials. A total of 152 highly cited papers in transportation equity were extracted from the Web of Science. A prompt was then prepared and supplied to ChatGPT to produce abstracts given the title of the manuscripts. Document similarity analysis and text network analysis, were applied to determine the relationship between human-written and ChatGPT-generated content.

It was found that on average, the human-written and ChatGPT-generated content are about 54% similar, with a minimum score of 1.4% and a maximum score of 96.7%. The content analysis of the abstracts showed a significant difference. In this case, the ChatGPT-generated materials appear to be more generic, especially when the collocation metric is considered.

Based on these findings, this study concludes that at this moment the transportation equity information retrieved through ChatGPT greatly differs from the human-written content. Therefore, researchers and the general community should exercise care when retrieving

transportation equity information from ChatGPT. It is advised that researchers perform a comprehensive review of the information retrieved through ChatGPT with other human-written information. Although the two sources differ significantly, there were observed a few cases where the resulting content had high similarity scores. This is an indication that there are a few topics within transportation equity that ChatGPT performed relatively better. Thus, for such topics, the information retrieved can be used even without conducting a comprehensive review.

Although this study successfully showed the similarities and differences between the information retrieved through ChatGPT against the one written by humans, several limitations exist. First, this study utilized transportation equity-related studies. These studies are not very common for the general community, thus, the chance that their key themes were used to train ChatGPT is relatively lower. That being the case, retrieving such information using ChatGPT is relatively lower. Future studies may consider studies in areas that are common to the general community, such as social science studies. Further, this study considered the abstract section of the published papers. Abstracts tend to contain various information including the objective, methodology, and key findings. It is relatively difficult for even a human to provide an abstract when given a title. Future studies may focus on other easily generatable parts of the manuscripts given a title, such as the study objectives.

REFERENCES

- Alba, D. (2023). ChatGPT, Open AI's Chatbot, Is Spitting Out Biased, Sexist Results - Bloomberg. Bloomberg. <https://www.bloomberg.com/news/newsletters/2022-12-08/chatgpt-open-ai-s-chatbot-is-spitting-out-biased-sexist-results#xj4y7vzkg>.
- APTA (American Public Transportation Association). (2023). Home - American Public Transportation Association. <https://www.apta.com/>.
- Aydin, Ö., and Karaarslan, E. (2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. SSRN Electronic Journal.
- Bass, D. (2023). ChatGPT maker OpenAI promises to fix the chatbot's biases | Fortune. <https://fortune.com/2023/02/16/chatgpt-openai-bias-inaccuracies-bad-behavior-microsoft/>.
- Biddle, S. (2023). The Internet's New Favorite AI Proposes Torturing Iranians and Surveilling Mosques. The Intercept. <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>.
- Bindra, J. (2023). Will ChatGPT replace Google as our go-to web search platform? | Mint. <https://www.livemint.com/opinion/columns/will-chatgpt-replace-google-asour-go-to-web-search-platform-11671733523981.html>.
- ChatGPT, O. A. A., and Perlman, A. (2022). The Implications of OpenAI's Assistant for Legal Services and Society. SSRN Electronic Journal.
- Clarivate. (2023). Web of Science Platform - Web of Science Group. <https://clarivate.com/webofsciencegroup/solutions/webofscience-platform/>.
- Das, S., and Dutta, A. (2020). Characterizing public emotions and sentiments in COVID-19 environment: A case study of India.
- di Ciommo, F., and Shiftan, Y. (2017). Transport equity analysis. *Transport Reviews*, 37(2), 139–151.
- Frye, B. L. (2022). Should Using an AI Text Generator to Produce Academic Writing Be Plagiarism? <https://papers.ssrn.com/abstract=4292283>.

- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., and Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *BioRxiv*, 2022.12.23.521610.
- Greene, T. (2023). ChatGPT will not replace Google Search | Impact of Social Sciences. <https://blogs.lse.ac.uk/impactofsocialsciences/2023/01/27/chatgpt-will-not-replace-google-search/>.
- Hunter, S. (2014). A Novel Method of Network Text Analysis. *Open Journal of Modern Linguistics*, 4, 350–366.
- Jiang, C., Bhat, C. R., and Lam, W. H. K. (2020). A bibliometric overview of Transportation Research Part B: Methodological in the past forty years (1979–2019). *Transportation Research Part B: Methodological*, 138, 268–291.
- Kim, J., and Lee, J. (2023). How does ChatGPT introduce transportation problems and solutions in North America? SSRN Electronic Journal.
- Kutela, B., Combs, T., John Mwekh'iga, R., and Langa, N. (2022). Insights into the long-term effects of COVID-19 responses on transportation facilities. *Transportation Research Part D: Transport and Environment*, 111, 103463.
- Kutela, B., Das, S., and Dadashova, B. (2021). Mining patterns of autonomous vehicle crashes involving vulnerable road users to understand the associated factors. *Accident Analysis & Prevention*, 106473.
- Kutela, B., Kitali, A. E., Kidando, E., Langa, N., Novat, N., and Mwende, S. (2023). Exploring commonalities and disparities of seattle residents' perceptions on dockless bike-sharing across gender. *City, Culture and Society*, 32, 100503.
- Kutela, B., Langa, N., Mwende, S., Kidando, E., Kitali, A. E., and Bansal, P. (2021). A text mining approach to elicit public perception of bike-sharing systems. *Travel Behaviour and Society*, 24, 113–123.
- Kutela, B., Magehema, R. T., Langa, N., Steven, F., and Mwekh'iga, R. J. (2022). A comparative analysis of followers' engagements on bilingual tweets using regression-text mining approach. A case of Tanzanian-based airlines. *International Journal of Information Management Data Insights*, 2(2), 100123.
- Kutela, B., Msechu, K., Das, S., and Kidando, E. (2023). Chatgpt's Scientific Writings: A Case Study on Traffic Safety. SSRN Electronic Journal.
- Kutela, B., Novat, N., Adanu, E. K., Kidando, E., and Langa, N. (2022). Analysis of residents' stated preferences of shared micro-mobility devices using regression-text mining approach. *Transportation Planning and Technology*, 45(2), 159–178.
- Kutela, B., Novat, N., and Langa, N. (2021). Exploring geographical distribution of transportation research themes related to COVID-19 using text network approach. *Sustainable Cities and Society*, 67.
- Kutela, B., and Teng, H. (2021). The Use of Dynamic Message Signs (DMSs) on the Freeways: An Empirical Analysis of DMSs Logs and Survey Data. *Journal of Transportation Technologies*, 11(01), 90–107.
- Kwayu, K. M., Kwigizile, V., Lee, K., Oh, J.-S., and Oh, J.-S. (2021). Discovering latent themes in traffic fatal crash narratives using text mining analytics and network topology. *Accident Analysis and Prevention*, 150, 105899.
- Martens, K., Bastiaanssen, J., and Lucas, K. (2019). Measuring transport equity: Key components, framings and metrics. *Measuring Transport Equity*, 13–36.

- Mobility-Innovators. (2023). How ChatGPT will be game changer for Public Transit agencies? - Mobility Innovators. <https://mobility-innovators.com/how-chatgpt-will-be-game-changer-for-public-transit-agencies/>.
- NACTO. (2023). National Association of City Transportation Officials | National Association of City Transportation Officials. <https://nacto.org/>.
- NCMM. (2023). Home - National Center for Mobility Management. <https://nationalcenterformobilitymanagement.org/>.
- Noever, D., and Ciolino, M. (2022). The Turing Deception. ArXiv.
- Paranyushkin, D. (2011). Identifying the Pathways for Meaning Circulation using Text Network Analysis. *Venture Fiction Practices*, 2(4). www.noduslabs.com.
- Qadir, J. (2022). *Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education*.
- Selivanov, D. (2018). Documents similarity. <https://text2vec.org/similarity.html>.
- TRID. (2023). About TRID | Information Services. <https://www.trb.org/InformationServices/AboutTRID.aspx>.
- USDOT. (2023). Equity | US Department of Transportation. <https://www.transportation.gov/priorities/equity>.
- Yalalov, D. (2022). 100 Best ChatGPT Prompts to Unleash AI's Potential | Metaverse Post. <https://mpost.io/100-best-chatgpt-prompts-to-unleash-ais-potential/>.
- Zach. (2020). How to Calculate Jaccard Similarity in R. <https://www.statology.org/jaccard-similarity-in-r/>.

Enhancing AV Safety: A Bagging Classifier Approach for Predicting Crash Outcomes

Sai Sneha Channamallu¹; Deema Almaskati²; Sharareh Kermanshachi, Ph.D., P.E.³;
and Apurva Pamidimukkala, Ph.D.⁴

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.

Email: sxc5928@mavs.uta.edu

²Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.

Email: deemanabeel.almaskati@mavs.uta.edu

³Associate Vice Chancellor and Associate Dean of Research, Pennsylvania State Univ., State College, PA. Email: svk5464@psu.edu

⁴Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

ABSTRACT

Safety is the predominant concern in the advancement of automated vehicles (AVs); therefore, extensive real-world testing is underway to ensure their secure operation. Despite the widespread belief that they will decrease the frequency of accidents, there remains uncertainty about their impact on the severity of crashes in which they are involved. The primary objective of this study is to use the bagging classifier technique to predict the likelihood of injuries in accidents involving AVs. This was accomplished by conducting an in-depth examination of a wide range of independent variables and an analysis of injuries sustained in crash incidents involving AVs from 2014 to July 2023. The bagging classifier model showed notable effectiveness, achieving a balanced accuracy of 0.59, along with high precision and recall values of 0.94 and 0.97, respectively. These metrics indicate the model's strong capability for accurately identifying severe crash outcomes and minimizing false positives. The precision-recall curve and a modified F1 score of 2.39 further endorse the model's performance, particularly highlighting its efficiency in handling the class imbalance present in the dataset. The validation and learning curves underscore the model's optimal complexity, displaying its proficiency in identifying essential patterns without succumbing to overfitting. Collectively, these metrics underscore the model's success in predicting injury outcomes in AV crashes with a high level of accuracy. This study contributes to the literature on AV safety by providing valuable insights for manufacturers and policymakers that will enable them to develop effective safety features and strategies, thereby enhancing traffic safety.

KEYWORDS: Autonomous vehicles, Crash injury, Predictive modeling, Crash patterns, Machine learning, Bagging classifier.

INTRODUCTION

Autonomous vehicles (AVs) have the potential to transform safety in the transportation sector by minimizing human involvement, the primary factor contributing to vehicle accidents (Novat et al., 2023). They are also expected to enhance mobility, decrease emissions and energy consumption, and reduce congestion (Song et al., 2021; Houseal et al., 2022; Pamidimukkala et al., 2023a). Autonomous driving systems are intended to support and assist human drivers by

managing and controlling driving tasks (Favarò et al., 2017; Etminani-Ghasroddashti et al., 2022a). The Society of Automotive Engineers (SAE) outlines six levels of autonomy (L0-L5) that assess how much support and assistance AVs can provide for driving tasks (Das et al., 2020; Channamallu et al., 2023; Khan et al., 2022a). Though it is generally accepted that they hold great promise for reducing injuries and fatalities due to roadway accidents, more testing is necessary to gain a better understanding of their capabilities and to pinpoint areas that require technological and operational improvements prior to commercialization (Kutela et al., 2022; Pamidimukkala et al., 2023b; Patel et al., 2022a). Automated driving systems are currently being tested globally, including in California, where the state has mandated that all manufacturers testing on public roads must submit reports on disengagements and collisions to the California Department of Motor Vehicles (CA DMV) (Ye et al., 2021; Khan et al., 2022b; Etminani-Ghasroddashti et al., 2022b).

Researchers globally have utilized a wide range of statistical models and descriptive analyses to evaluate the CA DMV data. For instance, Novat et al. (2023) developed a Bayesian network, using the Markov Chain Monte Carlo to establish a comparative analysis between collisions of AVs and conventional vehicles, and Ding et al. (2023) evaluated the variables impacting crash severity by utilizing a multinomial logit model. Other researchers investigated crash severity factors, utilizing ordinal regression, binary logistic regression, and probit models. While the findings of these studies are essential for a better understanding of the performance of AVs in a mixed traffic environment, statistical models can often be characterized by imbalanced data. This study attempts to address this issue by incorporating class weighting strategies. Sinha et al. (2021) is one of the few who has developed a crash model utilizing a bagging classifier; however, they only assessed datapoints from 2014 to 2019, whereas this paper provides a comprehensive assessment of the data from 2014 to July 2023.

The primary objective of this study is to develop a bagging classifier crash model that can predict the outcome of injuries resulting from collisions involving AVs. This paper will contribute to the current literature on autonomous driving systems and CA DMV data by providing a comprehensive analysis of AV crash predictions, using a bagging classifier model. The inferences from the model will be beneficial to AV manufacturers for creating more adaptable and responsive safety features, which may lessen the severity of crashes and lead to increased traffic safety in general. The findings will also equip transportation professionals and policymakers with a thorough understanding of AV collisions and help them develop measures and policies that will reduce their severity.

LITERATURE REVIEW

In 2014, the CA DMV authorized manufacturers to test AVs on public roads to expedite the development of autonomous driving systems (Ye et al., 2021; Khan et al., 2023a; Patel et al., 2023a). Since then, more crashes and disengagements have been reported to the DMV due to the increase of AV roadway testing, and researchers have analyzed the data to evaluate AV performance.

Previous studies have employed a range of analytical techniques, including a hierarchical Bayesian approach (Ren et al., 2022; Patel et al., 2022b), classification tree (Wang and Li, 2019; Zhu and Meng, 2022; Etminani-Ghasroddashti et al., 2023), and probit model (Yuan et al., 2022; Khan et al., 2023b); others have attempted to determine the risk factors connected to AV collisions. Leilabadi and Schmidt (2019) examined how environmental conditions affect the

performance of AVs, and the results showed that there is no significant relationship between the severity of crashes and the roadway conditions or the weather (Leilabadi and Schmidt, 2019; Khan et al., 2022c). The lack of correlation may not be representative of the performance of AVs in a different climate, however, since California is not characterized by its adverse weather conditions. Similarly, there was minimal association with lighting conditions, a factor which Ye et al. (2021) considered to be the most significant in traffic injuries. The Pearson Chi-Square test conducted by Leilabadi and Schmidt (2019) found roadway surfaces to have the strongest correlation with AV performance. Houseal et al. (2022) examined AV collisions and, through logistic regression and decision trees, identified the type of collision and movements of AVs and other vehicles as significant contributing factors to AV-involved accidents. Similarly, Wang and Li (2019) utilized crash database and statistical modeling approaches such as ordinal logistic regression and CART classification tree to investigate the contributing factors and mechanisms of AV crashes and found that crash severity significantly increased when the AV was primarily responsible for the collision. Wang and Li (2019) and Patel et al (2023b) determined that individuals are more likely to suffer severe injuries when accidents occur on highways, possibly due to the high travel speeds. In contrast, Favarò et al. (2017) found that only 20% of accidents occur on highways; most (48%) had taken place on suburban roads at low relative speeds. By employing text analytics and a hierachal Bayesian regression to analyze crash reports, Boggs et al. (2020) concluded that AVs are more likely to be involved in a collision in areas of mixed land use in comparison to other land uses, such as public areas and school zones.

METHODOLOGY

The methodology adopted in this study to develop a bagging classifier model to predict the outcomes of injuries sustained in AV crashes is presented in Figure 1. Data collection and data processing, including cleaning, feature elimination, and transformation of categorical data into numerical formats were employed, and evaluation metrics were selected to provide a nuanced assessment of the model's predictive performance in the context of skewed data distributions.



Figure 1. Methodology

Data Collection and Pre-processing

The study leverages the dataset originally compiled from the CA DMV, mandating that companies testing AVs on public roads submit accident reports in PDF format within 10 business days. The process included accumulating specifics from every PDF report on the CA DMV's website, spanning from 2014 to July 2023. After extracting these details, they were methodically compiled into an Excel file for a numerical examination. Figure 1 presents the yearly distribution of AV crashes in the data set.

To ensure the relevance and quality of the data, it was meticulously cleaned and preprocessed, including the removal of irrelevant features such as exact location coordinates that

do not directly contribute to crash outcomes. Missing values were addressed by excluding records with significant gaps, and categorical variables were transformed into numerical formats using techniques such as one-hot encoding, to facilitate computational analysis and model training. Figure 3 presents a correlation matrix of the selected dependent variables.

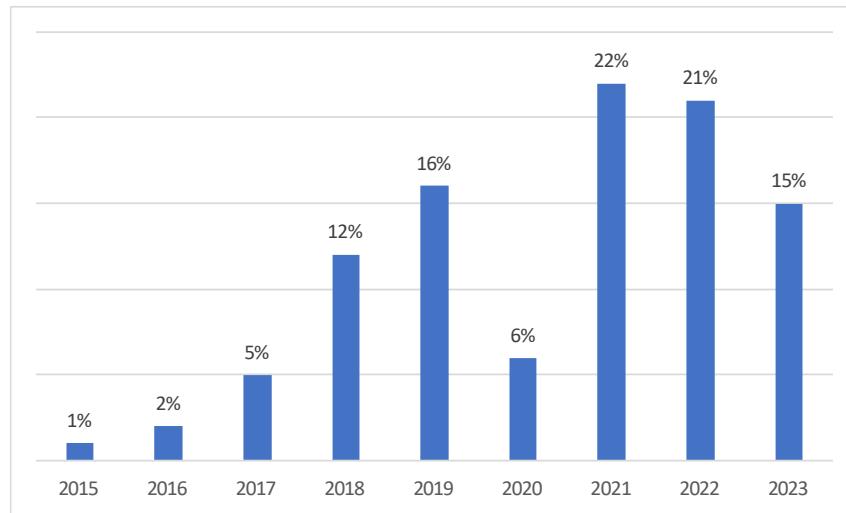


Figure 2. Yearly distribution of AV crashes

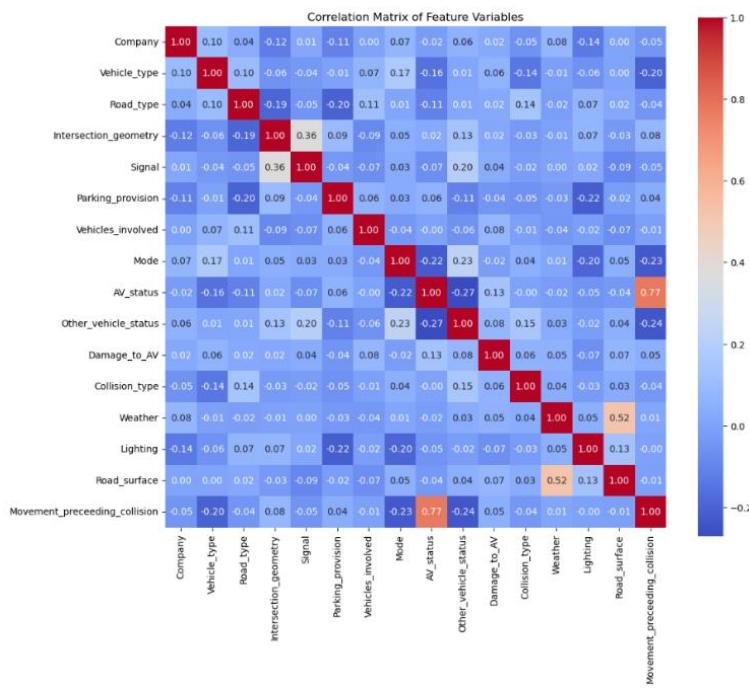


Figure 3. Correlation matrix of variables

The feature selection for the model was informed by the relevance and contribution of various factors to crash outcomes. These factors include the extent of vehicle damage, the AV manufacturer, the nature of the collision, and environmental conditions like weather and lighting,

along with the time attributes of the crash. The correlation matrix presents a quantitative analysis of how different variables relate to each other, ranging from technical aspects like vehicle type and AV status, to environmental conditions such as weather and lighting. Variables such as the status and mode of the AV have a high positive correlation (0.77), suggesting that the autonomous mode of the vehicle is strongly related to its operational status. Hence, mode was excluded from the variables.

Model Development

The crash prediction model that was developed included one dependent and 15 independent variables including vehicle-related, environmental, traffic-related and contextual factors for forecasting injury outcomes in crashes. The dependent variable was binary, coded as 0 for crashes without injuries and 1 for those with injuries. There was a notable imbalance in the sample sizes of the two categories, with 331 non-injury incidents compared to just 27 injury incidents. To counter this imbalance, the model employed class weighting techniques, assigning weights to each class inversely related to their occurrence in the dataset. The weighting formula shown in Equation 1 factors in the total number of samples (n_{sample}), the number of classes in the dataset ($n_{classes}$), and the count of each class ($np.bincount(y)$).

$$\text{Class weights of a specific class } y = \frac{n_{sample}}{n_{classes} \times np.bincount(y)} \quad (1)$$

The dataset was divided into training and testing portions, with a distribution of 80% for training purposes and 20% for testing. Stratified sampling was applied to counteract the imbalances in class distribution and maintain equal class proportions in both training and testing sets. A bagging classifier was selected as the main predictive model due to its ability to handle complex data and resist overfitting. During training, the model's parameters, including the number of decision trees and their depth, were fine-tuned through grid search and cross-validation, ensuring a balance between bias and variance.

Evaluation Metrics

The model's performance was assessed, using sophisticated metrics such as balanced accuracy that better reflect its effectiveness in datasets with skewed distributions than traditional accuracy measures. Precision and recall were deemed crucial metrics, given the importance of accurately predicting crash severity, and they provided insights into the model's proficiency in accurately identifying true positive cases, while reducing false negatives and positives. The F1-Score, which is the harmonic mean of precision and recall, was utilized to offer a comprehensive perspective on the model's predictive capabilities. Minimizing false positives was a key focus of the model, given the significant implications of incorrectly predicting non-injury incidents. Therefore, the study adopted an adapted evaluation approach that uses the modified F1 score as the principal metric, as recommended in prior research (Li et al., 2018; Sinha et al., 2021). This score effectively balances precision and recall, and provides reliable predictions across both classes, a crucial aspect in handling imbalanced datasets. The computation for the modified F1 score is detailed in Equation 2 where β is a parameter; When β is a parameter with a value between 0 and 1, the focus is more on precision, whereas a β greater than 1 indicates that recall is more important. In this study, a β of 0.5 is employed to place greater importance on precision.

$$\text{Modified F1 score} = (1+\beta^2) \times \frac{\text{precision} \times \text{recall}}{\beta^2 \times (\text{precision} + \text{recall})} \quad (2)$$

DISCUSSION

The bagging classifier's performance in predicting AV crash outcomes is presented in Table 1. The model achieved a balanced accuracy of 0.59, which is particularly significant as it indicates its ability to maintain an equilibrium between sensitivity (true positive rate) and specificity (true negative rate). Balanced accuracy is critical in AV crash predictions due to the dual importance of correctly identifying both non-severe and severe outcomes.

Table 1. Results of bagging classifier model

Confusion matrix		Evaluation metrics				
True labels	Predicted labels	Balanced accuracy	Precision	Recall	F1 score	Modified F1 score
Crash w/o Injury	Crash w/o Injury	0.59	0.94	0.97	0.96	2.39
Crash w/ Injury	Crash w/o Injury	65	4	1	2	

The precision score of 0.94 demonstrates the bagging classifier's adeptness for accurately classifying positive instances (severe crash outcomes) and is indicative of the model's ability to minimize false positives - a key factor in avoiding unnecessary alarms or interventions. The recall score of 0.97 highlights the model's proficiency in detecting actual positive cases and is essential in AV safety contexts, where failing to identify a severe crash outcome could have critical implications. The combination of high precision and recall in the bagging classifier is particularly noteworthy, as it suggests that the model is not only good at identifying severe crash outcomes when they occur but also in correctly recognizing non-severe outcomes, which minimizes the risk of misclassification. The modified F1 score of 2.39, adjusted to account for class imbalance, stands out as a commendable achievement. This adjustment is crucial because conventional F1 scores may not fully capture the model's performance in datasets where one class (e.g., severe crashes) is significantly underrepresented. This high modified F1 score indicates that the bagging classifier is effective in balancing precision and recall, even in an imbalanced dataset. This is a significant strength, as imbalanced data is a common challenge in AV crash datasets where severe outcomes are, fortunately, less frequent.

Threshold Analysis

The precision-recall curve presented in figure 4 is an essential visualization tool that offers profound insights into the predictive performance of our bagging classifier model, as it reflects the classifier's efficacy in distinguishing between severe and non-severe crash outcomes at varying threshold levels, which is paramount in the domain of AV safety. A notable observation from the precision-recall curve is the high precision that the model maintains even at high recall

levels, which is not commonly observed in typical models, especially in imbalanced datasets like AV crash data.

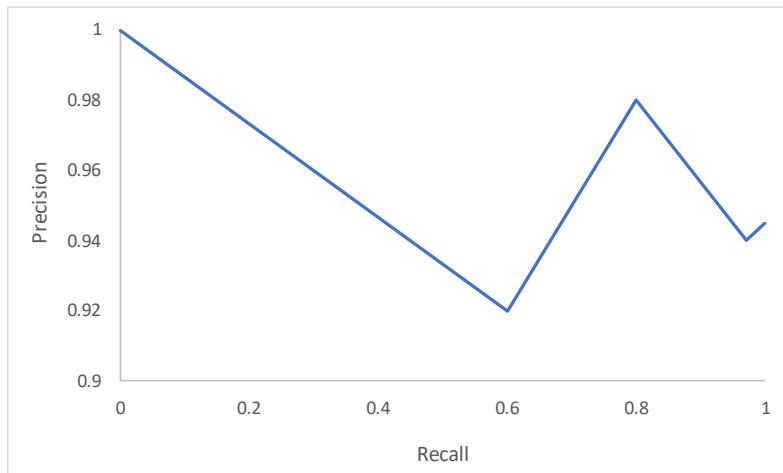


Figure 4. Precision-recall curve

The shape of the curve also provides valuable information. The initial high precision can be particularly useful for settings where the cost of false positives is high, but as we move towards the right on the curve, increasing recall, the precision moderately declines, indicating a trade-off that must be managed. The specific point on the curve where we decide to operate the model can be chosen based on the relative costs of false negatives (missed severe crashes) versus false positives (unnecessary interventions). The curve's progression also presents a nuanced picture of the classifier's performance across different operational points. In some regions, we observe that the precision can be recovered slightly after initial drops, which suggests that the classifier's performance is not monotonically decreasing with respect to recall; rather it exhibits a complex relationship with the threshold settings. This is indicative of the model's sophistication and its ability to adapt to the intricacies of the data.

Validation Curve

The validation curve presented in figure 5 further substantiates the efficacy of the model in predicting AV crash outcomes. The curve plots the training and cross-validation scores as a function of the number of estimators, providing insight into the model's learning dynamics and its generalization capabilities. The training score remains high as the number of estimators increases, suggesting that the model has sufficient capacity to capture the underlying patterns in the training data. Notably, there is a marginal diminishing return on accuracy as the number of estimators grows beyond 10, indicating that adding more estimators beyond this point does not significantly enhance the model's ability to fit the training data. The cross-validation score exhibits a more nuanced behavior. Starting from a lower bound, the accuracy improves sharply as the number of estimators increases to 5, signifying that the model benefits from additional complexity up to a certain point. Beyond this point, however, the curve begins to plateau, with a slight undulation that suggests an optimal range rather than a singular optimal point. This behavior is characteristic of the bagging classifier's ability to stabilize its predictions without

overfitting the data as more estimators are introduced, as is evidenced by the convergence of training and cross-validation accuracy.

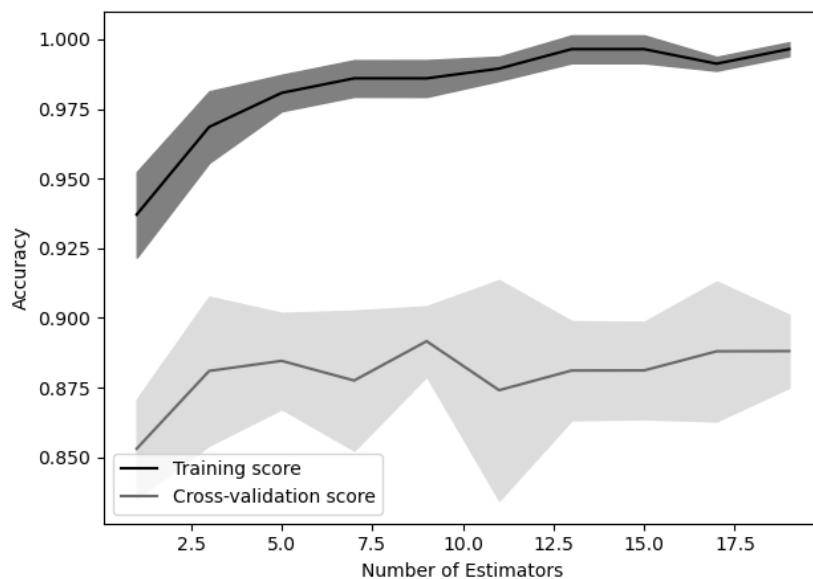


Figure 5. Validation curve

The shaded area representing the variability of the cross-validation score indicates the confidence interval of the generalization performance. The relatively tight confidence intervals around the cross-validation score as compared to the training score further emphasize the model's stability. It is worth noting that the model does not exhibit overfitting, as the training and cross-validation scores remain close across the spectrum of complexity. This demonstrates its robustness and ability to generalize to new, unseen data.

Learning Curve

The learning curve presented in figure 6 provides additional depth to our understanding of the bagging classifier's performance characteristics. It illustrates the relationship between the training size and the model's score for both the training and cross-validation sets and makes it evident that the training score starts high (close to 1.00) and remains relatively stable regardless of the increase in training size. This is indicative of the model's ability to learn effectively from the training data. The consistently high score across different training sizes underscores the model's low bias and its ability to capture the underlying data distribution without overfitting, as the training score does not decrease with the addition of more training data.

The cross-validation score starts at a lower point when the training size is small, suggesting underfitting where the model is not complex enough to capture the underlying pattern in the data. As the training size increases, however, the cross-validation score improves significantly, indicating that with more data, the model is better able to generalize its predictions to unseen data. The increase in the cross-validation score with training size is an encouraging sign of the model's learning capability. Notably, there is a dip in the cross-validation score as the training size increases from 100 to 150. This could be attributed to several factors, such as variability in the data, a particular subset of data that is more difficult to learn, or simply the randomness

inherent in the cross-validation process. However, as the training size further increases, the cross-validation score recovers, suggesting that the model is able to overcome these issues with more data.

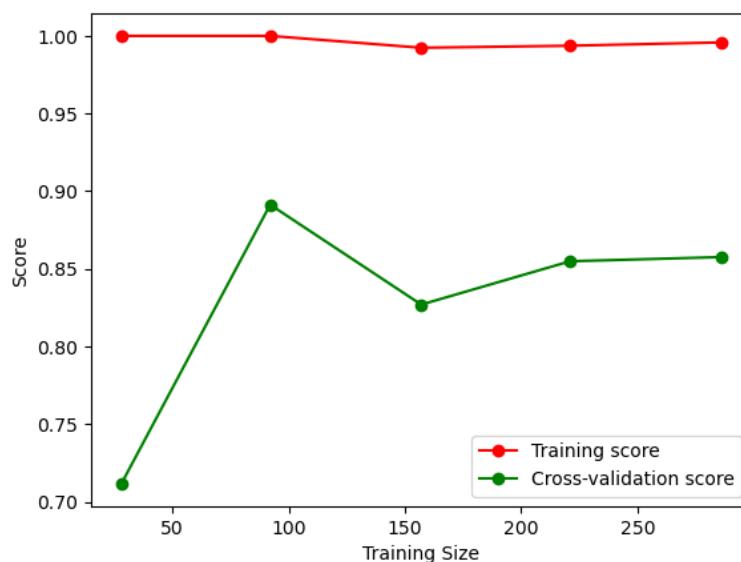


Figure 6. Learning curve

The plateauing of the cross-validation score for larger training sizes suggests that additional data beyond this point does not result in overfitting. This is a desirable property in a predictive model, as it indicates that the model has reached a level of complexity sufficient to capture essential patterns in the data, without being swayed by noise or outliers. The learning curve, therefore, complements our previous discussion on the precision-recall and validation curves by highlighting the bagging classifier's consistent performance and its ability to learn effectively as more data is provided.

CONCLUSION

This study uses a bagging classifier model to present a comprehensive analysis of the prediction of injuries from crashes involving AVs. The methodology, which encompasses data preprocessing, model development, and rigorous evaluation metrics, successfully addresses key challenges in AV crash data analysis, notably class imbalance and data complexity. The choice of a bagging classifier was a strategic decision that was based on its effectiveness with complex and imbalanced datasets, and it adeptly managed the data imbalance through class weighting and stratified sampling, ensuring equitable class representation.

The comprehensive evaluation of the bagging classifier model through various performance metrics and curves demonstrates its substantial potential for reliably predicting AV crash outcomes. The balanced accuracy of 0.59, while modest, ensures an equitable consideration of both severe and non-severe crash predictions, which is crucial for real-world applications where both classes are of significant interest. The precision of 0.94 and recall of 0.97, as illustrated by the precision-recall curve, underscore the model's ability to accurately identify severe crashes while minimizing false positives. This balance is crucial in AV contexts, where the costs of

misprediction can range from unnecessary alarms to missed opportunities for preventing accidents. The precision-recall curve not only validates the high performance of our bagging classifier, but also provides strategic guidance on selection of the threshold. The high modified F1 score of 2.39, which was adjusted for the class imbalance, also reflects the curve's indication that the model effectively balances the precision and recall, making it a trustworthy tool for critical safety applications. The validation curve suggests that the model is sufficiently complex to capture the necessary patterns in the data, while avoiding both underfitting and overfitting. The learning curve further reinforces these findings, showing that the model's ability to generalize improves with more data but stabilizes, indicating that the model is capturing the essential structure of the problem without being unduly influenced by noise.

Overall, the ability of the bagging classifier's performance to handle the intricacies of imbalanced AV crash data with a high degree of precision and recall positions it as a promising tool for enhancing safety in the burgeoning field of AVs. As autonomous technology continues to evolve, the deployment of such advanced predictive models will play a pivotal role in mitigating risks and improving outcomes in AV operations. The insights gained from the various performance evaluations will guide future refinements of the model, ensuring that they continue to serve the needs of an industry where the stakes are nothing short of life-altering.

The main limitation of this study is its reliance on geographically limited data from the CA DMV. Future studies could explore integrating additional data sources, real-time data, and varying environmental factors to further enhance the model's predictive capabilities. Investigating the application of the bagging classifier in different geographic and regulatory contexts would provide a more comprehensive understanding of its effectiveness across diverse environments.

ACKNOWLEDGMENT

We gratefully acknowledge the support and generosity of the Transportation Consortium of South-Central States (Tran-SET), without which the present study could not have been completed.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Boggs, A. M., Wali, B., and Khattak, A. J. 2020. Exploratory analysis of automated vehicle crashes in California: A text analytics & hierarchical Bayesian heterogeneity-based approach. *Accident Analysis & Prevention*, 135, p.105354.
- Channamallu, S. S., Kermanshachi, S., and Pamidimukkala, A. 2023. Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles. In *International Conference on Transportation and Development 2023* (pp. 39-50).
- Das, S., Dutta, A., and Tsapakis, I. 2020. Automated vehicle collisions in California: Applying Bayesian latent class model. *IATSS research*, 44(4), pp.300-308.

- Ding, S., Abdel-Aty, M., Wang, D., Wang, Z., Zheng, O., Barbour, N., and Yuan, R. 2023. Exploratory Analysis of the Crash Severity between Vehicular Automation (SAE L2-5) with Multi-Source Data.
- Etminani-Ghasrodeshti, R., Patel, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022a. Exploring Factors Affecting Willingness to Ride Shared Autonomous Vehicles. In *Proceedings of Transportation Research Board 101 Annual Conference*.
- Etminani-Ghasrodeshti, R., Ketankumar Patel, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022b. Modeling users' adoption of shared autonomous vehicles employing actual ridership experiences. *Transportation Research Record*, 2676(11), 462-478.
- Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2023. Exploring motivating factors and constraints of using and adoption of shared autonomous vehicles (SAVs). *Transportation Research Interdisciplinary Perspectives*, 18, 100794.
- Favarò, F. M., Nader, N., Eurich, S. O., Tripp, M., and Varadaraju, N. 2017. Examining accident reports involving autonomous vehicles in California. *PLoS one*, 12(9), p.e0184952.
- Houseal, L. A., Gaweesh, S. M., Dadvar, S., and Ahmed, M. M. 2022. Causes and effects of autonomous vehicle field test crashes and disengagements using exploratory factor analysis, binary logistic regression, and decision trees. *Transportation research record*, 2676(8), pp.571-586.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Identifying Usage of Shared Autonomous Vehicles (SAVs): Early Findings from a Pilot Project. In *Transportation Research Board 101st Annual Meeting* Washington, DC.
- Khan, M. A., Etminani-Ghasrodeshti, R., Shahmoradi, A., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022b. Integrating shared autonomous vehicles into existing transportation services: evidence from a paratransit service in Arlington, Texas. *IJCE*, 20(6), 601-618.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2022c. Do ridesharing transportation services alleviate traffic crashes? A time series analysis. *Traffic injury prevention*, 23(6), 333-338.
- Khan, M. A., Patel, R. K., Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., Hladik, G., and Foss, A. 2023a. Factors that determine a university community's satisfaction levels with public transit services. *Frontiers in Built Environment*, 9, 1125149.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2023b. Understanding Students' Satisfaction with University Transportation. In *International Conference on Transportation and Development 2023* (pp. 522-532).
- Kutela, B., Avelar, R. E., and Bansal, P. 2022. Modeling automated vehicle crashes with a focus on vehicle at-fault, collision type, and injury outcome. *Journal of transportation engineering, Part A: Systems*, 148(6), p.04022024.
- Leilabadi, S. H., and Schmidt, S. 2019. In-depth analysis of autonomous vehicle collisions in California. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (pp. 889-893). IEEE.
- Novat, N., Kidando, E., Kutela, B., and Kitali, A. E. 2023. A comparative study of collision types between automated and conventional vehicles using Bayesian probabilistic inferences. *Journal of safety research*, 84, pp.251-260.
- Pamidimukkala, A., Patel, R. K., Kermanshachi, S., Rosenberger, J. M., and Tanvir, S. 2023a. A Review on Shared Mobility and Electric Vehicles. In *International Conference on Transportation and Development 2023* (pp. 333-342).

- Pamidimukkala, A., Kermanshachi, S., and Patel, R. 2023b. An Exploratory Analysis of Crashes Involving Autonomous Vehicles. In *International Conference on Transportation and Development 2023* (pp. 343-350).
- Patel, R. K., Etminani-Ghasrodishti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Exploring willingness to use shared autonomous vehicles. *International Journal of Transportation Science and Technology*, 12(3), 765-778.
- Patel, R. K., Etminani-Ghasrodishti, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022b. How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems, ASCE International Conference on Transportation & Development*, pp. 81-93.
- Patel, R., Channamallu, S. S., Khan, M. A., Kermanshachi, S., and Pamidimukkala, A. 2023a. An Exploratory Analysis of Temporal and Spatial Patterns of Autonomous Vehicle Collisions. *Public Works Management & Policy*. DOI:
- Patel, R. K., Pamidimukkala, A., Kermanshachi, S., and Etminani-Ghasrodishti, R. 2023b. Disaster preparedness and awareness among university students: a structural equation analysis. *International journal of environmental research and public health*, 20(5), 4447.
- Ren, W., Yu, B., Chen, Y., and Gao, K. 2022. Divergent effects of factors on crash severity under autonomous and conventional driving modes using a hierarchical Bayesian approach. *International journal of environmental research and public health*, 19(18), p.11358.
- Sinha, A., Vu, V., Chand, S., Wijayaratna, K., and Dixit, V., 2021. A crash injury model involving autonomous vehicle: Investigating of crash and disengagement reports. *Sustainability*, 13(14), p.7938.
- Song, Y., Chitturi, M. V., and Noyce, D. A. 2021. Automated vehicle crash sequences: Patterns and potential uses in safety testing. *Accident Analysis & Prevention*, 153, p.106017.
- Wang, S., and Li, Z. 2019. Exploring the mechanism of crashes with automated vehicles using statistical modeling approaches. *PloS one*, 14(3), p.e0214550.
- Xu, C., Ding, Z., Wang, C., and Li, Z. 2019. Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. *Journal of safety research*, 71, pp.41-47.
- Ye, W., Wang, C., Chen, F., Yan, S., and Li, L. 2021. Approaching autonomous driving with cautious optimism: analysis of road traffic injuries involving autonomous vehicles based on field test data. *Injury prevention*, 27(1), pp.42-47.
- Yuan, Q., Xu, X., Wang, T., and Chen, Y. 2022. Investigating safety and liability of autonomous vehicles: Bayesian random parameter ordered probit model analysis. *Journal of intelligent and connected vehicles*, 5(3), pp.199-205.
- Zhu, S., and Meng, Q. 2022. What can we learn from autonomous vehicle collision data on crash severity? A cost-sensitive CART approach. *Accident Analysis & Prevention*, 174, p.106769.

Exploring Riders' Preferences of Using Shared Autonomous Vehicles

Apurva Pamidimukkala, Ph.D.¹; Sharareh Kermanshachi, Ph.D., P.E., F.ASCE²;
Jay Michael Rosenberger, Ph.D.³; and Ann Foss, Ph.D.⁴

¹Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Penn State Univ., Middletown, PA (corresponding author). Email: svk5464@psu.edu

³Professor, Dept. of Industrial, Manufacturing, and Systems Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: jrosenbe@uta.edu

⁴Office of Strategic Initiatives, City of Arlington, Arlington, TX.
Email: ann.foss@arlingtontx.gov

ABSTRACT

Public opinion regarding autonomous vehicles (AV) heavily influences how quickly the technology will be implemented and adopted in the future. This research utilized data collected from riders of Rideshare Automation Payment Integration Demonstration (RAPID), a shared autonomous vehicle (SAV) service in Arlington, Texas, and performed exploratory data analysis to identify the monthly ridership trends of the riders. Within the scope of this project, AVs provided service to the UTA campus and downtown Arlington. The findings revealed that students are the most frequent riders of the RAPID service when compared to non-students. It was also revealed that the usage frequency of RAPID service was more on Tuesdays and Wednesdays when compared to other working days of the week. Additionally, the findings indicate that the service was less utilized during early mornings before 9:00 am and evenings after 6:00 pm when compared to other times of the day. This study offers critical insights toward ridership trends of shared AVs that will assist local, state, and federal transit authorities and planners in formulating policies and transportation strategies to target SAV ridership when the service is more widely offered.

INTRODUCTION

Self-driving vehicles are a promising technology that can reduce urban mobility and accessibility barriers by providing effective transportation options (Etminani-Ghasrodasthi et al., 2021; Khan et al., 2022; Patel et al., 2021; Pamidimukkala et al., 2023a). Rapid advances in technology and self-driving shuttles offer shared rides to citizens who have limited access to transportation, such as elderly people, individuals with disabilities and low-income people (Krueger et al. 2016; Etminani-Ghasrodasthi et al., 2022; Khan et al., 2023a). By providing first/last mile rides on low-demand routes, shared autonomous vehicles (SAVs) can be a complimentary option to existing public transportation services (Khan et al., 2023b; Patel et al., 2022).

Currently in the United States, 41 states and DC have enacted AV legislation, 6 states have issued executive orders, and 5 states have passed laws and executive orders related to autonomous driving, allowing testing or deployment of autonomous vehicles on public roads (NCSL 2020). In addition, multiple Level 4 AV pilot projects have been initiated in various

cities throughout the United States, like Portland, OR; Chandler, AZ; San Jose, CA; Boston, MA; Pittsburgh, PA; and Arlington, TX (Perkins et al. 2018; Channamallu et al., 2023; Etmian-Ghasrodasti et al., 2023). The SAV pilot projects are usually begun by a public-private partnership (Kerlin 2019) and when integrated with existing on-demand ridesharing services have the potential to expand accessibility to individuals with different mobility needs. Accordingly, public inclination, attitudes, preferences, and adoption of AV and SAV technology can result in the successful implementation of pilot projects.

Understanding public acceptance and adoption of such vehicles is essential; consequently, the research on this topic is rapidly growing (Acheampong and Cugurullo 2019; Etzioni et al. 2021; Yuen et al. 2020; Pamidimukkala et al., 2023b; Patel et al., 2023a). Although empirical evidence of the association among factors such as attitudes and preferences towards AV technology and acceptance of this new mobility mode has been substantiated, several gaps in the literature exist.

Most studies explore the public's interest in vehicle technology based on data gathered from surveys (Haboucha et al. 2017; Krueger et al. 2016; Bansal et al. 2016; Lu et al. 2017; Patel et al., 2023b), and they simulate and evaluate the integration of AVs into an existing public transit system (Shen et al. 2018; Wen et al. 2018; Levin et al. 2019; Pamidimukkala et al., 2023c). However, it is still uncertain whether the results and findings about the adoption of SAVs through simulation methods will be consistent when autonomous vehicles operate on the road in the near future. Additionally, most of the existing literature is based on the researchers conducting an investigation into the acceptance of SAV technology by modeling data collected from a sample of people with no SAV riding experience (Dichabeng et al. 2021; Hossain and Fatmi 2022; Fu et al. 2021; Pamidimukkala et al., 2023d; Wang et al. 2020). This is inevitable, due to the limited number of AVs and SAVs on the road, which means that due to the intangibility of this new technology, the results from the literature regarding perceptions towards SAVs may not be reliable. Furthermore, most studies on the adoption of SAVs were developed in metropolitan areas that provide access to a variety of transportation modes (Soltani et al. 2021). However, there is an essential need to understand how an SAV service can improve the mobility, accessibility, and equity of those who live in distant and sprawling areas with limited or no access to fixed-route transit.

Given the importance of SAV acceptance and adoption, this research aims to use the data obtained from riders of an SAV service called RAPID (Rideshare, Automation, and Payment Integration Demonstration) in Arlington, Texas, and perform exploratory data analysis to identify the monthly ridership trends of the riders.

The RAPID demonstration was funded to support the goals of the federal transit administration (FTA) and the USDOT's Mobility Innovation initiative by integrating Level 4 autonomous vehicles (AVs) into the existing public on-demand rideshare service. The project aims to increase access for students and individuals with limited personal mobility, improve equity and accessibility to public transit, and enhance safety and efficiency. Therefore, this study aims to explore the ridership trends of this SAV service to understand better how riders use the SAV service in small communities such as university campuses and city downtowns.

METHODOLOGY

The primary objective of this study is to examine the trends of riders using RAPID service, which is a trial initiative involving autonomous shuttles in Arlington, Texas. The Arlington

RAPID programmed, which aims to integrate an on-demand rideshare service (Via) with Shared Autonomous Vehicle service in the city, is being financially supported by the Federal Transit Administration. The on demand rideshare service commenced operating inside a restricted area of Arlington in December 2017 and subsequently expanded its service throughout the entire city in January 2021. The app is a ridesharing service that operates on-demand and offers transportation to consumers in 6-passenger vans within the designated service borders of the City of Arlington in 2019. The primary stakeholders in the RAPID SAV pilot project included May Mobility, Via Transportation, the University of Texas at Arlington (UTA), and the city of Arlington. The city provides autonomous shuttles throughout the Downtown Arlington region and the UTA campus as part of the RAPID initiative. These shuttles are integrated with the Via platform for ride-sharing services. The service is completely based on immediate requests and offers transport services from Monday to Friday, operating between 7 a.m. and 7 p.m. To achieve the equitable objectives of the pilot initiative, RAPID offers complimentary transportation services to college students and operates vehicles that are accessible for individuals with disabilities (City of Arlington 2021). The service borders encompass a region characterized by a poverty rate of 39% and 11% of households lacking private vehicles, in order to fit with the project's objectives of promoting transportation equity and accessibility.

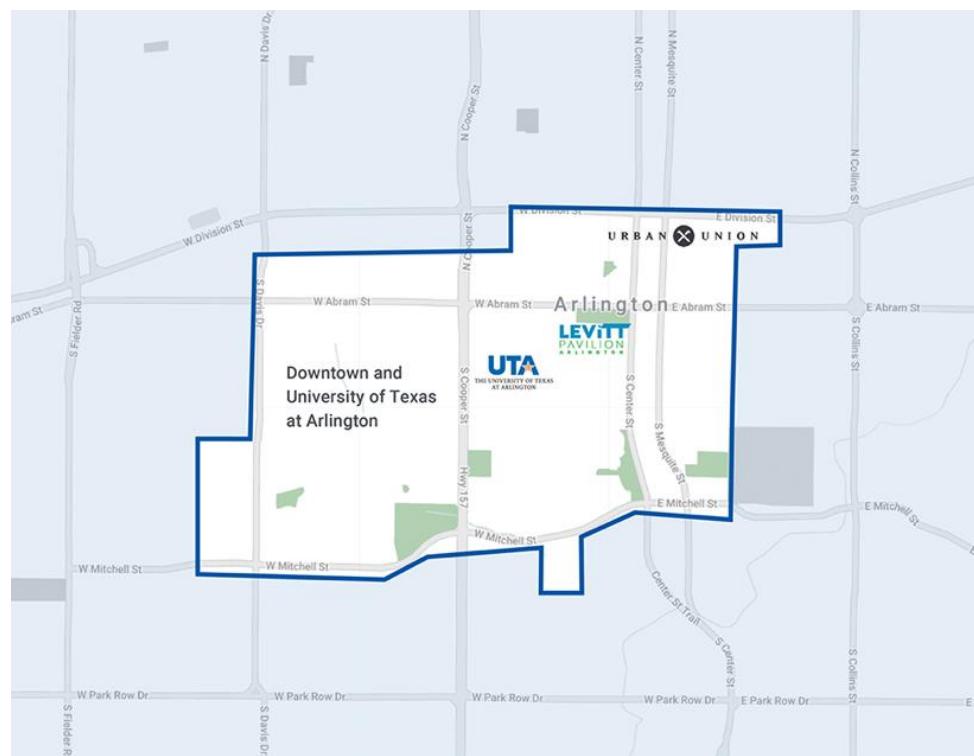


Figure 1. RAPID SAV Service Area

RESULTS AND DISCUSSION

Although the RAPID service is made available from 2021, this study uses data from June 2022 to May 2023 to analyze the ridership trends of passengers. The service shuttles are operated in Downtown Arlington area and the UTA campus and majority of the riders are UTA students.

Rides Based on Day of the Week

Figure 2 illustrates the monthly ride variations for the shared autonomous vehicle (SAV) service under the RAPID project from June 2022 to May 2023. As the service area of RAPID is downtown and the University of Texas at Arlington, most service riders are UTA students. As shown in Figure 2, the service in June 2022 begins with a notable number of rides, suggesting an immediate adoption and utilization by university students and riders from downtown Arlington. July maintains a solid level of ridership, showcasing continued interest and reliance on the shared autonomous vehicle service during the summer break. Rides surge in August, likely as students return for the fall semester, indicating a heightened demand for convenient transportation. September maintains a robust ridership, reflecting sustained popularity among students as they settle into their academic routines. The rides from October to December 2022 show a decrease, possibly due to variations in academic schedules, holidays and altered travel plans, aligning with the winter break when many students leave campus. The rides in January 2023 start with reduced rates, possibly as students return from the winter break and academic activities slowly resume. February sees a moderate increase, indicating a return to regular transportation patterns as the spring semester progresses. March witnesses a notable uptick, potentially influenced by increased campus activities or special events. April experiences a substantial surge, possibly due to heightened academic engagements, occurrences, or favorable weather conditions. The data concludes with a peak in May, indicating a significant demand associated with final exams and end-of-semester activities.

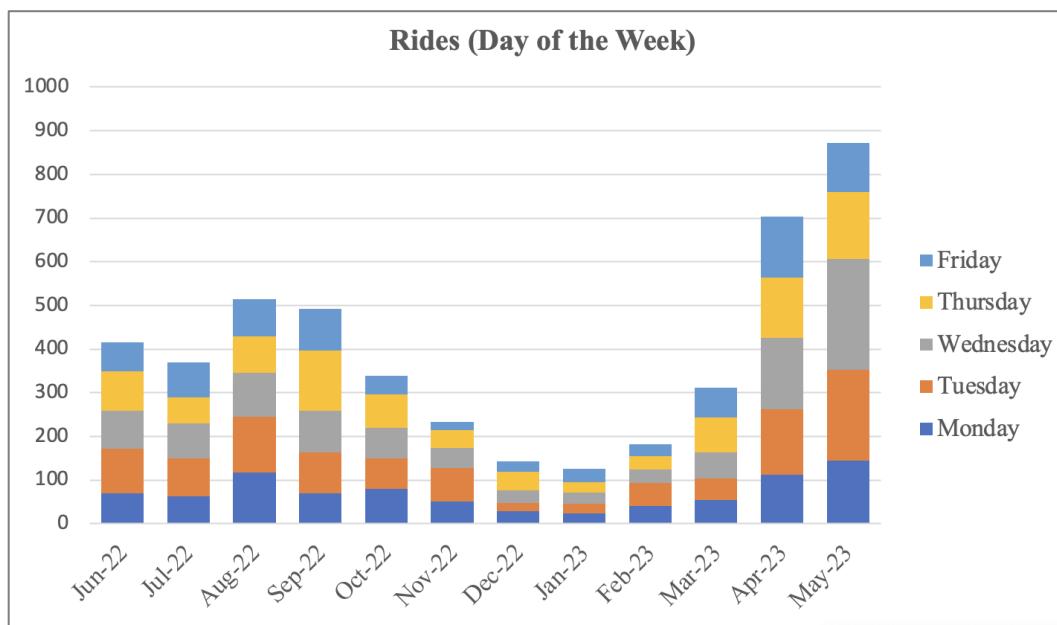


Figure 2. Monthly Rides Analysis for RAPID Service Based on Day of the Week

This analysis also delves into the weekly ride patterns, exploring the number of rides recorded from Monday to Friday. The dataset comprises ride counts for each respective day, providing insights into the fluctuations and trends within the week. As shown in Figure 2, the week begins with a robust number of rides on Mondays, indicating a significant demand for transportation services as individuals kickstart their workweek. Tuesdays also see a notable

increase in rides, suggesting heightened commuting activity and potential mid-week engagements. Wednesdays maintain a high level of ridership, reflecting consistent demand and attention throughout the middle of the week. Rides slightly decrease on Thursdays, but the numbers remain substantial, possibly influenced by varied weekday routines. The week concludes with a dip in Friday rides, indicative of reduced commuter activity as individuals transition into the weekend.

Rides Based on Time of the Day

The Figure below provides ride numbers categorized into early afternoon (12 pm – 2:59 pm), late afternoon (3 pm – 5:59 pm), and evening (6 pm – 6:59 pm), offering valuable insights into the varying demands across different times of the day. Although the service was offered during early morning (7 am – 8:59 am), and morning (9 am – 11:59 am) hours, no rides were recorded during these times.

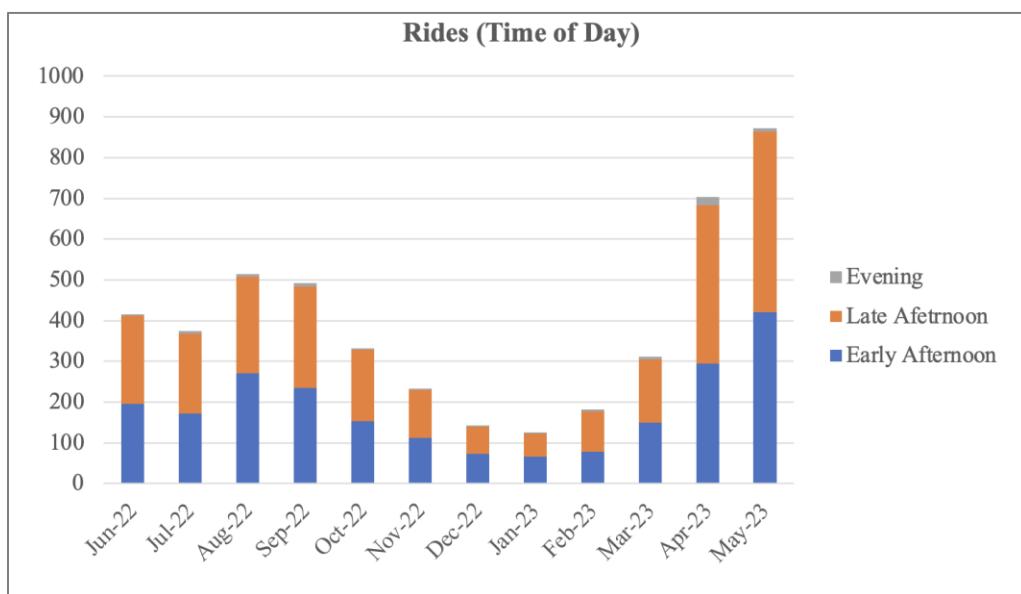


Figure 3. Monthly Rides Analysis for RAPID Service Based on Time of Day

As presented in Figure 3, the day starts with substantial number of rides during the early afternoons, highlighting a significant demand for transportation services. This could be indicative of midday commuting or other daytime activities. The ride counts peak during the late afternoons, indicating a surge in demand as the workday progresses. This period may align with the typical rush hour, with commuters utilizing the RAPID service for their afternoon travels. Furthermore, evenings see a comparatively lower number of rides, suggesting a decline in demand as the day transitions into nighttime. This could be influenced by factors such as reduced work-related travel and a shift towards more localized activities.

Rides Based on Number of Passengers in a Single Ride Request

As presented in Figure 4, the majority of rides fall into the category of single-passenger requests, indicating a prevalent trend toward individualized transportation. This suggests that

commuters often prefer personalized rides, aligning with the convenience and flexibility of the RAPID service. Furthermore, a smaller yet noteworthy portion of rides accommodates two passengers, highlighting instances where users opt for shared rides. This could represent instances of shared commutes or travel companionship, showcasing the versatility of the RAPID service to cater to various passenger needs.



Figure 4. Monthly Rides Analysis for RAPID Service Based on Number of Passengers in a Single Ride Request

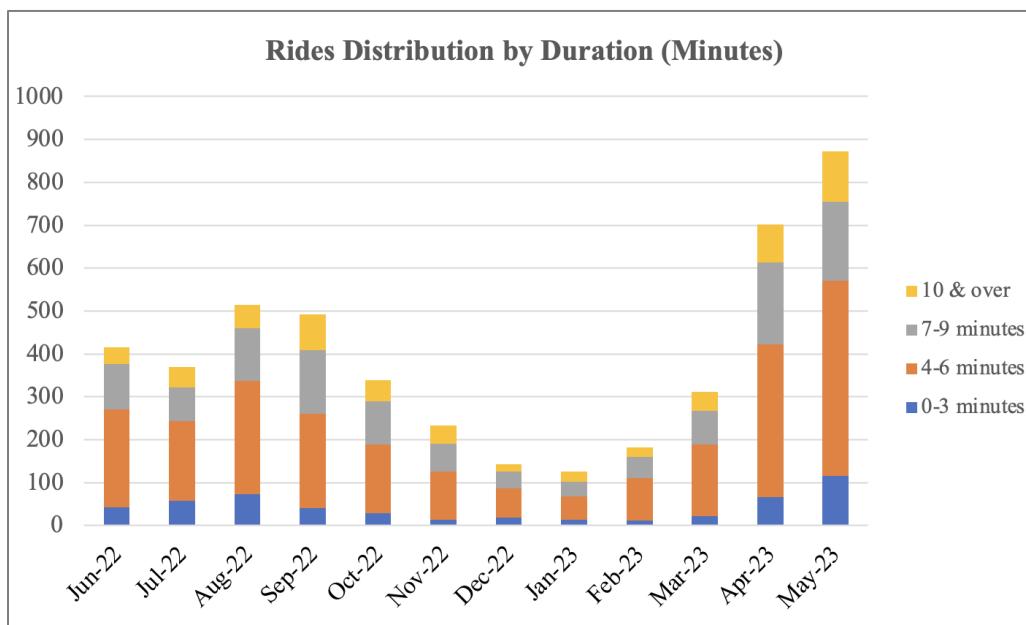


Figure 5. Monthly Rides Analysis for RAPID Service Based on Duration (Minutes)

Rides Distribution by Duration (Minutes)

As shown in Figure 5, A notable portion of rides falls within the 0–3-minute duration, indicating that a significant number of users are opting for short-distance trips. This suggests that RAPID effectively caters to quick and convenient transportation needs within the service boundaries. Most rides fall within the 4–6-minute duration, showcasing the efficiency and accessibility of RAPID for moderately short to medium-distance journeys. This range likely represents a mix of local commutes and slightly longer trips. A substantial number of rides fall within the 7–9-minute duration, reflecting the service's ability to accommodate longer journeys while maintaining a balance between speed and accessibility. A segment of rides extends beyond 10 minutes, indicating that RAPID serves users with longer-distance transportation needs. This may include trips from one end of the service boundaries to another, providing comprehensive coverage for diverse travel requirements.

CONCLUSION

This study utilized the data from RAPID service in Arlington, Texas, and performed exploratory data analysis to identify the monthly ridership trends of the riders. The results revealed that highlight the consistent demand for RAPID services from Monday to Thursday, with a slight dip on Fridays as individuals transition into the weekend. This observation aligns with expected commuter behaviors, offering insights that can inform service optimization and resource allocation strategies. Additionally, the exploration of ride patterns based on different times of the day, as demonstrates the adaptability of RAPID to varying demands throughout the day. The service effectively caters to midday commuting, experiences a surge during the late afternoon rush hour, and witnesses a decline in the evening, reflecting changing travel patterns as the day progresses. The results also revealed that the majority of rides accommodate a single passenger, emphasizing the convenience and flexibility that users associate with personalized transportation. This research providing crucial insights into the patterns of SAV ridership. These findings will be valuable for local, state, and federal transit authorities and planners, enabling them to develop policies and transportation strategies that specifically address and promote SAV ridership as the service becomes more widely available.

REFERENCES

- Acheampong, R. A., and F. Cugurullo. (2019). "Capturing the behavioural determinants behind the adoption of autonomous vehicles: Conceptual frameworks and measurement models to predict public transport, sharing and ownership trends of self-driving cars." *Transportation research part F: traffic psychology and behaviour* 62: 349-375.
- Bansal, P., K. M. Kockelman, and A. Singh. (2016). "Assessing public opinions of and interest in new vehicle technologies: An Austin perspective." *Transportation Research Part C: Emerging Technologies* 67: 1-14.
- City of Arlington. (2019). "Data in Action: Arlington's Via Rideshare Service Enhancing Mobility for Citizens." Office of Communication. Accessed 17 November 2020.
https://www.arlingtontx.gov/news/my_arlington_t_x/news_stories/datainaction_via.

- City of Arlington. (2021). "City of Arlington Launches First-of-its-Kind, On-Demand Self Driving Shuttle Service with RAPID Program." https://www.arlingtontx.gov/news/my_arlington_t_x/news_stories/arlington_rapid_launch.
- Channamallu, S. S., S. Kermanshachi, and A. Pamidimukkala. (2023). "Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles." In *International Conference on Transportation and Development 2023* (pp. 39-50), Reston, VA: ASCE.
- Dichabeng, P., N. Merat, and G. Markkula. (2021). "Factors that influence the acceptance of future shared automated vehicles—A focus group study with United Kingdom drivers." *Transportation research part F: traffic psychology and behaviour* 82: 121-140.
- Etmiani-Ghasrodashti, R., R. K. Patel, S. Kermanshachi, J. M. Rosenberger, D. Weinreich, and A. Foss. (2021). "Integration of shared autonomous vehicles (SAVs) into existing transportation services: A focus group study." *Transportation Research Interdisciplinary Perspectives*, 12, p.100481.
- Etmiani-Ghasrodashti, R., G. Hladik, S. Kermanshachi, J. M. Rosenberger, M. Arif Khan, and A. Foss. (2022). "Exploring shared travel behavior of university students." *Transportation Planning and Technology*, 46(1), pp. 22-44.
- Etmiani-Ghasrodashti, R., M. Khan, R. K. Patel, S. Kermanshachi, J. M. Rosenberger, A. Pamidimukkala, G. Hladik, and A. Foss. (2023). "Measuring students' satisfaction levels for transit services: An application of latent class analysis." *International Journal of Transportation Science and Technology*. (In Press).
- Etzioni, S., R. A. Daziano, E. Ben-Elia, and Y. Shiftan. (2021). "Preferences for shared automated vehicles: A hybrid latent class modeling approach." *Transportation Research Part C: Emerging Technologies* 125: 103013.
- Fu, X., Q. Nie, J. Liu, Z. Zhang, and S. Jones. (2021). "How Do College Students Perceive Future Shared Mobility with Autonomous Vehicles? A Survey of the University of Alabama Students." *International Journal of Transportation Science and Technology*.
- Haboucha, C. J., R. Ishaq, and Y. Shiftan. (2017). "User preferences regarding autonomous vehicles." *Transportation Research Part C: Emerging Technologies* 78: 37-49.
- Hossain, M. S., and M. Rahman Fatmi. (2022). "Modelling the adoption of autonomous vehicle: How historical experience inform the future preference." *Travel Behaviour and Society* 26: 57-66.
- Kerlin, K. (2019). "Automated Shuttle Comes to UC Davis West Village." <https://climatechange.ucdavis.edu/news/automated-shuttle-comes-to-uc-davis-west-village/>.
- Khan, M. A., R. Etmiani-Ghasrodashti, A. Shahmoradi, S. Kermanshachi, J. M. Rosenberger, and A. Foss. (2022). "Integrating shared autonomous vehicles into existing transportation services: evidence from a paratransit service in Arlington, Texas." *International Journal of Civil Engineering*, 20(6), pp.601-618.
- Khan, M. A., R. Etmiani-Ghasrodashti, A. Pamidimukkala, S. Kermanshachi, and J. M. Rosenberger. (2023a). "Impacts of on-demand ride services on the number of traffic crashes—A case study of RideAustin in Austin, TX." *Transportation Research Interdisciplinary Perspectives*, 22, p.100966.
- Khan, M. A., R. K. Patel, A. Pamidimukkala, S. Kermanshachi, J. M. Rosenberger, G. Hladik, and A. Foss. (2023b). "Factors that determine a university community's satisfaction levels with public transit services." *Frontiers in Built Environment*, 9, p.1125149.
- Krueger, R., T. H. Rashidi, and J. M. Rose. (2016). "Preferences for shared autonomous vehicles." *Transportation research part C: emerging technologies* 69: 343-355.

- Levin, M. W., M. Odell, S. Samarasena, and A. Schwartz. (2019). "A linear program for optimal integration of shared autonomous vehicles with public transit." *Transportation Research Part C: Emerging Technologies* 109: 267-288.
- Lu, Z., R. Du, E. Dunham-Jones, H. Park, and J. Crittenden. (2017). "Data-enabled public preferences inform integration of autonomous vehicles with transit-oriented development in Atlanta." *Cities* 63: 118-127.
- NCSL (National Council of State Legislature). (2020). *Autonomous Vehicles / Self-Driving Vehicles Enacted Legislation*. National Council of State Legislature.
- Pamidimukkala, A., S. Kermanshachi, J. M. Rosenberger, and G. Hladik. (2023a). "Evaluation of barriers to electric vehicle adoption: A study of technological, environmental, financial, and infrastructure factors." *Transportation Research Interdisciplinary Perspectives*, 22, p.100962.
- Pamidimukkala, A., S. Kermanshachi, J. M. Rosenberger, and G. Hladik. (2023b). "Barriers and Motivators to the Adoption of Electric Vehicles: A Global Review." *Journal of Green Energy and Intelligent Transportation* (In Press).
- Pamidimukkala, A., R. K. Patel, S. Kermanshachi, J. M. Rosenberger, and S. Tanvir. (2023c). "A Review on Shared Mobility and Electric Vehicles." In *International Conference on Transportation and Development 2023* (pp. 333-342), Reston, VA: ASCE.
- Pamidimukkala, A., S. Kermanshachi, and R. Patel. (2023d). "An Exploratory Analysis of Crashes Involving Autonomous Vehicles." In *International Conference on Transportation and Development 2023* (pp. 343-350), Reston, VA: ASCE.
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. M. Rosenberger, and D. Weinreich. (2021). "Exploring preferences towards integrating the autonomous vehicles with the current microtransit services: A disability focus group study." In *International Conference on Transportation and Development 2021* (pp. 355-366). Reston, VA: ASCE.
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. Rosenberger, and A. Foss. (2022). "How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems*." *ASCE International Conference on Transportation & Development*, pp. 81-93, Reston, VA: ASCE.
- Patel, R. K., S. S. Channamallu, M. A. Khan, S. Kermanshachi, and A. Pamidimukkala. (2023a). "An Exploratory Analysis of Temporal and Spatial Patterns of Autonomous Vehicle Collisions." *Public Works Management & Policy*, 0(0).
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. M. Rosenberger, A. Pamidimukkala, and A. Foss. (2023b). "Identifying individuals' perceptions, attitudes, preferences, and concerns of shared autonomous vehicles: During-and post-implementation evidence." *Transportation Research Interdisciplinary Perspectives*, 18, p.100785.
- Perkins, L., N. Dupuis, and B. Rainwater. (2018). *Autonomous Vehicle Pilots Across America*. <https://www.nlc.org/resource/autonomous-vehicle-pilots-across-america>.
- Shen, Y., H. Zhang, and J. Zhao. (2018). "Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore." *Transportation Research Part A: Policy and Practice* 113: 125-136.
- Soltani, A., D. Ananda, and M. Rith. (2021). "University Students' Perspectives on Autonomous Vehicle Adoption: Adelaide Case Study." *Case Studies on Transport Policy*.
- Wang, S., Z. Jiang, R. B. Noland, and A. S. Mondschein. (2020). "Attitudes towards privately-owned and shared autonomous vehicles." *Transportation research part F: traffic psychology and behaviour* 72: 297-306.

- Wen, J., Y. X. Chen, N. Nassir, and J. Zhao. (2018). "Transit-oriented autonomous vehicle operation with integrated demand-supply interaction." *Transportation Research Part C: Emerging Technologies* 97: 216-234.
- Yuen, K. F., D. T. K. Huyen, X. Wang, and G. Qi. (2020). "Factors influencing the adoption of shared autonomous vehicles." *International journal of environmental research and public health* 17 (13): 4868.

Automated Vehicles vs. Human Drivers: Modeling Driving Behavior Using Data from Field Experiments

Pei Li, Ph.D.¹; Steven T. Parker, Ph.D.²; and David A. Noyce, Ph.D.³

¹Traffic Operations and Safety Laboratory, Dept. of Civil and Environmental Engineering, Univ. of Wisconsin–Madison, Madison, WI. Email: pei.li@wisc.edu

²Traffic Operations and Safety Laboratory, Dept. of Civil and Environmental Engineering, Univ. of Wisconsin–Madison, Madison, WI. Email: sparker@engr.wisc.edu

³Traffic Operations and Safety Laboratory, Dept. of Civil and Environmental Engineering, Univ. of Wisconsin–Madison, Madison, WI. Email: danoyce@wisc.edu

ABSTRACT

As automated vehicles (AVs) gradually gain prevalence on public roads, understanding their distinctive driving behavior is crucial for traffic management and planning. This study conducted field experiments using an SAE Level-3/4 AV and collected driving data of AVs and human drivers on public roads using sensors including GPS, radar, camera, and LiDAR. The Wilcoxon rank-sum test is used to identify the difference in the behavior between AVs and human drivers. In addition, logistic regression and Extreme Gradient Boosting (XGBoost) are used to classify AVs and human drivers. Results suggest that there exists a significant difference in driving behavior between AVs and human drivers. Moreover, features including the mean speed and the distance from the vehicle to the detected objects are positively related to the probability of the vehicles being AVs, while the standard deviation of speed and the mean acceleration are negatively associated with it. Furthermore, XGBoost accurately identifies AVs and human drivers using the extracted features with an average area under the curve of 0.92. Results from interpreting results from XGBoost indicate that it performs better when the mean speed is either in the low or high ranges. Moreover, AVs and human drivers are hard to differentiate using the model when the vehicle is too far from other objects. This study underscores the substantial divergence in driving behavior between AVs and human drivers, offering valuable insights for the evaluation of the impact of AVs on traffic conditions.

INTRODUCTION

Automated Vehicle (AV), as an emerging technology, has drawn a lot of attention from researchers, policymakers, vehicle manufacturers, and the public due to its potential to improve traffic safety and efficiency significantly. However, the era of AVs is yet to come and there will be a long period when AVs and Human-Driven Vehicles (HVs) coexist. Mixed traffic flow created by this coexistence would have a substantial impact on the existing traffic conditions. Hence, the driving behavior of AVs and human drivers needs to be comprehensively investigated to better understand AVs' behavior, identify the difference between them, and thus facilitate the evaluation of AVs' impact on traffic conditions.

Most existing studies (Beauchamp et al., 2022; Jayaraman et al., 2019; Mahadevan et al., 2019; Palmeiro et al., 2018; Soni et al., 2022; Wen et al., 2022; Zhao et al., 2020) have focused on investigating the reactions of other road users, including drivers, pedestrians, and cyclists to AVs using techniques such as surveys, simulators, closed test tracks, etc. The findings derived

from these studies indicated a notable difference in people's reactions to AVs and HVs. For example, Zhao et al. (2020) experimented with a closed test track to investigate the car-following behavior of human drivers when the lead vehicles were either AVs or HVs. Results suggested that drivers could be divided into three groups based on their car-following behavior, AV-believers, AV-skeptics, and AV-insensitive. Similarly, Soni et al. (2022) have studied the gap acceptance, car-following, and overtaking behavior of human drivers while interacting with AVs and HVs. Results suggested that drivers would adopt smaller critical gaps while interacting with AVs than interacting with HVs. Drivers would maintain much smaller headway after overtaking AVs, compared to overtaking HVs. Moreover, this study suggested that positive information about the behavior of AVs could lead to closer interactions with AVs. Wen et al. (2022) have analyzed the safety of human drivers' behavior while interacting with AVs and HVs using the Waymo Open Dataset. Two driving volatility measures, time headway and Time-To-Collision (TTC) were used to evaluate the behavior of human drivers. Experimental results suggested that human drivers had lower driving volatility (i.e., smaller time headways and larger TTCs) while following AVs. Similar results have been found by Beauchamp et al. (2022), who have studied interactions between automated shuttles and other road users. Trajectory data of shuttles and other road users were extracted using video data, which were then used to estimate different surrogate safety measures and other metrics. Results suggested that automated shuttles have safer interactions with lower speeds and higher TTCs compared to human drivers. Jayaraman et al. (2019) conducted a simulator study with 30 participants in a virtual reality environment to investigate pedestrians' trust in AVs that had defensive, normal, and aggressive behavior. Results suggested that pedestrians' trust would decrease significantly if AVs behaved aggressively, such as accelerating and decelerating abruptly.

In summary, results from existing studies indicated that road users would act differently to AVs compared to their reactions to HVs. Furthermore, the behavior of AVs has a substantial influence on the reactions of road users toward them. Therefore, it is crucial to examine the difference in driving behavior between AVs and HVs, and then develop methods to differentiate them. In addition, Fuest et al. (2020) conducted a survey involving 40 participants about their attitudes regarding marking AVs to help people differentiate them from HVs. 66 % of the participants would prefer AVs to be marked as it would be easier to assess the AV's driving behavior and to adapt their behavior to the presence of AVs. Nevertheless, only one study has analyzed the difference in driving behavior between AVs and HVs using a driving simulator and has indicated a notable difference in driving behavior between them (Remonda et al., 2021). For example, AVs were found to have a shorter reaction time compared to human drivers.

The objective of this study is to enhance the comprehension of the difference in driving behavior between AVs and human drivers. Moreover, in contrast to most AV-related studies that used surveys, simulators, and closed test tracks, this study has conducted experiments on real-world open roads to improve the fidelity of the results. This study aims to answer three questions. First, what is the difference between the driving behavior of AVs and human drivers? Second, can driving behavior-related features be used to identify AVs and human drivers? Third, what are the key features for identifying AVs and human drivers? To answer these questions, an SAE Level-3/4 AV was used for data collection in real-world driving conditions. Features related to driving behavior have been extracted from the data, such as the mean speed, the standard deviation of speed, the mean acceleration, etc. Statistical tests have been conducted to investigate the difference in driving behavior between AVs and human drivers. Moreover, logistic regression was applied to quantify the contribution of different features to the driving modes

(i.e., AVs and human drivers), and an XGBoost model was developed to classify AVs and human drivers using the extracted features. The model was further interpreted to investigate the effects of different features on its results.

METHODS

Apparatus

Data used in this study are collected using an SAE Level-3/4 AV manufactured by Perrone Robotics as shown in Figure 1. The AV is equipped with several sensors (e.g., GNSS, camera, radar, and LiDAR) for localization, perception, and sensing. It has two modes. Specifically, the automated mode enables the AV to operate automatically on pre-mapped routes, while the manual mode allows the driver to take control of the AV and drive it like a normal vehicle. The operation of the AV requires two operators who occupy the driver and passenger seats, one as the driver, and another as the observer. The driver can switch the mode of the AV between the AV and manual modes when it is necessary, or as required. The operator is responsible for using a tablet/laptop to run the program that allows the automated operation of the vehicle.



Figure 1. The AV and sensors in the AV.

Experimental Design

Experiments were conducted in the City of Racine, WI, USA, from February 2022 to June 2023. Data from both AV and manual modes were gathered on the same day, at close intervals, with the aim of reducing the influence of varying traffic conditions on the behavior of both autonomous vehicles and human drivers. This approach ensures a more equitable comparison between the two modes. Several routes were mapped for operating the vehicle, covering a wide range of infrastructure, including parking lots, signalized intersections, stop-sign-controlled intersections, unsignalized intersections, etc. Figure 2 gives an example of four routes. The first and third routes are examples of routes in parking lots. They are designed to test the AV's ability to perform simple driving tasks, i.e., driving through static objects (e.g., parked vehicles, trash cans, etc.). The second and fourth routes are examples of routes that are on public routes. They are designed to test the AV's ability to navigate through complex traffic conditions and interact with other road users, including vehicles, pedestrians, and bicyclists. Six drivers are trained to operate the AV. All drivers have valid US driver's licenses and are of normal or corrected-to-normal vision.

Detection, vehicle dynamics, and location data are saved automatically into the AV's data logger during its operation. Specifically, each trip's data, including the log, camera, and LiDAR data are saved into an individual folder. The log is a text file containing detailed information about the vehicle per cycle (i.e., 100 milliseconds), including the location, speed, heading, driving mode, and information about objects detected by sensors. The radar, LiDAR, and camera data are collected as auxiliary data to investigate the vehicle's surrounding environment. For example, video data could be used to verify objects detected by radar and LiDAR sensors.

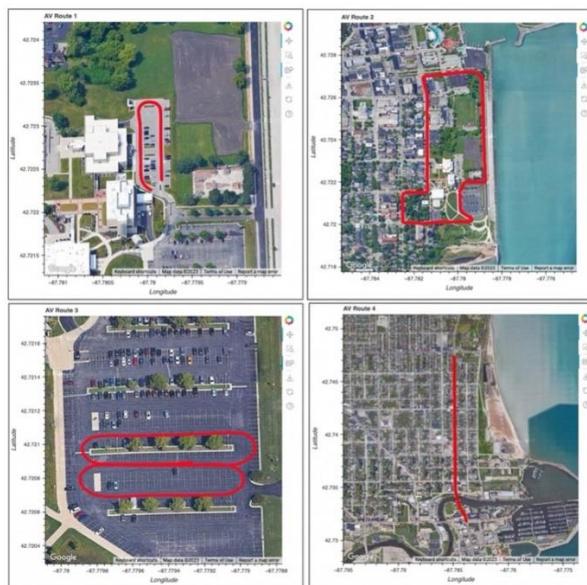


Figure 2. The designed AV routes in the City of Racine, USA.

Data Preparation

This study uses trip logs to investigate the driving behavior of AVs and human drivers. The collected logs are parsed to convert the unstructured text data into structured location, vehicle dynamics, and detection data as shown in Figure 3. A Python script is created to parse a log file into two tabular files. Since the log data are saved per cycle, the number of cycles is used to match the structured location, vehicle dynamics, and detection data. A detailed data description is shown in Table 1. Each trip's data are further segmented into small trips based on the value of driving modes. Trips that last less than ten seconds are removed as they are too short to reflect useful information about driving behavior. After trip segmentation, several statistical features are estimated in a 1-second time interval to describe driving behavior, such as the mean speed, the standard deviation of speed, the mean acceleration, etc.

ANALYZING DRIVING BEHAVIOR

The difference in driving behavior between AVs and human drivers

The Wilcoxon rank-sum test is used in this study to examine the difference in driving behavior between AVs and human drivers. It is a non-parametric test that does not require the

data to be normally distributed (Woolson, 2007). This test is more robust against violations of normality assumptions that are required by parametric tests such as the T-test. The null hypothesis of the Wilcoxon rank-sum test is that two populations have the same distribution with the same median.

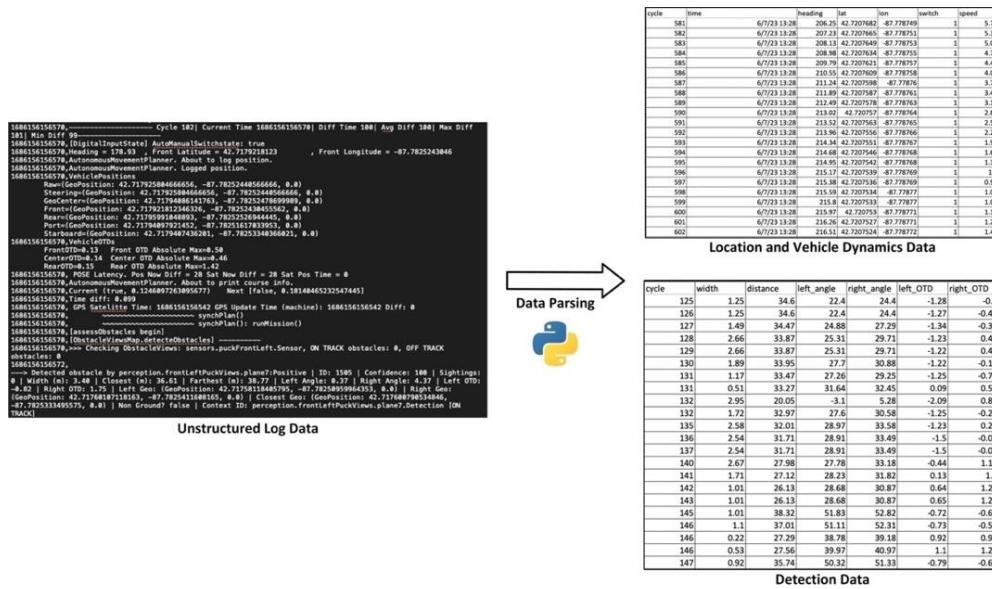


Figure 3. The process of data preparation.

Table 1. Data Description

Type	Name	Description
General	Cycle	The number of the cycle
	Mode	The driving mode of the vehicle, which is either automated or manual
	Time	The timestamp of the cycle
Vehicle	Latitude & Longitude	The latitude and longitude of the vehicle in world geodetic system 1984 coordinate
	Speed	The speed of the vehicle in m/s
	Acceleration	The acceleration of the vehicle in m ² /s
	Heading	The angle from the true north to the forward heading of the vehicle in degrees
Detection	Width	The detected width of the object in meters
	Distance	The distance between the closest point of the object and the associated sensor in meters

Classifying AVs and human drivers based on driving behavior

Two models, logistic regression and XGBoost, are used in this study for classifying AVs and human drivers and estimating the relationships between different features and driving modes. Given a driving record i that has a vector of features represented by $X_i = (x_1, x_2, \dots, x_n)$, its

corresponding driving mode y_i is either 1 (AV) or 0 (human driver). Let p_i be the probability of the driving record i is in AV mode, a logistic regression model can be built to estimate p_i using the following equations:

$$y_i \sim Bernoulli(p_i) \quad (1)$$

$$\text{logit}\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \quad (2)$$

where β_0 is the intercept term for the driving record i , $\beta = (\beta_0, \beta_1, \dots, \beta_n)$ is the vector of regression coefficients for the independent variables, which are the driving behavior-related features of the driving record i . Analyzing features' regression coefficients can help understand their relationships with the driving mode.

While logistic regression can help further understand the relationships between driving behavior-related features and driving modes, it may not be able to achieve highly accurate classification results compared with machine learning models. Therefore, this study utilizes XGBoost to classify AVs and human drivers using the features in Table 2. XGBoost was proposed as a tree-boosting method (Chen & Guestrin, 2016). Each tree in XGBoost is developed upon its previous tree and affects its next tree. For a given data, it uses K additive functions to estimate \hat{y} as shown in Equation (3), where \mathcal{F} is the space of regression trees. Equation (4) gives an example of a tree in XGBoost, where q represents the structure of each tree and w represents the leaf weights. The objective of XGBoost is to minimize the value of the regularized objective function in Equation (5), where l is a differentiable convex loss function that estimates the difference between y_i and \hat{y}_i . Ω is a regularization term that penalizes the model's complexity as shown in Equation (6). The usage of the regularization function prevents the model from becoming over-complicated, where T is the number of leaves, and λ and γ are the regularization parameters.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F} \quad (3)$$

$$f_k(x) = w_{q(x)} \quad (4)$$

$$\text{obj}(\theta) = \sum_i^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (5)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (6)$$

XGBoost is a popular machine learning model due to its high scalability and efficiency. Moreover, since XGBoost is a tree-based model, it is easy to interpret the contributions of features on its results. To develop an XGBoost model using the prepared data, data are first randomly divided into training, validation, and test using a ratio of 7:1:2. The model is trained on the training data with its hyperparameters (e.g., learning rate, the number of trees, etc.) optimized

based on its performance on the validation data. Then, the model is evaluated on the test data. This study uses the Area Under the Curve (AUC) for evaluating the model's performance. AUC is estimated as the size of the area under the Receiver Operating Characteristic (ROC) curve, which plots a model's true positive rate (i.e., sensitivity) against its false positive rate (i.e., false alarm rate) at various threshold settings. The estimation of the true and false positive rates can be referred to (Wikipedia, 2023). ROC and AUC are optimal choices for evaluating a model's performance on classification, especially when the data are imbalanced as these two measures are not biased toward models that perform well on the minority class at the expense of the majority class. Moreover, since the data contained more AV samples than human driver samples, directly developing the model using the data would affect its ability to identify human drivers. A data re-sampling method named Synthetic Minority Over-sampling Technique (SMOTE) is used in this study to balance the training data (Chawla et al., 2002). The validation and test data are not balanced as they are designed to reflect the distribution of real-world data.

RESULTS

Data Summary

In total, around 10.8 hours of driving data have been extracted from the log file, which contained 394,576 data points. Moreover, 78% of the driving data was collected while the vehicle was in the automated mode, and the rest 22% was collected while the vehicle was in the manual mode. Table 2 provides detailed information about features used in this study, which were estimated in 1-minute intervals. Most features, such as the mean speed, standard deviation speed, and mean acceleration, were widely used in existing studies to depict driving behavior (Eren et al., 2012; Li et al., 2021; Linkov et al., 2019; Toledo et al., 2007). In addition, the mean distance between the vehicle and the detected objects was used in this study to represent the driver's ability to navigate through the encountered objects.

Table 2. Descriptive Statistics of the Extracted Features

Name	Description	AV Mode		Manual Mode	
		(Mean STD)	(Min, Max)	(Mean STD)	(Min, Max)
Mean_speed	The mean speed	(4.51, 2.74)	(0.00, 11.08)	(3.77, 3.48)	(0.00, 11.11)
Std_speed	The standard deviation of speed	(0.08, 0.11)	(0.00, 0.95)	(0.08, 0.10)	(0.00, 0.71)
Mean_acc	The mean acceleration	(0.00, 0.43)	(-2.70, 1.63)	(0.01, 0.41)	(-1.89, 2.34)
Std_acc	The standard deviation of acceleration	(0.17, 0.16)	(0.00, 2.14)	(0.12, 0.14)	(0.00, 2.36)
Mean_dist	The mean distance between the vehicle and the detected objects	(13.83, 17.37)	(0.00, 99.58)	(11.57, 15.34)	(0.00, 98.22)

Analyzing the Driving Behavior of AVs and Human Drivers

Figure 4 gives an example of the distributions of the mean speed of automated and manual driving modes, and both have clear non-normal distributions. Similar observations have been found in the distributions of other features, which justify the usage of the Wilcoxon test to identify the difference in the distributions of different features between the automated and manual driving modes. Results from the Wilcoxon tests indicated that the difference in the distributions of the mean speed, the standard deviation of speed, the standard deviation of acceleration, and the mean distance from the vehicle to objects between the AV and human drivers were significant at the 5% level. Therefore, the null hypothesis that two populations have the same distribution with the same median was rejected. This finding suggested that it would be feasible to use these features to identify different driving modes.

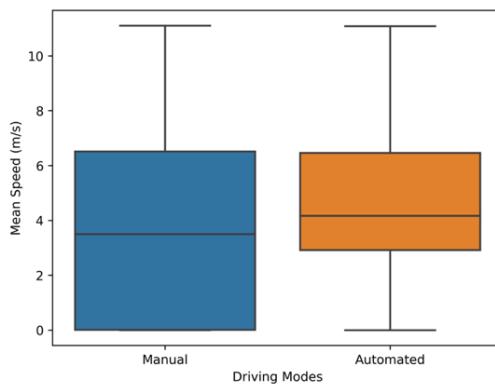


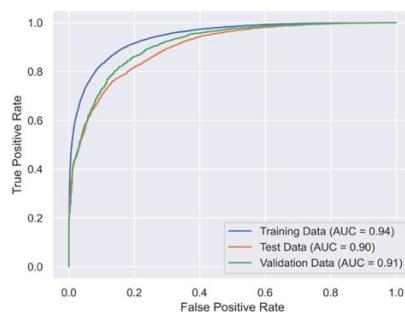
Figure 4. The boxplots of the mean speed of different modes.

The results of the logistic regression model are shown in Table 3. First, four features were significant at the 0.05 level, except the mean acceleration, which confirmed the difference in driving behavior between AVs and human drivers was overall significant. The results were consistent with the results of the Wilcoxon tests as the difference in the distributions of the same four features between AVs and human drivers was significant at the 0.05 level. Moreover, higher values of the mean speed, standard deviation of acceleration, and mean distance were likely to increase the probability of drivers being AVs. While higher values of the standard deviation of speed and mean acceleration were likely to decrease the probability of drivers being AVs. These results suggested that AVs would be able to drive at a higher speed and maintain a relatively larger distance to objects, compared to human drivers. Moreover, AVs were able to change their speed more smoothly than human drivers, as indicated by the relationships between the driving mode and the standard deviation of speed.

An XGBoost model was developed and tested using the method presented in the previous section. The model's performance on training, validation, and test data is shown in Figure 5. The model has achieved promising results in terms of classification accuracy. Besides, the model had an AUC of 0.94 on training data and an AUC of 0.90 on test data, indicating that the model did not have an over-fitting issue. Overall, results from the developed model suggested that it was feasible to use driving-related features to accurately classify AVs and human drivers, which corresponded to previous statistical findings that the difference in most features was significant between AVs and human drivers.

Table 3. Results of the Logistic Regression Model

Variable	Mean	STD Error	P-value	95% CI	
Intercept	0.5459	0.028	0	0.492	0.6
Mean_speed	0.0859	0.004	0	0.077	0.095
Std_speed	-1.1518	0.126	0	-1.399	-0.905
Mean_acc	-0.0452	0.029	0.122	-0.102	0.012
Std_acc	2.581	0.117	0	2.351	2.811
Mean_dist	0.0096	0.001	0	0.008	0.011

**Figure 5. ROC curves and AUCs of the model on training, validation, and test data.**

Machine learning models, especially complicated models, are often referred to as "black-box" models that are hard to interpret. Poor interpretability would make it hard to comprehend the relationships between features and the model's output it has learned. Moreover, if the interpretation of a model does not make sense, for example, giving high importance to features that are practically irrelevant to the output, this model should not be used even if it has high accuracy. Therefore, this study used the SHapley Additive exPlanations (SHAP) framework to explain the developed model. SHAP is a game theory-based method that assigns each feature an importance value to represent its contribution to the model's output (Lundberg & Lee, 2017). The SHAP values are additive, which means that the contribution of each feature to the final prediction can be computed independently. For a binary classification XGBoost model, the SHAP values are the log odds ratios estimated for the model.

First, the Beeswarm plot was used to examine the overall relationships between the features and the model's output. Each instance of the given explanation was represented by a single dot on each feature row. The horizontal position of the dot was determined by the SHAP value. Figure 6 shows the beeswarm plot of the five features which are ordered by their mean absolute SHAP values. Specifically, the standard deviation of speed and the mean acceleration were negatively correlated with the SHAP value, suggesting that the decrease of these features would increase the probability of the drivers being AVs. The standard deviation of acceleration and the mean distance were positively related to the SHAP value, suggesting the increase of these features would increase the probability of the drivers being AVs. The relationship between the mean speed and the SHAP value was more complicated and did not have a clear linear indication. Overall, the model interpretation results corresponded to the parameters of features estimated using the logistic regression and confirmed the previous findings.

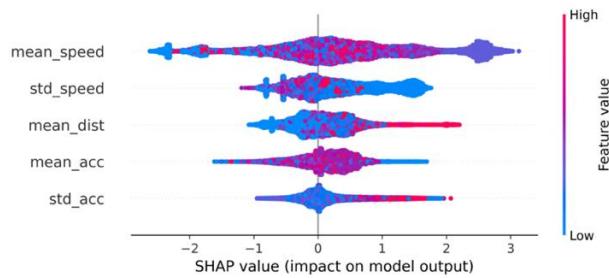


Figure 6. Beeswarm plot.

Second, to measure the impact of each feature on the model's output, the dependency plot was used as shown in Figure 7. A dependency plot shows the change in the SHAP values as a feature changes horizontally. The vertical spread of the SHAP values at a fixed feature value is due to the interaction effects with other features. Figure 7 (a) suggested that the relationships between the mean speed and the output of the model fluctuated over different values of the mean speed. This was consistent with the results in Figure 6, where the mean speed had a mixed impact on the model output, while Figure 7 (a) provided detailed information about it. Specifically, Figure 7 (a) indicated a positive impact of the mean speed on the model's output when the speed was within the range from 0 m/s to 3 m/s, while it was negative when the speed was larger than 7 m/s. The relationship when the speed was within 3 m/s to 7 m/s was not clear and more complicated. This finding suggested that the model would have better performance while the speed was low or high but may have trouble identifying AVs and human drivers while the speed was in mid-range.

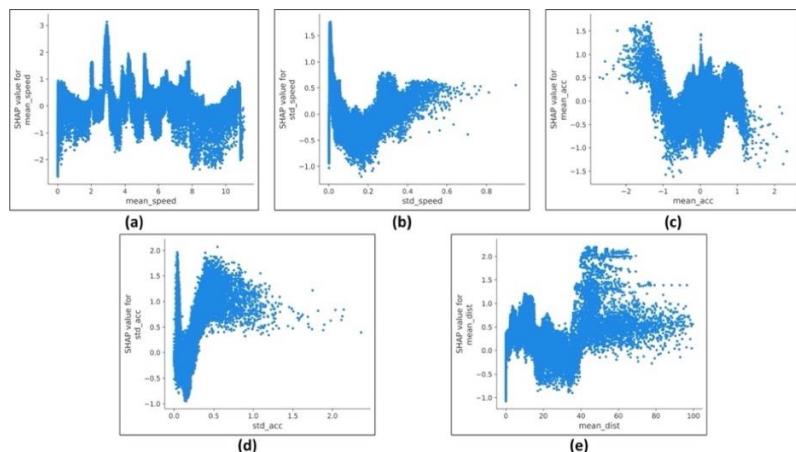


Figure 7. Dependency plots of the five features: (a) mean speed, (b) standard deviation of speed, (c) mean acceleration, (d) standard deviation of acceleration, and (e) mean distance between the vehicle and the detected objects.

Figures 7 (b) and (c) suggested the negative effects of the standard deviation of speed and mean acceleration on the model's output. Furthermore, the values of the model's output decreased significantly when the standard deviation of speed increased from 0 m/s to 0.2 m/s. Although they recovered after the standard deviation of speed was larger than 0.2 m/s, the overall

relationships were still negative. The relationships between the mean acceleration and the model's output had a similar trend as the standard deviation of the speed. For example, the values of the model's output dropped significantly when the mean acceleration increased from $-2 \text{ m}^2/\text{s}$ to $-1 \text{ m}^2/\text{s}$ and recovered after the mean acceleration exceeded $-1 \text{ m}^2/\text{s}$. Moreover, the positive effects of the standard deviation of acceleration and mean distance on the model's output were suggested in Figures 7 (d) and (e). An interesting finding from Figure 7 (d) was that the model's output had wide vertical spreads when the standard deviation of acceleration was within the range from $0 \text{ m}^2/\text{s}$ to $0.25 \text{ m}^2/\text{s}$, indicating the interaction effects from other features were relatively strong. Moreover, the values of the model's output had a substantial increase when the standard deviation of acceleration increased from $0.25 \text{ m}^2/\text{s}$ to $0.5 \text{ m}^2/\text{s}$. Lastly, Figure 7 (e) indicates the positive impact of the mean distance on the model's output. However, this relationship became weaker as the mean distance was larger than 40 m. This was consistent with the reality as it will be hard to identify AVs and human drivers if the distance is too large.

CONCLUSIONS

Extensive studies have been conducted to investigate the reactions of other road users to AVs over the past decades and have suggested a significant impact of AVs on road users' reactions. Moreover, as the penetration rate of AVs increases, their behavior will have substantial effects on existing traffic conditions. Therefore, it is crucial to analyze the driving behavior of AVs and identify the difference in driving behavior between AVs and HVs. However, only a limited number of studies have been conducted on this topic using only simulated data.

This study aims to analyze the driving behavior of AVs and human drivers using real-world data collected on public roads. An SAE Level-3/4 AV was used for data collection, and over 10 hours of driving data were collected along different routes that went through parking lots, local roads, major roads, and different types of intersections. Several features, including the mean speed, the standard deviation of speed, and the mean acceleration, have been extracted to describe the driving behavior of AVs and human drivers. A statistical measure was used to estimate the difference in driving behavior between AVs and human drivers. Logistic regression and XGBoost were used to classify AVs and human drivers, and SHAP was used to estimate the contribution of each feature to the probability of the drivers being AVs or humans. Several key findings have been summarized as follows, first, the difference in driving behavior between AVs and human drivers was significant, indicated by results from the Wilcoxon tests. Second, results from logistic regression suggested that the mean speed, the standard deviation of acceleration, and the mean distance were positively related to the probability of drivers being AVs, while the mean acceleration and the standard deviation of speed were negatively related to it. Third, the developed XGBoost was able to classify AVs and human drivers successfully, indicated by an average AUC of 0.92. Lastly, results from the SHAP values estimated using the XGBoost model have confirmed the relationships between features and driving modes suggested by the logistic models. Furthermore, the mean speed has been found to have a mixed effect on the XGBoost's output, suggesting that the model would have better performance while the speed was in the low and high ranges. Moreover, AVs and human drivers would become hard to differentiate when the distance from the vehicles to other objects was too large.

In the future, further research can be conducted while considering the surrounding and traffic environments when assessing the driving behavior of AVs and human drivers. This includes examining the impact of factors such as signalized intersections, stop-sign-controlled

intersections, the presence of nearby drivers, and interactions with pedestrians. Utilizing the cameras and LiDARs equipped in AVs, valuable information can be acquired through perception and sensing techniques to enhance the understanding of these specific driving scenarios.

ACKNOWLEDGEMENT

The Badger Automated Vehicle Research project is supported by the City of Racine, Wisconsin in collaboration with the Gateway Technical College. The ideas and views expressed in this paper are strictly those of the Traffic Operations and Safety (TOPS) Laboratory at the University of Wisconsin-Madison.

REFERENCES

- Beauchamp, É., Saunier, N., and Cloutier, M.-S. (2022). Study of automated shuttle interactions in city traffic using surrogate measures of safety. *Transportation Research Part C: Emerging Technologies*, 135, 103465.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Chen, T., and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Eren, H., Makinist, S., Akin, E., and Yilmaz, A. (2012). Estimating driving behavior by a smartphone. *2012 IEEE Intelligent Vehicles Symposium*, 234–239.
- Fuest, T., Feierle, A., Schmidt, E., and Bengler, K. (2020). Effects of marking automated vehicles on human drivers on highways. *Information*, 11(6), 286.
- Jayaraman, S. K., Creech, C., Tilbury, D. M., Yang, X. J., Pradhan, A. K., Tsui, K. M., and Robert, L. P., Jr. (2019). Pedestrian trust in automated vehicles: Role of traffic signal and AV driving behavior. *Frontiers in Robotics and AI*, 6, 117.
- Li, P., Abdel-Aty, M., and Yuan, J. (2021). Using bus critical driving events as surrogate safety measures for pedestrian and bicycle crashes based on GPS trajectory data. *Accident Analysis & Prevention*, 150, 105924.
- Linkov, V., Zaoral, A., Řezáč, P., and Pai, C.-W. (2019). Personality and professional drivers' driving behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 105–110.
- Lundberg, S. M., and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 4765–4774). Curran Associates, Inc. <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>.
- Mahadevan, K., Sanoubari, E., Somanath, S., Young, J. E., and Sharlin, E. (2019). AV-Pedestrian interaction design using a pedestrian mixed traffic simulator. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 475–486.
- Palmeiro, A. R., van der Kint, S., Vissers, L., Farah, H., de Winter, J. C., and Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 1005–1020.

- Remonda, A., Veas, E., and Luzhnica, G. (2021). Comparing driving behavior of humans and autonomous driving in a professional racing simulator. *PLoS One*, 16(2), e0245320.
- Soni, S., Reddy, N., Tsapi, A., van Arem, B., and Farah, H. (2022). Behavioral adaptations of human drivers interacting with automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 86, 48–64.
- Toledo, T., Koutsopoulos, H. N., and Ben-Akiva, M. (2007). Integrated driving behavior modeling. *Transportation Research Part C: Emerging Technologies*, 15(2), 96–112.
- Wen, X., Cui, Z., and Jian, S. (2022). Characterizing car-following behaviors of human drivers when following automated vehicles using the real-world dataset. *Accident Analysis & Prevention*, 172, 106689.
- Wikipedia. (2023). Receiver operating characteristic. https://en.wikipedia.org/wiki/Receiver_operating_characteristic.
- Woolson, R. F. (2007). Wilcoxon signed-rank test. *Wiley Encyclopedia of Clinical Trials*, 1–3.
- Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., and Qu, X. (2020). Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. *Transportation Research Part C: Emerging Technologies*, 114, 205–224.

Evaluating Autonomous Vehicles' Impact on Emergency Evacuation Clearance Time: A Comparative Study

Asad Ali¹; Mingwei Guo²; Ying Huang³; Talha Ahmed⁴; and Pan Lu⁵

¹Dept. of Civil, Construction, and Environmental Engineering, North Dakota State Univ., Fargo, ND. Email: asad.ali@ndsu.edu

²Dept. of Transportation, Logistics, and Finance, North Dakota State Univ., Fargo, ND. Email: mingwei.guo@ndsu.edu

³Dept. of Civil, Construction, and Environmental Engineering, North Dakota State Univ., Fargo, ND. Email: ying.huang@ndsu.edu

⁴Dept. of Civil, Construction, and Environmental Engineering, North Dakota State Univ., Fargo, ND. Email: talha.ahmed@ndsu.edu

⁵Dept. of Transportation, Logistics, and Finance, North Dakota State Univ., Fargo, ND. Email: pan.lu@ndsu.edu

ABSTRACT

Natural disasters like hurricanes, wildfires, and floods pose immediate hazards. Such events necessitate prompt and well-planned emergency evacuations to save lives and reduce fatalities, injuries, and property damage. This comparative study explores the potential impact of autonomous vehicles on evacuation travel times, citing the significance of swift evacuations amid the increasing frequency, severity, and global reach of natural disasters. The study rigorously models evacuation scenarios using GIS and VISSIM simulations to assess the feasibility of integrating autonomous cars into evacuation strategies for various natural disasters. The methodology focuses on evaluating a historical wildfire evacuation event, by employing a connected vehicle dataset, to showcase how autonomous vehicles navigate various emergency conditions. The results indicate a promising reduction in evacuation travel times and highlight the potential of autonomous vehicles in optimizing emergency evacuations and reducing clearance time. The findings also highlight the importance of integrating emerging technologies like connected and autonomous vehicles (CAVs) to enhance evacuation strategies and preparedness for emergencies.

INTRODUCTION

Natural disasters such as earthquakes, volcanic eruptions, hurricanes, wildfires and floods are very powerful and unpredictable events. These forces of nature, intensified by the effects of climate change, often cause widespread damages, injuries, and fatalities (Grünthal et al., 2006; Zuzak et al., 2022). According to statistical data, the United States witnesses an alarming frequency of natural disasters annually. In 2021 alone, the country encountered a staggering total of 97 natural disasters. Most were severe thunderstorms, wildfires, heat waves, and floods (Statista, 2023). The country has sustained 373 weather and climate disasters since 1980, where overall damages reached or exceeded \$1 billion. Hurricane Katrina in 2005, the costliest disaster at \$182.5 billion, resulted in nearly 2,000 fatalities, making it one of the deadliest in U.S. history (Smith, 2022). Similarly, the 2018 Camp Fire was the most destructive wildfire in modern history. The fire caused \$12.5 billion total insured losses and burned over 150,000 acres, forcing

at least 52,000 people to evacuate. Tragically, 85 people lost their lives in the fire, making it the deadliest wildfire in U.S. history in the past 100 years(Camp Fire, 2022; Hamideh et al., 2022).

One of the critical aspects in confronting these natural calamities is the timely evacuation of individuals residing in vulnerable areas. Whether it's the advance warning of a hurricane or the rapid spread of wildfires, the ability to evacuate swiftly and efficiently can significantly reduce the loss of life and property(Thompson et al., 2017). Among the various natural disasters that demand immediate evacuation measures, wildfires stand out due to their rapid and often unpredictable spread. Wildfire is a raging, uncontrolled fire that erupts in natural environments like forests, grasslands, or prairies, and poses a severe threat to property, lives, and the ecological balance of the area (McKenzie et al. 2004).The urgency of evacuation in wildfire-prone regions cannot be overstated, given the speed at which these infernos can engulf vast swathes of land, homes, and natural habitats. Such Large-scale wildfires frequently result in mass evacuations, which cause societal upheaval, long-term infrastructure damage, and injuries and deaths among evacuees and first responders (Beverly & Bothwell, 2011; Caton et al., 2017). According to (U.S.Fire Administration 2019) USFA, the country saw an estimated yearly average of 1,344,100 wildfires between 2008 and 2017, resulting in over 3,000 deaths, 16,225 injuries, and \$14.7 billion in direct property losses.

Communities residing in wildland-urban interface (WUI) zones, characterized by undeveloped wildland, confront the greatest wildfire risk due to their proximity to highly flammable vegetation(Ahmad et al., 2024). Many of these communities are experiencing rapid population growth, but their transportation infrastructure remains inadequate to meet the growing traffic demand (Theobald & Romme, 2007). The inadequacy of the transportation system causes congestion and traps the evacuees especially during wildfire evacuations. Moreover, delayed evacuation trigger alarms or delays in evacuation advice implementation are other problems that force locals to stay until the last minute and face hazardous situations. This highlights the significance of a well-coordinated, timely evacuation during a wildfire, as a disorganized evacuation is likely to result in loss of lives and damages (Beyki et al., 2023).

The primary focus of wildfire evacuation research has centered on human behavioral studies reliant on social science-based surveys. This emphasis stems from the lack of empirical data on traffic movement and clearance time in existing evacuation studies (Kuligowski, 2021). However, in recent years, substantial datasets on human mobility have surfaced due to the widespread use of sensor technologies, i.e. Global Positioning System (GPS) devices, in various modes of transportation and the ubiquitous presence of mobile phones. This abundance of data has become a cornerstone in shaping the evolving standard of smart cities (Ramezani and Geroliminis 2015). Multiple independent third-party entities, such as Waze, INRIX, TomTom, and HERE, are now actively accumulating extensive crowdsourced data to offer accurate real-time traffic insights (Ahmad et al., 2023). Additionally, prominent big data firms such as Otonomo and Wejo have emerged, who specializes in organizing and standardizing historical data obtained from connected vehicles (CVs). These companies share such data to support research and development endeavors among vehicle manufacturers, researchers, and technology developers in the industry(Zanwar et al., 2021).

Although this anonymized CV dataset has recently become available to explore historical traffic patterns at the location of interest, several researchers have already effectively tested Wejo's vehicle movement dataset(Desai et al., 2021; Li et al., 2021). Therefore, given the lack of empirical data and the limitations of existing travel data collection technologies in acquiring such information, the datasets derived from connected vehicles (CVs) presents an exceptional

opportunity for accurately analyzing and modeling historical wildfire events. The travel time information obtained from CV data can be utilized to strategize and coordinate evacuation routes and timings, pinpoint and resolve potential bottlenecks, and communicate accurate estimates of travel time and fastest evacuation routes to the public (Bahaaldin et al., 2017).

The advent of autonomous vehicles (AVs) marks a groundbreaking shift in modern transportation. These vehicles operate without direct human input, relying on advanced sensors and AI algorithms(Tyagi & Aswathy, 2021). The potential impact of AVs extends across multiple industries, offering prospects of safer roads, more efficient transportation systems, and potential environmental advantages. Unmanned aerial autonomous vehicles have already found application in disaster management, especially in wildfire monitoring (Saffre et al., 2022). AVs present substantial potential in wildfire evacuations by potentially furnishing real-time updates on road conditions, adeptly navigating evacuation routes, and maneuvering through challenging terrains, such as areas with limited visibility due to smoke(Liu et al., 2023). However, it is crucial to thoroughly investigate the potential application of AVs in emergency scenarios, considering the complexities associated with the interaction of AVs with other vehicles and the environment. Advanced data technologies like CV data can be instrumental in assessing the effectiveness of AVs in evacuation strategies and enhancing emergency responses during wildfires.

Thus, the primary objective of this study is to utilize a CV dataset for a comparative analysis between the actual travel time observed during a real-life wildfire incident and the travel time data provided by the State authorities. Additionally, this study aims to conduct simulations of a historical wildfire event to assess the potential impact of Autonomous Vehicles (AVs) on clearance time.

MATERIAL AND METHODS

Case Study of Knolls Fire 2020. The fire event selected for evaluation and comparison is the Knolls Fire 2020 that occurred in Saratoga Springs, Utah on 28 June 2020. The fire initially broke out between 2:00 pm and 2:30 pm to the east of Lake Mountain, south of Saratoga Springs, and rapidly spread towards the city, fueled by strong 60 mph gusts of wind (Tribune, 2020). The city is bordered by Utah Lake on its eastern side and Lake Mountain on its western side. State Route 68 (SR-68), commonly known as 'Redwood Road,' serves as the primary exit route from the city. After the fire ignited, mandatory evacuation orders were issued for over 3,100 homes, impacting approximately 13,000 residents, which equates to nearly a third of the city's total population. The evacuation commenced at 2:45 pm, starting with the southern neighborhoods of the city. The residents in these areas were asked to evacuate immediately amidst strong winds, smoke, and dust. Following that, later in the afternoon, all individuals living south of Grandview Boulevard on the western stretch of Redwood Road were advised to evacuate their houses due to the rapid spread of the fire (Fox13, 2020). A shelter area was established at Westlake High School for the evacuees on Redwood Road, which was used to redirect Northbound traffic. Moreover, southbound traffic was blocked by downed power lines near the fire perimeter (Fox13, 2021; Radio, 2020). The evacuation traffic on NB SR-68 gradually resulted in heavy traffic congestion and vehicle queues, as shown by the operator response notes obtained from Utah department of Transportation (UDOT) in Table 1.

Data collection and processing. This study used WEJO's vehicle trajectory data which was collected during a 10-day period for the study area, i.e. from June 20th to June 29th, 2020, to

evaluate the travel time variation between evacuation and non-evacuation days. A description of key attributes of the dataset can be found in Table 2.

Table 1. Utah DOT operator response notes, Knolls Fire 2020

Time Stamp	Operator Note
6/28/2020 3:48:04 pm	A fire has started on the West side of Utah Lake on Lake Mountain. Redwood Road is being closed both directions due to the fire. Northbound is being closed at milepost 0 at the US-6 junction. Southbound is being closed at milepost 28 at Lake View Terrace in Saratoga Springs. Reason(s) for IPR: Media Attention. Route(s) Affected: Redwood Rd, US-6, I-15, SR-145. On Scene: UHP, Local PD, Fire. Incident Mgr: TOC. Action(s) Taken: Notified Region 3, Primary PIO and TOC Mgmt. Queue Length: None at this time. Delays: To detour around Redwood Road onto I-15 and US-6 adds 10 minutes of travel time. Estimated Duration: Unknown. Next update in ~60 mins
6/28/2020 4:42:56 pm	The UDOT Maintenance Sheds are setting up hard closures at the closure points. Traffic is congested throughout the Saratoga Springs area due to evacuations in the southern section of Saratoga. Delays in the area are 5-10 minutes at this time.
6/28/2020 5:44:11 pm	NB Redwood Rd is congested for 4 miles in Saratoga Springs. Delays are 10 minutes.
6/28/2020 6:34:08 pm	NB Redwood Rd in Saratoga is now congested for 4 miles with 20-minute delays.

Table 2. Description of WEJO Data Attributes

Data attributes	Description
Datapoint ID	A unique identifier for an individual captured datapoint every 3 seconds.
Journey ID	A unique identifier for an individual vehicle's movement through to an ignition off event happening.
Timestamp	Time and date of each datapoint along with location time zone offset.
Heading	Heading of each datapoint with 0 = North moving clockwise to 359°.
Speed	The speed of vehicle at each datapoint.
Latitude	North-South positioning of the vehicle on the Earth's surface.
Longitude	East-West positioning of the vehicle on the Earth's surface.

The study area for travel time calculation consisted of 5 miles of NB SR-68 roadway between mile markers (MM) 25-30. This section of SR-68 consisted of several connecting roads

and five traffic signals allowing vehicular traffic to enter and exit the affected neighborhoods as shown in Figure 1. The raw dataset comprised of more than eleven million datapoints covering the entire city, which had to be segregated to include only datapoints from the studied section of SR-68. The heading attribute associated with each datapoint was used to filter out the traffic heading south.

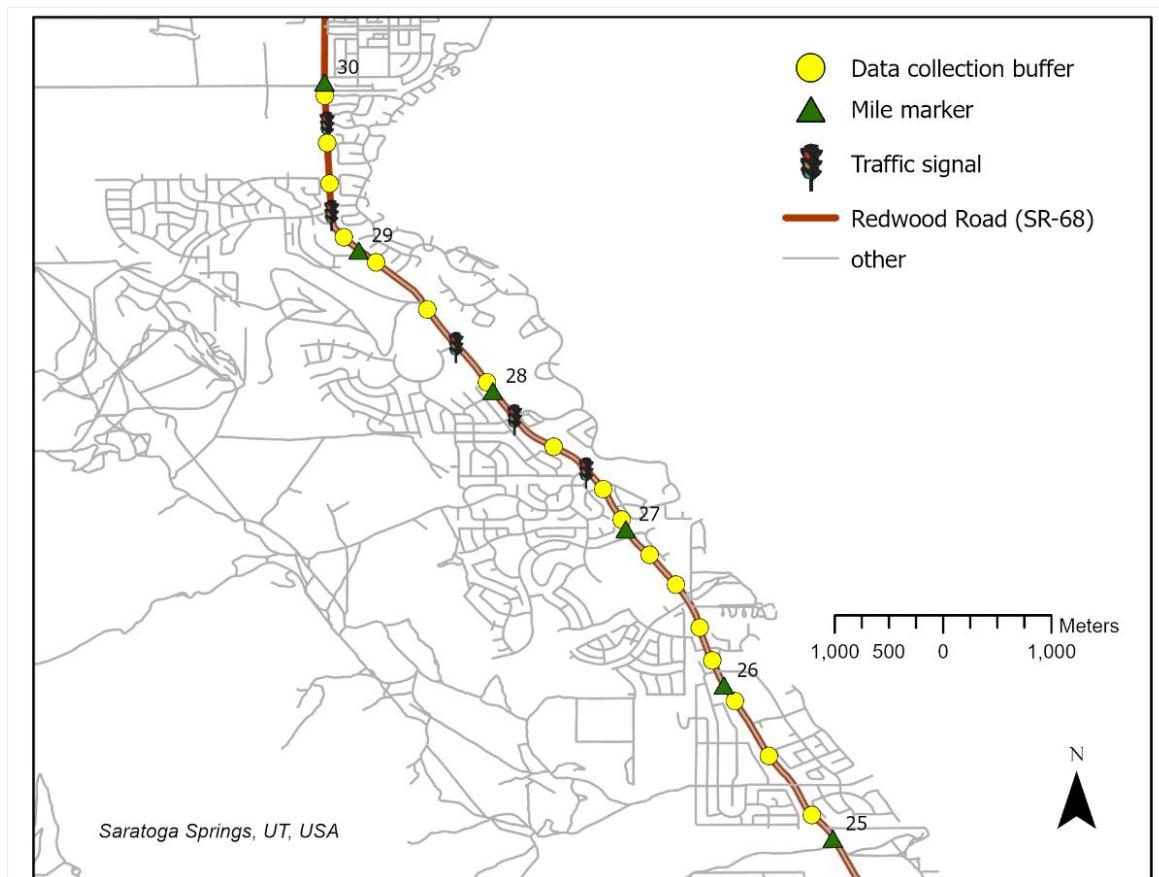


Figure 1. Study Area Map

Data Analysis. The processed data was subsequently analyzed using geographic information system (GIS) and Microsoft excel to perform a comparative analysis of travel time between non-evacuation and evacuation days.

Scenario Simulation. To comprehensively evaluate the potential impacts of autonomous vehicles in wildfire evacuation scenarios, we established a base scenario using PTV VISSIM, which is a powerful microsimulation tool for analyzing networks at intersections, freeways, and different segments (VISSIM, 2020). The study utilized a VISSIM model previously employed in our study (Ahmad et al., 2023), to simulate conditions along a 3-mile segment of SR-68, spanning from Mile Marker 27 to 30, including five traffic signals and interconnected roads. The model was validated with a similarity assessment by comparing its outcomes with historical performance metrics obtained from the Utah Department of Transportation's (UDOT) ATSPM website. The validation process confirmed that the baseline scenario accurately replicated travel times comparable to those observed on the actual evacuation day.

Following the establishment of the baseline scenario, a case study was conducted to evaluate the potential outcomes of an evacuation scenario exclusively utilizing autonomous vehicles. After a thorough examination of car-following models, which involved rigorous adjustments and recommendations, the car-following models identified in accordance with (Chen et al., 2019), were selected for this study. The adopted values are listed in Table 3 and Table 4.

Table 3. Adopted values for the W99 car-following model for human-driven vehicles

W99 Parameters Human-Driven	Definition	VISSIM Default	Adjusted Value
CC0 (m)	Standstill Distance	1.5	4.45
CC1(s)	Headway Time	0.9	0.87
CC2 (m)	Following Variation	4	5.28
CC3 (s)	Threshold for Entering Following	-8	-7.92
CC4 (m/s)	Negative Following Threshold	-0.35	-1.52
CC5 (m/s)	Positive Following Threshold	0.35	1.52
CC6 (-)	Speed Dependency of Oscillation	11.44	0.71
CC7 (m/s ²)	Oscillation Acceleration	0.25	0.31
CC8 (m/s ²)	Standstill Acceleration	3.5	1.03
CC9 (m/s ²)	Acceleration at Speed of 80 km/h	1.5	0.33

Table 4. Adopted values for the W99 car-following model for autonomous vehicles

W99 Parameters Autonomous Vehicles	Definition	VISSIM default Values for all Weather
CC0 (m)	Standstill Distance	1.5
CC1(s)	Headway Time	0.9
CC2 (m)	Following Variation	0
CC3(s)	Threshold for Entering Following	-8
CC4 (m/s)	Negative Following Threshold	-0.1
CC5 (m/s)	Positive Following Threshold	0.1
CC6 (-)	Speed Dependency of Oscillation	0
CC7 (m/s ²)	Oscillation Acceleration	0.1
CC8 (m/s ²)	Standstill Acceleration	3.5

RESULTS AND DISCUSSION

Comparative Analysis Results. This section shows a comparative analysis of the average travel time for the 5-mile stretch during the 10-day study period when NB SR-68 was affected by the Knolls Fire in 2020, as shown in Figure 2. The analysis reveals a consistent travel time on non-evacuation days, whereas a substantial increase in travel time was noticed on the evacuation day.

The most significant traffic delays occurred during evening peak hours, notably shown by slow-moving traffic between MM 26-30 after 6 o'clock. The observed traffic conditions corresponded to the operator response notes provided by UDOT in table 1, which highlights significant traffic delays on the analyzed section of NB SR-68. This shows that the evacuees

were stuck on SR-68 for a considerable amount of time, which completely jeopardized their safety.

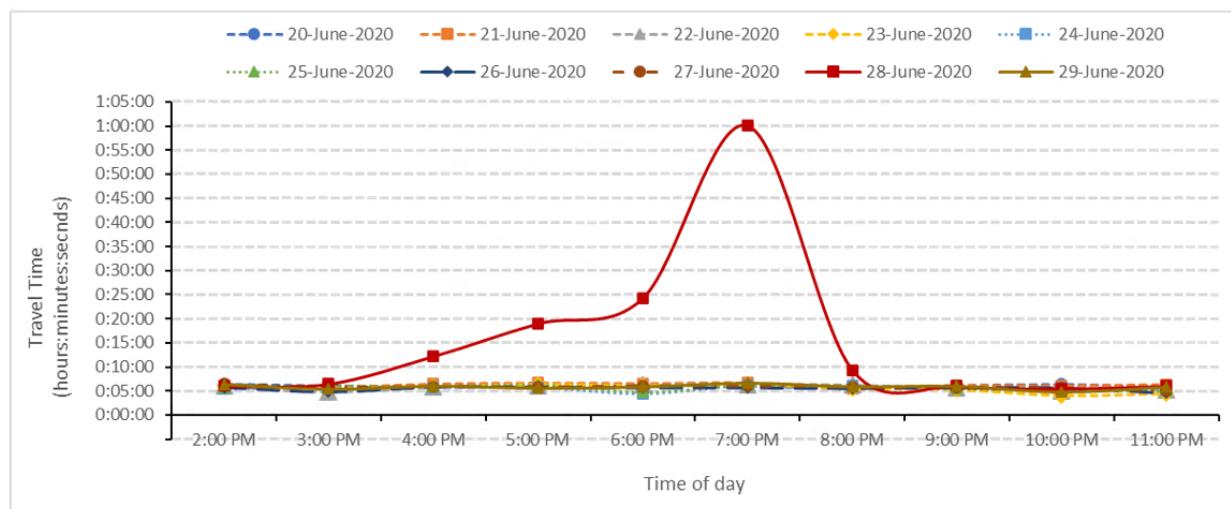


Figure 2. CV travel time calculation results

VISSIM Simulation Results. The autonomous vehicle simulation showcased favorable trends compared to the base scenario as depicted in Figure 3. Among the 9 travel-time measuring segments, autonomous options exhibited an advantage in six segments. In the remaining segments where human-driven vehicles performed better than AVs, the travel time difference was consistently under 10%.

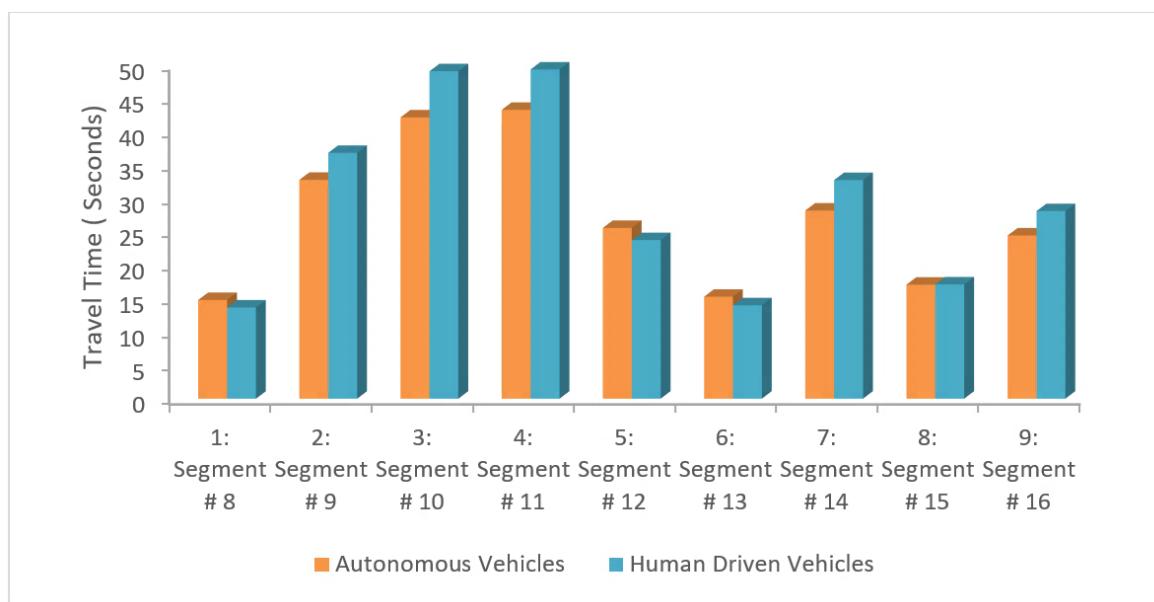


Figure 3. Simulation travel time comparison

CONCLUSION

In conclusion, our comparative analysis, especially through simulations of emergency evacuation scenarios, provides insight into the potential impact of autonomous vehicles. Our findings emphasize significant reductions in travel time by integrating autonomous vehicles into evacuation traffic. By comparing different scenarios i.e. with and without autonomous vehicles, our study demonstrated positive outcomes, notably in decreasing evacuation travel times. These results underscore the promising role that autonomous vehicles could play in enhancing emergency evacuation efforts. The tangible benefits highlighted in our research strongly advocate for integrating autonomous technologies as a crucial component of future emergency response strategies. Future studies can expand on these findings by examining broader implications of integrating autonomous vehicles in emergency wildfire evacuations. In particular, a comprehensive exploration of safety aspects, including accident prevention, adaptability to dynamic conditions, and coordination with emergency response systems, would contribute to a more thorough understanding of the potential benefits and challenges associated with autonomous vehicles in such events.

ACKNOWLEDGEMENT

This research was funded by the U.S. Department of Transportation under the University Transportation Center (UTC), Contract Agreement No. 131987-Z912530 through CMMM project No. FAR0037360 and UTC Mountain Plains Consortium.

REFERENCES

- Ahmad, S., Ahmed, H. U., Ali, A., Yang, X., Huang, Y., Guo, M., Ren, Y., and Lu, P. (2024). Evaluating driving behavior patterns during wildfire evacuations in wildland-urban interface zones using connected vehicles data. *Fire Safety Journal*, 142, 104015.
- Ahmad, S., Ali, A., Ahmed, H. U., Huang, Y., and Lu, P. (2023). Evaluating traffic operation conditions during wildfire evacuation using connected vehicles data. *Fire*, 6(5), 184.
- Bahaaldin, K., Fries, R., Bhavsar, P., and Das, P. (2017). A Case Study on the Impacts of Connected Vehicle Technology on No-Notice Evacuation Clearance Time. *Journal of Advanced Transportation*, 2017.
- Beverly, J. L., and Bothwell, P. (2011). Wildfire evacuations in Canada 1980-2007. *Natural Hazards*, 59(1), 571–596.
- Beyki, S. M., Santiago, A., Laím, L., and Craveiro, H. D. (2023). Evacuation Simulation under Threat of Wildfire—An Overview of Research, Development, and Knowledge Gaps. *Applied Sciences*, 13(17), 9587.
- Camp Fire. (2022). California Department of Forestry and Fire Protection.
<Https://Www.Fire.ca.Gov/Incidents/2018/11/8/Camp-Fire/>.
<https://www.fire.ca.gov/incidents/2018/11/8/camp-fire/>.
- Caton, S. E., Hakes, R. S. P., Gorham, D. J., Zhou, A., and Gollner, M. J. (2017). Review of Pathways for Building Fire Spread in the Wildland Urban Interface Part I: Exposure Conditions. *Fire Technology*, 53(2), 429–473.

- Chen, C., Zhao, X., Liu, H., Ren, G., Zhang, Y., and Liu, X. (2019). Assessing the influence of adverse weather on traffic flow characteristics using a driving simulator and VISSIM. *Sustainability*, 11(3), 830.
- Desai, J., Saldivar-Carranza, E., Mathew, J. K., Li, H., Platte, T., and Bullock, D. (2021). Methodology for Applying Connected Vehicle Data to Evaluate Impact of Interstate Construction Work Zone Diversions. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2021-Septe*, 4035–4042.
- Fox13. (2020). Thousands evacuated as “Knolls Fire” threatens Saratoga Springs homes. <https://www.fox13now.com/news/local-news/knolls-fire-forces-evacuations-in-saratoga-springs>.
- Fox13. (2021). Human-caused wildfires continue to spread in Utah. <https://www.fox13now.com/news/local-news/human-caused-wildfires-continue-to-spread-in-utah>.
- Grünthal, G., Thielen, A. H., Schwarz, J., Radtke, K. S., Smolka, A., and Merz, B. (2006). Comparative Risk Assessments for the City of Cologne – Storms, Floods, Earthquakes. *Natural Hazards*, 38(1), 21–44.
- Hamideh, S., Sen, P., and Fischer, E. (2022). Wildfire impacts on education and healthcare: Paradise, California, after the Camp Fire. *Natural Hazards*, 111(1), 353–387.
- Kuligowski, E. (2021). Evacuation decision-making and behavior in wildfires: Past research, current challenges and a future research agenda. *Fire Safety Journal*, 120, 103129.
- Li, X., Martin, M., Dadashova, B., Salgado, D., and Samant, S. (2021). *Exploring Crowdsourced Big Data To Estimate Border Crossing Times*.
- Liu, J., Liu, Z., Jia, B., Zheng, S., and Ji, H. (2023). Managing traffic evacuation with multiclass connected and autonomous vehicles. *Physica A: Statistical Mechanics and Its Applications*, 625, 128985.
- McKenzie, D., Gedalof, Z., Peterson, D. L., and Mote, P. (2004). Climatic change, wildfire, and conservation. In *Conservation Biology* (Vol. 18, Issue 4, pp. 890–902).
- Radio, K. S. L. N. (2020). Knolls Fire Update: Evacuation order lifted but still possible, only 25% contained. <https://kslnewsradio.com/1928199/knolls-fire-update-evacuation-orders-still-in-effect-10000-acres-burned-25-contained/>.
- Ramezani, M., and Geroliminis, N. (2015). Queue Profile Estimation in Congested Urban Networks with Probe Data. *Computer-Aided Civil and Infrastructure Engineering*, 30(6), 414–432.
- Saffre, F., Hildmann, H., Karvonen, H., and Lind, T. (2022). Monitoring and Cordoning Wildfires with an Autonomous Swarm of Unmanned Aerial Vehicles. *Drones*, 6(10), 301.
- Smith, A. B. (2022). *US Billion-dollar weather and climate disasters, 1980-present*.
- Statista. (2023). Natural Disasters in the U.S. <Https://Www.Statista.Com/Topics/1714/Natural-Disasters/#topicOverview>.
- Theobald, D. M., and Romme, W. H. (2007). Expansion of the US wildland-urban interface. *Landscape and Urban Planning*, 83(4), 340–354.
- Thompson, R. R., Garfin, D. R., and Silver, R. C. (2017). Evacuation from natural disasters: a systematic review of the literature. *Risk Analysis*, 37(4), 812–839.
- Tribune, T. S. L. (2020). 1 home destroyed, 12 damaged as Knolls Fire grows to 10,000 acres; thousands still under evacuation in Saratoga Springs. <https://www.sltrib.com/news/2020/06/28/draper-lehi-residents/>.

- Tyagi, A. K., and Aswathy, S. U. (2021). Autonomous Intelligent Vehicles (AIV): Research statements, open issues, challenges and road for future. *International Journal of Intelligent Networks*, 2, 83–102.
- US Fire Administration. (2019). *Fire in the United States 2008-2017*.
- PTV VISSIM. (2020). *5.30-05 User Manual, 2011*. PTV AG. Karlsruhe, Germany.
- Zanwar, P., Kim, J., Kim, J., Manser, M., Ham, Y., Chaspari, T., and Ahn, C. R. (2021). Use of Connected Technologies to Assess Barriers and Stressors for Age and Disability-Friendly Communities. *Frontiers in Public Health*, 9, 578832.
- Zuzak, C., Mowrer, M., Goodenough, E., Burns, J., Ranalli, N., and Rozelle, J. (2022). The national risk index: establishing a nationwide baseline for natural hazard risk in the US. *Natural Hazards*, 114(2), 2331–2355.

Optimizing Lane Configuration for Efficient Platoon Control during Lane Closures

Vamsi Krishna Maddineni¹ and Mohamed M. Ahmed, Ph.D., P.E.²

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: maddinva@mail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamed.ahmed@uc.edu

ABSTRACT

In the era of connected and autonomous vehicles (CAVs), optimizing transportation infrastructure is crucial to unlock the full potential of connected and autonomous technology (CAT), particularly in the context of advanced operations like platooning. Considering the case, this study focused on infrastructure modifications, with a specific focus on low-automation CAVs. The study employed PTV VISSIM to model a variety of scenarios on a freeway segment, considering six distinct lane configuration scenarios, varying market penetration rates (MPRs), and factors such as platoon allowance/restriction within a 1-mi buffer upstream of lane closures, as well as open or closed lane conditions (dedicated lanes). The findings revealed that, for a two-lane segment, restricting platoons to open lanes and allowing them on upstream 1-mi resulted in a 28% reduction in the total number of conflicts, with time-to-collision (TTC) serving as the conflict indicator. The worst-case scenario involves allowing platoons without any lane restrictions, leading to a substantial increase in conflicts at higher MPRs.

KEYWORDS: CAV, CAT, TTC, Lane closures, PTV VISSIM, platooning, Traffic safety, Early merge

INTRODUCTION

In recent years, the field of transportation has witnessed rapid technological advancements, marked by successive revolutionary innovations. While intelligent transportation systems (ITS), such as automated vehicles, have been in development for several decades, recent initiatives indicate an accelerated timeline for widespread implementation. Researchers have shown considerable interest in various applications of vehicle automation, seeking to identify those with the most profound implications for the future of transportation (Fagnant and Kockelman, 2015; Wadud et al., 2016). Fundamentally, connected and automated vehicles (CAV) are poised to assist drivers by delegating some or all driving tasks to computerized systems. Additionally, CAVs leverage vehicle-to-vehicle (V2V) connectivity to exchange information with nearby vehicles and adapt their behavior accordingly. The Society of Automotive Engineers has established a framework comprising six levels of vehicle automation, ranging from level zero (no automation) to level five (full automation) (SAE. 2018). These levels categorize vehicles based on their implementation of various advanced driving assistance (ADA) technologies, with level five designating vehicles capable of executing all driving functions without human intervention.

During the early phases of Connected and Automated Vehicle (CAV) deployment, when both CAVs and human-driven vehicles share the same lanes, several concerns arise. Given that CAV platoons tend to follow shorter headways, this can potentially create significant challenges

for other vehicles when encountering platoons or attempting lane changes nearby. To effectively address these challenges, it is imperative to conduct a comprehensive assessment of platoon behavior across different freeway sections. This assessment encompasses the implementation of various lane-based configurations and strategies aimed at managing CAV operations and fostering a safer traffic environment. One frequently considered approach is the allocation of Dedicated Lanes (DLs) exclusively for CAVs, which reduces interactions between CAV platoons and human-driven vehicles while improving the possibility of platoon formation (Dehman et al. 2021), leading to a higher prevalence of CAV platoons in these lanes. Nonetheless, it is crucial to recognize that Dedicated Lanes may become less efficient when the Market Penetration Rate (MPR) is very low (Razmi Rad et al. 2020). A proposed solution could be Human-Driven Vehicles (HDVs) sharing dedicated lanes with CAVs during periods of very low MPR, with a subsequent transition to fully dedicated CAV lanes once a predetermined MPR threshold is reached (Razmi Rad et al. 2020).

While the driver behavior is improved with CAV's, it is still not clear on how these vehicles when form into platoons react to work zones. It was reported that a work zone fatality occurs every 15 h and this also involved construction workers. Although many studies have utilized microsimulation to assess the impacts of CAVs near work zones. Most of them focused on operational advantages and safety improvements due to better communication V2I and V2V. But, there has been a dearth in the number of studies that took lane configurations as a key factor in their assessment. Additionally, there is a need for comparative analysis between mixed lanes, shared lanes and dedicated lanes at various MPRs. This paper aims to fulfill this gap by focusing on different lane configurations and their role in improving both safety and operations of platoons near work zones, this in a way helps to provide more insights for Infrastructure Owners and Operators, Traffic agencies for coming up with better ways to adopt the upcoming advances in CAT technology.

BACKGROUND

Microscopic traffic simulation models have been extensively utilized for conducting impact analyses of CAVs. Considerable efforts have been dedicated to modeling mixed traffic flow incorporating CAVs. Previous studies have employed various simulation software platforms for CAV simulation, including VISSIM, AIMSUN, PARAMICS, and SUMO. These software tools have proven useful in simulating and analyzing the behavior and interactions of CAVs within a broader traffic context, allowing researchers to explore the potential impacts of CAVs on traffic flow, efficiency, and safety (Fountoulakis et al. 2017; Shi et al. 2019; Yang et al. 2017). Among these packages, VISSIM software developed by PTV group is one of most utilized simulation platform among the researchers (Gora et al. 2020).

Platooning logic

This study extensively explores lane configuration strategies aimed at mitigating safety issues resulting from mandatory lane merges involving conventional vehicles and CACC platoons. This will place greater emphasis on the development of various scenarios and assumptions, all of which are informed by the internal platooning logic in PTV VISSIM.

The maximum platoon size is limited to five vehicles, with a gap time of 0.6 seconds between CACC vehicles in a platoon. This aggressive following behavior assumption aims to

emphasize safety concerns in platoons while preserving their operational and environmental advantages through close spacing.

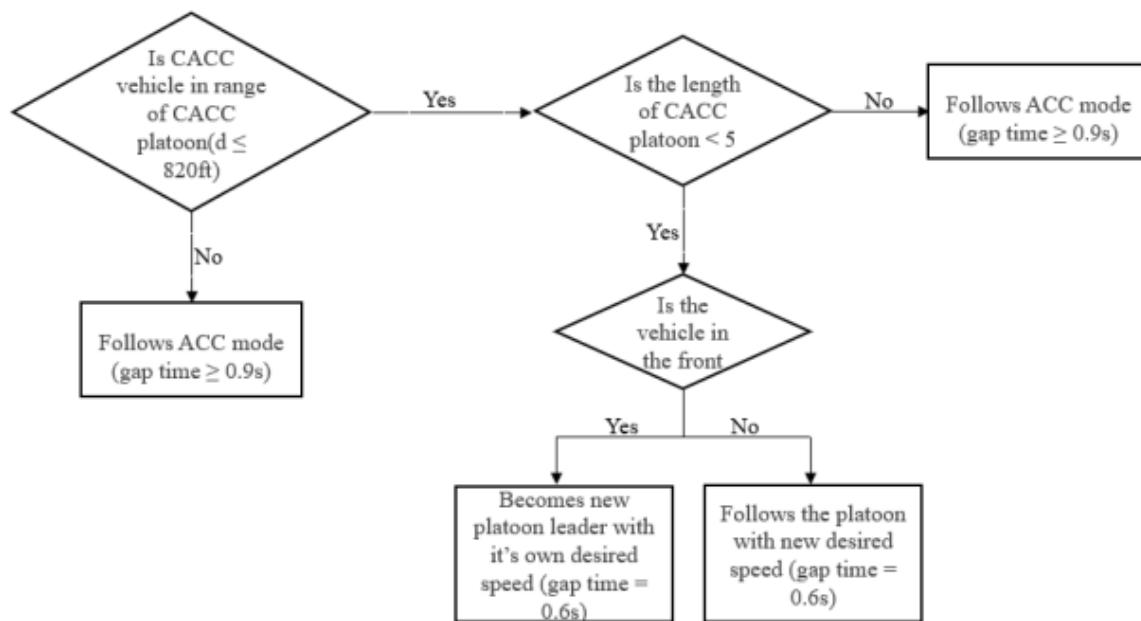


Figure 1: Internal platooning logic (PTV VISSIM)

Vissim model

The simulations were conducted on a two-lane section of a standard freeway. This segment consisted of an unobstructed freeway stretch spanning three miles, followed by an upstream work zone extending for one mile, and then the actual work zone, also covering one mile in length. The parameters for vehicle inputs were established at the freeway's entry point, with a traffic volume of 1000 vehicles per hour. The composition of vehicles in this flow was modeled after that of rural freeway I-80, with 40% heavy goods vehicles (HGVs) and 60% passenger cars.

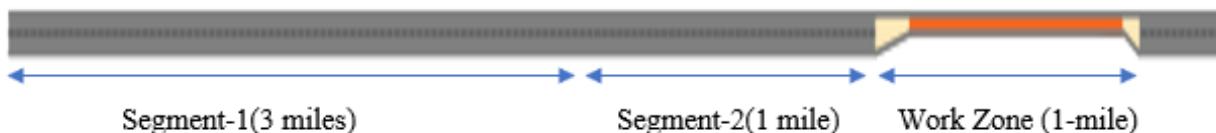


Figure 2: Freeway section modelled in VISSIM

Simulation parameters and other settings

The study adheres to the car-following parameters established by Wiedemann in 1999, utilizing internal settings within VISSIM. Additionally, lane-change parameters have been refined and updated based on the insights gained from the SHRP2 project. A comprehensive summary of these modifications is provided in the accompanying table.

The number of simulation runs for each scenario were kept to 10 with different seeds and simulation resolution was 20 timesteps/simulation sec.

Table 1: Driver behavior parameters

Car following parameters		Lane change parameters	
CC0 (Standstill distance (ft))	4.92	Own vehicle	
CC1 (Gap time distribution (s))	0.9	Maximum deceleration(ft/s ²)	-11.38
CC2 ('Following' distance oscillation (ft))	0	Accepted deceleration(ft/s ²)	-0.11
CC3 (Threshold for entering 'following' (s))	-8	Trailing vehicle	
CC4 (Negative speed difference (ft/s))	-0.1	Maximum deceleration(ft/s ²)	-1.78
CC5 (Positive speed difference (ft/s))	0.1	Accepted deceleration(ft/s ²)	-0.09
CC6 (Distance dependency of oscillation)	0	Other lane change Parameters	
CC7 (Oscillation acceleration (ft/s ²))	0.33	Minimum clearance(ft)	104.41
CC8 (Acceleration from standstill (ft/s ²))	11.48	Cooperative Lane change	Yes
CC9 (Acceleration at 50 mph (ft/s ²))	4.92	Advanced merging	Yes

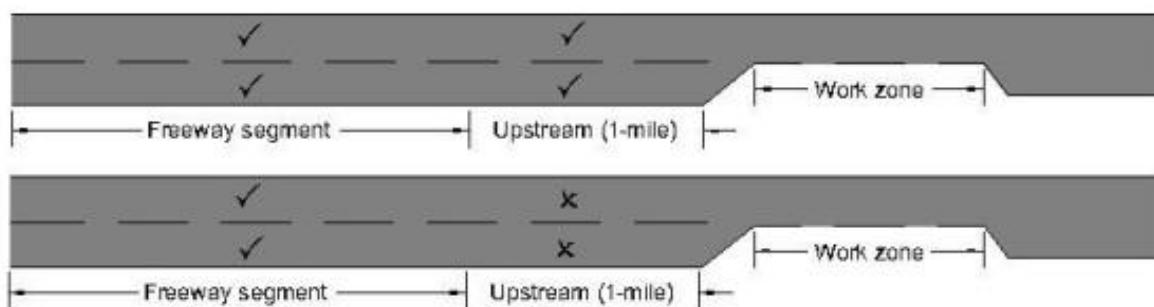
Scenario Development

In this study, we implemented work zone warnings one mile ahead of the transition zone, prompting vehicles to follow conventional early merge strategies. We examined platoons under six conditions, based on three lane configurations. The first condition allowed platoons to use both lanes on the freeway without restrictions. In the second and third conditions, platoons were confined to a single lane due to one lane being closed in the work zone, resulting in platoons using either the open or closed lane. Within these three conditions, we also considered whether platoons were permitted to merge upstream, resulting in six distinct conditions for a two-lane freeway with one lane closed.

Base scenario: All the vehicles are human driven with right lane closure at the work zone

Lane configuration based platooning scenarios:

- 1) Platooning allowed on both lanes (platooning allowed upstream (✓), not allowed upstream (✗))

**Figure 3: No restrictions for platooning scenario**

- 2) Platooning allowed only on open lane (platooning allowed upstream (✓), not allowed upstream (✗))

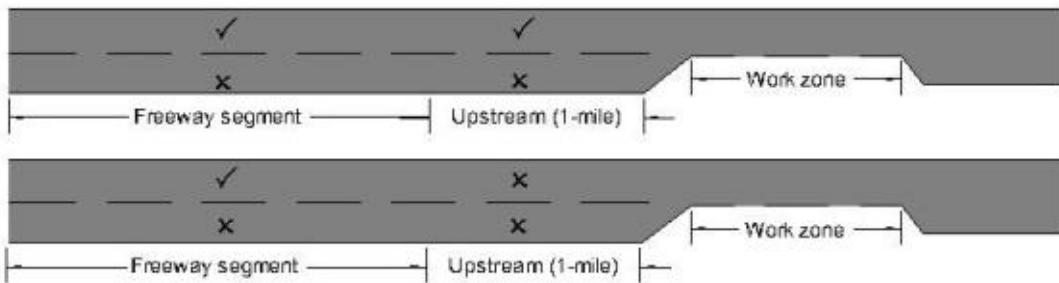


Figure 4: Platooning lane is open

- 3) Platooning allowed only on closed lane (platooning allowed upstream (✓), not allowed on upstream (✗))

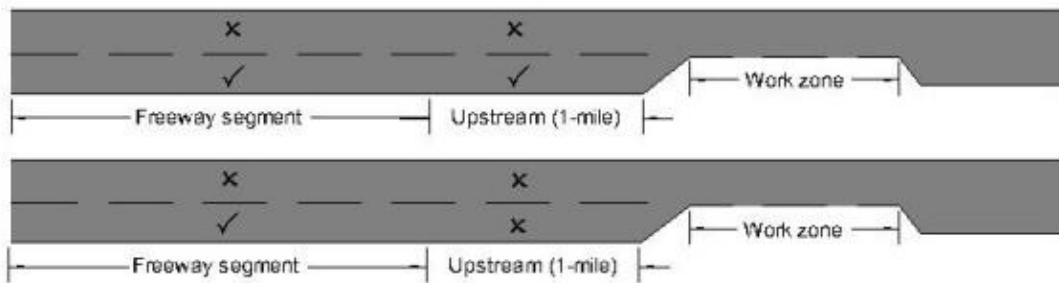


Figure 5: Platooning lane is closed

RESULTS AND DISCUSSION

Traffic conflict analysis

This study utilized TTC threshold of 1.5 s and traffic conflicts were extracted using SSAM tool based on the trajectory outputs from microsimulation.

Table 2: Traffic conflict frequency distribution for various MPR (%CACC)

Lanes used by platoons, upstream (1-mile) condition for platooning	Average no. of conflicts(hr) for different MPR(%CACC)				
	0%	20%	40%	60%	80%
Both lanes, Yes	149.4	109	158	198	303
Both lanes , No	149.4	170.4	166.2	181.4	268.8
Open lane, Yes	149.4	139	112.2	104.6	114
Open lane, No	149.4	163.8	131.6	125.2	126.6
Closed lane, Yes	149.4	136.2	137.5	138.6	158.3
Closed lane, No	149.4	136	137	129	142.5

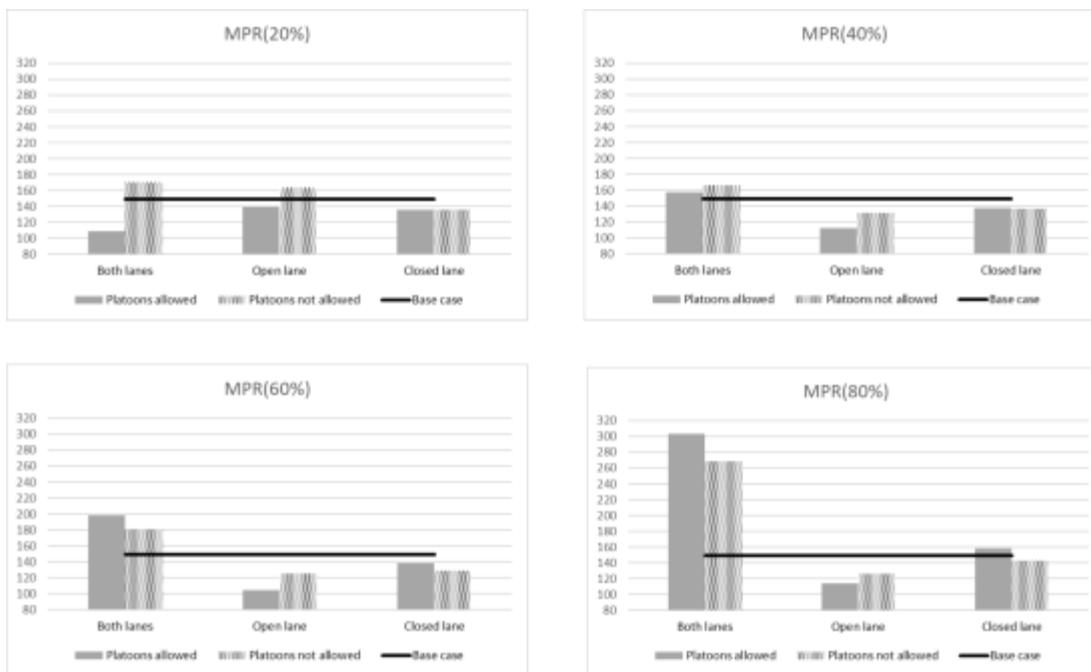


Figure 6: Traffic conflict frequency distribution vs MPR(%CACC)

The results indicate that utilizing both lanes for platooning may compromise safety as the Minimum Platoon Ratio (MPR) increases, primarily due to increased interactions between Heavy-Duty Vehicles with Cooperative Adaptive Cruise Control (HDV-CACC) and Cooperative Adaptive Cruise Control with Cooperative Adaptive Cruise Control (CACC-CACC) vehicles during lane merges. Interestingly, in both open and closed lanes, there was no significant increase in traffic conflicts with increasing MPR, likely because platooning in a single lane reduced interactions between CACC-CACC vehicles. Among the scenarios considered, platooning in open lanes during lane closures was the most suitable, as conflicts decreased, suggesting fewer interactions between CACC vehicles and HDV, thus improving safety.

To determine the overall percentage of CACC vehicles participating in platoons across these scenarios, we processed the raw ".fzp" files from VISSIM using Python's PANDAS library. This involved calculating the percentage of CACC vehicles involved in platooning at each time step and subsequently computing the average platoon percentage across all time steps.

Table 3: Total percentage of CACC vehicles involved in platooning for various MPR (%CACC)

Lane utilized by platoons & upstream platooning condition	% of CACC vehicles in platooning for different MPR(%CACC)			
Both lanes, Yes	20%	40%	60%	80%
Both lanes , No	23.4	37.4	53.1	59.6
Open lane, Yes	11.0	17.3	27.7	33.2
Open lane, No	13.7	17.3	25.9	30.2
Closed lane, Yes	7.1	8.4	13.2	14.8
Closed lane, No	9.3	17.4	20.8	25.4
Both lanes, Yes	7.7	10.2	13.1	16.0

Analysis of the results from both tables reveals a clear trend: a higher percentage of platoons occurs when both lanes are utilized for platooning, especially at higher Minimum Platoon Ratios (MPR), compared to using a single lane. Notably, an intriguing finding emerges; in the open lane condition at 80 MPR, there is a greater number of Cooperative Adaptive Cruise Control (CACC) platoons (30.2%) than in the closed lane condition (25.4%), while the number of conflicts is notably lower by 28%. This suggests that the open lane configuration for platooning is the most favorable scenario in terms of safety near work zones.

Operational characteristics evaluation

From the simulation runs, Average speeds, Delay were calculated and to monitor the Queue characteristics, a Queue detector was placed at 600 ft before complete lane closure in the transition zone where vehicles unable to merge on upstream have to stop and make necessary lane changes. All of these results were observed for various MPR of CACC vehicles in different lane configurations

Table 4: Avg. speed, Total Avg. delay for different MPR (%CACC)

Lane configuration & platooning condition	Average speed(mph) for different MPR(%CACC)					Total Avg. delay(hr) for different MPR(%CACC)				
	0%	20%	40%	60%	80%	0%	20%	40%	60%	80%
Both lanes, Yes	55.9	58.6	59.6	61.4	61.8	71.6	57.6	52.1	42.6	39.5
Both lanes , No	55.9	55.9	56.3	57.9	58.5	71.6	72	70.0	61.4	58.7
Open lane, Yes	55.9	56.9	58.9	60.5	61.4	71.6	67.2	56.8	48.6	44.6
Open lane, No	55.9	56.2	57.6	58.6	59.2	71.6	71	64.0	59.6	56.8
Closed lane, Yes	55.9	57.9	59.3	60.1	61.1	71.6	62.1	55.2	51.9	47.2
Closed lane, No	55.9	57.9	59.1	59.1	59.9	71.6	62	56.9	57.1	53.6

Table 5: Avg. Queue length observed in transition zone vs MPR(%CACC)

Lane configuration & platooning condition	Avg. Queue length observed before lane closure segment for different MPR(%CACC)				
	0%	20%	40%	60%	80%
Both lanes, Yes, B(Y)	346	167.5	71	27.4	8.7
Both lanes, No, B(N)	346	307.8	245.6	156.3	121.5
Open lane, Yes, O(Y)	346	273.4	133.7	47.9	23.6
Open lane, No, O(N)	346	315.1	207.7	135.8	109.4
Closed lane, Yes, C(Y)	346	239.4	118.3	115.7	84.5
Closed lane, No, C(N)	346	220.7	137	119.5	102.9

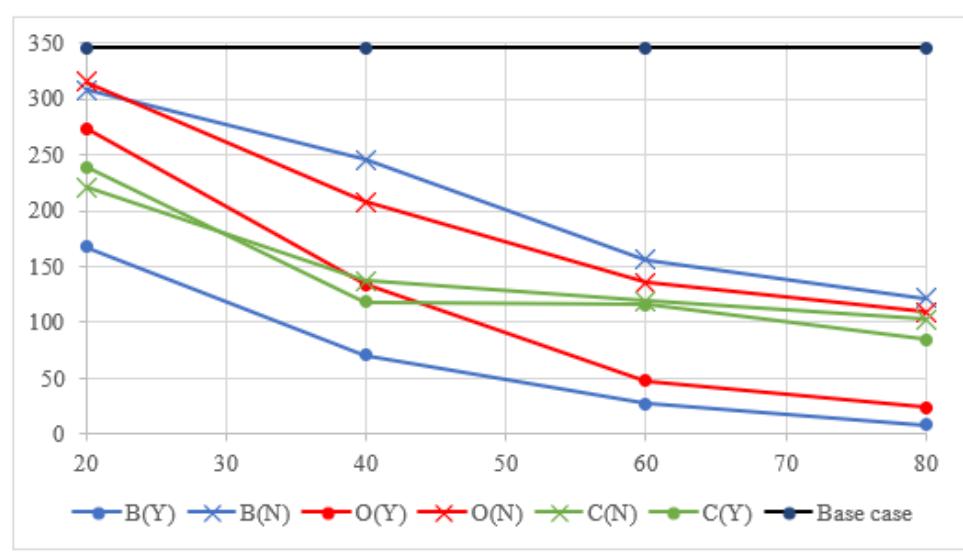


Figure 7: Avg. queue length (ft) vs MPR (%CACC)

Analysis of the tables indicate that as the Market Penetration Rate (MPR) increases, operational performance also sees improvements. The scenario allowing platooning on both lanes, including the upstream approach, exhibits the best operational results, but at the compromise of safety performance as it has resulted in highest number of traffic conflicts. Surprisingly, the open lane scenario closely follows the former in terms of operational performance, with marginal differences. At the same time, the open lane scenario consistently excels in safety compared to all other scenarios. These results suggest that permitting platooning on open lanes near work zones is not problematic. However, for any other lane configurations, it is recommended that vehicles disengage from platooning at least one mile upstream of the work zone to ensure improved safety. Remarkably, platoons allowed on open lanes demonstrate superior performance in both operational and safety aspects compared to other alternate scenarios.

CONCLUSIONS

This study aimed to determine the optimal lane configurations for enhancing the adoption of platoons on two-lane freeways, focusing on both safety and operational efficiency. Given the forthcoming challenge of accommodating low automation Cooperative Adaptive Cruise Control (CACC) platoons on freeways, this investigation assumed Level 1 platooning, wherein only longitudinal behavior was automated. Lane parameters were derived from NDS SHRP2 data (Chapter 3) for clear weather conditions to update lateral behavior. To simulate interactions between CACC-equipped vehicles and conventional vehicles, a work zone was introduced at the end of the freeway, necessitating mandatory lane changes for merging into a single lane. The simulations encompassed three distinct lane configurations (both lanes open, open lane, closed lane) for platooning, each with two scenarios based on upstream platooning conditions (1-mile), resulting in six scenarios. Initial safety evaluations revealed that the open lane condition for platooning exhibited the lowest number of conflicts, outperforming other scenarios. Conversely, when platooning was unrestricted across all lanes and upstream (1-mile), it yielded the highest

conflict count. Operational performance improved across all scenarios with increasing Market Penetration Rate (MPR). The best operational performance was observed when platooning was allowed on both lanes and upstream, followed closely by the open lane condition. Overall, the analysis indicated that the open lane scenario offered the best combination of safety and operational performance. Unrestricted platooning enhanced operations but compromised safety, while the closed lane condition proved beneficial when platooning was dissolved before reaching the upstream work zone.

The findings from this study highlight the benefits of utilizing lane specific platooning and could benefit traffic agencies, Infrastructure Owners and Operators (IOOs) in planning efficient strategies for adopting CAT infrastructure in the near future.

REFERENCES

- Fagnant, D. J., and K. Kockelman. "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations." *Transportation Research Part A: Policy and Practice* 77 (2015): 167-181.
- Wadud, Z., D. MacKenzie, and P. Leiby. "Help or Hindrance? The Travel, Energy and Carbon Impacts of Highly Automated Vehicles." *Transportation Research Part A: Policy and Practice* 86 (2016): 1–18.
- Fountoulakis, M., N. Bekiaris-Liberis, C. Roncoli, I. Papamichail, and M. Papageorgiou. "Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing." *Transportation Research Part C: Emerging Technologies* 78 (2017): 13-33.
- Shi, Y., Q. He, and Z. Huang. "Capacity analysis and cooperative lane changing for connected and automated vehicles: Entropy-based assessment method." *Transportation research record* 2673, no. 8 (2019): 485-498.
- Yang, H., Z. Wang, and K. Xie. "Impact of connected vehicles on mitigating secondary crash risk." *International journal of transportation science and technology* 6, no. 3 (2017): 196-207.
- Gora, P., and I. Rüb. "Traffic models for self-driving connected cars." *Transportation Research Procedia* 14 (2016): 2207-2216.
- Rad, S. R., H. Farah, H. Taale, B. van Arem, and S. P. Hoogendoorn. "Design and operation of dedicated lanes for connected and automated vehicles on motorways: A conceptual framework and research agenda." *Transportation research part C: emerging technologies* 117 (2020): 102664.
- Dehman, A., and B. Farooq. "Are work zones and connected automated vehicles ready for a harmonious coexistence? A scoping review and research agenda." *Transportation research part C: emerging technologies* 133 (2021): 103422.
- Das, A., and M. M. Ahmed. "Adjustment of key lane change parameters to develop microsimulation models for representative assessment of safety and operational impacts of adverse weather using SHRP2 naturalistic driving data." *Journal of safety research* 81 (2022): 9-20.
- Rad, S. R., H. Farah, H. Taale, B. van Arem, and S. P. Hoogendoorn. "Design and operation of dedicated lanes for connected and automated vehicles on motorways: A conceptual framework and research agenda." *Transportation research part C: emerging technologies* 117 (2020): 102664.

Safety and Operational Impacts of Varying Market Penetration and Level of Autonomy of Connected and Automated Vehicles on I-80 Freeway Segment

Mandip Sigdel¹ and Mohamed M. Ahmed, Ph.D., P.E.²

¹Graduate Research Assistant, Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: sigdelmp@mail.uc.edu

²Professor and Director, Dept. of Civil and Architectural Engineering and Construction Management, College of Engineering and Applied Science Transportation Center, Univ. of Cincinnati, Cincinnati, OH. Email: ahmed2mh@ucmail.uc.edu

ABSTRACT

The ongoing progress in automation technologies holds the potential to mitigate human errors and create a safer transportation environment. Soon, the roads will see a mix of human driven vehicles alongside connected and automated vehicles (CAVs) with varying market penetration and level of autonomy (MPLA). As technology advances, it becomes clear that as market penetration of CAV increases, so does their level of autonomy. This study employs a VISSIM-based microsimulation model to mimic these mixed traffic scenarios and investigate its effect on safety and operational efficiency. Altogether 27 scenarios with varying traffic volumes and MPLA were explored, assessing safety using the time to collision indicator with three different thresholds. The findings showed that CAVs with mid to high autonomy and a market penetration of about 50% reduced conflicts by about 50% in medium to high traffic volumes while improving network efficiency. At higher MPLA scenarios, the system performed consistently well, with minimal variations in safety and operational impacts across different traffic volumes.

INTRODUCTION

Connected and Automated Vehicles (CAVs) technology holds promising potential in diminishing collisions and potential disputes on road networks by enhancing drivers' comprehension of their surroundings and eliminating human errors. This is achieved through the real-time delivery of environmental data to road users, thereby increasing their situational awareness (Amini et al. 2021, Viridi et al. 2019). A crucial aspect of this evolving approach to traffic safety lies in recognizing the importance of modifying driver behavior—a Human Factor pivotal in ensuring roadway safety, accountable for over 90% of all road crashes (Hosseinzadeh et al. 2021, Singh 2015). The Society of Automotive Engineers (SAE International 2021) has established a six-level classification for driving automation, ranging from level zero (providing no assistance) to level 5 (complete driver disengagement). Currently, technologies aligning with levels 1 and 2 are prevalent, while levels 3 to 5 are undergoing testing in various regions. Market research surveys have been employed to forecast the expected market penetration of different levels of CAVs according to the SAE classification. The initiation year, which marks the commercial introduction and public utilization commencement, for Level 4 and Level 5 CAV operations, as well as their initial market shares and subsequent growth patterns, are crucial factors that significantly influence infrastructure planning and operation related to CAVs. Thus, understanding the effect on traffic due to these mixed vehicles of different levels of autonomy is

equally crucial (Gkartzonikas and Gkritza 2019, Saeed et al. 2020). Many current studies in the CAV field solely focus on the one-dimensional aspect of CAV market penetration, while the level of autonomy is left unattended (Adomah et al. 2021, Liu and Fan 2020, Mousavi et al. 2012, Yu et al. 2019). It is certain that the future transportation landscape will consist of mixed traffic ranging from Human Driven Vehicles (HDVs) to highly automated CAVs and thus it is vital to understand the safety and operational impact we might face during these mixed traffic conditions. Therefore, this study adopts the concept of changing Market Penetration and Level of Autonomy (MPLA) of CAVs from Saeed et. al (2021) and seeks to contextualize both the aspects of Market Penetration and Level of Autonomy to address the safety and operational impacts due to mixed traffic conditions.

To delve deeper into the MPLA concept, it is crucial to identify a suitable corridor for analysis. Interstate-80 (I-80) in Wyoming emerges as an ideal choice for various reasons. Primarily, I-80 in Wyoming is already part of the USDOT FHWA Connected Vehicle Pilot Deployment Program, which focuses on implementing and testing CAV technologies. This program provides a strong foundation for future CAV integration in the region. Despite its low traffic, I-80 is distinguished by high truck volumes and adverse weather conditions, posing significant safety concerns (Bakhshi and Ahmed 2021, Gaweesh et al. 2021). Highways with low traffic volume, like certain sections of I-80 in Wyoming, exhibit substantial variations in driver behavior influenced by factors such as road geometry and large distances between vehicles at the desired speed (Eftekharzadeh and Khodabakhshi 2014). Fewer interactions among vehicles on low-volume highways may lead to decreased driver awareness of conditions and potential downstream hazards, contributing to multi-vehicle collisions and severe crashes (Nilsson 2004). In line with the research of (Das and Ahmed 2022), this paper builds upon their findings and employs their lane change parameters from the SHRP2 NDS database, along with parameters defined by the CoExist project (Sukennik 2020), for VISSIM microsimulation. The subsequent sections of this paper provide an overview of existing literature on CAV simulation for safety assessment, highlighting the novelty of this research. This is followed by a comprehensive explanation of the case studies, including CAV behavior in the simulation, sensitivity analysis and calibration process, and the definition of simulation scenarios. A thorough discussion of results follows, leading to the conclusion section, which summarizes the study's findings and draws overall conclusions.

BACKGROUND

Microscopic traffic simulation models have been extensively utilized for conducting impact analyses of CAVs and various simulation software platforms, such as VISSIM, AIMSUN, PARAMICS, and SUMO, have been utilized to analyze CAV behavior within broader traffic contexts, allowing researchers to explore their potential impacts on traffic flow, efficiency, and safety. (Morando et al. 2018) utilized VISSIM® software to simulate Autonomous Vehicles and employed the Surrogate Safety Assessment Model (SSAM) to evaluate the potential for crashes. The result of the study revealed a significant decrease in the average number of crashes within the simulated signalized intersection and four-legged roundabout as the penetration rate of AVs increased. (Hartmann et al. 2017) used VISSIM® to assess the effects of automated vehicles on freeway capacity, finding a decrease of up to 7% with partially and highly automated vehicles, but a substantial increase of up to 30% with a high penetration rate of CAVs. (Mousavi et al. 2019) investigated the impact of autonomous vehicles on the safety of unsignalized intersections,

revealing a potential decrease in conflicts near intersections across different levels of service. (Arvin et al. 2019) used VENTOS to model low and highly automated vehicles and introduced a new safety assessment indicator called “driving volatility” introducing the safety indicator “driving volatility”. The results indicated a significant reduction in accidents as the number of AVs increased. (Ye and Yamamoto 2018) found that dedicating lanes for CAVs improved traffic flow throughput when the Market Penetration Rate (MPR) exceeded 40%. (Deluka-Tibljaš et al. 2018) utilized VISSIM® to quantify safety effects of CAVs at roundabouts and found a slight increase in conflicts with the introduction of CAVs. Another study by (Rahman and Abdel-Aty 2018) focused on the impact of platooning autonomous vehicles on expressways using VISSIM® software. By utilizing Time to Collision (TTC) as one of the Surrogate Measures of Safety (SMoS), they found that forming platoons could improve safety by reducing the total time in which vehicles are running below the TTC threshold by 19% to 28%.

Previous research has examined the efficiency and safety of highway segments and large-scale networks in relation to Connected and Automated Vehicle (CAV) traffic. These studies have investigated the impact of various factors, such as truck ratios and exclusive lane policies, on CAV operations. While the majority of these studies have primarily focused on the changing market penetration rate of CAVs, there is limited research that considers both Market Penetration and Level of Autonomy simultaneously. This highlights a gap in the literature and emphasizes the need for further studies that explore the combined effects of MPLA on mixed traffic. This study aims to fill this gap by assessing the impact of varying CAV Market Penetration Rates and levels of autonomy on the safety and operation of an Interstate freeway (I-80 corridor) of Wyoming. The following sections will detail the methodology employed and the analysis of data.

METHODOLOGY AND SIMULATION FRAMEWORK

This study utilizes VISSIM® software to simulate the designated transportation network, and SSAM is employed to evaluate the impact of CAVs on the safety of the specific section under investigation. The selected freeway segment spans I-80 in Wyoming from milepost (MP) 357 to MP 359, covering a 2-mile eastbound stretch with two lanes, various on- and off-ramps, and significant merging and diverging traffic movements. This segment necessitates frequent lane changes, making it more suitable for the study than a Basic Freeway Segment (BFS) which has comparatively fewer lane change maneuvers. The layout of the selected freeway segment is illustrated in Figure 1.

To gather information about roadway geometries, posted speed limits, and other network characteristics, Google Earth Pro and Google Maps were utilized as additional data sources. Certain initial assumptions are established in this study to define the scope and context of the research objectives and to provide clarity regarding the specific focus and conditions under study.

Simulation Settings

Assumptions

As time progresses, the level of autonomy and market penetration of CAVs are expected to change. With increasing automation, CAVs are anticipated to adopt more assertive behaviors,

maintaining shorter following distances. Additionally, elevated communication levels are projected to facilitate cooperative lane changes, advanced merging, platoon formation, predictive collision avoidance, and cooperative speed harmonization. In this study, the potential factors of automation delay, communication lag, cybersecurity vulnerabilities, and hardware and software malfunctions are presumed to be absent due to their complexity in replication and the lack of definitive values available for their inclusion. Additionally, it is also assumed that vehicles designated to operate at the same level of automation exhibit consistent adoption of automation across their respective fleets. Furthermore, ideal operational conditions of clear weather and well-marked roads have been assumed. The study replicates CAV behavior using three levels: Low-Level CAVs, Mid-Level CAVs, and High-Level CAVs, which are simulated using parameters defined by the CoExist project (Sukennik 2020) utilizing the different aspects of Cautious, Normal and Aggressive Automation levels. Low-Level CAVs are automated but lack sophisticated communication, relying heavily on human drivers and thus the communication features have been turned off for them. Mid-Level CAVs are intended to simulate non-aggressive automation levels and connectivity while still involving some human control. These correspond to the automation levels, where the system assists the driver, but the driver remains partially responsible for vehicle control. High-Level CAVs on the other hand, represent higher automation capabilities with aggressive behavior, shorter following distances, and superior communication abilities, relying significantly less on human control. However, HDVs were simulated using the Real-time traffic data from the Wyoming Department of Transportation (WYDOT) traffic database, and lane change parameters and desired speed distribution for clear weather condition from the paper by (Das and Ahmed 2022). CAVs adhere to CoExist project's guidelines for speed distribution, aligning with posted speed limits on Interstate 80 (I-80). Figure 2 illustrates the desired speed distribution of HDVs and different CAV levels. In scenarios with HDVs, Low-Level CAVs follow their speed distribution; in their absence, they follow Mid-Level CAVs. A mixed traffic environment is created, encompassing both Non-CAV and CAV cars and trucks. Platooning was intentionally excluded in this study to avoid complexities arising from coordinated travel of multiple vehicles closely together which could introduce additional variables. Disabling platooning in the simulation allows the study to precisely examine how changing MPLA influences traffic behavior and performance.



Figure 1. Layout of Selected Study Segment (captured from VISSIM® background map)

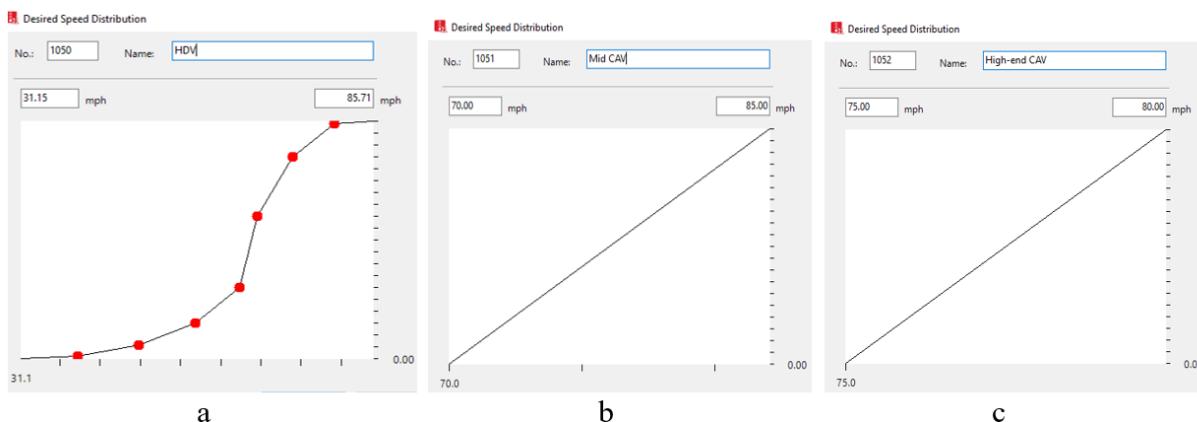


Figure 2. Screenshots of Desired Speed Distribution in VISSIM®. (a) HDV. (b) Mid-Level CAV. (c) High-Level CAV

Simulation Parameters

VISSIM® uses the default Wiedemann car-following model (Higgs et al. 2011), with two variations: Wiedemann 1974 for urban conditions and Wiedemann 1999 for freeways. This study employed the Wiedemann 99 model to simulate human drivers and Connected and Automated Vehicles (CAVs) due to its comprehensive behavioral capturing. The model comprises ten parameters (CC0 to CC9). The PTV VISSIM® User Manual provides detailed documentation on these parameters. The PTV Group and CoExist offer recommendations for configuring the internal model in VISSIM® specifically for CAVs, involving adjustments to car-following and lane-changing behavior parameters (Sukennik 2020). This study made the following modifications to the internal models in VISSIM®, as shown in Table 1.

Simulation Scenarios

Using VISSIM® background maps, the freeway segments' layout was created to mirror real-world conditions, verified for accuracy through cross-checking with Google Street View. A total of 27 microsimulation scenarios were developed, covering three traffic volumes (600, 800, and 1000 vehicles per hour (vph)) tested with nine scenarios each for varying MPLA conditions (Table 2). Simulations ran for 70 minutes, including a 10-minute warm-up period for stabilization. Each scenario underwent 10 simulation runs to account for seed variations. Data collected from each run assessed safety and operational aspects under diverse MPLA conditions across the network.

Calibration and Validation

The baseline model underwent calibration and validation before examining existing freeway models. VISSIM® parameters were fine-tuned for accurate representation of real-life driving scenarios. A custom cumulative distribution function established desired speeds for human-driven vehicles (HDVs) based on field speed data, and other vehicle characteristics were adjusted to reflect real-world conditions. Microsimulation models were validated by comparing simulated traffic volumes (averaged over 10 runs) with real-life volumes using the Geoffrey E. Havers

(GEH) statistic test (PTV Group 2018). GEH values were below 2.5 indicating a strong agreement between simulated and field volumes, confirming successful model calibration. Furthermore, travel times from the simulation model were compared with WYDOT field measurements, yielding a mean absolute percentage error (MAPE) value of 2.95%, meeting acceptable thresholds and validating the baseline model for the freeway segment based on travel times (Moreno et al. 2013).

TABLE 1. Car-following and lane-change parameters of HDVs and CAVs

Driving Behavior	Parameters	HDVs	Low-Level CAVs	Mid-Level CAVs	High-Level CAVs
Following Parameters	Look-ahead distance (ft)	0-820.21	0-820.21	0-820.21	0-984.25
	Look-back distance (ft)	0-492.13	0-492.13	0-492.13	0-492.13
	Number of interacting objects	2	2	2	10
Car following Parameters	CC0, Standstill distance (ft)	4.92	4.92	4.92	3.28
	CC1, Gap time distribution (s)	0.9	1.5	0.9	0.6
	CC2, 'Following' distance oscillation (ft)	13.12	0	0	0
	CC3, Threshold for entering 'following' (s)	-8	-10	-8	-6
	CC4, Negative speed difference (ft/s)	-0.35	-0.1	-0.1	-0.1
	CC5, Positive speed difference (ft/s)	0.35	0.1	0.1	0.1
	CC6, Distance dependency of oscillation	11.44	0	0	0
	CC7, Oscillation acceleration (ft/s ²)	0.82	0.33	0.33	0.33
	CC8, Acceleration from standstill (ft/s ²)	11.48	9.84	11.48	13.12
	CC9, Acceleration at 50 mph (ft/s ²)	4.92	3.94	4.92	6.56
lane change Parameters	Maximum deceleration, own vehicle (ft/s ²)	-11.38	-11.48	-13.12	-13.12
	Maximum deceleration, trailing vehicle (ft/s ²)	-1.78	-8.2	-9.84	-13.12
	-1 ft/s ² per distance, own vehicle, and trailing vehicle (ft)	200	80	100	100
	Accepted deceleration, own vehicle (ft/s ²)	-0.11	-3.28	-3.28	-3.28
	Accepted deceleration, trailing vehicle (ft/s ²)	-0.09	-3.28	-3.28	-4.92
	Waiting time before diffusion (s)	200	60	60	60
	Minimum Clearance, front/rear (ft)	104.41	1.64	1.64	1.64
	Safety distance reduction factor	0.6	1	0.6	0.75
	Maximum deceleration for cooperative braking (ft/s ²)	-1.7	-8.2	-9.84	-19.69
	Advanced Merging	No	Yes	Yes	Yes
Autonomous Driving Parameters	Cooperative lane change	No	No	Yes	Yes
	Maximum speed difference (mph)	-	-	6.71	6.71
	Maximum collision time (s)	-	-	10	10
	Enforce absolute braking distance	No	Yes	No	No
Autonomous Driving Parameters	Use implicit stochastics	Yes	No	No	No
	Platooning	No	No	No	No

Safety Performance Assessment

Time-To-Collision (TTC)

This study employed the Surrogate Safety Assessment Model (SSAM) to assess the safety of simulated scenarios by analyzing trajectory data for potential conflicts in the traffic. Time-To-

Collision (TTC) served as a surrogate measure of safety, evaluating crash risk across various scenarios in the study segment. TTC is defined as the time remaining until a potential collision if interacting road users maintain the same speed and direction (Hayward 1972). TTC can be expressed using the following Equation:

$$\text{TTC} = \begin{cases} \frac{D_{1-2}}{V_2 - V_1}, & \text{if } V_2 > V_1 \\ \infty, & \text{Otherwise} \end{cases}$$

where D_{1-2} represents the gap distance, V_1 and V_2 represent the speeds of the leading and following vehicles, respectively. Typically, TTC thresholds range from 1 to 5 seconds, with 1.5 seconds commonly used. In this study, three specific TTC thresholds were employed: 1.5 seconds for high-risk, 3 seconds for medium-risk, and 5 seconds for low-risk. These thresholds facilitated the classification and analysis of collision risk levels in simulated scenarios. The study aimed to assess safety implications under different traffic conditions and evaluate the effectiveness of implemented measures in reducing collision risk.

TABLE 2. Simulation Scenarios

SN	Scenario	Proportion of vehicles			
		HDV	L-CAV	M-CAV	H-CAV
1	All vehicles are human-driven	100	0	0	0
2	Low-Level CAVs are introduced alongside HDV	75	25	0	0
3	Mid-Level CAVs are introduced with HDV and Low-Level CAV has a higher percentage	25	50	25	0
4	Mid-Level CAV takes over in the presence of Low-Level CAV and some HDV	25	25	50	0
5	High-Level CAV is introduced, Low-Level CAVs speed follows HDVs speed	0	25	50	25
6	High-Level CAV is introduced, Low-Level CAVs speed follows Mid-Level CAVs speed	0	25	50	25
7	High-Level CAV alongside Mid-Level CAV	0	0	50	50
8	High-Level CAV takes over Mid-Level CAV	0	0	25	75
9	All vehicles are High-Level CAVs	0	0	0	100

L-CAV = Low-Level CAV, M-CAV = Mid-Level CAV and H-CAV = High-Level CAV.

RESULTS AND DISCUSSION

Traffic conflict Analysis

Table 3 summarizes conflicts from SSAM for different MPLA scenarios, considering various TTC thresholds. In Scenario 1 (baseline), featuring only HDVs, few conflicts occurred at low volumes, indicating good network performance. The number of conflicts were found to be increased for scenario 1 as the volume was increased. Scenario 2 (mixed traffic with low-level CAVs) showed improved safety at low volumes but worsened at higher volumes due to limited communication and HDV-like behavior of low-level CAVs. Scenarios 3 and 4 (Mid-Level CAVs, Low-Level CAVs, and HDVs) performed slightly better than the baseline for all TTC thresholds and volumes, suggesting improved safety with higher automation and market

penetration. Scenarios 5 and 6 explored Low-Level CAVs with different speed distributions, with Scenario 6 demonstrating better safety, emphasizing the benefits of uniform CAV speeds. Scenarios 4 and 5 showed about 50% conflict reduction with 50% market penetration of Mid to High-Level CAVs, especially at higher volumes. Scenarios 7, 8, and 9, comprising only Mid and High-Level CAVs, yielded the best results. In almost all volumes and TTC thresholds, it can be clearly seen that in Scenarios 7, 8, and 9, the conflicts were reduced by more than 85%. At higher TTC thresholds, Scenario 9 (High-Level CAVs only) showed slightly more conflicts than Scenario 8 (High-Level and Mid-Level CAVs). This may be due to the aggressive nature of High-Level CAVs, using shorter following distances and more assertive lane changes compared to the less aggressive Mid-Level CAVs. The safety impact is clearly nonlinear as we move through different scenarios. For low volume condition, the safety impacts are slightly noticed in the initial scenarios up to Scenario 5 where we can see that the conflicts are slightly decreasing. However, as we reach high automation stages comprising high market penetration, the conflicts are significantly decreased, and the effect of CAVs appears to be profound. This observation aligns with the complex interactions in transitions across levels of market penetration rate and levels of autonomy. It is imperative to recognize the dynamic nature of these transitions, as factors such as mixed traffic scenarios, varying communication capabilities among vehicles, and human driver behavior can influence safety and mobility outcomes in unpredictable ways. Acknowledging the nonlinear nature of these transitions underscores the importance of comprehensive analysis and modeling techniques to capture the dynamic interactions and complexities inherent in connected and automated vehicle environments. The MPLA's impact was more pronounced at higher volumes, showcasing the effectiveness of Mid and High-Level CAVs in reducing conflicts. In general, both the rear-end and lane change conflicts followed a similar trend with the lane change conflict always being less than or equal to the rear-end conflicts.

Operational Characteristics Analysis

Table 4 presents performance measures, including average speed, total travel time, and total delay, to assess the operational performance of the simulated network. Generally, a notable increase in average speed was seen with higher MPLA, except for Scenario 2 where it slightly decreased. This decrease may be attributed to Low-Level CAVs following HDVs' speed distribution and lacking advanced communication and lane change capabilities. In Scenario 5, which included only CAVs, the average speed exceeded the baseline by 12% across all traffic volumes. Furthermore, the Average Speed of the vehicles reached around 77 mph at later stages of the scenario irrespective of the volume. Similarly, Total travel time and Total Delay showed improvement with higher MPLA values, though slight degradation occurred in the initial stages. In line with the safety findings, the operational results also demonstrated nonlinearity, which can be attributed to the dynamic nature of these transitions. Factors such as mixed traffic scenarios, diverse communication capabilities among vehicles, and human driver behavior contribute to unpredictable influences on safety and mobility outcomes. The results highlight the efficiency of CAVs in enhancing freeway operational characteristics, particularly with high-level to medium-level automation and higher market penetration. Conversely, Low-Level CAVs with low market penetration exhibited limited effectiveness, underscoring the importance of communication technology in automated vehicle performance.

TABLE 3. Simulation results of conflicts for TTC threshold of 1.5, 3, and 5 s

Volume	Scenario	TTC = 1.5s			TTC = 3s			TTC = 5s		
		Total	rear end	lane change	Total	rear end	lane change	Total	rear end	lane change
Low (600 vph)	1	28	23	5	60	37	23	91	52	39
	2	25	23	2	51	31	20	87	50	37
	3	23	13	10	44	23	21	77	50	27
	4	25	17	8	46	28	18	83	54	29
	5	27	20	7	44	25	19	68	37	31
	6	7	4	3	16	9	7	14	8	6
	7	3	3	0	5	5	0	8	4	4
	8	4	4	0	3	3	0	9	6	3
	9	4	3	1	2	2	0	9	6	3
Medium (800 vph)	1	90	57	33	180	96	84	315	160	155
	2	110	67	43	170	103	67	286	164	122
	3	57	30	27	88	51	37	195	100	95
	4	56	32	24	92	49	43	179	95	84
	5	41	24	17	53	29	24	138	85	53
	6	17	8	9	36	21	15	47	32	15
	7	8	6	2	12	8	4	19	14	5
	8	5	4	1	6	4	2	11	8	3
	9	2	1	1	11	8	3	17	10	7
High (1000 vph)	1	197	115	82	398	215	183	793	448	345
	2	243	148	95	422	240	182	763	468	295
	3	146	94	52	221	135	86	447	275	172
	4	100	51	49	155	75	80	379	201	178
	5	96	60	36	153	88	65	352	217	135
	6	44	26	18	88	60	28	119	84	35
	7	15	12	3	19	15	4	26	19	7
	8	7	6	1	13	8	5	26	18	8
	9	5	4	1	19	12	7	33	23	10

1: HDV = 100%; 2: L-CAV = 25%, HDV = 75%; 3: M-CAV=25%, L-CAV= 50%, HDV=25%; 4: M-CAV= 50%, L-CAV=25%, HDV= 25%; 5: H-CAV= 25%, M-CAV= 50%, L-CAV= 50% (L-CAVs speed follows HDVs speed); 6: H-CAV= 25%, M-CAV= 50%, L-CAV= 50%(L-CAVs speed follows M-CAVs speed); 7: H-CAV= 50%, M-CAV=50%; 8: H-CAV= 75%, M-CAV= 25%; 9: H-CAV= 100%

TABLE 4. Simulation Results of Operational Characteristics

Volume	Scenario	Average Speed (mph)	Total Travel Time (h)	Total Delay (h)
Low (600 vph)	1	63.15	0.629	0.014
	2	62.89	0.632	0.017
	3	65.22	0.609	0.019
	4	68.59	0.577	0.015
	5	72.44	0.546	0.009
	6	76.62	0.517	0.005
	7	76.96	0.515	0.003
	8	77.12	0.514	0.002
	9	77.18	0.513	0.002
Medium (800 vph)	1	62.25	0.844	0.031
	2	61.91	0.849	0.035
	3	64.31	0.817	0.037
	4	67.74	0.775	0.028
	5	71.91	0.729	0.017

6	76.36	0.687	0.008
7	76.82	0.683	0.004
8	77.01	0.681	0.003
9	77.11	0.680	0.003
High (1000 vph)			
1	61.16	1.075	0.058
2	60.77	1.082	0.065
3	63.18	1.039	0.065
4	66.92	0.980	0.048
5	71.11	0.921	0.031
6	75.82	0.864	0.016
7	76.63	0.855	0.007
8	76.87	0.852	0.005
9	77.01	0.851	0.004

1: HDV = 100%; 2: L-CAV = 25%, HDV = 75%; 3: M-CAV=25%, L-CAV= 50%, HDV=25%; 4: M-CAV= 50%, L-CAV=25%, HDV= 25%; 5: H-CAV= 25%, M-CAV= 50%, L-CAV= 50% (L-CAVs speed follows HDVs speed); 6: H-CAV= 25%, M-CAV= 50%, L-CAV= 50%(L-CAVs speed follows M-CAVs speed); 7: H-CAV= 50%, M-CAV=50%; 8: H-CAV= 75%, M-CAV= 25%; 9: H-CAV= 100%

CONCLUSIONS

It is highly likely that in the near future, Connected and Automated Vehicles (CAVs) will be introduced to the market and share the roads with conventional vehicles (HDVs). Initially, the proportion of CAVs on the roads will be low, and their level of autonomy will be limited. However, over time, the number of CAVs and their level of autonomy is expected to increase. This study aimed to investigate the effects of this mixed traffic environment, considering different levels of CAV market penetration and autonomy, alongside HDVs.

The main focus of this study was to evaluate the safety and operational impacts of varying CAV market penetration and level of autonomy (MPLA) on freeways. By analyzing the simulation data, the study aimed to gain insights into how these factors affect traffic dynamics and identify any potential challenges or benefits associated with the introduction of CAVs. Operational performance in this study was assessed by analyzing parameters such as average speed, total travel time, and total delay. Safety analysis, on the other hand, utilized Time-to-Collision (TTC) as a surrogate measure to evaluate the number of conflicts. Results show that, in the long term, as the market penetration and level of autonomy of CAVs increase, they are expected to have positive effects on the highway's traffic system, including increased average speed, decreased travel time, reduced delay, and fewer conflicts. However, during the transitional phase of mixed operation of manual vehicles and low-end CAVs, there may be negative effects to consider. The findings underscore that achieving significant benefits, such as a notable 50% reduction in conflicts through CAV deployment, necessitates a 50% market penetration of Mid to High-Level CAVs, particularly at higher traffic volumes. Moreover, in scenarios 7, 8, and 9, which exclusively feature Mid and High-Level CAVs, conflicts were mitigated by more than 85% across varying traffic volumes. It's imperative to consider the interplay between the level of autonomy and market penetration, as analyzing these factors in isolation may yield misleading results. This highlights the importance of contextualizing research within the dynamic relationship between both parameters. Failure to do so raises significant concerns about the validity of the findings and underscores the need for future research to provide comprehensive context explaining the concurrent changes in both variables.

The findings of this study have implications for traffic engineers, car manufacturers, and stakeholders, as they provide insights into how different levels of connectivity and autonomy in vehicles can affect both operational and safety characteristics. This knowledge can be valuable in improving the manufacturing processes of connected and autonomous vehicles and enhancing

traffic management strategies. It can aid in elucidating the optimal balance between autonomy levels and market penetration that leads to beneficial outcomes, while also identifying configurations that are less favorable. It is important to note that this study specifically focuses on the impact of Connected and Automated Vehicles (CAVs) on a rural freeway segment and does not consider other roadway sections or the concept of platooning. Future research should aim to investigate more complex road sections, such as urban freeways and other roadway facility types, while also considering the integration of platooning into the traffic system.

ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials (AASHTO) and the Wyoming Department of Transportation (WYDOT). The writings of this work were enhanced by using ChatGPT which helped in checking for grammatical errors as well as helping the authors in paraphrasing their sentences in a more concise manner.

AUTHOR CONTRIBUTIONS

The authors confirm the contribution to the paper as follows: study conception and design: Mandip Sigdel and Mohamed M. Ahmed; data collection: Mandip Sigdel and Mohamed M. Ahmed; analysis and interpretation of results: Mandip Sigdel and Mohamed M. Ahmed; draft manuscript preparation: Mandip Sigdel and Mohamed M. Ahmed. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Adomah, E., A. K. Bakhshi, and M. M. Ahmed. 2021. "Safety Impact of Connected Vehicles on Driver Behavior in Rural Work Zones under Foggy Weather Conditions." *Transportation Research Record*, 2676 (3): 88–107.
- Amini, E., A. Omidvar, and L. Elefteriadou. 2021. "Optimizing operations at freeway weaves with connected and automated vehicles." *Transportation Research Part C: Emerging Technologies*, 126: 103072.
- Arvin, R., A. Khattak, and J. R. Torres. 2019. "Evaluating Safety with Automated Vehicles at Signalized Intersections: Application of Adaptive Cruise Control in Mixed Traffic."
- Bakhshi, A. K., and M. M. Ahmed. 2021. "Practical advantage of crossed random intercepts under Bayesian hierarchical modeling to tackle unobserved heterogeneity in clustering critical versus non-critical crashes." *Accident Analysis & Prevention*, 149: 105855.
- Das, A., and M. M. Ahmed. 2022. "Weather-Based Lane-Change microsimulation parameters for safety and operational performance evaluation of weaving and basic freeway segments." *Transportation Research Record*, 2676 (12): 550–563.
- Deluka-Tibljaš, A., T. Giuffrè, S. Šurdonja, and S. Trubia. 2018. "Introduction of Autonomous Vehicles: Roundabouts design and Safety Performance Evaluation." *Sustainability*, 10 (4): 1060.
- Eftekharpazadeh, S. F., and A. Khodabakhshi. 2014. "Safety evaluation of highway geometric design criteria in horizontal curves at downgrades." *International Journal of Civil Engineering*, 12 (3): 326–332.
- Gaweesh, S., A. K. Bakhshi, and M. M. Ahmed. 2021. "Safety Performance Assessment of connected vehicles in mitigating the risk of secondary crashes: a Driving Simulator study." *Transportation Research Record*, 2675 (12): 117–129.

- Gkartzonikas, C., and K. Gkritza. 2019. "What have we learned? A review of stated preference and choice studies on autonomous vehicles." *Transportation Research Part C: Emerging Technologies*, 98: 323–337.
- Hartmann, M., S. Krause, S. Hoffmann, N. Motamedidehkordi, P. Vortisch, and F. Busch. 2017. *Impact of Automated Vehicles on Capacity of the German Freeway Network*.
- Hayward, J. C. 1972. *Near-miss determination through use of a scale of danger*. Highway Research Record.
- Higgs, B., M. Abbas, and A. Medina. 2011. "Analysis of the Wiedemann Car–Following Model over Different Speeds Using Naturalistic Data." *3rd International Conference on Road Safety and SimulationPurdue UniversityTransportation Research Board*.
- Hosseinzadeh, A., A. Moeinaddini, and A. Ghasemzadeh. 2021. "Investigating factors affecting severity of large truck-involved crashes: Comparison of the SVM and random parameter logit model." *Journal of Safety Research*, 77: 151–160.
- Liu, P., and W. Fan. 2020. "Exploring the impact of connected and autonomous vehicles on freeway capacity using a revised Intelligent Driver Model." *Transportation Planning and Technology*, 43 (3): 279–292.
- Morando, M. M., Q. Tian, L. T. Truong, and H. L. Vu. 2018. "Studying the safety impact of autonomous vehicles using Simulation-Based Surrogate Safety Measures." *Journal of Advanced Transportation*, 2018: 1–11.
- Moreno, J. J. M., A. L. P. Pol, A. Sesé, and B. C. Blasco. 2013. "Using the R-MAPE index as a resistant measure of forecast accuracy." *PubMed*, 25 (4): 500–6.
- Mousavi, S. M., O. A. Osman, and D. Lord. 2019. "Impact of Urban Arterial Traffic LOS on the Vehicle Density of Different Lanes of the Arterial in Proximity of an Unsignalized Intersection for Autonomous Vehicle vs. Conventional Vehicle Environments." *International Conference on Transportation and Development*.
- Mousavi, S. M., O. A. Osman, D. Lord, K. Dixon, and B. Dadashova. 2021. "Investigating the safety and operational benefits of mixed traffic environments with different automated vehicle market penetration rates in the proximity of a driveway on an urban arterial." *Accident Analysis & Prevention*, 152: 105982.
- Nilsson, G. 2004. "Traffic safety dimensions and the power model to describe the effect of speed on safety."
- PTV Group. 2018. *PTV VISSIM 11 User Manual*. Planung Transport Verkehr AG, Karlsruhe, Germany, pp. 1–1219.
- Rahman, S., and M. Abdel-Aty. 2018. "Longitudinal safety evaluation of connected vehicles' platooning on expressways." *Accident Analysis & Prevention*, 117: 381–391.
- Saeed, T. U., B. N. T. Alabi, and S. Labi. 2021. "Preparing road infrastructure to accommodate connected and automated vehicles: System-Level Perspective." *Journal of Infrastructure Systems*, 27 (1).
- Saeed, T. U., M. Burris, S. Labi, and K. C. Sinha. 2020. "An empirical discourse on forecasting the use of autonomous vehicles using consumers' preferences." *Technological Forecasting and Social Change*, 158: 120130.
- SAE International. 2021. "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles." https://www.sae.org/standards/content/j3016_202104/. Accessed Jun. 14, 2023.
- Singh, S. 2015. "Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey." *Traffic Safety Facts - Crash Stats*.

- Sukennik, P. 2020. Micro-Simulation Guide for Automated Vehicles-Final. <https://www.h2020-coexist.eu/wp-content/uploads/2020/04/D2.11-Guide-for-the-simulation-of-AVs-with-microscopic-modelling-tool-Final.pdf>. Accessed Jun. 19, 2023.
- Virdi, N., H. Grzybowska, S. T. Waller, and V. Dixit. 2019. "A safety assessment of mixed fleets with Connected and Autonomous Vehicles using the Surrogate Safety Assessment Module." *Accident Analysis & Prevention*, 131: 95–111.
- Ye, L., and T. Yamamoto. 2018. "Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput." *Physica A: Statistical Mechanics and Its Applications*, 512: 588–597.
- Yu, H., S. Tak, M. Park, and H. Yeo. 2019. "Impact of Autonomous-Vehicle-Only lanes in mixed traffic conditions." *Transportation Research Record*, 2673 (9): 430–439.

Safety Assessment of Cooperative Adaptive Cruise Control Truck Platooning on Freeway Work Zones

Vamsi Krishna Maddineni¹ and Mohamed M. Ahmed, Ph.D., P.E.²

¹Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: maddinva@mail.uc.edu

²Dept. of Civil and Architectural Engineering and Construction Management, Univ. of Cincinnati, Cincinnati, OH. Email: mohamed.ahmed@uc.edu

ABSTRACT

Freeway work zones, particularly in merging areas, experience serious traffic safety and operational difficulties. The advent of connected and autonomous vehicles (CAVs), capable of complex maneuvers like platooning, adds new layers of complexity. Despite numerous studies exploring operational aspects of heavy vehicle platooning in work zones, safety evaluations remain scant. This research, using VISSIM microsimulation, addresses this gap by examining the safety implications of cooperative adaptive cruise control (CACC) truck platooning in such areas. The study simulated varying scenarios by positioning lane change warning signs at distinct distances upstream of work zones. Findings reveal that early merge strategies—signs at least 1 mi upstream—decreased traffic conflicts by 300% compared to late merge conditions. This study also emphasizes that while platooning provides significant benefits in controlled traffic flow situations, infrastructure owners and operators (IOOs) should work closely with other stakeholders such as trucking companies, and law enforcement agencies to make necessary changes to their infrastructure to accommodate truck platooning.

KEYWORDS: Work Zones, Truck Platooning, Cooperative Adaptive Cruise Control, Lane-change, SHRP2 Naturalistic Driving Study, VISSIM, Freeway weaving, Early and Late Merge, Platoon Configuration

INTRODUCTION

In recent years, work zones have become an integral and common feature of the United States' highway infrastructure. Work zones often necessitate the closure of certain travel lanes and the reduction of speed limits, leading to diminished roadway capacity. This typically results in congestion and delays, causing inconvenience for road users (Mitigating Work Zone Safety and Mobility Challenges through Intelligent Transportation Systems Case Studies 2014). Approximately 10% of road traffic delays are caused by roadway construction (FHWA Work Zone Facts and Statistics - FHWA Office of Operations 2023). Compounding the issue, a substantial increase in the volume of heavy vehicles utilizing the freeways exacerbates the situation further, leading to an escalation in traffic delays (Zhou et al. 2019). Work zone fatalities increased by 10.8% between 2020 and 2021, while overall roadway fatalities increased by 10.3%. In work zones, fatal crashes involving rear-end collisions and Commercial Motor Vehicles (CMVs) have notably risen. From 2020 to 2021, fatal crashes involving CMVs escalated from 210 to 291, a stark 39% increase, compared to a minor 2% rise in non-CMV related fatalities (FHWA Work Zone Facts and Statistics - FHWA Office of Operations 2023).

In recent years, the development of Connected and Automated Vehicles (CAVs) has advanced significantly. These sophisticated systems not only facilitate autonomous driving, but also seamlessly integrate with other elements of the transportation ecosystem, including other vehicles and road infrastructure (Guanetti et al. 2018). Numerous studies indicate that Connected and Automated Vehicles (CAVs) can significantly enhance road safety, as they hold the potential to eliminate the human errors involved in roughly 94% of crashes. However, full automation remains a distant goal. The journey towards complete automation will necessitate an extended period of mixed traffic, with automated and non-automated vehicles cohabitating on the roadways. CAVs also possess the capacity to form platoons—groups of vehicles where behavior is dictated by the lead vehicle. This platooning ability yields several advantages, such as allowing subsequent vehicles to maintain shorter time headways than non-CAVs. This, in turn, results in increased capacity, reduced delay, and fuel savings (Haque et al. 2023) and environmental benefits (Wadud et al. 2016).

The platooning of heavy trucks equipped with CAV technologies has gained momentum over the past few years (Kang et al. 2019; Haque et al. 2021; Haque et al. 2023). Specifically, the concept of Cooperative Adaptive Cruise Control (CACC) has been extensively studied in real-time scenarios. Building upon the foundation of Adaptive Cruise Control (ACC), CACC introduces a cooperative approach through vehicle-to-vehicle (V2V) communication. Prominent projects, such as the California PATH project, have dedicated substantial efforts to exploring the potential of CACC technology, particularly in the context of trucks. With a primary focus on level 1 automation, these initiatives have aimed to demonstrate the feasibility and benefits of CACC implementation in real-world scenarios. There has been a great deal of attention and funding devoted to Connected and Autonomous Heavy Vehicle deployment with an emphasis on platooning by the Federal Highway Administration (FHWA), numerous state transportation departments, and local transportation agencies (Hallmark 2019; Mchale, n.d.)

Most existing studies concentrate on the operational impacts and benefit-cost analysis of heavy truck platooning in freeway work zones. One area of particular concern is the execution of lane changes by CACC trucks when encountering work zones. Such lane changes within a platoon could pose safety risks and disrupt the close following distances that characterize CACC operations. It's critical to comprehend how CACC trucks safely navigate lane changes in work zones to ensure their effective integration and deployment.

To the best of authors' knowledge, very few studies have examined the impact of heavy truck platooning in work zones, with a majority focusing on operational characteristics. A gap exists in the literature regarding safety evaluations of truck platoons in work zones. Consequently, this paper seeks to assess the safety impacts of Connected and Automated Heavy Vehicles (CAHV) platooning on highway work zones using microsimulation tools, such as VISSIM, and Surrogate Measures of Safety (SMoS).

Using SMoS for CAV safety assessment

In the past decade, CAV safety assessment has attracted substantial attention. The latent causal relationship between traffic conflicts and crashes has made SMoS derived from traffic conflicts a popular option among a variety of modeling methodologies (Dai et al. 2023). There are three types of SMoS: time-based, deceleration-based, and energy-based. In particular, energy-based SSMs are used to determine the severity of a potential crash. In comparison with the other two types of SSM, they require assumptions regarding the characteristics of the

vehicles involved (e.g., mass and kinetic energy). Thus, most of the CAV safety studies conducted so far have employed a time-based or deceleration-based SSM methodology (Wang et al. 2021). Deceleration-based SMoSs assess how vehicle deceleration (or acceleration) can prevent a crash from occurring while time-based SMoSs, assess the risk of an interaction between two vehicles based on their proximity to a collision.

This study is primarily focused on two major objectives. Firstly, it investigates the influence of level 1 and level 2 Cooperative Adaptive Cruise Control (CACC) trucks on the safety and operational efficiency of work zone merge areas. In particular, the research delves into the impact of early merge scenarios in freeway weaving segments that involve lane restrictions. The goal is to understand how trucks with CACC, operating at different levels of automation, affect safety and operational parameters in these specific areas. Secondly, the study conducts a safety evaluation of work zones, considering the various lane change configurations adopted by truck platoons on the upstream of work zones.

METHODOLOGY

The first step of this study was to implement truck platoons in early merge situation with one mile of Advance warning area provided from transition zone. To achieve this, certain assumptions have been made regarding the nature of platoon configuration. These include the idea that a vehicle can join a platoon from either the front or the rear when it is within a suitable range to catch up, and that vehicles do not make lane changes to join a platoon. It is also assumed that the platoon formation operates under a flexible configuration system, wherein the platoon leader changes if a vehicle joins from the front, with the new vehicle becoming the leader. Lastly, the study assumes that when a lane change is performed, the platoon dissolves, allowing each vehicle to make individual lane changes.

Driving Behavior

According to the Society of Automotive Engineers (SAE), there are six levels of automation that are possible for vehicles, while heavy vehicle platoon systems have a total of four phases with final phase being driverless platoon system. While both phase 1 and 2 of platoon systems require lead driver to at least perform lateral operations. This study focused on mainly two types of truck platoons systems. For the low automation level (SAE level 1&2 or CACC Level-1), the trucks are assumed to have only longitudinal control and lane changes are still performed by human drivers, while for high automation level (SAE level 4 or CACC Level-2), both lateral and longitudinal operations are automated. In this study, to investigate the effects of platoons on work zones, all trucks are assumed to have been equipped with CACC level 1 and CACC level 2 technology and other conventional vehicles were driven by human drivers. The lane change parameters for low automation scenarios were updated based on SHPR 2 values extracted by the study (Das and Ahmed 2022) and for CACC level 2 they were based on parameters defined coexist for cautious mode (Micro-Simulation Guide for Automated Vehicles 19AD).

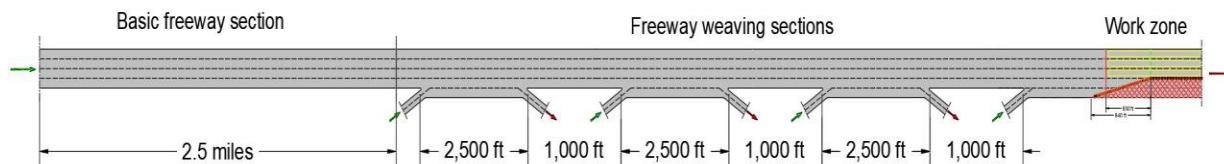
In the Level 1 CACC scenario, it is assumed that all trucks are equipped with CACC technology, maintaining a time gap of 1.5 seconds between each other. However, upon encountering other trucks, they possess the capability to form platoons, given certain conditions are met. Within these platoons, the gap time is reduced to 0.6 seconds.

Table 1: Lane change Parameters for different driving behaviors

Parameters	CACC Level-1	CACC Level-2
Own vehicle		
Maximum deceleration(ft/s^2)	-11.38	-11.48
Accepted deceleration(ft/s^2)	-0.11	-3.28
Trailing vehicle		
Maximum deceleration(ft/s^2)	-1.78	-8.2
Accepted deceleration(ft/s^2)	-0.09	-0.33
Other Parameters		
Minimum clearance(ft)	104.41	1.64
Cooperative Lane change	Yes	Yes
Advanced merging	Yes	Yes

VISSIM Model Preparation

The VISSIM traffic simulation tool was utilized to create a freeway segment that consists of a combination of a basic freeway and weaving segments (Figure 1). The vehicle inputs were programmed at the beginning of the basic freeway segment, with vehicles traveling from the basic freeway segment to the freeway weaving section. At the end of the freeway weaving section (Figure 3), a work zone section was introduced. The entire analysis segment had a total length of 4.5 miles. Within the configured segment, the rightmost lane was designated as the priority lane for trucks, while the middle lane served as the priority lane for other conventional vehicles. Given that the truck composition accounted for 40% of the total vehicles, the rightmost two lanes were predominantly occupied by trucks.

**Figure 1: Basic Freeway and Weaving Segments Modeled in VISSIM**

The modelled roadway section has a basic freeway segment of length (2.5 miles) followed by multiple freeways weaving sections and work zone is added at the end of freeway weaving segment (Figure 2). The weaving segments were mainly included to capture the effects of conflicting movements that occur between necessary lane changes performed by vehicles due to work zone condition and lane changes by vehicles that exit the freeway, and the distances between each interchange are maintained at 1000 ft and length of interchange is 2500 ft as per guidelines provided by the Highway Capacity Manual (HCM). The freeway weaving section compromises of four entry ramps and three exit ramps. The work zone segment is modelled at the end of the fourth entry point.

Given that the upstream area of a work zone necessitates a substantial number of lane changes, a situation that is particularly hazardous when truck platoons are involved, this study focuses specifically on the impact of such platoons on work zone merge areas. Therefore, the

work zone section is designed with only merge section and other parts of work zone were omitted as they have less significant impact on safety compared to merge section on the upstream of Work zone area.

Work zone development

In the work zone area, the total number of lanes was reduced from five to three using a tapered section. To ensure safe utilization of the tapered section, an emergency stop position was established 600 feet upstream of the lane closure connector. To denote the reduced speed areas within the work zone, corresponding speed limit reductions were implemented in the VISSIM simulation. The speed limit within these areas was decreased by 15 mph compared to the speed limit in the freeway weaving section. It is important to note that the work zone section and Lane change distances were updated in VISSIM based on this study (Jameis et al., n.d.). In the simulation, the lane change decision for CACC trucks was activated one mile upstream from the start of the transition zone. This allowed platoons to disperse and merge early within the work zones. The activation of lane change decisions for platoons at this specific distance was facilitated by the provision of work zone information and vehicle-to-infrastructure communication.

Since the study is interested in rural interstates, specific parameters such as speed distributions and vehicle composition are coded based on field observations and conditions on Interstate 80 (I80) in Wyoming. The study acknowledges that truck percentages can reach as high as 50% on I-80 in Wyoming, and therefore, a truck percentage of 40% is assumed for all scenarios in the simulations. Traffic volume was determined based on preliminary runs and density calculations, with specific volumes set to ensure they did not surpass a Level of Service B (LOS B) within the freeway weaving section, as defined by the Highway Capacity Manual (HCM). The traffic input assumptions are based on the premise that rural I-80, known for adverse weather conditions and high truck composition, does not experience extremely high traffic volumes. It's worth noting that a higher percentage of trucks is a characteristic feature of this route. Three different volume levels: 1700 vph, 2100 vph, and 2500 vph were selected after preliminary analysis.

DATA ANALYSIS AND RESULTS

Exploring the Influence of Traffic Volume, Platoon Length, and Driving Behavior on Safety in Work Zones

This study's primary objective is to scrutinize the influence of Cooperative Adaptive Cruise Control (CACC) trucks on traffic safety. As previously established, significant safety enhancements are only attained at higher Market Penetration Rates (MPR) of trucks. This section, however, explores the effects of other parameters—volume, maximum platoon size, and driving behavior—on work zone safety, with a constant truck MPR of 100%.

(Table 2) offers a detailed review of the conflict numbers under varied volumes, platoon sizes, and driving behaviors. Each scenario is evaluated against the base case without platooning, determining statistical significance through t-tests with a critical value derived from a significance value (p-value) of 0.05. This t-test analysis was performed using SSAM software.

Analyzing (Table 2), several patterns emerge correlating volume and conflicts within work zones. As the volume escalates, so does the number of conflicts in all scenarios. However, the

benefits of CACC-based platooning become increasingly prominent with volume escalation. In scenarios of low volume and automation (CACC1), conflict numbers rise by 20% as the maximum platoon size extends to seven. Yet, in high-volume situations, this trend reverses dramatically, showing a 50% decrease in conflicts when the maximum platoon size reaches seven. This suggests that in high-volume conditions, platooning significantly improves safety by effectively mitigating conflicts.

Table 2: Conflict frequency and risk comparison for various platoon configurations

CACC Level 1 Automation											
Volume (vph)	Max. Platoon size	Avg no. of conflicts (1 hr.)	Relative risk	t-statistic	Rear end Conflicts	Relative risk	t-statistic	Lane change Conflicts	Relative risk	t-statistic	t-critical
1700	zero	129	1	-	71.25	1	-	48	1	-	-
	two	128	1.26	2.41	97.6	1.37	3.21	53	1.1	0.92	1.89
	five	139.4	1.17	3.84	95.4	1.34	4.78	44	0.92	-1.82	1.89
	seven	146	1.22	3.74	97.6	1.37	5.38	47.4	0.99	-0.14	1.89
2100	zero	256.6	1	-	167.6	1	-	89	1	-	-
	two	282.2	1.1	1.79	185.2	1.11	1.77	97	1.09	1.5	1.86
	five	248.2	0.97	-0.62	165.4	0.99	-0.2	82.6	0.93	-1.41	1.86
	seven	254.6	0.99	-0.13	169	1.01	0.13	85.6	0.96	-0.62	1.86
2500	zero	542	1		542.25	1		181.75	1		-
	two	506	0.93	-0.56	494.5	0.91	-0.6	176	0.97	-0.32	1.94
	five	383	0.66	-2.74	339.5	0.63	-2.81	141.5	0.78	-2.2	1.94
	seven	339	0.54	-3.87	309	0.57	-3.27	84.75	0.47	-6.27	1.94
CACC Level 2 Automation											
1700	zero	92.75	1	-	55.5	1	-	37.25	1	-	-
	two	91.25	0.98	-0.34	56.75	1.02	0.25	34.5	0.93	-1.07	1.94
	five	86.25	0.93	-1.24	59.5	1.07	0.78	26.75	0.72	-3.26	1.94
	seven	82.25	0.89	-1.77	54	0.97	-0.23	28.25	0.76	-3.31	1.94
2100	zero	180	1	-	109.5	1	-	70.5	1	-	-
	two	150.25	0.83	-2.37	98.75	0.9	-0.83	50.75	0.72	-8.13	1.94
	five	146.75	0.82	-2.62	101.75	0.93	-0.64	45	0.64	-14.06	1.94
	seven	131.5	0.73	-4.33	90	0.82	-1.78	41.5	0.59	-9.17	1.94
2500	zero	375.25	1	-	251.5	1	-	123.75	1	-	-
	two	301.25	0.8	-3.35	214.5	0.85	-2.33	86.5	0.7	-4.15	1.94
	five	226.25	0.6	-6.31	152.75	0.61	-6.42	73.5	0.59	-4.83	1.94
	seven	254.5	0.68	-6.03	187	0.74	-4.74	67.5	0.55	-5.96	1.94

Conversely, in scenarios with Level 2 CACC automation where platooning is permitted, the conflict numbers consistently remained below those of the base scenario. The safety benefits of platooning in this case grow progressively more pronounced with increasing volume, leading to a larger reduction in conflicts relative to the base scenario.

An important observation from the (Table 3) previously mentioned is the rising mean Time-to-Collision (TTC) as platoon length extends, while Delta V (change in velocity) recedes. This

pattern indicates a decrease in conflict severity with platooning implementation, becoming more substantial at higher traffic volumes.

These insights underscore the safety benefits related to platooning, as it aids in lowering conflict severity. Furthermore, the results suggest these safety advantages are amplified as traffic volumes rise, indicating platooning's efficacy in managing conflicts in high-volume situations.

Table 3: Comparison of Mean TTC and Delta V parameters for various scenarios

CACC level 1 Automation						
Volume (vph)	Max. Platoon size	Mean TTC(s)	t-statistic	Delta V (ft/s)	t-statistic	t-critical
1700	zero	0.42	-	5.93	-	-
	two	0.46	1.21	6.1	0.49	1.66
	five	0.49	2.13	4.64	-3.78	1.66
	seven	0.48	1.72	4.28	-4.97	1.66
2100	zero	0.41	-	5.64	-	-
	two	0.45	1.67	5.27	-1.7	1.66
	five	0.49	3.4	4.58	-4.86	1.66
	seven	0.46	1.67	5.48	-0.51	1.66
2500	zero	0.62	-	3.89	-	-
	two	0.63	0.59	4.23	2.88	1.66
	five	0.56	-3.16	4.52	4.51	1.66
	seven	0.71	4.78	2.85	-8.89	1.66
CACC level 2 Automation						
1700	zero	0.44	-	6.72	-	-
	two	0.62	3.82	5.22	-3.79	1.66
	five	0.64	4.33	3.99	-7.4	1.66
	seven	0.64	4.23	3.92	-7.31	1.66
2100	zero	0.48	-	5.32	-	-
	two	0.59	3.01	4.61	-2.63	1.66
	five	0.64	4.53	4.05	-4.89	1.66
	seven	0.63	4.06	4.2	-4.12	1.66
2500	zero	0.5	-	5.09	-	-
	two	0.63	5.08	4.21	-4.93	1.66
	five	0.6	3.79	4.22	-4.49	1.66
	seven	0.72	8.73	3.45	-9.63	1.66

Distribution of speeds along the Upstream of Work zone

To comprehend the variations in speed across the upstream work zone, a heatmap is generated for each individual scenario, capturing data from every timestamp. This visual representation illustrates speed changes along the length of different zones observed upstream of the work zone.

Heatmap analysis identifies two prominent regions where speed significantly reduces. The first region, approximately 2000 feet from the beginning of Zone 1—an interchange—shows a speed drop. This decrease is ascribed to the concurrent lane changes involving CACC trucks shifting left and other vehicles exiting the freeway that necessitate right lane changes. These simultaneous lane alterations result in speed reductions as vehicles adjust to these transitions. The second region with speed reduction manifests in the transition zone, where the areas of

reduced speed begin. The presence of these speed deceleration zones consequently leads to slower speeds.

The speed distribution plots provide clear evidence that platooning has a significant impact on reducing speeds. This trend is particularly observed in the case of CACC level 1 automation, where operational speeds are decreased but safety is improved. On the other hand, in the case of CACC level 2 automation, the speed distribution within platoons is almost similar to the base case without platooning. However, safety advantages are still observed without compromising the operational speed on the freeway. These findings indicate that as the automation level increases, both safety and operational performance can be improved.

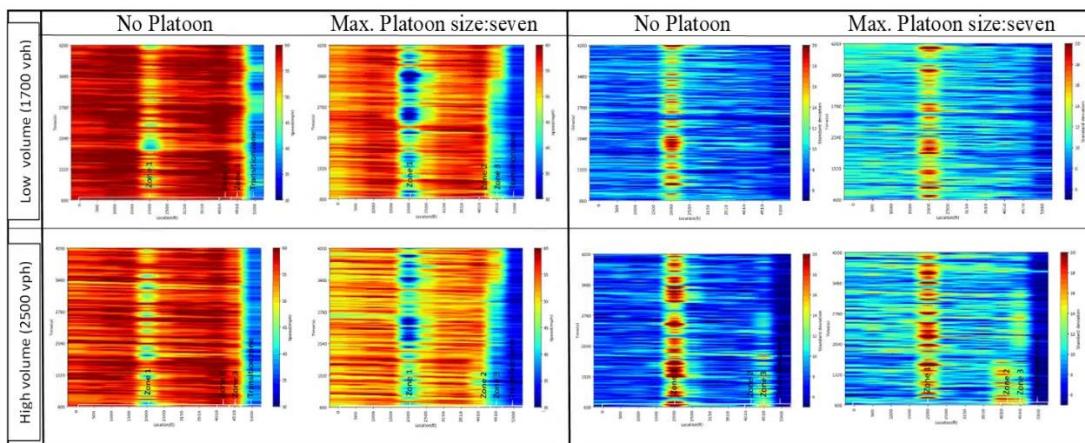


Figure 2: Spatiotemporal distribution of speeds and standard deviation for CACC Level 1 Automation

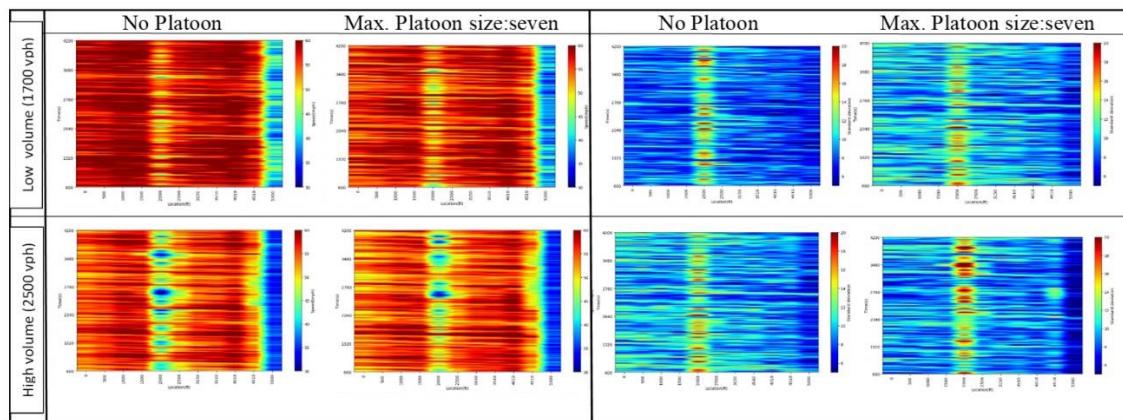


Figure 3: Spatiotemporal distribution of speeds and standard deviation for CACC Level 2 Automation

Effect of early merge and late merging of platoons in work zones

This investigation aims to understand how platoon-required lane changes affect the upstream area of a work zone. Simulations were conducted using various lane merge initiation distances: 2 miles, 1 mile, 0.5 mile, and immediately before the transition zone (0.2 miles from the lane

closure connector). All trucks participating in the study were automated with CACC Level 1, implementing advanced merge and cooperative lane change behaviors during the lane change process.

It was observed that the Traffic safety was critical when platoons merge near the transition zone. The no. of conflicts increased by 300% - 400 % (Figure 4) as compared to the case when there were no platoons. On the other hand, the longer truck platoons improved safety along the freeway weaving section during early merge situations. It was observed that no. of conflicts decreased by 40% at high traffic volumes with respect to base case scenario without platoons. The results reveal that the location of advanced warning signs up to one mile upstream does not significantly impact safety outcomes. However, after the lane change signs were placed at distances of 0.5 miles for platoons and, more significantly, at 0.2 miles just before the transition zone, the total number of conflicts increased noticeably. This is due to huge increase in number of Rear end conflicts as lane changes happen closer to work zone area and this led to shockwave development just before the transition zone. It suggests that while platooning can be advantageous in controlled flow conditions, it poses a considerable risk to traffic safety when sudden lane changes or merges are involved. To understand the traffic flow variation on upstream of work zones, the speed distribution was plotted for various lane merge scenarios.

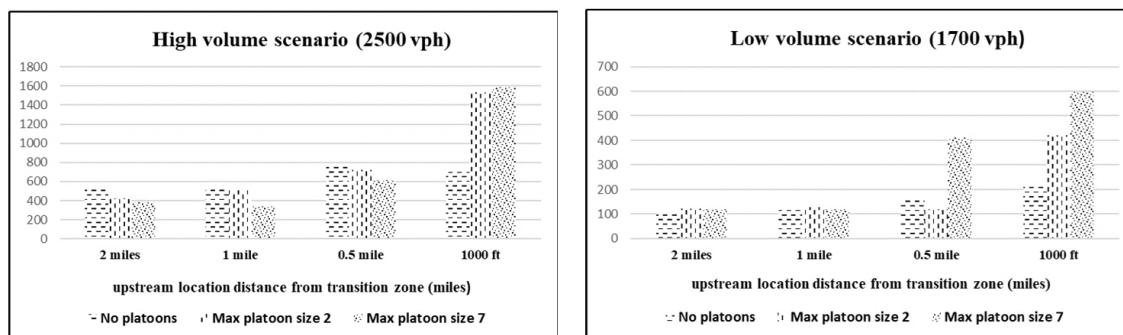


Figure 4: Conflict frequency distribution with various lane change configurations

CONCLUSIONS AND RECOMMENDATIONS

The study's primary objective was to evaluate the safety of truck platooning when encountering critical scenarios like work zones on freeway segments. The analysis mainly focused on the low level of automation in trucks to reflect real-life conditions and the possibility of trucks being connected and automated before other conventional vehicles. This study employed various MPR's 100% MPR on trucks to investigate safety impacts, examining both safety and operational aspects. The simulation integrated a freeway weaving segment and a Basic freeway segment in VISSIM, with vehicle inputs at the start of the Basic freeway segment to facilitate platoons forming before encountering the freeway weaving segment, the work zone was modeled at the end of freeway weaving segment with both the ramp and one freeway segment lane closed. Safety assessment involved calculating indicators such as TTC, Delta V, and critical conflicts, filtered based on a TTC threshold of 1.5 seconds, using trajectory data from VISSIM and the FHWA SSAM tool for analysis.

Based on the analysis, it was evident that in early merge situations on the upstream of work zones, truck platooning has a positive impact on both safety and operational performance as well.

The safety advantages were clearly observed for both CACC level 1 and Level 2 automation. Safety improved as the automation level increased from level 1 to level 2, and the benefits of platooning became more significant in higher traffic volumes. The first section of this study performed various scenarios with only one assumption which is the early merge case, and lane change sign was encountered by trucks one mile before the transition zone. To look at the operational characteristics, the speeds were observed from 1 mile upstream of the transition zone, and speeds became more harmonized because of Platooning.

In the second part of the study, the impact of both early merge and late merge scenarios for truck platoons approaching work zones were examined. Surprisingly, platooning exhibited negative effects in late merge situations, resulting in nearly a 300% increase in critical conflicts and significant reductions in speeds near the transition zone. This effect was particularly pronounced at higher traffic volumes, primarily due to the additional time platoons took to dissolve and execute lane changes compared to conventional trucks. Consequently, long queues formed and unsafe lane changes occurred in the transition zone, highlighting the importance of carefully considering the implementation of platooning strategies in work zones to avoid such adverse outcomes.

The study showcased both positive and negative safety impacts of truck platooning on work zones. However, a notable limitation was the focus on low automation scenarios and cautious driving behavior, which restricted the assessment of operational characteristics such as speed, total delay, and travel times. Future research should explore high levels of automation and include operational characteristics for platoons in different merge situations. Despite these limitations, the study offered valuable insights regarding the various lane merge configurations that could be adopted by platoons in the near future by comparing and highlighting the safety aspects for different platoon sizes and volume levels in the work zones.

ACKNOWLEDGMENTS

This work was conducted under the second Strategic Highway Research Program (SHRP2), which is administrated by the Transportation Research Board of the National Academies of Sciences, Engineering, and Medicine, and it was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials (AASHTO) and the Wyoming Department of Transportation (WYDOT).

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: Study design by Dr. Mohamed M. Ahmed, Data Analysis and Interpretation by Vamsi Krishna. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Adomah, E., A. Khoda Bakhshi, and M. M. Ahmed. 2022. "Safety Impact of Connected Vehicles on Driver Behavior in Rural Work Zones under Foggy Weather Conditions." In *Transportation Research Record*, 2676:88–107. SAGE Publications Ltd.

- Dai, Y., C. Wang, and Y. Xie. 2023. "Explicitly Incorporating Surrogate Safety Measures into Connected and Automated Vehicle Longitudinal Control Objectives for Enhancing Platoon Safety." *Accident Analysis and Prevention* 183 (April).
- Das, A., and M. M. Ahmed. 2022. "Adjustment of Key Lane Change Parameters to Develop Microsimulation Models for Representative Assessment of Safety and Operational Impacts of Adverse Weather Using SHRP2 Naturalistic Driving Data." *Journal of Safety Research* 81 (June): 9–20.
- "FHWA Work Zone Facts and Statistics - FHWA Office of Operations." 2023.
- Guanetti, J., Y. Kim, and F. Borrelli. 2018. "Control of Connected and Automated Vehicles: State of the Art and Future Challenges." *Annual Reviews in Control*. Elsevier Ltd.
- Hallmark, S. 2019. "Preparing Local Agencies for the Future of Connected and Autonomous Vehicles." <http://mndot.gov/research/reports/2019/201918.pdf>.
- Haque, M. M. S., L. R. Rilett, and L. Zhao. 2023. "Impact of Platooning Connected and Automated Heavy Vehicles on Interstate Freeway Work Zone Operations." *Journal of Transportation Engineering, Part A: Systems* 149 (3).
- Hoque, M. M., Q. Lu, A. Ghiasi, and C. Xin. 2021. "Highway Cost Analysis for Platooning of Connected and Autonomous Trucks." *Journal of Transportation Engineering, Part A: Systems* 147 (1).
- Tsai, Y. J., R. Turochy, and N. Jehn. 2022. "Quantitatively Evaluate Work Zone! Driver Behavior Using 2D Imaging, 3-D Lidar & Artificial Intelligence in Support of Congestion Mitigation IV1odel Calibration & Validation."
- Kang, S., H. Ozer, and I. L. Al-Qadi. 2019. "Benefit Cost Analysis (BCA) of Autonomous and Connected Truck (ACT) Technology and Platooning."
- Mchale, G. n.d. "US DOT Truck Platooning Research Program."
- "Micro-Simulation Guide for Automated Vehicles." 19AD.
- "Mitigating Work Zone Safety and Mobility Challenges through Intelligent Transportation Systems Case Studies." 2014.
- Wadud, Z., D. MacKenzie, and P. Leiby. 2016. "Help or Hindrance? The Travel, Energy and Carbon Impacts of Highly Automated Vehicles." *Transportation Research Part A: Policy and Practice* 86 (April): 1–18.
- Wang, C., Y. Xie, H. Huang, and P. Liu. 2021. "A Review of Surrogate Safety Measures and Their Applications in Connected and Automated Vehicles Safety Modeling." *Accident Analysis and Prevention* 157 (July).
- Zhou, J., L. Rilett, and E. Jones. 2019. "Estimating Passenger Car Equivalent Using the HCM-6 PCE Methodology on Four-Lane Level Freeway Segments in Western U.S." *Transportation Research Record*.

Dynamically Expanding Capacity of Autonomous Driving with Near-Miss Focused Training Framework

Ziyuan Yang¹; Zhaoyang Li²; Jianming Hu³; and Yi Zhang⁴

¹Dept. of Automation, Tsinghua Univ. Email: yangziyu22@mails.tsinghua.edu.cn

²Dept. of Automation, Tsinghua Univ. Email: lizhaoya21@mails.tsinghua.edu.cn

³Associate Professor, Dept. of Automation, Tsinghua Univ. (corresponding author). Email: hujm@mail.tsinghua.edu.cn

⁴Professor, Dept. of Automation, Tsinghua Univ. Email: zhyi@tsinghua.edu.cn

ABSTRACT

The long-tail distribution of real driving data poses challenges for training and testing autonomous vehicles (AV), where rare yet crucial safety-critical scenarios are infrequent. And virtual simulation offers a low-cost and efficient solution. This paper proposes a near-miss focused training framework for AV. Utilizing the driving scenario information provided by sensors in the simulator, we design novel reward functions, which enable background vehicles (BV) to generate near-miss scenarios and ensure gradients exist not only in collision-free scenes but also in collision scenarios. And then leveraging the robust adversarial reinforcement learning (RARL) framework for simultaneous training of AV and BV to gradually enhance AV and BV capabilities, as well as generating near-miss scenarios tailored to different levels of AV capabilities. Results from three testing strategies indicate that the proposed method generates scenarios closer to near-miss, thus enhancing the capabilities of both AVs and BVs throughout training.

INTRODUCTION

Autonomous driving technology has got rapid development these years. However, deploying it in the real world for widespread commercial use still poses safety challenges. Testing the safety of autonomous driving systems in real-world scenarios would require covering billions of miles (Feng et al. 2023). This is attributed to the phenomenon of a "*long-tail distribution*" in natural driving datasets, where the occurrence probability of useful safety-critical scenarios for testing autonomous driving systems is typically low. A primary approach to addressing this issue is through simulation platforms, artificially generating safety-critical scenarios, thereby expediting the training and testing of autonomous driving systems. However, in most current work, the Background Vehicle (BV) acting as the aggressor is aimed at generating collisions with the Autonomous Vehicle (AV), which serves as the subsequent subject for testing and training. Within the process of scenario generation, AV remains a fixed agent. This approach limits the capabilities of the trained BV as it cannot adapt its attack performance adaptively based on different abilities exhibited by AV. Moreover, the generated scenarios focus solely on generating collisions, with the gradient of danger existing primarily in factors such as the distance between BV and AV or their relative velocities, limited to the moments just before an actual collision occurs. However, collision scenarios also vary in severity. For instance, if an AV is surrounded by four BVs approaching from different directions simultaneously, the AV would need to practically take off to escape. Therefore, there should be a distinction in the severity of collision scenarios to measure the level of danger in each specific scenario.

Considering that the AV operates as a complicated learner, we propose integrating relevant concepts from educational psychology to aid in formulating training strategies. In educational psychology, there exists a concept known as the "Zone of Proximal Development" (ZPD), introduced by the Soviet psychologist Lev Vygotsky. It refers to the gap between a learner's current abilities and their potential abilities with guidance. Learners require assistance to bridge this gap in their abilities. Thus, a plausible learning model involves collaborative learning between two learners: one slightly more capable than the other (Wertsch et al. 1984). Analogous to the field of autonomous driving, the range of safe driving scenarios that an AV can handle independently represents the Zone of Actual Development (ZAD). The critical scenarios where the AV and BV experience severe collisions stand as the boundary beyond which the capabilities of the AV lie. Meanwhile, the scenarios where the BV poses dangerous yet non-lethal attacks on the AV, allowing the AV a possibility of escape, represent the zone within the AV's proximal development range.

Hence, we designate BV and AV in our training model as learners with slightly differing capabilities. Leveraging the framework of Robust Adversarial Reinforcement Learning (RARL) (Pinto et al. 2017), we devise a training framework where the AV and BV undergo alternating training sessions. Throughout each iteration, BV is trained to attack AV, but with the imposition of rewards at each step to regulate the intensity of the attack. This compels the generation of near-miss scenarios, subsequently training the AV to evade the BV's attacks and proceed with normal driving. Additionally, constraints on speed and steering are implemented to ensure that both the AV and BV adopt behaviors that closely mimic real-world driving situations.

Our primary contributions include: 1. Proposing a quantification method for evaluating the criticality of near-miss scenarios, thereby delineating the severity of critical situations. 2. Introducing a dynamic training approach for both the AV and BV, aiming to consistently expand the boundaries of their respective capabilities while maintaining stability.

RELATED WORK

In the field of autonomous driving scenario generation, adversarial scenario generation stands out as a widely applied method. BV acting as the controlled agent, launch adversarial attacks against AV, thereby creating safety-critical scenarios used in the training and safety testing of autonomous driving system. Modeling scenes as Markov Decision Processes (MDP) and solving them through reinforcement learning (RL) represents a more commonly adopted approach. Feng et al. (2021) and Sun et al. (2021) utilizes Deep Q-Networks to generate discrete adversarial traffic scenarios. Kuutti et al. (2020) employs the Advantage Actor Critic (A2C) algorithm to control a BV in a following-car scenario. Niu et al. (2021) reinforces AV using adversarial models within deep reinforcement learning, gradually exposing them to increasingly challenging scenarios. (Re)2H2O proposed by Niu et al. (2023) employs a hybrid offline-online RL to generate safety-critical scenarios, enhancing scenario authenticity by integrating real-world data. Rempe et al. (2022) introduces the STRIVE framework, optimizing gradients to generate critical scenarios, increasing scenario diversity through adversarial learning methods.

Most existing scenario generation work aims to create collision scenarios, where the gradients of the scenarios exist primarily in non-collision scenarios. This is assessed through physical quantities like the minimum relative distance and relative velocity between two vehicles (Ding et al. 2020). Severity assessment for collision scenarios often centers on analyzing accident data, categorizing collision incidents based on the resulting property and personnel

damage (Hungar et al. 2017), an aspect less explored in scenario generation efforts. However, in scenario generation, it's not that more dangerous scenarios are more beneficial, and AV needs to have a certain probability of escaping safety-critical scenarios in order to improve their capabilities. Hence, generating near-miss scenarios holds substantial value (Feng et al. 2023). Calò et al. (2020) design a search-based method to generate avoidable collision scenarios. Tuncali et al. (2019) propose a scenario generation method based on rapid random searches, guiding near-miss scenario generation through cost function design. In these approaches, AV is a fixed agent, limiting the transferability of generated scenarios to accommodate AV with different skill levels.

In order to dynamically generate corresponding near-miss scenarios during the process of enhancing AV capabilities, we propose a novel framework in this paper focusing on the generation of near-miss scenarios for autonomous driving training. Leveraging RARL, our framework dynamically trains AV and BVs. We measure the criticality of scenarios using collision momentum and obstacle frames, guiding BVs to generate near-miss scenarios. Through experiments, we demonstrate that our approach trains AV to drive more safely. Moreover, BVs can generate near-miss scenarios for AV of specific skill levels, assisting in the steady improvement of AV capabilities.

METHODOLOGY

Problem Formulation

a) Scenario Formulation

Our scenarios take place within a specified length of a straight highway section. The traffic participants within the scenarios encompass one AV (V_0) and N BVs ($V_i, i = 1, \dots, N$) traveling on this road segment. For a specific moment t , the scene at the moment includes the state and action of all vehicles involved:

$$s_t = [s_t^0, s_t^1, \dots, s_t^N]^T, \quad a_t = [a_t^0, a_t^1, \dots, a_t^N]^T$$

For each vehicle at moment t , the state vector comprises its lateral and longitudinal position coordinates (x, y), velocity v , heading angle θ , acceleration α , and angular velocity ω . The action vector consists of brake/throttle control parameter p and steering wheel control parameter δ :

$$s_t^i = [x_t^i, y_t^i, v_t^i, \theta_t^i, \alpha_t^i, \omega_t^i], \quad a_t^i = [p_t^i, \delta_t^i]$$

According to the definitions in (Feng et al., 2020), we depict a scenario as a finite sequence of H -frame scenes:

$$s_0 \rightarrow u_0 \rightarrow s_1 \rightarrow u_0 \rightarrow \dots \rightarrow s_H \rightarrow u_H$$

To simulate the dynamic state transitions in driving behavior as accurately as possible, we utilize the CARLA simulator (Dosovitskiy et al. 2017). We employ the kinematic model within CARLA to facilitate state transitions, denoted as K , to mimic dynamics of the real world:

$$s_{t+1} = \mathbf{K}(s_t, a_t)$$

b) Driving Policy Optimization

In the problem formulation of RL, traffic scenarios can also be described using an MDP, often represented as a sextuple $(S, A, R, P, \rho, \gamma)$, where S and A are state space and action space that have been outlined above. P, ρ and γ are the state transition probability, initial state distribution and discount factor. The reward function at a given moment t is defined as $r_t = r(s_t, a_t)$. Due to different objectives in policy optimization, the reward function is categorized into two types: r_{drive} and r_{attack} . These correspond to training objectives for safe driving and adversarial attack, respectively. r_{drive} includes the following items:

$$r_{drive} = \sum r_{collision, speed, yaw, cooperation}$$

- *Collision*: Prevent collisions between vehicles:

$$r_{collision} = -\alpha_1 \mathbf{I}_{\{\text{if collision happens}\}}$$

- *Speed*: Force speed fluctuates above and below the average speed:

$$r_{speed} = -\alpha_2 |v - \bar{v}|$$

where \bar{v} represents the preset target average velocity.

- *Heading Direction*: Instruct vehicle to go straight:

$$r_{yaw} = -\alpha_3 |\theta + 90|$$

where heading angle for straight driving is -90 degrees.

- *Cooperation*: Encourage driving states among multiple vehicles to have minimal differences:

$$r_{cooperation} = -\alpha_4 (v_{max} - v_{min})$$

Due to different training driving objectives, the composition of items in r_{attack} and r_{drive} differs:

$$r_{attack} = \sum r_{speed, distance, lane, obstacle, impulse}$$

- *Speed*: Encourage speed fluctuates above and below the average speed and drive more smoothly:

$$r_{speed} = -\alpha_5 |v_{adv} - \bar{v}|$$

where \bar{v} represents the preset target average velocity.

- *Distance*: Encourage BVs to approach AV during driving:

$$r_{distance} = -\alpha_6 d_{min}$$

where d_{min} represents the minimum distance between multiple BVs and AV.

- *Lane*: Force vehicles to maintain lane travel:

$$r_{lane} = -\alpha_7 \omega_{adv}$$

where ω_{adv} is the angular velocity of BV.

- *Obstacle*: Encourage BV to occlude AV for a longer duration of frames :

$$r_{obstacle} = \alpha_8 f + \alpha_8 (1 - \frac{d}{5}) \mathbf{I}_{\{if\ d < d_{min}\}}$$

where f represents the number of frames in each step during which BV occludes AV, and d is the current minimum distance between BV and AV in this step. If this distance is smaller than the historical minimum distance d_{min} , an additional reward is provided.

- *Impulse*: Instruct BVs to avoid collisions with AV that are too forceful:

$$r_{impulse} = -\alpha_9 J_{max}$$

where J_{max} represents the maximum impulse of the collision between BVs and AV in current step.

We choose the offline RL algorithm Soft Actor-Critic (SAC) (Haarnoja et al. 2018) for iterating the driving strategy. In conventional Actor-Critic framework, Q-function approximation \hat{Q} is obtained by minimizing the standard Bellman error. Meanwhile, the policy $\hat{\pi}$ is improved by maximizing the Q-function:

$$\hat{Q} \leftarrow \arg \min_Q \mathbf{E}_{\mathbf{s}, \mathbf{a}, \mathbf{s}' \sim \mathcal{U}} \left[\frac{1}{2} \left((Q - \hat{B}^\pi \hat{Q})(\mathbf{s}, \mathbf{a}) \right)^2 \right]$$

$$\hat{\pi} \leftarrow \arg \max_\pi \mathbf{E}_{\mathbf{s}, \mathbf{a} \sim \mathcal{U}} \left[\hat{Q}(\mathbf{s}, \mathbf{a}) \right]$$

where \mathcal{U} represents the data buffer produced by an earlier iteration of policy $\hat{\pi}$ via simulated interactions online. The Bellman evaluation operator \hat{B}^π is defined as $\hat{B}^\pi \hat{Q}(\mathbf{s}, \mathbf{a}) = r(\mathbf{s}, \mathbf{a}) + \gamma \mathbb{E}_{\mathbf{a}' \sim \hat{\pi}(\mathbf{a}' | \mathbf{s}')} \hat{Q}(\mathbf{s}', \mathbf{a}')$. While the objective function of SAC includes an entropy term of the policy distribution:

$$J_\pi(\phi) = \sum_{t=0}^H \mathbf{E}_{\mathbf{s}_0 \in \rho, \mathbf{a}_t \sim \pi_\phi(\cdot | \mathbf{s}_t), \mathbf{s}_{t+1} \sim P(\cdot | \mathbf{s}_t, \mathbf{a}_t)} \left[\gamma^t r(\mathbf{s}_t, \mathbf{a}_t) + \alpha H(\pi_f(\cdot | \mathbf{s}_t)) \right]$$

In this context, π_ϕ signifies the driving policy defined by ϕ , with H denoting the entropy of distribution. The parameter α serves as the temperature hyperparameter within the algorithm.

This setup facilitates exploring a broader range of potential actions while maximizing the expected reward discounted by γ , which boosts the model's resilience and aligns better with real-world traffic scenario applications.

To represent randomness, we depict the driving policy as a Gaussian distribution $\mathcal{N}(\mu, \sigma)$, where μ and σ are approximated by a neural network \mathcal{W}_θ (Niu et al. 2023). Due to the constraints in CARLA, where throttle/brake and steering control inputs are limited between -1 and 1, we incorporate tanh and linear mapping functions to fit the driving policy.

Algorithmic Framework

Borrowing the idea of RARL (Pinto et al. 2017), we divide the training of AV and BVs into two stages. Initially, the policy parameters for both are sampled from random distributions.

Table 1 shows the process of algorithmic framework. In each of the N_{iter} iterations, we alternately execute a two-step optimization process. Firstly, for N_μ iterations, the AV policy parameters remain unchanged as θ^ν , while the parameters θ^μ of BVs' policy are optimized to maximize r_{attack} . For the next step, BV policy parameters θ^μ are held constant for the next N_ν iterations. Before each parameter update, the *roll* function samples N_{step} step samples under the current policy functions of AV and BV at that time, where ε represents the driving environment in the simulator. This alternating process repeats for N_{iter} iterations.

Table 1: Pseudo code of the proposed algorithm

Algorithm 1 Algorithmic Framework	
Input:	Driving environment ε ; BV policy μ and AV policy ν
Initialize:	Learnable parameters θ_0^μ for μ and θ_0^ν for ν
for $i = 1, 2, \dots, N_{iter}$ do	
$\theta_i^\mu \leftarrow \theta_{i-1}^\mu$	
for $j = 1, 2, \dots, N_\mu$ do	
$\{(s_t^i, a_t^{1i}, a_t^{2i}, r_t^{1i}, r_t^{2i})\} \leftarrow roll(\varepsilon, \mu_{\theta_i^\mu}, \nu_{\theta_i^\mu}, N_{step})$	
$\theta_i^\mu \leftarrow policyOptimizer(\{(s_t^i, a_t^{1i}, r_t^{1i}), \mu, \theta_i^\mu\})$	
end for	
$\theta_i^\nu \leftarrow \theta_{i-1}^\nu$	
for $j = 1, 2, \dots, N_\nu$ do	
$\{(s_t^i, a_t^{1i}, a_t^{2i}, r_t^{1i}, r_t^{2i})\} \leftarrow roll(\varepsilon, \mu_{\theta_i^\nu}, \nu_{\theta_i^\nu}, N_{step})$	
$\theta_i^\nu \leftarrow policyOptimizer(\{(s_t^i, a_t^{2i}, r_t^{2i}), \nu, \theta_i^\nu\})$	
end for	
end for	
Return: $\theta_{N_{iter}}^\mu, \theta_{N_{iter}}^\nu$	

The diagram in Figure 1 provides a more intuitive representation of the training process. In the illustration, the grey-colored vehicles represent BVs, while the blue-colored one represent

AV. Yellow markings around the vehicles indicate ongoing training, and the heptagonal stars represent the training aggressiveness strategy. The process flows from left to right. The first AV on the left has already undergone pretraining and convergence. Based on this, BVs are trained to learn normal driving strategies. Subsequently, the training alternates between BV aggressive strategies and AV normal driving strategies.

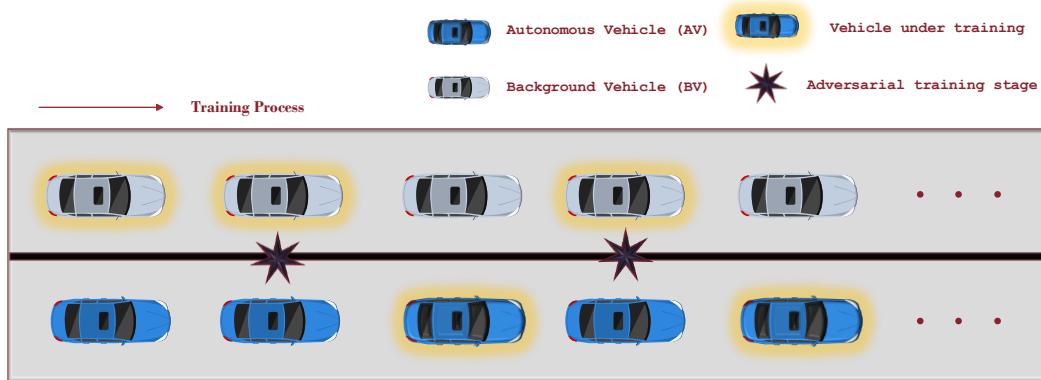


Figure 1. Training process schematic diagram.

EXPERIMENTS

Experiment Settings

a) Environment Setup

Due to the various sensors available in CARLA, providing diverse types of information for driving scenarios, and the closer resemblance of vehicle types and behavioral patterns to real-world driving, we choose CARLA as our simulation environment.

We select a straight segment of road in the “Town04” map of CARLA as our experimental area. Within 20 different starting point combinations, we randomly generated n vehicles (1 AV, $n - 1$ BVs) to drive in CARLA’s *autopilot* mode to the departure area, which are considered as the initial state, ensuring both a reasonable and randomized starting point for the experiment. As for experimental parameters, according to experimental requirements and experience, we assign $\alpha_{1\sim 9}$ values in reward function sequentially as [40, 0.5, 1.2, 0.5, 0.4, 1.5, 3, 10, 0.01].

b) Evaluation Metrics

To evaluate the driving performance of AV and BV, we select several evaluation metrics. To measure the temporal and spatial density of AV-BV collisions, we employ the metrics of **Average Collision Per Second (CPS)** and **Average Collision Frequency Per 100 Meter (CPM)** (Niu et al. 2023). These two metrics signify the average number of AV-BV collisions N_{col} over the total testing time T_{tt} and the total driving distance D_{tt} of AV, respectively: $CPS = N_{col}/T_{tt}$, $CPM = N_{col}/D_{tt}$. Moreover, most existing work primarily design gradients in scenarios where collisions don’t occur. To quantify gradients in collision scenarios, we use the **maximum impulse J_{max}** in collision scenarios to gauge the severity of the scenarios. And we use **obstacle frames (OBF)** in which BVs occlude AV to describe the extent of BVs’ impact on AV’s driving, thus measuring the criticality of scenarios from the perspective of AV driving, rather than simply based on collision occurrence.

c) Testing Policies

To test the policy models of AV and BVs, we select various testing strategies. The initial model for AV was a pretrained AV model (**RL_AV**), while for BV, we considered the BV model from CARLA autopilot mode (**auto_BV**), a pretrained BV model (**RL_BV**), and a BV model trained for 10 rounds with our proposed method (**trained_BV**). Thus we conduct three sets of experiments: 1) **RL_AV** vs **auto_BV**; 2) **RL_AV** vs **RL_BV** (double RL); 3) **RL_AV** vs **trained_BV**

Experimental Results

a) Overall Experiments

In this segment, we analyze the maximum impulse of collisions and OBF under different testing strategies. All RL AV and RL BV models undergo training and convergence before being configured to correspond to their respective training strategies. Figure 2 illustrates the comparison among three testing strategies and the proposed method when there are 3 BVs, and all results are computed using exponential smoothing with a 1.00 coefficient and averaged over 3 random seeds.

The maximum impulse reflects the intensity of collisions caused by BV to AV—larger impulses denote more severe collisions, while smaller impulses represent milder disturbances, indicating more closer to near-miss scenarios. OBF measures the degree of interference of BV to AV driving by reflecting the number of frames in which BV obstructs AV's view. A higher OBF suggests more obstruction to AV's normal driving by BV during operation.

In Figure 2(a), the proposed method during training maintains consistently low values for maximum impulse, indicating relatively mild collision impacts. **RL_AV** vs **auto_BV**, due to BV being set in autopilot mode, actively avoids collisions with AV, hence maintaining lowest values for maximum impulse. In the double RL setting, where both AV and BV are RL models, the collision frequency is higher due to the less efficient simultaneous training of multiple RL models, resulting in consistently higher impulse values. For the BV model trained with the proposed method for 10 rounds, which means $N_{iter} = 10$ in Table 1, its capabilities are enhanced, causing initially higher impulse values when paired with the AV pretrained model. However, over the course of training, these values gradually decrease and converge to a relatively high but stable level.

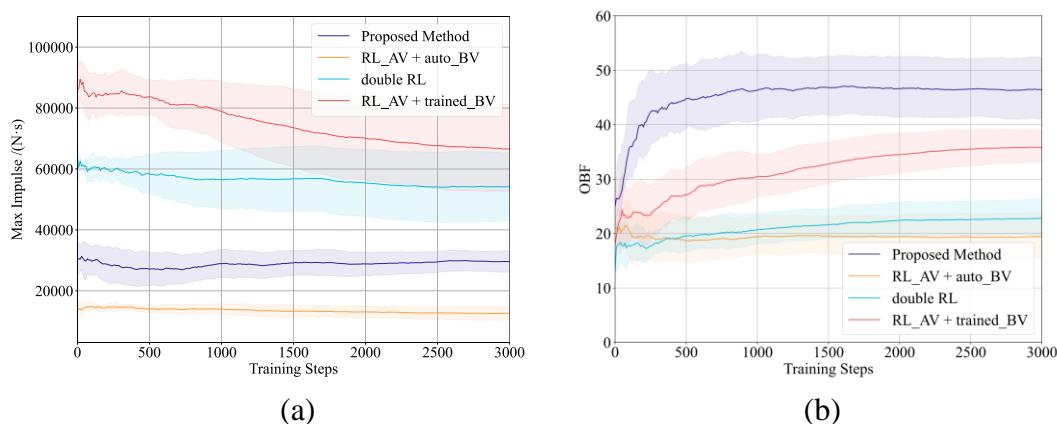


Figure 2. Comparison between four training strategies.

In Figure 2(b), the BV trained using the proposed method demonstrates a notable capability in creating substantial occlusion interference to AV, maintaining a comparatively high OBF after an initial rise. When BV operates in autopilot mode, resulting in no aggressive behavior towards AV, the OBF remains relatively low. The double RL model tends to end scenarios early due to collisions occurring at an earlier stage, resulting in lower OBF. When BV is the trained model, initially, due to its relatively aggressive nature towards AV, collisions occur earlier, leading to a lower OBF. As training steps increase and AV's capabilities improve, scenario durations increase, subsequently elevating the OBF, eventually converging to a relatively high level.

b) Controlled Experiments

In this section, we conduct tests using BV trained over different rounds (Round = 1, 2, ..., 15) along with a pretrained AV and AV trained over different rounds along with a BV after training 15 rounds. Figure 3 illustrates the changes in CPS (Collisions per Scene) and CPM (Collisions per Minute) as BV or AV training rounds increase.

From Figure 3(a), it is evident that as the BV training rounds increase, its aggressiveness towards pretrained AV also escalates. Both CPS and CPM show an upward trend, indicating an increase in the frequency of collisions per second and per hundred meters. And from Figure 3(b), when pretrained AV encounters trained 15 rounds BV, it is obvious that CPS and CPM decrease as AV training rounds increase, which means that BV will cause more collisions to AV at lower training procedure.

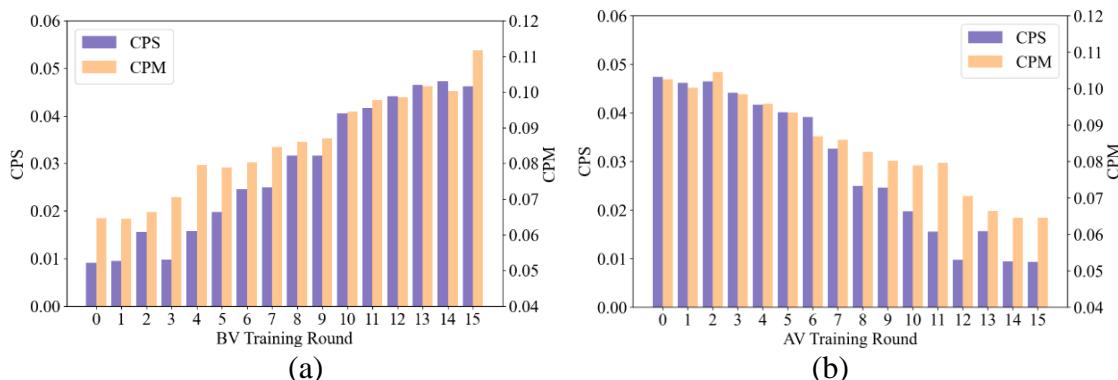


Figure 3. Controlled experiment results of varying-capability AV and BV

c) Scenarios in Simulation

In this section, we present some scenarios during training rendered in the CARLA simulator to better illustrate the training outcomes. The blue one in Figure 4 is AV, and grey ones are BVs.

Figure 4(a) illustrates the BV in the right front of AV changing lanes, obstructing AV. Subsequently, the BV on the left rear side of AV accelerates, slightly turns towards AV, closely follows it, interfering with its lane-changing to evade the front BV. In Figure 4(b), both the BV on the right front and left rear of AV change lanes simultaneously. Meanwhile, the BV on the right rear accelerates, blocking AV from three directions. Figure 4(c) portrays a scenario where a vehicle in the right front of AV changes lanes. AV attempts to swerve to the left, but then a BV behind AV's original lane accelerates, obstructing AV, preventing it from returning to its initial lane.



Figure 4. Typical scenarios in simulation

CONCLUSION

This paper introduces a near-miss focused autonomous driving training framework. By designing reinforcement learning rewards, it establishes the foundation for generating near-miss scenarios, ensuring gradients exist not only in collision-free scenes but also in collision scenarios. Leveraging the RARL framework for training both AV and BV, it consistently enhances AV capabilities while enabling BV to purposefully generate near-miss scenarios to aid AV training. Experimental results demonstrate the effectiveness of this method in significantly improving AV capabilities. Moreover, BV's aggressiveness gradually increases during training. Compared to other training strategies, the proposed method generates scenarios that closely resemble near-miss situations. Future research could consider incorporating a wider array of traffic participants to enhance the similarity between simulated scenarios and real-world settings.

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China under Grant No. 62333015 and Beijing Natural Science Foundation L231014.

REFERENCES

- Calò, A., Arcaini, P., Ali, S., Hauer, F., and Ishikawa, F. (2020, October). Generating avoidable collision scenarios for testing autonomous driving systems. In *2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST)* (pp. 375-386). IEEE.
- Ding, W., Xu, M., and Zhao, D. (2020, May). Cmts: A conditional multiple trajectory synthesizer for generating safety-critical driving scenarios. In *2020 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 4314-4321). IEEE.

- Ding, W., Xu, C., Arief, M., Lin, H., Li, B., and Zhao, D. (2023). A survey on safety-critical driving scenario generation—A methodological perspective. *IEEE Transactions on Intelligent Transportation Systems*.
- Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017, October). CARLA: An open urban driving simulator. In *Conference on robot learning* (pp. 1-16). PMLR.
- Feng, S., Feng, Y., Yu, C., Zhang, Y., and Liu, H. X. (2020). Testing scenario library generation for connected and automated vehicles, part I: Methodology. *IEEE Transactions on Intelligent Transportation Systems*, 22(3), 1573-1582.
- Feng, S., Yan, X., Sun, H., Feng, Y., and Liu, H. X. (2021). Intelligent driving intelligence test for autonomous vehicles with naturalistic and adversarial environment. *Nature communications*, 12(1), 748.
- Feng, S., Sun, H., Yan, X., Zhu, H., Zou, Z., Shen, S., and Liu, H. X. (2023). Dense reinforcement learning for safety validation of autonomous vehicles. *Nature*, 615(7953), 620-627.
- Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018, July). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning* (pp. 1861-1870). PMLR.
- Hungar, H., Köster, F., and Mazzega, J. (2017). Test specifications for highly automated driving functions: Highway pilot.
- Kuutti, S., Fallah, S., and Bowden, R. (2020, May). Training adversarial agents to exploit weaknesses in deep control policies. In *2020 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 108-114). IEEE.
- Niu, H., Hu, J., Cui, Z., and Zhang, Y. (2021, October). Dr2l: Surfacing corner cases to robustify autonomous driving via domain randomization reinforcement learning. In *Proceedings of the 5th International Conference on Computer Science and Application Engineering* (pp. 1-8).
- Niu, H., Chen, Q., Li, Y., and Hu, J. (2023). Stackelberg Driver Model for Continual Policy Improvement in Scenario-Based Closed-Loop Autonomous Driving. arXiv preprint arXiv:2309.14235.
- Niu, H., Ren, K., Xu, Y., Yang, Z., Lin, Y., Zhang, Y., and Hu, J. (2023). (re)2h2o: Autonomous driving scenario generation via reversely regularized hybrid offline-and-online reinforcement learning. In *2023 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1-8).
- Niu, H., Xu, Y., Jiang, X., and Hu, J. (2023). Continual Driving Policy Optimization with Closed-Loop Individualized Curricula. arXiv preprint arXiv:2309.14209.
- Pinto, L., Davidson, J., Sukthankar, R., and Gupta, A. (2017, July). Robust adversarial reinforcement learning. In *International Conference on Machine Learning* (pp. 2817-2826). PMLR.
- Sun, H., Feng, S., Yan, X., and Liu, H. X. (2021). Corner case generation and analysis for safety assessment of autonomous vehicles. *Transportation research record*, 2675(11), 587-600.
- Tuncali, C. E., and Fainekos, G. (2019). Rapidly-exploring random trees-based test generation for autonomous vehicles. arXiv preprint arXiv:1903.10629.
- Wertsch, J. V. (1984). *The zone of proximal development: Some conceptual issues. New directions for child development*.

Delivering Tomorrow: Analyzing Automated Delivery Vehicle Narratives through Media Mining

Subasish Das, Ph.D.¹; and Boniphace Kutela, Ph.D., P.E.²

¹Dept. of Civil Engineering, Texas State Univ., San Marcos, TX. Email: subasish@txstate.edu

²Texas A&M Transportation Institute, Bryan, TX. Email: b-kutela@tti.tamu.edu

ABSTRACT

Amid the COVID-19 pandemic, automated delivery vehicles (ADVs) have gained even more significance due to their touchless capabilities, aligning perfectly with the need for contactless delivery services to safeguard public health. This study collected and conducted quantitative analysis on the most relevant 86 news articles focused on ADVs by employing sophisticated text network and collocation analyses to extract valuable insights. The findings indicated a noteworthy trend: express delivery services have forged partnerships with various ADV companies, leveraging their capabilities to tackle the challenging “last mile” delivery stage. Furthermore, the study revealed a growing integration of ADVs into diverse delivery sectors, particularly in food, grocery, and other on-demand services. Overall, this study provides invaluable insights into the current state of ADVs and their widespread adoption across various delivery sectors.

INTRODUCTION

E-commerce has sparked a transformation in the retail industry, resulting in a surge in demand for efficient delivery services. This surge has paved the way for the development of Automated Delivery Vehicles (ADVs) tailored to meet these delivery demands. ADVs have captured significant attention due to their potential to revolutionize the delivery of goods, offering numerous advantages over conventional delivery methods.

A crucial milestone advancing ADV technology was reached with the National Highway Traffic Safety Administration's (NHTSA) approval of ADV deployment at low-speed thresholds. This approval enables the use of ADVs in specific controlled environments, such as college campuses, business parks, and other restricted areas, where they can operate safely and effectively according to AASHTO (2010). Unlike traditional automated vehicles (AVs), ADVs are exclusively designed for delivery purposes and have no human occupants. This design grants them several key advantages over conventional delivery methods. Firstly, ADVs are projected to significantly reduce delivery costs by eliminating the need for human drivers. Secondly, their operations are optimized for efficiency, as they can plan and execute delivery routes without requiring breaks or rest periods. Finally, ADVs boast eco-friendliness by releasing fewer harmful emissions than traditional delivery vehicles.

An essential advantage of ADVs is their potential to address the issue of crashes caused by human error. The Highway Safety Manual (HSM) reveals that 93% of crashes result from human errors. By eliminating the need for human drivers, ADVs remove this risk factor, contributing to safer roads and enhancing overall road safety. During the COVID-19 pandemic, ADVs have become even more critical due to their touchless capability. Online shopping has increased significantly, with people avoiding public places to maintain social distancing. ADVs allow

packages to be delivered without human interaction, reducing the risk of virus transmission. This has increased the demand for ADVs and highlighted their importance in maintaining a safe and healthy environment.

As a relatively new technology, there are still many uncertainties and doubts regarding ADVs. While there are significant benefits to using ADVs for delivery services, there needs to be a greater understanding of the current state of research and development in this field. Very few studies have focused on ADVs, instead typically focusing on general AVs without accounting for the unique aspects and challenges that emerge with ADV adoption. Berendt (2017) mentioned that news media plays a crucial role in communicating further developments on emerging technologies, including ADVs, and is a beneficial tool for learning more about the state of development of significant research areas.

However, analyzing news articles is challenging due to the vast amount of information available online and the lack of structured data such as topic labels and semantic structure. This study conducted a comprehensive literature review and text mining of news media to collect relevant articles on ADVs, and text network and collocation analyses were employed to identify current trends and topics surrounding ADVs.

LITERATURE REVIEW

ADVs have garnered significant attention for their potential to revolutionize the logistics industry. Flämig (2016) conducted a comprehensive study outlining strategies to integrate AVs into road freight transportation systems within public facilities. The research aimed to determine the feasibility of employing AVs in logistic systems and provided historical insights into in-house logistics, shedding light on companies' reasons for adopting AVs. Additionally, the study addressed essential aspects such as navigation, safety, and control requirements for ADVs. On the other hand, Paddeu and Parkhurst (2020) investigated the developments surrounding ADVs and presented an extensive review focusing on their current and future state of development. They identified research gaps concerning assessing ADVs' economic benefits and development costs. Furthermore, practice and policy barriers were highlighted as significant challenges that still need to be addressed. Both studies contribute valuable information to the ongoing exploration of ADVs' potential impact on the logistics sector.

Crash prevention and safety

Crash prevention is a critical concern for ADVs, as it is necessary to consider the importance of safety when introducing new technology. Operating primarily in densely populated neighborhoods, ADVs must prioritize safety, particularly in ensuring smooth interactions with pedestrians. Strauss et al. (2021) underscored in the safety report their focus on minimizing physical harm in case of ADV-pedestrian collisions. In this pursuit, NHTSA has identified 12 safety elements for autonomous driving systems, and Nuro's safety report addresses their approach to these elements. ADVs offer a significant advantage in reducing fatalities and injuries related to vehicle occupants, primarily due to their occupant-less nature. Note that Strauss et al. (2021) specializes in robotics and is at the forefront of developing ADVs. It also is the first company to be granted an autonomous exemption by the NHTSA. Witcher et al. (2021) conducted a study demonstrating that full market penetration of occupant-less ADVs could potentially reduce fatalities by 58.2% and injuries by 61.8%. Based on current safety data, Dorr

and Seba (2020) predicted a 90% decrease in crashes involving AVs compared to conventional vehicles. Additionally, Hawkins (2017) wrote that Tesla's introduction of autopilot in 2015 resulted in a 40% drop in crash rates. However, some critics doubt the feasibility of a 90% crash reduction and raise concerns about potential risks AVs might introduce to traffic safety. Mueller et al. (2020) suggested that AVs could reduce up to 34% of traffic crashes, with even more significant improvements if technology could eliminate all traffic violations.

Though direct studies on ADV safety impact are limited, safety features of ADVs bear similarities to those of general AVs. Morando et al. (2018) utilized a simulation-based surrogate safety measure approach to study AV safety impacts, revealing substantial overall safety improvements at high market penetration rates, despite AVs operating with smaller headway to enhance roadway capacity. Ye and Yamamoto (2019) used a heterogeneous flow model to explore the impact of connected autonomous vehicles (CAVs) on traffic safety, finding that higher market penetration rates yielded additional benefits for traffic safety, especially when employing a more cautious car-following strategy. Papadoulis et al. (2019) developed a decision-making CAV control algorithm and implemented the Surrogate Safety Assessment Model (SSAM) to evaluate its safety effects. The results indicated that even at low market penetration rates, CAVs could significantly reduce traffic conflicts. Katrakazas et al. (2019) developed a novel risk assessment approach using interaction-aware motion models and Dynamic Bayesian Networks (DBN), which showed an up to 10% improvement in the interaction-conscious model, particularly in collision-prone traffic conditions.

Public Acceptance of ADV

Public acceptance of ADVs is crucial as this emerging technology still faces skepticism and uncertainties among consumers. Understanding the factors influencing public acceptance is essential to effectively deploy ADV services, allowing companies to address concerns and tailor their approach accordingly. Pani et al. (2020) conducted a comprehensive analysis of public feedback from 483 Portland customers, examining desires, attitudes, and willingness to pay (WTP) for ADVs. Their study provides practical insights to promote the mass adoption of eco-friendly delivery vehicles by identifying latent class WTP determinants. On the other hand, Kapser and Abdelrahman (2020) highlighted that ADVs may only reach their full potential if they are not widely embraced as a viable alternative. Additionally, Kapser et al. (2021) investigated the gender differences in ADV acceptance during the COVID-19 pandemic, incorporating gender as a moderator within the Unified Theory of Acceptance and Use of Technology (UTAUT2) framework. Employing structural equation modeling on questionnaire data, they concluded that price sensitivity is a significant factor influencing ADV acceptance among consumers in Germany, while perceived risk plays a decisive role in the acceptance of ADVs among female consumers in the country.

Despite their increasing significance in the delivery industry, ADVs have yet to receive substantial research attention. As the potential for mass deployment of ADVs becomes more imminent, researchers are increasingly invested in developing efficient and safe ADV networks and operation systems. However, a critical aspect that requires attention is the lack of studies on public perception of ADVs. Understanding how the public perceives this emerging technology is essential for its successful integration into society. Public perception plays a pivotal role in shaping the acceptance and adoption of new technologies. By examining the public perception of ADVs, this study aims to identify potential barriers, concerns, and misconceptions that might

hinder their widespread adoption. Additionally, understanding the factors contributing to positive or negative perceptions of ADVs can aid in developing targeted strategies to foster public trust and confidence in this technology.

As ADVs promise to transform the delivery landscape and significantly impact daily life, studying public perception is vital to ensure their successful and harmonious integration into our communities. Such research can contribute to developing effective communication strategies, policy frameworks, and educational initiatives that address public concerns and facilitate the responsible deployment of ADVs. Ultimately, a well-informed and positive public perception of ADVs can maximize their benefits while minimizing potential challenges during their implementation and deployment.

METHODOLOGY

Data Collection

In this study, data was collected using Google News Alert and entering relevant keywords related to ADVs to gather pertinent Uniform Resource Locators (URLs) pointing to webpages with related information. To automate the web scraping process, the popular R package Quicknews (<https://github.com/jaytimm/quicknews>) was utilized. Through this approach, 86 articles were extracted and chosen for subsequent analysis. Table 1 presents the selected news articles chronologically, covering the period from January 30, 2018, to July 1, 2021. The most recent articles are listed at the bottom of the table. To ensure the accuracy and relevance of the collected data, all news articles were manually inspected to verify their connection to ADVs. The headlines of these articles encompass a wide range of topics, providing insights into various aspects of ADV development and implementation. Some highlighted subjects include the progress made by companies like Nuro and Google in creating self-driving delivery vehicles, collaborative efforts between retailers and AV startups, and regulatory changes aimed at facilitating the operation of AVs on public roads.

Table 1. Sample Headlines of News Articles

No.	date	News Headlines
1	01/30/18	Ex-Google Engineers Raise \$92 Million To Roll Out Robot Delivery Vehicles This Year
2	01/30/18	Nuro raises \$92 million to develop autonomous delivery vehicle
3	01/04/19	DoorDash eyes driverless vehicles for grocery delivery
4	01/07/19	udelv to begin autonomous last-mile delivery pilot in Houston
5	02/06/20	U.S. lets autonomous vehicle bypass human-driver safety rules
6	02/06/20	Nuro's R2 autonomous delivery vehicle gets official OK to ditch steering wheel
7	03/12/20	Delivery robot firm Neolix closes RMB 200 million Series A+
8	07/06/21	Yandex Self-Driving Group partners with GrubHub to bring robotic delivery to college campuses
9	10/07/22	Amazon shelves pilot of robotic delivery vehicle
10	11/09/22	Waymo can now charge for fully driverless services in San Francisco

Text Network Analysis

Text mining, a popular technique in data mining and a branch of natural language processing (NLP), permits exploratory analysis of unstructured datasets. It can be extended to complex statistical analysis based on specific research questions. A powerful tool by (Hunter, 2014; Kwayu et al. 2021) that was used within text mining is text network analysis (TNA), which uncovers hidden trends in unstructured text data. TNA has been used across several fields, including literature and linguistics, traffic safety and operations, and the bibliometrics of transportation studies by (Paranyushkin, 2011; Kutela et al. 2021; Kutela et al., 2022). TNA uses nodes and edges to establish relationships between keywords within a corpus (a large, structured body of text).

Several steps are undertaken to process the data during TNA. Normalization is the first step, in which unstructured data is converted into a structured format, symbols are removed, and all text is converted to lowercase. Then, this processed data is used to develop a matrix representing keywords and their respective frequencies of occurrence. Next, the matrix is visualized, with node sizes representing the keywords' frequency. Two common metrics for comparative analysis are employed: document frequency, which measures the number of documents containing the keyword, and keyword frequency, which counts how often the keyword appears in the documents. Collocation frequency is another important metric that assesses the proximity between two keywords in the corpus. This metric is crucial in forming text clusters or communities of keywords.

These communities represent groups of clustered keywords in the text network, and a single network can have several of these communities. Barnier and Privé (2023) implemented a simple Reinert textual clustering method which was employed to further assess the key clusters for each source in the introduction section. This method utilizes the same data preparation approach as TNA but results in several clusters, each representing a specific theme. The following section presents the results and discussions of this analysis.

RESULTS AND DISCUSSIONS

Qualitative Data Analysis

This study used Taguette, a free and open-source tool designed for qualitative research, for quantitative data analysis. Rampin and Rampin (2021) allows users to import and tag a variety of formats, including text, images, and PDF documents. Several major steps were conducted during the qualitative data analysis. The first step was identifying key themes or topics that need to be explored in the news media reports and creating tags to represent these themes. Some of the developed tags were: 'ADV Development,' 'Partnerships,' 'Safety Concerns,' 'Regulatory Changes,' and 'Public Perception.' The second step was reading the news media reports, identifying text segments corresponding to each tag, and highlighting and applying the relevant tags to each segment by selecting the appropriate tag from the Taguette interface. The third step was annotating the tagged segments and adding notes in Taguette if additional context or insights needed to be provided. Annotations and notes can aid in the later stages of analysis and reporting. The final step was analyzing the tagged data to identify trends, patterns, and insights related to ADVs. Some of the main themes identified from the analysis are discussed below.

Several news articles reported on COVID-19 and delivery services made by ADVs. Amidst COVID-19 restrictions and limited entertainment options, there has been a notable surge in

online shopping, leading to reduced driving activity and increased reliance on delivery services. As a result, delivery vans have been covering unprecedented distances. The landscape of autonomous or self-driving cars is transforming due to the COVID-19 crisis. Several companies are now exploring the utilization of these vehicles for delivery purposes. Among them, Nuro is a well-known company developing its autonomous delivery vehicles (ADVs). The autonomous mobile robot industry is experiencing significant growth, with the e-commerce sector witnessing a surge, particularly during the pandemic. Despite the recession, the market is expected to continue expanding due to the increasing demand for automated delivery solutions. Several news articles reported on food delivery services made by ADVs. ADVs have emerged as an innovative solution for food delivery, offering efficient and timely transportation from restaurants to customers' doorsteps without the need for human drivers. By programming these vehicles, they can take the most optimized routes to avoid traffic congestion, ultimately reducing delivery times. The reports reveal the significant interest in ADVs within the food delivery industry. Companies like Domino's Pizza, DoorDash, and Uber Eats have been actively testing and implementing ADV systems to enhance delivery services. The ADV companies are being targeted for the last-mile delivery of local goods and services in collaboration with major grocery or wholesale firms. For instance, Nuro, a company specializing in local deliveries, has joined forces with the long-established grocery retailer Kroger to offer convenient same-day delivery services.

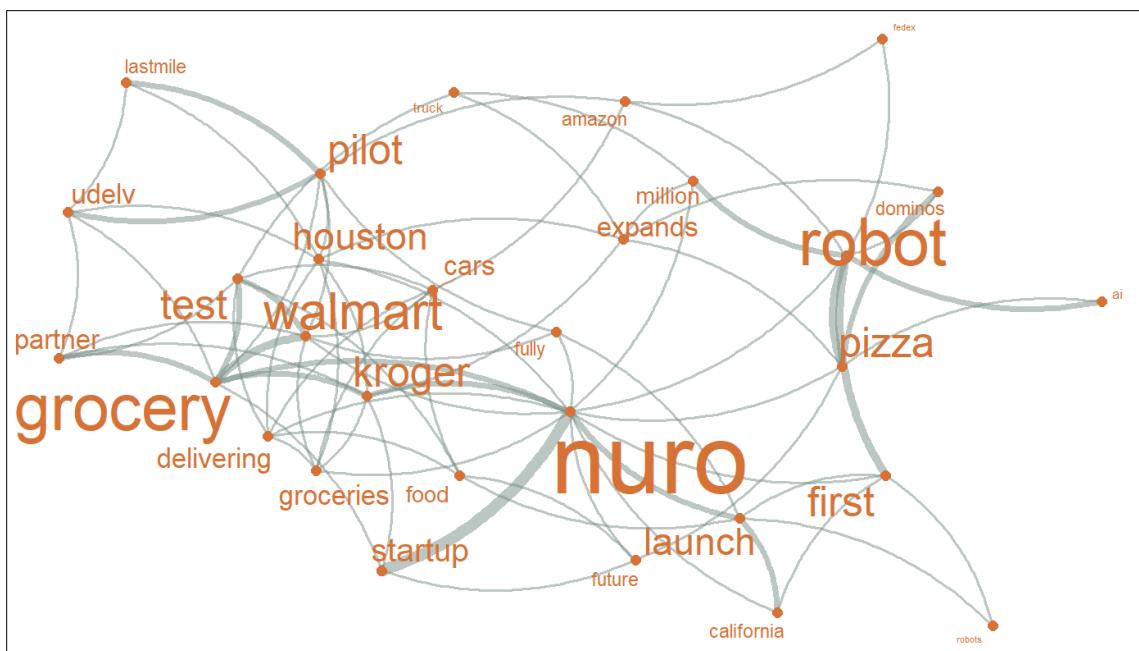
Text Network Analysis

Content of the News Headlines

The contents of the news headlines of ADVs were evaluated, and the text network of news headlines is presented in Figure 1, while Table 2 presents the metrics of the text networks' top frequent words and collocations. The text network of news headlines is heavily centered on the keywords *nuro*, *grocery*, *robot*, *walmart*, and *kroger*. This is because all of the news collected in this paper was ADV-related. In addition, some keywords of news headlines with a relatively low representation are presented, like *amazon*, *launch*, *udelv*, *partner*, and *pizza*, which indicates other industries involving ADVs. Observing the text network, some linked keywords include *startup* and *nuro*, *pizza* and *robot*, *dominos* and *pizza*, *cars* and *delivering*, and *coronavirus* and *tests*.

Although Figure 1 shows the text networks, a comparative analysis of the networks can be performed using the keyword and collocation frequencies. According to the results in Table 2, among the top 20 keywords, the most frequent word in this metric is *nuro* (an American robotics company), appearing 15 times. The top 5 most frequent words include *robot*, *grocery*, *pizza*, and *first*, which appeared 12, 7, 6, and 6 times, respectively. There are also some less frequent words like *walmart*, *amazon*, *kroger*, *udelv*, and some companies with ADVs.

In addition to the individual keywords, the collocated keywords' results can be presented in the metrics. Among the top 20 linked keywords, the most frequent collocation in this metric is *startup nuro*, appearing 4 times. The top 5 most frequent collocations also have *pizza robot*, *domino's pizza*, *cars delivering*, and *coronavirus tests*, which appeared 3, 2, 1, and 1 times, respectively. There are also some less frequent collocations like *cars groceries*, *expands program*, *food program*, *gatik expands*, and *delivering food*, among others.

**Figure 1.** Text networks from News Headlines**Table 2.** Top frequent words and collocations in the news headlines.

Word	Count	Collocation	count	lambda	z
nuro	15	startup nuro	4	6.0918	5.3602
robot	12	pizza robot	3	4.6030	4.4441
grocery	7	dominos pizza	2	4.9162	4.3258
pizza	6	cars delivering	1	5.5148	3.8969
first	6	coronavirus tests	1	5.5148	3.8969
walmart	5	groceries arizona	1	5.5148	3.8969
startup	5	industry chain	1	5.5148	3.8969
kroger	5	roll robots	1	5.5148	3.8969
amazon	5	service safety	1	5.5148	3.8969
pilot	5	bring robotic	1	5.0013	3.7958
launch	5	cars groceries	1	5.0013	3.7958
million	4	expands program	1	5.0013	3.7958
future	4	food program	1	5.0013	3.7958
food	4	gatik expands	1	5.0013	3.7958
partner	4	groceries houston	1	5.0013	3.7958
california	4	middlemile truck	1	5.0013	3.7958
udelv	4	truck begins	1	5.0013	3.7958
truck	4	delivering food	1	4.4877	3.7022
test	4	delivering groceries	1	4.4877	3.7022
cars	3	expands houston	1	4.4877	3.7022

Content of the News Reports

Figure 3 presents the four clusters based on various statistical methods, including chi-square, likelihood, frequency, and document proportion. According to the clusters for chi-square (Figure 3(a)), the first cluster contains keywords such as *zerooccupant*, *windshield*, *mirrors*, *Prius*, and *nuro*, which are associated with the automobile topic. For the second cluster in Figure 3(a), the keywords mainly cover the field of business and economy, including the keywords *facilities*, *Nevada*, *track*, *jobs*, *manufacturing*, and *facility*. Some retail companies are mentioned by Cluster 3, such as *Walmart*, *Kroger*, *grocery*, *Domino's*, and *food*. The last cluster in this figure, Cluster 4, includes the keywords: *drone*, *robots*, *cameras*, *lidar*, and *autonomy*, which are associated with the technology and innovation of ADVs.

Figure 2(b) presents clusters derived from likelihood. From the figure, these keywords may be related to automobiles in ADV, which is like cluster 1 in Figure 3(a); such keywords are *zerooccupant*, *windshield*, *Prius*, *cameras*, *mirrors*, and *seat*. On the other hand, the second cluster contains keywords such as *facility*, *facilities*, *Nevada*, *tax*, *manufacturing*, and *million*. These keywords are linked to business and economy. Keywords relating to some companies with delivery services can be observed in the third cluster: *walmart*, *kroger*, *doordash*, *domino's*, and *fedex*. The fourth cluster includes keywords such as *lidar*, *robots*, *cameras*, *drone*, *truck*, and *areas*, explaining the technology and innovation in the delivery services.

Figure 2(c) presents clusters showing the frequency of keywords in these news media. The figure shows some keywords in the first cluster, like *Nuro*, *zerooccupant*, *pilot*, *road*, *testing*, and *mirrors*, which are related to technology and innovation of new mobility services such as Nuro. In the second cluster, the keywords are associated with the analysis of business and economy, with *facility*, *track*, *tax*, *manufacturing*, *capital*, *build*, and *plans* included in this cluster. Keywords in the third cluster, like *walmart*, *amazon*, *waymo*, *nuro*, *pilot*, and *Kroger* show that this cluster is mainly about the industry of grocery and delivery services. The last cluster includes *ai*, *robots*, *robot*, *technology*, and *industry*, which show the cluster is associated with technology in the field of ADVs.

Figure 2(d) presents clusters on keywords proportion in the document. The figure shows that some keywords in the first cluster are mainly related to technology used in ADVs, like *generation*, *pilot*, *mirrors*, *nuro*, and *zerooccupant*. In the second cluster, the keywords are associated with the development of business in the field of industry, as *goed*, *million*, *manufacturing*, *capital*, *plans*, and *track* are included in this cluster. Keywords in the third cluster like *kroger*, *customers*, *services*, *food*, and *goods* show that this cluster is mainly about the services in the grocery company. The fourth cluster contains keywords like *technology*, *robots*, *robot*, *safety*, and *drivers*, related to the technology used in the delivery services.

The content of the keywords in news media was evaluated. Figure 3 presents the text network of the keywords section in the four clusters, while Table 2 and Table 3 present the performance metrics of the text networks. The text network of keywords in cluster 1 (Figure 3(a)) is heavily centered on the words *nuro*, *technology*, *zerooccupant*, *first*, and *pilot*, which are associated with the advanced technology in ADVs. Further, the title sections constitute keywords such as *mountain*, *mirrors*, *permit*, and *commercial*, indicating the news also covers some information about business and the economy. Some other keywords have a relatively low representation, including *testing*, *generation*, *eleven*, and *road*.

Figure 3(b) presents the text network for cluster 2. The text network is heavily centered on the words *nuro*, *testing*, *track*, and *facility*, indicating this cluster is mainly related to business

and economy in ADVs. Observing the text network (Figure 4(b)), some words are linked, such as *get* and *permit*, *first* and *generation*, *permit*, and *test*, and *mountain* and *view*. Figure 3(c) presents the text network for cluster 3, which indicates the keywords in the field of delivery services in some companies, especially grocery companies. The text network is heavily centered on *grocery*, *amazon*, and *service*. Some words are significantly linked, including *nuro* and *service*, *pilot* and *program*, and *electric* and *technology*. These linked keywords show the core of the research. Figure 4(d) presents the text network for cluster 4. The text network is heavily centered on the keywords *technology*, *robots*, *drivers*, *like*, and *amazon*. These words are associated with innovation in the delivery services of ADVs. There are also some words with low presentations, like *road*, *drone*, *robotics*, and *truck*.

Although Figure 3 shows that the four networks share different aspects in news media about ADVs, a comparative analysis of the four networks can be performed using the keyword frequencies. According to the results in Figure 3, among the top 20 keywords there are some common for different sides, like *nuro*, which is ranked first in Clusters 1, 2, & 3. There are also other common keywords with various ranks. For instance, the keyword *technology*, ranked 13th in cluster 1, appearing 13 times, is ranked fourth in cluster 3, appearing 67 times, and is ranked second in cluster 4, appearing 76 times. Keywords often appear in multiple clusters, including *pilot* in clusters 1 & 3, *like* in clusters 1 & 4, *testing* in clusters 1, 2, & 3, and *road* in clusters 1 & 4. There are several common words in clusters 3 & 4, such as *service*, *first*, *amazon*, *food*, and *goods*.

Table 3. Top frequent words in news report contents by clusters.

Word	Cluster 1 (count)	Word	Cluster 2 (count)	Word	Cluster 3 (count)	Word	Cluster 4 (count)
nuro	48	nuro	26	nuro	113	robots	93
mountain	29	nevada	19	customers	79	technology	76
view	26	las	11	service	75	robot	58
zerooccupant	23	vegas	11	technology	67	people	55
thats	22	fully	10	walmart	60	amazon	54
make	20	testing	10	pilot	59	first	50
nuros	18	goed	10	safety	57	food	49
like	15	jobs	10	first	53	many	46
way	15	facility	8	amazon	51	like	44
pilot	15	track	8	test	46	ai	44
permit	15	press	8	grocery	45	around	41
see	13	include	8	gatik	43	human	40
technology	13	tax	8	food	41	business	39
testing	13	abatements	8	kroger	40	goods	38
eleven	13	according	8	goods	39	industry	36
generation	12	wage	8	waymo	39	road	35
mirrors	12	application	8	trucks	38	last	33
road	12	test	6	program	37	need	33
really	12	southern	6	testing	37	robotics	33
get	11	state	6	cruise	37	service	32

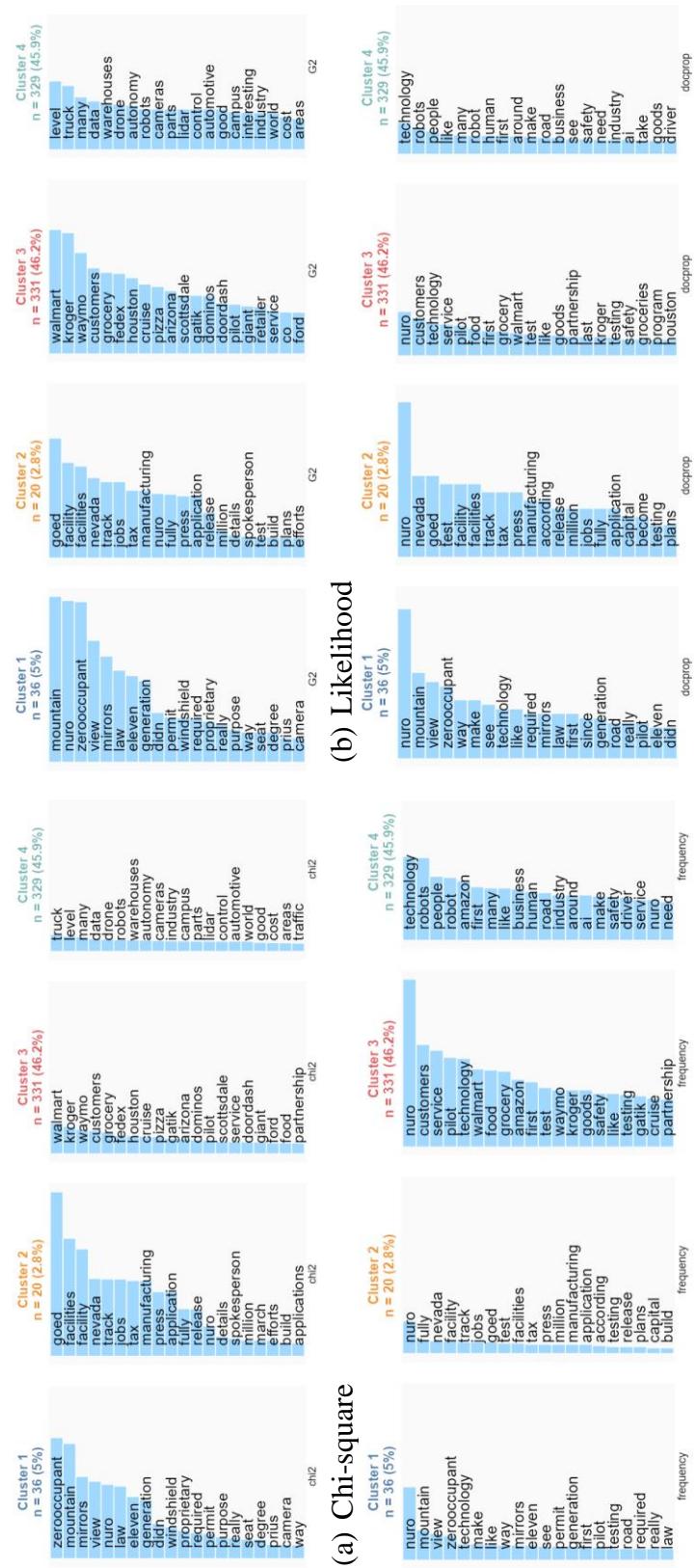


Figure 2. Clusters generated from the news reporting contents.

Downloaded from ascelibrary.org by Texas State University on 06/14/24. Copyright ASCE. For personal use only; all rights reserved.

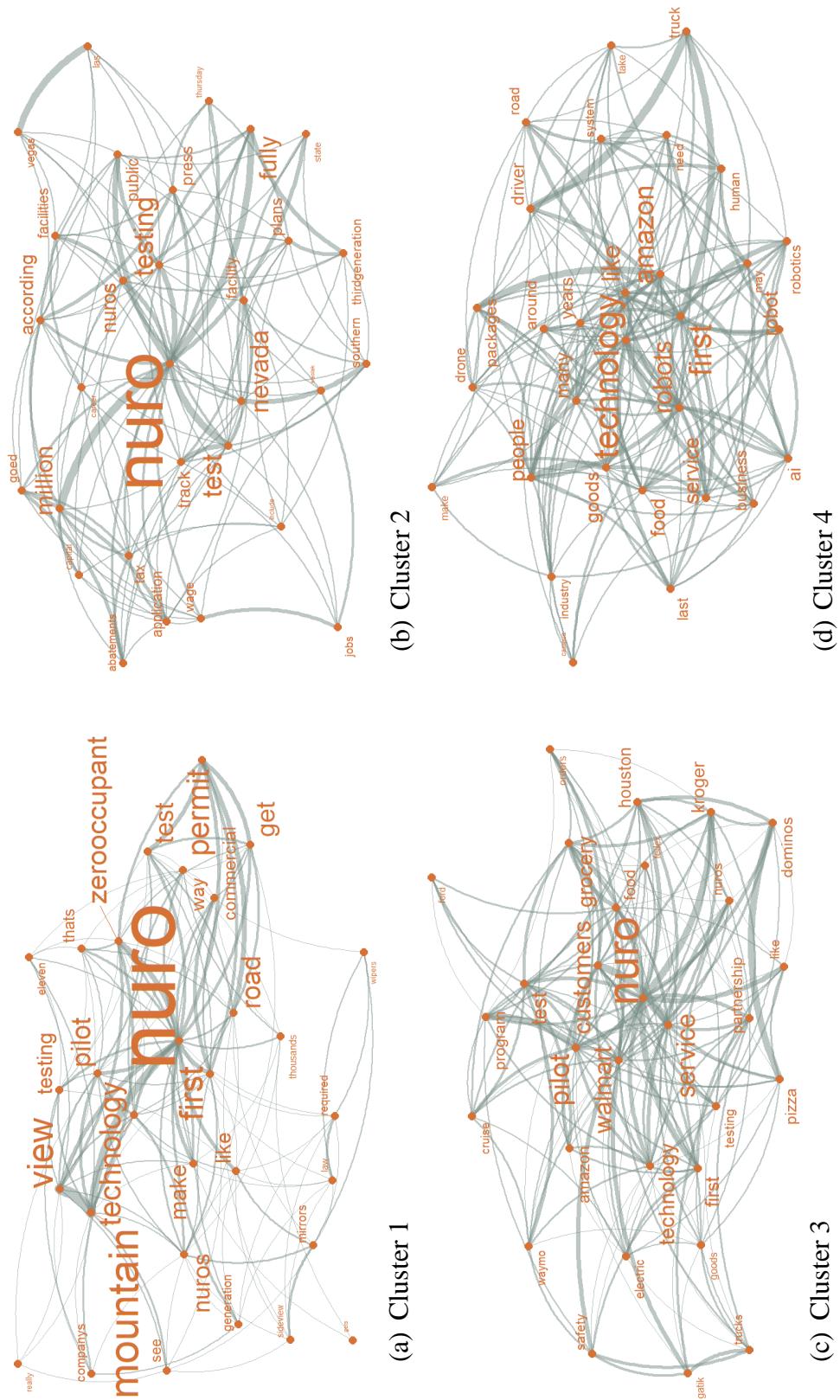


Figure 3. Text networks from the news reporting contents.

CONCLUSIONS

ADVs are gaining significant attention and are expected to revolutionize how goods are delivered. The analysis of news media indicated that there are some specific topics mainly discussed regarding ADVs, including express delivery service, Nuro and other ADV companies, robot delivery, and food delivery. The emerging companies developing ADVs, such as Nuro, Waymo, Udelv, Quadrobot, Einride, REV-1, and Kar-go, offer tremendous potential to revolutionize the delivery industry. The results show that express delivery services have begun cooperating with various ADV companies to complete the last mile, including DHL Express, FedEx, and the U.S. Postal Service. ADVs are increasingly being used in food, grocery, and other delivery services, with companies like Domino's Pizza, DoorDash, and Walmart partnering with ADV companies to offer same-day deliveries. The use of ADVs can result in reduced vehicle use, emissions, and traffic, owing to increased efficiency for vehicle use. Overall, ADVs are becoming an increasingly popular option for delivery services, providing a more efficient and convenient way to deliver goods while reducing costs and improving road safety.

This study's unique contributions lie in its innovative use of TNA to better understand ADVs' current trends, adoption, and impact as portrayed in news media reports. By employing TNA, the study presents valuable insights that can guide further research, policymaking, and industry developments in autonomous delivery vehicles. The study's policy implications highlight the need for clear regulations and safety standards for ADVs, incentives for sustainable ADV adoption, and public-private partnerships to facilitate responsible deployment. Policymakers should also address employment challenges, promote accessibility and equity, ensure data privacy and security, and invest in public awareness and education to support the successful integration of ADVs into transportation systems.

REFERENCES

- AASHTO. 2010. *Highway Safety Manual*.
- Barnier, J., and Privé, F. 2023. *rainette: The Reinert Method for Textual Data Clustering*.
- Berendt, B. 2017. Text Mining for News and Blogs Analysis. *Encyclopedia of Machine Learning and Data Mining* 1247–1255.
- Dorr, A., and Seba, T. 2020. RethinkX Energy Disruption Methodology. *Rethinking Energy 2020-2030: 100% Solar, Wind, and Batteries is Just the Beginning*.
- Flämig, H. 2016. Autonomous Vehicles and Autonomous Driving in Freight Transport. *Autonomous Driving* 365–385.
- Hawkins, A. 2017. Tesla's crash rate dropped 40 percent after Autopilot was installed, Feds say - The Verge [WWW Document]. URL <https://www.theverge.com/2017/1/19/14326258/teslas-crash-rate-dropped-40-percent-after-autopilot-was-installed-feds-say> (accessed 11.11.23).
- Hunter, S. 2014. A Novel Method of Network Text Analysis. *Open Journal of Modern Linguistics* 04, 350–366.
- Kapser, S., and Abdelrahman, M. 2020. Acceptance of Autonomous Delivery Vehicles for Last-Mile Delivery in Germany – Extending UTAUT2 with Risk Perceptions. *Transportation Research Part C: Emerging Technologies* 111, 210–225.
- Kapser, S., Abdelrahman, M., and Bernecker, T. 2021. Autonomous delivery vehicles to fight the spread of Covid-19 - How do men and women differ in their acceptance? *Transp Res Part A Policy Pract* 148, 183–198.

- Katrakazas, C., Quddus, M., and Chen, W.-H., 2019. A New Integrated Collision Risk Assessment Methodology for Autonomous Vehicles. *Accident Analysis & Prevention* 127, 61–79.
- Kutela, B., Das, S., and Dadashova, B. 2022. Mining patterns of autonomous vehicle crashes involving vulnerable road users to understand the associated factors. *Accident Analysis & Prevention* 165, 106473.
- Kutela, B., Langa, N., Mwende, S., Kidando, E., Kitali, A. E., and Bansal, P. 2021. A text mining approach to elicit public perception of bike-sharing systems. *Travel Behaviour and Society* 24, 113–123.
- Kwayu, K. M., Kwigizile, V., Lee, K., and Oh, J.-S. 2021. Discovering latent themes in traffic fatal crash narratives using text mining analytics and network topology. *Accident Analysis & Prevention* 150, 105899.
- Morando, M. M., Tian, Q., Truong, L. T., and Vu, H. L. 2018. Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *Journal of Advanced Transportation* 2018, 1–11.
- Mueller, A. S., Cicchino, J. B., and Zuby, D. S. 2020. What humanlike errors do autonomous vehicles need to avoid to maximize safety? *Journal of Safety Research* 75, 310–318.
- Paddeu, D., and Parkhurst, G. 2020. Chapter Twelve - The Potential for Automation to Transform Urban Deliveries: Drivers, Barriers and Policy Priorities, in: Milakis, D., Thomopoulos, N., van Wee, B. (Eds.), *Advances in Transport Policy and Planning, Policy Implications of Autonomous Vehicles*. Academic Press, pp. 291–314.
- Pani, A., Mishra, S., Golias, M., and Figliozzi, M. 2020. Evaluating Public Acceptance of Autonomous Delivery Robots during COVID-19 Pandemic. *Transportation Research Part D: Transport and Environment* 89, 102600.
- Papadoulis, A., Quddus, M., and Imprailou, M. 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. *Accident Analysis & Prevention* 124, 12–22.
- Paranyushkin, D. 2011. *Identifying the Pathways for Meaning Circulation using Text Network Analysis* 2.
- Rampin, R., and Rampin, V. 2021. Taguette: open-source qualitative data analysis. *Journal of Open Source Software* 6, 3522.
- Strauss, L., Tatar, K., and Nuro, S. 2021. instance: Soma-based multi-user interaction design for the telematic sonic arts. *Organised Sound* 26, 390–402.
- Witcher, C., Henry, S., McClafferty, J. A., Custer, K., Sullivan, K., Sudweeks, J. D., and Perez, M. A. 2021. Estimating Crash Consequences for Occupantless Automated Vehicles (Report). Virginia Tech Transportation Institute.
- Ye, L., and Yamamoto, T. 2019. Evaluating the impact of connected and autonomous vehicles on traffic safety. *Physica A: Statistical Mechanics and its Applications* 526, 121009.

Impact of Automated Driving under Virtual Traffic Signals on Traffic Resilience of Urban Road Networks

Yankai Zhang¹; Jiajing Chen²; Bo Liang³; Zihuan Xu⁴; Lei Mo⁵; Zhao Zhang⁶;
and Ruijie “Rebecca” Bian⁷

¹Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China.

Email: 21374148@buaa.edu.cn

²Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China.

Email: 21376373@buaa.edu.cn

³Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China.

Email: 21375271@buaa.edu.cn

⁴Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China.

Email: 21376222@buaa.edu.cn

⁵Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China.

Email: molei@buaa.edu.cn

⁶Professor, Key Laboratory of Intelligent Transportation Technology and System, Ministry of Education, and School of Transportation Science and Engineering, Beihang Univ., Beijing, China. Email: zhaozhang@buaa.edu.cn

⁷Louisiana Transportation Research Center, Dept. of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana. Email: rbian1@lsu.edu

ABSTRACT

With the rapid development of autonomous driving technology, it is imperative to study the impact of this technology on the resilience of urban road transportation networks. This study focuses on virtual traffic signal control (VTS) as a representative of autonomous driving technology and proposes a multi-dimensional resilience evaluation system to compare the impacts of three different signal-timing strategies on the resilience of urban traffic networks under various scenarios. Simulation results demonstrate that the average vehicle speed under VTS control is 38.1% and 40.1% higher than the other two control modes, respectively. In addition, the VTS approach significantly reduces delays and minimizes fuel consumption compared to the other two approaches. Moreover, this paper delves into the intrinsic mechanisms by which VTS enhances traffic resilience. It highlights the importance of integrating autonomous driving technology into traffic strategies to achieve carbon emissions reduction and promote sustainable transportation. The network resilience enhanced by VTS enables cities to effectively respond to unexpected events, thereby ensuring smooth traffic flow and safety. This research provides valuable insights into existing knowledge and promotes further research in the field of traffic resilience.

Keywords: Autonomous driving technology, Resilience, Fixed-time control, Actuated control, Virtual traffic signal control

INTRODUCTION

In the last decade, resilience in transportation systems, emphasizing their capacity to withstand and recover from disruptions, has gained significant attention. Traffic signalization is crucial in this regard, especially in urban road networks, for managing disturbances and maintaining traffic flow continuity. However, the impact of traffic control methods on network resilience remains under-explored.

The emergence of Connected and Autonomous Vehicles (CAV) introduces new dynamics. CAVs offer enhanced safety, efficiency, reduced congestion, and mobility for non-drivers. (Al-Fuqaha, F. et al. 2018) Their ability for infrastructure communication is vital, especially in emergencies, aiding in traffic optimization and contributing to network resilience in high-demand and unexpected scenarios.

VTS control (Z. Zhang et al. 2020), a post-CAV innovation, utilizes CAV capabilities to optimize traffic flow and improve resilience. This paper introduces a Multi-dimensional Resilience Evaluation System to assess autonomous driving effects under VTS control, aiming to understand CAV integration's impact on road network resilience. It covers an overview of transportation resilience and CAV concepts, problem identification, methodological framework, and simulation scenarios for model validation, and concludes with research outcomes and future directions.

LITERATURE REVIEW

Recent urban traffic resilience research has delved into diverse aspects such as congestion patterns, road network resilience distribution, and evaluation methodologies. Limiao Zhang's (Zhang, L. et al. 2019) research highlighted the spatiotemporal clusters of congestion. Xiaoyong Ni (Ni, X. et al. 2021) proposed a method focusing on a traffic flow assignment model to assess urban road traffic resilience. Maria Nogal (Nogal, M et al. 2016) introduced a methodology for employing a dynamic equilibrium-restricted assignment model to evaluate network performance under stress. These studies offered insights into the complex resilience behavior of urban traffic systems.

In urban road traffic optimization, Wang Tao's (Tao, W. et al. 2022) research introduced a two-stage accident resilience triangle framework and a traffic signal optimization model. Simeon C. (Calvert et al. 2018) developed the Link Performance Index for Resilience (LPIR) to measure the resilience levels of individual road sections. Wei Meng (Wei, M. et al. 2022) assessed the resilience of road networks post-Wenchuan earthquake, emphasizing the importance of resilience in emergency scenarios. Gopal R. Patil's (Patil, G. R. et al. 2016) computational experiments explored the relationship between network resilience and traffic demand. However, most studies have not fully explored the role of connected and autonomous vehicles (CAVs) in enhancing road network resilience.

Regarding CAVs, Lanhang Ye (Ye, L. et al. 2019) investigated CAVs' impact on traffic safety, finding improvements with higher CAV penetration. Alireza Talebpour (Talebpour, A. et al. 2016) studied the effects of CAVs on highway driving, showing that higher CAV penetration might increase throughput and prevent shockwave formation. Montanaro (Montanaro et al. 2019) analyzed CAVs' potential benefits in various traffic aspects. Zongyuan Wu (Wu, Z. et al. 2021) focused on enhancing intersection management efficiency through CAVs, and Shofiq Ahmed (Ahmed, S. 2018) developed a resilience model for mixed-traffic environments, assessing traffic

system resilience across different CAV penetration scenarios during large-scale natural disasters. In addition, due to the high reliance on the VTS control (Z. Zhang et al. 2020) mode on CAV for its implementation, and the ability of the VTS control mode to demonstrate most of the characteristics of CAV, analyzing the impact of VTS control on traffic resilience enables us to infer, to some extent, the influence of CAV on traffic resilience.

These studies underscore the potential of CAVs in improving traffic flow stability and throughput but leave the impact of CAVs on traffic resilience and the role of traffic control largely unexplored. Therefore, the impact of CAVs on traffic resilience needs more comprehensive studies.

This study set Virtual Traffic Signal Control (VTS) as a representative of autonomous driving technology and analyzes the impact of autonomous driving technology on traffic resilience under three control methods: Fixed-time control, actuated control, and VTS control, using five parameters to measure road network resilience.

METHODS

In this paper, increased demand and decreased supply are regarded as special events. Five parameters are used to build a Multi-dimensional Resilience Evaluation System (MDRES): namely average speed before a special event (V), the difference between the average speed before special cases and the point of lowest average speed (H), the time difference from the start of special cases to the complete recovery of the road network (L), the area of the resilience triangle (S), and average fuel consumption per vehicle (C).

This research considers the abnormal conditions that occur in the road network as different levels of perturbations to the network. The response of the road network after the perturbation reflects the network's resistance and recovery, which is regarded as the network's resilience. If the network's resistance is high, the reduction in the network's capacity after the perturbation is minimal. If the network's recovery is high, it takes less time to restore the network's initial capacity after being perturbed.

Therefore, this research quantifies the road network's resistance by measuring the degree of decline in its traffic capacity, and its recovery by measuring the time it takes for the network to return to its original state after the decline.

As illustrated in Figure 1, the road network's passage ability decreases near the beginning of the perturbation and gradually recovers as time progresses. This research denotes the road network's resistance as H and its recovery as L.

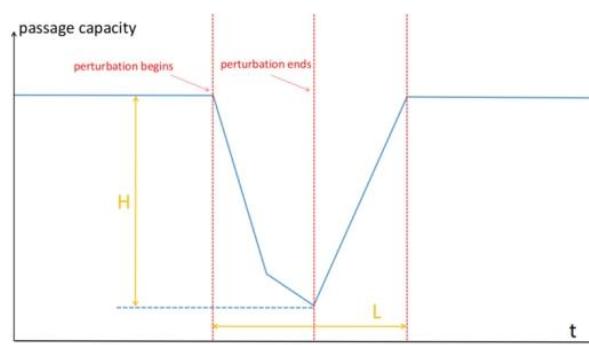


Figure 1. Resilience curve

The remaining three indicators can be obtained directly from the simulated data. Since the five indicators have different units and different scales of measurement, we normalize the data by using the sigmoid function. Additionally, the MK trend test method is utilized to find special points. In the following section, experiments under various scenarios are conducted and the results are analyzed to evaluate the resilience of the road network.

CASE STUDY

In this section, we first present the entire transportation road network structure, followed by two separate scenarios from the simulation experiments: increased demand and decreased supply. As shown in Figure 2, a 3x3 grid network with a traffic signal control system at each intersection is created for simulation purposes.

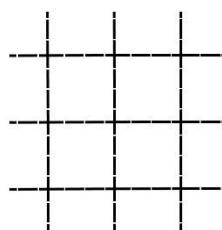


Figure 2. Road network

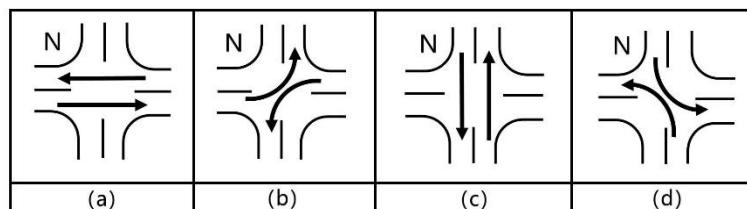


Figure 3. The phase diagram of fixed-time control

Fixed-time control is a conventional approach that utilizes predetermined signal timings to regulate traffic flow. Its phase diagram is shown in Figure 3. The parameters associated with fixed-time control are presented in Table 1.

Table 1. Signal control scheme

Phase	Traffic flow	Green time	Yellow time	Signal cycle
1	East-West straight	15	3	72
2	East-West left-turn	15	3	
3	North-South straight	15	3	
4	North-South left-turn	15	3	

Actuated control is a traffic management strategy that exhibits strong adaptability by utilizing real-time traffic flow data to continuously adjust traffic signal duration. In this study, a maximum

green time of 30s, a minimum green time of 6s, and an extension time of 3s were proposed for the actuated control condition.

Virtual traffic signal control is a novel method that optimizes signal timings by utilizing communication between vehicles and traffic signals. In intersections, by autonomously coordinating and synchronizing the arrival and departure of vehicles under one operating rule, VTS can ensure continuous vehicle movement with guaranteed safe spacing, effectively eliminating the need for traffic signals. For more details on this approach, please refer to the previous work by Zhang et al. (Z. Zhang et al. 2020). The driving mode used in this article is the Cooperative Adaptive Cruise Control (CACC) model. (Milanés et al. 2014)

In this paper, the simulation of the traffic road network under two scenarios of increased demand and decreased supply was conducted on the SUMO simulation platform. In each scenario, this research extracted the speed of each road segment and visualized the data after performing Singular Spectrum Analysis (SSA) filtering analysis.

1. Increased Demand:

To simulate the situation of increased demand, simulation experiments are conducted to set up the traffic flow on this road network in sumo by time period. Table 2 shows Traffic flow per lane in different durations.

Table 2. Traffic flow per lane in different durations.

Time	Traffic demand
0~1800s	200veh/h/lane
1800~3600s	500veh/h/lane
3600~7200s	200veh/h/lane

As shown in Figure 4, the average speed change curves for the three control modes show similar trends. The average speeds for fixed-time control, actuated control, and VTS control were 11.8m/s, 12.4m/s, and 15.1m/s, respectively. However, when demand increases, all three control modes experience a drop in average speed, with VTS showing a smaller decrease compared to the other modes. The recovery times for average speed also differ significantly, with fixed-time control, actuated control, and VTS control recovering at around 8000s, 6000s, and 4000s, respectively.

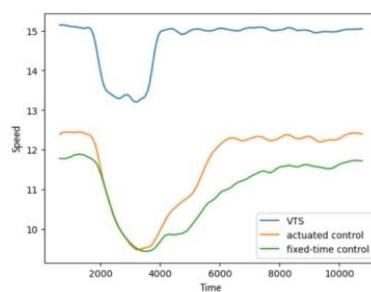


Figure 4. Average vehicle speed for the three control modes

We use the Mann-Kendall trend test to detect changes to create an elastic triangle region for the three control modes, as demonstrated in Figure 5.

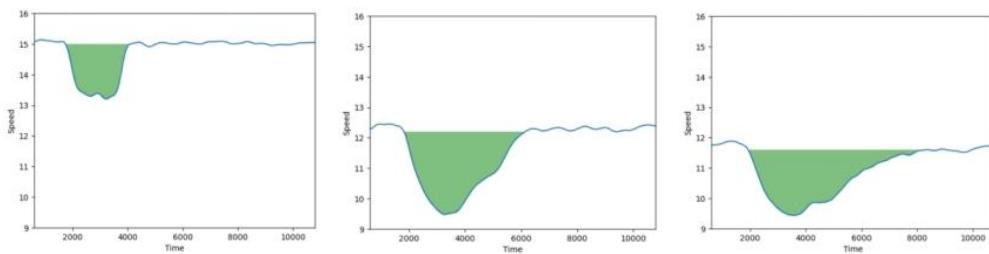


Figure 5. Resilience triangle areas in the case of increased demand

Table 3 shows the results with a confidence level of 0.05 by combining the data results of Figure 5 with the MK trend test function.

Table 3. The parameters in the three control modes

	V	H	L	S	C
VTS	15.1	2.0	2280	2987.6	152941.1
actuated control	12.4	3.1	4380	7118.7	208683.0
fixed-time control	11.8	2.5	6480	6929.2	258802.9

2. Decreased Supply:

In the decreased supply scenario, one road segment is assumed to be damaged between intersections 4 and 5. The damaged road segment is shown in Figure 6.

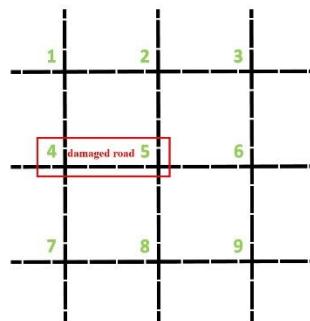


Figure 6. The road network of decreased supply

With a traffic flow of 300veh/h, this research set a maximum limit for a section of the road network to simulate decreased supply, as shown in Table 4.

As shown in Figure 7, the average speed change curves of the three modes exhibit similar trends. VTS has a higher average speed of around 14.5 m/s compared to the fixed time control and actuated control at approximately 11.5 m/s and 11.6 m/s respectively when there is no supply reduction. When supply decreased in the 1800s, the average speeds of all three modes

were affected and decreased. VTS shows a significantly smaller decrease in average speed compared to the other modes. The recovery time for average speed also differs among the modes, with fixed time control, actuated control, and VTS control taking around 7560s, 7620s, and 5580s to recover, respectively.

Table 4. Maximum speed limit in different durations.

Time	Maximum speed limit
0~1800s	15m/s
1800~3600s	5m/s
3600~7200s	15m/s

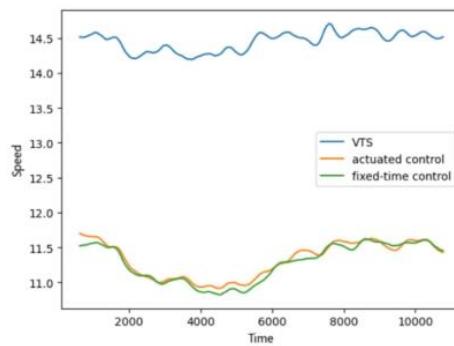


Figure 7. The average speed of decreased supply

Using the MK trend test method, the curve of the resilience triangle area for the three control modes can be plotted in Figure 8. By examining Figure 8 and using the MK trend test function, the results with a confidence level of 0.05 can be presented in Table 5.

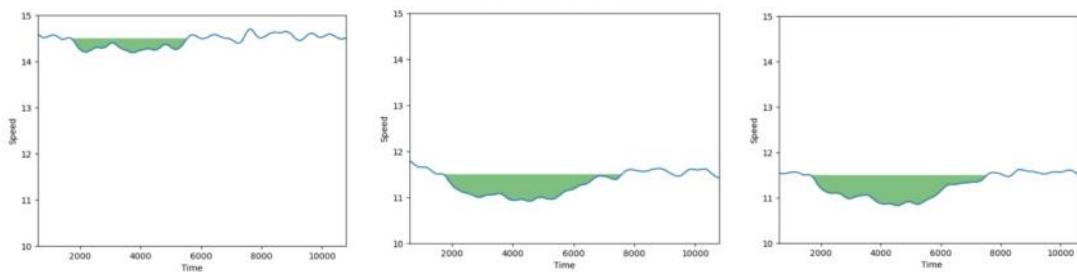


Figure 8. Resilience triangle areas of decreased supply (whole road network)

Table 5. The parameters in the three control modes

	V	H	L	S	C
VTS	14.5	0.5	3780	790.3	140281.6
actuated control	11.6	0.8	5820	2119.0	202293.9
fixed-time control	11.5	0.9	5760	2433.7	207866.3

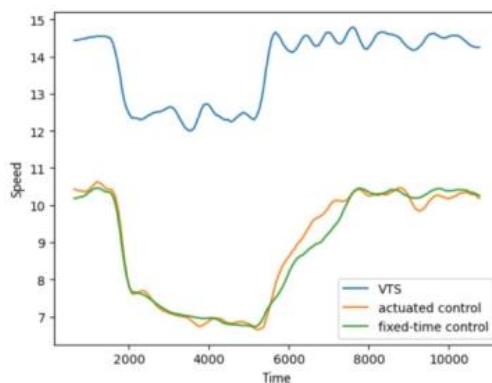


Figure 9. Average speed of decreased supply on speed limit section

As shown in Figure 9, VTS has a higher average speed of around 14.5 m/s compared to the actuated control and fixed-time control modes at approximately 10.5 m/s and 10.3 m/s respectively when there is no supply reduction. When supply decreased in the 1800s, VTS experienced a smaller decrease in average speed compared to the other modes. The recovery time for average speed also differs among the modes, with VTS recovering around 5640s, actuated control around 7680s, and fixed-time control around 7620s.

Similarly, the results in the case of decreased supply only in the speed limit road are shown in Figure 10, Table 6.

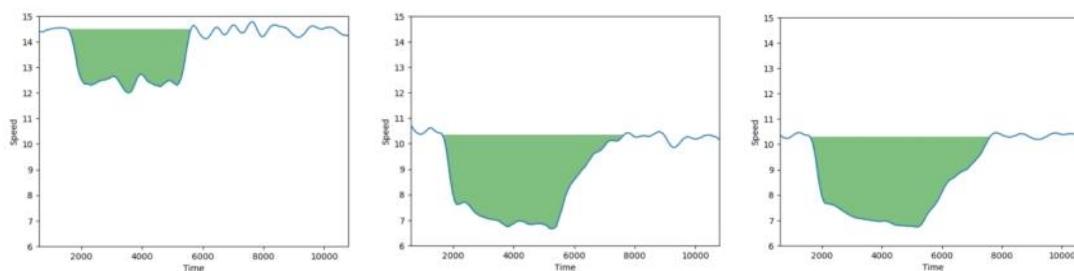


Figure 10. Resilience triangle areas of decreased supply (the speed limit road)

Table 6. The parameters in the three control modes

	V	H	L	S	C
VTS	14.5	3.2	3840	7613.1	140326.1
actuated control	10.5	4.0	5880	14623.5	202571.7
fixed-time control	10.3	3.8	5820	15255.1	208091.4

RESULT ANALYSIS

In this section, we first analyze the simulation results, then compare the three signal control methods and analyze the possible reasons for the results, and finally present the impact of using the VTS control method.

1. Analysis of simulation results:

The radar chart of the multi-dimensional resilience assessment results in the case of increased demand is shown in Figure 11.

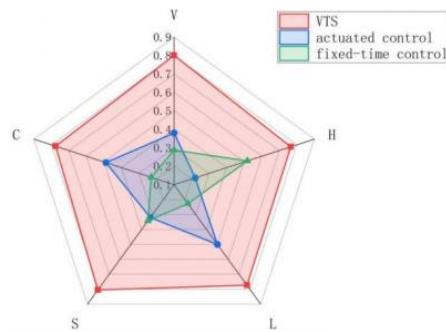


Figure 11. Resilience assessment in the case of increased demand

Based on Figure 11, we can find that the pentagon area under VTS control is larger than the other modes, indicating better overall performance.

Regarding resilience, VTS control shows a smaller average speed drop before and after demand increase compared to the other modes, with a decrease of 35.5% and 20.0% respectively. The traffic flow speed under VTS control remains relatively stable during the elastic recovery process, indicating better resistance to demand increase. VTS also has an earlier time point for demand increase recovery, shortened by 47.9% and 64.8% compared to the other modes. The resilience triangle area under VTS control is 2987.6m², reducing by 58.0% and 56.9% compared to the other modes. In terms of efficiency, VTS control results in a higher average speed, with a 17.9% and 21.9% increase compared to the other modes.

Regarding fuel consumption, VTS control leads to lower average fuel consumption, with a reduction of 26.7% and 40.9% compared to the other modes.

Figure 12 shows the multi-dimensional resilience assessment results in a decreased supply of the whole road network (left) and speed limit road (right).

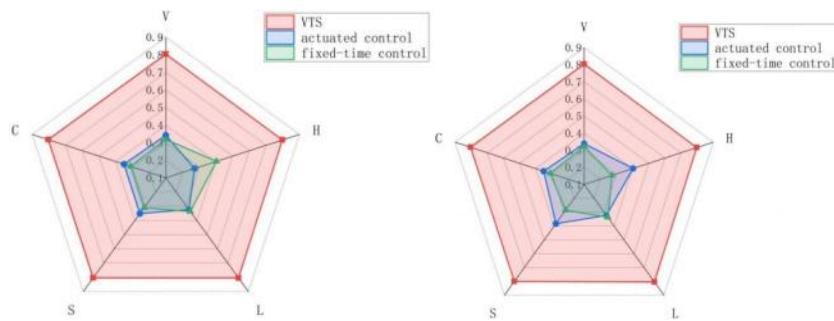


Figure 12. Resilience assessment in the case of decreased supply

In terms of resilience, VTS control shows better resistance to demand changes compared to the other two control modes. The decrease in average vehicle speed under VTS control is 37.5% and 44.4% compared to the other modes. VTS also has an earlier time point for supply reduction

and recovery, shortening by 62.7% and 67.5% respectively. The resilience triangle area under VTS control is 790.3m, reducing by 62.7% and 67.5% compared to the other modes.

Similar trends are observed in damaged sections, with a 20% decrease in average vehicle speed under VTS control compared to 15.8% and 20% under the other modes. The time point for supply reduction and recovery is earlier, shortening by 34.7% and 34.0% respectively. The resilience triangle area under VTS control is 7613.1m, reducing by 47.9% and 50.1% compared to the other modes.

In terms of efficiency, VTS control results in higher average vehicle speed, with a 25.0% and 26.1% increase compared to the other modes for the entire network. In damaged sections, the increase is 38.1% and 40.1% respectively. The average vehicle speed under VTS control decreases little, while it decreases by 9.5% and 10.4% under the other modes.

Regarding fuel consumption, VTS control leads to lower average fuel consumption, with a reduction of approximately 30.7% and 32.5% compared to the other modes. This is because VTS control reduces fuel consumption by minimizing unnecessary engine start and stop, making it more environmentally friendly.

2. Comparative analysis of control models:

Fixed-time control provides traffic efficiency under stable conditions but lacks the flexibility to adapt to changing traffic demands, leading to congestion and delays.

Actuated control adjusts signal durations based on real-time data. Still, it has limited response time and can be affected by malfunctioning or interference, limiting its ability to improve traffic resilience, it falls short compared to VTS's flexibility, accuracy, wide coverage, and utilization of advanced scheduling algorithms for comprehensive optimization. VTS optimizes vehicle trajectories and signal timings in real-time, eliminating the impact of signals and stop signs, and enhancing traffic resilience.

Compared to fixed-time control and actuated control, VTS exhibits advantages in terms of average speed, resistance, recovery, the area of the resilience triangle, and average fuel consumption per vehicle, displaying superior resistance and recovery capabilities in scenarios of increased demand and reduced supply. Vehicles under VTS control experienced lower delays and higher speeds, minimizing unnecessary stops, reducing fuel consumption, enhancing responsiveness to unexpected events and demand, and promoting eco-friendliness.

3. Impacts of the findings:

For short-term impacts, the findings of this article spark interest in transportation engineering, encouraging further research in the academic community. In addition, it provides insights for policy and decision-makers on immediate traffic management decisions and autonomous driving integration.

For long-term impacts, authorities may consider incorporating VTS control systems into their future road network designs during infrastructure planning to enhance resilience, reduce congestion, and improve overall traffic flow. Automakers and technology companies may prioritize the integration of VTS capabilities into autonomous vehicles as they develop autonomous driving technologies, recognizing the advantages it offers in terms of traffic efficiency, resilience, and safety.

For broader impacts, the findings of this article inspire the idea that we can reduce fuel consumption by reducing unnecessary vehicle starts and stops, aligning with broader sustainability goals and initiatives aimed at reducing carbon emissions and promoting sustainable mobility in the traffic system. In addition, this also enhances urban infrastructure resilience during emergencies or peak periods.

CONCLUSION

Given the limited research on the connection between autonomous driving technology and traffic resilience, this paper introduces a multidimensional evaluation system to assess the impact of autonomous driving technology on traffic resilience. Simulations are conducted using three signal control methods to represent various traffic scenarios. Simulation results reveal VTS's superiority in improving traffic resilience of urban road networks by comparing the other two control methods. The findings are significant for academia, decision-makers, infrastructure planning, and the development of autonomous driving technology, driving the future transportation systems' resilience, efficiency, and sustainability.

DECLARATIONS

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 52232011. The authors would also like to thank the insightful and constructive comments from anonymous reviewers.

REFERENCES

- Ahmed, S. (2018). Resilience modeling of the surface transportation system in the mixed traffic environment. West Virginia University. [URL <https://researchrepository.wvu.edu/d/6155/>].
- Al-Fuqaha, F., Guizani, M., Mohammadi, M., Aledhari, M., and Ayyash, M. (2018). Autonomous Vehicles: A Survey on Technologies, Applications, and Challenges. *IEEE Communications Surveys & Tutorials*, 20(1), 423-452.
- Calvert, S. C., and Snelder, M. (2018). A methodology for road traffic resilience analysis and review of related concepts. *Transportmetrica A: transport science*, 14(1-2), 130-154.
- Milanés, V., and Shladover, S. E. (2014). "Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data." *Transportation Research Part C: Emerging Technologies* 48:285-300.
- Montanaro, U., Dixit, S., Fallah, S., Dianati, M., Stevens, A., Oxtoby, D., and Mouzakitis, A. (2019). Towards connected autonomous driving: a review of use-cases. *Vehicle System Dynamics*, 57(6), 779-814.
- Ni, X., Osaragi, T., Huang, H., Li, R., and Chen, A. (2021). Resilience-oriented performance assessment method for the road-traffic system: a case study in Beijing, China. *KSCE Journal of Civil Engineering*, 25(10), 3977-3994.
- Nogal, M., O'Connor, A., Caulfield, B., and Martinez-Pastor, B. (2016). The resilience of traffic networks: From perturbation to recovery via a dynamic restricted equilibrium model. *Reliability Engineering & System Safety*, 156, 84-96.

- Patil, G. R., and Bhavathrathan, B. K. (2016). Effect of traffic demand variation on network resilience. *Advances in complex systems*, 19(01n02), 1650003.
- Talebpour, A., and Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71, 143-163.
- Tao, W., Wang, Z. Z., Liu, C. J., Lu, Y. N., Zhang, Y., and Jiang, Z. H. (2022). Resilience assessment and enhancement of urban road networks subject to traffic accidents: a network-scale optimization strategy. *Journal of Intelligent Transportation Systems*, 1-17.
- Wei, M., Fang, S. J., Chen, S., Huang, Y., and Yang, L. (2022). Resilience Assessment of Road Networks in the Extremely Severe Disaster Areas of the Wenchuan Earthquake. *Frontiers in Earth Science*, 10, 834302.
- Wu, Z., and Waterson, B. (2021). Urban Intersection Management Strategies for Autonomous/Connected/Conventional Vehicle Fleet Mixtures. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12084-12093.
- Ye, L., and Yamamoto, T. (2019). Evaluating the impact of connected and autonomous vehicles on traffic safety. *Physica A: Statistical Mechanics and its Applications*, 526, 121009.
- Zhang, L., Zeng, G., Li, D., Huang, H. J., Stanley, H. E., and Havlin, S. (2019). Scale-free resilience of real traffic jams. *Proceedings of the National Academy of Sciences*, 116(18), 8673-8678.
- Zhang, Z., Liu, F., Wolshon, B., and Sheng, Y. (2020). "Virtual Traffic Signals: Safe, Rapid, Efficient and Autonomous Driving Without Traffic Control," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6954-6966, Nov. 2021.

Assessing Cybersecurity Risks and Traffic Impact in Connected Autonomous Vehicles

Saurav Silwal¹; Lu Gao, Ph.D.²; Yunpeng Zhang, Ph.D.³; Ahmed Senouci, Ph.D.⁴;
and Yi-Lung Mo, Ph.D., P.E.⁵

¹Dept. of Construction Management and Dept. of Civil and Environmental Engineering, Univ. of Houston, Houston, TX. Email: ssilwal@cougarnet.uh.edu

²Dept. of Construction Management, Univ. of Houston, Houston, TX.
Email: lgao5@central.uh.edu

³Dept. of Information Science Technology, Univ. of Houston, Sugar Land, TX.
Email: yzhang119@uh.edu

⁴Dept. of Construction Management, Univ. of Houston, Houston, TX. Email: asenouci@uh.edu

⁵Dept. of Civil and Environmental Engineering, Univ. of Houston, Houston, TX.
Email: yilungmo@central.uh.edu

ABSTRACT

Given the promising future of autonomous vehicles, it is foreseeable that self-driving cars will soon emerge as the predominant mode of transportation. While autonomous vehicles offer enhanced efficiency, they remain vulnerable to external attacks. In this research, we sought to investigate the potential impact of cyberattacks on traffic patterns. To achieve this, we conducted simulations where cyberattacks were simulated on connected vehicles by disseminating false information to either a single vehicle or vehicle platoons. The primary objective of this research is to assess the cybersecurity challenges confronting connected and automated vehicles and propose practical solutions to minimize the adverse effects of malicious external information. In the simulation, we have implemented an innovative car-following model for the simulation of connected self-driving vehicles. This model continually monitors data received from preceding vehicles and optimizes various actions, such as acceleration, and deceleration, with the aim of maximizing overall traffic efficiency and safety.

Keywords: Autonomous Vehicle; Connected Vehicle; Car Following Model; Cyberattacks; Traffic flow; Simulation.

INTRODUCTION

Technological advancements consistently bring forth new innovations across diverse fields. In recent years, researchers have made significant strides by introducing semi-autonomous vehicles, fulfilling humanity's longstanding aspiration for autonomous travel. Moreover, the escalating issues of traffic congestion and accidents highlight a pressing need for innovative remedies, propelling this sector toward further evolution. A compelling aspect of autonomous vehicles is Connected Autonomous Vehicles (CAVs), where vehicles interact on the road, exchanging critical data to enable autonomous driving. Bajpai (2016) explores the promising effects of CAVs, demonstrating their potential to mitigate congestion, enhance safety, and reduce fuel usage.

CAVs employ adaptive cruise control (ACC) and its advanced iteration, connected adaptive cruise control (CACC). ACC is a driving assistant and functions using the vehicle's sensors to automatically adjust its speed relative to the preceding vehicle, whereas CACC leverages

vehicle-to-vehicle(V2V) communication, sharing data like the vehicle's speed, location, and activity, facilitating a more synchronized and seamless driving experience (Van Arem et.al 2006). Over the years, researchers have examined how implementing ACC or CACC in vehicles enhances traffic dynamics. VanderWerf et.al. (2002) employed Monte Carlo simulations to conduct a quantitative assessment of the effectiveness of ACC and concluded that ACC has the potential to enhance highway traffic throughput, particularly when CACC is integrated into vehicles. Similarly, Ploeg et.al. (2011) conducted experiments involving six CACC-equipped vehicles, affirming that CACC improves traffic flow.

CAVs present numerous promising benefits for the future of transportation, yet they also bring forth significant vulnerabilities. Among these vulnerabilities, one of the primary concerns is the susceptibility to cyber-attacks. The inclusion of wireless communication between vehicles exposes them to various types of attacks that could potentially cause harm or damage. Given the pervasive use of wireless technology in our daily lives, we are aware of the damaging potential of cyber threats. It is crucial to recognize that vehicles equipped with wireless communication are susceptible to attacks from individuals with malicious intentions. Addressing such threats requires a proactive approach: first identifying potential attack vectors and understanding their possible impacts. Armed with this knowledge, efforts can be directed toward developing defenses to safeguard these vehicles against potential cyber-attacks. Yağdereli et.al. (2015) explored the evolving landscape of autonomous transportation modes, detailing diverse categories of cyber threats: passive and active attacks. They offered potential mitigation strategies while emphasizing the pressing need for an update to the Control Area Network (CAN) standard, which currently stands as outdated in the light of emerging cyber threats. Similarly, Parkinson et.al. (2017) conducted a comprehensive analysis of existing literature in this domain, revealing a significant gap in knowledge within the field and the need for extensive research endeavors to fill this void and fortify cybersecurity measures.

Numerous researchers have proposed diverse approaches to comprehensively grasp the cyber threats faced. They have delineated different types of cyber-attacks, ranging from sending false messages, and replaying stored messages, to manipulating acceleration or deceleration, among others. By employing different car-following models, researchers aimed to simulate traffic flow scenarios and scrutinize vehicle behavior when subjected to these cyber-attacks, whether it's an individual vehicle or multiple vehicles affected. Wang et.al. (2018), for instance, utilized an extended optimal car-following model. Their study focused on analyzing the impacts of cyber-attacks that altered spacing and velocity, evaluating how these disruptions impacted traffic flow stability. Similarly, Sun et.al. (2023) conducted microsimulation modeling to examine time-delay and disturbance attacks at freeways and unsignalized intersections, uncovering the risk posed by these attacks, leading to collisions, congestion, and reduced road capacity. Furthermore, Khattak et.al. (2021) modeled multiple CAV platoons and simulated lane-changing dynamics. They found that cyber-attacks significantly jeopardized platoon stability and safety, with lane-changing conflicts posing greater challenges than rear-end collisions. In an extensive study examining the impact of cyber-attacks on CAV platoon stability, Masoud et.al. (2023) incorporated sensors along with an augmented state extended Kalman filter to detect and filter anomalous readings. They conducted sensitivity analysis and obtained a critical detection sensitivity, which ensures the effect of detected faulty reading are not amplified as information is relayed to subsequent vehicles, thereby maintaining platoon stability.

The journey toward achieving complete autonomy in vehicles is still in its early stages, and conducting extensive research into the potential weakness of autonomous vehicles remains

crucial. Current research on the cybersecurity threats to CAVs is somewhat limited. This paper seeks to add to this body of knowledge. Our study involves the simulation of a car-following model where we have introduced attacks such as injecting false messages, replaying stored data, and manipulating acceleration to assess their impact on the traffic flow.

METHODOLOGY

This section presents the car following model utilized for simulating traffic flow. Following the establishment of this traffic model, cyber-attacks were instigated to observe their impact on the traffic flow.

Car Following Model

The car-following model is the mathematical representation of the behavior of vehicles in a traffic stream. Pipes (1953) was the pioneer in introducing a car-following model to describe the characteristics of vehicles moving in a stream. The model described how a vehicle reacts to the changes in the vehicle in front of it. The development of a car-following model as a mathematical representation has made it possible to conduct stability analysis and made it possible to gauge the safety and traffic disruptions and propel the study toward autonomy (Rothery 1992). At its core, the basic car-following model operates on the principle that the vehicle's acceleration, i.e., the response is determined by the stimulus, which is the velocity difference between the vehicle and its leader (Chandler et.al. 1958). Over time researchers have progressively refined the car-following model, striving for greater realism to mirror real-life scenarios in traffic. Treiber et.al (2000) proposed a new model called the intelligent driver model (IDM) that is an improvement on the classical models. This new model demonstrates the capacity to portray realistic vehicle behavior within single-lane traffic scenarios by incorporating relevant parameters.

Intelligent Driver Model

The IDM computes the subject vehicle's acceleration utilizing data from the lead vehicle, as expressed in the following equation:

$$\dot{v}_n(t) = a \left(1 - \left(\frac{v_n(t)}{v_0} \right)^{\delta} - \left(\frac{s * (v_n(t), \Delta v_n(t))}{s_n(t)} \right)^2 \right) \quad (1)$$

$$s * (v_n(t), \Delta v_n(t)) = s_0 + T v_n(t) + \frac{v_n(t) * \Delta v_n(t)}{2\sqrt{ab}} \quad (2)$$

where, $s_n(t) = x_{n-1} - x_n - l$, where $n - 1$ represents the preceding vehicle, n represents the subject vehicle, l is the length of the vehicle, and $\Delta v_n(t) = v_n(t) - v_{n-1}(t)$.

Here, v_0, s_0, a, b and T are model parameters.

v_0 = desired velocity, the velocity of the vehicle in free traffic

s_0 = minimum spacing between the vehicles

a = maximum acceleration

b = maximum comfortable deceleration

T = minimum possible time to vehicle in front

δ = acceleration exponent

The IDM has been widely employed in various studies to simulate longitudinal vehicle dynamics. Despite its effectiveness as a model, the classical IDM's limitation lies in its focus solely on the influence of the immediately preceding vehicle. Previous studies have generalized the IDM to accommodate connected vehicle settings. For example, Wang et.al (2020) introduced a new approach within the classical IDM by integrating weighing coefficients for spacing and velocity differences, thus accommodating the impact of multiple leading vehicles. In this method, the acceleration of the target vehicle is influenced by the gap and speed difference relative to the vehicles in front of it. Although using weights to determine the degree of influence of preceding vehicles is good, this approach fails to represent a true-to-life situation. According to this model, if the nearest leading vehicle is abruptly halted, the subject vehicle will decelerate but, due to information from other vehicles ahead, it will not come to a complete stop, eventually leading to a collision. Thus, in this study, we developed a novel approach to extend the IDM to allow the subject vehicle to gather velocity and position data from preceding vehicles. Leveraging this information, the subject vehicle utilizes IDM to calculate acceleration with respect to each of these leading vehicles and select the one with the safest consequence. Furthermore, for simplicity, in this study reaction time was not included.

$$\dot{v}_n(t) = \min(\dot{v}_n^{n-1}(t), \dot{v}_n^{n-2}(t), \dots, \dot{v}_n^{n-m}(t)) \quad (3)$$

where,

$\dot{v}_n^j(t)$ is the acceleration calculated using Equation (1) and (2) by assuming vehicle j is the vehicle right in front of vehicle n . Figure 1 shows a simplified representation of connected automated vehicles setting considered in this paper. On a single-lane road, multiple autonomous vehicles are in motion. These vehicles have established a wireless network to share information about their speed and position, assisting each vehicle in deciding its next move. This study focuses exclusively on the scenario where vehicles communicate only with those ahead of it.

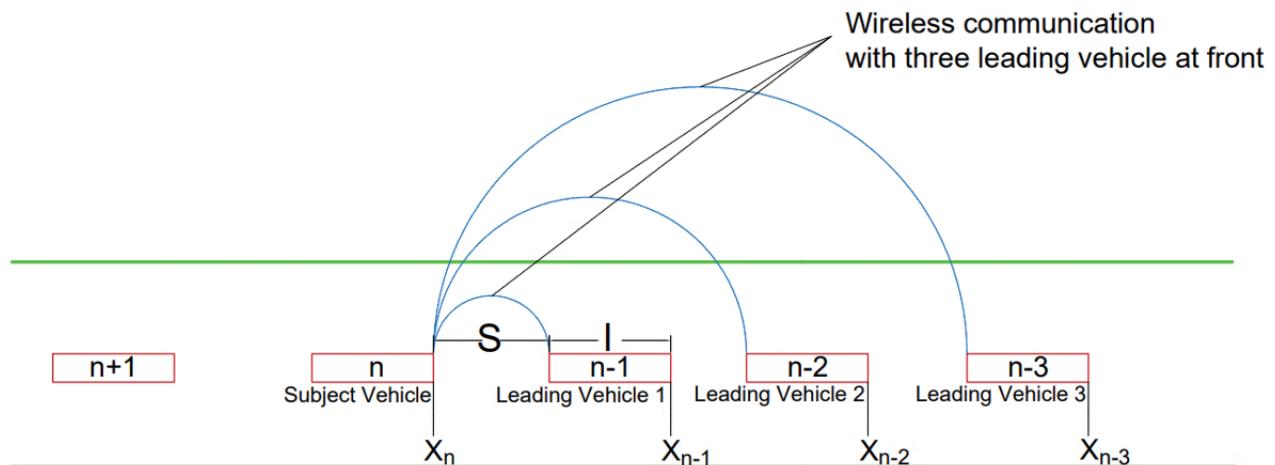


Figure 1. Typical Representation of CAV

Cyberattacks

CACC-equipped vehicles engage in wireless communication among themselves. However, in the event of a compromise or cyber-attack on this communication network, vehicles still rely on in-vehicle sensors to gather essential information. Our study aims to investigate the ramifications of cyber-attacks by exploring a scenario where both wireless communication and sensors are compromised. This worst-case scenario leaves the vehicle highly susceptible to malicious attacks, enabling us to evaluate the severe impact such vulnerabilities might cause. The cyber-attacks that have been implemented in this study are as follows.

- i. False Message: The attacked vehicle is fed erroneous information regarding the leader vehicle's position and velocity.
- ii. Platoon Leader Identity Attack: This cyberattack targets a vehicle, causing it to no longer recognize any vehicles positioned ahead of it, consequently perceiving itself as the leading vehicle within the platoon.
- iii. Acceleration Manipulation: This attack involves replacing the acceleration value calculated by Equation (3) with an alternative value.
- iv. Multiple Vehicle Attack: Simultaneously targeting multiple vehicles by providing them with misleading or false information.

These attacks cause disturbances in traffic flow, as demonstrated in the subsequent results section. Rear-end collisions or travel delays have been identified as key indicators of system failure.

Simulation Setup

Python programming language was used to create a simulation platform. A total of nine vehicles were defined to set up a vehicular platoon on a single-lane road. Overtaking or lane-changing behaviors were not considered for this study. Initially, specific positions and velocities were assigned to each vehicle. However, during simulation, the vehicles adapted their characteristics following Equation (3). Table 1 outlines the parameters we employed and their respective values for the study.

Table 1. Simulation Parameters

Parameters	Values
Number of lanes	1
Length of road segment	1000 m
Desired speed (v_0)	10 m/s
Length of vehicles (l)	5 m
Minimum gap (s_0)	2 m
Headway time (T)	1.5 sec
Maximum acceleration (a)	0.73 m/s^2
Maximum deceleration (b)	1.67 m/s^2
Acceleration exponent (δ)	4

To simplify our study, the vehicle's communication range was limited to a maximum of three vehicles ahead. This decision was based on the understanding that distant vehicles exert less

influence on the decision made by the subject vehicle. This range can be easily extended to more vehicles if needed.

RESULT AND DISCUSSION

Following the development of a simulation platform, various cyber-attacks were initiated to assess their influence on traffic flow. Figure 2 shows the CAV platoon in the absence of any attacks. This figure serves as a basis in comparing cyberattack scenarios. In the absence of an attack, the graphs exhibit smoothness, showcasing the vehicles' uninterrupted travel. The graph illustrating car positions over time demonstrates their smooth functioning without any deviations. Initially, the cars accelerate, then enter a phase of stability. The graphs of velocity and acceleration against simulation time illustrate this behavior. As the lead vehicle approaches the road's end, it slows down, causing the vehicles behind it to also decelerate and reduce their speed. The graph showing the gap to the leading vehicle over time represents the changing distance between each vehicle and the one in front of it over time. The blue line on this graph shows the gap perceived by the first vehicle, which initially is large due to the open road ahead but decreases as the vehicle progresses and reaches the end of the road. This decreasing blue line represent this trend.

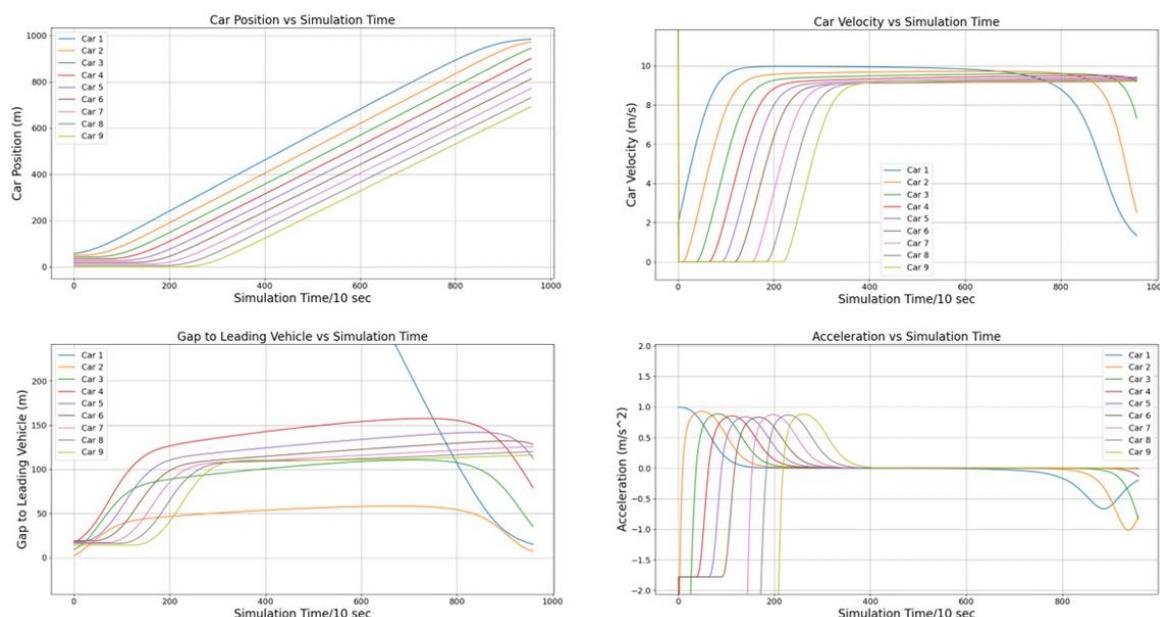


Figure 2. CAV Platoon without Cyber Attack

Attack Type: False Message

In this attack scenario, the fifth car within the platoon received two types of false messages. Initially, it received erroneous data related to spacing, indicating that the three vehicles it was meant to follow were positioned 80 meters farther ahead than their actual respective positions. This deceptive information led the targeted vehicle to perceive a larger gap in front, prompting it to accelerate. The attack commenced 10 seconds after the simulation began and persisted for 75 seconds in total. Within the attack timeframe, a rear-end collision occurred between vehicle no. 5

and vehicle no. 4 at 63 seconds into the simulation. Figure 3 depicts the disruption in traffic dynamics caused by this attack.

Another attack was aimed at manipulating the speed of the leading vehicles, specifically targeting the fifth vehicle once again. This attack commenced 10 seconds after the simulation began and persisted for 70 seconds. The vehicle was compromised to perceive its lead vehicles' velocities as tenfold less than their actual speed, resulting in a significant slowdown. The sixth vehicle, trailing the fifth vehicle was examined. Under normal circumstances, this vehicle should have reached the 600-meter mark within 75.3 seconds from the start of the simulation. However, due to the attack, a delay of 3.87 seconds was incurred in reaching the same position. Figure 4 illustrates the changes in the traffic flow under this type of attack.

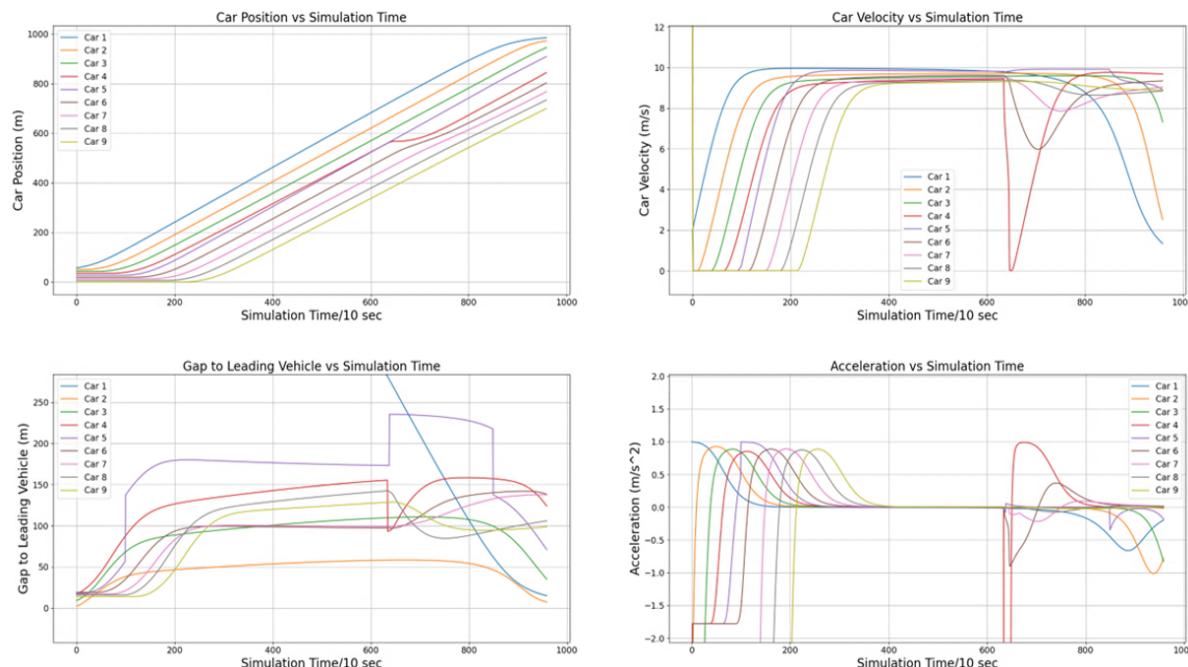


Figure 3. CAV Platoon under Manipulated Spacing Data

Attack Type: Platoon Leader Identity Attack

In this attack scenario, within the CAV platoon, the fifth vehicle was selected, and its communication was tampered with, causing a disruption in its connectivity with the preceding vehicles. The vehicle's perception was manipulated to disregard any vehicles ahead, prompting it to assume the role of the platoon leader. Consequently, the vehicle interpreted an open road ahead, initiating an acceleration that led to a rear-end collision with vehicle no. 4 at 53.4 seconds into the simulation. Figure 5 demonstrates the resultant alterations in traffic flow that ensued once the attack commenced 10 seconds into the simulation and lasted for 70 seconds.

Attack Type: Acceleration Manipulation

In this scenario, at the 30-second mark within the simulation, an attack was initiated which disrupted vehicle number 5's expected acceleration, introducing a new value. This alteration in

acceleration is depicted in Figure 6, displaying a sudden decline. This disruptive attack persisted for 10 seconds, enforcing a deceleration of 2m/s^2 . Consequently, the targeted vehicle experienced an abrupt slowdown, causing all the subsequent vehicles behind it to also slow down. A travel delay became apparent, as under normal conditions, vehicle number 6 would have reached the 600-meter mark in 75.33 seconds, but due to the attack, a delay of 11.43 seconds was observed.

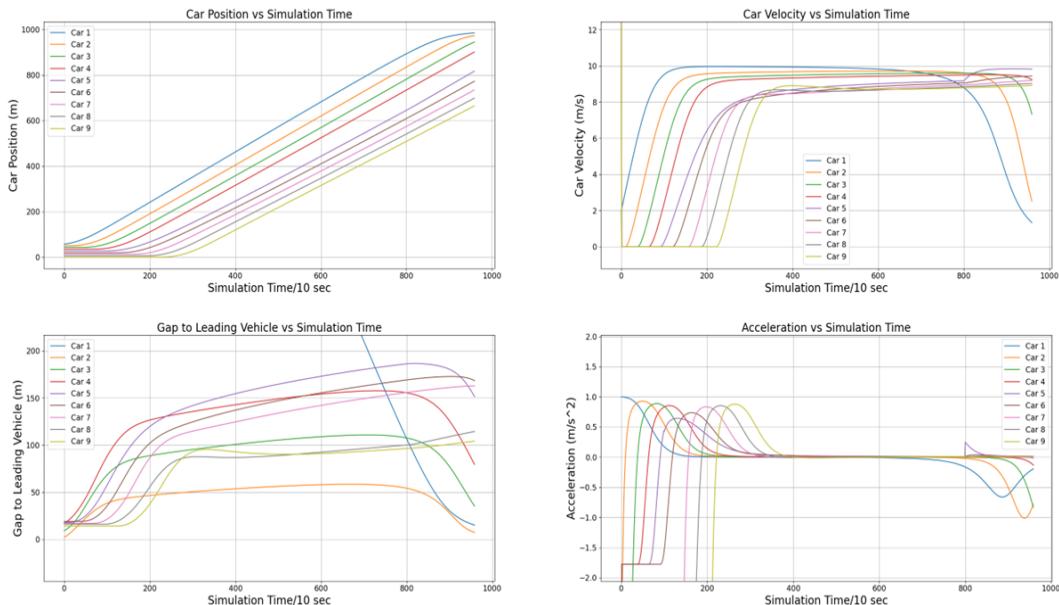


Figure 4. CAV Platoon under Manipulated Velocity Data

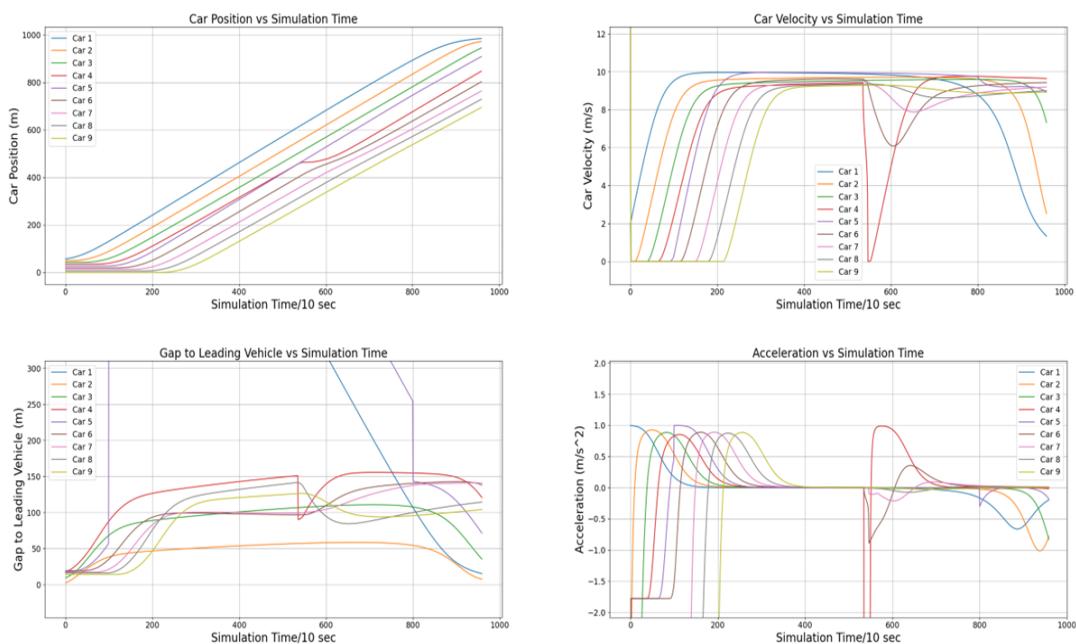


Figure 5. CAV Platoon with Platoon Leader Identity Attack

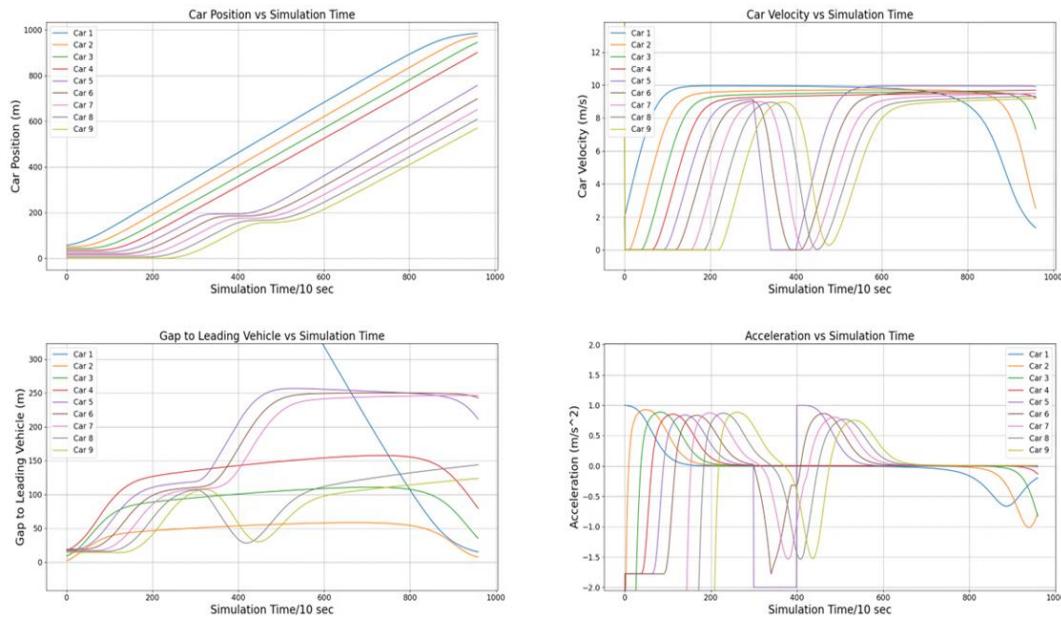


Figure 6. CAV Platoon under Acceleration Manipulation Attack

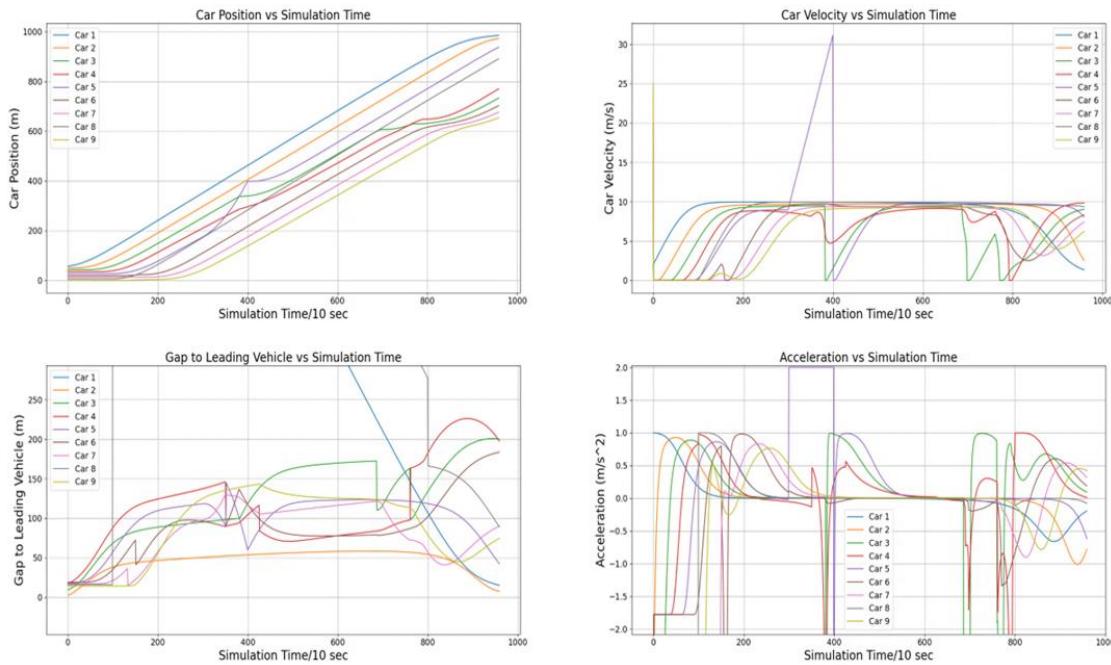


Figure 7. CAV Platoon under Multiple Vehicle Attack

Attack type: Multiple Vehicle Attack

Previously, a single vehicle was selected and was subjected to a cyber-attack. Now, three different vehicles were selected and attacked simultaneously. This coordinated attack led to significant disruption in the flow of traffic. Just 10 seconds into the simulation, vehicles 4, 5, and 8 were manipulated with false inputs. Specifically, vehicle 4 received inaccurate data, perceiving

its lead vehicle as 60 meters closer than reality. Meanwhile, vehicle 5 has its acceleration tampered with, set to a fixed value of 2m/s^2 . The consequences unfolded rapidly: vehicle 5 collided with vehicle 4, 25 seconds after the attack commenced. Furthermore, vehicle 8 encountered an attack that rendered it incapable of recognizing any vehicles ahead. As a result, it assumed the role of platoon leader and accelerated, leading to a collision with vehicle 7 within just 5 seconds of the attack's initiation. This disruption caused by this attack on traffic dynamic is portrayed in Figure 7.

CONCLUSION

CAVs bolster many advantages ranging from reduced congestion, and safer traffic to reduced fuel consumption, making it the goal for the future mode of transportation. However, it is very vulnerable to cyber-attacks. This paper aims to investigate various potential cyber threats and their repercussions on the traffic flow. Focusing on the worst-case scenario where a vehicle's inbuilt sensor and wireless communication are compromised, the study showcases the consequences of such attacks- collision and delay.

In this study, we explored attacks in two scenarios: when a single vehicle is targeted and when multiple vehicles are subjected to attacks. The outcomes were notably more severe when multiple vehicles were affected. Among the instances where a single vehicle was targeted, those inducing acceleration were identified as particularly dangerous, leading to rear-end collisions with the vehicle ahead. On the other hand, attacks causing the targeted vehicle to decelerate didn't result in collisions but extended travel time. This delay markedly exceeded the normal duration expected without any attacks. Although CAVs are envisioned as a safe and efficient mode of travel, these cyber threats pose a great challenge that needs to be tackled and overcome.

In this paper, we have tested an innovative car-following model for connected vehicles where the subject vehicle selects acceleration with the safest expected outcome. Through our analysis, a crucial finding was identified: attacks causing a vehicle to decelerate had a milder impact compared to those inducing acceleration. When a vehicle slowed down, its followers also decelerated, aligning with the principles of the car following model wherein subsequent vehicles adopt the minimum calculated acceleration. Conversely, vehicles that accelerated ultimately collided with the vehicle ahead.

One of the primary limitations of this study is that the subject vehicle is only concerned with the leading vehicles. Establishing communication with the trailing vehicles could mitigate the accident risks. The velocity and position data of the attacked vehicle can help the leading vehicle to either accelerate or exit the platoon and give way, to avoid collision.

Although the foundation of the proposed model lies in the Intelligent Driver Model (IDM) for vehicle following, its novel application to multi-anticipative car following sets it apart. Currently, the research is confined to simulations. To confirm the findings, it is necessary to compare them against real-world data obtained from external companies specializing in autonomous vehicle trials. Future research could explore these aspects further, offering valuable insights into enhancing vehicle autonomy.

REFERENCES

- Bajpai, J. N. (2016). Emerging vehicle technologies & the search for urban mobility solutions. *Urban, Planning and Transport Research*, 4(1), 83-100.

- Chandler, R. E., Herman, R., and Montroll, E. W. (1958). Traffic dynamics: studies in car following. *Operations research*, 6(2), 165-184.
- Sun, Z., Liu, R., Hu, H., Liu, D., and Yan, Z. (2023). Cyberattacks on connected automated vehicles: A traffic impact analysis. *IET Intelligent Transport Systems*, 17(2), 295-311.
- Khattak, Z. H., Smith, B. L., and Fontaine, M. D. (2021). Impact of cyberattacks on safety and stability of connected and automated vehicle platoons under lane changes. *Accident Analysis & Prevention*, 150, 105861.
- Masoud, N., Wang, Y., Zhang, R., and Liu, H. (2023). *Road-side based cybersecurity in connected and automated vehicle systems*.
- Parkinson, S., Ward, P., Wilson, K., and Miller, J. (2017). Cyber threats facing autonomous and connected vehicles: Future challenges. *IEEE transactions on intelligent transportation systems*, 18(11), 2898-2915.
- Pipes, L. A. (1953). An operational analysis of traffic dynamics. *Journal of applied physics*, 24(3), 274-281.
- Ploeg, J., Scheepers, B. T., Van Nunen, E., Van de Wouw, N., and Nijmeijer, H. (2011, October). Design and experimental evaluation of cooperative adaptive cruise control. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 260-265). IEEE.
- Rothery, R. W. (1992). Car following models. *Trac Flow Theory*.
- Treiber, M., Hennecke, A., and Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2), 1805.
- Van Arem, B., Van Driel, C. J., and Visser, R. (2006). The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on intelligent transportation systems*, 7(4), 429-436.
- Vander Werf, J., Shladover, S. E., Miller, M. A., and Kourjanskaia, N. (2002). Effects of adaptive cruise control systems on highway traffic flow capacity. *Transportation Research Record*, 1800(1), 78-84.
- Wang, P., Yu, G., Wu, X., Qin, H., and Wang, Y. (2018). An extended car-following model to describe connected traffic dynamics under cyberattacks. *Physica A: Statistical Mechanics and Its Applications*, 496, 351-370.
- Wang, P., Wu, X., and He, X. (2020). Modeling and analyzing cyberattack effects on connected automated vehicular platoons. *Transportation research part C: emerging technologies*, 115, 102625.
- Yağdereli, E., Gemci, C., and Aktaş, A. Z. (2015). A study on cyber-security of autonomous and unmanned vehicles. *The Journal of Defense Modeling and Simulation*, 12(4), 369-381.

Prediction of Daily Disengagements of Automated Vehicles Using Explainable Machine Learning Approach

Boniphace Kutela, Ph.D.¹; Sunday Okafor²; Norris Novat³; Tumlumbe Juliana Chengula⁴;
and John Kodi, Ph.D.⁵

¹Roadway Safety Program, Texas A&M Transportation Institute, Bryan, TX.

Email: b-kutela@tti.tamu.edu

²Univ. of Alabama, Tuscaloosa, AL. Email: scokafor1@crimson.ua.edu

³Iteris, Inc., Fairfax, VA. Email: nnovat@iteris.com

⁴South Carolina State Univ., Orangeburg, SC. Email: tchengul@scsu.edu

⁵HNTB Corporation, Tallahassee, FL. Email: jkodi@hntb.com

ABSTRACT

This study presents a comprehensive analysis of automated vehicle (AV) disengagements in California, leveraging traffic conditions and employing the XGBoost model and explainable AI techniques. The study investigates six distinct categories of disengagement initiation: driver-initiated, system-initiated, freeway-based, arterial-based, driverless-capable, and driverless-incapable. The findings reveal that predictors of each disengagement type differ significantly. Generally, the day of the week and vehicle miles of travel showed a great contribution to the prediction of overall AV disengagements. On the other hand, human-driven VMT was the main predictor for driver-initiated disengagement, while human-driven vehicle crashes were the highest-ranked factor for system-initiated, arterial roadways, and driverless-capable AV disengagement. For freeway-based disengagements, the year of operation was the major factor. The study contributes to the existing body of knowledge by providing a nuanced understanding of AV disengagements, emphasizing the need for targeted interventions and strategies. The research serves as a pioneering step toward unveiling hidden patterns in AV disengagements, showcasing the potential of explainable AI in unraveling complex AV disengagement scenarios, and providing a solid foundation for future advancements in AV technology and policy formulation. The findings point to further investigation of AV disengagements to understand the safe operation of AVs and better predict their maturity, thereby aiding policymakers in making informed decisions about AV testing and deployment.

INTRODUCTION

Testing of Automated Vehicles (AVs) on public roads has been underway in several states across the United States. California, Florida, Texas, Pennsylvania, and Nevada are among the states that have granted permission to AV operators for conducting tests on public roadways (Banner, 2023). California took the lead in AV testing back in 2014 (California-DMV, 2023). As of December 2022, the California Department of Motor Vehicles (CA DMV) has issued Automated Vehicle Testing Permits (with a driver) to 43 companies (California-DMV, 2023). The CA DMV mandates AV operators to report all cases of AV disengagements and crashes on public roads so as. This is done to promote innovation and gain better insights into the technology's performance and shortcomings. As of June 2023, there were 604 AV-related crashes reported in California (California-DMV, 2023). While crash data is recorded and

reported within ten days of an incident, disengagement reports are published annually (California DMV, 2023).

The availability of crash and disengagement reports have ignited the curiosity of researchers keen to understand AV operations. The significance of disengagement data lies in its role as a key aspect of the safe system approach. Unlike crashes, disengagements illustrate the actions taken by the AV in advance to prevent potential crashes. As such, analyzing cases of disengagements is essential for comprehending the specific conditions that trigger automatic disengagement or necessitate manual disengagement by the driver. Analyzing the cases of disengagements facilitates a deeper understanding of AV operations and the identification of opportunities for optimizing AV technology.

Most early studies centered on examining the descriptive statistics of AVs' disengagements. For instance, a study conducted between 2014 and 2017, using data from the CA DMV, revealed that 144 AVs were tested, covering a cumulative distance of 1,116,605 miles on public roads. Over this duration, a total of 5,328 disengagements were recorded (Banerjee et al., 2018). Some other studies have adopted complex methodologies to assess the factors associated with AV disengagements (Houseal et al., 2022). Recently, researchers have used disengagement data to gain insights into the maturity of AV technology (Guo & Zhang, 2022), and to identify key determinants influencing disengagement occurrences (Houseal et al., 2022).

While several efforts have been made to understand the nature of AV disengagement, the link to the traffic condition and weather data is not well understood. To be specific, the influence of the traffic and weather conditions on the overall and specific type of disengagement is not well known. For instance, does the factor associated with human-initiated disengagement also apply to system-initiated disengagements? What about freeways against arterial roadways, do they share the same sets of factors? The advantage of utilizing traffic and weather conditions to predict AV disengagement lies in the easy availability of such data. Therefore, this study utilized traffic and weather data to predict specific types of disengagement using an explainable machine-learning approach.

METHODOLOGY

In this study, the investigation focused on AV disengagements in the state of California, utilizing the powerful XGBoost model for predictive analysis. Subsequently, the XAI technique was employed to unravel the underlying factors contributing to the daily disengagements. The combination of XGBoost with XAI is expected to shed light on the factors of AV disengagements that would unveil further safety measures and AV system optimization. The types of disengagement whose prediction was computed were driver-initiated and system-initiated disengagement, freeway-based and arterial-based disengagements, and driverless capable and driverless incapable disengagements.

Analytical Approach

XGBoost Model

This model was employed for the black box predictive modeling for AV disengagements due to its high reliability in predicting outcomes. The model auto-breaks down the disengagement data into batches of multiple learners which through ensemble learning provided one coherent

model with an improved predicting power of each disengagement type. XGBoost has proven successful in various predictive models due to its renowned capability of employing a gradient-boosting framework that sequentially trains decision trees, with each tree attempting to correct the errors of the previous one (Ruseruka et al., 2023; Shi et al., 2021; Vanichrujee et al., 2018; Yang et al., 2021).

The objective function for this ensemble model consists of two components: a loss function that measures the difference between predicted and actual values, and a regularization term that controls the complexity of the model to prevent any possible overfitting. Inside the objective function, $\hat{y}_i^{(t)}$ is the prediction of the i^{th} disengagement record for a respective t^{th} iteration. f_t is added into the equation to minimize the objective function to improve the predictive power of the model. The objective function explains how the model iteratively builds decision trees and adjusts their hyperparameters to minimize the objective function (Yang et al., 2021). This results in an optimized ensemble model that can accurately predict AV disengagements based on input features from the dataset. The function can be represented as per Equation 1:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (1)$$

Where:

n is the number of training instances

y_i is the actual value of the target variable for sample i

\hat{y}_i is the predicted value of the target variable for sample i .

t is the number of weak learners (trees).

f_t represents the t^{th} iteration.

$\Omega(f_t)$ is the regularization term for an iteration t .

Explainable Artificial Intelligence (XAI)

The XAI technique was employed to interpret and understand the predictions made by the XGBoost model. Given its black-box nature, XAI methods help uncover the reasons behind model decisions toward disengagement and identify important features contributing to those disengagements. SHAP, a subdivision of XAI provides individual feature attributions for each prediction.

In this study, SHAP values were computed based on the distinct types of disengagement present in the dataset. The dataset was categorized into six specific types of disengagements, namely driver-initiated disengagements, system-initiated disengagements, freeway-based disengagements, arterial-based disengagements, driverless-capable disengagements, and driverless-incapable disengagements. By stratifying the dataset in this manner, the analysis provided a fine-grained understanding of how different disengagement scenarios influenced the model's predictions (Yang et al., 2021). The Shapley value (ϕ_i) on the model output is allocated based on their marginal contribution of feature i across all possible feature combinations. SHAP values were then calculated for each type of disengagement, providing valuable insights into the contribution of individual features to each specific disengagement type. Mathematically, SHAP values for a prediction can be represented as per Equation 2:

$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [f(S \cup \{j\}) - f(S)] \quad (2)$$

Where:

N is the set of all features.

i is the index feature of interest.

$f(S)$ is the model output for a specific instance with features S .

$|S|$ denotes the number of features in set S .

$|N|$ is the total number of features.

Data Description

This study utilized disengagement data collected from the California Department of motor vehicles (CA DMV). The DMV has authorized more than 30 companies to operate their AVs in California, including 7 companies that can operate driverless AVs. Further, the DMV been collecting this type of data since 2014 whereby each AV operator is supposed to report the annual disengagements. Additionally, the traffic condition and human driven crash data were collected from the Transportation Injury Mapping System (TIMS, 2023) website. The weather data on the other hand was collected from the weather underground website for San Francisco International Airport Station (*San Bruno, CA Weather History / Weather Underground*, 2023).

The collected data were cleaned and summarized (see Table 1). It can be observed that the maximum number of daily disengagements was 165 while the minimum was zero. For the specific type of disengagements, the maximum number of driver-initiated was 129 while system-initiated disengagements were 107, with an average of 12.7 and 3.3 disengagements, respectively. The maximum for arterials/streets was 163 while that of the freeways was 85 per day. Further, driverless capable AVs have only a maximum of 28 daily disagreements compared to 158 driverless incapable AVs. For the traffic-related variables, on average the weekly HDV crashes and HDV VMT are 1,084 crashes and 2,688 miles, respectively. The average monthly AV mileage was 291,500 miles with a minimum of 6,604 and a maximum of 666,454 miles. Other variables of interest are shown in Table 1.

RESULTS AND DISCUSSION

Overall Disengagements

Figure 1 (a) presents the feature importance graph plotted in terms of their RMSE, offering valuable insights into the most influential features across all disengagement types. Notably, disengagements on Saturday and Sunday, monthly AV VMT in ten thousand (M_AV_VMT_10k), Human -driven vehicle VMT in hundreds (HDV_VMT_100), and total HDV crashes (TotalHDV_Crashes) exhibit high feature importance. These features are critical in determining disengagement occurrences across different types. The time factors observed here might suggest that the day of the week may significantly influence the occurrence of disengagements. This may be due to the differences in traffic patterns, driver behaviors or existing road conditions. The presence of HDV and AV VMTs signify that overall travel distance of both AVs and HDVs has an impact on disengagement. Moreover, total human-driven vehicle crashes emerge as an important feature, implying that the safety conditions and accident rates of human-driven vehicles are linked to AV disengagements.

Table 1. Descriptive statistics of the model variables

Variable	Driver-initiated		System-initiated		On Freeways		On Streets		Driverless		Non-driverless	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
Year												
2019	6386	70.3%	2700	29.7%	1112	12.2%	7995	88.0%	547	6.0%	8560	94.0%
2020	2513	78.1%	705	21.9%	352	10.9%	2882	89.6%	43	1.3%	3191	98.7%
2021	2232	82.8%	463	17.2%	368	13.7%	2327	86.3%	32	1.2%	2663	98.8%
2022	7201	88.8%	909	11.2%	3386	41.8%	4724	58.2%	3	0.0%	8107	100.0%
Month of the year												
January	1711	80.2%	423	19.8%	227	10.6%	1911	89.6%	136	6.4%	2002	93.6%
February	1205	75.3%	395	24.7%	265	16.6%	1338	83.6%	52	3.2%	1551	96.8%
March	955	84.2%	179	15.8%	313	27.6%	822	72.5%	17	1.5%	1118	98.5%
April	1201	82.8%	250	17.2%	413	28.5%	1038	71.5%	16	1.1%	1435	98.9%
May	1084	88.6%	139	11.4%	253	20.7%	970	79.3%	17	1.4%	1206	98.6%
June	1111	79.4%	288	20.6%	269	19.2%	1131	80.8%	14	1.0%	1386	99.0%
July	1360	83.7%	264	16.3%	263	16.2%	1361	83.8%	13	0.8%	1611	99.2%
August	2525	86.3%	401	13.7%	737	25.2%	2190	74.8%	45	1.5%	2882	98.5%
September	2368	84.2%	444	15.8%	925	32.9%	1891	67.2%	94	3.3%	2722	96.7%
October	2209	73.9%	779	26.1%	728	24.4%	2262	75.7%	99	3.3%	2891	96.7%
November	1990	64.3%	1104	35.7%	759	24.5%	2335	75.5%	79	2.6%	3015	97.4%
December	613	84.7%	111	15.3%	66	9.1%	679	93.8%	43	5.8%	702	94.2%
Day of the week												
Sunday	163	78.4%	45	21.6%	60	28.8%	148	71.2%	2	1.0%	206	99.0%
Monday	3487	79.5%	897	20.5%	861	19.6%	3524	80.4%	144	3.3%	4241	96.7%
Tuesday	3751	81.3%	861	18.7%	1156	25.1%	3460	75.0%	85	1.8%	4531	98.2%
Wednesday	3584	80.4%	871	19.6%	1025	23.0%	3438	77.2%	84	1.9%	4379	98.1%
Thursday	3810	77.9%	1084	22.1%	1000	20.4%	3901	79.7%	150	3.1%	4751	96.9%
Friday	3200	80.8%	760	19.2%	1010	25.5%	2965	74.9%	149	3.7%	3826	96.3%
Saturday	337	56.5%	259	43.5%	106	17.8%	492	82.6%	11	1.8%	587	98.2%
Precipitation												
Yes	2536	82.5%	538	17.5%	697	22.7%	2385	77.6%	109	3.5%	2973	96.5%
No	15796	78.8%	4239	21.2%	4521	22.6%	15543	77.6%	516	2.6%	19548	97.4%

In Figure 1 (b), the Shapley values reveal that maximum humidity, total human-driven vehicle crashes, and monthly AV VMT are the features with the highest contributions to the model's predictions. These features likely play crucial roles in determining when and why disengagements occur across all types. High maximum humidity might indicate that adverse weather conditions can significantly impact AV performance and lead to disengagements. The number of human-driven vehicle crashes could be associated with the overall safety conditions and potential risks on the road. Additionally, the AV's monthly VMT might reflect the total travel distance of AVs, which can be related to their exposure to different driving conditions and scenarios.

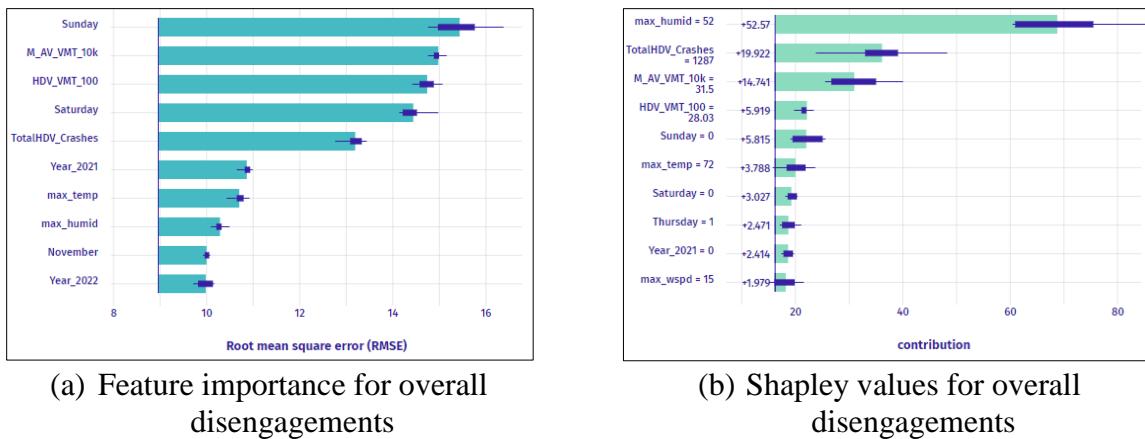


Figure 1. Feature importance and Shapley values for overall AV disengagements

Disengagement Initiation

In Figure 2(a), the shift in feature importance for driver-initiated disengagements highlights that human-driven vehicles VMT and Sunday remain significant, consistent with their importance in overall disengagements. This suggests that the travel distance of human-driven vehicles and the day of the week continue to influence the occurrence of driver-initiated disengagements. However, Saturday emerges as a more important feature, indicating that disengagements occurring on Saturdays specifically have a higher impact on driver-initiated scenarios. Additionally, the inclusion of the year 2022 as an important feature suggests that temporal trends or changes in driving conditions may have affected driver-initiated disengagements in that year. In Figure 2(b), for system-initiated disengagements, total human-driven vehicles and Sunday retain their importance from overall disengagements, signifying their consistent influence across disengagement types. Notably, November and maximum humidity now emerge as influential features. November could represent specific seasonal effects or changes in driving conditions during that month. The high importance of maximum humidity might indicate that weather conditions play a significant role in triggering system-initiated disengagements, as adverse weather can challenge AV systems' performance and decision-making.

For Figure 2(c), in contrast to overall disengagements, the shapley values for driver-initiated disengagements show that human-driven vehicle VMT, year 2022, and monthly VMT are the features with the highest contributions. While human-driven vehicle VMT remains important, the inclusion of year 2022 as a significant contributor suggests that temporal variations or specific events during that year might have influenced driver-initiated disengagements.

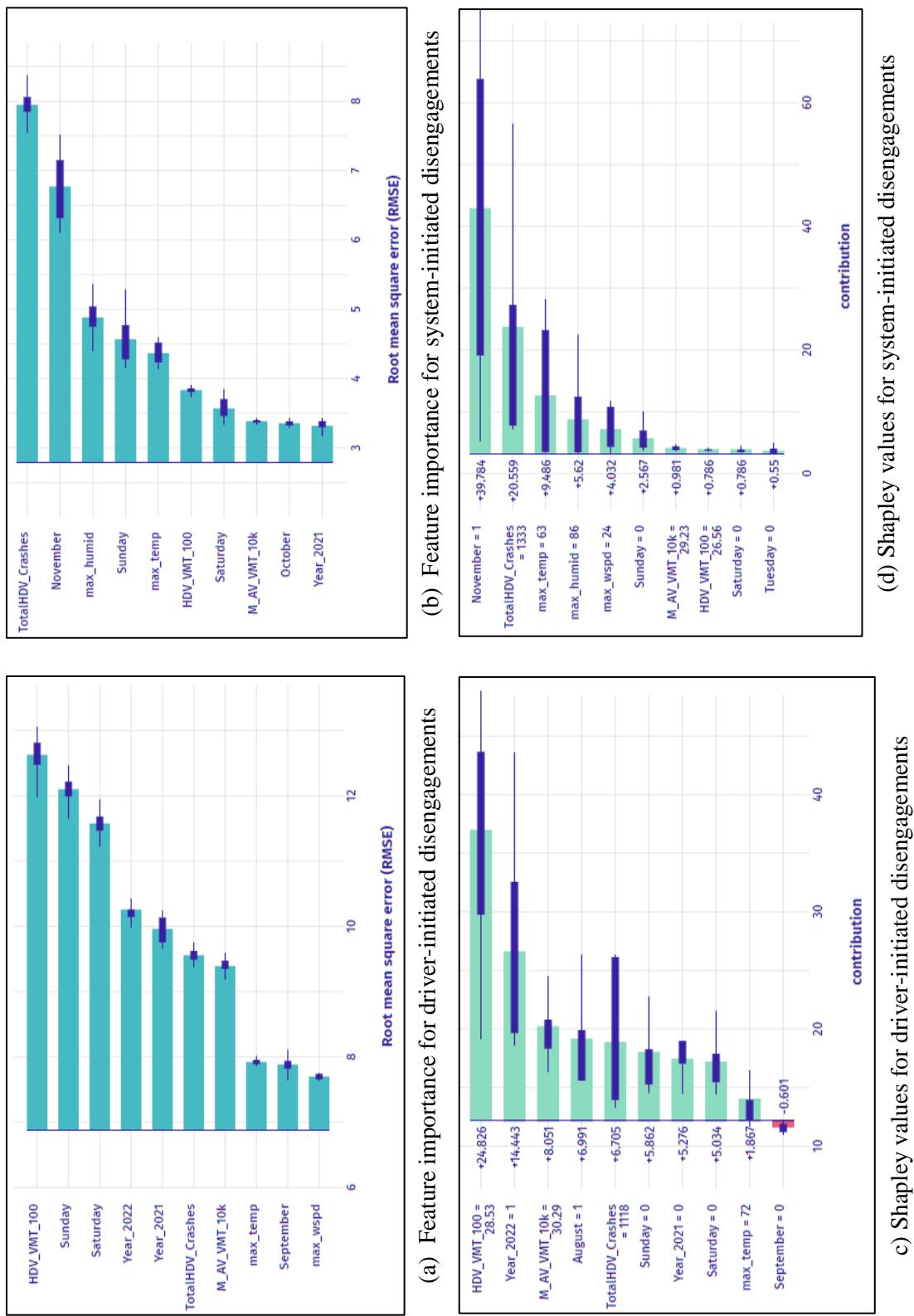
Moreover, monthly VMT's continued importance indicates that the travel distance of AVs remains a crucial factor in predicting driver-initiated disengagements. For system-initiated disengagements, Figure 2(d) shows November, total human-driven vehicle crashes, and maximum temperature are the features with the highest Shapley values. Notably, November's importance points to potential seasonality effects or specific events during that month that influenced system-initiated disengagements. The continued significance of total human-driven vehicle crashes suggests that overall safety conditions and risks on the road are critical factors. Additionally, the high importance of maximum temperature implies that extreme weather conditions can impact AV systems' behavior and trigger system-initiated disengagements.

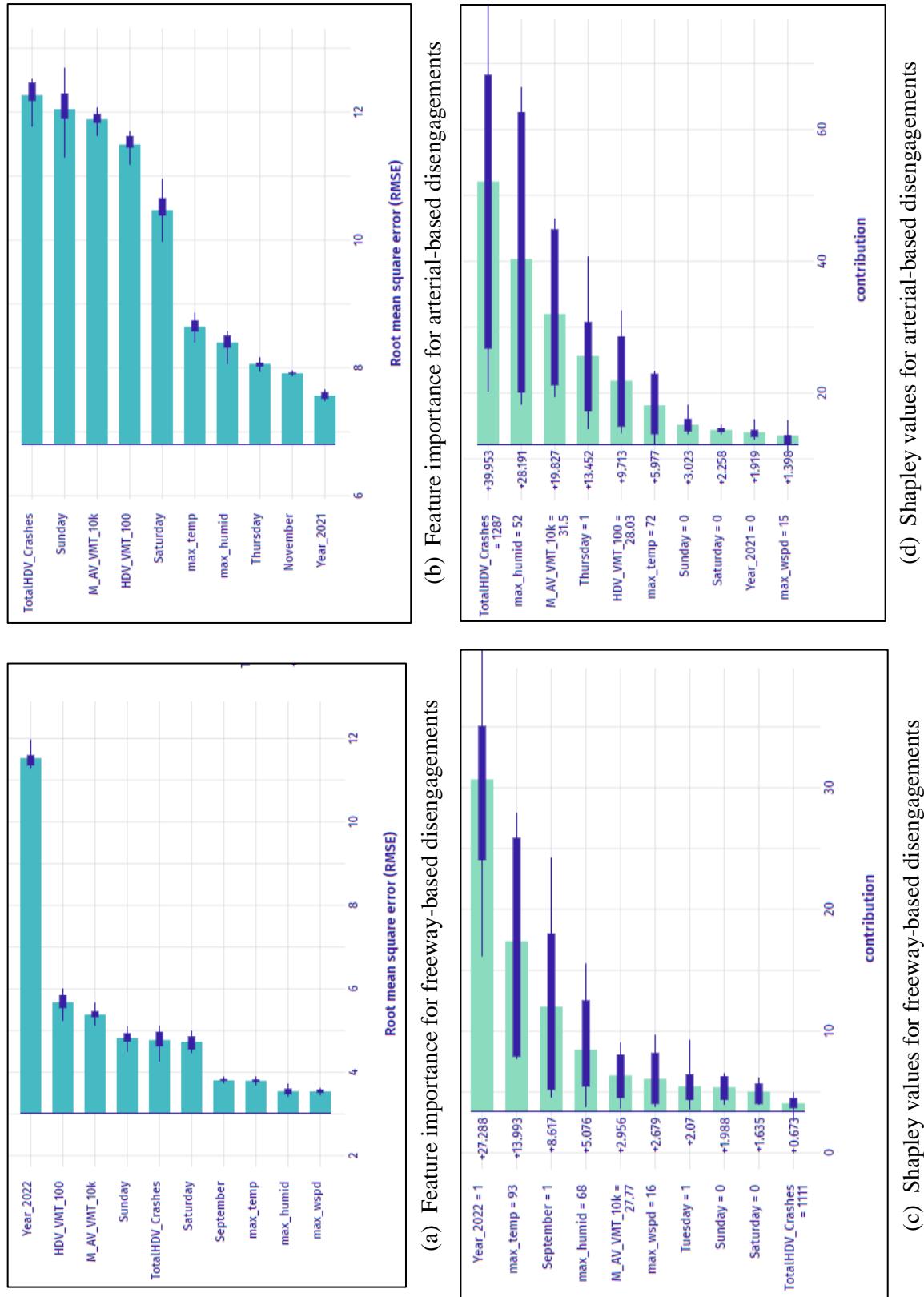
Facility Type

In Figure 3 (a) the change in feature importance for freeway-based disengagements reveals that the year 2022, human-driven vehicle VMT, and monthly AV VMT are the most important features. Notably, these features differ from the overall disengagements, where maximum humidity and total human-driven vehicle crashes were more influential. The inclusion of the year 2022 as an important feature suggests that temporal variations specific to that year may have played a significant role in freeway-based disengagements. Additionally, the prominence of human-driven vehicle VMT and monthly AV VMT indicates that the total travel distance of AVs and human-driven vehicles on freeways is crucial in predicting such disengagements (F. M. Favarò et al., 2017; Madadi et al., 2020). Figure 3 (b) for arterial-based disengagements shows that the most important features are total human-driven vehicle crashes, Sunday, and monthly AV VMT, which differs from the overall disengagements, where maximum humidity and human-driven vehicle VMT were more prominent. The higher importance of total human-driven vehicle crashes indicates that safety conditions and accident rates on arterial roads significantly impact AV disengagements in these scenarios. The continued relevance of human-driven vehicle VMT suggests that the total travel distance of both AVs and human-driven vehicles remains a crucial factor.

In Figure 3 (c), the Shapley values for freeway-based disengagements further reinforce the significance of the year 2022 feature, while also highlighting the contributions of maximum temperature and September as the temporal feature contributor for disengagement prediction. The influence of maximum temperature indicates that weather conditions on freeways could be critical factors affecting AV disengagements. The higher contribution of the September feature suggests that specific events or seasonal variations during that month might have impacted disengagements on freeways. Figure 3 (d) illustrates the Shapley values for arterial-based disengagements. This gives significance to total human-driven vehicle crashes and monthly AV VMT. Additionally, the contribution of maximum humidity suggests that weather conditions on arterial roads might play a role in predicting disengagements.

The change in feature importance and Shapley values for freeway-based and arterial-based disengagements reveals that road characteristics, driving conditions, environmental factors, and safety concerns play significant roles in influencing AV disengagements in each setting. Freeway-based and arterial-based roads exhibit distinct traffic flow, speed limits, and roadway designs, leading to different sets of important features for AV disengagements (Khattak et al., 2021). Driving conditions, including congestion, road geometry, and speed variations, create varying scenarios and challenges for AVs, influencing the importance of different features. Weather conditions, such as maximum temperature and humidity, can have varying impacts on AV disengagements based on the road type.

**Figure 2. Feature importance and Shapley values for AV disengagement by initiation**

**Figure 3. Feature importance and Shapley values for AV disengagement by facility type**

AV Capability

In Figure 4 (a), the feature importance for driverless-capable disengagements indicate that total human-driven vehicle crashes, maximum humidity, and maximum temperature are the most important features. These features differ from the overall disengagements, where maximum humidity, total human-driven vehicle crashes, and monthly AV VMT were more prominent. The inclusion of maximum temperature as an important feature for driverless-capable disengagements suggests that specific weather conditions might have a significant impact on AV disengagements in situations where the AVs can operate without human intervention (Walker et al., 2020; Zang et al., 2019). Figure 4 (b) shows that for driverless-incapable disengagements, the most important features are Sunday, human-driven vehicle VMT, Saturday, and monthly AV VMT. This differs from the overall disengagements, where maximum humidity, total human-driven vehicle crashes, and monthly AV VMT were more influential. The higher importance of specific days of the week, such as Sunday and Saturday, indicates that certain temporal patterns or driving behaviors on weekends might contribute to AV disengagements in driverless-incapable scenarios.

The Shapley values for driverless-capable disengagements in Figure 4 (c) reinforce the importance of maximum humidity, total human-driven vehicle crashes, and maximum temperature, indicating their strong contributions to the model's predictions. The influence of maximum humidity and temperature further supports the role of weather conditions in influencing disengagements in driverless-capable scenarios. The Shapley values for driverless-incapable disengagements from Figure 4 (d) highlight the significance of maximum humidity and total human-driven vehicle crashes, which are consistent with their importance in overall disengagements. Additionally, the contribution of monthly AV VMT suggests that the total travel distance of AVs remains relevant in predicting driverless-incapable disengagements.

The change in feature importance and Shapley values for driverless-capable and driverless-incapable disengagements highlights the influence of AV capability, weather conditions, and temporal patterns on AV disengagements. Driverless-capable AVs may face unique challenges and triggers specific to their higher level of autonomy, with weather conditions being particularly critical in these scenarios. On the other hand, driverless-incapable AV disengagements may be influenced by specific temporal patterns, such as weekends, and potentially related driving behaviors or road conditions.

Different types of disengagements were explained by different feature contributions. From the main disengagement to the defined categories of disengagements, feature importance has been observed to interchangeably contribute to the model prediction. The change of feature importance and contribution between overall disengagements and the distinguished types of disengagement can be attributed to the distinct driving factors and underlying patterns associated with each type. Results summaries indicate that weather conditions play a major role in AV disengagement, and crashes caused by human-driven vehicles also greatly contributed to AV disengagement prediction (F. M. Favarò et al., 2017; Walker et al., 2020; Zang et al., 2019). Another persistent factor that is applied in most categories of disengagement is the monthly AV vehicle miles of travel. Factors contributing to arterial disengagements were highly related to VMT (Khattak et al., 2021; Madadi et al., 2020). Time elements were observed as mild contributory factors towards disengagement prediction, while they topped as important features, their contribution towards the model performance was mediocre. One important observation was that 2022 had the most disengagements.

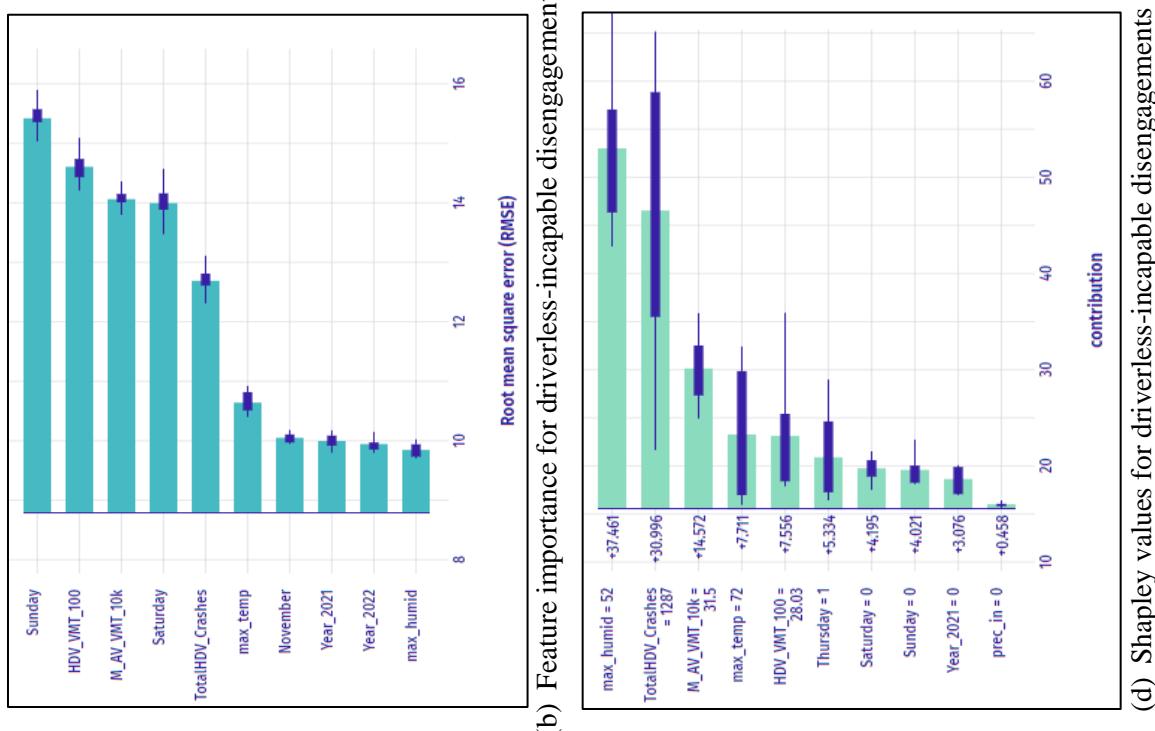
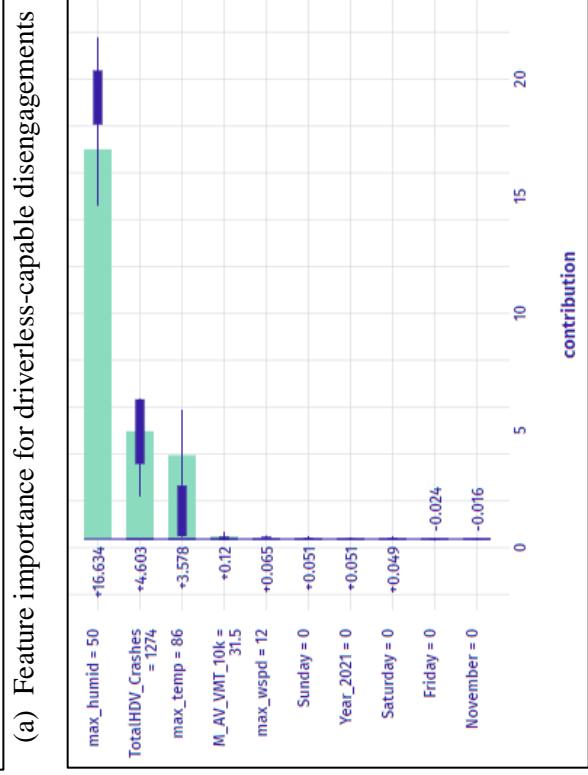
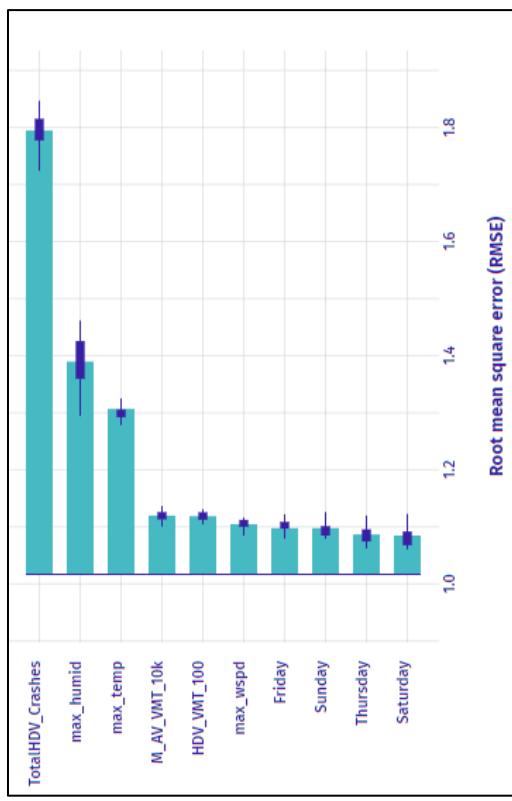


Figure 4. Feature importance and Shapley values for AV disengagement by capability

In conclusion, this analysis convincingly argues for a focus on specific types of disengagement rather than treating all disengagements as a uniform phenomenon. Understanding the discrepancies between the various types will enable researchers and engineers to develop targeted interventions, optimize AV systems, and improve their overall reliability and safety on the road. By recognizing the uniqueness of each disengagement type, the research contributes to the advancement of autonomous driving technology and the realization of a safer and more efficient future for self-driving vehicles.

CONCLUSION AND FUTURE WORKS

This research study performed an in-depth analysis of AV disengagements in California using the powerful XGBoost black box model and XAI techniques, specifically Shapley values. By investigating six distinct categories of disengagement initiation, namely driver-initiated, system-initiated, freeway-based, arterial-based, driverless-capable, and driverless-incapable disengagement, valuable insights were gained into the varied driving factors and patterns influencing each disengagement type. The change in feature importance and Shapley values between overall disengagements and specific types underscored the uniqueness of each category, driven by factors such as road characteristics, driving conditions, weather, and AV capability.

The findings of this study reveal that AV disengagements are influenced by a variety of factors, including road characteristics, driving conditions, weather, and AV capability. The study identified that different types of disengagements - driver-initiated, system-initiated, freeway-based, arterial-based, driverless-capable, and driverless-incapable - each have unique driving factors and patterns. Similarly, freeway-based and arterial-based disengagements arose from distinct driving conditions and traffic flow patterns. The study also found that weather conditions and crashes caused by human-driven vehicles greatly contributed to AV disengagement prediction. These findings underscore the complexity of AV disengagements and the need for a nuanced approach to their analysis and mitigation.

The study contributes significantly to the existing body of knowledge by providing a detailed and nuanced understanding of AV disengagements. It goes beyond the traditional approach of treating all disengagements as a uniform phenomenon and instead focuses on the unique characteristics and underlying causes of different types of disengagements. This approach allows for the development of more targeted interventions and strategies for addressing AV disengagements, thereby enhancing the safety and reliability of autonomous driving systems.

This study serves as a pioneering step towards unveiling hidden patterns hence leading to more reliable autonomous driving systems. By leveraging on the dynamics of AV disengagements and presenting study findings for the uniqueness of each disengagement type, the study emphasizes the necessity for tailor-made strategies and approaches in addressing AV disengagements. The convergence of sophisticated machine learning models with interpretability techniques, as demonstrated in this study, showcases the potential of Explainable AI in unraveling complex AV disengagement scenarios. Ultimately, this research provides a solid foundation for future advancements in AV technology, policy formulation, and the realization of a future where autonomous vehicles play a pivotal role in enhancing road safety and transforming transportation ecosystems.

REFERENCES

- Banerjee, S. S., Jha, S., Cyriac, J., Kalbarczyk, Z. T., and Iyer, R. K. (2018). Hands off the wheel in autonomous vehicles?: A systems perspective on over a million miles of field data. *Proceedings - 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2018*, 586–597.
- Banner, J. (2023). Are Self-Driving Vehicles Legal in My State? <https://www.motortrend.com/features/state-laws-autonomous-self-driving-driverless-cars-vehicles-legal/>.
- Boggs, A. M., Wali, B., and Khattak, A. J. (2020). Exploratory analysis of automated vehicle crashes in California: A text analytics & hierarchical Bayesian heterogeneity-based approach. *Accident Analysis and Prevention*, 135, 105354.
- California-DMV. (2023). Autonomous Vehicle Collision Reports - California DMV. <https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-collision-reports/>.
- CaliforniaDMV. (2023). Disengagement Reports - California DMV. <https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/disengagement-reports/>.
- Das, S., Dutta, A., and Tsapakis, I. (2020). Automated vehicle collisions in California: Applying Bayesian latent class model. *IATSS Research*.
- Dixit, V. V., Chand, S., and Nair, D. J. (2016). *Autonomous Vehicles: Disengagements, Accidents and Reaction Times*.
- Favarò, F., Eurich, S., and Nader, N. (2018). Autonomous vehicles' disengagements: Trends, triggers, and regulatory limitations. *Accident Analysis and Prevention*, 110(October 2017), 136–148.
- Favarò, F. M., Nader, N., Eurich, S. O., Tripp, M., and Varadaraju, N. (2017). Examining accident reports involving autonomous vehicles in California. *PLOS ONE*, 12(9), e0184952.
- Guo, X., and Zhang, Y. (2022). Maturity in Automated Driving on Public Roads: A Review of the Six-Year Autonomous Vehicle Tester Program. *Transportation Research Record*, 2676(11), 352–362.
- Houseal, L. A., Gaweesh, S. M., Dadvar, S., and Ahmed, M. M. (2022). Causes and Effects of Autonomous Vehicle Field Test Crashes and Disengagements Using Exploratory Factor Analysis, Binary Logistic Regression, and Decision Trees. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(8), 571–586.
- Khattak, Z. H., Fontaine, M. D., and Smith, B. L. (2020). Exploratory Investigation of Disengagements and Crashes in Autonomous Vehicles Under Mixed Traffic: An Endogenous Switching Regime Framework. *IEEE Transactions on Intelligent Transportation Systems*, 22(12), 7485–7495.
- Khattak, Z. H., Fontaine, M. D., and Smith, B. L. (2021). Exploratory Investigation of Disengagements and Crashes in Autonomous Vehicles Under Mixed Traffic: An Endogenous Switching Regime Framework. *IEEE Transactions on Intelligent Transportation Systems*, 22(12), 7485–7495.
- Kutela, B., Avelar, R. E., and Bansal, P. (2022). Modeling Automated Vehicle Crashes with a Focus on Vehicle At-Fault, Collision Type, and Injury Outcome. *Journal of Transportation Engineering, Part A: Systems*, 148(6).

- Kutela, B., Das, S., and Dadashova, B. (2021). Mining patterns of autonomous vehicle crashes involving vulnerable road users to understand the associated factors. *Accident Analysis & Prevention*, 106473.
- Lv, C., Cao, D., Zhao, Y., Auger, D. J., Sullman, M., Wang, H., Dutka, L. M., Skrypchuk, L., and Mouzakitis, A. (2018). Analysis of autopilot disengagements occurring during autonomous vehicle testing. *IEEE/CAA Journal of Automatica Sinica*, 5(1), 58–68.
- Madadi, B., Nes, R., Snelder, M., and Arem, B. (2020). A bi-level model to optimize road networks for a mixture of manual and automated driving: An evolutionary local search algorithm. *Computer-Aided Civil and Infrastructure Engineering*, 35(1), 80–96.
- Nordhoff, S., and de Winter, J. (2023). *Why do drivers and automation disengage the automation? Results from a study among Tesla users.*
- Novat, N., Kidando, E., Kutela, B., and Kitali, A. E. (2022). A comparative study of collision types between automated and conventional vehicles using Bayesian probabilistic inferences. *Journal of Safety Research*.
- Ruseruka, C., Mwakalonge, J., Comert, G., Siuhi, S., Ngeni, F., and Major, K. (2023). Pavement Distress Identification Based on Computer Vision and Controller Area Network (CAN) Sensor Models. *Sustainability*, 15(8), 6438.
- San Bruno, CA Weather History | Weather Underground. (2023). <https://www.wunderground.com/history/daily/us/ca/san-bruno/KSFO>.
- Shi, R., Xu, X., Li, J., and Li, Y. (2021). Prediction and analysis of train arrival delay based on XGBoost and Bayesian optimization. *Applied Soft Computing*, 109, 107538.
- Sinha, A., Vu, V., Chand, S., Wijayaratna, K., and Dixit, V. (2021). A Crash Injury Model Involving Autonomous Vehicle: Investigating of Crash and Disengagement Reports. *Sustainability* 2021, Vol. 13, Page 7938, 13(14), 7938.
- Solís Marcos, I. (2018). *Utaningar inom delvis automatiserad körning : Ett human factors-perspektiv*. Linköping University Electronic Press.
- TIMS (Transportation Injury Mapping System). (2023). TIMS - Transportation Injury Mapping System. <https://tims.berkeley.edu/>.
- Vanichrujee, U., Horanont, T., Pattara-atikom, W., Theeramunkong, T., and Shinozaki, T. (2018). Taxi Demand Prediction using Ensemble Model Based on RNNs and XGBOOST. *2018 International Conference on Embedded Systems and Intelligent Technology & International Conference on Information and Communication Technology for Embedded Systems (ICESIT-ICICTES)*, 1–6.
- Walker, C. L., Boyce, B., Albrecht, C. P., and Siems-Anderson, A. (2020). Will Weather Dampen Self-Driving Vehicles? *Bulletin of the American Meteorological Society*, 101(11), E1914–E1923.
- Wang, J., Zhang, L., Huang, Y., and Zhao, J. (2020). Safety of Autonomous Vehicles. *Journal of Advanced Transportation*, 2020.
- Wang, S., and Li, Z. (2019). Exploring causes and effects of automated vehicle disengagement using statistical modeling and classification tree based on field test data. *Accident Analysis & Prevention*, 129, 44–54.
- Xu, C., Ding, Z., Wang, C., and Li, Z. (2019). Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. *Journal of Safety Research*, 71, 41–47.

- Yang, C., Chen, M., and Yuan, Q. (2021). The application of XGBoost and SHAP to examining the factors in freight truck-related crashes: An exploratory analysis. *Accident Analysis & Prevention*, 158, 106153.
- Zang, S., Ding, M., Smith, D., Tyler, P., Rakotoarivelo, T., and Kaafar, M. A. (2019). The Impact of Adverse Weather Conditions on Autonomous Vehicles: How Rain, Snow, Fog, and Hail Affect the Performance of a Self-Driving Car. *IEEE Vehicular Technology Magazine*, 14(2), 103–111.

Machine Learning Strategies for Optimizing Urban Parking: A Comparative Evaluation

Sai Sneha Channamallu¹; Sharareh Kermanshachi, Ph.D., P.E.²;
Jay Michael Rosenberger, Ph.D.³; and Apurva Pamidimukkala, Ph.D.⁴

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: sxc5928@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Pennsylvania State Univ., State College, PA. Email: svk5464@psu.edu

³Professor, Dept. of Industrial Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: jrosenbe@uta.edu

⁴Assistant Professor, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: apurva.pamidimukkala@mavs.uta.edu

ABSTRACT

Parking management presents a complex challenge in urban cities, as a scarcity of parking spaces and the ever-increasing vehicular traffic have led to congestion, environmental pollution, and overall reduced urban productivity. Addressing the problem requires predicting the exact number of available parking spaces and categorizing parking occupancy levels. This study aims to achieve these tasks by employing machine learning models to accurately predict occupancy, thus optimizing parking resource allocation and enhancing the urban parking experience. A dataset derived from a college campus garage for a period spanning from January 2022 to June 2023 was used to analyze the performance of various predictive models, including random forest, decision tree, linear regression, and support vector machine. The models were compared using multiple evaluation metrics, and the results revealed that the random forest model was the most reliable. Its strong performance in regression analysis translated into precise estimations of available parking spaces. Similarly, its capability in classification analysis proved essential for categorizing parking occupancy into distinct levels, enhancing communication and streamlining decision-making processes. These findings are significant for improving parking management systems and contributing to the development of efficient and sustainable parking solutions in urban environments.

KEYWORDS: Parking Management, Predictive Modeling, Random Forest Model, Decision Tree, Support Vector Machine, Linear Regression.

INTRODUCTION

Parking management is becoming an increasingly challenging task in urban areas and on college campuses due to the growing number of vehicles and limited number of parking spaces. The rise in economic development and urbanization has led to a significant increase in car ownership, further exacerbating the mismatch between parking supply and demand (Yang et al., 2019; Pamidimukkala et al., 2023; Patel et al., 2022a). Drivers spend an average of 3.5 to 14 minutes searching for a parking spot, depending on the location and time of day, which not only leads to frustration and delays, but also wastes fuel and contributes to pollution (Zheng et al., 2015; Channamallu et al., 2023; Khan et al., 2022a). Minimizing the amount of time spent

searching for parking can help reduce the number of cars circulating in parking lots and decrease traffic congestion and noise pollution, particularly in the vicinity of parking lot entrances (Caicedo et al., 2012; Kotb et al., 2017; Etminani-Ghasroodashti et al., 2022).

The inefficient methods employed to search for parking are due to a lack of real-time and near-future information on the availability of spaces (Yang et al., 2019; Khan et al., 2022b). This information can be made accessible to drivers by implementing real-time parking maps delivered through their navigation systems that would provide information about parking availability (Huang et al., 2018; Sester, 2020; Patel et al., 2022b) and could also be used by parking facility managers and city planners to better manage and allocate parking spaces based on the predicted demand.

Research has shown that drivers with access to parking space information are 45% more likely to secure a spot (Caicedo et al., 2006; Khan et al., 2022c), and the realm of influence extends beyond that of relieving the driver of frustration; it is also crucial for reducing traffic and energy use (Shoup, 2006; Patel et al., 2023c) and for improving transportation management (Lin et al., 2017; Etminani-Ghasroodashti et al., 2023). Ignoring parking predictions increases stress and wastes time and fuel. In the US, drivers spend approximately 17 hours annually looking for a place to park, costing \$345 each (INRIX Research, 2017; Patel et al., 2022a); in the U.K. and Germany, the amounts of time and costs are even higher. This highlights the importance of addressing the problem.

Parking occupancy predictions have traditionally been based on static rules and heuristics that often lack precision and flexibility and rely on manual observations, limited historical data, and/or basic statistical techniques that may not be accurate or scalable enough for complex urban parking dynamics (Channamallu et al., 2023a). The development of advanced data-gathering technologies, availability of extensive parking datasets, and introduction of machine learning techniques, however, have precipitated a shift towards more data-driven approaches. Utilizing machine learning models allows parking management systems to better predict parking needs, optimize space allocation, and provide precise information to drivers, which leads to less congestion, increased user satisfaction, and greater overall efficiency (Liu et al., 2019; Sun et al., 2019; Khan et al., 2023a; Patel et al., 2022b).

This study focuses on meeting the challenges presented by the absence of real-time parking information in urban environments. It centers on analyzing historical occupancy data, time patterns, and events related to urban parking spaces to develop an intelligent model that accurately forecasts the availability of parking spaces. The primary objective of the paper is to gain an understanding of the strengths and limitations of machine learning models by analyzing their ability to successfully perform both regression and classification tasks within the context of parking occupancy. The study's goal is to bridge the gap between current parking management systems and the evolving demands of urban environments by leveraging the predictive power of machine learning. The results of this research hold practical significance for parking management systems in urban environments and can aid in creating efficient and sustainable parking solutions.

LITERATURE REVIEW

A significant body of research explores various methodologies for investigating the intricate nature of parking patterns and providing precise estimations of available parking spaces. It highlights the effectiveness of various machine learning models, with a particular focus on random forest and decision tree algorithms. Yalcin and Zeydan (2016) discovered that the

random forest model is more accurate than both decision trees and SVR, although decision trees are more computationally efficient. Wu et al. (2017) explored the suitability of decision tree-based techniques for accurately predicting parking occupancy and affirmed their effectiveness. Dey and Nath (2019) recognized the potential of decision tree models for accurately estimating parking availability; Liu et al. (2019) emphasized the efficacy of gradient boosting decision tree algorithms in forecasting short-term parking occupancy; and Sun et al. (2019) reported that decision tree algorithms, particularly when used in conjunction with genetic algorithms, can provide reliable parking availability estimates. A comparative analysis performed by Farooq et al. (2019) revealed that random forest and decision tree models consistently achieved higher accuracy and a higher F1-score than SVR and linear regression models, with random forest also showing superior precision and recall. Anwar et al. (2019) noted that while SVR demonstrated better recall, linear regression models had the lowest mean absolute error. He et al. (2020) further corroborated the superior performance of random forest and decision tree models in terms of accuracy. More recent studies by Srinivasan et al. (2021), Wu et al. (2020), Kim and Kim (2021), Zhang et al. (2021), and Shi et al. (2021) consistently validate the greater accuracy of random forest and decision tree models over SVR and linear regression, affirming their superiority in the field of parking occupancy predictions.

Random forest and decision tree algorithms have consistently demonstrated superior performance over SVR and linear regression in accuracy and precision (Farooq et al., 2019; He et al., 2020; Srinivasan et al., 2021). Random forest models are noted for their computational efficiency while still offering reliable parking availability predictions (Yalcin and Zeydan, 2016; Wu et al., 2020; Subramanya et al., 2022; Channamallu et al., 2023b), and decision tree models are also acknowledged for their ability to provide trustworthy forecasts of parking availability (Wu et al., 2017; Dey and Nath, 2019; Liu et al., 2019; Sun et al., 2019; Khan et al., 2023b; Channamallu et al., 2023c). This study aims to leverage these insights to further the development of effective and efficient models for predicting parking occupancy. The collective wisdom from these studies guides the selection of methodologies, features, and evaluation metrics, paving the way for a holistic approach to parking management.

METHODOLOGY

The first step of this research entailed an extensive literature review to thoroughly understand the current state of knowledge in the field. This critical phase laid the groundwork for the study. A detailed search was performed across various academic databases, including Google Scholar, IEEE Xplore, Springer, ProQuest, Science Direct, Scopus, and Web of Science, using carefully selected keyword combinations to precisely target the research topic. The aim of this literature search was to compile a wide range of relevant scholarly materials, such as conference proceedings, journal articles, and technical papers, that shed light on the methods and developments in parking availability prediction. This involved not only finding pertinent publications but also comprehensively analyzing and integrating the information they provided to form a well-rounded understanding of the subject. This extensive review informed the subsequent proposal for the parking occupancy prediction framework depicted in Figure 1.

Data Collection

The dataset was comprised of 13,104 data points and included the date and hourly occupancy from January 2022 to June 2023. Figure 2 presents the hourly occupancy in percentages. To

determine historical occupancy rates, a 60-minute lookback window was employed that encompassed various inputs that provided the model with information regarding occupancy rates 60 minutes before an attempt was made to make a prediction.

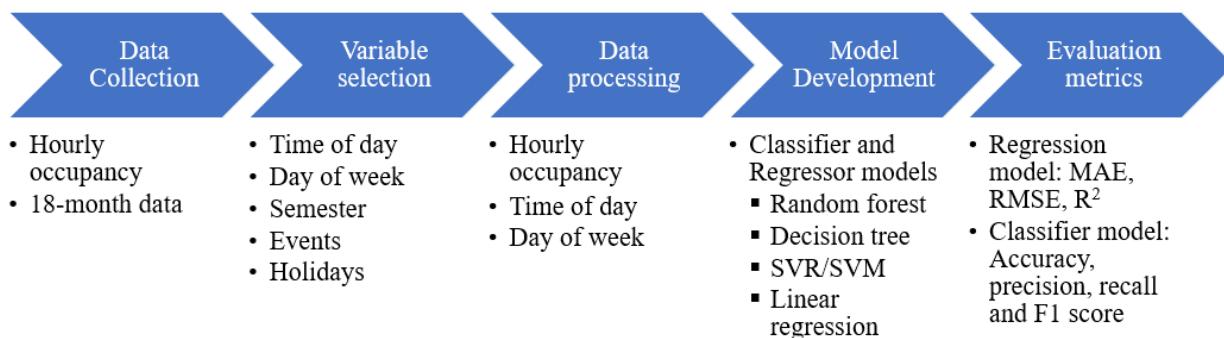


Figure 1. Research framework

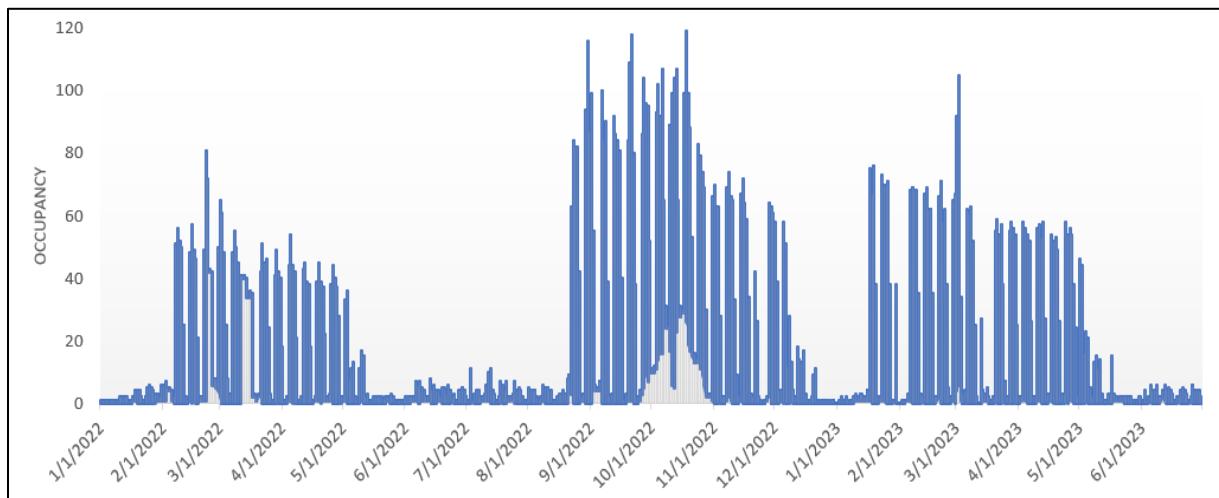


Figure 2. Data set

Variable selection

Relevant features, including the day of the week, time of day, and other contextual information such as semester, holidays, and special events, were selected based on their potential impact on parking occupancy. Figure 3 presents the variations in occupancy by the variables employed. The time of day was revealed to be a crucial factor in predicting parking occupancy as it varied significantly, for example, between early morning hours and midday. The day of the week was also shown to be important, as parking demands on weekends and weekdays can be vastly different on college campuses, with weekdays experiencing higher occupancy due to classes and staff presence. This study considered the impact of academic semesters on parking occupancy and analyzed the data separately for fall, spring, and summer semesters, as parking patterns differ across semesters due to variations in class schedules. For example, summer semesters have much lower occupancy than the fall and spring semesters.

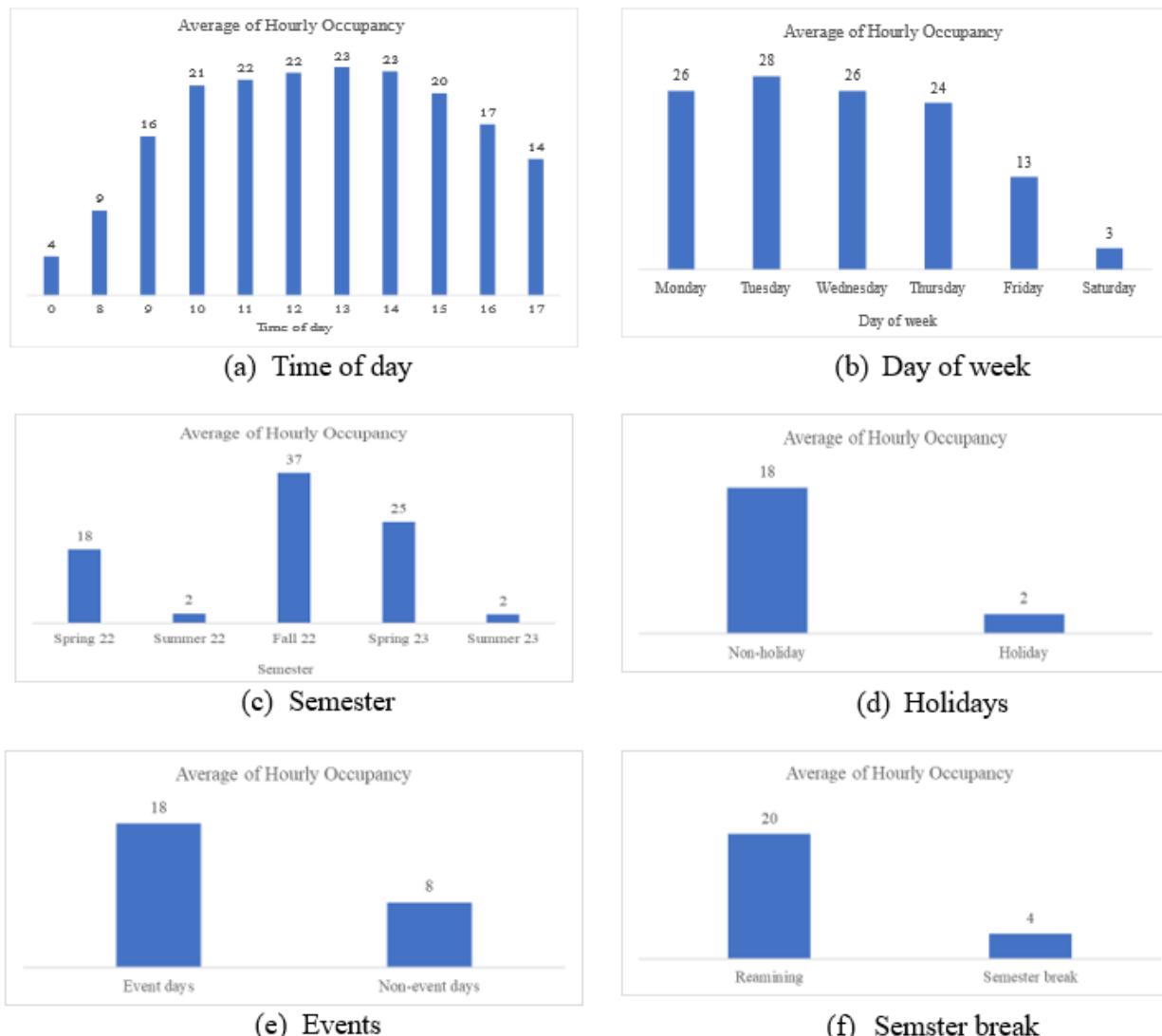


Figure 3. Variations in average hourly occupancy

Data pertaining to holidays was manually gathered from different sources to account for significantly reduced parking occupancy on these dates that would otherwise be outliers in predictive models. The analysis also considered semester breaks, recognizing that the parking occupancy patterns would be different from those of the regular academic sessions, as lower occupancy is typical during these breaks due to the lack of regular classes and reduced presence of students. Critical event dates included exams, highly attended sports matches, and concerts, as such events can lead to significant deviations in normal parking patterns, often causing spikes in occupancy.

Data Preprocessing

The analysis considered hourly occupancy data as a continuous and categorical variable. For classification purposes, the data was re-categorized into three distinct groups (low, medium, and high), based on tercile distributions, with specific threshold values set for each category to

provide clearer insights and more accurate predictions. Given the focus on providing accurate occupancy predictions during typical college operational hours, which are usually from 8 am to 5 pm, the hourly occupancy data for the hours outside this range were averaged and treated as equivalent to the 0-hour mark. This adjustment was made to refine the analysis by concentrating on the most relevant time periods. Data from Saturdays and Sundays was grouped together and categorized as the weekend to simplify the analysis by treating the weekend as a single entity, acknowledging the distinct parking patterns typically observed on these days.

Prediction methods

Based on the aforementioned review and relevant information, we selected linear regression, random forest, SVM, and decision tree models for our research, as they exhibit exceptional prediction performance and practical feasibility. By comparing the performance of these models, we aimed to determine their suitability for accurate parking occupancy predictions for a college campus garage.

Random forest models are particularly notable for their robust handling of complex variable relationships and efficacy in minimizing overfitting, a common challenge in predictive modeling. As an ensemble method that amalgamates multiple decision trees, it offers more generalized and dependable predictions than what might be achieved with individual decision trees. The decision tree model stands out for its simplicity and ease of interpretation. It is especially beneficial in contexts where understanding the model's decision-making process is crucial, as it delineates a clear, straightforward decision pathway based on input features, and it is user-friendly for stakeholders without a technical background. Linear regression is appreciated for its straightforward approach in modeling linear relationships between predictors and a target variable and is particularly effective in scenarios where the linear assumption is valid. It is easy to implement and understand, and the models are ideal for initial analyses or situations where data relationships are linear. SVR is lauded for its ability to manage nonlinear relationships, which is vital for complex scenarios like parking occupancy predictions. Its proficiency in dealing with nonlinearities makes it a strong candidate for cases where predictor-target relationships are intricate.

The models were implemented using Python libraries, then were trained and tested by dividing the dataset into subsets: 70% training and 30% testing. As the input data follows a sequential nature, such as a time series, we preserved its chronological order during the testing phase to allow us to evaluate the model's sensitivity to seasonal patterns.

Evaluation metrics

The mean absolute error (MAE), root mean square error (RMSE), and R-squared (R^2) were employed as key metrics in the regression analysis for assessing the performance of predictive models for parking occupancy. RMSE and MAE are absolute indicators, and comparing their values helped determine which model had the smallest average prediction error. While RMSE and MAE focus on the absolute magnitude of errors, R^2 captures the proportion of variability that the model can account for in the target variable. In a classification analysis, precision measures the proportion of true positives against all positive predictions, recall assesses the model's ability to correctly identify all the actual positives, and the F1 Score provides a balance between precision and recall, serving as a harmonic mean of the two.

RESULTS & DISCUSSION

The comparative analysis of the random forest, decision tree, linear regression, and SVR models revealed their respective strengths and weaknesses in predicting parking occupancy in a college campus garage.

Regression analysis

The machine learning models were analyzed, using hourly occupancy as a continuous variable, and the results highlighted the differences in performance and suitability of each model. Figure 4 presents the R^2 score of the models before pre-processing the hourly occupancy.

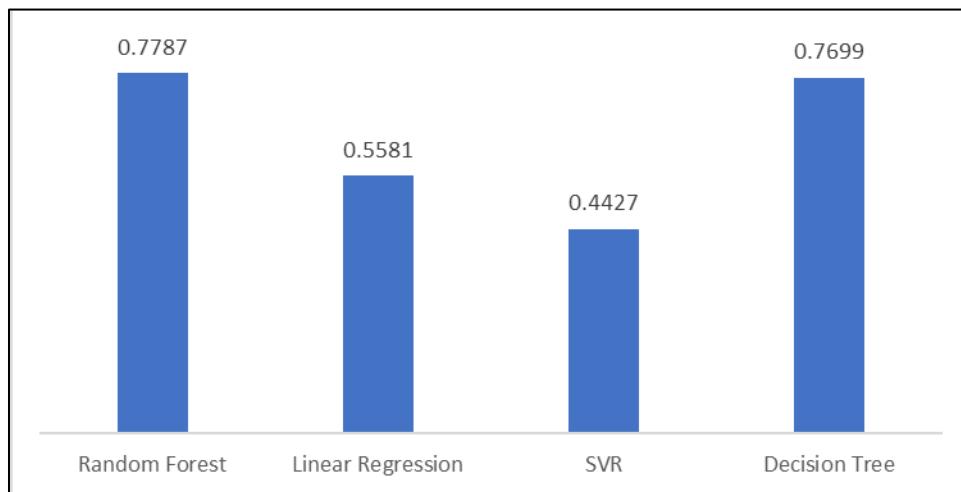


Figure 4. R^2 score of models

In the analysis of the models, the random forest model emerged as the standout performer for predicting hourly parking occupancy, thus highlighting its robustness in capturing the complexities of parking occupancy data. It exhibited an MAE of 6.26, indicating its high accuracy in predictions with minimal deviation from actual values; its RMSE of 11.3506 underscored its precision in predictions, showing minimal variance in errors; and its impressive R^2 score of 0.7787 signified its ability to explain approximately 78% of the variances in the dependent variable. The decision tree model was closely aligned with the random forest model and also displayed a commendable performance. Its slightly higher MAE of 6.3054 indicates strong predictive accuracy, a comparable RMSE suggests its effectiveness in making precise predictions, and the R^2 score of 0.7699 emphasizes its ability to explain a substantial portion of the data's variances.

Conversely, the linear regression model exhibited a weaker performance with a notably higher MAE of 12.63, which indicated less precise predictions. This was further supported by a higher RMSE value of 16.038, suggesting greater variance and larger errors in its predictions. The R^2 score of 0.558 was lower than that of the tree-based models, indicating that it is less effective for capturing the variability of parking occupancy data. The SVR model displayed a unique set of results. Its MAE of 11.1, though lower than that of linear regression, was still higher than the tree-based models, indicating moderate precision; however, its 18.01 RMSE was

higher than all of the models, implying significant variance in its predictive errors. Its R^2 score of 0.4427 indicates that it struggles to adequately capture underlying patterns in the data for this application.

Classification analysis

The bar chart presented in Figure 5 shows comparisons of the performances of four machine learning models (random forest, linear regression, SVM, and decision tree), based on four key metrics: precision, recall, F1 score, and accuracy.

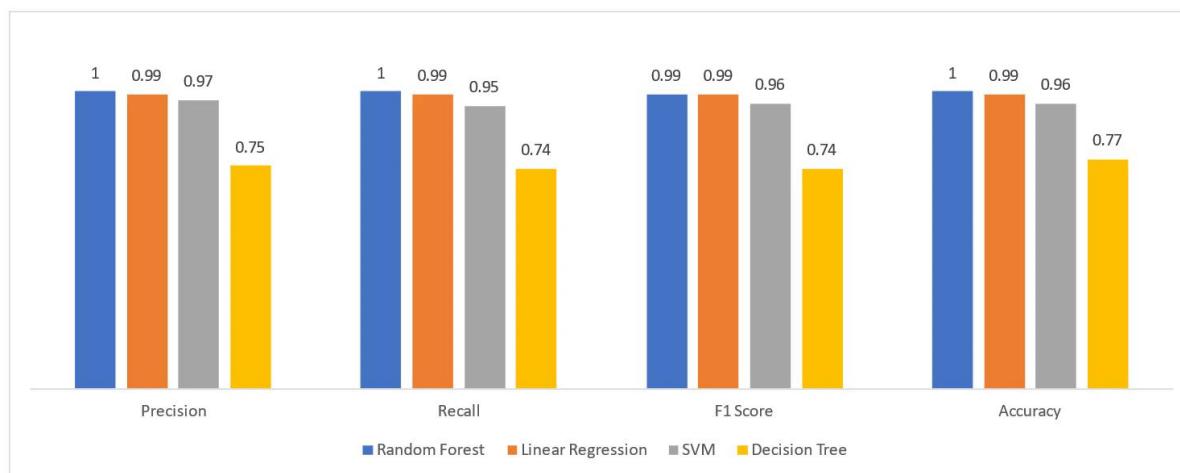


Figure 5. Results of prediction models

In terms of precision, which measures the proportion of true positive predictions out of all positive predictions made, random forest achieved a perfect score of 1.00. Linear regression followed closely behind with a score of 0.99, indicating that it also makes highly precise predictions. SVM's score was slightly lower at 0.97, and decision tree's was the lowest at 0.75, which suggests that it may generate more false positives than the other models.

All the models performed exceptionally well for recall, which assesses the proportion of actual positives that are correctly identified. Both random forest and linear regression had perfect recall scores of 1.00, meaning they correctly identified all the actual positives. SVM also performed well with a recall of 0.99, and the decision tree's score, although the lowest among the four, was still high at 0.95.

Accuracy measures the overall correctness of the model, and random forest and linear regression both scored 0.99, suggesting that they accurately predict the majority of instances. SVM and decision tree scores were slightly lower at 0.96 and 0.77, respectively, but it should be noted that although decision tree is less accurate than the others, an accuracy of 0.77 means that it correctly predicts more than three-quarters of the instances.

The F1 Score is the harmonic mean of precision and recall and provides a single measure of a test's accuracy. Random forest and linear regression again led with a perfect score of 1.00, denoting an excellent balance between precision and recall. SVM's F1 score of 0.98 was slightly lower but still indicates a strong performance. The decision tree's score of 0.74 is respectable but indicates room for improvement in balancing precision and recall.

To summarize, random forest and linear regression demonstrated superior performance across all metrics, indicating that they are highly reliable models for the dataset in question. SVM also showed strong performance, particularly in recall and the F1 Score. The decision tree model, while lagging behind the others, still presents a respectable performance, especially in recall, but may benefit from further fine tuning to improve its precision and overall accuracy.

CONCLUSION

This paper presents a comprehensive analysis of machine learning techniques for predicting parking occupancy in a college campus garage. A comparative analysis was conducted of four models (random forest, decision tree, SVM, and linear regression) by gathering and evaluating data and assessing the performance and precision of each model. The evaluation utilized metrics like MAE, RMSE, and R2 for regression analysis, along with accuracy, precision, recall, and the F1 score for classification analysis.

The in-depth evaluation of both regression and classification models provided substantial insights. The random forest model's performance was outstanding, as it achieved perfect or near-perfect scores in both classification metrics and performance in regression analysis. The linear regression model showed robust results in classification, but its performance in regression analysis was not as strong. SVM displayed consistent results in classification but ranked lowest in regression analysis, and the decision tree model, while having the least impressive classification metrics, demonstrated a strong performance in regression analysis.

It is clear from integrating these insights that the random forest model's robust performance in both classification and regression analysis establishes it as the premier choice for predicting parking occupancy. This dual strength allows for both detailed, number-specific predictions and categorizations that serve the diverse needs of real-time parking management, enabling users in making quick assessments and facility managers in effective strategic planning and resource distribution.

While extreme care was taken in conducting this comprehensive analysis of parking occupancy predictions, it is important to acknowledge the limitations of the study. This research focused on a specific urban environment, a parking garage, and the results may not seamlessly apply to different parking scenarios with unique characteristics. Therefore, applying the random forest model developed in this study to other contexts should be done with caution. Future research could expand on these findings by incorporating real-time data, adjusting models dynamically, and implementing predictive models in practical parking management systems. Further enhancements in model accuracy and prediction could be achieved through more detailed feature selection, refining hyperparameters, and adding more contextual variables.

ACKNOWLEDGEMENT

We gratefully acknowledge the support and generosity of the North Central Texas Council of Governments (NCTCOG), without which the present study could not have been completed.

REFERENCES

- Anwar, T., Asadullah, S., Azzam, M., and Abbas, K. 2019. Comparison of different machine learning models for parking occupancy prediction. In *2019 International Conference on Frontiers of Information Technology (FIT)* (pp. 59-64). IEEE.

- Caicedo, F., Blazquez, C., and Miranda, P. 2012. Prediction of parking space availability in real time. *Expert Systems with Applications*, 39(8), 7281-7290.
- Caicedo, F., Robuste, F., and Lopez-Pita, A. 2006. Parking management and modeling of car park patron behavior in underground facilities. *Transportation Research Record: Journal of the Transportation Research Board*, 1956(1), 60-67.
- Channamallu, S. S., Kermanshachi, S., and Pamidimukkala, A. 2023. Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles. In *International Conference on Transportation and Development 2023* (pp. 39-50).
- Channamallu, S. S., Kermanshachi, S., Rosenberger, J. M., and Pamidimukkala, A. 2023a. A review of smart parking systems. *Transportation Research Procedia*, 73, 289-296.
- Channamallu, S. S., Kermanshachi, S., Rosenberger, J. M., and Pamidimukkala, A. 2023b. Parking occupancy prediction and analysis - a comprehensive study. *Transportation Research Procedia*, 73, 297-304.
- Channamallu, S. S., Padavala, V. K., Kermanshachi, S., Rosenberger, J. M., and Pamidimukkala, A. 2023c. Examining parking occupancy prediction models: a comparative analysis. *Transportation Research Procedia*, 73, 281-288.
- Dey, K., and Nath, B. 2019. Prediction of parking occupancy using decision tree and support vector machine. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 131-134). IEEE.
- Etminani-Ghasradashti, R., Hladik, G., Kermanshachi, S., Rosenberger, J. M., Arif Khan, M., and Foss, A. 2022. Exploring shared travel behavior of university students. *Transportation planning and technology*, 46(1), 22-44.
- Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2023. Exploring motivating factors and constraints of using and adoption of shared autonomous vehicles (SAVs). *Transportation Research Interdisciplinary Perspectives*, 18, 100794.
- Farooq, S., Khaksar, W., Ullah, M. A., and Yaqoob, I. 2019. Comparative analysis of machine learning algorithms for parking occupancy prediction. *Journal of Ambient Intelligence and Humanized Computing*, 10(4), 1343-1355.
- He, Q., Li, Z., Huang, C., Zhang, M., and Li, D. 2020. Comparison of machine learning algorithms for parking occupancy prediction in urban areas. *Journal of Advanced Transportation*, 2020, 1-12.
- Huang, H., Gartner, G., Krisp, J., Raubal, M., and Van De Weghe, N. 2018. "Location Based Services: Ongoing Evolution and Research Agenda." *Journal of Location Based Services* 12 (2): 63–93.
- INRIX Research. 2017. Searching for parking costs Americans \$73 Billion a year. Retrieved from <http://inrix.com/press-releases/parking-pain-us>.
- Khan, M. A., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2022a. Do ridesharing transportation services alleviate traffic crashes? A time series analysis. *Traffic injury prevention*, 23(6), 333-338.
- Khan, M., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022b. Identifying Usage of Shared Autonomous Vehicles (SAVs): Early Findings from a Pilot Project. In *Transportation Research Board 101st Annual Meeting* Washington, DC.
- Khan, M. A., Etminani-Ghasradashti, R., Shahmoradi, A., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022c. Integrating shared autonomous vehicles into existing transportation services: evidence from a paratransit service in Arlington, Texas. *International Journal of Civil Engineering*, 20(6), 601-618.

- Khan, M. A., Patel, R. K., Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., Hladik, G., and Foss, A. 2023a. Factors that determine a university community's satisfaction levels with public transit services. *Frontiers in Built Environment*, 9, 1125149.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2023b. Understanding Students' Satisfaction with University Transportation. In *International Conference on Transportation and Development 2023* (pp. 522-532).
- Kim, J. H., and Kim, J. Y. 2021. A comparative analysis of machine learning algorithms for parking occupancy prediction: Focusing on different time horizons. *Sustainability*, 13(10), 5557.
- Kotb, A., Shen, Y., and Huang, Y. 2017. Smart parking guidance, monitoring and reservations: a review *IEEE Intell. Transp. Syst. Mag.* 9(2) pp. 6–16, 2017.
- Lin, T., Rivano, H., and Le Mouel, F. 2017. A survey of smart parking solutions. *IEEE Transactions on Intelligent Transportation Systems*, 18(12), 3229-3253.
- Liu, Z., Li, Y., Liu, Y., and Cai, H. 2019. Parking occupancy prediction based on gradient boosting decision tree. In *2019 IEEE 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE)* (pp. 358-362). IEEE.
- Pamidimukkala, A., Kermanshachi, S., and Patel, R. 2023. An Exploratory Analysis of Crashes Involving Autonomous Vehicles. In *International Conference on Transportation and Development 2023* (pp. 343-350).
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Exploring willingness to use shared autonomous vehicles. *International Journal of Transportation Science and Technology*, 12(3), 765-778.
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022b. How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems, ASCE International Conference on Transportation & Development*, pp. 81-93.
- Patel, R. K., Pamidimukkala, A., Kermanshachi, S., and Etminani-Ghasrodeshti, R. 2023a. Disaster preparedness and awareness among university students: a structural equation analysis. *International journal of environmental research and public health*, 20(5), 4447.
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Michael Rosenberger, J., and Foss, A. 2023b. Exploring Factors Affecting Shared Autonomous Vehicles Adoption: A Structural Equation Modeling Analysis, *Transportation Research Board 102 Annual Meeting* (pp. 23-00216).
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Michael Rosenberger, J., and Foss, A. 2023c. Users' and Nonusers' Attitudes and Perceptions of Shared Autonomous Vehicles: A Case Study in Arlington, Texas. In *International Conference on Transportation and Development 2023* (pp. 241-252).
- Sester, M. 2020. "Analysis of Mobility Data – A Focus on Mobile Mapping Systems." *Geo-spatial Information Science* 23 (1): 68–74.
- Shi, W., Chen, X., Zhang, Y., Guo, R., and Xiong, Y. 2021. Comparative analysis of machine learning algorithms for parking occupancy prediction in smart city. *Wireless Communications and Mobile Computing*, 2021, 5551553.
- Shoup, D. C. 2006. Cruising for parking. *Transport Policy*, 13(6), 479-486.
- Srinivasan, S., Soundararajan, R., Venkata Chalapathy, K., and Balasubramanian, R. 2021. Comparative analysis of machine learning techniques for parking occupancy prediction. *Journal of Ambient Intelligence and Humanized Computing*, 12(5), 5789-5804.

- Subramanya, K., Kermanshachi, S., and Patel, R. K. 2022. The Future of Highway and Bridge Construction: Digital Project Delivery Using Integrated Advanced Technologies. In *International Conference on Transportation and Development 2022* (pp. 14-25).
- Sun, Y., Wang, Y., Chen, L., and Gao, Y. 2019. Short-term parking occupancy prediction based on decision tree and random forest. In *Proceedings of the 2019 3rd International Conference on Industrial and Business Engineering* (pp. 135-139). ACM.
- Wu, F., Du, Z., Zhang, J., Zhang, S., and Wang, Y. 2017. Parking occupancy prediction based on decision tree and random forest. In *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)* (pp. 234-238). IEEE.
- Wu, J., Yang, B., and Zhu, D. 2020. A comparative study of machine learning algorithms for parking occupancy prediction. *IEEE Access*, 8, 104431-104439.
- Yalcin, M. E., and Zeydan, M. 2016. Comparative analysis of machine learning algorithms for parking occupancy prediction. *Expert Systems with Applications*, 63, 249-259.
- Yang, Y., Sun, S., Qiao, Z., and Zhu, H. 2019. Parking Services in Smart Cities: A Comprehensive Survey. *IEEE Access*, 7, 64271-64292.
- Yang, S., Ma, W., Pi, X., and Qian, S. 2019a. A deep learning approach to real-time parking occupancy prediction in transportation networks incorporating multiple spatiotemporal data sources. *Transportation Research Part C: Emerging Technologies*, 107, 248–265.
- Zhang, K., Zheng, S., Zhao, S., and Li, P. 2021. Comparative study of machine learning algorithms for parking occupancy prediction: A case study in Beijing. *Sustainable Cities and Society*, 75, 103287.
- Zheng, Y., Rajasegarar, S., and Leckie, C. 2015. Parking availability prediction for sensor-enabled car parks in smart cities, in: *Proceedings of the Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, IEEE, 2015, pp. 1–6.

Ego-Centric Pedestrian Trajectory Prediction Considering Camera Motion Parameters

Yufei Ji¹; Pei Liu²; Nanfang Zheng³; Siwen Liu⁴; and Ziyuan Pu⁵

¹School of Transportation, Southeast Univ. Email: 213201789@seu.edu.cn

²School of Transportation, Southeast Univ. Email: liupeiluna@foxmail.com

³School of Transportation, Southeast Univ. Email: 213202292@seu.edu.cn

⁴School of Transportation, Southeast Univ. Email: 213202087@seu.edu.cn

⁵School of Transportation, Southeast Univ. Email: ziyuan.pu@monash.edu

ABSTRACT

Accurate pedestrian trajectory prediction is crucial for enhancing pedestrian and autonomous vehicle safety. In comparison to the bird-eye view, ego-centric trajectory prediction is more challenging because the ego-motion of the camera can dynamically alter the scale of pedestrian bounding boxes. As the distance between the vehicle and pedestrians decreases, pedestrian bounding boxes are magnified, subsequently influencing the predicted pedestrian trajectories within the image plane. To address this challenge, the paper proposes a deep learning network based on LSTM encoder-decoder architecture that leverages trajectory, ego-vehicle motion information, and local-visual context as the additional input information. Specifically, the motion parameters of the in-vehicle camera are introduced to represent self-vehicle motion information at a coarse granularity level, reducing the impact of noise when representing self-vehicle speed by optical flow. Meanwhile, due to the disproportionate impact of pedestrians at different distances on the optimization process, the loss function is normalized using the actual width and height of the pedestrian's bounding box, effectively reducing the weight assigned to pedestrians in proximity during optimization. The attention mechanism is utilized to better capture the long-term temporal variations. Several typical pedestrian trajectory prediction benchmark datasets are used for the algorithm validation, such as JAAD. According to the results, the proposed method outperformed the selected state-of-the-art algorithms with about 19% improvements in trajectory prediction accuracy.

INTRODUCTION

The rapid development of autonomous driving highlights the significance of pedestrian trajectory prediction. Accurately anticipating the future trajectory of pedestrians plays a crucial role in providing valuable insights for vehicle path planning and collision prevention with pedestrians. Most current researches focus on predicting pedestrian trajectories from a bird-eye view (Alahi et al., 2016; Gupta et al., 2018) instead of ego-centric view (Bhattacharyya et al., 2018; Rasouli et al., 2019). Compared to the bird-eye view, data from the ego-centric view is easily gathered from front-facing cameras mounted on vehicles and can be applied in an autonomous driving setting. However, because of the optical axis of camera being parallel to the ground, the captured images exhibit a perspective effect, causing discrepancies in object displacement compared to their actual displacement in the real world (Yagi et al., 2018). This introduces challenges in ego-centric pedestrian trajectory prediction.

A recent study (Rasouli et al., 2019) predicts the future pedestrian trajectories from an ego-centric view. And it only involves simple learning of pedestrian and vehicle features, with the

feature of historical trajectories being directly connected to other factors such as vehicle speed. Additionally, few studies (Czech et al., 2023; Girase et al., 2021) consider the issue of pedestrian scale from an ego-centric view, neglecting changes in pedestrian bounding box scale and position when vehicles approach. However, Given the complex interaction between vehicles and pedestrians, observing vehicle motion, local visual information, and bounding boxes scale about pedestrians, etc., contributes to more precise trajectory prediction.

In the paper, we propose a deep learning network structure based on the LSTM encoder-decoder, which jointly encodes pedestrian historical trajectory, self-vehicle motion information, and local-visual context. Specifically, the motion of the vehicle influences pedestrian behavior; for instance, pedestrians may quickly pass when the approaching vehicle decelerates. Therefore, we introduce motion parameters from the onboard camera to represent the self-vehicle motion information at a coarse granularity, reducing the impact of noise when using optical flow to represent vehicle speed. Additionally, due to the scale variation of pedestrians, the influence of pedestrians at different distances on the training process is disproportionate. Hence, we redesign the loss function—normalizing it using the actual width and height of the pedestrian bounding boxes, effectively reducing the weight assigned to close-distance pedestrians during the optimization process. Eventually, we evaluate the proposed method's performance using publicly available pedestrian behavior benchmark dataset, such as JAAD, demonstrating improvement over selected state-of-the-art algorithms in predicting pedestrian trajectories.

LITERATURE REVIEW

Pedestrian trajectory prediction modeling methods: Existing pedestrian trajectory prediction methods can be broadly categorized into two types: expert-based and data-driven approaches (Golchoubian et al., 2023). In terms of prediction accuracy, data-driven approaches significantly outperform expert-based models. This paper primarily focuses on deep learning models for pedestrian trajectory prediction.

Due to the temporal nature of pedestrian trajectories, many models are based on Recurrent Neural Networks (RNNs), which can capture dependencies between time series data. Various types of RNNs are employed, such as Gated Recurrent Units (GRU) (Girase et al., 2021), or Long Short-Term Memory (LSTM) (Chen et al., 2022; Rasouli et al., 2019). For improved trajectory prediction, a common approach is the use of an encoder-decoder architecture (Cai et al., 2022; Rasouli et al., 2019; Ridel et al., 2019). In this architecture, the historical pedestrian trajectory and other related features are encoded as hidden states within the encoder module. The decoder module then generates predictions for the future pedestrian trajectory based on these encoded states. While LSTM exhibits powerful memory and forgetting capabilities, it is limited in its ability to remember states from distant past moments, as it tends to focus more on the immediate preceding moment. In this paper, we utilize an LSTM encoder-decoder with an attention mechanism to establish the correlation between the trajectory at the current moment and those from past moments.

Pedestrian-vehicle interaction: For papers focused on ego-centric trajectory prediction, pedestrian interaction with the ego-vehicle is the primary focus. These interactions are typically modelled by inputting some features of the ego-vehicle motion and pedestrians (e.g., the head orientation and posture) (Czech et al., 2023; Ridel et al., 2019).

One of the common features is the speed of the ego-vehicle (Rasouli et al., 2019, 2021). Additionally, some approaches (Bhattacharyya et al., 2018; Czech et al., 2023) employ an

encoder-decoder network to predict ego-vehicle speed of the next moment. The output of the ego-motion predictor is then used to forecast pedestrian trajectories. Considering that some ego-centric view benchmark datasets (Rasouli et al., 2017) lack information on ego-vehicle speed, some studies (Quan et al., 2021; Yin et al., 2021) resort to using optical flow and monocular depth neural networks to estimate the ego-vehicle speed. This approach captures high-dynamic motion information but is susceptible to image noise. However, our method considers the complex interaction between pedestrian and the ego-vehicle based on the motion parameters between consecutive camera frames and the local visual features of pedestrians.

Pedestrian scale issue: Pedestrian trajectory prediction is typically conducted from a bird-eye view or an ego-centric view. Compared to the bird-eye view, predicting pedestrian trajectories from an ego-centric view is generally more challenging because of the scale issue with traffic agents. Specifically, vehicle motion not only affects pedestrian behavioral decisions, but also alters the scale and relative position of the pedestrian bounding box. As the distance between the pedestrian and the vehicle decreases, the pedestrian bounding box is enlarged, thereby impacting the precise prediction of pedestrian trajectories at the image level.

To address the aforementioned challenges, (Yagi et al., 2018) captures the scale of individuals and implicitly predicts pedestrian trajectories. Holistic LSTM (Quan et al., 2021) integrates vehicle speed variations into the LSTM output gate, dynamically adjusting the scale of pedestrian bounding box. (Neumann & Vedaldi, 2021) explicitly separates the real-world motion of pedestrians from the ego-motion of vehicles, enabling the observation and prediction of intrinsic pedestrian motion in standardized views across multiple frames. A new metric, termed “scaled distance error”, is proposed by him (Rasouli, 2023) to better rank models under each scenario. In the paper, we scale the loss function based on the size of the annotated pedestrian bounding boxes to mitigate the impact of the pedestrian scale issue.

METHODOLOGY

Problem Formulation

Figure 1 illustrates an example of our task, where a pedestrian is crossing the road, and our goal is to predict the future bounding boxes of pedestrians in the next n frames given the bounding boxes of pedestrians in the past m frames. The bounding box of the i th pedestrian at time step t can be described using the pixel coordinates of the top-left and bottom-right points: $l_i^t = \{(x_{tl}, y_{tl}), (x_{br}, y_{br})\}$, where (x_{tl}, y_{tl}) is the coordinate of the top-left point, and (x_{br}, y_{br}) is the coordinate of the bottom-right point. Formally, we formulate pedestrian behavior prediction as an optimization process in which the goal is to learn distribution $p(L_p | L_0, E_0, C_0)$, where $L_p = \{l_i^{t+1}, l_i^{t+2}, \dots, l_i^{t+n}\}$ are the future sequences of pedestrians, $L_0 = \{l_i^{t-m+1}, l_i^{t-m+2}, \dots, l_i^t\}$ are the observed sequences of pedestrians, $E_0 = \{e_i^{t-m+1}, e_i^{t-m+2}, \dots, e_i^t\}$ are the observed ego-vehicle motion parameters and $C_0 = \{c_i^{t-m+1}, c_i^{t-m+2}, \dots, c_i^t\}$ are the sequence of local-visual context.

Architecture

We propose a deep learning network structure based on the LSTM encoder-decoder, as depicted in Figure 2. Next, we discuss the different modules of the proposed model:

Ego-Motion Parameters: Considering the intricate interaction between pedestrians and vehicles, the self-motion of vehicles may impact the movement of the target pedestrian. We represent vehicle self-motion using rotation and translation between consecutive frames. Rotation is represented by quaternions ($q_t \in \mathbb{R}^4$), and translation is represented using a three-dimensional vector ($v_t \in \mathbb{R}^3$), both denoting local motion in the camera coordinate system from frame $t-1$ to frame t . We concatenate rotation vector q_t and translation vector v_t to form a 7-dimensional feature vector $e_t (e_t \in \mathbb{R}^7)$ for each frame. Finally, these vectors are stacked to create an input sequence $E_0 = \{e_i^{t-m+1}, e_i^{t-m+2}, \dots, e_i^t\}$.

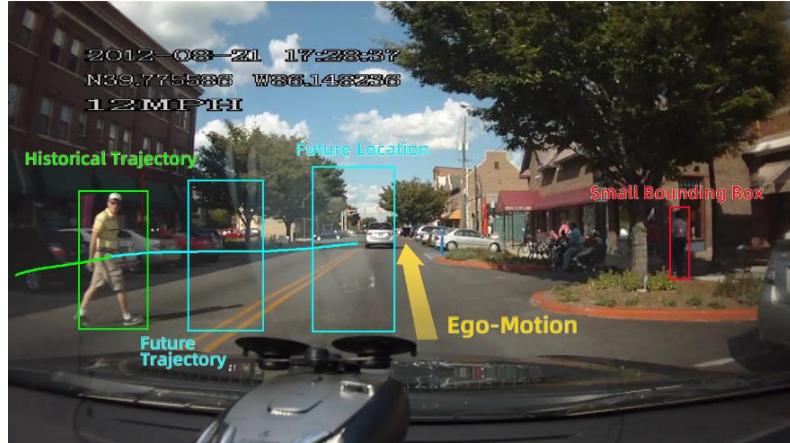


Figure 1: Illustration of pedestrian trajectory prediction task

Local-Visual Context: To integrate visual information about pedestrian appearance, pose, and the surrounding environment, we employ ResNet-50 (Koone, 2021) to extract local-visual context features. Specifically, we input the locally cropped region, which is twice the size of the pedestrian bounding box, into ResNet-50, generating a 2048-dimensional feature vector $c_t (c_t \in \mathbb{R}^{2048})$.

Context Representation: In order to reduce the influence between different data modalities, this module jointly encodes ego-motion parameters and local-visual context, projecting them into a common feature space through a four-layer LSTM network.

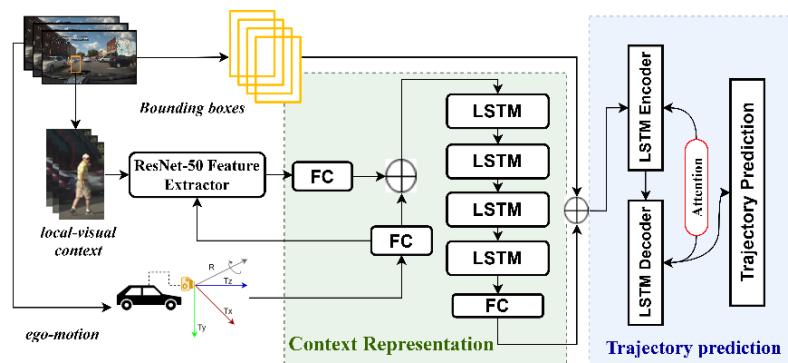


Figure 2: Framework of the pedestrian trajectory prediction

Trajectory Prediction: We use an LSTM encoder-decoder structure to output the future trajectory. The encoder takes as input the observed bounding boxes concatenated with the context representation. An attention mechanism identifies the most relevant frames from the observed data. The decoder connects to a fully connected layer, ultimately outputting the future coordinates of pedestrians.

Learning Objective

For learning objective, the most commonly employed metric is distance-based, such as the mean squared error (MSE) (Ridel et al., 2019; Yagi et al., 2018) or the average Euclidean distance (Wang, 2021) between predicted value and ground truth. As illustrated in Figure 1, pedestrians at different distances from the vehicle exhibit bounding box sizes that are not uniform. This discrepancy causes the larger-scale pedestrians dominating the overall loss function during training. To address the issue, we scale the distances based on the actual width and height of the pedestrian bounding boxes:

$$L_l = \frac{1}{N \times \tau} \sum_{i=1}^n \sum_{j=t}^{t+\tau} \left\| \frac{y_i^j - \hat{y}_i^j}{S} \right\|_1 \quad (1)$$

Where y represents the coordinates of the bounding box, S represents actual width and height.

EXPERIMENT DESIGN

Datasets

JAAD Dataset (Rasouli et al., 2017): This dataset is designed to understand pedestrian intent in various traffic scenarios. It comprises 346 video segments ranging from 5 to 15 seconds in duration, featuring 2.2k unique pedestrian samples and annotated with 337k bounding boxes. We follow the same train/test set division ratio as in (Rasouli et al., 2018).

Training Details

We employ a pre-trained ResNet-50 on ImageNet(Deng et al., 2009), replacing the final fully connected layer with an identity mapping layer to extract local visual context for pedestrians. We use LSTMs for all extractions, encoders, and decoders with hidden size of 256, L2 regularization of 0.001 and dropout of 0.3. The output size of fc layers in Context Features Extraction are set to 64.

For training, we use the Adam optimizer with an initial learning rate of 0.01 and a batch size of 256. We conduct training for 100 epochs and implement a learning rate reduction with a decay factor of 0.9 by exponential decay strategy.

Evaluation metrics and baselines

To facilitate a fair comparison, we adopt standardized metrics to evaluate the performance of our model in predicting pedestrian trajectories from an ego-centric view video. These metrics are outlined below:

- 1) Average displacement error (ADE): The average Euclidean distance between corresponding predicted center coordinates and ground truth across all frames,

$$\text{ADE} = \frac{\sum_{i=1}^N \sum_{j=t+1}^{\tau} \|y_i^j - \hat{y}_i^j\|_2}{N \times \tau} \quad (2)$$

- 2) Final displacement error (FDE): The Euclidean distance between the predicted final position and the truth final position in a sequence,

$$\text{FDE} = \frac{\sum_{i=1}^N \|y_i^{t+\tau} - \hat{y}_i^{t+\tau}\|_2}{N} \quad (3)$$

Where N is the number of pedestrian trajectories, τ is the prediction time length, and y and \hat{y} represent the ground truth and predicted value on the pixel level. On the JAAD dataset, in line with previous work (Rasouli et al., 2021), the prediction length is 1.0s. The values obtained after the convergence of the model in each experiment are considered as the result in the table.

We compare our proposed method with other methods that are trained with ego-centric view data. These methods are **Future Person Localization (FPL)** (Yagi et al., 2018), **Bayesian LSTM (B-LSTM)** (Bhattacharyya et al., 2018), **FOL** (Yao et al., 2019), and **PIE_{traj}** (Rasouli et al., 2019).

RESULTS

Comparison to State-of-the-Art

As shown in the Table 1, ours outperforms all other selected methods on the JAAD dataset. We improve the performance by up to 6% than PIE_{traj} on ADE (from 23.49 to 21.85). Also, is better than (by about 19%) PIE_{traj} on FDE (from 50.18 to 40.38).

Table 1: Prediction errors for different methods at a prediction length of one second

Method	ADE↓	FDE↓
FOL	61.39	126.97
FPL	42.24	86.13
B-LSTM	28.36	70.22
PIE _{traj}	23.49	50.18
ours	21.85	40.38

Influence of Different Inputs and Loss Functions

We conduct an ablation study on the JAAD dataset to investigate how ego-motion parameters, Local-visual context, and the standardized loss function contribute to the overall predictive performance.

Different information: The utilization of three types of information, namely historical trajectory, ego-motion parameters, and local-visual context, results in the creation of four distinct inputs. The experimental outcomes for these four inputs are presented in Table 2 (see Figure 3

for some examples). When historical trajectory and local-visual context are employed as inputs, the scores for ADE1.0s and FDE1.0s are respectively 23.18 and 42.21. Comparatively, there is no improvement in performance when solely using historical trajectory, indicating that local-visual context alone contributes minimally. Enhanced performance is observed when the proposed operation is applied to incorporate both historical trajectory and ego-motion parameters. Finally, with the integration of local-visual context, the ADE1.0s and FDE1.0s scores are reduced to 21.85 and 40.38, respectively. All the results emphasize the efficacy of the integrated ego-motion parameters and local-visual context in predicting pedestrian trajectories, with ego-motion parameters being more effective.

Table 2: Ablation results of different inputs

input	0.5s		1.0s	
	ADE↓	FDE↓	ADE↓	FDE↓
L_{in}	14.80	20.68	23.21	42.45
$L_{in} + C_{in}$	14.85	20.95	23.18	42.21
$L_{in} + E_{in}$	13.96	19.73	22.21	41.07
$L_{in} + C_{in} + E_{in}$	13.85	19.70	21.85	40.38

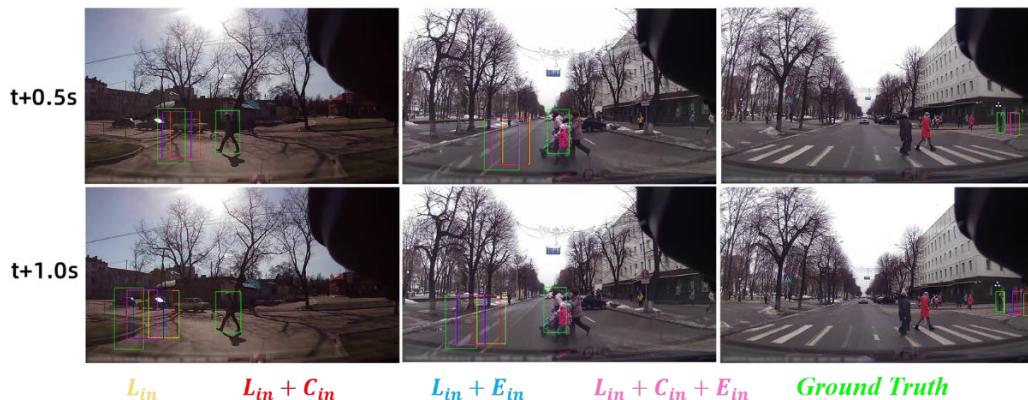


Figure 3: Qualitative results of the proposed method on JAAD showing the predicted bounding boxes using L_{in} (yellow), $L_{in} + C_{in}$ (red), $L_{in} + E_{in}$ (blue) and $L_{in} + C_{in} + E_{in}$ (pink) against ground truth (green).

Loss function: In Table 3, we train the model using Mean Squared Error (MSE) or proposed loss function based on scaled distance. With all the information ($L_{in} + C_{in} + E_{in}$) as input, all evaluation metrics for the scaled-distance loss function outperform MSE.

Table 3: Ablation results of different loss functions

input	loss function	0.5s		1.0s	
		ADE↓	FDE↓	ADE↓	FDE↓
$L_{in} + C_{in} + E_{in}$	MSE	14.42	20.13	22.39	41.03
	Scaled Distance	13.85	19.70	21.85	40.38

CONCLUSION

With the increasing research on autonomous driving vehicles, our goal is to enable these vehicles to navigate through future congested urban roads. As cameras serve as the primary perception tool for autonomous vehicles, the scale issue of pedestrians in video data collected by onboard cameras makes it challenging to accurately infer pedestrian positions. In this paper, we develop a deep learning network structure based on LSTM encoder-decoder and validate it on the benchmark dataset JAAD. Compared to other selected methods, our proposed model shows a significant improvement (reaching 19%) in predicting the final position of pedestrians. We also discuss the impact of different inputs and loss functions on model performance. Specifically, ego-motion parameters prove to be a more effective factor compared to local-visual context. Additionally, using a loss function based on scaled distance can enhance model prediction accuracy.

In the future, comprehensive experiments on the deep learning model can be conducted to compare their performance in more complex road environments, such as non-structured settings like parks or campuses. We will also explore incorporating pedestrian crossing intent to assist in predicting pedestrian trajectories. Designing a lightweight model for deployment on mobile phones or edge computing modules is a promising avenue for future work.

REFERENCES

- Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S. (2016). Social lstm: Human trajectory prediction in crowded spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 961-971).
- Bhattacharyya, A., Fritz, M., and Schiele, B. (2018). Long-term on-board prediction of people in traffic scenes under uncertainty. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4194-4202).
- Cai, Y., Dai, L., Wang, H., Chen, L., Li, Y., Sotelo, M. A., and Li, Z. (2021). Pedestrian motion trajectory prediction in intelligent driving from far shot first-person perspective video. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 5298-5313.
- Chen, T., Jing, T., Tian, R., Chen, Y., Domeyer, J., Toyoda, H., and Ding, Z. (2021). Psi: A pedestrian behavior dataset for socially intelligent autonomous car. arXiv preprint arXiv:2112.02604.
- Czech, P., Braun, M., Kreßel, U., and Yang, B. (2023). Behavior-aware pedestrian trajectory prediction in ego-centric camera views with spatio-temporal ego-motion estimation. *Machine Learning and Knowledge Extraction*, 5(3), 957-978.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.
- Girase, H., Gang, H., Malla, S., Li, J., Kanehara, A., Mangalam, K., and Choi, C. (2021). Loki: Long term and key intentions for trajectory prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9803-9812).
- Golchoubian, M., Ghafurian, M., Dautenhahn, K., and Azad, N. L. (2023). Pedestrian trajectory prediction in pedestrian-vehicle mixed environments: A systematic review. *IEEE Transactions on Intelligent Transportation Systems*.

- Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., and Alahi, A. (2018). Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2255-2264).
- Koonce, B., and Koonce, B. (2021). ResNet 50. *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, 63-72.
- Neumann, L., and Vedaldi, A. (2021). Pedestrian and ego-vehicle trajectory prediction from monocular camera. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10204-10212).
- Quan, R., Zhu, L., Wu, Y., and Yang, Y. (2021). Holistic LSTM for pedestrian trajectory prediction. *IEEE transactions on image processing*, 30, 3229-3239.
- Rasouli, A. (2023). A Novel Benchmarking Paradigm and a Scale-and Motion-Aware Model for Egocentric Pedestrian Trajectory Prediction. arXiv preprint arXiv:2310.10424.
- Rasouli, A., Kotseruba, I., Kunic, T., and Tsotsos, J. K. (2019). Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6262-6271).
- Rasouli, A., Kotseruba, I., and Tsotsos, J. K. (2017). Are they going to cross? a benchmark dataset and baseline for pedestrian crosswalk behavior. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 206-213).
- Rasouli, A., Rohani, M., and Luo, J. (2021). Bifold and semantic reasoning for pedestrian behavior prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 15600-15610).
- Ridel, D. A., Deo, N., Wolf, D., and Trivedi, M. (2019, June). Understanding pedestrian-vehicle interactions with vehicle mounted vision: An LSTM model and empirical analysis. In *2019 IEEE Intelligent Vehicles Symposium (IV)* (pp. 913-918). IEEE.
- Wang, Y. (2021, September). Ltn: Long-term network for long-term motion prediction. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)* (pp. 1845-1852). IEEE.
- Yagi, T., Mangalam, K., Yonetani, R., and Sato, Y. (2018). Future person localization in first-person videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7593-7602).
- Yao, Y., Xu, M., Choi, C., Crandall, D. J., Atkins, E. M., and Dariush, B. (2019, May). Egocentric vision-based future vehicle localization for intelligent driving assistance systems. In *2019 International Conference on Robotics and Automation (ICRA)* (pp. 9711-9717). IEEE.
- Yin, Z., Liu, R., Xiong, Z., and Yuan, Z. (2021, August). Multimodal Transformer Networks for Pedestrian Trajectory Prediction. In *IJCAI* (pp. 1259-1265).

Prototype Data Dashboard for Multi-Source Transportation and Community Health Data Analytics

Chengyue Wang¹; Talha Azfar²; and Ruimin Ke³

¹Dept. of Civil Engineering, Univ. of Texas at El Paso, El Paso, TX.

Email: cwang5@miners.utep.edu

²Dept. of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY.

Email: azfar@rpi.edu

³Dept. of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY.

Email: ker@rpi.edu

ABSTRACT

This study had a primary objective of creating a versatile and scalable data dashboard prototype that highlights transportation and community health data, using El Paso as a case study. The dashboard serves as a centralized hub for various data sources within the city, facilitating visual analysis to uncover the intricate connections between these two crucial domains. It's important to note that the primary aim was not to develop a final product for the city, but rather to establish a proof-of-concept, engage potential users, and conduct scenario analysis. The true impact of this study is expected to extend beyond the creation of the dashboard itself, as it has the potential to influence significant city-wide or regional policy changes based on the valuable insights derived from the comprehensive data analysis. This initiative aims to provide a robust tool for decision-makers and stakeholders to make informed choices that could ultimately benefit the community's transportation and health outcomes.

INTRODUCTION

City and county-wide data on transportation and community health are currently dispersed across various agency repositories, each with its own data format and resolution. This fragmentation hinders the analysis of cross-departmental or interdisciplinary indicator relationships. To address this, independent dashboards for transportation (WSDOT, 2023; Zuo et al., 2020), environment (Oberlin, 2023), and community health (EPTX, 2023) have been developed. This study aims to amalgamate, house, and synchronize spatial and temporal data from these three sectors into a unified dashboard.

A dashboard serves as a centralized data hub, streamlining the analysis of data spanning transportation, environment, and community health. It is more than just a repository; it combines open data sources with tools for data management, mapping, visualization, and modeling. Typically, a dashboard operates on three layers of data processing. The first layer focuses on data collection, preprocessing incoming data, and storing it in databases. The second layer is tasked with computing intermediate performance metrics or indicators, linking diverse data sets using SQL (Structured Query Language), and executing spatial-temporal modeling. The third layer is dedicated to application-specific tasks, such as visualization and analysis.

The study's goal is to create a customizable and scalable prototype data dashboard, facilitating the visualization of data related to transportation and community health. Using El Paso, Texas as a case study, the research team demonstrated how the prototype dashboard could

be utilized to intersect different data sets, enabling the analysis of the interplay among transportation and community health sectors.

LITERATURE REVIEW

A city dashboard is a sophisticated tool that aggregates, examines, and illustrates key performance indicators of a city's or region's infrastructure and services. It serves to enlighten both the public and policymakers about the current state and efficiency of vital resources such as transportation, healthcare, energy, and water, among others. Presented through an interactive and user-friendly graphical interface, these dashboards often utilize maps, classifying them under geospatial dashboards. They are instrumental in delivering real-time data, with some applications requiring updates as frequently as every 24 hours to maintain their relevance and accuracy.

According to Farmanbar and Rong (2020), city dashboards can be categorized into two distinct types: city-level and analytical dashboards. City-level dashboards provide a comprehensive view of a city's overarching performance, intended for a general audience. In contrast, analytical dashboards are designed for in-depth, specialized analysis, enabling users to delve into data across various layers, catering to analysts and decision-makers.

Globally, major cities have developed their dashboards to address diverse applications. However, not all platforms labeled as dashboards meet the criteria of showcasing live, operational data. The criteria for inclusion in an overview of such dashboards, as summarized in the referenced Table 1, are stringent: they must display live data, and should not be mistaken for mere landing pages of broader smart city initiatives. An often-cited exemplar of a city dashboard is Rio de Janeiro, Brazil's Operation Center, although, as of Mattern's 2015 publication, its online presence was not accessible.

Zuo et al. (2020) crafted dashboards for New York City and Seattle to explore the impact of COVID-19 social distancing policies on various traffic metrics including volume, travel time, average speed, crash occurrences, and bikeshare usage, using data provided by local agencies. They also incorporated additional datasets pertaining to pedestrian density and social distancing adherence. The dashboards were utilized to analyze changes in social distancing compliance, traffic speed variations, crash severity, and bicycle usage during the pandemic. The dashboards developed by Zuo et al. feature a tripartite structure comprising:

1. **Data Integration and Access Layer:** This foundational layer consolidates and facilitates access to the collected data.
2. **Data Mining and Cloud Computing Layer:** At this intermediate level, the data undergoes processing and analysis, leveraging cloud computing resources.
3. **All-in-One Data Dashboard and Analytics Layer:** The topmost layer presents a comprehensive view of the analytics, offering users a unified interface for data interpretation.

Farmanbar and Rong (2020) introduced the Triangulum City Dashboard, designed to assist residents and policymakers in discerning the interconnections among mobility, energy usage, and environmental stewardship to inform policy decisions aimed at enhancing quality of life. This cloud-based platform also operates on a three-tier architecture:

- **Data Layer:** Engages in the acquisition, storage, and preliminary handling of data.
- **Application Layer:** Processes the data from the lower layer to extract actionable insights.
- **Presentation Layer:** Displays the processed data in a user-friendly format, facilitating an intuitive understanding for end-users.

Table 1. Operational Dashboards in Selected Cities

City	Dashboard name	Reference	Application
Baltimore, U.S.	CitiStat	City of Baltimore (2022), Gullino (2009)	Community development, environment, public safety
Boston, U.S.	Boston COVID-19	City of Boston (2023a)	COVID-19 cases and vaccinations
Boston, U.S.	Imagine 2030 Boston Metrics Dashboard	City of Boston (2023b)	A variety of city's KPIs
Charlotte, U.S.	Capital Project Dashboard	City of Charlotte (2022)	Building projects
Dublin, Ireland	Building City Dashboard	Maynooth University (2023)	City infrastructures' KPIs
Las Vegas, U.S.	Community Dashboard	City of Las Vegas (2023)	A variety of city's KPIs
London, U.K.	CityDashboard	CASA (2023)	A variety of city's KPIs
Melbourne, Australia	Melbourne Dashboard	Crisis Dashboard (2021)	Traffic, weather
New York, NY	NYC Open Data	City of New York (2022)	Various city services
New York City and Seattle, U.S.	none	Zuo et al. (2020)	Transportation and sociability
Singapore	LIVE Singapore	Ratti et al. (n.d.)	Communication, environment, transportation
Washington D.C., U.S.	DCStat	ADT (2006)	Property maintenance, public safety

The data layer integrates five distinct datasets: traffic, electricity, parking management, energy, and cargo e-bike data. Farmanbar and Rong also navigate through challenges such as varying data resolutions and formats, incomplete datasets due to sensor limitations, and the need to uphold privacy laws.

DATA COLLECTION

The foundational layer of the dashboard comprises demographic data, predominantly sourced from the U.S. Census website (Census, 2023). The most up-to-date data available was from the 2020 American Community Survey. As the official results of this survey were not yet public, estimates spanning six years, from 2016 to 2021, were utilized. The U.S. Census provides data at the census tract level, and for the City of El Paso, Texas, this equates to 128 distinct tracts. The acquired data included estimates of population size and median household income for each tract over the years 2016 through 2020. This data was subsequently consolidated by zip code. In the City of El Paso, there are 23 zip codes, which are identified by numbers ranging from 79901 to 79936, albeit not sequentially.

The crash data set encompasses 35,422 recorded traffic collisions that took place within the City of El Paso, Texas, from January 1, 2016, to October 18, 2021. These records originated

from CR-3 crash report forms, completed by police officers at the scene of each incident, and were subsequently uploaded to the Texas Department of Transportation (TxDOT) Crash Report Information System (CRIS). This data compilation is a crucial element of the City of El Paso's Intersection Safety Improvement Plan (Cheu et al., 2021). It specifically includes crashes that occurred at city-managed street intersections and excludes incidents on highways under TxDOT jurisdiction, non-intersection-related accidents, and those on private properties. The dataset was furnished in a Microsoft Excel CSV format, with each row detailing the attributes of an individual crash.

The research team acquired a comprehensive set of vehicle trajectory data from Wejo (2023), capturing vehicular movements across El Paso County, Texas. This dataset spans from October 25, 2021, to November 21, 2021, and records round-the-clock vehicle activity, amassing a total of 968,715,648 individual records. Each record documents various attributes of a vehicle, with data points collected at three-second intervals. Wejo provided this extensive dataset via the Amazon Web Service S3 database in an efficiently compressed JSON format. To verify the accuracy of the data's geolocation, the points were plotted on a street map.

In addition to traffic data, the dashboard integrates community health metrics, particularly those pertaining to COVID-19 within the City of El Paso, Texas. The local health data, available to the public on the El Paso Strong website (EPTX, 2023), align with the standard statistics requested by the CDC. This includes the community risk level, new and cumulative counts of positive test results, new and cumulative deaths attributed to COVID-19, along with the number of hospitalizations, ICU admissions, and ventilator utilizations.

PROTOTYPE DATA DASHBOARD FRAMEWORK

This paper presents a web-based dashboard prototype designed to perform cross-layer analyses of extensive datasets, facilitating the tracking and identification of trends within transportation, environment, and community health in the City of El Paso, Texas. The proposed dashboard's architecture, depicted in Figure 1, is structured into three principal layers: data collection, modeling, and application, aligned in a bottom-up configuration.

The foundational data collection layer is tasked with the aggregation, preprocessing, and storage of data from diverse infrastructural systems. This includes data quality assurance measures such as outlier removal, deduplication, and gap filling to ensure the data's consistency, accuracy, completeness, and timeliness. This refined data is then stored across various layers, utilizing cloud-based solutions or local drives for the prototype.

The intermediary modeling layer serves as the nexus between the raw data and the user interface. It processes data from the collection layer, applying algorithms and models to derive key metrics like incident frequencies, safety indices, congestion levels, vehicle trajectories, and health indices. Integration within this layer is pivotal, merging multi-source data into formats suitable for modeling. This layer can employ pre-programmed or user-interactive models, leveraging predefined models from Python libraries like Scikit-Learn or custom models tailored to specific requirements, ensuring efficient resource utilization and minimal response latency.

At the forefront is the application layer, which acts as the interactive visual interface of the dashboard. It offers users the capability to navigate and manipulate displayed information through a one-page or drill-down design, constructed on a web-based application framework. Depending on the programming language, this framework may vary, such as Vaadin for Java or Streamlit for Python. It incorporates common interface elements like buttons, dropdowns, tabs,

checkboxes, date selectors, sliders, and text boxes, with a unique layout for each function driven by the modeling layer algorithms. Interactivity between modules, along with the integration of charts and maps for data representation, is crucial. Tools like Google Map, Open Street Map, and the Folium library in Python are instrumental in visualizing geospatial data. The design of this layer is focused on the optimal visualization of model outputs, including map overlays, heatmaps, and a variety of charts, with the flexibility for data to be displayed on the map, in pop-up windows, or as downloadable files.

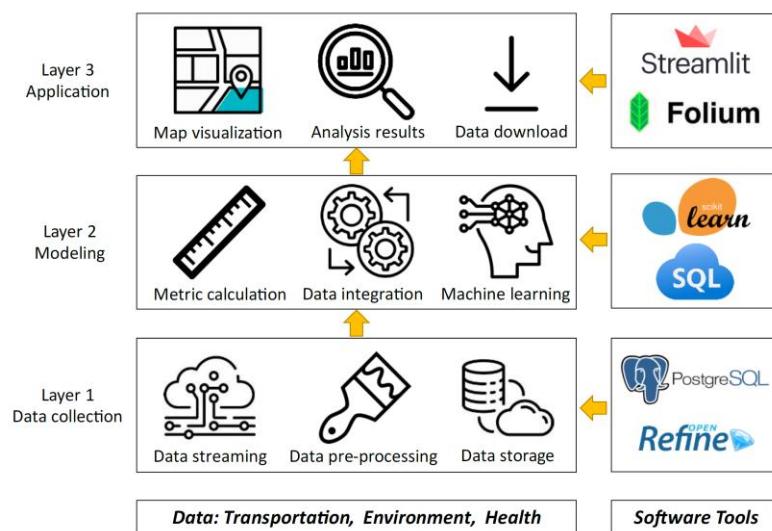


Figure 1. Concept Diagram of the Dashboard

PROTOTYPE DATA DASHBOARD DEVELOPMENT

The prototype data dashboard is now live and can be visited at <https://chengyuew-el-paso-dashboard-case-study-dashboard-5ukgqq.streamlit.app/>. This section will guide you through the dashboard's essential functionalities. Upon arrival, the landing page greets visitors with a welcome message, illustrated in Figure 2, and offers three informative tabs: About, People, and Disclaimer. The About tab provides insight into the dashboard's background and aims. The People tab credits the research team behind the study, while the Disclaimer tab mentions the funding source and clarifies any data usage and interpretation constraints.

To dive into the dashboard, users are prompted to double-click the [Double Click Enter El Paso Dashboard] button at the landing page's foot. This action transitions the user to the main interface of the dashboard, which maintains a consistent layout across all data layers. The interface is strategically partitioned into three sections:

- The left panel, known as the Main Menu, outlines the four data layers available: Demographic, Crash Data, Traffic Data, and Health Data. Users have the option to focus on one data layer at a time. The lower segment of this panel presents dropdown menus for selecting or inputting attributes relevant to the chosen data layer.
- The right-top panel is dedicated to the dashboard's title block, "El Paso Data Dashboard," accompanied by the subtitle "Transportation, Environment and Community Health," establishing the dashboard's thematic focus.

- The right-bottom panel, the primary display area, showcases the maps or visualizations corresponding to the selections made in the left panel, enabling users to visualize the data objects or attributes in detail.

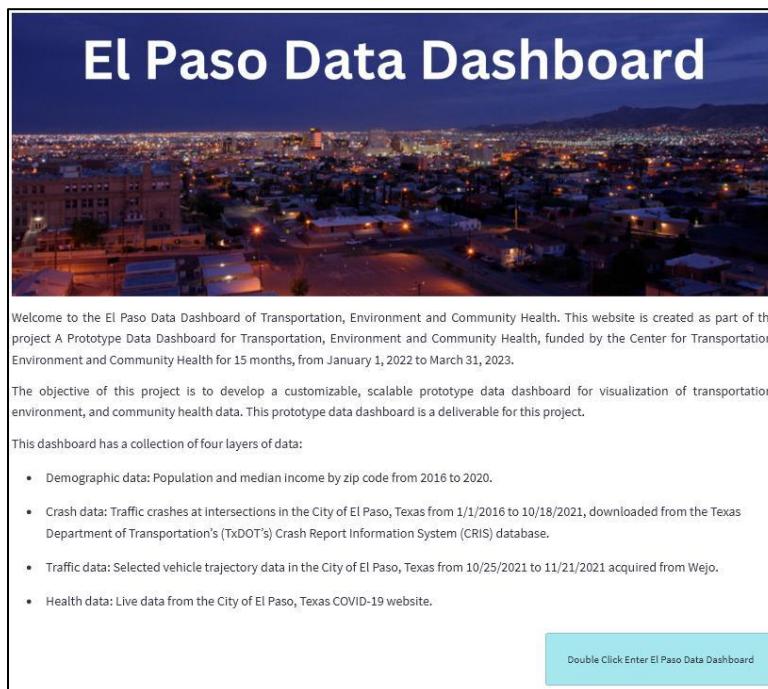


Figure 2. The landing page of the dashboard

Figures 3 contains screenshots as examples of the Demographic layer. The base map divides the City of El Paso by zip code. Users can select to display either Population or Median Household Income by a calendar year from 2016 to 2020. The data of the selected Attribute and Year are plotted by zip code on a map. Figure 4 (left) displays the Population in 2016. Figure 3 (right) displays the Median Household Income in 2016.

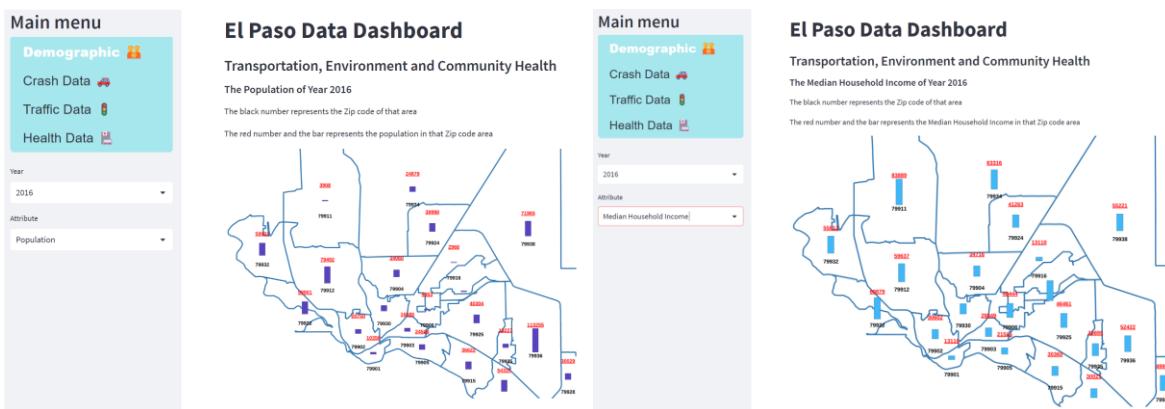


Figure 3. Demographic – population and median household income by zip code

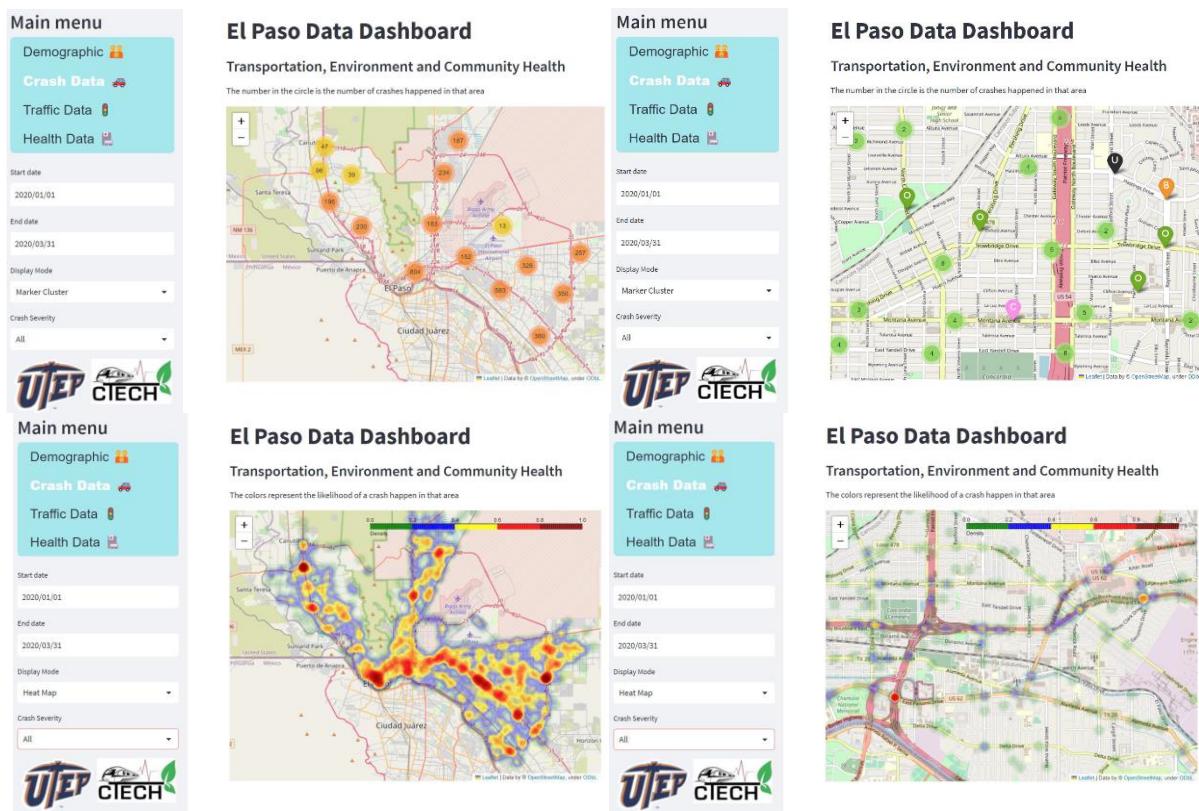


Figure 4. Crash data display on the dashboard with interactive functionality

Figure 4 demonstrates the diverse methodologies for visualizing crash data within the dashboard. Users are presented with three distinct display options for crash data visualization: Cluster Marker, Heat Map, and Marker. These can be selected via the dropdown menu labeled 'Display Mode' found in the left panel. For each mode, there are additional customization options available, such as specifying the Start Date, End Date, and Crash Severity.

When it comes to Traffic Data, there are three modes of display: Table, Density Heat Map, and Trajectory Visualization. The dashboard allows users to analyze traffic data by mode and by zip code, offering a granular view of traffic patterns.

The default screen for Traffic Data, shown in Figure 5, starts in 'Table' mode. This table format displays individual data points for each vehicle, including details like journey ID, latitude, longitude, speed, and heading. The data for each vehicle is sampled at three-second intervals. As an example, Figure 5 illustrates vehicle trajectory data from zip code 79912, captured on November 10, 2021, from 10:00:00 a.m. to 10:02:59 a.m. The table header indicates the total number of data points, while key statistics such as average, maximum, minimum speeds, and speed standard deviation are calculated for all data points within the selected zip code and displayed beneath the table.

Additionally, Figure 5 showcases a density heat map representing all zip codes. This visualization can be accessed by selecting 'Density Heat Map' in the Display Mode and opting for 'All' under Zip Code in the left panel. The density metric is derived from the number of vehicle trajectory data points in a given zip code, adjusted for the area of the zip code. It's important to note that this density represents the vehicles tracked by Wejo and does not directly equate to the actual number of vehicles per mile or kilometer. The density values are normalized

to a range between 0 and 1, in increments of 0.2, with varying colors depicting different density levels.

Figure 5 (bottom) displays the sampled vehicles' trajectories in a selected area. This is obtained when the Display Mode is changed to Trajectory Visualization. The trip origins and destinations are represented by the “O” markers and the “D” markers, respectively. The origin and destination of a trip, and the travel path are indicated by one “O” maker, one “D” marker and a color-coded line. Only trips that have at least an “O” or a “D” marker in the selected display area are shown on the map. The data points for the same vehicles are collected at 3 second intervals.

The Health Data section of the dashboard presents the City of El Paso's COVID-19 data in a variety of visual formats: Case Table, Bar Chart, Cumulative Positive Curve, Cumulative Recovery Curve, and Cumulative Death Curve. Figure 6 showcases the default view of the Health Data, which is a table format detailing COVID-19 statistics across three columns: cumulative positive cases, cumulative recovered cases, and cumulative deaths. Each row corresponds to a different zip code in El Paso, and the numbers represent the total new cases accumulated between selected start and end dates.

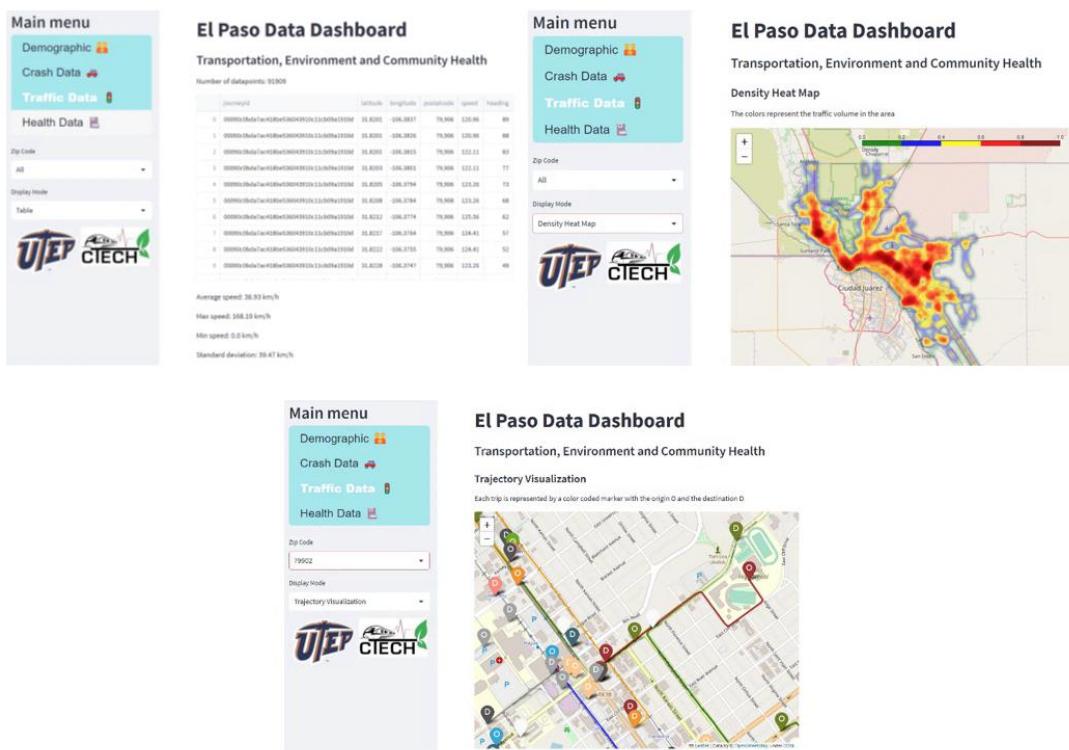


Figure 5. Traffic data displayed on the dashboard

In the example illustrated in the top left of Figure 6, the dashboard is set to display the data in Table mode, encompassing all zip codes, spanning from September 1, 2021, to March 31, 2022. This view highlights nine zip codes, with the option to scroll down for more. The top right section of Figure 6 displays a bar chart visualizing cumulative positive cases by zip code, offering a clear perspective on which areas experienced the highest number of cases during the specified timeframe.

The Health Data feature enables users to delve into detailed statistics about cumulative positive cases, recoveries, and deaths within specific timeframes. Users can tailor the display by selecting the desired zip code, choosing the type of curve under Display Mode, and setting the start and end dates. For instance, the bottom left of Figure 6 presents a Cumulative Positive Cases curve for zip code 79902, charting the period from March 1, 2020, to March 31, 2021. Similarly, the bottom right part of the figure illustrates the Cumulative Recovery Curve for the same timeframe, but for zip code 79911. These visualizations provide an insightful overview of the pandemic's impact across different areas of El Paso.

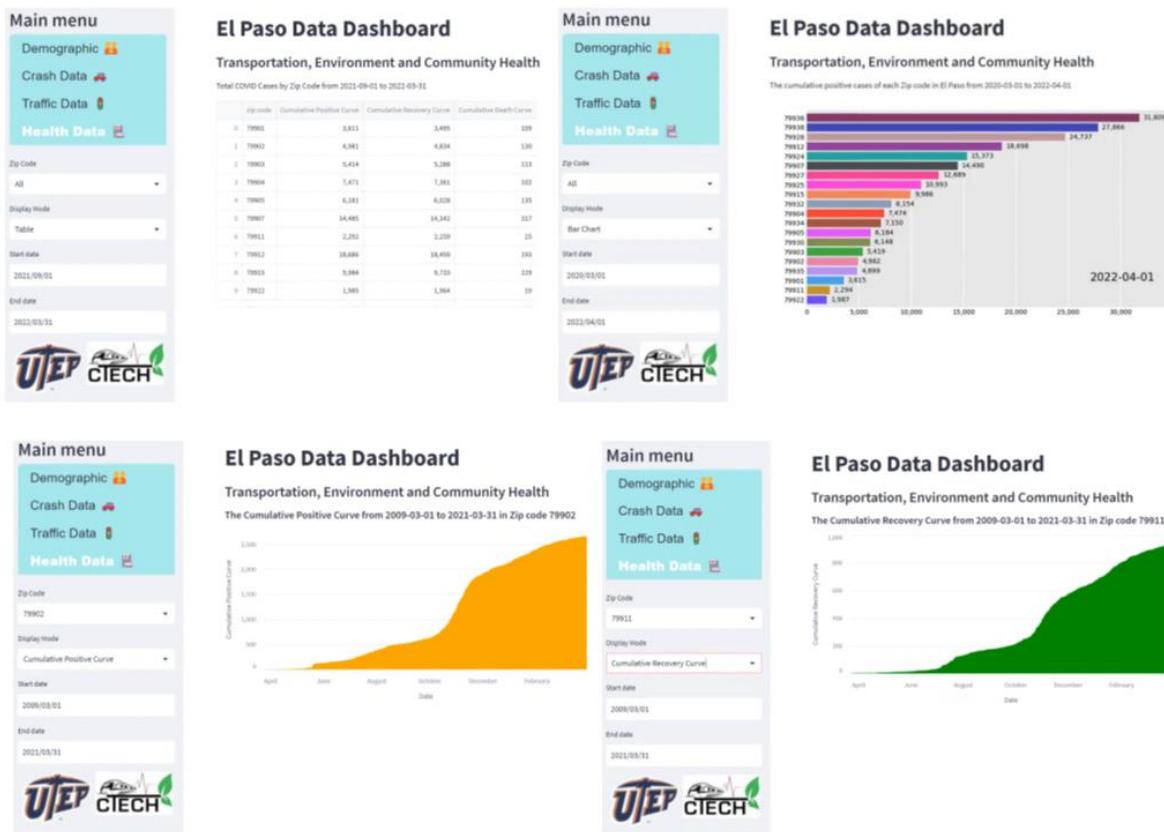


Figure 6. Health Covid-19 data displayed on the dashboard

CONCLUSION

In this study, a prototype dashboard has been developed that integrates demographic, crash, traffic, and health data within the City of El Paso. This tool was employed to engage stakeholders from El Paso and solicit their feedback, demonstrating its customizability and scalability through two case studies. The dashboard showcases the possibility of hosting, visualizing, and analyzing diverse datasets, illustrating the integration and cross-layer analysis of data as its principal output. This capacity makes it an indispensable tool for urban decision-makers, allowing them to discern patterns and trends through a convenient and intuitive interface. Beyond the dashboard itself, this research has yielded valuable processed data. Despite originating from disparate sources, the dashboard unifies these data streams, offering filtering

and visualization options to clarify emerging patterns. The integration and analytic capabilities offered by the dashboard could revolutionize urban decision-making, providing a holistic tool that benefits decision-makers and, by extension, all city residents. This research holds profound implications for the use of data in informed urban planning and policy. Future enhancements could include:

- Direct connections between the dashboard and data sources for real-time updates.
- An expanded dataset timeframe and scope, such as including post-October 13, 2021, crash data and incidents on TxDOT-managed properties.
- Enhanced data layers with additional attributes, like demographic distributions by gender and age, and traffic volume for a richer analysis.
- Integration of additional layers covering transportation and infrastructure data, such as bus stops, parking, and traffic signal timings.
- Implementation of new applications within the dashboard, like service level and equity analyses.
- Application of machine learning techniques to analyze and integrate data layers, revealing novel insights and patterns.

REFERENCES

- ADT (Application Development Trend). (2006). In depth DCStat. Application Development Trend Magazine. April 5, 2—6.
<https://adtmag.com/Articles/2006/04/05/DCStat.aspx?Page=2&p=1>. Accessed January 26, 2023.
- CASA (Centre for Advanced Spatial Analysis). (2022). CityDashboard: London. Centre for Advanced Spatial Analysis, University College London. <https://citydashboard.org/london>.
- Census. (2023). Data and tables. US Census Bureau. <https://www.census.gov/data.html>. Accessed January 26, 2023.
- City of Baltimore. (2022). About CitiStat. Office of Performance and Innovation. City of Baltimore, Maryland. <https://www.baltpi.com/about-citistat>. Accessed January 26, 2023.
- City of Boston. (2023a). COVID-19 in Boston. City of Boston.
<https://www.boston.gov/government/cabinets/boston-public-health-commission/covid-19-boston>. Accessed January 26, 2023.
- City of Boston. (2023b). Imagine Boston 2030 metrics dashboard. City of Boston.
<https://data.boston.gov/showcase/imagine-boston-2030-metrics-dashboard>. Accessed January 26, 2023.
- City of Charlotte. (2022). Capital projects. City of Charlotte, North Carolina.
https://charlottenc.gov/charlottefuture/Pages/Capital_Project.aspx. Accessed January 26, 2023.
- City of Las Vegas. (2023). City of Las Vegas community dashboard.
<https://communitydashboard.vegas>. Accessed January 26, 2023.
- City of New York. (2022). NYC open data dashboard. City of New York.
<https://opendata.cityofnewyork.us/dashboard>. Accessed January 26, 2023.
- Crisis Dashboards. (2021). <https://www.citydashboard.com.au/melbourne.php>. Accessed January 26, 2022.
- Delpeuch, A. (2022). OpenRefine User Manual. <https://openrefine.org/docs>. Accessed February 23, 2023. 70.

- EPTX. (2023). El Paso Strong. City of El Paso, Texas. <https://epstrong.org>. Accessed February 15, 2023.
- Farmanbar, M., and Rong, C. (2020). Triangulum city dashboard: an interactive data analytic platform for visualizing smart city performance. *Processes*, 8(2), 250.
- Jing, C., and Du, M. (2019). Geospatial dashboards for monitoring smart city performance. *Sustainability*, 11, 5648.
- Fischer, F. (2021). UTEP, City of El Paso opening new COVID-19 mega-testing site. KFOX14 News, August 19, 2021. <https://kfoxtv.com/news/local/utep-city-of-el-paso-opening-new-covid-19-mega-testing-site>.
- Microsoft. (2023). Azure Database for PostgreSQL. Microsoft Corp., <https://azure.microsoft.com/en-us/products/postgresql/#overview>. Accessed February 23, 2023.
- Mattern, S. (2015). Mission control: a history of the urban dashboard. *Places Journal*, March 2015. Accessed 24 Jan 2023.
- Maynooth University. (2023). Building city dashboard. <https://dashboards.maynoothuniversity.ie>. Accessed January 26, 2023.
- Oberlin. (2023). Oberlin environment dashboard - bringing environmental dashboard to your community. Oberlin College, Ohio. https://environmentaldashboard.org/bring-dashboard-to-your-community#_Toc446430146. Accessed February 15, 2023.
- Ratti, C., et al. (n.d.). LIVE Singapore! Senseable City Lab, Massachusetts Institute of Technology. <https://senseable.mit.edu/livesingapore/visualizations.html>. Accessed 26 January 2023.
- Sanromā, L. (2021). The principles of visual design for dashboards. Adnia Solutions. Retrieved November 18, 2022, from <https://adniasolutions.com/dashboard-design-principles/the-principles-of-visual-design-for-dashboards>.
- Scikit-Learn. (2023). Scikit-learn, machine learning in Python. Scikit-Learn Developers. <https://scikit-learn.org/stable/> Accessed February 23, 2023. 71.
- Streamlit. (2023). Streamlit documentation. Streamlit Inc. <https://docs.streamlit.io>. Accessed February 23, 2023.
- Story, R. (2013). Folium 0.14.0 documentation. <https://python-visualization.github.io/folium>. Accessed February 23, 2023.
- Wejo. (2023). Trafic Movements. Wejo Limites. <https://www.wejo.com/products/vehicle-movements>. Accessed February 15, 2023.
- WSDOT (Washington State DOT). (2023). COVID-19 Multimodal transportation system performance dashboard. Washington State Department of Transportation. <https://www.wsdot.wa.gov/about/covid-19-transportation-report>. Accessed 2/15/2023.
- Zuo, F., Wang, J., Gao, J., Ozbay, K., Ban, X. J., Shen, Y., Yang, H., and Iyer, S. (2020). An interactive data visualization and analytics tool to evaluate mobility and sociability trends during COVID-19. *UrbComp 2020: The 9th SIGKDD International Workshop on Urban Computing*, San Diego, California, USA.

Leveraging Computer Vision for Transportation Infrastructure Health Monitoring

Claudia Marin-Artieda, Ph.D., P.E., M.ASCE¹; Ogheneruona Akpareva²;
and Stephen Arhin, Ph.D., P.E., M.ASCE³

¹Professor, Dept. of Civil and Environmental Engineering, Howard Univ., Washington, DC.
Email: cmarin@howard.edu

²Undergraduate Student, Dept. of Civil Engineering, Howard Univ., Washington, DC.
Email: ogheneruona.akparev@bison.howard.edu

³Professor, Dept. of Civil and Environmental Engineering, Howard Univ., Washington, DC.
Email: saarhin@howard.edu

ABSTRACT

The paper discusses a vision-based sensor methodology for tracking displacements in structures using videos and computer vision algorithms to detect and track displacements at various scales. This method is tested on a full-scale bridge column footage from earthquake simulations, demonstrating its ability to capture both global and large-scale displacements and detailed local deformations. The paper compares these video-based measurements with traditional sensor data, proving the method's effectiveness. The study highlights the potential of video analysis in expanding data collection to areas not typically covered by sensors. It suggests that this method can enrich datasets and improve the accuracy of structural health monitoring.

INTRODUCTION

Accurate and reliable methods for evaluating and monitoring structural conditions are essential for transportation infrastructure maintenance, repair, or replacement decisions. Structural Health Monitoring (SHM), predominantly based on structural dynamics, continuously monitors and assesses structural integrity. It involves data acquisition, system identification, and condition assessment, primarily through frequency-domain and time-domain methods. Traditional SHM, using contact-based sensors, faces challenges like optimal sensor placement, complex analysis, and limited damage detection. Computer vision (CV), particularly vision-based sensing, offers a solution with noncontact, wide-range sensing, high accuracy, and improved damage detection. Advances in deep learning and cloud computing further enhance data analysis and interpretation in SHM.

Fujino et al. 1993 pioneered the use of video recordings and digital analysis to assess the synchronization of walking patterns and their impact on pedestrians' bridge lateral vibrations. Since then, advanced camera technology and computer vision techniques have guided vision-based sensing as an alternative to traditional sensors for SHM. It utilizes video imagery of structural elements to derive displacement time histories at multiple points concurrently. Since the 2010s, vision-based sensor technology has been used for displacement measurement, proving its efficacy mostly in laboratory settings. This technology facilitates modal analysis using displacement data, aiding in natural frequency and mode shape identification, model updating, damage detection, and structural integrity assessment.

Several studies have explored vision-based techniques. Choi et al. 2011 introduced a dynamic displacement vision system, proving its effectiveness in varied conditions. Chen et al. 2016 developed vibration measurement in structures using motion magnification, which was

applied to an antenna tower. Kim et al. 2017 validated a vision-based displacement sensor, emphasizing its high accuracy and practicality through lab and field tests. Concurrently, Chen et al. 2017 merged optical flow with Kalman filters for noncontact damage detection. Shishido et al. 2017 developed a technique for estimating the 3D positions of fast-moving objects, notably enhancing training and tactical analysis in sports.

Chen et al. 2018 compared camera-based measurements with traditional sensors, affirming their effectiveness in capturing bridge dynamics. Park et al. 2018 combined vision-based data with acceleration data to improve accuracy, the method was validated through numerical and experimental testing. Narazaki et al. 2019 focused on 3D displacement and strain estimation in miter gates, stressing the importance of accurate measurement planning. Spencer et al. (2019) provided an overview of recent advances in vision-based monitoring, identifying key challenges and ongoing research areas. Narazaki et al. 2020 explored automated bridge component recognition using vision and the development of advanced vision-based 3D displacement measurement algorithms, enhancing the accuracy and efficiency of SHM. Zhang et al. 2021 used semantic image segmentation for pixel-level damage identification in bridges. Karameşe et al. 2023 demonstrated the potential of using smartphone videos to measure structures' dynamic responses.

The literature showcases a range of applications of computer vision technologies in SHM. Dynamic response measurement is identified as a significant application area, enabling the evaluation of structural behavior under various conditions. The modal analysis benefits from these technologies, facilitating the identification and analysis of modal properties in structures. Structural damage detection is improved by implementing vision-based systems, enhancing the ability to identify and evaluate local damage. Additionally, the literature explores the potential of computer vision for estimating cable forces in structures.

Overall, the literature consistently acknowledges the potential inherent in vision-based sensor technology for SHM. It highlights the technology's capacity to measure structural displacement, strain/stress, and dynamic responses. The literature focuses on recognizing and addressing errors when applying vision-based measurement systems. Integration of vision-based sensors and traditional sensor-based approaches enhances the overall capabilities of SHM systems. This fusion allows more effective and accurate measurement of structural displacements, strain/stress, and dynamic responses. Furthermore, the literature addresses the importance of managing environmental factors influencing measurement accuracy. Techniques to mitigate these influences, such as compensatory methods, are explored to enhance the robustness of vision-based sensors in diverse environmental conditions. These compensatory strategies encompass software-based algorithms to physical shielding and multi-sensor integration. These methods aim to mitigate the adverse effects of lighting, temperature, noise, and other environmental variables to enhance the robustness of vision-based sensor systems in field applications.

The literature also emphasizes several challenges associated with computer vision technologies in SHM. Addressing systematic errors and the need for robust error assessment is emphasized. The impact of environmental conditions, such as varying lighting and temperature changes, challenges measurement accuracy. Camera-related issues, including resolution, focus, and stability, are identified as factors affecting measurement accuracy. Managing and integrating the substantial volumes of data generated and ensuring accuracy and resolution in data collection is also emphasized.

The paper discusses a vision-based sensor methodology for tracking displacements of full-scale structures. Video and computer vision algorithms are used to detect and track

displacements at various scales. The paper validates this method by applying it to video footage from earthquake simulations on a full-scale reinforced concrete (RC) bridge column. It compares video-based displacement measurements with traditional sensor data, discussing the method's accuracy in capturing both global and local structural responses. The study highlights the potential of video analysis in expanding data collection to areas not covered by sensors. It suggests that this method can enrich datasets and improve the accuracy of structural health monitoring.

VISION-BASED SENSOR CONCEPTUALIZATION TO TRACK DISPLACEMENTS

In CV, images are conceptualized as a grid of pixels, each representing a discrete point in the image. The resolution, defined by the pixel count along the image's width and height, dictates the captured detail level. Higher-resolution images containing more pixels can capture finer details. CV algorithms utilize this pixel grid to analyze color, intensity, and texture attributes. Figure 1 illustrates this concept, showing a pedestrian bridge over a city represented as a grid, with grid lines placed every 50 pixels. Such a representation enhances spatial resolution and detail, which is particularly beneficial in local damage assessment tasks. Tracking numerous points across the grid achieves a higher spatial resolution of the structural behavior.



Figure 1. A pedestrian bridge is represented as a grid, with grid lines placed every 50 pixels.

To calculate the pixel displacements, optical flow is used to describe the apparent motion pattern of objects, surfaces, or edges within a structure resulting from the relative movement between a camera and the structure.

The Lucas-Kanade method (Lucas and Kanade 1981), a differential approach, supports tracking fine details in images. This method operates under the assumption that the motion of pixels is small and smooth within a localized area. It determines pixel displacements between image frames by analyzing gradients in image intensity and employing a least-squares fitting technique. Particularly effective in regions rich in texture or distinct features, the Lucas-Kanade method captures subtle movements. Figure 2 illustrates this method's application in estimating optical flow between two frames. It presents the initial image frame and overlays the optical flow

vectors generated by the subsequent frame. These vectors graphically depict the tracked pixel displacements, visualizing motion between the frames.



Figure 2. Illustration of the Lucas-Kanade method estimating optical flow between two frames.

Integrated with the Lucas-Kanade method (Lucas and Kanade 1981), optical flow pyramids (Horn and Schunck 1981, NVIDIA 2021) can handle large displacements across multiple pixels by representing an image at multiple scales.

The pyramid approach in image processing is adept at detecting and tracking displacements by utilizing a multi-scale strategy. It begins at the highest levels of the pyramid with lower-resolution images, which are effective for spotting larger displacements that are more discernible with less image detail. The method progressively shifts focus toward the lower levels of the pyramid, which feature higher-resolution images for pinpointing smaller, more localized displacements. As depicted in Figure 3, the technique involves sequentially resizing the original image to capture an extensive range of motion. The pyramid's top level offers a broad overview, suitable for observing large-scale movements. In contrast, the bottom level provides detailed, high-resolution imagery necessary for detecting small displacements. This methodology effectively visualizes and calculates motion differences between two consecutive video frames at the pixel level, ensuring a thorough and multi-faceted motion analysis across different image scales.

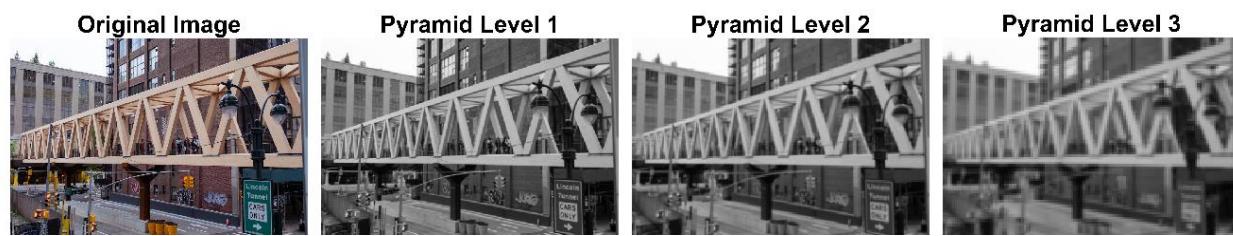


Figure 3. Multi-scale analysis through pyramid-based optical flow: from broad views to fine details

A two-directional prototype Python code integrating the Lucas-Kanade method with optical flow pyramids for tracking pixel displacements in video sequences was developed and validated

to calculate structural displacements. This integration supports tracking over long distances and handling large displacements. The script relies on user input to determine which points to track.

ENRICHING A DATA SET OF A FULL-SCALE BRIDGE COLUMN EXPERIMENT

The prototype code was applied to analyze video footage from a series of earthquake simulations of a study of a full-scale RC bridge column conducted by Schoettler et al. 2015 at the University of California, San Diego, that involved a 1.22-meter diameter, 7.32-meter high cantilevered reinforced concrete bridge column. The column was subjected to ten earthquake simulations on a shake table, testing its limits up to near collapse. During testing, a 2322 kN mass was placed atop the column to induce inertia forces. This column displayed ductile behavior with a plastic hinge forming near its base, enduring concrete spalling, bar buckling, and fractures. It remained stable even beyond design-level events. The experiment featured detailed instrumentation, including linear variable differential transformers (LVDTs), string potentiometers, strain gauges, accelerometers, GPS for 3D displacement, and six video cameras focused on the plastic hinge region, complemented by additional cameras for diverse angle captures. Figure 4 displays two images of the columns and sensors and sample point-tracked coordinates of the prototype code.

The prototype video measurements were used to assess global and local responses at specific locations equipped with sensors for evaluating performance under a series of tests. The series included ten primary seismic motions labeled EQ1 to EQ10, followed by white noise testing. Unfortunately, video data from these white noise tests were unavailable. This report presents selected local and global response analysis results as a sample of the 2D prototype's measurements. Specific tests, such as EQ2 (Coralitos record from the 1989 Loma Prieta earthquake), EQ4 (repetition of Coralitos record post-EQ3 Los Gatos Presentation Center record from the same earthquake), EQ6, EQ7, EQ8, and EQ9 (Takatori records from the 1995 Kobe earthquake at different amplitudes) were chosen for their available parallel view video footage of the column responses.

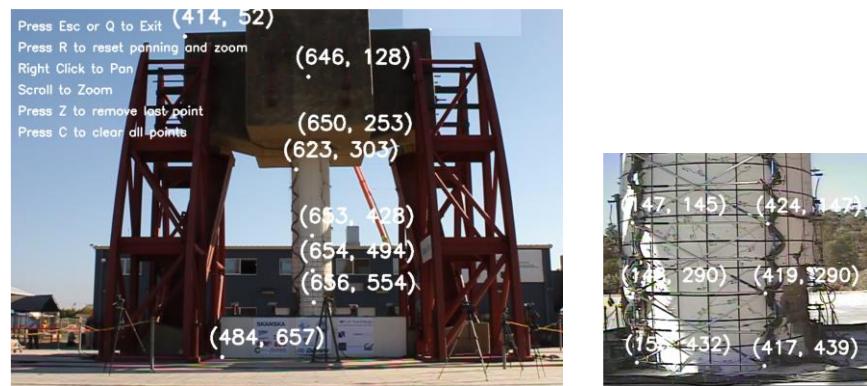


Figure 4. Images of the columns and sensors and sample point-tracked coordinates

The column diameter of 1219 mm was used as a reference for converting pixel measurements to millimeters. However, the image clarity and the capture of the 2D plane could affect measurement accuracy, primarily when significant displacement differences existed between the camera and the columns. The program assumes the column plane is parallel to the image plane,

with the column's longitudinal axis aligned with the image's axes. Corrections are necessary for structures not parallel to the image plane.

Synchronizing video and sensor data involved aligning their peak responses due to different start times. Sensor data sometimes showed initial deformations not visible in the video, likely because sensors recorded non-physical changes. For accurate comparison, signal means were subtracted to highlight differences between these measurement methods.

Capturing global responses

The column was instrumented with several potentiometers along its height, and accelerometers, and GPS devices were positioned at the top. Multiple points were identified for tracking displacements. The results presented focus on the analysis from the top potentiometer and the displacements measured by an accelerometer located at the top, as shown in Figure 5. This figure illustrates the view of the column and identifies the sensors relevant to the analysis presented here. The displacement results obtained from the video are compared with the data from the experimental data for experiment EQ6.

For the EQ6 test, the displacement at the location of the string potentiometer was determined using video analysis (MVI_2825 with resolution 1280 x 720 pixels, 30 frames per second (fps), Schoettler et al. 2010) and compared with measurements from the potentiometer. Figure 6 illustrates the measured and estimated displacement at this location. Analysis of these peaks shows a minimal discrepancy difference in the peak response of up to 2%. Figure 6 also presents the frequency analysis of the displacement data over time. This was achieved using the wavelet synchrosqueezing transform, a technique effective for extracting instantaneous frequencies from displacement data (Mallat 2008). There were moderate frequency variations, particularly between 28 to 34 seconds. However, outside of this interval, the frequency variations of both signals closely align. Given the simplicity of calculating these displacements, such differences are considered relatively minor. The displacement measurements obtained from the videos for other tests, which were aligned with the location of the string potentiometer S5E, yielded similar results. Details of these results are omitted here for brevity.

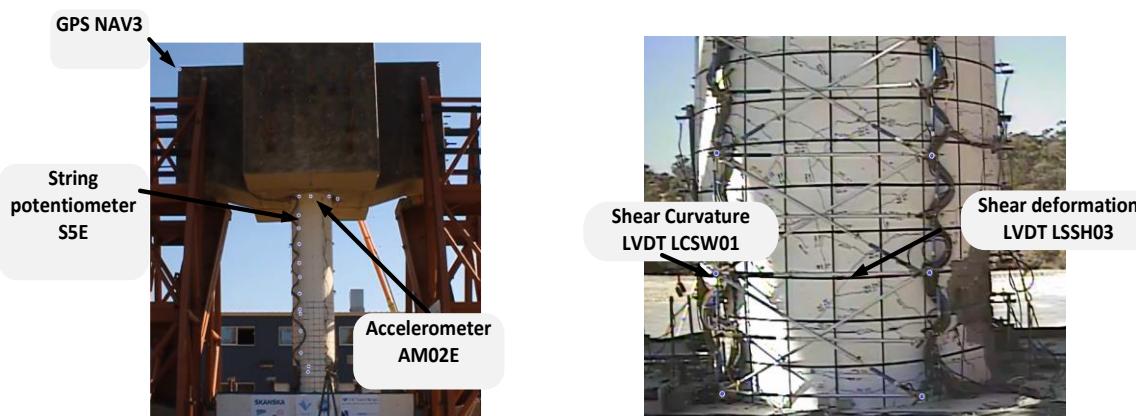


Fig. 5. View of the column identifying the sensors relevant to the analysis.

The displacement at the GPS NAV3 location atop the columns (see Fig. 5) was evaluated through video analysis and then compared with GPS measurements (50 fps). Figure 7 depicts

both the GPS and video-derived displacement at this point, emphasizing the peak values recorded in each dataset. Analysis of these peaks shows a minimal discrepancy difference in the peak response of up to 3%, signifying very similar frequency variations over time for both signals. Observing the comparable sampling rates, with videos being captured at 30 frames per second (fps) and GPS data recorded at 50 fps. This comparison underscores the codes' effectiveness in accurately capturing large-scale displacements.

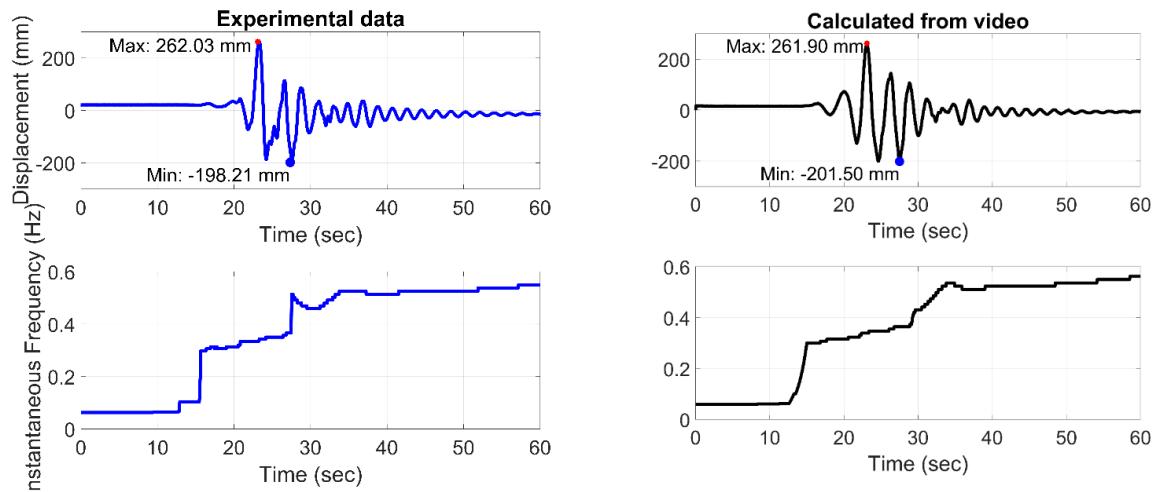


Figure 6. Measured and estimated displacement at the S5E location.

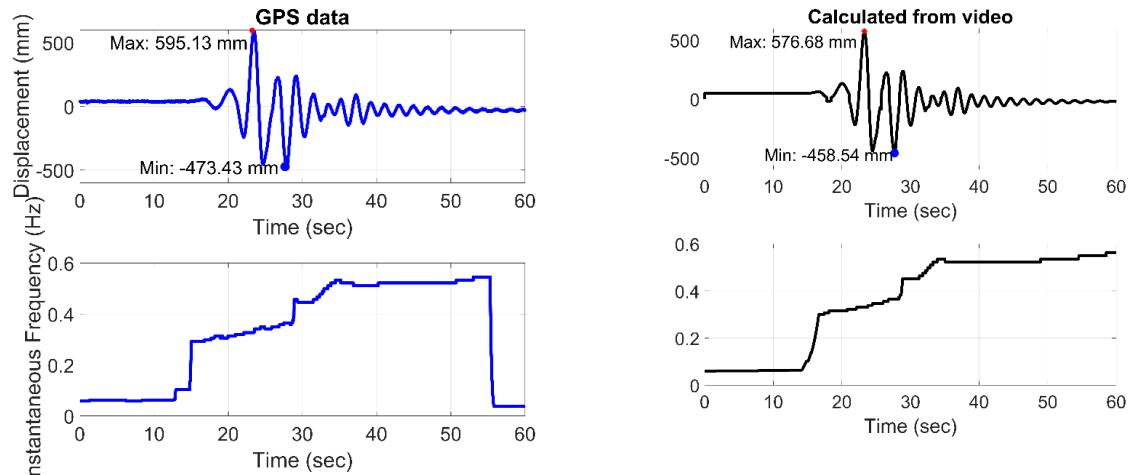


Figure 7. Measured and estimated displacement at the GPS NAV3 location

Capturing local responses

The column in the study was equipped with various sets of LVDTs to measure curvatures and shear in the hinge area. Multiple points for displacement tracking were established. The results discussed herein are primarily derived from the LVDTs highlighted in Figure 5. This figure provides a view of the column, pinpointing the sensors pertinent to the analysis. The aim

was to calculate the actual displacement between the two points defined by the LVDT and then compare it with the data from the LVDTs. This analysis utilized data from EQ6.

The displacement at two points linked to the LVDT locations LCSW01 and LSSH03 were assessed through video analysis (using Camera 1 South with a resolution of 704 x 480 pixels and at 30 fps, Schoettler et al. 2010). The results for these two LVDTs are shown in Figure 8. This figure shows the relative displacements measured between the points corresponding to the LVDTs. The video measurements of the vertical LVDT LCSW01 show good agreement with the LVDT's measured deformations. There is a minor difference in the minimum values, with a divergence of approximately 13%, but the maximum peaks are nearly identical. The general trend of both signals is very similar, the video measurements and the LVDT readings are in agreement.

Regarding the horizontal LVDT LSSH03, the maximum values are similar, with a marginal difference of about 2 % on the positive side. However, there is a noticeable discrepancy on the negative side, with a larger value observed by the video measurement. Based on the results using this experimental dataset, it is challenging to determine which sensor is more accurate, as both sensors have certain limitations.

The results confirm that video-based displacement measurements align closely with traditional sensor data, highlighting the method's accuracy in tracking structural responses at various scales. Despite considerable distances between the specimen and the camera and the low resolution of some of the videos, the video-derived displacement data showed high accuracy. However, variations in frequency consistency were observed, which was expected given the video's frame rate of 30 frames per second. The analysis was particularly effective in evaluating local displacements, especially at points where LVDT sensors were used for monitoring shear deformations. These results suggest the potential of video analysis in expanding data collection to regions not covered by sensors during experiments. This methodology can enrich data sets, even without sensor placement.

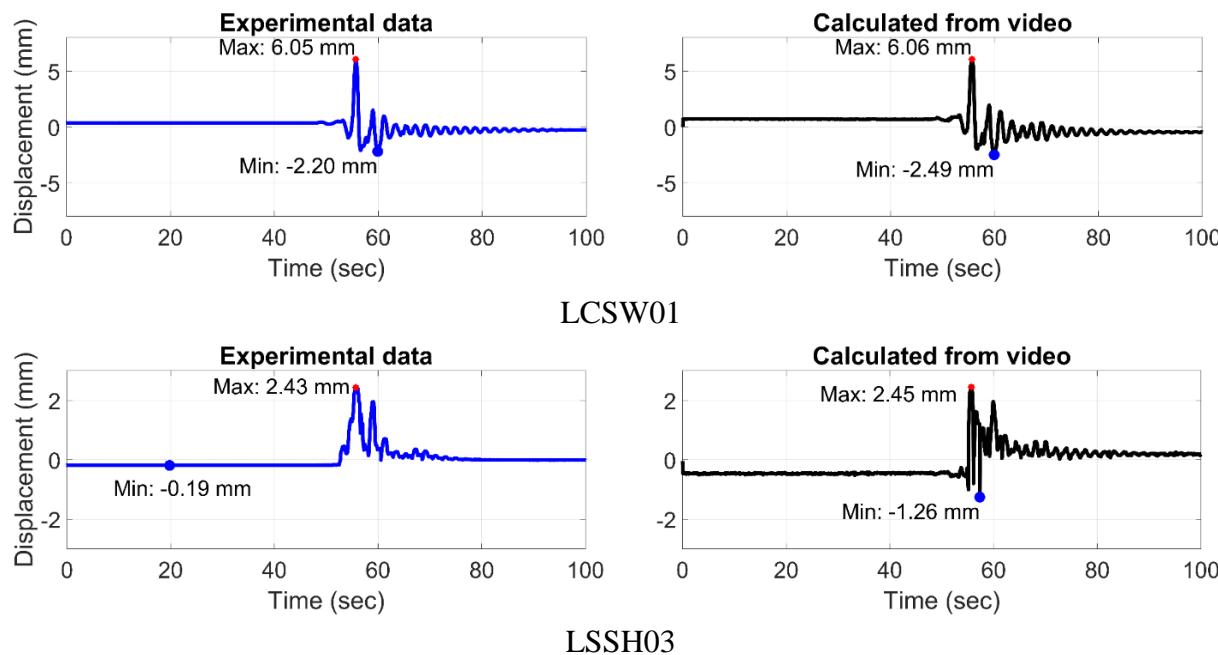


Figure 8. Deformation results of the LVDTs

MULTI-POINT DISPLACEMENT MEASUREMENT BENEFITS

Most SHM studies and applications have focused on acceleration measurements. However, displacement responses are more directly indicative of structural conditions. Multi-point displacement measurements allowed by vision-based sensors can improve vibrational analysis aiding in assessing structural integrity. The vision-based measurements can capture the structure's responses with spatial resolution and enable dense spatial sampling. Such details can support understanding a broad spectrum of structural responses, including local deformations that might go unnoticed with fewer accelerometers.

Free vibration of the bridge column

To validate the efficacy of video analysis in capturing effective viscous damping and the fundamental periods of bridge columns discussed in the previous section, displacement data measured from video recordings that parallel the columns during various excitations were used. The focus was on analyzing the displacement responses at the top of the columns, particularly at location S5E as indicated in Figure 5. The top displacements during the excitations and at the end of the response, the free vibration, were measured. Figure 9 illustrates the tail-end (free vibration) of these displacement responses at the top column location (S5E) for different excitations, including the exponential decay and the corresponding damping and period of the columns following each seismic shaking test. This aspect of the study is particularly useful, as it shows that even with small vibrations and taking the end of the displacement responses, the video measurements effectively captured the fundamental period and damping characteristics of the columns. This shows that video recordings of any structure can be used to identify viscous damping and the fundamental periods. Video-based measurements allow data collection in hard-to-reach or unsafe areas, overcoming the limitations of traditional accelerometers, and can be used to identify effective viscous damping and the fundamental periods of structures. This is especially beneficial for monitoring large, slender structures or those in hazardous environments.

Estimation of accelerations from displacements

Direct displacement measurements make estimating structural conditions easier by directly providing deflections and drifts, unlike acceleration-based methods that require double integration and can increase noise and errors.

Calculating displacement from acceleration requires unknown initial values and can increase low-frequency noise, affecting accuracy. Choosing the right low-cut filter is crucial for reducing errors. Conversely, deriving accelerations from displacement doesn't need initial values but can amplify high-frequency noise.

Displacement measurements from videos at points matching an accelerometer's location on an RC bridge column were analyzed. The displacement at accelerometer AM02E, recorded by Camera 1 South (Schoettler et al. 2010), was differentiated twice and compared with the experimental acceleration data. Figure 10 displays calculated displacements from the video at accelerometer AM02E's location and the resulting accelerations. A moving average low pass filter (window size 4) was applied to velocity and acceleration signals to minimize noise from differentiation. The figure shows general trend similarities in both calculated and measured accelerations, with up to 15% larger peaks in the calculated data. Filter parameter selection

significantly affects acceleration estimates, a balance is needed between noise reduction and maintaining true signal characteristics.

Integration for displacement from acceleration needs initial conditions and manages low-frequency noise, while differentiation for acceleration from displacement amplifies high-frequency noise. Both processes require careful signal processing to balance true signal preservation and noise reduction.

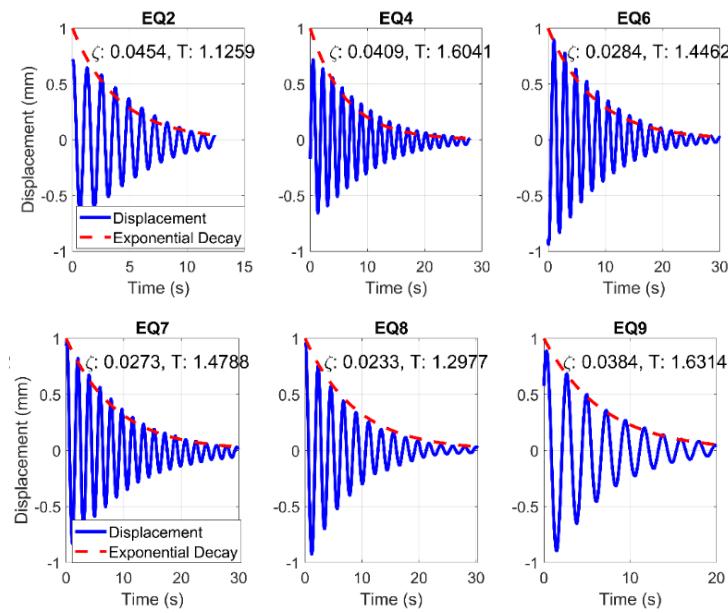


Figure 9. Free vibration displacement responses at the top column (S5E) for different excitations

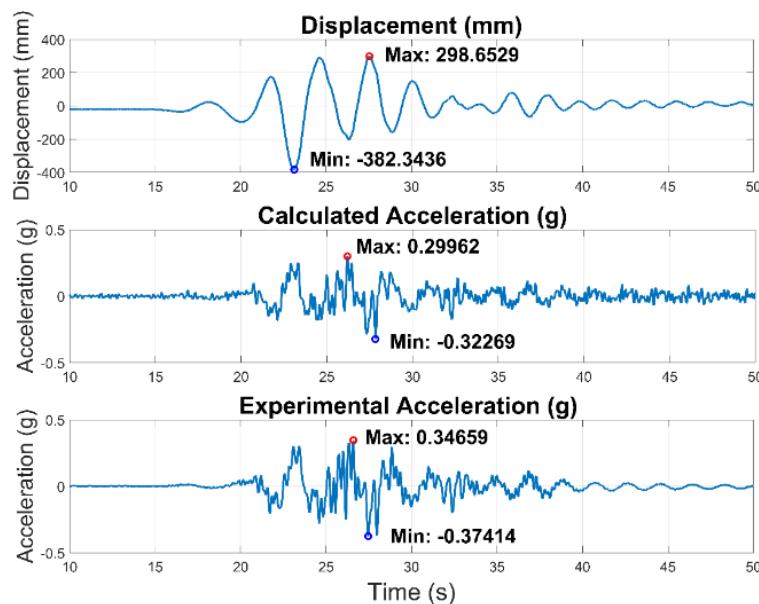


Figure 10. Displacements at accelerometer AM02E's location and measured accelerations

CONCLUSION

The results of this research show that video-based displacement measurements offer an accurate alternative for assessing structural responses, both globally and locally. It highlights the effectiveness of this method in capturing large-scale displacements and small local deformations. The study validates the accuracy of video analysis in comparison to traditional sensors, emphasizing its potential to expand data collection, especially in areas not covered by sensors. The approach can be particularly beneficial for monitoring structures in hard-to-reach or hazardous environments.

Vision-based sensors can offer the opportunity to monitor general infrastructure, such as traffic signals, lighting poles, and communication antennas. These structures, often overlooked due to their low profile, can lead to significant disruptions if they fail.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. 2040665 and by U.S. DOT 69A3552348339 Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of supporting agencies.

REFERENCES

- Horn, B. K. P., and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3), 185-203.
- Lucas, B. D., and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, 674-679.
- Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., and Büyüköztürk, O. (2016). Video camera-based vibration measurement for civil infrastructure applications. *Journal of Structural Engineering*, 142(10), 04016108.
- Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., and Büyüköztürk, O. (2017). Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters. *Journal of Structural Engineering*, 143(1), 04016163.
- Chen, J. G., Adams, T. M., Sun, H., Bell, E. S., and Büyüköztürk, O. (2018). Camera-based vibration measurement of the World War I Memorial Bridge in Portsmouth, New Hampshire. *Journal of Structural Engineering*, 144(11), 04018200.
- Choi, H.-S., Cheung, J.-H., and Kim, S.-H. (2011). Structural dynamic displacement vision system using digital image processing. *Journal of Computing in Civil Engineering*, 25(5), 375-383.
- Fujino, Y., Pacheco, B. M., Nakamura, S., and Warnitchai, P. (1993). Synchronization of human walking observed during lateral vibration of a congested pedestrian bridge. *Journal of Sound and Vibration*, 167(2), 347-364.
- Karameşe, G., Yardımcı, F., Güllü, A., Yüksel, E., Yalçın, C., and Büyüköztürk, O. (2023). Vision-based dynamic response measurements of structures by using smartphone videos and alternative techniques. *Journal of Structural Engineering*, 149(1), 04022100.

- Kim, Y., Kim, J., and Lee, H. (2017). Experimental validation of cost-effective vision-based structural health monitoring. *Journal of Computing in Civil Engineering*, 31(2), 04017072.
- Mallat, S. (2008). A wavelet tour of signal processing: The sparse way. In *A wavelet tour of signal processing: The sparse way*. Cambridge, MA: Academic Press.
- Narazaki, Y., Gomez, F., Hoskere, V., Smith, M. D., and Spencer, B. F., Jr. (2020). Efficient development of vision-based dense three-dimensional displacement measurement algorithms using physics-based graphics models. *Journal of Computing in Civil Engineering*, 34(2), 04019057.
- Narazaki, Y., Hoskere, V., Eick, B. A., Smith, M. D., and Spencer, B. F. (2019). Vision-based dense displacement and strain estimation of miter gates with the performance evaluation using physics-based graphics models. *Journal of Computing in Civil Engineering*, 33(2), 04018097.
- NVIDIA. (2021). Pyramidal LK optical flow. Vision Programming Interface.
- Park, J.-W., Moon, D.-S., Yoon, H., Gomez, F., Spencer, B. F., and Kim, J. R. (2018). Visual-inertial displacement sensing using data fusion of vision-based displacement with acceleration. *Structural Control and Health Monitoring*, 25(3), e2122.
- Schoettler, M. J., Restrepo, J. I., Guerrini, G., Duck, D., and Carrea, F. (2010). *Full scale bridge column test*. DesignSafe-CI [publisher].
- Schoettler, M. J., Restrepo, J. I., Guerrini, G., Duck, D. E., and Carrea, F. (2015). A full-scale, single-column bridge bent tested by shake-table excitation. PEER Report 2015-02. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.
- Shishido, H., Kameda, Y., and Kato, T. (2017). Visual tracking method of a quick and anomalously moving badminton shuttlecock. *ITE Transactions on Media Technology and Applications*, 5(3), 110-120.
- Spencer, B. F., Jr., Hoskere, V., and Narazaki, Y. (2019). Advances in computer vision-based civil infrastructure inspection and monitoring. *Journal of Computing in Civil Engineering*, 33(4), 04019012.
- Zhang, Y., Li, J., and Wang, Y. (2021). A hierarchical semantic segmentation framework for computer vision-based bridge damage detection. *Automation in Construction*, 126, 103757.

Harnessing Synthetic Image Datasets for Enhanced Scene Understanding in Construction Work Zones

Sultan Al Shafian, S.M.ASCE¹; and Da Hu, Ph.D., A.M.ASCE²

¹Dept. of Civil and Environmental Engineering, Kennesaw State Univ., Marietta, GA.
Email: salshaf1@students.kennesaw.edu

²Dept. of Civil and Environmental Engineering, Kennesaw State Univ., Marietta, GA.
Email: dhu3@kennesaw.edu

ABSTRACT

This study pioneers the use of synthetic image datasets, generated via the Unity game engine, to train deep learning models for construction work zone scene understanding. This innovative approach simplifies data acquisition and ensures a rich, diverse dataset that includes various construction scenarios, both hazardous and typical. We created 12,360 images with accurate bounding box annotations, ensuring high-quality, consistent data crucial for model training and validation, and effectively addressing the ambiguities often found in real-world datasets. An object detection benchmark was established using this dataset alongside eight state-of-the-art object detectors. This benchmark thoroughly evaluates the performance of these detectors on a wide range of construction site images, enabling comparisons and analyses of different models. It highlights their respective strengths and weaknesses in construction site applications. Notably, YOLOv8-L demonstrated exceptional performance, achieving a mean average precision (mAP) of 70.7% on the validation set and 69.8% on the testing set. These results underscore the efficacy of synthetic datasets in training models for complex scene understanding. This integration of synthetic and real-world imagery has the potential to revolutionize scene comprehension in construction zones, significantly enhancing safety and efficiency in the construction industry.

INTRODUCTION

The construction industry is vast and multifaceted, encompassing a wide range of projects, each with its unique challenges and requirements. The equipment and methodologies differ significantly from erecting a high-rise building to excavating an underground tunnel. As a result, each construction site, widely recognized as one of the most hazardous working environments, has its distinctive complexities and hazards (Mahmoud et al. 2020; Timofeeva et al. 2017). Injuries in construction zones not only impact the well-being of the workers but can also lead to delays, increased costs, and reduced efficiency. According to data from the U.S. Bureau of Labor Statistics, the construction and extraction occupations recorded the second-highest count of occupational fatalities in 2021, totaling 951 cases, alongside a notable 8.9% increase in the total number of fatal work injuries across the United States, culminating in 5,190 recorded incidents (U.S Bureau of Labor Statistics 2022). As each construction site is unique, given the dynamic nature of construction environments, where machinery, materials, and manpower constantly interact, a holistic grasp of the site can significantly reduce errors, prevent accidents, and streamline operations. By accurately interpreting and predicting the interplay of various elements in real-time, stakeholders can optimize resource allocation, ensure safety protocols, and expedite project timelines, ultimately leading to cost savings and timely project delivery. Hence, a

comprehensive understanding of these construction work zones through advanced technological methods like artificial intelligence, virtual reality, augmented reality, etc., is deemed vital for various interested parties in the industry.

The potential of such advanced technology in transforming construction scene comprehension is vast and transformative. These tools offer real-time data and insights, predict potential hazards, optimize resource allocation, and automate routine tasks. By integrating these technologies, construction professionals can achieve a deeper, more accurate understanding of their sites, leading to safer environments, reduced inefficiencies, and a more streamlined construction process, specifically, image processing methods, which have the potential to transform construction scene comprehension. However, image analysis requires vast datasets to ensure accuracy and robustness, especially when powered by machine learning. This insatiable demand for data presents a significant drawback. Firstly, acquiring large volumes of high-quality, relevant images can be time-consuming, expensive, and logistically challenging. Additionally, the need for diverse samples to account for various scenarios further complicates data collection. Once obtained, these datasets require meticulous annotation, which can be labor-intensive. Furthermore, the storage and processing of such extensive data necessitate substantial computational resources. This data-centric nature of image analysis can often impede rapid deployment and scalability, especially for smaller organizations with limited resources.

To address this challenge, the researchers adopt an innovative strategy utilizing synthetic image datasets to train deep learning models specifically designed to understand construction work zone scenes. This method streamlines the data collection process while ensuring the dataset captures a wide range of construction scenarios, including potential hazards and common work zone activities.

LITERATURE REVIEW

Understanding the dynamics of a civil engineering construction site is challenging due to its complexity, which involves unpredictable worker movements, equipment placement, and various other factors (Park and Cho 2017). Factors such as the unpredictable movement of workers, placement of equipment, variable material deliveries, weather fluctuations, and the terrain of the site, including uneven ground and subsurface conditions, contribute to this complexity. These factors necessitate constant adjustments in work schedules and plans, and they are further compounded by the involvement of specialized trades and subcontractors, the transient nature of the workforce, and the need to adhere to safety standards and regulatory compliance. Additionally, the impact of local community and stakeholder interests, like noise restrictions and environmental protections, also affects the pace and quality of work.

Artificial intelligence (AI) serves as a powerful tool in comprehending the dynamic environment of construction sites and analyzing and interpreting the complex data they generate. AI systems are equipped to identify patterns and make decisions with minimal human intervention. One of the most significant recent advancements in AI is the application of machine learning and deep learning to address real-world challenges in civil engineering. These technologies are revolutionizing the field, offering sophisticated tools for data analysis and complex problem-solving. Machine learning algorithms are utilized to predict structural integrity, optimize construction workflows, minimize risks, and develop civil engineering software, among other applications (Hu 2022; Stone et al. 1991). Deep learning, with its advanced neural networks, excels in processing vast amounts of data from various sources, such

as sensors and satellite images, aiding in urban planning, environmental monitoring, infrastructure management, and locating underground structures (Wang et al. 2023b). The integration of AI into civil engineering is not only streamlining existing processes but also facilitating innovative approaches in design, maintenance, and construction that were previously unattainable. As these technologies continue to evolve, they promise to transform civil engineering further, making it more efficient, sustainable, and resilient.

However, the effectiveness of machine learning and deep learning in civil engineering is heavily reliant on the availability of diverse and extensive data sets, which are often difficult to obtain due to various constraints. For example, robots used in construction safety need large image datasets for hazard recognition, but contractual limitations concerning privacy and proprietary information can impede data collection. Structural health monitoring also requires detailed historical data on material behavior that may be restricted due to commercial or security concerns. Traffic management applications face similar challenges, as they require comprehensive traffic data to improve flow and reduce congestion, yet such data collection can be obstructed by practical or regulatory issues. In such cases, synthetic data sets emerge as a highly beneficial solution. These artificially created data sets can be generated under controlled lab conditions to simulate a wide range of scenarios and variables that models may encounter in the real world. Researchers and engineers can effectively train and refine their AI models by utilizing synthetic data, ensuring they are robust and well-equipped to handle complex tasks (Song et al. 2023). This approach not only circumvents the issue of data scarcity but also allows for the exploration of data-driven insights in environments where real data may be too sensitive or expensive to collect. The use of synthetic data is particularly advantageous in civil engineering, where real-world data can be fraught with privacy issues and proprietary concerns or be too hazardous to collect *in situ*. With the strategic use of synthetic data, the potential for AI to innovate and improve civil engineering practices is greatly amplified, paving the way for more advanced, safe, and efficient infrastructure systems.

DATASET PREPARATION

The Unity game engine was employed for this research to create a comprehensive dataset. Unity's advanced features, particularly its versatile lighting and shading options, make it ideal for constructing detailed and realistic virtual environments. These environments are skillfully crafted to mirror real-world conditions, proving invaluable for machine and deep learning applications. By using Unity, we can effectively replicate real-world objects and scenarios in a controlled virtual setting, offering a unique advantage for our research purposes.

In order to achieve enhanced accuracy and realism in the dataset, this project has integrated authentic construction site elements into the Unity-based synthetic environment. This method entails an in-depth examination of real-life construction environments and a precise recreation of their essential elements within the Unity 3D modeling platform. For instance, typical construction site elements such as containers, concrete blocks, road barriers, and electric cables (as illustrated in Figure 1) have been incorporated. These components were meticulously modeled in Unity to resemble their real-life equivalents closely. Such detailed replication is vital to ensure that the synthetic dataset accurately reflects the complexity and specific characteristics of real construction sites, making it a valuable tool for our research.

To enhance the diversity of the datasets, key variables within the construction scenes, including lighting conditions, object positioning, and camera perspectives, have been

meticulously adjusted. The dataset emulates different times of the day, featuring varying daylight conditions such as morning, afternoon, and night, as depicted in Figure 2. This diversity in lighting is essential to create a dataset that authentically represents the dynamic visual environment of a construction site across different times. The unique lighting and shading challenges associated with various times of the day are critical for training deep-learning models to perform effectively and adaptively in a range of lighting conditions. This careful engineering of scene variables significantly contributes to the robustness and applicability of the dataset in real-world scenarios.

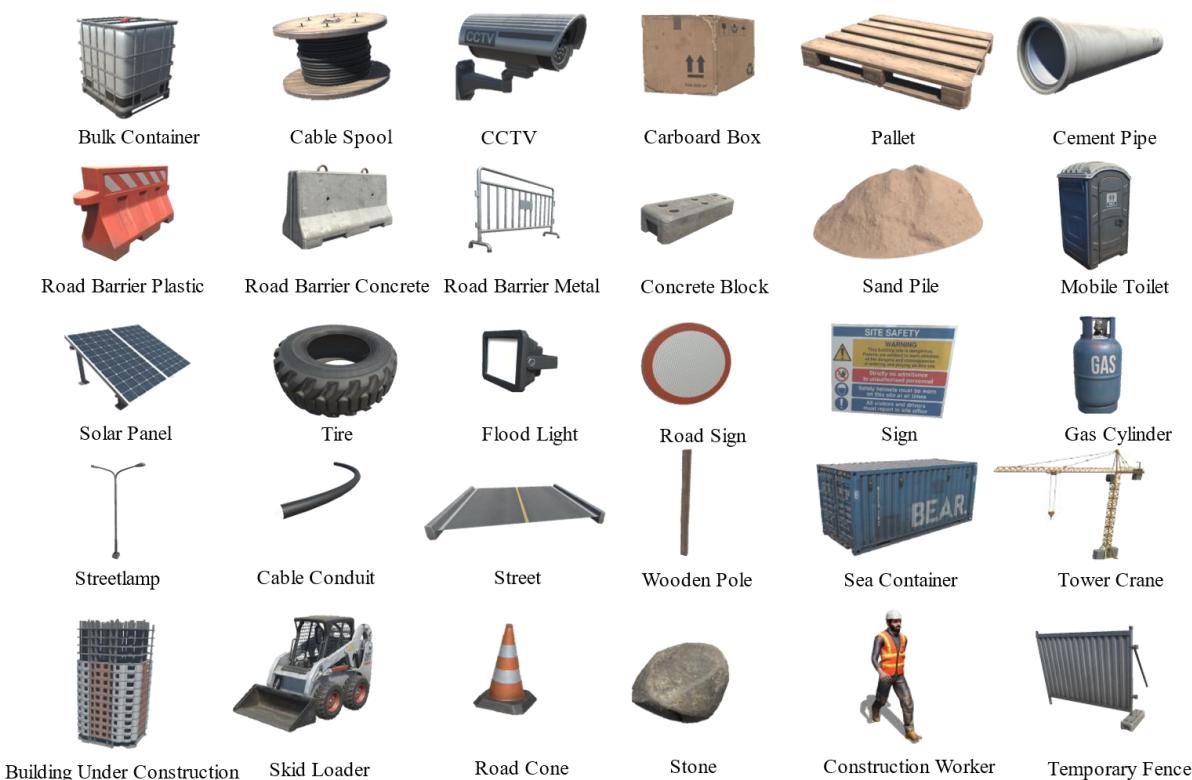


Figure 1. Objects in Unity for dataset creation.

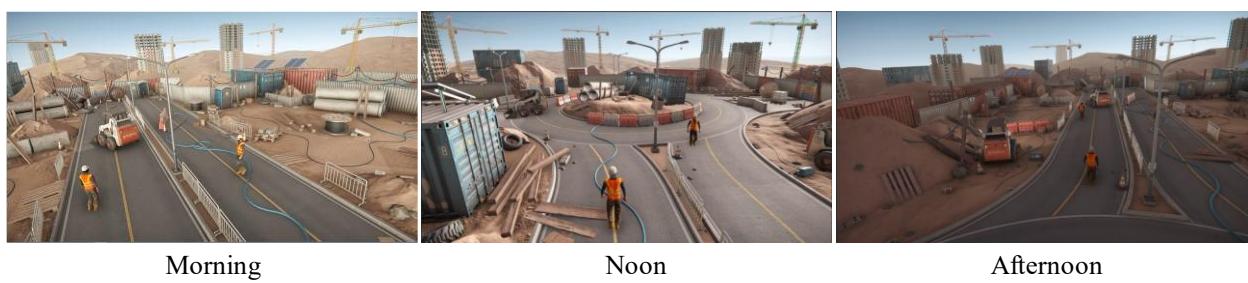


Figure 2. Different daylight conditions simulated in Unity.

Diverse camera angles have been utilized in the Unity Perception package's image capture process to augment the dataset's realism, as illustrated in Figure 3. This approach ensures that the dataset accurately represents the varied perspectives and vantage points typically encountered in

a real construction environment. The employment of multiple camera angles allows for a thorough representation of the construction site, capturing detailed views of equipment and environmental conditions from various angles. This method facilitates the creation of a broad spectrum of scenarios, including some that may be uncommon or challenging to document in actual construction settings but are essential for an exhaustive understanding of these environments. Moreover, the integration of animated worker figures within the Unity 3D model substantially enhances the dataset. The inclusion of these dynamic human models introduces a range of human poses and movements, thereby adding a critical dimension of complexity and authenticity. This addition of animated workers enriches the dataset, making it a more effective tool for studying and understanding the intricacies of construction site dynamics and human interactions within these settings.



Figure 3. Different camera perspectives used for image capturing.

The creation of the dataset culminates in utilizing the bounding box object detection feature of the Unity Perception package. This phase is crucial as it involves identifying and labeling various objects within the 3D environment. Object detection is essential in tailoring the dataset to the needs of deep learning applications. The dataset provides well-structured and clear data that is optimally suited for training and evaluating deep-learning models by precisely delineating objects with bounding boxes. These models are then equipped to recognize and interpret construction objects in diverse situations and settings, drawing on the varied scenarios depicted in the dataset. The dataset's inclusion of various scenarios, such as differing lighting conditions and camera perspectives, presents a comprehensive challenge to the deep learning algorithms (see Figure 4). These varied scenarios simulate a broad spectrum of real-world conditions, enhancing the models' adaptability and robustness. This step ensures the dataset's relevance for practical applications and significantly contributes to the development of advanced deep-learning models capable of effectively navigating and understanding complex real-world construction environments.



Figure 4. Sample annotations in the dataset.

The object detection dataset includes a diverse range of construction-related objects. As indicated in Figure 5, key elements such as 'wooden_pole', 'cable_conduit', and 'stone' are prominently featured, indicating their importance in model training. The dataset also contains significant representations of 'sand_pile', 'temporary_fence', 'concrete_block', and 'street', contributing to its diversity. Although less frequent, objects like 'sea_container' and 'road_barrier_metal' are still well-represented, ensuring a comprehensive training scope. The inclusion of the 'worker' category highlights human activity in construction environments. Rarer objects such as 'road_sign' and 'gas_cylinder' are also present, aiding in fine-tuning the model's accuracy. The dataset thoroughly compiles construction-specific objects, balancing common and uncommon scenarios for effective object detection training and evaluation. The accompanying visualizations, including a boxplot and heat maps, reveal key insights. The boxplot shows a high concentration of bounding box overlaps around the median, suggesting that objects are often centered within the image frame. This underscores the need for algorithms capable of distinguishing closely spaced items. The heat maps further illustrate this central concentration of object locations and the dataset's typical dimensions of bounding boxes. Most objects tend to be wider than tall, with a relative scarcity of very large objects. These insights are crucial for understanding the dataset's composition and guiding the development of more effective object detection algorithms.

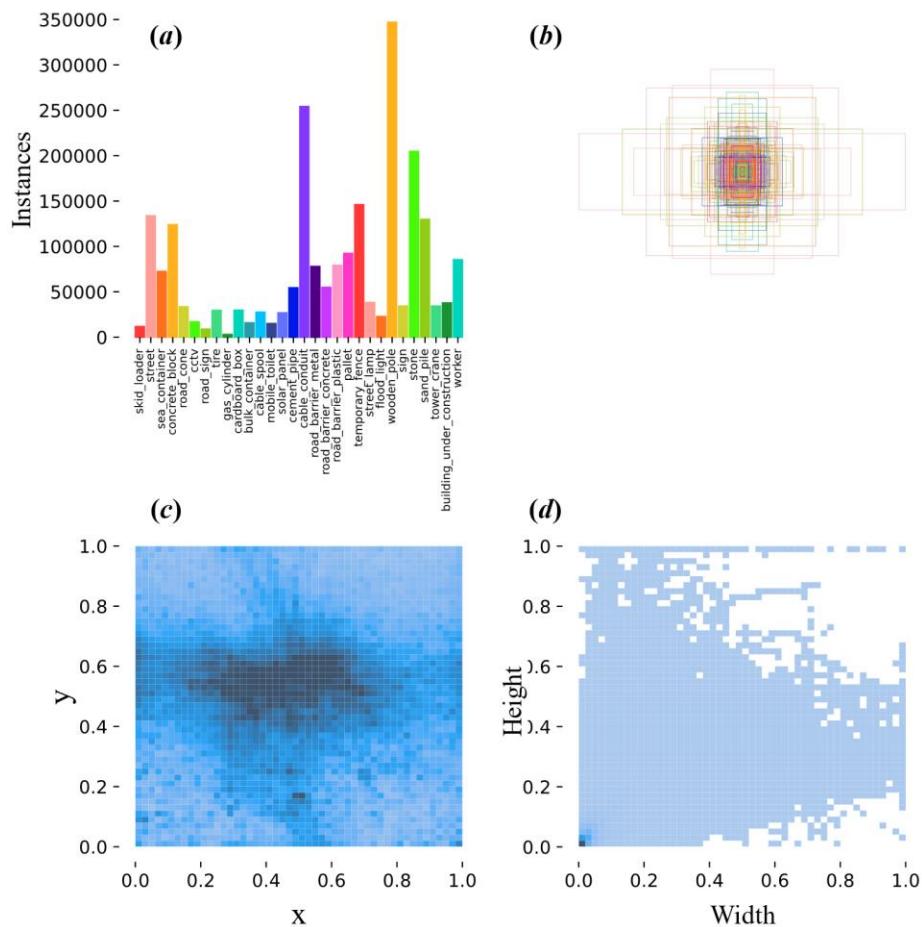


Figure 5. Dataset statistics

DATASET PERFORMANCE

Network selection. In this study, we have conducted an analysis of the performance of YOLOv5, YOLOv6 (Li et al. 2022), YOLOv7 (Wang et al. 2023a), YOLOv8, and RT-DETR networks on the construction scene understanding dataset. Our selection criteria were carefully considered, with a focus on their real-time processing abilities. This is of utmost importance for dynamic and unpredictable construction sites where rapid processing and object detection are crucial. The YOLO series has always been known for its speed and efficiency in real-time object detection, making it the perfect choice for such settings. Each version, from YOLOv5 to YOLOv8, offers iterative improvements in speed and accuracy, providing a wide range of performance characteristics for analysis. Additionally, RT-DETR complements the YOLO models with its transformer-based approach, which has proven to be highly effective in real-time detection tasks. These fast-processing abilities are critical for practical applications in construction scenes, where timely and accurate object detection significantly impacts safety and operational efficiency. Our second criterion was to ensure that these algorithms represent state-of-the-art object detection performance. The YOLO series has consistently set new benchmarks in the field, pushing the limits of object detection. Their innovative architectures and advanced learning algorithms consistently deliver top-tier performance across various settings. RT-DETR introduces a novel perspective with its transformer-based approach. By incorporating these state-of-the-art models, our analysis benefits from their advanced capabilities and provides a comprehensive overview of current technological advancements in object detection.

Evaluation metrics. The accuracy of the network's detections on the benchmark dataset is measured by the metric of precision. This metric is calculated by dividing the number of true positives (correct detections) by the total number of detections, including both true and false positives. A high precision score indicates that the network can identify objects with minimal false detections, ensuring that only relevant objects are detected. Another important metric is recall, which measures the network's ability to identify all relevant objects in an image. This metric is calculated by dividing the number of true positives by the sum of true positives and false negatives. Essentially, recall measures how many actual objects the network can detect. A high recall rate suggests that the network captures a significant portion of the relevant items in the image, missing a few objects.

The mAP₅₀ (mean Average Precision at IoU=0.50) thoroughly evaluates a model's object detection capabilities. This metric calculates the average precision for detections at an Intersection over Union (IoU) threshold of 0.50 across various categories. IoU measures the extent to which the predicted bounding box overlaps with the ground truth box, and a threshold of 0.50 indicates a minimum of 50% overlap. The mAP₅₀ is an average of these precision values across all categories, providing insight into the model's ability to accurately detect objects at this level of overlap. The mAP₅₀₋₉₅ (mean Average Precision at IoU=0.50:0.95) expands on this concept by covering a wider range of IoU thresholds, from 0.50 to 0.95, in increments of 0.05. This metric is a comprehensive measure that averages the mean average precision across these varying IoU levels. It offers a more complete assessment of the model's object detection performance, taking into account different levels of bounding box accuracy and the challenges of detecting objects with varying sizes and degrees of overlap. The mAP₅₀₋₉₅ is therefore considered a more rigorous and all-encompassing metric, reflecting the model's proficiency in accurately detecting objects under a broader range of conditions.

Performance. The comparison of YOLOv5L, YOLOv6L, YOLOv7L, YOLOv8L, and RT-DETR-L models on both the validation set (Table 1) and testing set (Table 2) offers valuable

insights into their respective object detection performance. YOLOv8L stands out as the most efficient model on the validation set, achieving the highest precision (0.961) and mAP₅₀ (0.802), as well as a strong mAP₅₀₋₉₅ score (0.707), all while operating at a moderate frame rate 71 FPS and 44M parameters. YOLOv5L strikes a good balance with a precision of 0.957 and mAP₅₀ of 0.789, operating at 54 FPS. Although YOLOv6L has the highest FPS (99), it exhibits a noticeable drop in precision (0.936) and mAP₅₀ (0.708) and has the largest number of parameters (59M). YOLOv7L, with the least number of parameters (36M), records lower scores across precision, mAP₅₀, and mAP₅₀₋₉₅. RT-DETR-L shines in its high FPS of 114 but exhibits moderate scores in precision and mAP₅₀.

Table 1. Network performance on the validation set. (The FPS is obtained from (Lv et al. 2023))

Model	Param. (M)	FPS _{bs=1} *	Precision	Recall	mAP ₅₀	mAP ₅₀₋₉₅
YOLOv5L	46	54	0.957	0.709	0.789	0.686
YOLOv6L	59	99	0.936	0.628	0.708	0.596
YOLOv7L	36	55	0.894	0.616	0.673	0.516
YOLOv8L	44	71	0.961	0.719	0.802	0.707
RT-DETR-L	32	114	0.909	0.608	0.722	0.552

The testing set (Table 2) shows similar trends as the validation set, with YOLOv8L maintaining its lead in precision (0.960) and mAP₅₀ (0.795), demonstrating its robustness and consistency across different datasets. YOLOv5L closely follows, displaying only a slight decrease in metrics compared to its validation set performance. YOLOv6L and YOLOv7L continue to exhibit lower efficiency in precision and mAP₅₀, while RT-DETR-L, despite having the highest FPS, lags slightly in precision and mAP scores compared to the leading models. These results highlight the trade-offs between model size, speed, and detection accuracy, with YOLOv8L and YOLOv5L presenting a more balanced profile suitable for efficient real-time object detection tasks.

Table 2. Network performance on the testing set. (The FPS is obtained from (Lv et al. 2023))

Model	Param. (M)	FPS _{bs=1}	Precision	Recall	mAP ₅₀	mAP ₅₀₋₉₅
YOLOv5L	46	54	0.954	0.696	0.781	0.676
YOLOv6L	59	99	0.929	0.617	0.701	0.588
YOLOv7L	36	55	0.884	0.601	0.653	0.499
YOLOv8L	44	71	0.960	0.709	0.795	0.698
RT-DETR-L	32	114	0.911	0.621	0.729	0.559

Figure 6 demonstrates the YOLOv8L model's effectiveness in detecting objects within construction site settings. Bounding boxes with associated confidence scores are visible around each item the model recognizes, showcasing its ability to identify and classify various elements within these complex scenes. The range of confidence scores reflects the model's precision, and the successful identification of numerous objects across different scenes underlines the model's robustness and versatility in practical applications.



Figure 6. Example of YOLOv8L detection on construction sites.

After analyzing the performance of different variations of the YOLOv8 model (YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l), it's clear that there is a direct correlation between model size and detection efficiency. The YOLOv8n model, the smallest with only 3M parameters, shows respectable performance with a precision of 0.853 and a mAP50 of 0.640, though it has the lowest scores in recall and mAP50-95. As the model size increases to YOLOv8s (11M parameters), all metrics see a notable improvement, with precision rising to 0.916 and mAP50 to 0.716. This trend continues with YOLOv8m (26M parameters), which achieves even higher precision (0.945) and mAP50 (0.767), indicating a more robust detection capability. The largest model, YOLOv8l (44M parameters), stands out with the highest scores across all metrics: precision peaks at 0.960, recall at 0.709, mAP50 at 0.795, and mAP50-95 at 0.698. This progression demonstrates that as the model size increases, the YOLOv8 network becomes more proficient in accurately detecting and classifying objects, achieving higher precision and mean average precision scores. It's important to keep in mind that a balanced approach in model design is crucial, where the increase in parameters contributes to enhanced detection capabilities while considering computational resources and processing speed.

Table 3. YOLOv8 performance on the testing set

Model	Param. (M)	Precision	Recall	mAP ₅₀	mAP ₅₀₋₉₅
YOLOv8n	3	0.853	0.548	0.640	0.478
YOLOv8s	11	0.916	0.626	0.716	0.579
YOLOv8m	26	0.945	0.683	0.767	0.654
YOLOv8l	44	0.960	0.709	0.795	0.698

CONCLUSION

The research conducted in this study has been successful in showcasing the immense potential of synthetic datasets, which have been generated using the Unity game engine, in the field of civil engineering. Specifically, the study highlighted the benefits of using such datasets for the advancement of deep learning applications in construction site understanding. The creation of the dataset was carried out with meticulous attention to detail, incorporating a broad range of scenarios, lighting conditions, and camera angles. Moreover, the bounding box object detection technique was employed with precision, resulting in an all-encompassing dataset that mirrors the complexity and variability of actual construction sites. This dataset not only helps to overcome the limitations of traditional data collection methods but also provides a sturdy platform for training and validating deep learning models. The use of these models has shown promising results in accurately interpreting and navigating construction environments, providing valuable insights that can lead to further innovations in construction site management, safety, and efficiency. Overall, this research has opened up exciting new avenues for leveraging the power of artificial intelligence and synthetic data to transform the landscape of civil engineering.

REFERENCES

- Hu, P. 2022. "Application of Machine Learning Algorithms in the Development of Civil Engineering Software." *2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI)*, 975–978. Changchun, China: IEEE.
- Li, C., et al. 2022. "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications." arXiv.
- Lv, W., Y. Zhao, S. Xu, J. Wei, G. Wang, C. Cui, Y. Du, Q. Dang, and Y. Liu. 2023. "DETRs Beat YOLOs on Real-time Object Detection." arXiv.
- Mahmoud, A. S., M. H. Ahmad, and Y. M. Yatim. 2020. "Overview of Safety Performance in the Construction Industry." *Proceedings of International Structural Engineering and Construction*, (Y. Vacanas, C. Danezis, A. Singh, and S. Yazdani, eds.), 7 (1).
- Park, J., and Y. K. Cho. 2017. "Development and Evaluation of a Probabilistic Local Search Algorithm for Complex Dynamic Indoor Construction Sites." *J. Comput. Civ. Eng.*, 31 (4): 04017015.
- Song, Z., et al. 2023. "Synthetic Datasets for Autonomous Driving: A Survey." arXiv.
- Stone, J. R., D. I. Blockley, and B. W. Pilsworth. 1991. "Managing risk in civil engineering by machine learning from failures." [1990] *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*, 255–259. College Park, MD, USA: IEEE Comput. Soc. Press.
- Timofeeva, S. S., D. V. Ulrikh, and N. V. Tsvetkun. 2017. "Professional Risks in Construction Industry." *Procedia Engineering*, 206: 911–917.
- US Bureau of Labor Statistics. 2022. *National Census of Fatal Occupational Injuries in 2021*. 2022.
- Wang, C.-Y., A. Bochkovskiy, and H.-Y. M. Liao. 2023a. *YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors*. 7464–7475.
- Wang, M., D. Hu, J. Chen, and S. Li. 2023b. "Underground infrastructure detection and localization using deep learning enabled radargram inversion and vision based mapping." *Automation in Construction*, 154: 105004.

Deep Learning Enhanced Crack Detection for Tunnel Inspection

Sultan Al Shafian, S.M.ASCE¹; Da Hu, Ph.D., A.M.ASCE²; and Wenbing Yu, Ph.D.³

¹Dept. of Civil and Environmental Engineering, Kennesaw State Univ., Marietta, GA.
Email: salshaf1@students.kennesaw.edu

²Dept. of Civil and Environmental Engineering, Kennesaw State Univ., Marietta, GA.
Email: dhu3@kennesaw.edu

³School of Civil Engineering, Chongqing Jiaotong Univ., Chongqing, China.
Email: yuwb@cqjtu.edu.cn

ABSTRACT

The study utilizes deep learning technology to improve crack detection in tunnels, surpassing the limitations of manual inspection methods. The deep learning model identifies cracks in high-resolution tunnel images, detecting even minor defects often overlooked in manual inspections. The model ensures unparalleled accuracy and consistency and is customized to tackle the unique challenges of tunnel environments, such as complex surface textures and varying lighting. The approach's effectiveness is evident from its remarkable performance metrics, including an F1 score of 0.873, a mean Intersection over Union (mIoU) of 0.794, a precision of 0.760, and a recall of 0.746. These exceptional results represent a significant advancement in tunnel inspection, leading to more reliable and comprehensive evaluations of structural integrity. The integration of deep learning technology for tunnel safety marks a crucial milestone in enabling more effective maintenance strategies and reducing risks associated with structural defects.

INTRODUCTION

Underground tunnels play a crucial role in modern transportation systems, providing direct routes that bypass geographical barriers and dense urban infrastructure, facilitating efficient transit. They enhance connectivity by linking disparate regions, thus fostering economic development through improved access to markets, resources, and labor pools. By diverting traffic below ground, tunnels alleviate surface congestion, improving the overall quality of urban life and reducing environmental impact (Tender et al. 2017; Ahmad et al. 2018). Using tunnels for subterranean transport networks, such as metros and trains, supports mass transit solutions essential for sustainable urban growth. Furthermore, underground tunnels are integral to the strategic planning of future smart cities, enabling multi-level transportation frameworks that optimize space usage and traffic flow. Typically, the construction of tunnels incurs high costs, and they are generally engineered to fulfill their intended function for a minimum duration of a century. So, the structural integrity of underground tunnels is a paramount concern for civil engineers during their service life. As crucial conduits for transportation and logistics, their reliability directly impacts economic vitality and public safety.

However, many structural complications arise in existing tunnels, attributable to geological circumstances, construction methodologies, and various other factors, potentially precipitating grave safety concerns. Research indicates that the formation of lining cracks within tunnels is a prevalent issue, both during their construction phase and operational tenure, posing a substantial threat to the integrity and stability of the tunnel frameworks (Min et al. 2021). The presence of

cracks can detrimentally influence the longevity and load-bearing competency of the tunnel lining, potentially precipitating localized structural compromise or, in instances of pronounced severity, a comprehensive destabilization. The timely and accurate identification of lining cracks and their expedient rectification stands as a pivotal engineering challenge. At present, the conventional methodologies for tunnel inspections predominantly hinge on manual visual inspections—a process that is not only laborious and time-consuming but also marred by the fallibility of human judgment and the vagaries of environmental conditions within the subterranean passages.

Advanced technologies have been introduced in the field of civil engineering inspections, leading to the exploration of automated systems for crack detection. Deep learning technology, a subset of artificial intelligence, has emerged as a revolutionary tool in this domain, offering improved accuracy and efficiency over traditional inspection methods. This research paper focuses on the application of a sophisticated encoder-decoder network designed to detect structural cracks in tunnel environments. This approach utilizes high-resolution imagery, potentially sourced from drones, to enable comprehensive scans of tunnels' inner linings without the limitations of human inspectors. Additionally, the paper explores the limitations of manual inspection methods, emphasizing the need for an automated, deep learning-based solution. The subsequent discussion will detail the development of the proposed deep learning model, explain its operational framework, and highlight the importance of transitioning from manual to automated tunnel inspections.

LITERATURE REVIEW

The integrity of tunnel linings is paramount for the safe operation of transportation and utility conduits, and cracks within these structures can have far-reaching and detrimental impacts. Cracks may serve as precursors to more severe structural failures, compromising the tunnel's ability to withstand geotechnical pressures and potentially leading to collapses or breaches that can endanger human lives and disrupt vital services. They facilitate the ingress of water and chemicals that can accelerate the degradation of the concrete and corrosion of embedded reinforcement, thereby reducing the tunnel's lifespan and increasing maintenance costs. A crack in the tunnel lining can happen for various reasons. For example, during shield tunnel construction, the concrete segments that are adjoined by circular and lengthwise joints frequently succumb to the detrimental forces of jacking thrust and grouting pressure, which, while integral to their construction, often lead to prevalent cracking within the tunnel lining (Yuan et al. 2013; Shi and Li 2015). Investigations have also shown that inadequate lining thickness and empty spaces behind the lining at the top are the primary reasons for cracks in the lining for conventional sprayed concrete tunnels. These voids and uncompressed areas can reduce and alter the direction of the bending moment, potentially leading to cracking on the inner surface of the lining at the tunnel arch's shoulders (Yang and Fu n.d.).

Concrete cracking not only diminishes the structural load-bearing capability but also impacts the resistance to fatigue resistance (Zhao et al. 2019). Deep cracks are prone to widening and extending, potentially leading to new fractures in different areas, which can significantly jeopardize the tunnel's traffic safety (Xu et al. 2020). For this reason, engineers dedicate a considerable amount of time to meticulously inspecting and pinpointing cracks in tunnels to ensure timely repairs. Structural tunnel inspections, traditionally reliant on the periodic visual assessments of skilled inspectors to spot defects like cracks and spalling, are critical yet must be

non-invasive, hence the preference for non-destructive techniques over destructive ones. Such methods range from visual to sonic, electrical, and radar approaches, necessitating the physical presence of personnel in potentially dangerous environments, which is costly, time-consuming, and susceptible to human error.



Figure 1. Example of visual crack detection (Hirota et al. 2016)

Various alternative methods have been extensively explored to address the shortcomings of traditional visual inspections. Traditional measuring tools like strain gauges and Linear Variable Differential Transducers (LVDT) were initially examined for recording deformations and cracks (Belloni et al. 2020). Line scan cameras and laser scanning with 3D LiDAR were also used by some researchers to automate crack detection inside the tunnel (Lee et al. 2007; Zhang et al. 2014). Over the past twenty years, considerable research has been devoted to exploring the efficacy of image-based techniques in detecting and measuring cracks, establishing them as a powerful tool for monitoring infrastructure. Techniques like edge-detection algorithms, mathematical morphology, and various machine learning models, including Principal Component Analysis, Extreme Learning Machine, and Support Vector Machine, have been employed. Moreover, the advent of deep learning, specifically Convolutional Neural Networks (CNNs), has greatly enhanced image-based monitoring. Deep learning strategies like object detection, image classification, semantic segmentation, etc., have been utilized in the crack detection (Hu and Yee 2023; Zhou et al. 2023). While object detection identifies crack locations, image classification offers more detail, though its precision is contingent on image division.

Semantic segmentation delivers the most precise results by classifying each pixel associated with a crack (Liu et al. 2019). Recent advances in deep learning have enhanced crack segmentation in images, utilizing architectures like FCN and U-Net to extract spatial features for image models, leading to state-of-the-art results with models such as SDDNet, achieving high accuracy on complex backgrounds. However, these models require extensive datasets and labeling and sometimes inaccurately segment continuous cracks as disjointed, a problem that recent methods like the improved PANet with integrated A* algorithm address to reduce crack quantification errors (Dang et al. 2022). Recent advancements in computer vision have been driven by convolutional neural networks (CNNs), which excel in image classification and segmentation tasks; Fully Convolutional Networks (FCNs) and U-Net are notable examples that

have streamlined semantic segmentation by adapting CNNs to output segmentation maps for images of any size, enhancing speed and reducing data requirements (Ren et al. 2020). Despite their success, the unique morphological features of concrete cracks present distinct challenges that necessitate further refinement of these network architectures for optimal performance in structural health monitoring and crack detection.

TUNNEL CRACK SEGMENTATION NETWORK

The tunnel crack segmentation employs the SegNeXt architecture (Guo et al. 2022), which is intricately designed for precise semantic segmentation. This architecture adopts an encoder-decoder model, where the encoder, named MSCAN, is structured following a pyramid design. The standout feature of MSCAN is the multi-scale convolutional attention (MSCA) module. MSCA innovatively combines depth-wise convolution, multi-branch depth-wise strip convolutions, and a 1×1 convolution. The 1×1 convolution output serves as attention weights, a critical aspect differentiating MSCA from traditional self-attention mechanisms. MSCAN is structured hierarchically into four stages with progressively lower spatial resolutions. This structure is pivotal for handling different scales of image features, which is essential in detecting cracks under varied tunnel conditions. Each stage in MSCAN includes a down-sampling block followed by a stack of building blocks, with batch normalization used in each block to optimize segmentation performance.

The SegNeXt model includes four variations of the MSCAN encoder (Zhang et al. 2021) - MSCAN-T, MSCAN-S, MSCAN-B, and MSCAN-L - each tailored to different computational needs and segmentation accuracies. The SegNeXt decoder, distinct from its counterparts, is designed to process features from the last three stages of the encoder. This is a strategic choice to avoid the drawbacks of incorporating excessive low-level information and reducing computational overhead. The decoder utilizes a lightweight Hamburger module to model global context effectively, enhancing the model's overall efficiency in terms of performance and computation. This approach allows SegNeXt to outperform SegFormer in applications like tunnel crack segmentation, where accuracy and computational efficiency are crucial.

EXPERIMENT AND RESULTS

Dataset. Recently, there has been a lot of attention on creating algorithms that can identify and separate concrete cracks. The concrete crack segmentation dataset created in (Bianchi and Hebdon 2022) was utilized for this study. This dataset contains 9,584 images that have cracks, with images without cracks excluded from the dataset. The Labeled Cracks in the Wild dataset was also constructed in (Bianchi and Hebdon 2022), which features more global scenes obtained from bridge inspection reports. These two datasets were instrumental in detecting and locating concrete cracks, with the LCW dataset providing a more comprehensive perspective of the inspection process. In total, this study utilized a combined collection of 10,543 images. Figure 2 shows some example images with annotations in the dataset.

Implementation details. The network was trained on a workstation running Ubuntu 20.04 with an Intel i9-10900X CPU, 32 GB RAM, and NVIDIA RTX A5000 GPU for training and testing the models. The network was expertly trained using the AdamW optimizer, with a learning rate initialized at 1e-6 that increases to 6e-5 after 1500 iterations with a linear warmup. A polynomial decay learning rate schedule was employed to prevent overfitting with an end

learning rate of 0. The batch size was set to 12, and pre-trained weights from ImageNet were utilized for the backbone. The pre-trained weights were fine-tuned on the combined concrete crack dataset without freezing any weights. All images were resized to 512 x 512 pixels during training for efficient processing. The dataset was split into training and validation sets to ensure thorough testing, with 85% of the images used for training and 15% for validation. To enhance variability, the training dataset underwent random horizontal image flipping. We also utilized RandomDistort to introduce variations in image brightness, contrast, and saturation, with each parameter set at a range of 0.5. During training, the model weights were adjusted using the cross-entropy loss function, which is particularly suitable for tasks that require measuring the difference between predicted probabilities and actual labels for each pixel. This function heavily penalizes misclassifications, ensuring that the model is incentivized to accurately identify and segment intricate patterns of cracks in various images.

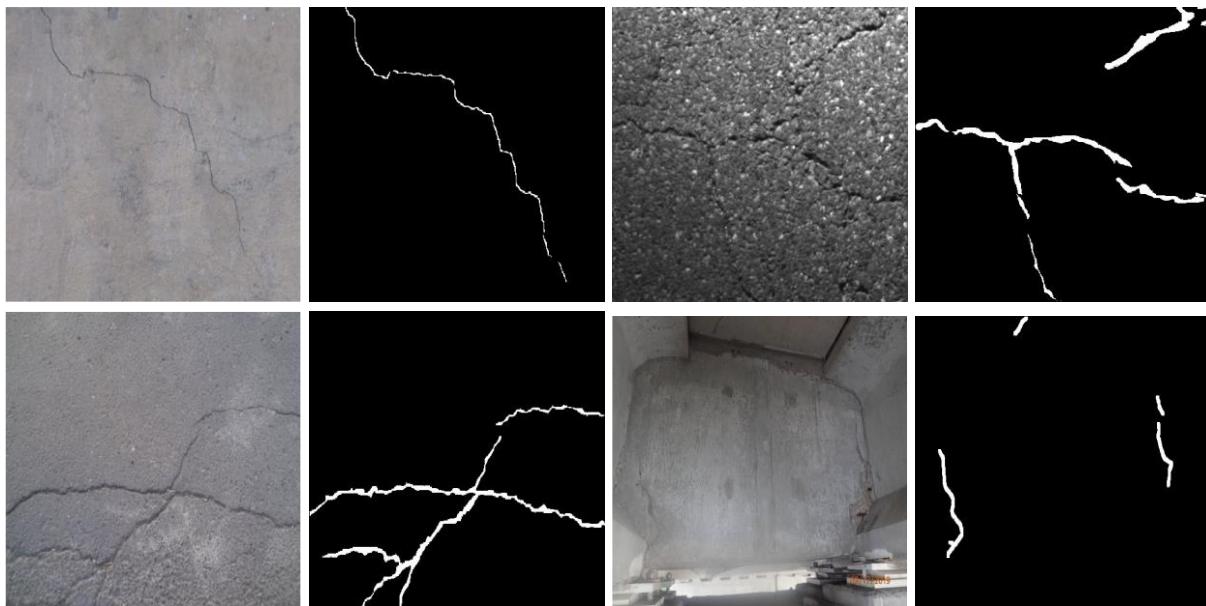


Figure 2. Example of images in annotation datasets.

Metrics. Evaluating the performance of the crack segmentation model is essential in determining its effectiveness. For quantitative accuracy assessment, metrics like mean intersection over union (mIoU), precision, recall, and dice coefficient (DSC) are utilized. The IoU metric, a standard in image segmentation, assesses the overlap between the predicted mask and the ground truth on a 0 to 1 scale, with a higher mIoU indicating better accuracy. mIoU is calculated in Eq. (1).

$$mIoU = \frac{1}{2} \sum_{i=1}^2 \frac{TP_i}{TP_i + FP_i + FN_i} \quad (1)$$

The F1, another key metric, measures the similarity between the predicted mask and the ground truth, ranging from 0 to 1, with higher values signifying superior performance. It is particularly useful for evaluating models in imbalanced classes or when precision and recall are critical. The F1 is defined in Eq. (2).

$$DSC = \frac{2TP}{2TP+FP+FN} \quad (2)$$

Precision and recall, also crucial in image segmentation, assess the model's accuracy in identifying true positives. Precision is the ratio of true positives to all positive predictions, while recall is the ratio of true positives to all actual positives in the dataset. Both range from 0 to 1, with higher values denoting better performance. The definitions of precision and recall are given in Eq. (3) and Eq. (4).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

Results. Figure 3 presents two subplots that track the progress of the segmentation model during its training phase. Subplot (a) showcases the loss function of the training set, demonstrating a consistent decrease from an initial value of approximately 0.6 to around 0.02 over 37,500 iterations. This suggests that the model has become increasingly accurate in its predictions of the training data. In contrast, Subplot (b) highlights the model's performance on the validation set, measured using the mIoU metric. The mIoU begins at just over 0.56 and shows a gradual upward trend, with minor fluctuations, ultimately stabilizing at approximately 0.79 after 37,500 iterations.

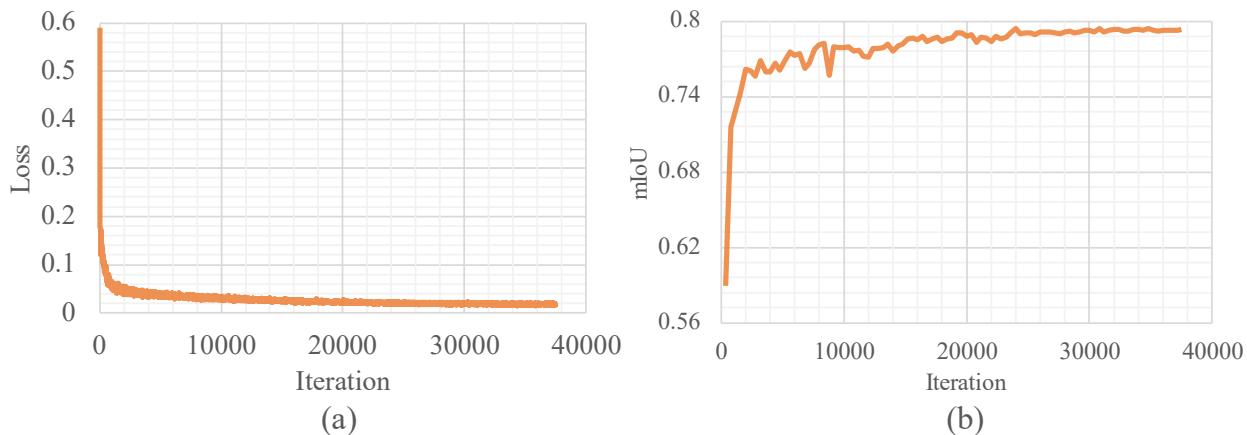


Figure 3. Convergence of the network during the training. (a) Loss value vs. iteration; (b) mIoU vs. iteration.

The network's crack segmentation achieved impressive results with an F1 score of 0.873, a mIoU of 0.794, a precision of 0.760, and a recall of 0.746, as shown in Figure 4. The network excelled in detecting and outlining cracks across various textures and substrates, even with complex background noise and lighting conditions. Its deep learning capabilities allowed for accurately identifying irregular crack patterns and differentiating them from surface textures. This approach demonstrates high accuracy in crack identification and robustness against false positives, which is essential for monitoring structural health. The model's consistent performance across different scenarios showcases strong generalization capabilities, making it a reliable tool for automated crack detection in tunnel monitoring.

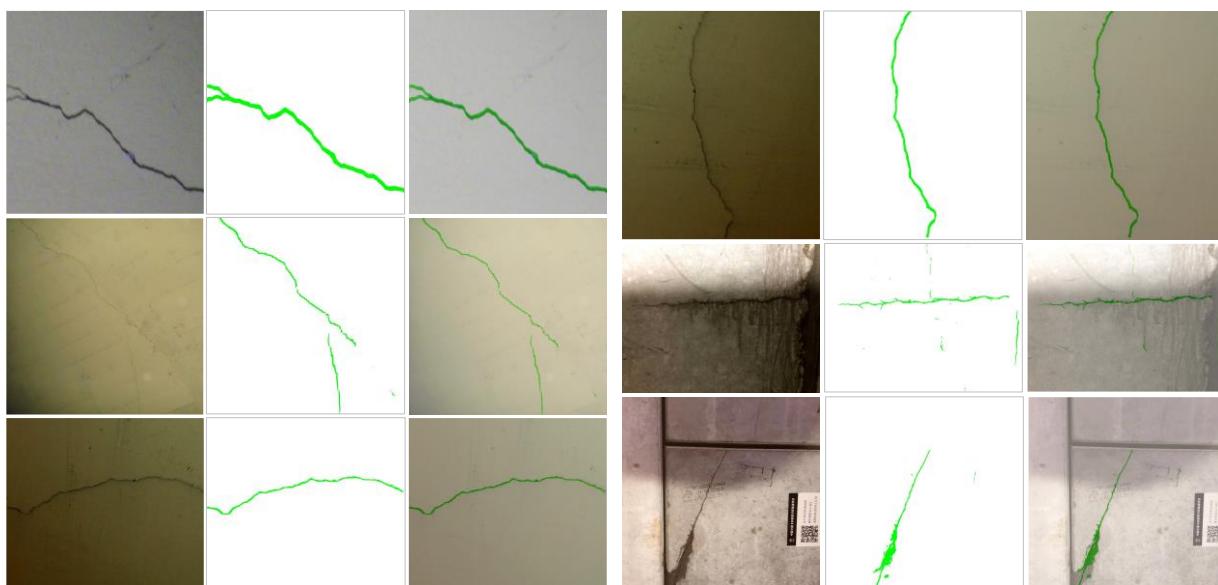


Figure 4. Examples of crack segmentation on the tunnel lining surface.

DISCUSSION

Backbone network. The SegNeXt architecture's performance was evaluated across different backbone networks, namely MSCAN-T, MSCAN-S, MSCAN-B, and MSCAN-L, to determine the impact of backbone complexity on segmentation outcomes. The largest model, SegNeXt-L with MSCAN-L backbone, achieved the highest mIOU of 0.798 and an F1 score of 0.875. This demonstrates the superior capability of the model to accurately segment cracks while balancing both the precision of predictions and the recall of crack features. SegNeXt-B, incorporating the MSCAN-B backbone, demonstrated a slightly lower mIOU of 0.794 and a decreased F1 score of 0.873, despite a reduction in parameters and GFLOPs. However, it maintained a high level of precision at 0.760 and recall at 0.746. The SegNeXt-S model with the MSCAN-S backbone showed a further reduction in mIOU to 0.784, with a precision of 0.764 and a recall of 0.713, resulting in an F1 score of 0.865. Finally, the smallest and least computationally intensive model, SegNeXt-T, with only 4.2 million parameters and 4.6 GFLOPs, registered an mIOU of 0.779, precision of 0.750, recall of 0.711, and an F1 score of 0.861.

Based on the results, the selection of MSCAN-B as the backbone for the SegNeXt model is justifiable due to its exceptional balance between segmentation performance and computational efficiency. Although MSCAN-L offers marginally higher accuracy, MSCAN-B achieves comparable performance metrics with significantly lower computational demands. This balance is critical in practical applications where the costs associated with computational resources can be prohibitive. Therefore, the efficiency of MSCAN-B makes it the perfect choice for deployment in environments where computational resources are limited, but high accuracy in crack detection remains a crucial requirement.

Learning rate variation. The efficiency and performance of neural networks are greatly influenced by the learning rate. Our analysis of the SegNeXt-B network shows that setting the learning rate to 0.0005 yields the best results, achieving an impressive mIOU of 0.801 and an F1 score of 0.878. This particular learning rate strikes a perfect balance, allowing the network to converge to an optimal solution without overfitting. This is evident from the improved precision

and recall. On the other hand, the default learning rate of 0.00006 used in initial training results in slightly suboptimal performance, as demonstrated by a decrease in mIOU and precision. Notably, reducing the learning rate further to 0.0001 closely approaches peak performance, indicating a narrow window for optimal learning rate settings in this context. Our findings highlight the importance of fine-tuning the learning rate to achieve the best possible segmentation accuracy with SegNeXt-B.

Table 1. Processing speed vs accuracy. The number of FLOPs (G) is calculated on the input size of 512×512

Model	Params. (M)	GFLOPs	mIOU	Precision	Recall	F1
SegNeXt-L	45.1	37.9	0.798	0.774	0.742	0.875
SegNeXt-B	26.7	21.9	0.794	0.760	0.746	0.873
SegNeXt-S	13.9	11.7	0.784	0.764	0.713	0.865
SegNeXt-T	4.2	4.6	0.779	0.750	0.711	0.861

Table 2. Impact of learning rate variation on SegNeXt-B performance metrics

Learning rate	mIoU	Precision	Recall	F1
0.001	0.791	0.772	0.725	0.870
0.0005	0.801	0.779	0.748	0.878
0.0001	0.799	0.778	0.741	0.876
0.00006	0.794	0.760	0.746	0.873

Limitations and future work directions. The successful implementation of the SegNeXt-B network for detecting cracks on concrete surfaces has shown potential, but it has also highlighted areas that require improvement. One primary limitation is the narrow scope of the dataset, which primarily features general crack patterns on concrete surfaces. This limited range may not represent the diverse crack types found in complex environments like tunnels. To enhance the model's generalizability and specificity, future research should aim to expand the dataset to include more varied scenarios, particularly from tunnel environments. Another significant limitation is the inability to measure crack dimensions, which is crucial for assessing structural integrity. Future network versions could integrate capabilities for dimensional analysis using advanced image processing techniques or additional sensor data.

Moreover, the network's variable performance under different lighting conditions poses a challenge. The model may struggle with consistent crack identification across lighting scenarios, affecting accuracy. Developing robust preprocessing algorithms to normalize lighting in input images or training the network to adapt to environmental variations could be useful to address this. Additionally, exploring the model's scalability and efficiency in processing large datasets or streaming data in real-time scenarios would be beneficial. Optimizing the model for real-time analysis would enhance its suitability for live infrastructure monitoring. Integrating the model with IoT devices and drones for automated monitoring and data collection would enable more frequent and widespread structural health assessments, particularly in hard-to-reach areas. Finally, enhancing the model's explainability is crucial, especially in safety-critical applications. Deep learning models must have transparent and interpretable decision-making processes. Therefore, future research should focus on making the SegNeXt-B network's predictions more

understandable to end-users. By addressing these limitations and exploring these directions, future research would improve the SegNeXt-B network's accuracy and applicability and advance AI-driven structural health monitoring, leading to more reliable and efficient infrastructure maintenance.

CONCLUSION

To summarize, this study has made significant progress toward automating the detection of cracks in tunnel environments. Through the use of advanced deep learning techniques, specifically the SegNeXt-B network, the authors have successfully demonstrated a reliable method for identifying and segmenting cracks on concrete surfaces, resulting in a high F1 score of 0.873 and mIoU of 0.794. This accuracy and efficacy are critical in the context of tunnel infrastructure assessment, where precise crack detection is crucial for preventing structural failures. These results have substantial implications for the field of structural health monitoring, suggesting that AI-driven methods can play a vital role in enhancing tunnel infrastructure safety and maintenance. However, challenges remain, such as diversifying the dataset, refining the model's adaptability to different lighting conditions, and integrating capabilities for quantifying crack dimensions. Optimizing the system for real-time analysis and extending its application to IoT and drone-based monitoring systems will also be key to realizing its full potential.

REFERENCES

- Ahmad, K. I., M. Abrar, S. Al Shafian, and H. M. Shahin. 2018. "Correlation Among the Soil Parameters of the Karnaphuli River Tunnel Project." *GEOMATE Journal*, 15 (48): 86–90.
- Belloni, V., A. Sjölander, R. Ravanelli, M. Crespi, and A. Nascetti. 2020. "Tack Project: Tunnel and Bridge Automatic Crack Monitoring Using Deep Learning and Photogrammetry." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B4-2020: 741–745.
- Bianchi, E., and M. Hebdon. 2022. "Development of Extendable Open-Source Structural Inspection Datasets." *Journal of Computing in Civil Engineering*, 36 (6): 04022039.
- Dang, L. M., H. Wang, Y. Li, Y. Park, C. Oh, T. N. Nguyen, and H. Moon. 2022. "Automatic tunnel lining crack evaluation and measurement using deep learning." *Tunnelling and Underground Space Technology*, 124: 104472.
- Guo, M.-H., C.-Z. Lu, Q. Hou, Z. Liu, M.-M. Cheng, and S.-M. Hu. 2022. "SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation." arXiv.
- Hirota, A., S. Nishiyama, T. Mizuguchi, and K. Ishimura. 2016. "Measurement of Crack Width Using Digital Photogrammetry for Evaluation of the Stability of Tunnel." *Journal of Japan Society of Civil Engineers, Ser. F1 (Tunnel Engineering)*, 72 (3).
- Hu, D., and T. Yee. 2023. "Integrated Framework for Bridge Crack Detection and Semantic BIM Model Generation Using Drone-Captured Imagery and Deep Learning Techniques." 167–175. American Society of Civil Engineers.
- Lee, S. Y., S. H. Lee, D. I. Shin, Y. K. Son, and C. S. Han. 2007. "Development of an Inspection System for Cracks in a Concrete Tunnel Lining." *Canadian Journal of Civil Engineering*, 34 (8): 966–975.
- Liu, Z., Y. Cao, Y. Wang, and W. Wang. 2019. "Computer Vision-Based Concrete Crack Detection Using U-Net Fully Convolutional Networks." *Automation in Construction*, 104: 129–139.

- Min, B., C. Zhang, W. Zhu, X. Zhang, and P. Li. 2021. "Influence of Cracks at the Invert on the Mechanical Behavior of the Tunnel Structures." *Thin-Walled Structures*, 161: 107405.
- Ren, Y., J. Huang, Z. Hong, W. Lu, J. Yin, L. Zou, and X. Shen. 2020. "Image-Based Concrete Crack Detection in Tunnels Using Deep Fully Convolutional Networks." *Construction and Building Materials*, 234: 117367.
- Shi, P., and P. Li. 2015. "Mechanism of Soft Ground Tunnel Defect Generation and Functional Degradation." *Tunnelling and Underground Space Technology*, 50: 334–344.
- Tender, M. L., J. P. Couto, and L. Bragança. 2017. "The Role of Underground Construction for The Mobility, Quality of Life And Economic and Social Sustainability of Urban Regions." *REM - International Engineering Journal*, 70 (3): 265–271.
- Xu, G., C. He, Z. Chen, C. Liu, B. Wang, and Y. Zou. 2020. "Mechanical Behavior of Secondary Tunnel Lining with Longitudinal Crack." *Engineering Failure Analysis*, 113: 104543.
- Yang, J. S., and J. Y. Fu. n.d. "Investigation and Rehabilitation of Lining Imperfections and cracks in an Operating Railway Tunnel."
- Yuan, Y., X. Jiang, and X. Liu. 2013. "Predictive Maintenance of Shield Tunnels." *Tunnelling and Underground Space Technology*, 38: 69–86.
- Zhang, W., Z. Zhang, D. Qi, and Y. Liu. 2014. "Automatic Crack Detection and Classification Method for Subway Tunnel Safety Monitoring." *Sensors*, 14 (10): 19307–19328.
- Zhang, X., X. Gao, L. He, and W. Lu. 2021. "MSCAN: Multimodal Self-and-Collaborative Attention Network for Image Aesthetic Prediction Tasks." *Neurocomputing*, 430: 14–23.
- Zhao, Y., J. Yang, and Y. Zhang. 2019. "Failure Behavior of Tunnel Lining Caused by Concrete Cracking: A Case Study." *Journal of Failure Analysis and Prevention*, 19 (4): 1158–1173.
- Zhou, Z., J. Zhang, C. Gong, and W. Wu. 2023. "Automatic tunnel lining crack detection via deep learning with generative adversarial network-based data augmentation." *Underground Space*, 9: 140–154. Elsevier.

Performance Evaluation of Bridge Approach Slabs and Joints through Sub-mm 3D Laser Imaging System

Guolong Wang, Ph.D., S.M.ASCE¹; Kelvin C. P. Wang, Dist.M.ASCE²;
Guangwei Yang, Ph.D., M.ASCE³; and Joshua Qiang, M.ASCE⁴

¹Research Associate, Western Transportation Institute, Montana State Univ., Bozeman, MT.
Email: guolong.wang@montana.edu

²Professor and Director, Western Transportation Institute, Montana State Univ., Bozeman, MT.
Email: kelvin.wang@montana.edu

³Research Associate, Virginia Tech Transportation Institute, Blacksburg, VA.
Email: garyyang@vt.edu

⁴Associate Professor, School of Civil and Environmental Engineering, Oklahoma State Univ.,
Stillwater, OK. Email: qiang.li@okstate.edu

ABSTRACT

Bridge approach slabs often experience differential settlement due to a higher settlement of embankment than bridge abutment under traffic. Severe differential settlement produces bumps and distress at approach slabs and joints, leading to riding discomfort and safety risks to public travelers. Surveying the condition of approach slabs and joints is crucial to evaluate the field performance of current designs and countermeasures to reduce approach bumps and distress throughout the lifespan of approach slabs. In this study, a total of 377 bridges located across Oklahoma with different designs and conditions of approach slabs were surveyed. The recent Pave3D 8K system featuring sub-mm 3D laser technology and inertial profiler was used to collect the 2D/3D images and inertial data from these bridges for evaluating approach bump and distress. Based on the bump levels of all approach slabs, the performance of the slab design, including slab length and skew angle, and bump mitigation method via asphalt overlay were evaluated. The results show that (1) approach slabs with lengths of 60–80 ft yield the least level of bumps; (2) approach slabs with a skew angle of 60° or departure slabs with 120° show less severe bumps than those with other angles; and (3) asphalt overlay is an effective method to minimize the bumps at bridge ends connecting with flexible pavement, while the road, approach slab, and bridge deck constructed entirely of concrete exhibit the smoothest transition.

Keywords: bridge approach slabs and joints, slab design, mitigation method, 2D/3D images, inertial data, field evaluation, bump.

INTRODUCTION

In the United States, around 25% of nationwide bridges have been reported suffering from differential settlement issues, which caused an annual maintenance cost of at least \$100 million (Briaud et al., 1997; Puppala et al., 2009). It is also reported that a considerable amount of the maintenance costs is primarily associated with the repair of the distress at joints that are more fragile and easily damaged by live loads than other components (Briaud, 1997; Arsoy, 1999). The differential settlement is primarily attributed to traffic-induced compression and settlement, resulting in bumps around the approach slab and bridge deck surface and finally leading to ride discomfort and safety risks to the traveling public.

The impact of bumps on approach slabs can range from causing discomfort to posing potential life-threatening risks, depending on the severity of the bumps. Up-to-date maintenance using remedial countermeasures or mitigation methods should be scheduled to treat the distressed approach slabs once the severity of the bumps becomes unacceptable. The maintenance countermeasures are variable and exhibit different levels of benefits and costs. In general, the replacement of the whole slabs is needed if the slabs are cracked to lose structural integrity. Otherwise, more economical remedial countermeasures can be applied.

Asphalt overlay is commonly used to mitigate excessive settlement at the end of approach slabs opposite the bridge end (Lu et al., 2018; Puppala et al., 2010). Moreover, laying asphalt overlay has even been a mandate for approach slab construction in Florida to ensure a smooth transition with flexible pavement (FDOT, 2012; Lu et al., 2018). Many departments of transportation (DOTs) such as Kentucky, Minnesota, Texas, and Oklahoma have adopted mud/slab jacking to quickly address the differential settlement problems of approach slabs (Yenigalla, 2012). Mud/slab jacking is a concrete leveling method that can lift the sunken slab to a desired height by pressure-injecting mud cement beneath the approach slabs. Without any adjustment of the approach slabs, the pressure grouting intends to fill the voids under slabs and normally acts as a preventive and short-term settlement countermeasure because it cannot prevent further settlement or erosion of backfill material (Abu-Hejleh et al., 2006; White et al., 2007).

Further, urethane injection and compaction grouting can significantly strengthen shallow and deep-seated soft layers with injected materials and thereby provide a long-term or permanent solution (Abu al-Eis and LaBarca, 2007; Miller and Roykroft, 2004). Particularly, the Urethane injection technique leverages the high-density polyurethane resin as the injected material which can harden up very quickly and gain 90% of maximum compressive strength within only 15 minutes. The Louisiana DOT identified the significance of using the Urethane injection technique to reduce approach bumps based on a four-year observation on two bridges, while the application of this technique in Oklahoma was not very successful (Abu al-Eis and LaBarca, 2007). Replacing construction embankments with lightweight fill materials such as Expanded Polystyrene (EPS) Geofoam is an alternative way that reduces the stress on the underlying subgrade to minimize differential settlement. Among the repairing methods, the asphalt overlay is comparable in cost to mud/slab jacking, while the Urethane injection method is much more expensive and followed by the costliest full replacement (Yenigalla, 2012).

Compared with those manually measured indices, such as differential settlement, approach-relative gradient, and approach index (Lu et al., 2018), the International Roughness Index (IRI) is a preferred engineering index used to evaluate bumps on and around approach slabs at highway speed using high-speed profilers (Abu-Farsakh and Chen, 2014; Olmedo et al., 2015). McGhee (2002) defined an IRI of 140 in./mi calculated for every 10 ft as the threshold to identify the approach slabs with severe bumps needing repairs. Olmedo et al. (2015) deemed that repairs need to be done on approach slabs with an IRI over 1,000 in./mi evaluated at the 1 ft interval. Because IRI is a measure of undulation accumulations of a given segment length, the calculation interval significantly affects the magnitude of IRI and exhibits different resolutions. Recently, Mishra et al. (2021) and Wang et al. (2023) demonstrated that the 1 ft calculation interval could best evaluate the bumps of the approach slabs considering their short length characteristics.

The evaluation of bumps on bridge approach/departure slabs via the classification criterion in Wang et al. (2023) was developed based on the advanced sub-mm 3D laser imaging technology for multi-source data collection at highway speeds, including 2D/3D pavement/bridge images,

roughness data, and right-of-way (ROW) images. The sub-mm 3D laser imaging technology offers a comprehensive approach to assessing the road and bridge surface condition in terms of roughness, rutting, faulting, texture, and various distress such as cracking, patching, pothole, and expansion joints, lane marking, and sealed cracking. The previous successful practice of 3D laser imaging technology for bridge approach bump evaluation motivated ODOT to use this non-destructive technology for field performance evaluation of approach slab designs and repair techniques applied on the mainline bridges in Oklahoma, with anticipated benefits of high efficiency, low labor cost, and no traffic control.

RESEARCH OBJECTIVE

This study intends to assess the field performance of existing approach slab designs and mitigation methods based on the severity levels of approach bumps using the latest generation of 3D laser technology. Because the detailed maintenance history on approach/departure slabs is unavailable to the research team, the field performance evaluation of the specific slab repair techniques could not be performed in this study. Instead, the performance of the design parameters of approach slabs in terms of slab length and skew angle, and the asphalt overlay as an observable mitigation method are studied. Accordingly, the proper geometric design of approach/departure slabs and whether the AC overlay should be laid on the approach/departure slabs have been recommended based on the analysis results.

FIELD DATA COLLECTION

In this study, the field data collections were conducted on 111 bridges along Interstate 35 (I-35) from Oklahoma City to the border with Texas, 264 bridges along Interstate 40 (I-40), one bridge on State Highway 4 (SH-4), and one bridge on State Highway 11 (SH-11) in Oklahoma, resulting in a total of 377 bridges surveyed. The bridge decks have lengths ranging from 85 ft to 3740 ft, with an average of 284 ft. There are 183 out of 377 bridge approach slabs (184 departure slabs) connected with Asphalt Concrete (AC) pavement, while the rest connect with Portland Cement Concrete (PCC) pavement. The approach slabs vary in length from 8 ft to 103 ft, with an average length of 30 ft. In contrast, the range of the departure slab lengths is between 8 ft and 119 ft, resulting in an average of 30 ft. Notably, 81 approach slabs and 79 departure slabs were laid with a layer of asphalt overlay while others were concrete slabs.

Figures 1 (a) - 1 (c) summarize the distributions of the deck lengths, approach/departure slab, and skew angles (defined in Figure 1 (d)) for all the surveyed bridges. From these figures, around 80% of bridges hold a deck length within 300 ft and more than 85% of bridges are associated with a length of less than 40 ft approach/departure slabs. Further, more than half of the bridges were constructed with a skew angle of zero degrees. To better evaluate the performance of bridge approach/departure slabs and joints, a 20-ft lead-in pavement, and a 20-ft lead-out pavement were additionally included in the data collection.

Figure 2 (a) shows the Pave3D 8K system in a data collection truck to collect sub-mm 2D/3D images and inertial profiles of the bridge approaches for bump and distress evaluation at highway speeds up to 60 mph. The Pave3D 8K system can capture high-quality 3D surface data with a resolution of 0.5 mm in the transverse direction with over 8,000 pixels for a lane width. In this study, the evaluation of the distressed bridge approach/departure slabs and joints was performed based on the collected 2D and 3D images. Notably, the WayLink Inertial System (WIS) along

with an Inertial Measurement Unit (IMU) is embedded into the Pave3D 8K system to continuously capture the longitudinal inertial profiles that are synchronized with the 2D/3D images for bump assessment of approach/departure slabs.

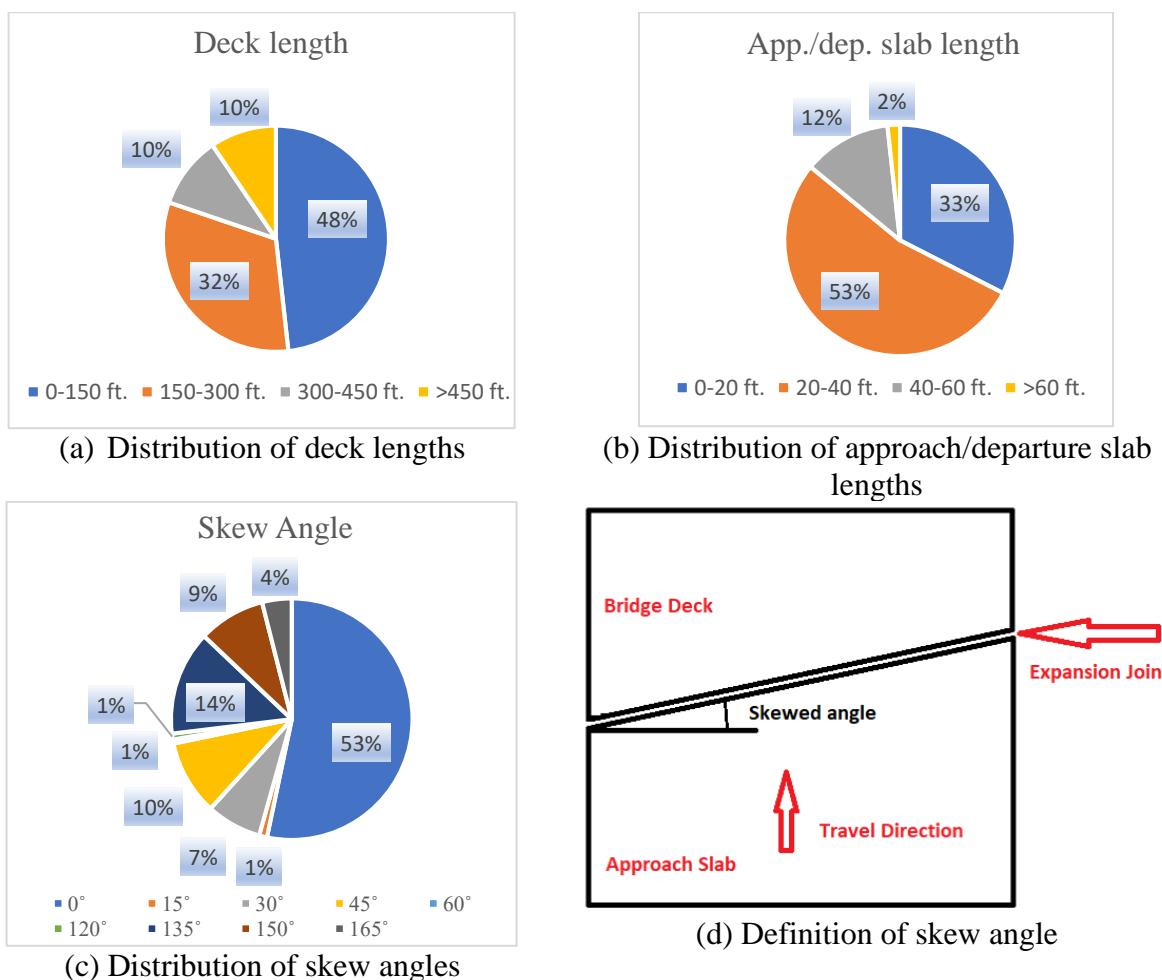


Figure 1. Distributions of Basic Geometry of Bridge Decks and Bridge Approaches

As shown in Figure 2 (b), the WIS integrates three major components: a Distance Measurement Instrument (DMI), an accelerometer, and a 3D height sensor. The DMI and the height sensor respectively play the roles of recording the travel distance of the vehicle traveling and height distance relative to the road surface, while the high-precision accelerometer collects the inertial profiles of ride surfaces. The collected inertial profiles were combined with the synchronized 3D height images to generate the longitudinal profiles along the left and right wheel paths, on which the International Roughness Index (IRI) values were calculated to quantify the bump levels of approach/departure slabs.

Figure 3 illustrates the 3D images of one typical transition scenario from the approach slab to the bridge deck at a skew angle of 45°. The 3D images demonstrated in the figure are (a) the rendered virtual 3D image from MHIS software; and (b) the top view 3D image at the same location from ADA3D software. The two types of 3D images essentially display the 3D surface at the same location but in different view angles. It is shown that the sub-mm 3D images can

clearly demonstrate the surface details of approach pavements, approach slabs, and bridge decks. Examples of more 3D images of bridge approaches can be found in our previous study (Wang et al., 2023).

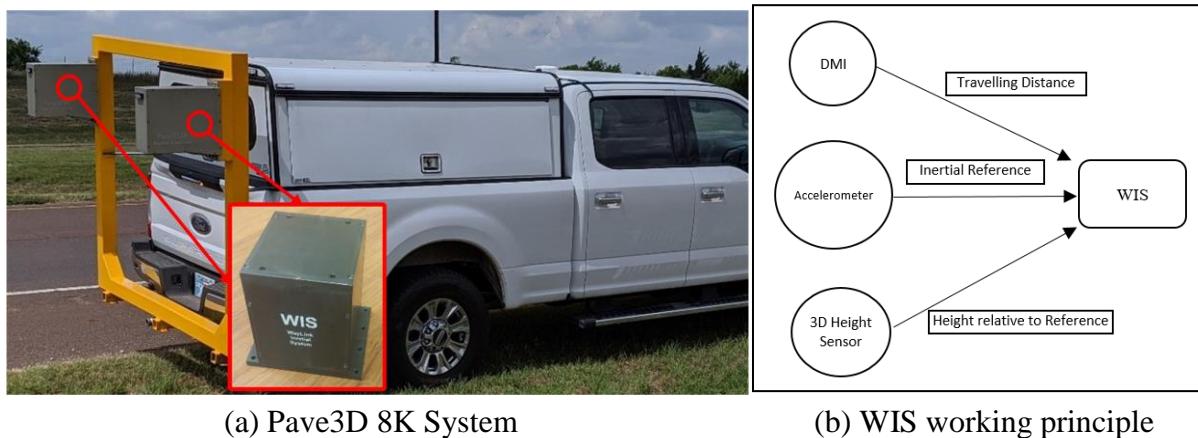


Figure 2. Data Collection Device

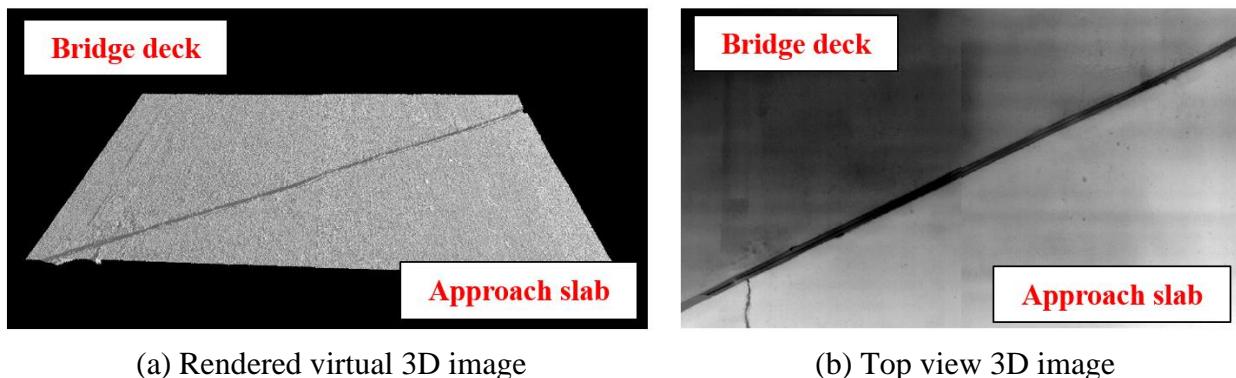


Figure 3. Approach Slab Connected with Bridge Deck at a Skew Angle of 45°

IRI CALCULATION

The IRI was computed from the left and right wheel paths for those field collections according to the standard (ASTM E1926-08, 2021) using an interval of 1 ft (Wang et al., 2023). Because approach slabs and departure slabs are too short in length to cut from the long data collection data for IRI calculation, five consecutive sections cut from the large data collections, including 20-ft approach pavement, approach slab, bridge deck, departure slab, and 20-ft departure pavement, were analyzed together during the IRI calculation phase. Considering the travelers' ride discomfort is mostly attributed to the location with the highest roughness for a given ride surface section, the maximum IRI was calculated within the local locations of the five sections for bump evaluation.

To better evaluate the bump levels of the approach/departure slabs, this project defines four major key locations for approach/departure slabs when calculating their maximum IRI values. As shown in Figure 4, on the approach slab, location B points to the connection joint between the

approach pavement and approach slab, and location C points to the joint between the approach slab and bridge deck. Further, location A is a point defined on the approach pavement that is 5 ft away from location B, while location D is a point defined on the bridge deck that is 5 ft away from location C. Finally, the maximum IRI for the approach slab is calculated on the left and right wheel paths of the Section-1 ranging from location A to location D instead of location B to location C. Similarly, the maximum IRI for the departure slab is calculated on Section-2.

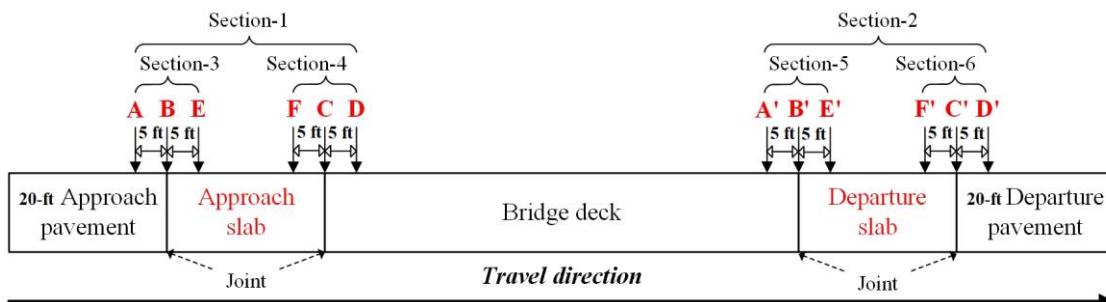


Figure 4. Locations of Calculating Maximum IRI for Approach/Departure Slabs

In addition, the maximum IRI was also calculated for four other sub-sections around the approach/departure joints: (1) the joint between the approach pavement and approach slab; (2) the joint between the approach slab and bridge deck; (3) the joint between the bridge deck and departure slab; and (4) the joint between departure slab and departure pavement. The four sub-sections are all 10 ft long and are symmetric to the approach/departure joints, which are Section-3 (location A to location E), Section-4 (location F to location D), Section-5 (location A' to location E'), and Section-6 (location F' to location D'), as shown in Figure 4. It should be mentioned that the maximum IRI within the sub-sections was calculated for the left and right wheel paths separately. The averages of maximum IRI of the left and right wheel paths of the four sub-sections are used to investigate the performance of the approach slab lengths, skew angles, and the asphalt overlay in affecting the bump levels of bridge approaches.

PERFORMANCE OF SLAB LENGTHS

To investigate how the lengths of approach/departure slabs affect bump levels, the approach/departure slab lengths of the 377 bridges are divided into four sub-ranges: (0, 20], (20, 40], (40, 60], and (60, 80]. In summary, there are 246, 390, 98, and 16 bridges whose approach slab lengths fall into the four sub-ranges, while 240, 408, 86, and 10 bridges have departure slab lengths falling into the four sub-ranges. Subsequently, the IRI values of six sub-sections of the bridges (Figure 4) were averaged accordingly for the four sub-ranges of approach/departure slab lengths. Figure 5 shows the boxplots of the IRI by different slab length sub-ranges for six sub-sections of bridges (Figure 4). Particularly, the number in or around each boxplot in Figure 5 indicates the average of the IRI for each sub-section.

It is seen from Figures 5 (a) and 5 (d) that both approach slabs and departure slabs exhibit a consistent decrease in the average IRI when the length of slabs increases. The IRI around the joint of the approach pavement and approach slab (Figure 5 (b)), and the joint of the approach slab and bridge deck (Figure 5 (c)) shows an overall decreasing trend as the approach slab length increases even though there is a slight increase in the IRI for the slab length between 40 ft and 60

ft. Particularly, the approach lab lengths of 60 ft to 80 ft can significantly decrease the bumps around the joint of the approach pavement and approach slab compared to other shorter approach slabs, as shown in Figure 5 (b).

From Figure 5 (e), the IRI around the joint of the bridge deck and departure slab experiences an initial moderate decrease for departure slab lengths changing from (0, 20] to (20, 40], while a very small increase as the departure slab length further increases. In contrast, The IRI around the joint of the departure slab and departure pavement (Figure 5 (f)) undergoes an initial increase for departure lengths ranging from (0, 20] to (20, 40], while a continuous decline as departure slab length further increases. Although the effect of the increase of departure slab lengths on bumps around the joint of the bridge deck and departure slab is not clear, the departure slab lengths of 60 ft to 80 ft can significantly decrease the bumps at the joint of the departure slab and departure pavement compared to shorter departure slabs.

Broadly speaking, the longer design of either the approach slab length or the departure slab length produces smaller bumps on and around the approach/departure slabs. This finding confirms the laboratory study by (Seo 2003) which also pointed out that a longer transition slab could reduce the pressure imposed on underlying soils. Therefore, it is recommended that the approach/departure slab with lengths of 60 ft to 80 ft should be used in place of currently frequently used transition slabs less than 40 ft to reduce the approach bumps when designing new transition slabs or replacing old slabs with new slabs for rehabilitation purposes.

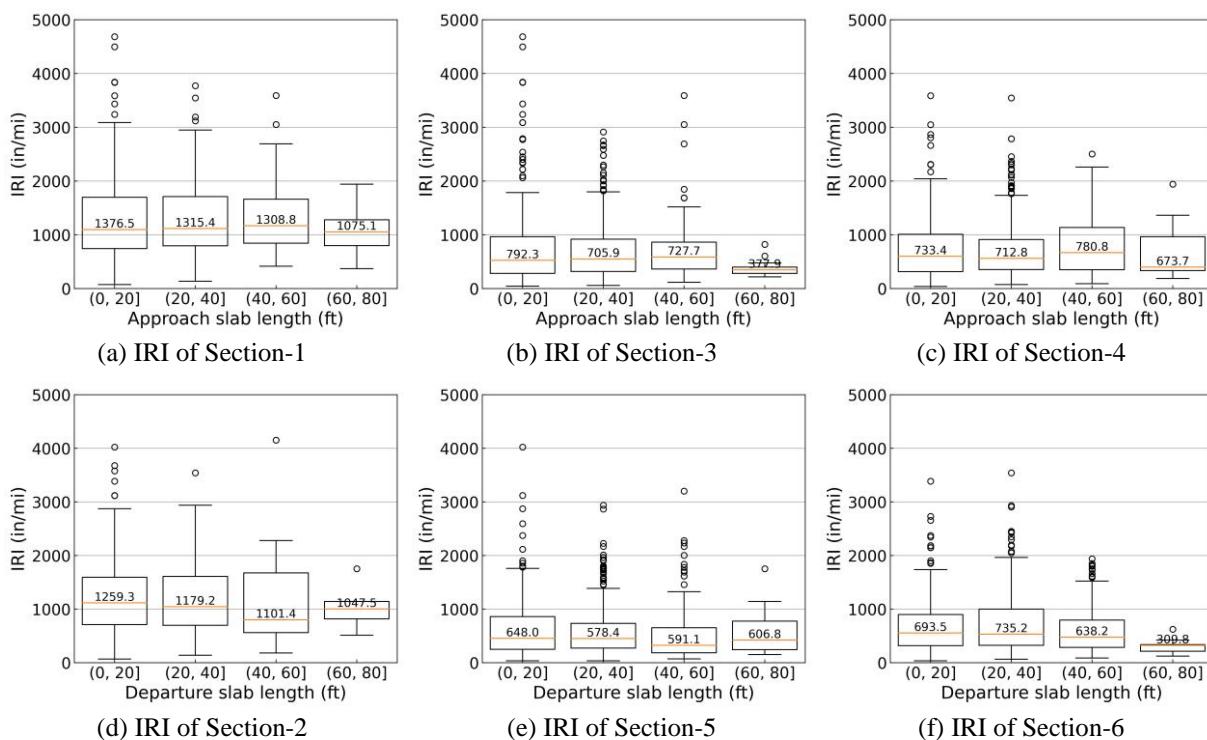


Figure 5. Average IRI of Six Bridge Sub-sections over Different Slab Lengths

PERFORMANCE OF SKEW ANGLES

There is a total of 9 discrete skew angles for the surveyed bridge approaches, which are 0° , 15° , 30° , 45° , 60° , 120° , 135° , 150° , and 165° in ascending order, corresponding to 402, 8, 56, 76,

4, 6, 108, 68, and 30 bridges, respectively. Accordingly, the IRI values of each of the six sub-bridge sections were summarized for the 9 skew angles for both approach slab and departure slab lengths. Figure 6 shows the boxplots of the IRI with the average IRI number in or around each boxplot by different skew angles for six sub-sections of bridges (Figure 4). It is seen from Figure 6 (a) that there is no significant difference in IRI among different skew angles for approach slabs except at 60° where the lowest IRI is produced. The significant decrease in IRI of approach slabs with 60° skew angles is attributed to the lowest IRI (369.8 in/mi) around the joint of approach pavement and approach slab (Figure 6 (c)) and the lowest IRI (527.7 in/mi) around the joint of approach slab and bridge deck (Figure 6 (d)).

Similarly, the IRI on departure slabs shows no obvious relationship with increased skew angles, as shown in Figure 6 (b). The significant drop in IRI on departure slabs occurs at the skew angle of 120°. The lowest IRI (614.7 in/mi) at this skew angle for departure slabs derives mainly from the smallest IRI (367.9 in/mi) around the joint of the bridge deck and departure slab (Figure 6 (e)). Therefore, it is tentatively suggested that the skew angle of 60° for approach slabs and 120° for departure slabs should be constructed for new slab design or old slab repairment to minimize the bumps at the ends of bridges. However, more efforts should be made in future work to verify if the suggested skew angles for approach/departure slabs are correct because the sample sizes of approach slabs at the angle of 60° and departure slabs at the angle of 120° are only 4 and 6, respectively, which may be too small to present unbiased results compared to the skew angle 0° with 402 samples or the skew angle 135° with 108 samples.

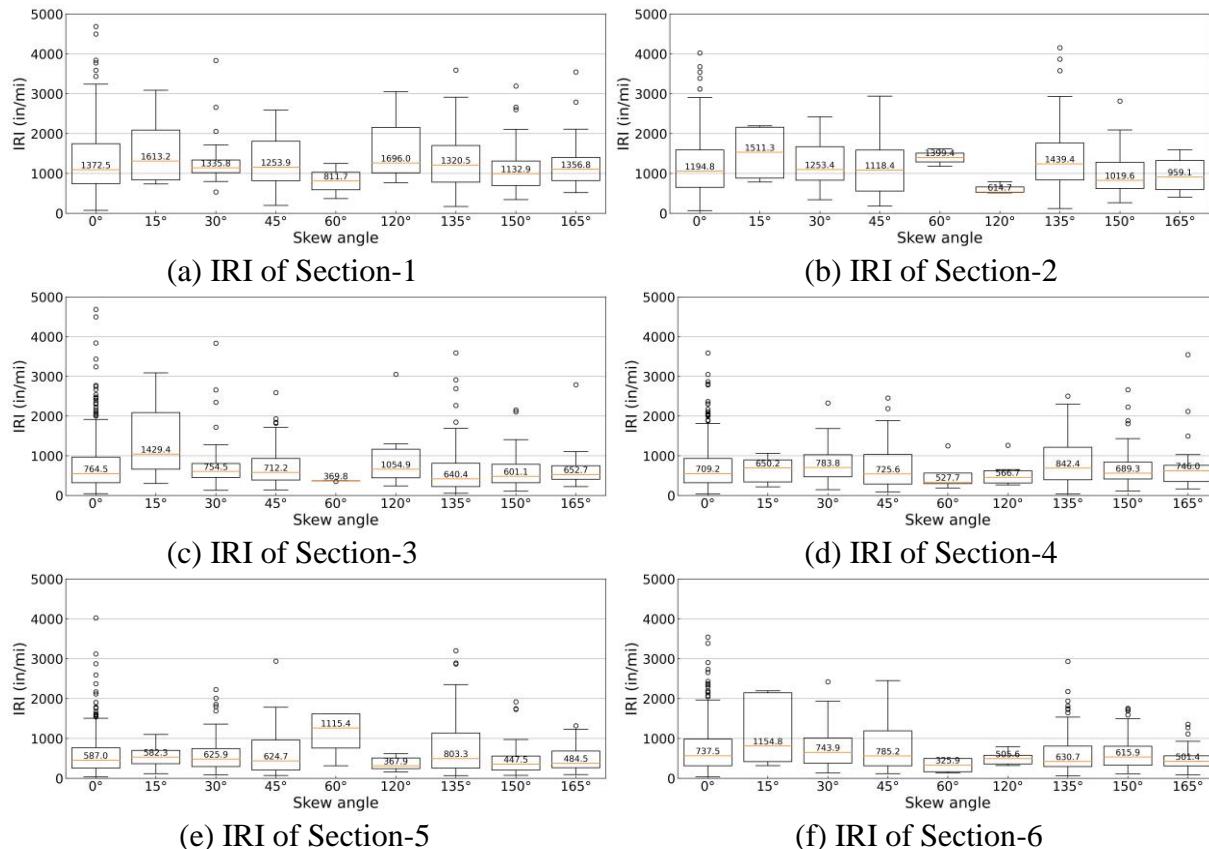
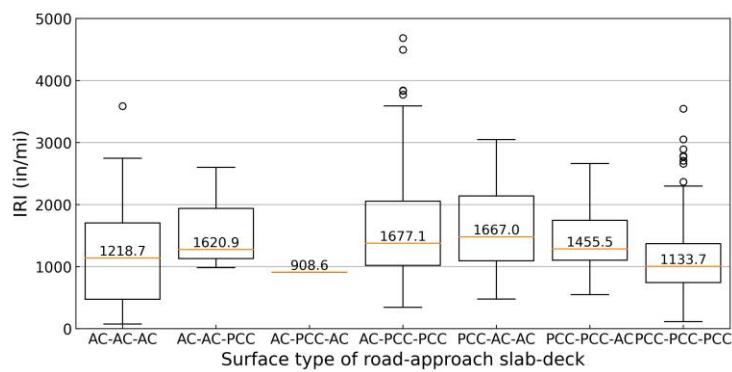
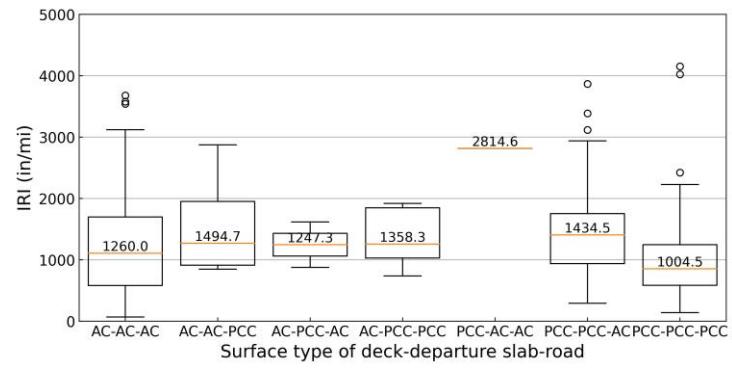


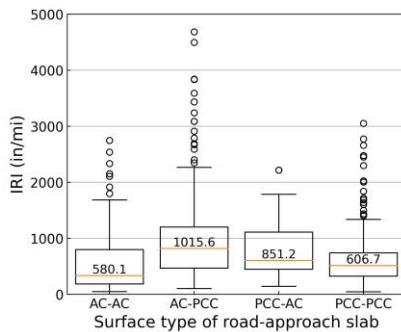
Figure 6. Average IRI of Six Bridge Sub-sections for Different Skew Angles



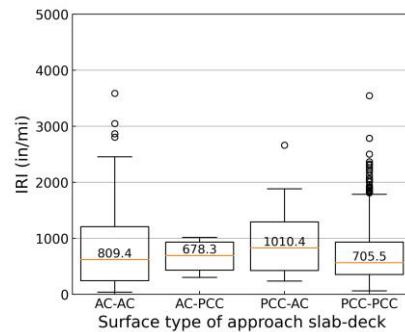
(a) IRI of Section-1 with different road-approach slab-deck surface types



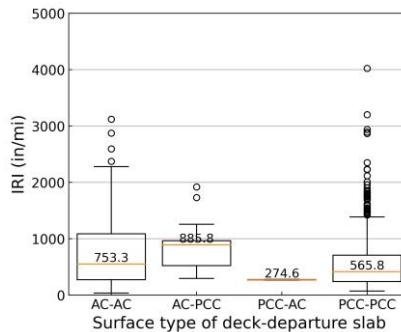
(b) IRI of Section-2 with different road-departure slab-deck surface types



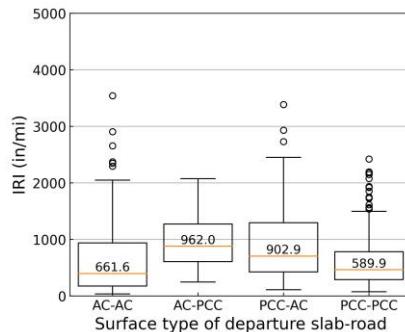
(c) IRI of Section-3 for different road-approach slab surface types



(d) IRI of Section-4 for different approach slab-deck surface types



(e) IRI of Section-5 for different deck-departure slab surface types



(f) IRI of Section-6 for different departure slab-road surface types

Figure 7. Average IRI of Six Bridge Sub-sections for Different Surface Transition Types

PERFORMANCE OF AC OVERLAY

The performance of AC overlay was studied from different aspects concerning different surface transition types of bridge sections. First, the surface transition types of the road-approach slab-bridge deck and the bridge deck-departure slab-road are summarized. Notably, there are seven types of surface transitions: AC-AC-AC, AC-AC-PCC, AC-PCC-AC, AC-PCC-PCC, PCC-AC-AC, PCC-PCC-AC, and PCC-PCC-PCC for bridge approaches. For example, the AC-AC-AC represents AC pavement connected with the approach slab with AC overlay and bridge deck with AC overlay in the case of road-approach slab-bridge deck transitions.

The seven types of surface transitions correspond respectively to 68, 3, 1, 111, 10, 6, and 180 approach slabs, and 68, 10, 2, 5, 1, 113, and 180 departure slabs. Accordingly, the IRI values of the seven transition types were summarized for both approach slabs and departure slabs. In addition, the IRI values of four other sub-sections around the joints of approach/departure slabs (Figure 4) were summarized according to the surface transition types of approach roadway to approach slab, approach slab to bridge deck, bridge deck to departure slab, and departure slab to departure roadway.

Figure 7 shows the boxplots of the IRI with the average IRI number in or around each boxplot by different skew angles for six sub-sections of bridges (Figure 4). It is seen from Figure 7 (a) that the lowest IRI (908.6 in/mi) is found at the AC-PCC-AC surface transitions. However, there is only one sample of such surface transitions from which the evaluated IRI may be biased. Particularly, AC-AC-AC (68 samples), AC-PCC-PCC (111 samples), and PCC-PCC-PCC (180 samples) are the three most frequently used surface transition types in the surveyed bridges for approach slabs. Among the three surface types, AC-PCC-PCC held the highest IRI (1677.1 in/mi), while PCC-PCC-PCC held the lowest IRI (1133.7 in/mi). The highest IRI of AC-PCC-PCC is attributed to the roughest transition from AC pavement to PCC approach slab where the average IRI is 1015.6 in/mi, as shown by AC-PCC in Figure 7 (c). It also means that the bump at the transition from AC pavement to PCC approach slab is more prone to occur than at the transition from approach slab to bridge deck. This is because most parts of the approach slab rest on the embankment that is more prone to settlement than the bridge abutment that supports the rest of the approach slab.

Compared to the transition from AC pavement to PCC approach slab, the transition from AC pavement to approach slab applied with AC overlay exhibits much lower IRI, as shown by AC-AC in Figure 7 (c). On the other hand, the approach slab laid with AC overlay also has minor improvement on the transition from approach slab to bridge deck compared to the PCC approach slab (Figure 7 (d)). For example, the approach slab laid with AC overlay connected with PCC bridge deck has a transition IRI of 678.3 in/mi, being slightly lower than 705.5 in/mi for the PCC approach slab connected with PCC bridge deck. However, the AC overlay applied on the bridge deck seems to have a negative effect on smoothing the transition from approach slab to bridge deck by comparing the PCC-AC (1010.4 in/mi) with PCC-PCC (705.5 in/mi) in Figure 7 (d).

There are three most frequently used surface transitions for departure slabs: AC-AC-AC (68 samples), PCC-PCC-AC (113 samples), and PCC-PCC-PCC (180 samples). It is seen from Figure 7 (b) that the lowest IRI (1004.5 in/mi) is found at the PCC-PCC-PCC surface transitions followed by the higher IRI at AC-AC-AC (1260.0 in/mi) and the highest IRI of PCC-PCC-AC (1434.5 in/mi). Among the three surface transitions, the highest IRI of PCC-PCC-AC derives from the rough transition from the PCC approach slab to AC departure pavement (902.9 in/mi), as shown by PCC-AC in Figure 7 (f). It also indicates that the transition from PCC departure slab

to AC departure pavement is more prone to bump than the transition from the bridge deck to the departure slab (Figure 7 (e)). Similarly, the AC overlay placed on the departure slabs can reduce bumps on the transition from the bridge deck to the departure slab (Figure 7 (e)) and the transition from departure slab to departure pavement only when the departure pavement is AC surface (Figure 7 (f)). It is shown in Figure 7 (e) that the bridge deck with AC overlay would increase the bump at the transition from the bridge deck to the departure slab with or without AC overlay. For example, in Figure 7 (e), the transition from the bridge deck with AC overlay to PCC departure slab has an average IRI of 885.8 in/mi, which is rougher than the transition from PCC bridge deck to PCC departure slab with an average IRI of 565.8 in/mi.

From the above observations, it is suggested that the placement of AC overlay on approach slabs should be performed as a routine mitigation method when the approach pavement is flexible pavement to minimize bumps at the ends of approach/departure slabs. If the approach/departure pavement is a PCC surface, it is seemingly needless to apply the AC overlay on the approach or departure slabs.

CONCLUSIONS AND FUTURE WORK

This project performed an efficient, accurate, and comprehensive performance evaluation of bridge approach/departure slabs and joints at a network level through a sub-mm 3D laser imaging system. The 3D/2D images, ROW images, and roughness profiles were collected simultaneously via the Pave3D 8K system from 377 bridges in Oklahoma. Based on the collected multi-source bridge data, the bumps present on approach/departure slabs were first evaluated via the maximum IRI calculated at the 1 ft interval from the left and right wheel paths. The average bump levels of different locations of approach/departure slabs were surveyed based on different slab design parameters to explore the performance of the approach slab lengths, skew angles, and asphalt overlay in affecting the bump levels of bridge approaches and presented the following findings.

- Longer approach/departure slabs help slow down the development of bumps at the ends of bridges. It is suggested that the approach/departure slab with lengths of 60 ft to 80 ft should be more frequently used to reduce the approach bumps when designing new transition slabs or replacing old slabs with new slabs in rehabilitation.
- It is identified that a skew angle of 60° for approach slabs and 120° for departure slabs are recommended in design to reduce bumps at the ends of bridges. However, more efforts should be devoted to validating this finding in future work.
- It is found that severe bumps are more likely to occur at the transition from AC pavement to PCC approach slab (or PCC departure slab to AC departure pavement) than at the transition from approach slab to bridge deck (or bridge deck to departure slab).
- AC overlay is an effective mitigation method to be applied on the approach/departure slabs to minimize the bumps at the ends of bridges when the approach/departure pavement is AC-based.
- The transition from PCC approach road to PCC approach slab and then to PCC bridge deck provides the smoothest ride for travelers compared to other surface transitions.

Unfortunately, only the AC overlay as a bump mitigation method has been evaluated in terms of its performance by comparing approach slabs with AC overlay with those without AC overlay. The field performance evaluation of other existing repair techniques could not be performed due to the unavailable detailed maintenance history on the approach/departure slabs.

To perform this type of evaluation would need the specific mitigation methods applied to the bridge approaches in the past, the time of applying these methods, and the related maintenance cost, etc.

REFERENCES

- Abu Al-Eis, K., and LaBarca, I. K. (2007). Evaluation of the URETEK Method of Pavement Lifting. Rep. No. WI-02-07. Wisconsin Department of Transprotation, Division of Transportation Systems Development, Madison, Wisconsin.
- Abu-Farsakh, M. Y., and Chen, Q. (2014). Field demonstration of new bridge approach slab designs and performance. Rep. No. FHWA/LA. 13/520. Louisiana Transportation Research Center.
- Abu-Hejleh, N., Hanneman, D., White, D. J., and Ksouri, I. (2006). Flowfill and MSE Bridge Approaches: Performance, Cost and Recommendations for Improvements. Report No. CDOT-DTD-R-2006-2. Colorado Department of Transportation, Denver.
- Arsoy, S., Barker, R. M., and Duncan, J. M. (1999). The Behavior of Integral Abutment Bridges. Rep. No. VTRC 00-CR3, Virginia Transportation Research Council, Charlottesville, VA.
- ASTM. ASTM E1926-08. (2021). *Standard practice for computing international roughness index of roads from longitudinal profile measurements*. West Conshohocken, PA.
- Briaud, J. L., James, R. W., and Hoffman, S. B. (1997). Settlement of Bridge Approaches (The Bump at the End of the Bridge). NCHRP Synthesis of Highway Practice 234. Transportation Research Board, Washington, DC.
- FDOT (Florida DOT). (2012). FDOT Design Standards FY 2012/2013. Index No. 20900. Florida Department of Transportation, Gainesville, Florida.
- Lu, Q., Li, M., Gunaratne, M., Xin, C., Hoque, M., and Rajalingola, M. (2018). *Best practices for construction and repair of bridge approaches and departures*. University of South Florida, Tampa, FL.
- McGhee, K. (2002). A new approach to measuring the ride quality of highway bridges. VTRC 02-R10. Charlottesville, VA.
- Miller, E. A., and Roykroft, G. A. (2004). "Compaction Grouting Test Program for Liquefaction Control." *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 355-361.
- Mishra, D., et al. (2021). Statistical analysis of pavement profiler data to evaluate the bump at the end of the bridge. FHWA-HRT-21-037. Research, Development, and Technology Turner-Fairbank Highway Research Center, Federal Highway Administration, Mclean, VA.
- Olmedo, C., Henderson, B., Cimini, G., and Springer, J. (2015). "Bump determination at bridge approach transitions using inertial profilers." In *Proc., Conf. of the Transportation Association of Canada*.
- Puppala, A. J., Saride, S., Archeewa, E., Nazarian, S., and Williammee, R., Jr. (2010). "Bridge approach settlements: lessons learned from present case studies and ground improvement solutions." In *Ground improvement and geosynthetics* (pp. 228-238).
- Seo, J. (2003). *The Bump at the End of the Bridge: An Investigation*. Dissertation, Texas, A&M University, College Station, Texas.
- Wang, G., Wang, K. C., Yang, G., Li, J. Q., and Peters, W. (2023). "Evaluation of Bridge Approach Slab and Dynamic Load Allowance Using Sub-mm 3D Laser Imaging Technology." *Journal of Bridge Engineering*, 28(7), 04023037.

- White, D., Mohamed, M., Sritharan, S., and Suleiman, M. (2007). "Underlying" Causes for Settlement of Bridge Approach Pavement Systems." *Journal of Performance of Constructed Facilities*, ASCE, 273-282.
- Yenigalla, R. V. (2012). *An overview of mitigation strategies for settlements under bridge approach slabs*. Dissertation, The University of Texas at Arlington, Texas.

A Review of Factors Affecting Severity of Autonomous Vehicle Crashes

Deema Almaskati¹; Sharareh Kermanshachi, Ph.D., P.E.²; and Apurva Pamidimukkala, Ph.D.³

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: deemanabeel.almaskati@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Penn State Univ.
Email: svk5464@psu.edu

³Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

ABSTRACT

Most of the researchers who have explored the factors that affect the severity of crashes involving autonomous vehicles (AVs) by analyzing the data made available by the California Department of Motor Vehicles (CA DMV) have done so in general terms. A comprehensive investigation is lacking in the literature, and the purpose of this study is to remedy that by examining accidents and resulting injuries, if any, that involved at least one AV and identifying and categorizing the factors that affected the severity of the crashes. Strategies for addressing challenges such as public acceptance and unavoidable accidents are also suggested. The findings support the frequently touted belief that AVs improve traffic safety and show that they are not responsible for most of the collisions in which they are involved. This study will benefit AV manufacturers and policymakers by providing them with a better understanding of AV crashes and recommending strategies that can help manage the challenges of AVs.

Keywords: autonomous vehicles; accidents; transportation safety; unavoidable collisions; ethics; policy

INTRODUCTION

The National Highway Traffic Safety Administration estimates that more than 6 million car crashes occur in the United States annually, resulting in 2 million injuries and 37,000 fatalities (Beck et al. 2023). Ninety-four percent (94%) of the fatalities result from human error (Combs et al. 2019), and this is an area where autonomous vehicles (AVs) have an opportunity to shine, as they lessen the possibility of human error and have the potential to reduce the number of accidents, injuries, and deaths by utilizing effective detection technology. Autonomous driving systems can improve highway safety by detecting their surroundings and responding safely, lowering pollution emissions, and increasing self-sufficiency for people with limited mobility (Robinson et al. 2022; Patel et al. 2023a). Public acceptance is required to fully realize these benefits (Evans et al. 2020), however, and this is often seen as the greatest barrier to their widespread adoption.

Six levels of automation are defined by the Society of Automotive Engineers (SAE): Level 0 is fully manual, levels 1 through 3 are partially autonomous, and levels 4 through 5 are fully autonomous (Etminani-Ghasroodashti et al. 2023a). The number of hazards decrease as the level of automation rises, as less human intervention and support are needed to operate safely (Ye et al. 2021), but safety remains a concern regardless of the level of automation. Extensive testing is being conducted in many states to determine whether AVs actually lower the number of traffic

accidents. (Sinha et al. 2021; Khan et al. 2022). The California Department of Motor Vehicles (CA DMV) allows manufacturers to test AVs on public roads, but they are required to gather data and publish reports on disengagements and collisions (Dixit et al. 2016; Sinha et al. 2021). Currently, this is the only dataset that is widely accessible to researchers who wish to evaluate data to analyze the effectiveness of AVs.

The literature contains results of studies that explored the various facets of AVs and the implications of utilizing them. Parekh et al. (2022) reviewed AV technologies and investigated their ability to detect their environment, including pedestrians, and plan the best route, but they neglected to address the ramifications of policies. Faisal et al. (2019) failed to take into account inevitable collisions or their ethical consequences and instead concentrated almost exclusively on policy issues and how to best prepare for the deployment of AVs. Robinson et al. (2022) examined the ethical conundrum that inevitable collisions present and discussed the implications of previous research but neglected to take into account the safety of AVs in real-world situations. The performance of AVs has been assessed by numerous researchers using the CA DMV data (Xu et al. 2019; Sinha et al. 2021). Yuan et al. (2022) and Ren et al. (2022) are among those that examined the data to identify the variables that influence the severity of crashes; however, they did not address the policy implications of inevitable collisions or the public's perception of AV safety; nor did they integrate the data with that derived from a literature review.

After carefully examining relevant articles in the literature, the following objectives were formulated to determine AVs' impact on traffic accidents and injuries: (1) recognize and categorize the variables that affect the severity of AV crashes, (2) discuss the challenges presented by autonomous driving technologies, and (3) specify management strategies intended to address public AV safety concerns and the moral ramifications of inevitable collisions. The results of this study will be beneficial to legislators and AV manufacturers, as they will act as a catalyst for additional research aimed at improving AV adoption and safety globally.

METHODOLOGY

A database of pertinent journal articles, technical reports, conference papers, and theses was created by entering various combinations of keywords, including autonomous vehicles, vulnerable road users, self-driving cars, traffic accidents, road safety, inevitable collisions, crash analysis, and CA DMV, into online search engines such as Google Scholar, Science Direct, and the ASCE Library. Articles from a broad range of fields, such as psychology, ethics, and transportation were studied and considered, and after applying the inclusion and exclusion criteria shown in Figure 1, 107 articles were selected for further review.

After the initial screening was performed to remove duplicates and inappropriate document types, the publications' titles and abstracts were examined to confirm their applicability. The 106 studies that met all of the requirements were thoroughly reviewed, then categorized into three groups: (1) those that analyzed CA DMV collision and disengagement reports, (2) those that addressed AV safety and public perception, and (3) those that focused on moral and ethical issues related to inevitable AV collisions. The database was designed to demonstrate how the state of AVs and traffic safety are represented in the literature.

RESULTS & DISCUSSIONS

AVs have the potential to completely transform transportation safety by removing the human error component from traffic accidents, thereby decreasing the number of accidents and related

injuries and fatalities. The following subthemes were addressed in the literature: (1) the severity of AV crashes, (2) the interactions between AVs and vulnerable road users, and (3) the ethical implications of autonomous driving technologies. This section examines the current status of AVs and traffic safety by addressing each of these subthemes and offering strategies for managing the inherent challenges of the technology.

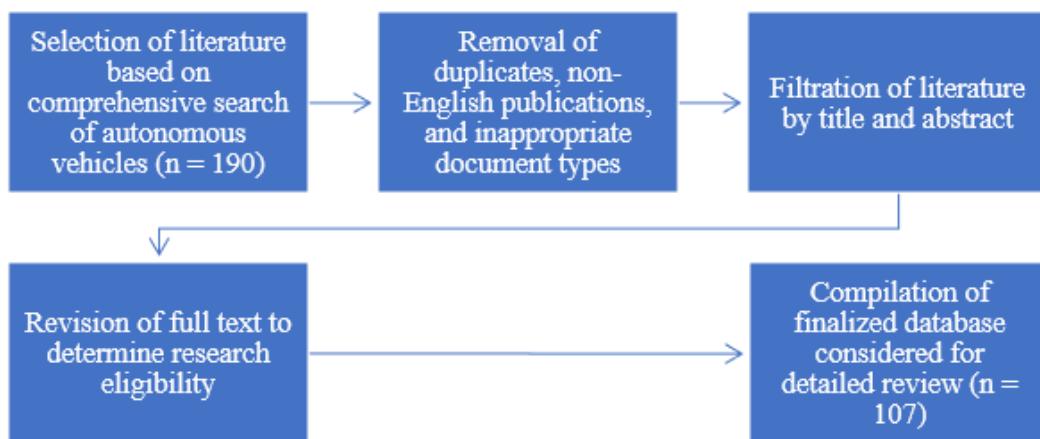


Figure 1. Screening process for identification and evaluation of relevant literature

AV Performance in a Mixed Traffic Environment

The utilization of AV technology has the potential to improve highway safety by reducing the impacts of human error, which account for the majority of conventional vehicle accidents (Pamidimukkala et al. 2023). A considerable amount of field and simulation research has been required to assess the safety level of AVs due to the lack of crash data and the small number of them on the roads (Song et al. 2021). California is one of many states that has passed legislation allowing AV testing on public roadways to help people understand how AVs perform in mixed traffic situations (Schoettle and Sivak 2015), but they are the only one that has made crash and disengagement data available to the public (Das et al. 2020). It is, therefore, the sole source of data for researchers, despite worldwide AV testing being conducted (Houseal et al. 2022).

Crash reports include the circumstances leading up to a collision, resulting injuries and/or property damage, geolocation, type of crash, weather, road conditions, and the time and date of the incident (Sinha et al. 2021). Through their analysis of the crash data, several authors found that rear-end and sideswipe collisions are the most common types of crashes (Favarò et al. 2017; Wang and Li 2019; Das et al. 2020; Petrović et al. 2020; Channamallu et al. 2023; Patel et al. 2023b). A study by Liu et al. (2021) distinguished between the pre-crash conditions of AVs and conventional cars and found that the two groups had different numbers of collisions under the same conditions. AVs are 1.6 times more likely than conventional vehicles to be involved in a rear-end collision, but they are not at fault in most of the incidents. This may be due to the behavior of drivers of conventional vehicles who drove too fast, follow the vehicle in front of them too closely, and are not accustomed to the dynamic driving style of AVs, which includes adhering to all traffic regulations and not being subject to human errors (Petrović et al. 2020). AVs are also less likely than conventional vehicles to be involved in sideswipes, head-on collisions, and other types of collisions, despite their higher involvement in rear-end collisions.

(Novat et al. 2023). Only about 6% of AV accidents involve pedestrians or bystanders, compared to 42% of conventional vehicle accidents (Petrović et al. 2020), and the difference that can be ascribed to AVs' enhanced perception of their environment and their ability to identify both moving or stationary objects through the application of sophisticated technological systems (Petrović et al. 2020).

Even though AVs are not at fault in 87% of the collisions they are involved in (Wang and Li 2019), it is crucial to identify the elements that raise the risk of severe collisions when they do occur. When a conventional car is involved in an accident, the degree of severity can vary from minimal damage to serious harm and even death to the car, driver, and other people involved (Yuan et al. 2022). An examination of CA DMV collision reports revealed that the majority of AV collisions cause only minor damage to the cars involved and the majority of personal injuries are related to back pain (Dixit et al. 2016). Ye et al. (2021) found that back, head, and neck injuries are the most common types of injuries, with 70.83% of injured people being AV occupants. Many consider safety to be the main benefit of AVs; therefore, factors linked to higher crash severity were identified through the literature review and are shown in Table 1.

Although the CA DMV data sheds light on how AVs function in a mixed traffic environment, the data's limitations may cause the findings to be inaccurate. Specifically, the data is restricted to a geographic area that is not representative of all driving conditions, and the sample size of AVs taken into account is small when compared to the number of conventional vehicles on the roads. For example, since California has a mild climate, the results don't show how AVs function in severe winters or other extreme weather events. A thorough analysis is difficult, if not impossible, given the scarcity of published data on crash reports and the complexity, dynamicity, and unpredictability of the driving environment. Beyond that, it is unlikely that all driving conditions have been represented. Future studies should take these shortcomings into account by examining various settings to determine how other variables might affect the technology's functionality.

Autonomous Vehicle Challenges

Designers and manufacturers of AVs are constantly striving to enhance the vehicles' ability to monitor their surroundings, identify nearby road users, and anticipate other drivers' movements (Alozi and Hussein 2022). These are vital tasks that are extremely challenging because interactions among road users are unpredictable, even in situations where all the other variables remain constant (Parekh et al. 2022; Chen et al. 2023). AV technology is based on effective detection of road elements such as curbs, lanes, traffic signs and signals, other road users, obstacles, and infrastructure (Hafeez et al. 2022), and a multi-sensor approach that includes visible light cameras, radar, and light detection and ranging is considered the best course of action (Combs et al. 2019; Khan et al. 2023).

Pedestrians are frequently considered the most vulnerable road users due to their frequent overrepresentation in traffic-related injuries and fatalities, (Alozi and Hussein 2022), and it can be difficult for AVs to identify them in low visibility situations and/or with low camera resolutions (Iftikhar et al. 2022; Parekh et al. 2022). Their actions and behaviors are highly unpredictable and probabilistic, and they frequently appear in a variety of settings that are influenced by the surrounding environment and the weather (Combs et al. 2019). According to Kaur and Rampersad (2018), safe pedestrian-AV interactions are essential to AV adoption and public trust; however, AVs have functional problems that make it challenging for them to identify pedestrians (Iftikhar et al. 2022).

Table 1. Identification and categorization of factors affecting severity of crashes involving AVs

ID	Characteristic	Crash Severity Factor	Study
C.1	Driving mode	AV driving in autonomous mode Disengagement failure	Wang and Li 2019 Khattak et al. 2020
C.2	Location of vehicle	Parked on the side of the road At signalized intersection Intersection	Xu et al. 2019 Sinha et al. 2021 Xu et al. 2019; Zhu and Meng 2022
C.3	Responsible party	Mid-block segments	Yuan et al. 2022
C.4	Type of road	AV One-way Highways	Wang and Li 2019 Xu et al. 2019 Wang and Li 2019
C.5	Collision type	Sideswipe Rear-end	Das et al. 2020 Das et al. 2020; Guo and Zhang 2022; Houseal et al. 2022
C.6	Vehicle movements	Frontal Turning Opposing car going straight and turning right In motion	Yuan et al. 2022 Das et al. 2020 Zhu and Meng 2022
C.7	Number of vehicles	Collision with multiple vehicles	Ren et al. 2022
C.8	Time	Nighttime	Das et al. 2020; Zhang and Xu 2021 Zhang and Xu, 2021; Ren et al 2022; Yuan et al. 2022
C.9	Lighting conditions	Streetlight illumination/poor lighting	Das et al. 2020; Ye et al. 2021; Zhu and Meng 2022
C.10	Weather conditions	Poor weather Rain	Houseal et al. 2022 Ren et al. 2022
C.11	Roadway surface conditions	Adverse	Leilabadi and Schmidt 2019
C.12	Mixed use traffic	Public transportation routes Bike lanes Sidewalks on both sides of road	Ren et al. 2022 Ren et al. 2022 Ren et al. 2022
C.13	Traffic conditions	High-density	Zhang and Xu 2021
C.14	Land use	Mixed	Ren et al. 2022

Several researchers have conducted surveys to assess public attitudes toward and acceptance of AVs in relation to safety (Pettigrew et al. 2019; Qu et al. 2019; Hasan et al. 2022; Patel et al. 2022; Patel et al. 2023c), and their findings have consistently shown that the public's perception, particularly with regard to safety, is the biggest obstacle to AV adoption (Etminani-Ghasrodashti et al. 2021). Thus, failing to ensure that road users feel safe interacting with AVs could negatively

impact other modes of transportation, like cycling and walking (Patel et al. 2023d), and make it more difficult for AVs to be integrated into public roads (Etminani-Ghasroddashti et al. 2023b).

The public's trust and acceptance of AV technology and confidence in the vehicles' safety are prerequisites for their full deployment (Patel et al. 2023e), and manufacturers play a major role in allaying fears about situations involving unavoidable collisions. The moral question of who bears the legal and ethical responsibility for a collision involving a fully autonomous vehicle (Hevelke and Nida-Rümelin 2015; Aguiar et al. 2022) has been the subject of numerous studies. Numerous studies have explored the question of who bears legal and ethical responsibility in the event of a collision involving a fully autonomous vehicle (Hevelke and Nida-Rümelin 2015; Aguiar et al. 2022). When faced with an unavoidable collision, the AV can either stay on course and crash into the oncoming car, or it can swerve and crash into a pedestrian. To further complicate the issue, AVs require more time than human drivers in conventional cars to make decisions that determine the vehicles' actions (Robinson et al. 2022). This dilemma has been addressed in great detail in the literature (Santoni de Sio 2017; Martinho et al. 2021), but the variety of possible outcomes, ambiguous nature of ethics, and differences in people's beliefs, values, opinions, and feelings make the decision-making process very difficult (Martinho et al. 2021; Robinson et al. 2022). One thing that seems certain, however, is that the public needs to be involved in the decision-making process (Robinson et al. 2022). Since public acceptance is believed to be the primary barrier to the adoption of AVs, a widely accepted solution to the challenge of unavoidable collisions should be implemented to boost public acceptance and trust. Table 2 provides a list of strategies that stakeholders can enact to mitigate the moral quandaries that unavoidable collisions present while simultaneously addressing AV public safety concerns.

Table 2. Strategies to help stakeholders manage AV challenges

ID	Strategy Recommendation	Challenge	Stakeholder	Previous Study
1	Educate the public about AVs with informational materials	Public perception	Policymakers	Pettigrew et al. 2019; Patel et al. 2023f
2	Impose responsibility on drivers for collisions that could have been avoided	Unavoidable collisions		Hevelke and Nida-Rümelin 2015
3	Establish tort liability for AV manufacturers that promotes the advancement and innovation of AVs	Unavoidable collisions		Hevelke and Nida-Rümelin 2015
4	Determine how AVs are designed to respond in the event of an unavoidable collision	Unavoidable collisions		Santoni de Sio 2017
5	Establish a safety assurance procedure for AVs operating at Levels 3 and above	Public perception		Jacobo et al. 2019
6	Give the general public opportunities to interact with AVs	Public perception		Xing et al. 2022; Pamidimukkala et al. 2023b

7	Program the AV's decision-making process by using shared ethical decisions	Unavoidable collisions	Policymakers and manufacturers	Bonnefon et al. 2020
8	Engage the public in the AV's decision-making process and establish common ethical standards	Unavoidable collisions		Robinson et al. 2022
9	Test AVs extensively in actual traffic situations	Public Perception		Pettigrew et al. 2019
10	Design AVs with efficient collision avoidance	Public perception and unavoidable collisions	Manufacturers	Luetge 2017; Bonnefon et al. 2020
11	Provide a computational framework that is adaptable enough to accommodate many different moral principles to direct an AV's actions	Unavoidable collisions		Evans et al. 2022
12	Provide a fundamental framework that allows the decision matrix algorithms of AVs to incorporate human values	Unavoidable collisions		Umbrello and Yampolskiy 2022
13	Use a multi-sensor approach and accurate positioning technologies to improve pedestrian detection	Public perception		Parekh et al. 2022
14	Use design strategies to stop vehicles from operating in an improper or dangerous manner	Public perception and unavoidable collisions		Luetge 2017; Bonnefon et al. 2020
15	Incorporate pedestrian detection into AV's decision-making process	Public perception and unavoidable collisions		Chen et al. 2023

CONCLUSION

The purpose of this study was to explore the effects of AVs on traffic safety by identifying the factors that affect the severity of crashes involving them, addressing public concerns about their safety, and recognizing the moral ramifications of inevitable collisions. In addition, 15 strategies are provided that manufacturers and policymakers can employ to address the moral implications of unavoidable accidents and improve the public's perception of AVs.

An analysis of the data made public by the CA DMV supports the premise that the use of AVs will improve road safety. Although they were found to be more likely to be involved in rear-end collisions, they were rarely at fault for the accidents, and the number of other types of

accidents they were involved in was much smaller than that of conventional vehicles. As all the available data was from California, where the climate is mild, it was not possible to determine the role that bad weather plays in the frequency or severity of collisions.

The study also explored public acceptance of AVs and addressed the ethical dilemma of unavoidable collisions, both of which are considered challenges to their widespread adoption. The results showed that people who trust the technology are more likely to be convinced of the efficacy and safety of AVs. The significance of vulnerable road users feeling safe when interacting with AVs in a mixed traffic environment has also been emphasized by a number of authors, who expressed belief that if the technology is not trusted, there will be a delay in the integration of AVs into public roads and/or a reduction in walking and cycling as modes of transportation.

The ethical implications of AVs and inevitable accidents were explored, and strategies that legislators and manufacturers can use to increase public acceptance and enhance society's perception of AVs were identified. This study will benefit AV manufacturers and policymakers by providing them with a better understanding of AV crashes and equipping them with strategies that can help manage the challenges of AVs.

REFERENCES

- Aguiar, F., Hannikainen, I. R., and Aguilar, P. (2022). "Guilt without fault: accidental agency in the era of autonomous vehicles." *Science and engineering ethics*, 28(2), 11.
- Alozi, A. R., and Hussein, M. (2022). "Evaluating the safety of autonomous vehicle–pedestrian interactions: An extreme value theory approach." *Analytic methods in accident research*, 35, 100230.
- Beck, J., Arvin, R., Lee, S., Khattak, A., and Chakraborty, S. (2023). "Automated vehicle data pipeline for accident reconstruction: new insights from LiDAR, camera, and radar data." *Accident Analysis & Prevention*, 180, 106923.
- Bonnefon, J. F., et al. (2020). "Ethics of connected and automated vehicles: Recommendations on road safety, privacy, fairness, explainability and responsibility."
- Channamallu, S. S., Kermanshachi, S., and Pamidimukkala, A. (2023). "Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles." In *International Conference on Transportation and Development 2023*, 39-50.
- Chen, Y., Li, S., Tang, X., Yang, K., Cao, D., and Lin, X. (2023). "Interaction-Aware Decision Making for Autonomous Vehicles." *IEEE Transactions on Transportation Electrification*.
- Combs, T. S., Sandt, L. S., Clamann, M. P., and McDonald, N. C. (2019). "Automated vehicles and pedestrian safety: exploring the promise and limits of pedestrian detection." *American journal of preventive medicine*, 56(1), 1-7.
- Das, S., Dutta, A., and Tsapakis, I. (2020). "Automated vehicle collisions in California: Applying Bayesian latent class model." *IATSS research*, 44(4), 300-308.
- Dixit, V. V., Chand, S., and Nair, D. J. (2016). "Autonomous vehicles: disengagements, accidents and reaction times." *PLoS one*, 11(12).
- Etmian-Ghasrodashti, R., Ketankumar Patel, R., Kermanshachi, S., Michael Rosenberger, J., and Weinreich, D. (2021). "Exploring Concerns and Preferences towards Using Autonomous Vehicles as a Public Transportation Option: Perspectives from a Public Focus Group Study." In *International Conference on Transportation and Development 2021*, 344-354.

- Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023a). "Capturing the Behavioral Determinant behind the Use and Adoption of Shared Autonomous Vehicles (SAVs)." In *Transportation Research Board 102 Annual Meeting*, 23-00181.
- Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023b). "Exploring motivating factors and constraints of using and adoption of shared autonomous vehicles (SAVs)." *Transportation Research Interdisciplinary Perspectives*, 18, 100794.
- Evans, K., de Moura, N., Chauvier, S., Chatila, R., and Dogan, E. (2020). "Ethical decision making in autonomous vehicles: The AV ethics project." *Science and engineering ethics*, 26, 3285-3312.
- Faisal, A., Kamruzzaman, M., Yigitcanlar, T., and Currie, G. (2019). "Understanding autonomous vehicles." *Journal of transport and land use*, 12(1), 45-72.
- Favarò, F., Eurich, S., and Nader, N. (2018). "Autonomous vehicles' disengagements: Trends, triggers, and regulatory limitations." *Accident Analysis & Prevention*, 110, 136-148.
- Favarò, F. M., Nader, N., Eurich, S. O., Tripp, M., and Varadaraju, N. (2017). "Examining accident reports involving autonomous vehicles in California." *PLoS one*, 12(9).
- Guo, X., and Zhang, Y. (2022). "Maturity in automated driving on public roads: a review of the six-year autonomous vehicle tester program." *Transportation research record*, 2676(11), 352-362.
- Hafeez, F., Ullah Sheikh, U., Mas' ud, A. A., Al-Shammari, S., Hamid, M., and Azhar, A. (2022). "Application of the theory of planned behavior in autonomous vehicle-pedestrian interaction." *Applied Sciences*, 12(5), 2574.
- Hasan, O., McColl, J., Pfefferkorn, T., Hamadneh, S., Alshurideh, M., and Kurdi, B. (2022). "Consumer attitudes towards the use of autonomous vehicles: Evidence from United Kingdom taxi services." *International Journal of Data and Network Science*, 6(2), 537-550.
- Hevelke, A., and Nida-Rümelin, J. (2015). "Responsibility for crashes of autonomous vehicles: An ethical analysis." *Science and engineering ethics*, 21, 619-630.
- Houseal, L. A., Gaweesh, S. M., Dadvar, S., and Ahmed, M. M. (2022). "Causes and effects of autonomous vehicle field test crashes and disengagements using exploratory factor analysis, binary logistic regression, and decision trees." *Transportation research record*, 2676(8), 571-586.
- Iftikhar, S., Zhang, Z., Asim, M., Muthanna, A., Koucheryavy, A., and Abd El-Latif, A. A. (2022). "Deep Learning-Based Pedestrian Detection in Autonomous Vehicles: Substantial Issues and Challenges." *Electronics*, 11(21), 3551.
- Jacobo, A. M., Nobuyuki, U., Kunio, Y., Koichiro, O., Eiichi, K., and Satoshi, T. (2019). "Development of a safety assurance process for autonomous vehicles in Japan." In *Proceedings of ESV Conference*.
- Kaur, K., and Rampersad, G. (2018). "Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars." *Journal of Engineering and Technology Management*, 48, 87-96.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Michael Rosenberger, J., and Foss, A. (2022). "Usage Patterns and Perceptions of Shared Autonomous Vehicles (SAVs): Empirical Findings from a Self-driving Service." In *Automated People Movers and Automated Transit Systems 2022*, 94-104.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023). "A User and Ridership Evaluation of Shared Autonomous Vehicles." *Journal of Urban Planning and Development*, 149(1), 05022048.

- Khattak, Z. H., Fontaine, M. D., and Smith, B. L. (2020). "Exploratory investigation of disengagements and crashes in autonomous vehicles under mixed traffic: An endogenous switching regime framework." *IEEE Transactions on intelligent transportation systems*, 22(12), 7485-7495.
- Leilabadi, S. H., and Schmidt, S. (2019). "In-depth analysis of autonomous vehicle collisions in California." In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 889-893.
- Liu, Q., Wang, X., Wu, X., Glaser, Y., and He, L. (2021). "Crash comparison of autonomous and conventional vehicles using pre-crash scenario typology." *Accident Analysis & Prevention*, 159, 106281.
- Luetge, C. (2017). "The German ethics code for automated and connected driving." *Philosophy & Technology*, 30, 547-558.
- Martinho, A., Herber, N., Kroesen, M., and Chorus, C. (2021). "Ethical issues in focus by the autonomous vehicles industry." *Transport reviews*, 41(5), 556-577.
- Novat, N., Kidando, E., Kutela, B., and Kitali, A. E. (2023). "A comparative study of collision types between automated and conventional vehicles using Bayesian probabilistic inferences." *Journal of safety research*, 84, 251-260.
- Pamidimukkala, A., Kermanshachi, S., and Patel, R. (2023a). "An Exploratory Analysis of Crashes Involving Autonomous Vehicles." In *International Conference on Transportation and Development 2023*, 343-350.
- Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., and Hladik, G. (2023b). "Evaluation of barriers to electric vehicle adoption: A study of technological, environmental, financial, and infrastructure factors." *Transportation Research Interdisciplinary Perspectives*, 22, 100962.
- Parekh, D., Poddar, N., Rajpurkar, A., Chahal, M., Kumar, N., Joshi, G. P., and Cho, W. (2022). "A review on autonomous vehicles: Progress, methods and challenges." *Electronics*, 11(14), 2162.
- Patel, R. K., Etminani-Ghasradashti, R., Kermanshachi, S., Michael Rosenberger, J., and Foss, A. (2022). "How Riders Use Shared Autonomous Vehicles." In *Automated People Movers and Automated Transit Systems 2022*, 81-93.
- Patel, R. K., Channamallu, S. S., Khan, M. A., Kermanshachi, S., and Pamidimukkala, A. (2023b). "An Exploratory Analysis of Temporal and Spatial Patterns of Autonomous Vehicle Collisions." *Public Works Management & Policy*, 1087724X231217677.
- Patel, R. K., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023f). "Exploring Factors Affecting Shared Autonomous Vehicle Adoption: A Structural Equation Modeling Analysis." In *Transportation Research Board 102 Annual Meeting*, 23-00216.
- Patel, R. K., Etminani-Ghasradashti, R., Kermanshachi, S., Michael Rosenberger, J., and Foss, A. (2023c). "Exploring People's Attitudes and Perceptions of Using Shared Autonomous Vehicles: A Focus Group Study." In *International Conference on Transportation and Development 2023*, 231-240.
- Patel, R. K., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023e). "Exploring willingness to use shared autonomous vehicles." *International Journal of Transportation Science and Technology*, 12(3), 765-778.
- Patel, R. K., Etminani-Ghasradashti, R., Kermanshachi, S., Rosenberger, J. M., Pamidimukkala, A., and Foss, A. (2023d). "Identifying individuals' perceptions, attitudes, preferences, and concerns of shared autonomous vehicles: During-and post-implementation evidence." *Transportation Research Interdisciplinary Perspectives*, 18, 100785.

- Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. (2023a). "Users' and Nonusers' Attitudes and Perceptions of Shared Autonomous Vehicles: A Case Study in Arlington, Texas." In *International Conference on Transportation and Development 2023*, 241-252.
- Petrović, Đ., Mijailović, R., and Pešić, D. (2020). "Traffic accidents with autonomous vehicles: type of collisions, manoeuvres and errors of conventional vehicles' drivers." *Transportation research procedia*, 45, 161-168.
- Pettigrew, S., Worrall, C., Talati, Z., Fritschi, L., and Norman, R. (2019). "Dimensions of attitudes to autonomous vehicles." *Urban, Planning and Transport Research*.
- Qu, W., Xu, J., Ge, Y., Sun, X., and Zhang, K. (2019). "Development and validation of a questionnaire to assess public receptivity toward autonomous vehicles and its relation with the traffic safety climate in China." *Accident Analysis & Prevention*, 128, 78-86.
- Ren, W., Yu, B., Chen, Y., and Gao, K. (2022). "Divergent effects of factors on crash severity under autonomous and conventional driving modes using a hierarchical Bayesian approach." *International journal of environmental research and public health*, 19(18), 11358.
- Robinson, J., Smyth, J., Woodman, R., and Donzella, V. (2022). "Ethical considerations and moral implications of autonomous vehicles and unavoidable collisions." *Theoretical issues in ergonomics science*, 23(4), 435-452.
- Santoni de Sio, F. (2017). "Killing by autonomous vehicles and the legal doctrine of necessity." *Ethical Theory and Moral Practice*, 20, 411-429.
- Schoettle, B., and Sivak, M. (2015). *A preliminary analysis of real-world crashes involving self-driving vehicles*. University of Michigan Transportation Research Institute.
- Sinha, A., Vu, V., Chand, S., Wijayaratna, K., and Dixit, V. (2021). "A crash injury model involving autonomous vehicle: Investigating of crash and disengagement reports." *Sustainability*, 13(14), 7938.
- Song, Y., Chitturi, M. V., and Noyce, D. A. (2021). "Automated vehicle crash sequences: Patterns and potential uses in safety testing." *Accident Analysis & Prevention*, 153, 106017.
- Umbrello, S., and Yampolskiy, R. V. (2022). "Designing AI for explainability and verifiability: a value sensitive design approach to avoid artificial stupidity in autonomous vehicles." *International Journal of Social Robotics*, 14(2), 313-322.
- Wang, S., and Li, Z. (2019). "Exploring the mechanism of crashes with automated vehicles using statistical modeling approaches." *PloS one*, 14(3).
- Xing, Y., Zhou, H., Han, X., Zhang, M., and Lu, J. (2022). "What influences vulnerable road users' perceptions of autonomous vehicles? A comparative analysis of the 2017 and 2019 Pittsburgh surveys." *Technological Forecasting and Social Change*, 176, 121454.
- Xu, C., Ding, Z., Wang, C., and Li, Z. (2019). "Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes." *Journal of safety research*, 71, 41-47.
- Ye, W., Wang, C., Chen, F., Yan, S., and Li, L. (2021). "Approaching autonomous driving with cautious optimism: analysis of road traffic injuries involving autonomous vehicles based on field test data." *Injury prevention*, 27(1), 42-47.
- Yuan, Q., Xu, X., Wang, T., and Chen, Y. (2022). "Investigating safety and liability of autonomous vehicles: Bayesian random parameter ordered probit model analysis." *Journal of intelligent and connected vehicles*, 5(3), 199-205.

- Zhang, J., and Xu, C. (2021). "Investigating the typical scenarios and contributory factors to crash severity of autonomous vehicle involved collisions using association rule analysis". In *Proceedings of the 100th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 5-29.
- Zhu, S., and Meng, Q. (2022). "What can we learn from autonomous vehicle collision data on crash severity? A cost-sensitive CART approach." *Accident Analysis & Prevention*, 174, 106769.

Autonomous Vehicle Safety: A Comprehensive Analysis of Crash Injury Determinants

Sai Sneha Channamallu¹; Deema Almaskati²; Sharareh Kermanshachi, Ph.D., P.E.³;
and Apurva Pamidimukkala, Ph.D.⁴

¹Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: sxc5928@mavs.uta.edu

²Ph.D. Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: deemanabeel.almaskati@mavs.uta.edu

³Associate Vice Chancellor and Associate Dean of Research, Pennsylvania State Univ., State College, PA. Email: svk5464@psu.edu

⁴Assistant Professor, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX.
Email: apurva.pamidimukkala@mavs.uta.edu

ABSTRACT

Traffic-related fatalities are a significant global challenge that is primarily attributed to human error. Autonomous vehicles (AVs) hold the promise of mitigating such fatalities by removing the potential for human error and thereby enhancing road safety; however, they too are involved in crashes that result in injuries. Prior research, limited by the scope of the dataset and depth of analysis, has not fully explored these factors, resulting in a lingering lack of clarity about the cause of the crashes. This study analyzed 358 AV crash records from California's Department of Motor Vehicles' database for the period spanning from 2014 to July 2023. A bagging classifier model was utilized to predict the likelihood of injuries occurring in AV crashes, and various features of the vehicles were evaluated to determine their impact on crash outcomes. The findings revealed vehicle damage as the key predictor of injury risk, with a non-linear damage-injury relationship. Characteristics specific to manufacturers were also found to play an important role, demonstrating their differences in technology and safety features. The study underscored the influence of vehicle and collision types on injury occurrences, particularly highlighting the vulnerability of two-wheeler drivers and the prevalence of sideswipe and rear-end collisions. Additionally, it revealed pre-collision movement patterns, especially those involving stationary vehicles, as crucial risk factors. The impact of different road types, like streets and avenues, on crash severity was also noted, further enriching the analysis. The findings offer valuable implications for AV manufacturers, policymakers, and urban planners in enhancing vehicle safety features.

KEYWORDS: Autonomous Vehicles, Crash Injuries, Bagging Classifier, Machine Learning, Road Safety, Vehicle Safety.

INTRODUCTION

Traffic accidents tragically claim lives and cause injuries to millions of people around the world each year, posing a significant threat to economic development (Governing, 2022; Channamallu et al., 2023a). In the United States, these incidents are a leading cause of death, especially among teenagers and young adults. The year 2020 alone saw nearly 36,000 fatalities and over 1.5 million injuries due to traffic accidents in the U.S. (NHTSA, 2022; Etminani-

Ghasrodashti et al., 2022a; Patel et al., 2022a), the economic burden of which was estimated at a staggering \$474 billion in 2022 (NSC, 2021; Khan et al., 2022a). A predominant factor in 94% of these accidents is driver error, with issues such as alcohol consumption, distractions, drug use, and fatigue contributing to more than 40% of the fatal crashes (Chen et al., 2020; Khan et al., 2022b; Patel et al., 2023a). In this context, autonomous driving technologies have emerged as a beacon of hope, drawing worldwide attention for their potential benefits, which range from reducing emissions to significantly enhancing traffic safety by removing the element of human error (Novat et al. 2023; Pamidimukkala et al., 2023b). Autonomous vehicles (AVs) are particularly promising in their ability to drastically lower accident rates by eliminating human errors in driving; however, they also present unique challenges and require extensive testing and evaluation before widespread deployment to ensure they truly contribute to safer roads (Sinha et al., 2021; Khan et al., 2023b).

To better understand AVs and ensure their safety, multiple states have permitted them to be tested on public and private roads (Song et al., 2021; Pamidimukkala et al., 2023a Channamallu et al., 2023b). Manufacturers who have been approved to test their vehicles on public California roads have been mandated to document AV-involved collisions and report this information to the California Department of Motor Vehicles (CA DMV) (Ren et al. 2022; Etminani-Ghasrodashti et al., 2022b; Patel et al., 2023b). Extensive research has been conducted on this data to assess the operational behavior of AVs, using both descriptive analysis (Ye et al., 2021; Patel et al., 2022b) and a range of statistical models (Alambeigi et al., 2020; Khan et al., 2022c). The primary objective of many of these studies, including those conducted by Das et al. (2020) and Houseal et al. (2022), was to identify the factors that contribute to the severity and frequency of crashes; however, the research was limited by the small scope of the datasets, as Das et al. evaluated data from 2014 to 2109 and Houseal et al. from 2017 to 2021.

This study aims to offer a timely and informed perspective that reflects the effect that the latest advancements in AV technology have on the factors that impact the severity of crashes. It employs sophisticated machine learning methods to scrutinize the most recent AV crash data from the CA DMV and focuses on identifying and ranking the factors that contribute to these accidents. The primary objective of this research is to determine and arrange these factors according to their predictive value in assessing AV crash outcomes. The findings are poised to be particularly beneficial for policymakers, AV manufacturers, and transportation experts, offering valuable insights into the performance and behavior of AVs on public roadways.

LITERATURE REVIEW

To support the advancement of autonomous driving technology, the CA DMV has permitted manufacturers to perform on-road testing, but they are required to document the occurrence of any crashes and disengagements (Das et al., 2020). Disengagement refers to instances when the automated driving system (ADS) either hands over control to a human driver or when a human driver feels the need to intervene and take control from the ADS (CA DMV, 2022; Channamallu et al., 2023). Researchers globally have utilized this open-source data for analyzing and interpreting the performance of AVs and how they can benefit road safety (Etminani-Ghasrodashti et al., 2023; Khan et al., 2023a). A summary of previous studies, their characteristics, and significant findings is shown in Table 1.

Although the findings of these studies provide illuminating insights into AV behaviors, they are constrained by the limited dataset and may not be representative of the recent developments in autonomous driving technology. Thus, this paper intends to supplement the existing body of

knowledge by providing updated insights on the prediction of injury likelihood in AV crashes, feature importance, and determination of the impact of various factors on crash outcomes.

Table 1. Previous studies and significant findings of AV crashes

Authors	Study Period	Analysis Method	Significant Findings
Wang and Li 2019	2014 – 2018	Ordinal and binary logistic regression and CART classification tree	<ul style="list-style-type: none"> ▪ Rear-end crashes were the most common type of collision. ▪ In 87% of AV-related accidents, the collision was not caused by the AV; when it was at fault, crash severity was shown to be lower. ▪ Highway accidents were more likely to result in severe injuries.
Alambeigi et al., 2020	October 2014 – June 2019	Probabilistic topic modeling analysis	<ul style="list-style-type: none"> ▪ Many sideswipe collisions involved motorcycles or were associated with manual disengagements.
Das et al., 2020	2014 – September 2019	Bayesian latent classification model	<ul style="list-style-type: none"> ▪ Rear-end crashes were the most common type of collision. ▪ Increased risk of serious injuries resulted from turning motions, multi-vehicle accidents, poor lighting, and sideswipe and rear-end collisions.
Song et al., 2021	2015 –2019	Sequence analysis and clustering	<ul style="list-style-type: none"> ▪ Disengagements were observed in 24% of crash sequences.
Ye et al., 2021	2017 – June 2019	Descriptive analysis and binary logistic regression	<ul style="list-style-type: none"> ▪ Head and neck injuries are the most frequent types of injuries in crashes involving AVs. AV occupants account for 71% of all injuries in AV-related accidents. Poor lighting is responsible for most AV injuries.
Houseal et al., 2022	2017 – 2021	Logistic regression and decision trees	<ul style="list-style-type: none"> ▪ Factors that contribute to AV collisions include the type of crash and movements of the AV or the other vehicle.
Ren et al., 2022	May 2018 – March 2021	Hierachal Bayesian approach	<ul style="list-style-type: none"> ▪ The driving mode significantly impacted the severity of the crashes. In autonomous driving mode, mixed land use was the major contributor.
Novat et al., 2023	2017 – 2020	Bayesian network model	<ul style="list-style-type: none"> ▪ Speeds under 45 mph and signalized junctions increase the likelihood of rear-end incidents for AVs by 43%. ▪ Crash types like head-on, sideswipe, and other types are less frequent with AVs.

METHODOLOGY

This section details the methodology used to analyze and predict the outcomes of AV crashes. It outlines the systematic approach, from the initial data gathering phase through identifying the key factors that influence the outcomes of crashes. The methodology is presented below in Figure 1.

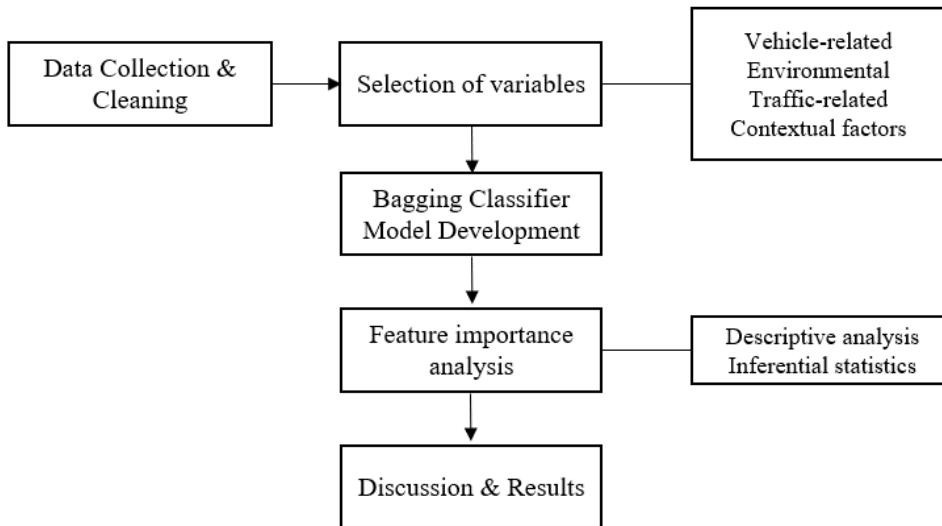


Figure 1. Methodology

The research utilized data sourced from reports submitted to the CA DMV by companies conducting AV on-road tests and included various attributes of the vehicles and a range of environmental, traffic, and contextual factors. The data underwent a rigorous pre-processing stage that involved cleaning, normalizing, and transforming the data to guarantee its quality and consistency. During this phase, outliers were managed, and missing values were eliminated to preserve the integrity of the dataset for subsequent model development.

A bagging classifier model was developed, with a key focus on analyzing the feature importance scores. This analysis was crucial in ranking the variables according to their impact on the model's predictions, thereby identifying the most pivotal factors in forecasting crash outcomes. An in-depth discussion ensued, interpreting the significance of the highest-ranked features. The study concluded with a synthesis of the findings from the most effective model, delineating a clear, evidence-based hierarchy of features that are influential in determining the probability of injuries resulting from AV crashes.

Dataset

The research utilized data obtained from the CA DMV, which mandated that the operators engaged in on-road AV testing submit accident reports in PDF format within 10 working days. The methodology entailed gathering information from each PDF report available on the CA DMV's website from 2014 to July 2023, extracting the data, then systematically arranging it in an Excel spreadsheet to facilitate quantitative analysis. Figure 2 illustrates the annual occurrence of AV collisions as recorded in the dataset. Following the data cleaning process, the dataset

consisted of 358 incident entries and 16 variables that relate to the attributes of the AV, the characteristics of the other party involved in the incident, the prevailing traffic conditions at the time, and additional details as mentioned in the accident reports. These variables are other party's vehicle status, Road Surface, Weather, AV status, Vehicles involved, Parking provision, Signal, Mode, Intersection Geometry, Lighting, Road type, Movement preceding collision, Collision type, Vehicle type, Company, and Damage to AV.

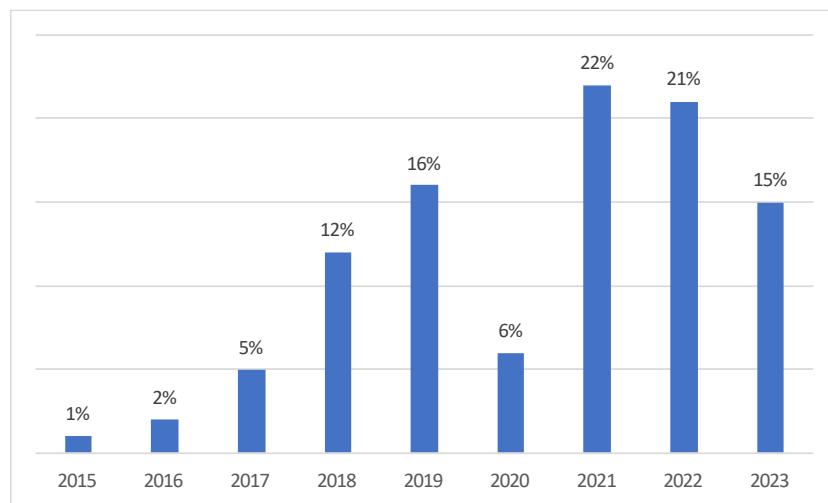


Figure 2. Yearly distribution of AV crashes

Model Development

The crash prediction model developed for the study included 16 independent variables and one dependent variable that predicts the likelihood of injury from a crash. The dependent variable was binary and coded as 0 for crashes without injuries and 1 for those with injuries. There was a notable discrepancy and uneven data distribution of the number of instances in each category, with 331 cases of non-injury crashes and only 27 involving injuries. To counteract this imbalance, the model used class weighting techniques and assigned weights inversely proportional to the frequency of each category's occurrence in the dataset. This approach was essential for addressing the imbalances, as it amplifies the significance of underrepresented categories and ensures a fairer and more balanced learning process.

Bagging classifiers are known for their ability to improve predictive accuracy, especially for minority classes such as injury-related crashes, which occur less frequently. Their ensemble methodology greatly diminished the likelihood of overfitting and enhanced the stability of the model, rendering it more dependable for our analysis. The resilience of this model amid noise and outliers, along with its ability to offer meaningful insights into feature importance, aligned well with the goals of our study. The practicality, straightforward implementation, and some level of interpretability of bagging classifiers also contributed to their selection, facilitating both understanding and an explanation of the outcomes. Additionally, our approach built on the research performed by Sinha et al. (2021), which utilized a bagging classifier but restricted their study to data prior to 2019. Our research expanded this timeframe to July 2023, yielding a more extensive view of AV crash-related injuries. The setup parameters for model are presented in table 2. For training and testing purposes, the data was divided, allocating 80% for training and

20% for testing. To counteract the bias due to uneven class distribution, stratified sampling was employed to divide the dataset, to ensure that the proportion of each class was maintained in both the training and testing sets.

Table 2. Setup parameters for bagging classifier model

Parameters	Assignment
Base estimator	Decision tree classifier
n_estimators (the number of base estimators in the ensemble)	9
Random state	42

RESULTS & DISCUSSION

The bagging classifier achieved a balanced accuracy of 59%, demonstrating its proficiency in striking a balance between sensitivity (true positive rate) and specificity (true negative rate). This balance is further highlighted by its high precision rate of 94% and recall rate of 97%, indicating its effectiveness in accurately classifying positive instances and efficiently identifying actual positive cases. The model attained a modified F1 score of 2.39, which, unlike the traditional F1 score, adjusts metric accounts for class imbalances and reflects the model's admirable performance in scenarios where data is unevenly distributed across different classes. A critical aspect of the model's analysis was understanding the varying influence of different features, which is crucial for prioritizing factors that significantly enhance road safety and crash prevention strategies. The varying degrees of influence of the model's features were quantified and ranked, and the results showed that the extent of the damage sustained by the AV is the most significant predictor of injuries. (See Figure 3.) This insight into the importance of features is vital for directing future efforts in AV safety and accident mitigation.

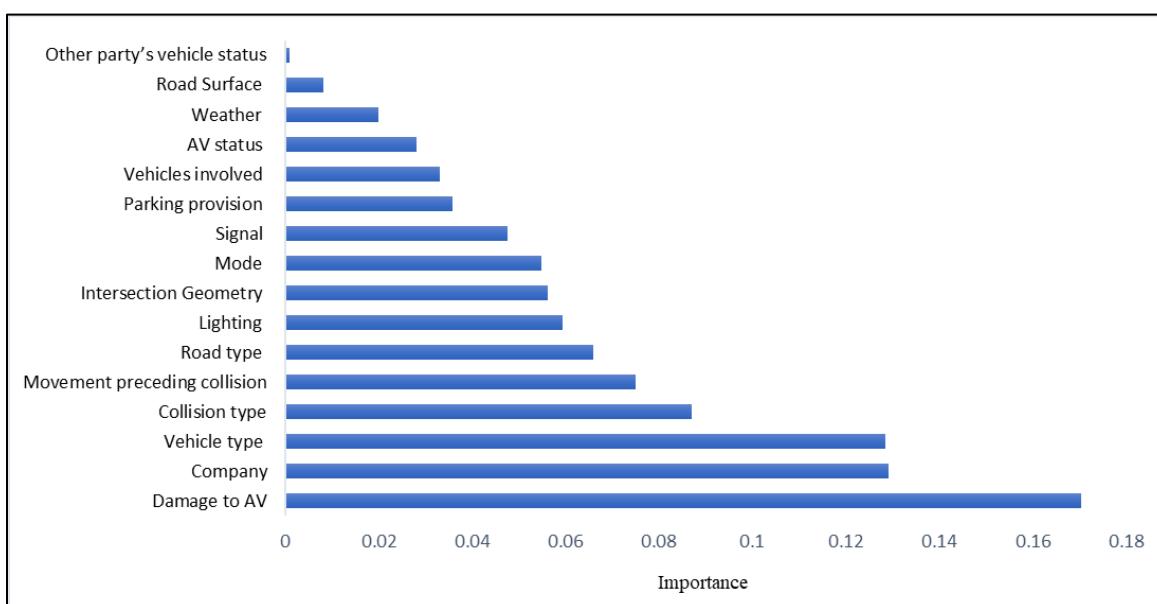


Figure 3. Feature importance score

The manufacturer of the AV was identified as the second significant factor in the model, underscoring the impact of the different technologies, safety features, and operational protocols employed by each. The type of vehicle ranked as the third most important factor, which aligns with the varied safety features that each type comes equipped with. Other key predictors are the type of collision and the movement patterns of the AV prior to the crash. The type of road on which the accidents occur is important, but less so than the vehicle-specific factors.

The geometry of intersections and lighting conditions were shown to be moderately important features, though not as pivotal as vehicle-related factors, and the presence and type of traffic control signals, along with parking provisions, were found to be also relevant to crash outcomes. Other features like the number of vehicles involved, the status of the AV and the other vehicle, and the weather and road surface conditions had a lesser impact on the model, and while they do play a role in crash predictions, their influence is relatively minor compared to other more prominent features.

The study uncovered complex interrelations of diverse elements by integrating empirical data with the results of predictive modeling to offer a detailed examination of the factors influencing injuries in AV crashes. Figure 4 provides a statistical overview of the critical determinants and highlights their significance in the context of AV crash injuries.

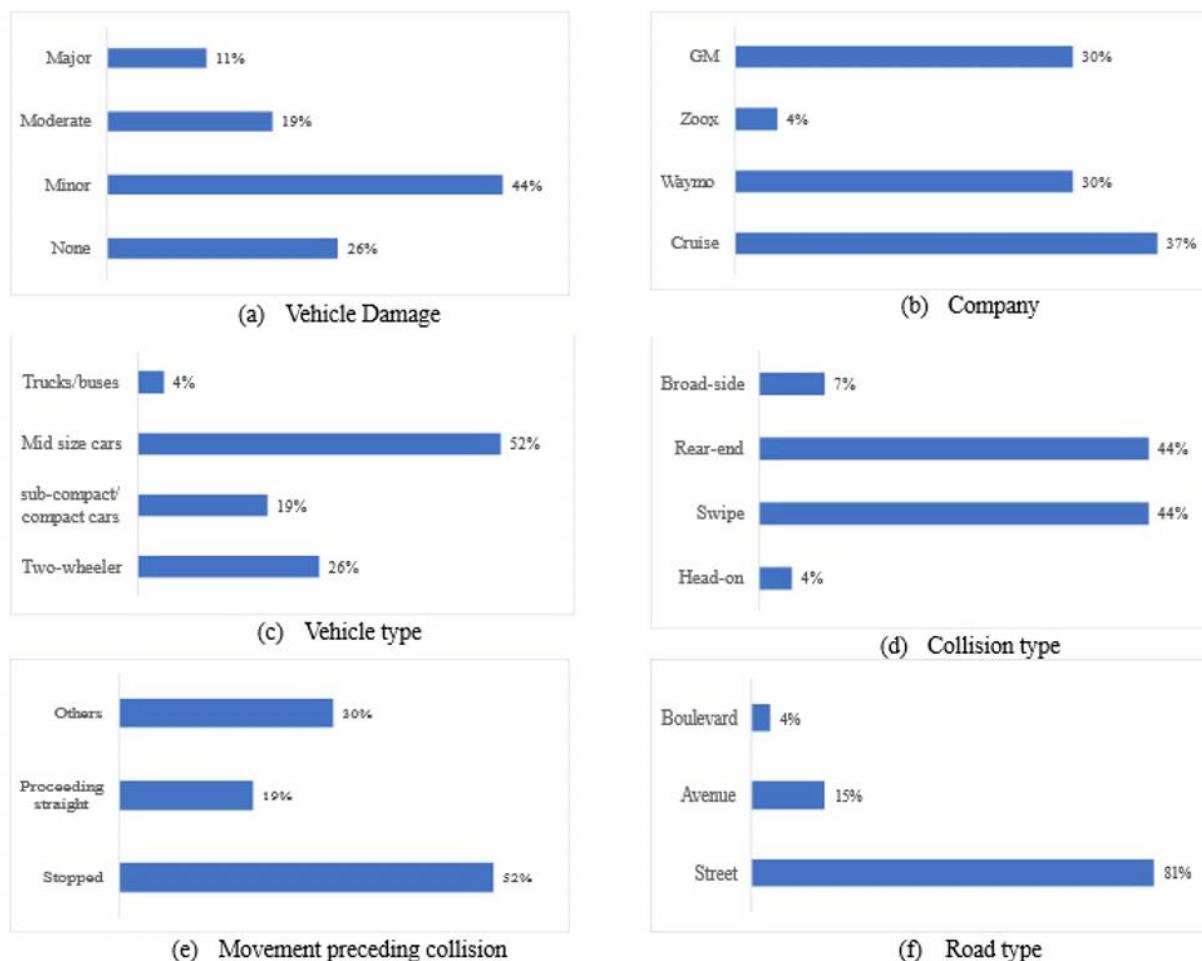


Figure 4. Statistics of critical determinants of AV crash injuries

Vehicle damage

An analysis of the vehicle damage sustained in AV crashes that resulted in injuries revealed a diverse pattern. There was no reported damage in 26% of the accidents, hinting that the injuries might have stemmed from indirect causes such as abrupt stops. Minor damage was noted in 44% of the cases, indicating that low-impact crashes can result in considerable injuries. Moderate damage was present in 19% of the cases and supports the idea that the degree of vehicle damage correlates with the risk of injuries, though not in a straightforward manner. Major vehicle damage occurred in 11% of the injury incidents, emphasizing the risk of severe injuries in high-impact scenarios and the necessity for effective vehicle safety features. The patterns of damage in crashes show that even minor damage can lead to significant injuries, which calls for enhanced occupant safety measures. This non-linear correlation between damage severity and injury probability underscores the importance of vehicle damage as a predictive element in the model.

Company

The variations in injury-related crash rates among companies like Cruise, GM, and Waymo imply that factors unique to each brand, such as their technology, safety protocols, or operational methods, might play a significant role. Therefore, the model's focus on company attributes could mirror the distinct quality and safety reputations of the various manufacturers. Cruise, GM, and Waymo have comparable injury crash distributions (37%, 30%, and 30%, respectively), while Zoox exhibits a markedly lower rate at 4%, possibly indicative of differing operational zones, activity levels, or more effective safety strategies. These statistics shed light on the relative safety efficacy of the AV technologies of each company, offering valuable insights for consumers and regulatory bodies.

Vehicle type

The differences in injury-related crash rates among various types of vehicles highlight the unique risks associated with each. The distribution of vehicle types in injury-inducing crashes shows notable variations in risk profiles. Two-wheelers, with their open structure that offers less protection, constituted 26% of such crashes, indicating a higher vulnerability for drivers or riders compared to those in more structurally shielded vehicles. Sub-compact and compact cars were involved in 19% of the incidents, which might be attributed to their smaller size and potentially smaller capacity than larger vehicles for absorbing an impact. Mid-size cars were involved in the majority of crashes (52%), which in addition to other factors could be due to the greater number of them on the roads. Trucks and buses accounted for just 4%, possibly because their larger size offers more protection, or they are less frequently involved in such incidents. This data is crucial for understanding traffic safety dynamics and has significant implications for vehicle design and safety strategies.

Collision type

The distribution of injuries among different types of collisions is vitally important in predicting crash severity. Injuries resulting from head-on collisions accounted for only 4% of the cases, which could be due to effective prevention strategies or their less frequent occurrence.

Conversely, sideswipe and rear-end collisions each made up 44% of injury-causing incidents, indicating a higher risk and prevalence, possibly linked to factors like distracted drivers or inadequate spacing between vehicles. Broad-side collisions, primarily occurring at intersections, also pose a significant risk but were less frequent, representing 7% of injury-related crashes. The model's emphasis on sideswipe and rear-end collisions underscores the necessity for safety measures that specifically target these high-frequency incidents to reduce the risk of injury. The infrequency of head-on collisions as a cause of injury suggests that they may be considered lower risk. The model also pays attention to the risks associated with broad-side collisions, especially at intersections, where the impact severity can be particularly high. These findings highlight the critical need for effective safety features to mitigate the risk of injuries across various collision types.

Movement preceding collision

The analysis of crashes resulting in injuries based on the vehicle's movement before the collision revealed that the majority (52%) of injuries occurred when the vehicles were stationary. This highlights the particular vulnerability of stopped vehicles to incidents like being rear-ended or otherwise involved in collisions. In comparison, injury-related crashes that occurred while the vehicles were moving in a straight direction accounted for only 19%, which might be attributed to the predictable nature of this motion. The remaining 30% of injury-inducing crashes fall into other categories and include more complex driving maneuvers, suggesting that these less typical movements put them at higher risk. The high incidence of injuries in situations where vehicles are stopped might indicate scenarios where AVs are especially prone to accidents, such as being rear-ended, a factor that the model appears to consider. The more infrequent injuries that resulted from crashes when the vehicles were traveling in a straight line might be due to the inherent predictability and control of this movement, potentially viewed by the model as a lower-risk situation. This breakdown of data is essential for developing safety measures and highlights the need for protective strategies across different vehicle movements, not just while in transit but also when stationary. This insight can guide the development of safety protocols and technologies that address the specific risks associated with each type of vehicle movement.

Road type

Road type plays an important role in assessing injuries from AV crashes. The majority of the crashes occurred on streets rather than highways, likely due to increased traffic, more frequent stopping, and the presence of pedestrians. Avenues contributed to 15% of these incidents, possibly because their broader layout enables higher speeds that lead to more severe crashes. Boulevards, with only 4% of injury-related crashes, perhaps see fewer incidents because their design supports safer travel at high speeds with less stopping. This trend of higher rates of injuries on streets necessitates a more thorough investigation into how AVs function within city settings and suggests that their navigational systems may need enhanced capabilities to deal with the complexities of urban traffic.

CONCLUSIONS

AVs offer a unique opportunity to reduce the primary cause of death globally by removing human error in driving and thus improving road safety. The primary objective of this study was

to gain a deeper understanding of the factors that contribute to injury-producing accidents involving at least one AV by assessing the predictive value of each for crash outcomes. Data from 358 incidents and 16 variables from the CA DMV was analyzed by moving beyond basic quantitative analysis to a deeper comprehension of crash dynamics. Utilization of a bagging classifier model illuminated the complex factors affecting AV crash injuries, and the model's balanced accuracy, high precision and recall, and an adjusted F1 score, confirmed its effectiveness in accurately classifying and forecasting injury-prone incidents, even with the challenges of imbalanced data.

The research methodically aligned statistical trends from the data with the feature importance scores from the model to reveal in-depth insights into how these elements collectively affect the likelihood of injuries in AV crashes. The bagging classifier's performance, with a balanced accuracy of 0.59, precision of 0.94, and recall of 0.97, highlights key factors in crash prediction: Vehicle damage patterns reveal a non-linear relationship between damage severity and injury risk, irrespective of impact level. Company-specific technologies and safety protocols significantly influence crash outcomes, as evidenced by varying injury rates among companies. The type of vehicle plays a crucial role, with different types showing varying levels of vulnerability. Collision types, particularly the prevalence of sideswipe and rear-end collisions, necessitate targeted safety measures. The movement preceding the collision, especially the high incidence of injuries when vehicles are stationary, points to unique risks in non-moving scenarios. Lastly, higher crash rates on streets as opposed to highways suggest the complexities AVs face in urban environments, highlighting the importance of a comprehensive approach in enhancing road safety and crash prevention strategies. This detailed understanding of various influencing factors can guide the development of more precise and effective safety and prevention measures. Taken together, these findings emphasize the necessity for a comprehensive approach to road safety that is responsive to the diverse conditions and challenges presented by the ever-changing field of autonomous mobility.

This study was limited by its dependence on data from the geographically limited and possibly incomplete California DMV reports, which could skew the representation of global AV performance amid varying driving conditions and reporting norms. A richer dataset would mitigate the substantial imbalance between non-injury and injury crash data. Future investigations into AV crash analysis would benefit from broadening the data collection to include a wider range of geographical locations, which would facilitate a more comprehensive evaluation of AV performance under different traffic conditions and regulatory frameworks. Enriching the datasets with more detailed information from telematics and sensors could also provide a deeper insight into the intricate factors influencing AV incidents.

ACKNOWLEDGMENT

We gratefully acknowledge the support and generosity of the Transportation Consortium of South-Central States (Tran-SET), without which the present study could not have been completed.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Alambeigi, H., McDonald, A. D., and Tankasala, S. R. 2020. "Crash themes in automated vehicles: A topic modeling analysis of the California Department of Motor Vehicles automated vehicle crash database." *arXiv preprint*.
- California DMV. 2022. "Disengagement Reports - California DMV", <www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/disengagement-reports> Accessed on Oct 26, 2023.
- Channamallu, S. S., Kermanshachi, S., and Pamidimukkala, A. 2023. Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles. In *International Conference on Transportation and Development 2023* (pp. 39-50).
- Channamallu, S. S., Kermanshachi, S., Rosenberger, J. M., and Pamidimukkala, A. 2023a. A review of smart parking systems. *Transportation Research Procedia*, 73, 289-296.
- Channamallu, S. S., Kermanshachi, S., Rosenberger, J. M., and Pamidimukkala, A. 2023b. Parking occupancy prediction and analysis - a comprehensive study. *Transportation Research Procedia*, 73, 297-304.
- Chen, H., Liu, Z., Sun, X., and Zhou, R. 2020. "Analysis of factors affecting the severity of automated vehicle crashes using XGBoost model combining POI data," *Journal of Advanced Transportation*, vol. 2020, Article ID 8881545, 12 pages.
- Das, S., Dutta, A., and Tsapakis, I. 2020. "Automated vehicle collisions in California: Applying Bayesian latent class model." *IATSS research*, 44(4), 300-308.
- Etminani-Ghasrodashti, R., Patel, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022a. Exploring Factors Affecting Willingness to Ride Shared Autonomous Vehicles. In *Proceedings of Transportation Research Board 101 Annual Conference*.
- Etminani-Ghasrodashti, R., Ketankumar Patel, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022b. Modeling users' adoption of shared autonomous vehicles employing actual ridership experiences. *Transportation Research Record*, 2676(11), 462-478.
- Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2023. Exploring motivating factors and constraints of using and adoption of shared autonomous vehicles (SAVs). *Transportation Research Interdisciplinary Perspectives*, 18, 100794.
- Governing. 2022. "Deaths and Injuries in Road Crashes Are a Silent Epidemic on Wheels", Retrieved from www.governing.com/now/deaths-and-injuries-in-road-crashes-are-a-silent-epidemic-on-wheels. Nov 2, 2023.
- Houseal, L. A., Gaweesh, S. M., Dadvar, S., and Ahmed, M. M. 2022. "Causes and effects of autonomous vehicle field test crashes and disengagements using exploratory factor analysis, binary logistic regression, and decision trees." *Transportation research record*, 2676(8), 571-586.
- Khan, M. A., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Identifying Usage of Shared Autonomous Vehicles (SAVs): Early Findings from a Pilot Project. In *Transportation Research Board 101st Annual Meeting* Washington, DC.
- Khan, M. A., Etminani-Ghasrodashti, R., Shahmoradi, A., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022b. Integrating shared autonomous vehicles into existing transportation services: evidence from a paratransit service in Arlington, Texas. *IJCE*, 20(6), 601-618.
- Khan, M. A., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2022c. Do ridesharing transportation services alleviate traffic crashes? A time series analysis. *Traffic injury prevention*, 23(6), 333-338.

- Khan, M. A., Patel, R. K., Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., Hladik, G., and Foss, A. 2023a. Factors that determine a university community's satisfaction levels with public transit services. *Frontiers in Built Environment*, 9, 1125149.
- Khan, M. A., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., Pan, Q., and Foss, A. 2023b. Understanding Students' Satisfaction with University Transportation. In *International Conference on Transportation and Development 2023* (pp. 522-532).
- Leilabadi, S. H., and Schmidt, S. 2019. "In-depth analysis of autonomous vehicle collisions in California." In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 889-893.
- NHTSA. 2022. "National Statistics". Retrieved from <https://cdan.nhtsa.gov/tsftables/National%20Statistics.pdf>. Oct 26, 2023.
- Novat, N., Kidando, E., Kutela, B., and Kitali, A. E. 2023. "A comparative study of collision types between automated and conventional vehicles using Bayesian probabilistic inferences." *Journal of safety research*, 84, 251-260.
- NSC. 2021. "Motor Vehicle Deaths in 2020 Estimated to Be Highest in 13 Years, Despite Dramatic Drops in Miles Driven". Retrieved from www.nsc.org/newsroom/motor-vehicle-deaths-2020-estimated-to-be-highest. Oct 22, 2023.
- Pamidimukkala, A., Patel, R. K., Kermanshachi, S., Rosenberger, J. M., and Tanvir, S. 2023a. A Review on Shared Mobility and Electric Vehicles. In *International Conference on Transportation and Development 2023* (pp. 333-342).
- Pamidimukkala, A., Kermanshachi, S., and Patel, R. 2023b. An Exploratory Analysis of Crashes Involving Autonomous Vehicles. In *International Conference on Transportation and Development 2023* (pp. 343-350).
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Exploring willingness to use shared autonomous vehicles. *International Journal of Transportation Science and Technology*, 12(3), 765-778.
- Patel, R. K., Etminani-Ghasrodeshti, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022b. How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems, ASCE International Conference on Transportation & Development*, pp. 81-93.
- Patel, R., Channamallu, S. S., Khan, M. A., Kermanshachi, S., and Pamidimukkala, A. 2023a. An Exploratory Analysis of Temporal and Spatial Patterns of Autonomous Vehicle Collisions. *Public Works Management & Policy*.
- Patel, R. K., Pamidimukkala, A., Kermanshachi, S., and Etminani-Ghasrodeshti, R., 2023b. Disaster preparedness and awareness among university students: a structural equation analysis. *International journal of environmental research and public health*, 20(5), 4447.
- Ren, W., Yu, B., Chen, Y., and Gao, K. 2022. "Divergent effects of factors on crash severity under autonomous and conventional driving modes using a hierarchical Bayesian approach." *International journal of environmental research and public health*, 19(18), 11358.
- Sinha, A., Vu, V., Chand, S., Wijayaratna, K., and Dixit, V. 2021. "A crash injury model involving autonomous vehicle: Investigating of crash and disengagement reports." *Sustainability*, 13(14), 7938.
- Song, Y., Chitturi, M. V., and Noyce, D. A. 2021. "Automated vehicle crash sequences: Patterns and potential uses in safety testing." *Accident Analysis & Prevention*, 153, 106017.

- Wang, S., and Li, Z. 2019. "Exploring the mechanism of crashes with automated vehicles using statistical modeling approaches." *PloS one*, 14(3).
- Ye, W., Wang, C., Chen, F., Yan, S., and Li, L. 2021. "Approaching autonomous driving with cautious optimism: analysis of road traffic injuries involving autonomous vehicles based on field test data." *Injury prevention*, 27(1), 42-47.

Computer Vision Based Detector Integrated with V2X Framework for Vulnerable Road User Safety

Sai Bonthus¹; Vaishak Gopalakrishna²; William Martin³; Nick Hegemier⁴; Victor Hunt⁵;
and Arthur Helmicki⁶

¹Research Assistant, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, OH.
Email: bonthusy@mail.uc.edu

²Research Assistant, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, OH.
Email: gopalavk@mail.uc.edu

³CAV R&D Engineer, Leidos; Transportation Research, Turner-Fairbank Highway Research Center. Email: william.martin@leidos.com

⁴Managing Director, Infrastructure, DriveOhio (Ohio Dept. of Transportation).
Email: nick.hegemier@dot.ohio.gov

⁵Professor, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, OH.
Email: huntvj@ucmail.uc.edu

⁶Professor, Dept. of Electrical and Computer Engineering, Univ. of Cincinnati, OH.
Email: helmicaj@ucmail.uc.edu

ABSTRACT

About 20% of total fatalities, reported by National Highway Traffic Safety Administration (NHTSA), in 2022 are pedestrians and pedal cyclists. In a lot of instances, pedestrian crashes and injuries are not accurately reported which resulted in lack of accurate vulnerable road user (VRU) data. As computer vision models evolve, there is an opportunity to collect more accurate road user data and develop real-time safety response systems. The Federal Highway Administration (FHWA) and Intelligent Transportation Systems Joint Program Office (ITSJPO) initiative, connected and automated vehicle education (CAVe-in-a-Box), allows for the seamless adaptation of existing computer vision and machine learning models used in transportation data analytics. At its core, it utilizes vehicle-to-everything (V2X) Hub and a cellular-vehicle-to-everything (C-V2X) road-side unit (RSU). The C-V2X framework supports the use of long-term evolution (LTE) and other wireless communication protocols using sidelink to establish robust communication between vehicles and infrastructure. Although in an early stage of implementation, C-V2X has been gaining popularity over the dedicated short-range communication (DSRC) protocol due to different channelization schemes, timing requirements, and a wireless architecture similar to that used by cell phones. Contrary to DSRC, C-V2X uses multiple carriers in blocks of smaller subframes. In this research study, CAVe-in-a-Box infrastructure and mobile kits are assembled and integrated with a real-time computer vision-based pedestrian detector to detect, localize, and classify VRUs, such as pedestrians and bike-pedestrians in addition to vehicles. Furthermore, real-time alerts are generated in two ways: (1) by sending pedestrian detection calls to a traffic signal controller via synchronous data link control (SDLC) and (2) by transmitting V2X personal safety messages (PSMs) to RSUs.

INTRODUCTION

The existing and highly sophisticated transportation infrastructure includes a complex network of freeways, roadways, intersections, roundabouts, and valuable assets such as traffic

signs. As intelligent transportation systems (ITS) technologies such as automated traffic control products, connected and automated vehicles (CAVs), fleet management software, and passenger information systems evolve to interact with the existing transportation infrastructure, there is a huge risk of potential security threats to infrastructure, humans, and wildlife. Specifically, due to limited historical data, access to advanced technology innovations powered by AI and potential exploitation of cyber security vulnerabilities, the risks are significantly raised to cause loss of life (Sheehan, B. 2019).

CAVs are transforming the evolution of ITS due to their ability to communicate through Vehicle to Everything (V2X) to send and receive data with other Vehicles (V2V), Devices (V2D), Pedestrians (V2P), Cloud (V2C), and Infrastructure (V2I), at the same time (USDOT 2020). The V2X framework in CAVs is complex due to the nature of diverse sensor fusion between different systems. Currently, the controller area network (CAN) bus, FlexRay, local interconnect networks (LIN), and Ethernet protocols are used to communicate with radio detection and ranging (RADAR), Light Detection and Ranging (LiDAR), ultrasonic sensors, and stereo cameras within the vehicle to navigate and localize the CAVs (A. Eskandarian 2021). Although these sensors work collaboratively with each other, their data formats are unique and require an efficient gateway to execute proper vehicular sensor fusion. For instance, advanced technologies such as real-time image processing to detect traffic signs, lane markers (Daniel Braunstein 1971), and obstacles are implemented through advanced driver assistance systems (ADAS) which use the sensor fusion of high-definition cameras and ultra-sonic sensors in the vehicle. In addition, 4G/5G wireless communications are used with Vehicular Ad Hoc Networks (VANETs) to connect with the internet, other vehicles in the network, access infrastructure data, and compute real-time solutions to cooperative intersection control or ramp merging problems (I. Ahmad 2018).

Transportation infrastructure, on the other hand, has been rapidly acclimatizing to estimate traffic flow and provide safer commutes to drivers and trucking services. Human driven vehicles (HDVs) are not fading away anytime soon due to several technical and political reasons and soon, a combination of HDVs and CAVs in the traffic flow will be a reality (K. Higashiyama 2020). However, limited methods were proposed to qualify existing infrastructures in terms of their ability to support HDVs and CAVs. For instance, as clearly discussed in (Dehman, A. 2021), assuming a coexistence of work zones and CAVs, there are several benefits to access work zone data and plan alternative routes through V2I, but it could potentially cause adverse effects to the existing infrastructure, drivers, and workers. Moreover, in recent developments, traffic monitoring solutions are adopting advanced technologies such as computer vision and AI-based deep learning algorithms to detect and classify vehicles, pedestrians, and bike-pedestrians in the traffic videos. Deep learning is being applied on aerial view imagery to proactively solve challenges such as single/multi-lane roundabout or intersection capacity evaluation (P. Khekare 2022). However, it is a challenge to provide real-time meaningful data of mobility, safety, and environment such as vehicle classification, vulnerable road user (VRU) detection such as pedestrians and bike-pedestrians, and road or weather conditions.

In this paper, CAVe-In-A-Box infrastructure and mobile kits are assembled and integrated with a real-time computer vision-based pedestrian detector to detect, localize, and classify VRUs such as pedestrians and bike-pedestrians, and vehicles. This setup generates real-time alerts in two ways: 1) by sending ped detection calls to signal controller via synchronous data link control (SDLC) and 2) by transmitting V2X personal safety messages (PSM) to RSUs.

CONNECTED AND AUTOMATED VEHICLE EDUCATION (CAVe-IN-A-BOX)

CAVe-In-A-Box. The CAVe-In-A-Box Infrastructure kit, as shown in Figure 1a, consists of a traffic signal controller (TSC), RSU, network switch, and V2X Hub computer. IP address configuration of all devices in the CAVe-In-A-Box is highlighted in Table 1. In this study, Bosch Autodome Inteox 7000i camera has been utilized to detect pedestrians in real-time and send SDLC ped detection calls to the signal controller while transmitting PSMs to the RSU. PSM messages are further broadcasted to on-board units (OBUs) which are part of the CAVe-In-A-Box Mobile kit, shown in Figure 1b.

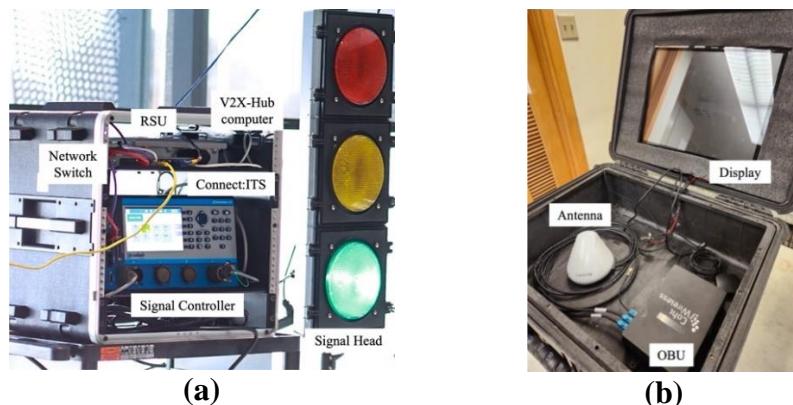


Figure 1. CAVe-In-A-Box (a) Infrastructure kit and (b) Mobile kit.

VRU SAFETY ALERTS WITH CAVE-IN-A-BOX

CAVe-In-A-Box supports several plug-ins including ‘immediate forward’ to transmit SPaT and PSM messages to RSUs and receive basic safety messages (BSMs) from OBUs as shown in Figure 2.

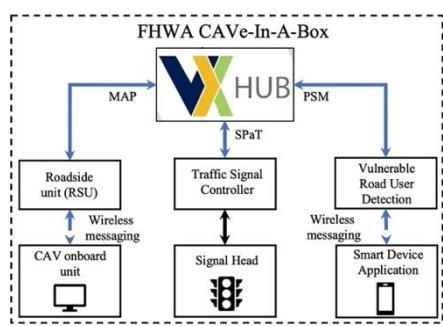


Figure 2. V2X Hub messaging architecture.

The V2X Hub software has been designed with plug-ins to support pedestrian detection and PSM forwarding. However, the software currently supports only a FLIR thermal camera. In this research, an open network video interface forum (ONVIF) supported device – Bosch Autodome 7000i Inteox camera is selected to increase the spectrum of supported cameras and further leverage advanced features such as intelligent video analytics (IVA).

Table 1. IP address configuration of devices

Device	IP Address
V2X Hub computer	192.168.0.146
RSU	192.168.0.40
OBUs	192.168.0.41
Signal controller	192.168.0.90
Camera	192.168.0.88
Connect:ITS	192.168.0.67

Development to include encoding of the SDSM in V2X Hub is currently ongoing. With this additional functionality, V2X Hub will be able to receive pedestrian detection information from various sensors and encode them into an SDSM for broadcast by RSUs. The SDSM will increase the efficiency of the V2X environment, by eliminating the need for separate pseudo-PSMs and pseudo BSMs to be generated by infrastructure and vehicles for each user that is detected and classified with their respective sensors. In other words, SDSM allows vehicles and infrastructure systems to package all detected pedestrians and vehicles into one single message to be broadcast.

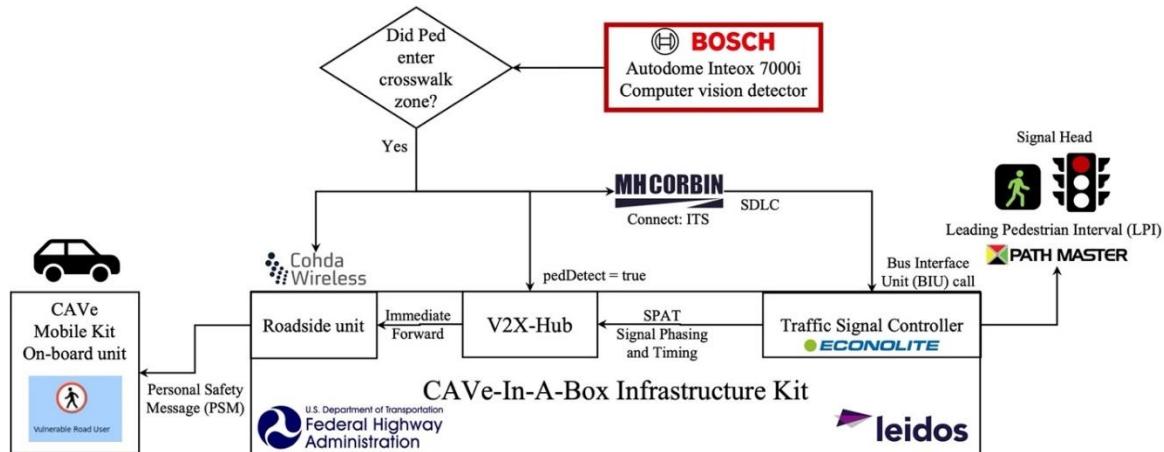


Figure 3. Real-time VRU safety alert architecture with CAVe-In-A-Box.

Object detection and ONVIF meta data. As shown in Figure 3, a crosswalk zone is created as a ‘object in field 1’ task in the Bosch configuration manager. When an object is detected in the field of interest, ONVIF metadata stream includes object id, object type (e.g., person/vehicle/bike), speed, location (latitude/longitude), and direction. Streamed metadata is processed and stored for use with downstream data analysis and V2X applications.

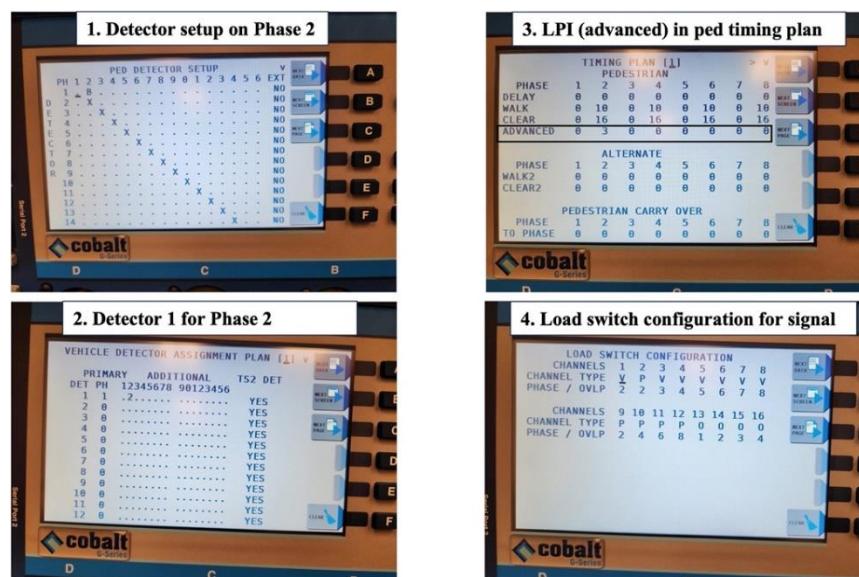


Figure 4. LPI programming on Econolite cobalt signal controller with ped detector calls.

V2X PSM messages. The object information must be parsed into a PSM and broadcasted to an RSU and OBU. In this research, MH Corbin Connect:ITS is utilized to parse the metadata to stream PSM messages to the RSU. Note that the V2X Hub software also has a plug-in to enable ‘immediate forward’ and stream PSM messages from vision-based detectors.

Leading pedestrian interval. Furthermore, Connect:ITS has a ‘stop bar’ application which allows for the detection of objects to set bus interface unit (BIU) calls via SDLC port to the signal controller. Considering the safety benefit of increased VRU safety at intersections, an LPI has been programmed in the signal controller (phase 2). Using the Connect:ITS BIU call, LPI will be latched in the corresponding phase which provides a 3 second head start to pedestrians detected at the intersection. The programming of LPI on a Econolite signal controller is depicted in Figure 4.

Outreach. To demonstrate real-world benefits of real-time computer vision-based pedestrian detection, LPI, and V2X messaging, several demonstrations have been conducted at 5-12 grade student summer camps and STEM activities. As shown in Figure 5, students are requested to volunteer as pedestrians walking on a crosswalk to understand technology operations. Demonstrations included explanations to important intersection traffic safety concerns and how LPI can benefit pedestrians when it is latched on a signal phase using real-time computer vision-based detection. Furthermore, V2X messaging such as broadcasting PSM messages to RSUs and OBUs has been demonstrated to provide knowledge on smart mobility advancements.

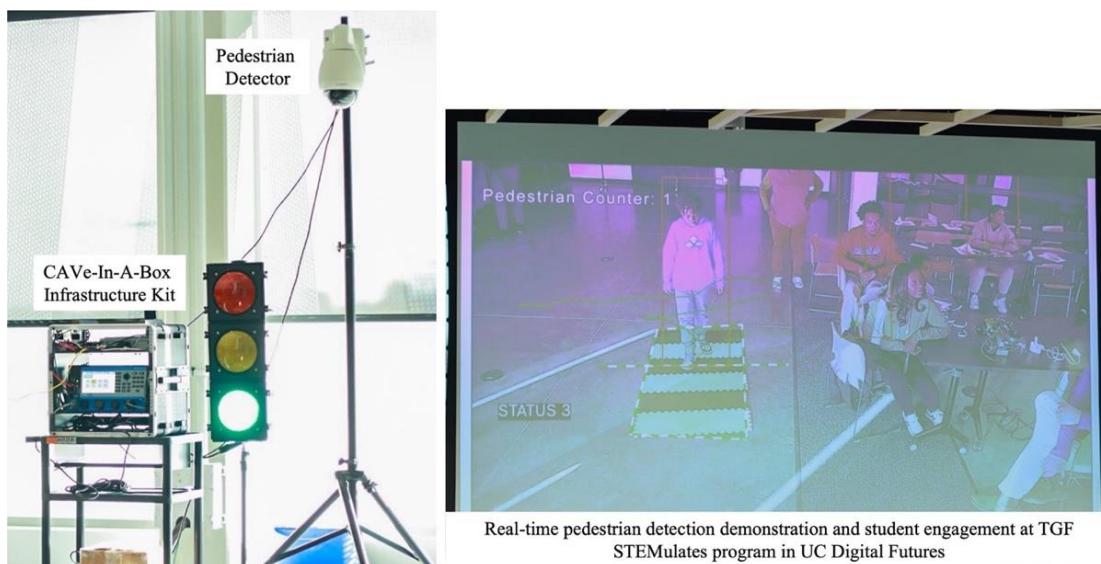


Figure 5. CAVe-In-A-Box Infrastructure kit demonstration and young learner engagement through VRU detection and intersection safety use case.

SENSOR DATA SHARING MESSAGE (SDSM)

SAE J3224. The Federal Communications Commission (FCC 2021) has revised rules to repurpose the lower 45 megahertz of the 5.850–5.925 GHz band (5.9 GHz band) for the expansion of unlicensed mid band spectrum operations, while retaining the upper 30 megahertz of spectrum in the 5.9 GHz band for ITS operations. As such, the bandwidth required to transmit V2X sensor data has to be minimized for efficient use of the allocated spectrum operations.

The SAE J3224 standard (SAE International 2022) for V2X sensor-sharing for cooperative and automated driving supports innovative use cases such as detection of an unequipped VRU by a RSU or host vehicle. The J3224 standard enables direct interface of RSUs with VRU detection technologies such as traffic cameras as shown in Figure 6. Furthermore, SDSM messages offer transmitting information on multiple detected objects (e.g., several pedestrians at a crosswalk). The SDSM also enables the use of much lower bandwidth required for audio and visual alerts (Figure 6) to drivers of approaching VRUs at the upcoming intersection.

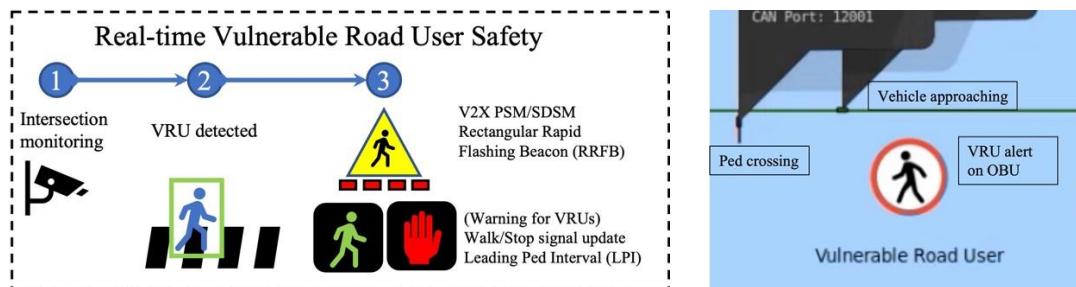


Figure 6. Real-time VRU detection and SDSM VRU alert (in Cohda Wireless VSIM).

Test setup. To transmit V2X messages via C-V2X radio, a GNSS fix is required. Hence, the setup required access to GNSS satellites for time synchronization across the RSU and OBU in this lab test. As shown in Figure 7, CAVe-In-A-Box infrastructure and mobile kits are connected to GPS antennas and the RSU is utilized to simulate a walking pedestrian at an intersection.

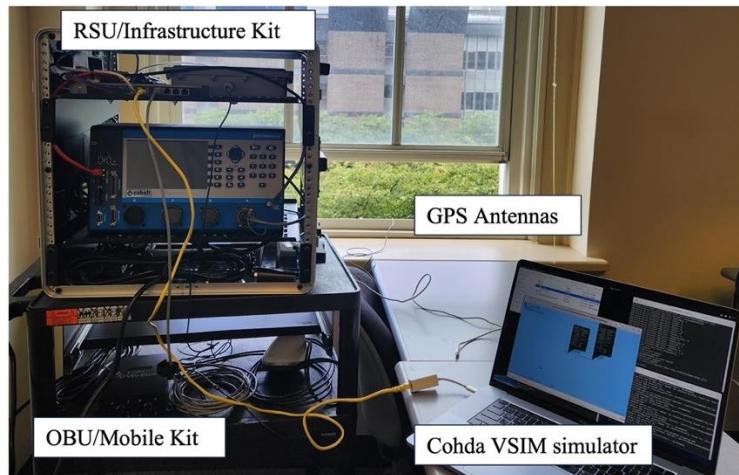


Figure 7. CAVe-In-A-Box V2X messaging test and simulation setup.

CONCLUSIONS

In this study, a Cohda Wireless MK6c C-V2X RSU is integrated with the CAVe-In-A-Box infrastructure kit and the Bosch Autodome 7000i Inteox camera to broadcast SAE J2735 PSM messages to nearby vehicle OBUs when VRUs are detected. Furthermore, V2X messages with the SAE J3224 (SDSM) standard have been tested and verified with lab tests. In addition, the CAVe-In-A-Box integrated with computer vision VRU detector final assembly is tested and

verified to detect VRUs (pedestrian and bike-pedestrians) and send a BIU call to the signal controller which is programmed with LPI.

First, PSM and SDSM V2X messages provide real-time safety benefits particularly in vehicle-VRU conflicts at intersections. Second, the LPI has been proven to significantly improve VRU safety. Future mobility, in terms of connected vehicles, will naturally benefit from this research project's outcomes. Combining V2X messages and infrastructure-level signal control timing changes will advance VRU safety at intersections.

ACKNOWLEDGEMENT

These preliminary results are shared for publication during an active research project sponsored by ODOT Office of Research. Authors thank ODOT Office of Research and the technical advisory committee for their support to this research project. Authors acknowledge technical and equipment support provided by Bosch, Cohda Wireless, Path Master, Inc., MH Corbin, and FHWA Saxton Lab.

REFERENCES

- Eskandarian, A., C. Wu, and C. Sun. (2021). "Research Advances and Challenges of Autonomous and Connected Ground Vehicles," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 683-711, Feb. 2021.
- Bonthu, et al. (2024). "Workforce Development Projects with FHWA CAVe-In-A-Box and CAVe-Lite: Lessons Learned" to be presented at *TRB Annual Meeting 2024*, Washington D.C.
- Damaj, I. W., D. K. Serhal, L. A. Hamandi, R. N. Zantout, and H. T. Mouftah. (2021). "Connected and Autonomous Electric Vehicles: Quality of Experience survey and taxonomy", *Vehicular Communications*, 28, 100312.
- Braunstein, D. (1971). "Autonomous Vehicle Speed Calibration," U.S. Patent US 2017/0008521 A1.
- Dehman, A., and B. Farooq. (2021). "Are work zones and connected automated vehicles ready for a harmonious coexistence? A scoping review and research agenda," *Transportation Research Part C: Emerging Technologies*, 133, 103422.
- FCC. (2021). "Use of the 5.850–5.925 GHz Band," Federal Register/Vol. 86, No. 83/Monday, May 3, 2021/Rules and Regulations.
- Ahmad, I., R. M. Noor, I. Ahmedy, S. Shah, I. Yaqoob, E. Ahmed, and M. Imran. (2018). "VANET-LTE based heterogeneous vehicular clustering for driving assistance and route planning applications", *Computer Networks* 145 (2018).
- Higashiyama, K., K. Kimura, H. Babakarkhail, and K. Sato. (2020). "Safety and Efficiency of Intersections with Mix of Connected and Non-Connected Vehicles," in *IEEE Open Journal of Intelligent Transportation Systems*, vol. 1, pp. 29-34.
- NHTSA (National Highway Traffic Safety Administration). *Early estimates of motor vehicle traffic fatalities and fatality rate by sub-categories in 2021*", May 2022.
- Khekare, P., S. Bonthu, V. Hunt, A. Helmicki, and K. Lee. "A Case Study on Multilane Roundabout Capacity Evaluation using Computer Vision and Deep Learning," in *ASCE Journal on Computing Systems*, vol. 36, no. 3, May 2022.
- SAE International. (2022). "V2X Sensor-Sharing for Cooperative and Automated Driving," SAE J3224.

- Sheehan, B., F. Murphy, M. Mullins, and C. Ryan. (2019). "Connected and autonomous vehicles: A cyber-risk classification framework," *Transportation Research Part A: Policy and Practice*, 124, 523–536.
- USDOT. (2020). *Ensuring American Leadership in Automated Vehicle Technologies: Automated Vehicles 4.0*, January 8, 2020, Las Vegas, NV.
- USDOT FHWA. "Connected and Automated Vehicle education (CAVe)- in-a-box System Design Document", April 2022.
https://www.pcb.its.dot.gov/CAVE/sys_design_doc/default.aspx.
- USDOT FHWA. "Connected and Automated Vehicle Education (CAVe)-in-a-box Infrastructure Kit Assembly Instructions", April 2022.
https://www.pcb.its.dot.gov/CAVE/sys_design_doc/default.aspx.

Examining Drivers of Electric Vehicle Adoption

Apurva Pamidimukkala, Ph.D.¹; Sharareh Kermanshachi, Ph.D.²;
Jay Michael Rosenberger, Ph.D.³; and Greg Hladik, Ph.D.⁴

¹Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Penn State Univ., Middletown, PA (corresponding author). Email: svk5464@psu.edu

³Professor, Dept. of Industrial, Manufacturing, and Systems Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: jrosenbe@uta.edu

⁴Executive Director, Auxiliary Services, Univ. of Texas at Arlington, Arlington, TX. Email: hladik@uta.edu

ABSTRACT

The escalating apprehension regarding environmental issues is driving nations to adopt renewable energy technologies to mitigate the release of greenhouse emissions from vehicles. Electric vehicles (EVs) provide a practical and eco-friendly alternative that can help transition to an environmentally friendly, low-emissions transportation system while protecting the environment. This study employed the theory of planned behavior (TPB) and incorporated additional factors such as price value and moral norms to examine consumers' intention to adopt EVs. A survey was administered to prospective consumers in March 2023, and the 733 responses were analyzed, using a structural equation model (SEM), to determine which factors have the greatest impact on EV adoption. The results revealed that attitudes, perceived behavior control, subjective norms, moral norms, and price value positively and significantly influenced consumers' intentions to adopt EVs.

Keywords: Electric vehicles, adoption intention, structural equation modeling

INTRODUCTION

Transport emissions experienced a consistent annual increase of 1.7% from 1990 to 2022, surpassing growth rates in energy consumption of all other sectors (Etminani-Ghasrodasthi et al., 2022; Khan et al., 2023a; Pamidimukkala et al., 2023a). It currently accounts for approximately 25% of carbon dioxide emissions globally, and that number is projected to increase to 50% by the year 2030 (IEA 2022). Electric vehicles (EVs) are viewed as a viable and sustainable option for the worldwide predicaments of energy scarcity and environmental pollution (Ali and Naushad, 2022; Feng et al., 2022), as they have the potential to reduce carbon emissions by approximately 30-50% and enhance fuel efficiency by around 40-60% in comparison to conventional vehicles (Channamallu et al., 2023; Etminani-Ghasrodasthi et al., 2021; Khan et al., 2022; Patel et al., 2021).

The 14% global increase in sales of EVs in 2019 captured the interest of the automotive industry. Europe's sales grew by 80% and Canada's grew by 43%, while sales in China and the United States remained constant (Etminani-Ghasrodasthi et al., 2023). Other nations, such as Norway (39.5%) and the United Kingdom (1.94%), joined the trend and also purchased more

electric vehicles. In 2021, a record 6.6 million EVs were sold globally, double that of the previous year, and the market share reached nearly 10%, which was four times higher than that recorded in 2019. Based on the research of Patyal et al. (2021), the global count of EVs reached approximately 16.5 million, indicating a threefold increase from 2018.

The considerable body of scientific studies on the advantages of EVs has not resulted in the expected rate of adoption (Pamidimukkala et al., 2023b, c). This may be attributed in part to the lack of studies on the factors that affect adoption (Kumar and Alok, 2020). These variables, however, vary among nations and cultures (Vafaei-Zadeh et al., 2022) and reveal the need to consider cross-cultural differences and how they impact EV adoption. Hence, the objective of this research is to determine and examine the factors that impact an individual's decision to purchase an EV.

LITERATURE REVIEW

The theory of planned behavior (TPB) is an influential psychological work that is often used to examine consumers' behavioral characteristics, including those pertaining to purchasing (Abou-Zeid and Ben-Akiva, 2011). Previous research has employed the TPB model to examine the environmentally conscious behavior of consumers. The discourse presented herein focuses on the use of a conceptual model that investigates the consumers' inclinations to adopt EVs. Figure 1 illustrates the hypothesized model developed for this research.

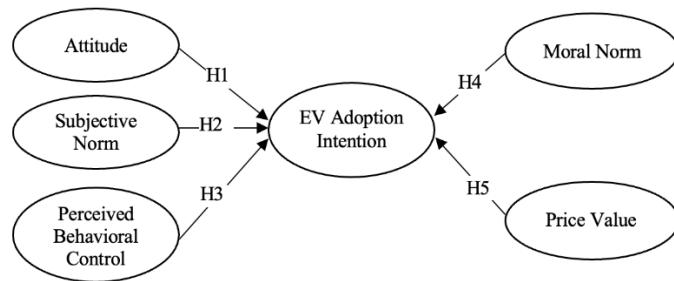


Figure 1. Hypothesized Model

Attitude

Attitude is a cognitive process that determines a person's positive or negative inclination about a particular object or concept (Sreen et al., 2018). According to the TPB, those who possess an optimistic disposition towards a particular task are more inclined to engage in the corresponding behavior (Collins et al., 2011). In other words, individuals with a positive perception of owning an EV are more inclined to purchase one. Recent studies have provided support for the assertion that there is a direct correlation between attitude and behavioral intention (Yarimoglu and Gunay, 2020). From this, the following hypothesis was developed:

Hypothesis H1: Attitude positively influences consumers' intention to adopt EVs

Subjective Norms

Subjective norms pertain to an individual's response to social influence that encourages or discourages them from performing a specific action (Collins et al., 2011). The concept of a social

norm pertains to the endorsement and promotion of a specific behavior by a particular individual or group (Shalender and Sharma, 2021). In essence, when a significant number of individuals who play a crucial role in an individual's life engage in a specific behavior, it becomes logical and emotionally compelling for that person to adopt similar actions (Patel et al., 2023a). The behavior of individuals is greatly influenced by their perception of their belief in a particular behavior, particularly when that behavior is pro-environmental and aligns with their preferences (Li et al., 2020). In this context of this paper, social norm is characterized as the consumers' perception that other esteemed individuals hold the belief that they should or should not acquire an electric vehicle. Therefore, it is assumed that

Hypothesis H2. Subjective norms positively influence consumers' intention to adopt EVs

Perceived Behavioral Control

Perceived behavioral control pertains to a person's subjective assessment of barriers they may encounter when they engage in a specific behavior. This is influenced by prior experiences or anticipated challenges, such as time limitations, convenience, and economic circumstances (Collins et al., 2011). Perceived behavioral control in the context of this paper pertains to the perceived level of difficulty or ease experienced by consumers while considering adopting an EV (Huang and Ge, 2019). The results of previous research have shown that there is a direct relationship between perceived behavioral control regarding green products and consumer adoption intention (Sreen et al., 2018; Xu et al., 2019). Therefore, it is assumed that

Hypothesis H3. Perceived behavioral control positively influences consumers' intention to adopt EVs

Moral Norm

The notion of moral norms relates to a person's sense of obligation in relation to engaging in specific types of behavior (Beck and Ajzen, 1991). The concept originates from the norm activation model and holds significant importance in the study of customers' psychological behavior (Schwartz, 1977). According to Graham-Rowe (2012), individuals who possess internal motivation and place importance on their societal responsibilities are more inclined to adopt EVs than those who do not exhibit these characteristics. According to the findings of Peters and Dutschke (2014), when there is alignment between individuals' value systems and their environmental values, they have a higher propensity to embrace EVs. The notion of moral norms is inherently intrinsic and distinct from subjective norms, as it lacks external influences and is solely guided by moral values and principles. Previous studies have found that individuals who possess a strong personal norm exhibit more favorable intentions towards the adoption of EVs compared to those who do not possess such a norm (Jansson et al., 2017; Patel et al., 2023b). From this, the following hypothesis was derived:

Hypothesis H4. Moral norm positively influences consumers' intention to adopt EVs

Price Value

The concept of price value pertains to the extent to which consumers derive utility from their expectations regarding cost management (Sweeney and Soutar, 2001). The cost-benefit relationship, as elucidated by Venkatesh et al. (2012) plays a crucial task in influencing a

person's behavioral intention to adopt new product and is contingent upon perceived benefits and the associated expenses incurred from its acquisition. Consumers form a perception of favorable value when the benefits derived from a particular product or service outweigh the associated costs (Pamidimukkala et al., 2023d). The current study ascribes significance to the advantages associated with the intention to acquire an EV. According to the US Department of Energy (Energy, 2020), EVs cost less and are more efficient than gasoline-fueled vehicles, making them a more economical choice in long term. According to Zhang et al. (2020) there is a positive correlation between the expected decrease in future expenses and the perceived present value. Consumers' knowledge of EVs' cost-saving benefits positively impacts their perceived value; therefore, the purchase is perceived as meaningful and suitable, and makes them inclined to purchase one. From this, the following hypothesis was derived:

Hypothesis H5. Price value positively influences consumers' intention to adopt EVs

METHODOLOGY

A survey questionnaire was designed and distributed electronically in March 2023 to potential consumers in Texas, using the online application QuestionPro to determine their perceptions of EVs. The questionnaire comprised 27 questions, which were categorized into two sections and took around 7 minutes to finish. After sending two reminder emails, 733 complete responses were received that were used for further analysis.

As illustrated in Figure 2, females comprised 52.9% of the responses, males 44.3%, those who identified as "Others" constituted 2.8%. Around 62% of the respondents possessed bachelor's degrees. Most participants (52.2%) were aged 35 years and above. A mere 17.6% of the participants belonged to households earning less than \$35,000 annually. Most of the respondents had more than one vehicle, and 77% of participants are experienced drivers.

Measures

The measurement items were developed through an extensive review of the current research on TPB, moral norms, and price value, and terminologies of some of the items were adjusted to better align with the electric vehicle context. The survey questions were in the form of five-point likert scale. The scale for attitude was taken from the study by Cheng et al. (2019); the scale for subjective norms and perceived behavioral control was taken from a study by Han et al. (2020). The scale for moral norms and price value were taken from the study of Wang et al. (2018). Finally, the three components of "EV adoption intention" were taken from He et al. (2018).

ANALYSIS AND RESULTS

First, a confirmatory factor analysis (CFA) was conducted to evaluate the measurement model. Table 1 displays the range of Cronbach's α values, which varied from 0.806 to 0.931, exceeding the threshold of 0.7 showing that the scales are reliable. The composite reliability values of the constructs were also above 0.7, suggesting that the scales exhibited strong internal consistency (Nunnally and Bernstein, 1978). Next, the soundness of the construct was examined through convergent validity to verify the degree to which the measurement items accurately represent the constructs. As illustrated in Table 1, the average variance extracted values varied

between 0.581 to 0.817, which were above the cut-off of 0.5, and the loadings of all the items were observed to be higher than the threshold value of 0.6 (Hair et al., 2015).

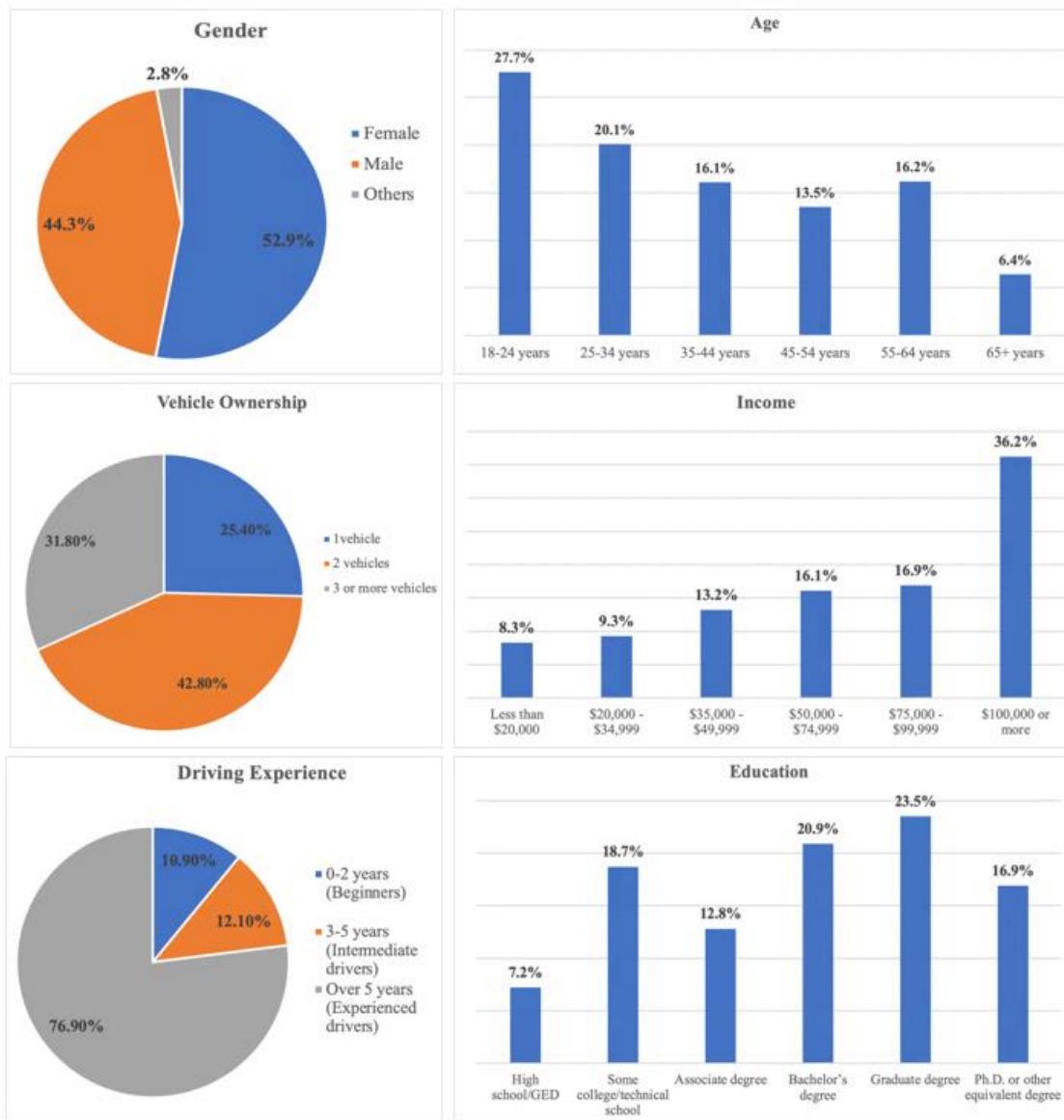


Figure 2. Demographics of Survey Respondents

Next, the structural model was assessed in AMOS v. 26.0. Before testing the hypothesis, the model fit was evaluated and was found to be acceptable. The (χ^2/df) was found to be 2.32, which is below the critical threshold of 3.0 as established by Hair et al. (2015), and the GFI, CFI, NFI, which were 0.908, 0.926, 0.960, respectively, met the minimum acceptable criteria of 0.90 as specified by Hair et al. (2015). The model fit indices, namely RMSEA=0.042, also exhibited values below the threshold of 0.08.

The results of the structural model were presented in Table 2. The hypothesis results revealed that attitude, subjective norms, perceived behavioral control, moral norms, and price value affect consumers' propensity to embrace EVs.

Table 1. Measurement Model

Construct	Item	Loadings	Cronbach's alpha	CR	AVE
Attitude (AT)	AT1	0.752	0.802	0.806	0.581
	AT2	0.823			
	AT3	0.708			
Subjective Norm (SN)	SN1	0.775	0.853	0.861	0.674
	SN2	0.860			
	SN3	0.826			
Perceived Behavioral Control (PBC)	PBC1	0.702	0.810	0.815	0.614
	PBC2	0.786			
	PBC3	0.822			
Moral Norms (MN)	MN1	0.867	0.874	0.876	0.702
	MN2	0.854			
	MN3	0.791			
Price Value (PV)	PV1	0.889	0.865	0.879	0.708
	PV2	0.838			
	PV3	0.795			
EV Adoption Intention (EVAI)	EVAI1	0.917	0.926	0.931	0.817
	EVAI2	0.902			
	EVAI3	0.894			

Table 2. Structural Model

Hypothesis	Path	Standard Coefficient (β)	Result
H1	AT → EVAI	0.426**	Supported
H2	SN → EVAI	0.164**	Supported
H3	PBC → EVAI	0.357**	Supported
H4	MN → EVAI	0.228**	Supported
H5	PV → EVAI	0.341**	Supported

Note: ** Coefficient is significant at 0.01

DISCUSSION AND CONCLUSION

This research utilized an extended TPB framework to examine the determinants that affect individual's inclination to embrace EVs. Data from 733 potential consumers was collected and analyzed, and the results revealed that attitudes, perceived behavioral control, moral norms, subjective norms, and price value exerted a positive and substantial influence. Table 2 summarizes the findings of the study and makes it clear that attitude, which has the highest path coefficient, has the largest influence on the inclination to adopt an EV. These results support the outcomes observed in the majority of studies on the TPB conducted in emerging markets and

suggest that individuals who hold a favorable disposition towards the utilization and acquisition of EVs are also more inclined to embrace their adoption (Alzahrani et al., 2019; Jaiswal et al., 2021). The results also revealed a positive correlation between subjective norms and the intention to embrace EVs, indicating that individuals' decisions about whether or not to adopt electric vehicles are influenced by the opinions of those others that they hold in high esteem.

The study revealed that perceived behavioral control had the second highest level of statistical significance in positively influencing the desire to adopt EVs, as consumers are more disposed to adopt EVs when they are self-confident in their purchasing decisions. The obtained outcome aligns with the initial TPB model, suggesting that perceived behavior control plays a significant role as a precursor to purchase intention. These results support the previous literature conducted in different contexts (Vafae-Zadeh et al., 2022), that indicated that perceived behavior control significantly influences pro-environmental consumer behavioral intention. If consumers perceive EVs as being convenient to possess, repair, and maintain, it is probable that they will be more inclined to adopt them.

Price value was found to be the third most influential factor in the model. Venkatesh et al. (2012) posits that the concept of price value pertains to the equilibrium between the perceived advantages of adopting novel technology and the monetary investment required to acquire it. This dynamic interplay between cost and benefit significantly influences an individual's inclination to engage with new technology, and any lack of sensitivity towards the cost can be attributed to a heightened awareness of the numerous benefits and intrinsic value associated with the technology.

These research findings have significant implications for those seeking to promote consumer adoption of EVs. It is suggested that governments and automakers give greater priority to enhancing consumer awareness of the user-friendliness and essential attributes of electric vehicles, which has the potential to positively influence their attitudes towards the vehicles and thus increase the likelihood that they will purchase one. Educational initiatives and media campaigns that highlight the environmental benefits of EVs are proven effective methods for accomplishing this. Similarly, car manufacturers can employ experiential marketing strategies, such as offering test drives at auto shows, to provide consumers with an opportunity to personally experience the benefits of owning and driving an EV.

This study will also serve as valuable information for electric vehicle manufacturers as they refine their products and marketing strategies, as it presents an overview of the various elements that can impact consumers' propensity to acquire an EV. The findings indicate that price value is the third most influencing factor in predicting an individual's intention to adopt an EV, as consumers compare the perceived value and benefits of owning an EV with the associated costs. As an example, automotive manufacturers have the capacity to improve the perceived pricing value for consumers by improving performance, ensuring reliability, enhancing fuel efficiency, and incorporating ecologically friendly features.

REFERENCES

- Ali, I., and M. Naushad. (2022). "A study to investigate what tempts consumers to adopt electric vehicles." *World Electric Vehicle Journal* 13(2): 26.
- Alzahrani, K., A. Hall-Phillips, and A. Z. Zeng. (2019). "Applying the theory of reasoned action to understanding consumers' intention to adopt hybrid electric vehicles in Saudi Arabia." *Transportation* 46: 199-215.

- Beck, L., and I. Ajzen. (1991). "Predicting dishonest actions using the theory of planned behavior." *Journal of research in personality* 25(3): 285-301.
- Channamallu, S. S., S. Kermanshachi, and A. Pamidimukkala. (2023). "Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles." In *International Conference on Transportation and Development 2023* (pp. 39-50), Reston, VA: ASCE.
- Cheng, P., Z. OuYang, and Y. Liu. (2019). "Understanding bike sharing use over time by employing extended technology continuance theory." *Transportation research part A: policy and practice* 124: 433-443.
- Collins, S. E., K. Witkiewitz, and M. E. Larimer. (2011). "The theory of planned behavior as a predictor of growth in risky college drinking." *Journal of studies on alcohol and drugs* 72(2): 322-332.
- Etminani-Ghasrodashti, R., R. K. Patel, S. Kermanshachi, J. M. Rosenberger, D. Weinreich, and A. Foss. (2021). "Integration of shared autonomous vehicles (SAVs) into existing transportation services: A focus group study." *Transportation Research Interdisciplinary Perspectives*, 12, p.100481.
- Etminani-Ghasrodashti, R., G. Hladik, S. Kermanshachi, J. M. Rosenberger, M. Arif Khan, and A. Foss. (2022). "Exploring shared travel behavior of university students." *Transportation Planning and Technology*, 46(1), pp. 22-44.
- Etminani-Ghasrodashti, R., M. Khan, R. K. Patel, S. Kermanshachi, J. M. Rosenberger, A. Pamidimukkala, G. Hladik, and A. Foss. (2023). "Measuring students' satisfaction levels for transit services: An application of latent class analysis." *International Journal of Transportation Science and Technology*. (In Press).
- Office of Energy Efficiency and Renewable Energy. (2020). *Saving on Fuel and vehicle Costs*. Office of Energy Efficiency and Renewable Energy.
- Feng, L., K. Liu, J. Wang, K.-Y. Lin, K. Zhang, and L. Zhang. (2022). "Identifying Promising Technologies of Electric Vehicles from the Perspective of Market and Technical Attributes." *Energies* 15(20): 7617.
- Graham-Rowe, E., B. Gardner, C. Abraham, S. Skippon, H. Dittmar, R. Hutchins, and J. Stannard. (2012). "Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: A qualitative analysis of responses and evaluations." *Transportation Research Part A: Policy and Practice* 46(1): 140-153.
- He, X., W. Zhan, and Y. Hu. (2018). "Consumer purchase intention of electric vehicles in China: The roles of perception and personality." *Journal of Cleaner Production* 204: 1060-1069.
- Huang, X., and J. Ge. (2019). "Electric vehicle development in Beijing: An analysis of consumer purchase intention." *Journal of cleaner production* 216: 361-372.
- Jaiswal, D., V. Kaushal, R. Kant, and P. K. Singh. (2021). "Consumer adoption intention for electric vehicles: Insights and evidence from Indian sustainable transportation." *Technological Forecasting and Social Change* 173: 121089.
- Jansson, J., A. Nordlund, and K. Westin. (2017). "Examining drivers of sustainable consumption: The influence of norms and opinion leadership on electric vehicle adoption in Sweden." *Journal of Cleaner Production* 154: 176-187.
- Khan, M. A., R. Etminani-Ghasrodashti, A. Shahmoradi, S. Kermanshachi, J. M. Rosenberger, and A. Foss. (2022). "Integrating shared autonomous vehicles into existing transportation services: evidence from a paratransit service in Arlington, Texas." *International Journal of Civil Engineering*, 20(6), pp.601-618.

- Khan, M. A., R. Etminani-Ghasroddashti, A. Pamidimukkala, S. Kermanshachi, and J. M. Rosenberger. (2023a). "Impacts of on-demand ride services on the number of traffic crashes—A case study of RideAustin in Austin, TX." *Transportation Research Interdisciplinary Perspectives*, 22, p.100966.
- Khan, M. A., R. K. Patel, A. Pamidimukkala, S. Kermanshachi, J. M. Rosenberger, G. Hladik, and A. Foss. (2023b). "Factors that determine a university community's satisfaction levels with public transit services." *Frontiers in Built Environment*, 9, p.1125149.
- Kumar, R. R., and K. Alok. (2020). "Adoption of electric vehicle: A literature review and prospects for sustainability." *Journal of Cleaner Production* 253: 119911.
- Li, L., Z. Wang, L. Chen, and Z. Wang. (2020). "Consumer preferences for battery electric vehicles: A choice experimental survey in China." *Transportation Research Part D: Transport and Environment* 78: 102185.
- Pamidimukkala, A., S. Kermanshachi, J. M. Rosenberger, and G. Hladik. (2023a). "Evaluation of barriers to electric vehicle adoption: A study of technological, environmental, financial, and infrastructure factors." *Transportation Research Interdisciplinary Perspectives*, 22, p.100962.
- Pamidimukkala, A., S. Kermanshachi, J. M. Rosenberger, and G. Hladik. (2023b). "Barriers and Motivators to the Adoption of Electric Vehicles: A Global Review." *Journal of Green Energy and Intelligent Transportation*, p.100153.
- Pamidimukkala, A., R. K. Patel, S. Kermanshachi, J. M. Rosenberger, and S. Tanvir. (2023c). "A Review on Shared Mobility and Electric Vehicles." In *International Conference on Transportation and Development 2023* (pp. 333-342), Reston, VA: ASCE.
- Pamidimukkala, A., S. Kermanshachi, and R. Patel. (2023d). "An Exploratory Analysis of Crashes Involving Autonomous Vehicles." In *International Conference on Transportation and Development 2023* (pp. 343-350), Reston, VA: ASCE.
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. M. Rosenberger, and D. Weinreich. (2021). "Exploring preferences towards integrating the autonomous vehicles with the current microtransit services: A disability focus group study." In *International Conference on Transportation and Development 2021* (pp. 355-366). Reston, VA: ASCE.
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. Rosenberger, and A. Foss. (2022). "How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems*." *ASCE International Conference on Transportation & Development*, pp. 81-93, Reston, VA: ASCE.
- Patel, R. K., S. S. Channamallu, M. A. Khan, S. Kermanshachi, and A. Pamidimukkala. (2023a). "An Exploratory Analysis of Temporal and Spatial Patterns of Autonomous Vehicle Collisions." *Public Works Management & Policy*, 0(0).
- Patel, R. K., R. Etminani-Ghasroddashti, S. Kermanshachi, J. M. Rosenberger, A. Pamidimukkala, and A. Foss. (2023b). "Identifying individuals' perceptions, attitudes, preferences, and concerns of shared autonomous vehicles: During-and post-implementation evidence." *Transportation Research Interdisciplinary Perspectives*, 18, p.100785.
- Patyal, V. S., R. Kumar, and S. Kushwah. (2021). "Modeling barriers to the adoption of electric vehicles: An Indian perspective." *Energy* 237: 121554.
- Peters, A., and E. Dütschke. (2014). "How do consumers perceive electric vehicles? A comparison of German consumer groups." *Journal of Environmental Policy & Planning* 16(3): 359-377.

- Schwartz, S. H. (1977). Normative influences on altruism. *Advances in experimental social psychology*, Elsevier. 10: 221-279.
- Shalender, K., and N. Sharma. (2021). "Using extended theory of planned behaviour (TPB) to predict adoption intention of electric vehicles in India." *Environment, Development and Sustainability* 23(1): 665-681.
- Sreen, N., S. Purbey, and P. Sadarangani. (2018). "Impact of culture, behavior and gender on green purchase intention." *Journal of retailing and consumer services* 41: 177-189.
- Sweeney, J. C., and G. N. Soutar. (2001). "Consumer perceived value: The development of a multiple item scale." *Journal of retailing* 77(2): 203-220.
- Vafaei-Zadeh, A., T.-K. Wong, H. Hanifah, A. P. Teoh, and K. Nawaser. (2022). "Modelling electric vehicle purchase intention among generation Y consumers in Malaysia." *Research in Transportation Business & Management* 43: 100784.
- Venkatesh, V., J. Y. Thong, and X. Xu. (2012). "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology." *MIS quarterly*: 157-178.
- Wang, S., J. Wang, J. Li, J. Wang, and L. Liang. (2018). "Policy implications for promoting the adoption of electric vehicles: Do consumer's knowledge, perceived risk and financial incentive policy matter?" *Transportation Research Part A: Policy and Practice* 117: 58-69.
- Xu, Y., W. Zhang, H. Bao, S. Zhang, and Y. Xiang. (2019). "A SEM-neural network approach to predict customers' intention to purchase battery electric vehicles in China's Zhejiang province." *Sustainability* 11(11): 3164.
- Yarimoglu, E., and T. Gunay. (2020). "The extended theory of planned behavior in Turkish customers' intentions to visit green hotels." *Business Strategy and the Environment* 29(3): 1097-1108.
- Zhang, Y., C. Xiao, and G. Zhou. (2020). "Willingness to pay a price premium for energy-saving appliances: Role of perceived value and energy efficiency labeling." *Journal of Cleaner Production* 242: 118555.

Real-Time Traffic Control and Safety Measure Analysis Using LiDAR Sensor during Traffic Signal Failures

A. Ansariyar, Ph.D.¹; and M. Jeihani, Ph.D.²

¹SABA Laboratory, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: alans2@morgan.edu

²SABA Laboratory, Dept. of Transportation and Urban Infrastructure, Morgan State Univ., Baltimore, MD. Email: mansoureh.jeihani@morgan.edu

ABSTRACT

This study investigates the impact of traffic signal failures on traffic flow, safety, and congestion at a signalized intersection with high rate of crashes in Baltimore City. Analyzing data from a malfunctioning signal, collected through a LiDAR sensor during a temporary outage, the research evaluates vehicle and pedestrian counts, V2V and V2P conflicts, and jaywalking events. Findings revealed that malfunctioning signals disrupt traffic flow and safety, emphasizing the necessity of adaptive signal systems. The analysis underscores the role of signal optimization in reducing congestion and improving traffic flow, promoting sustainability in urban transportation. Ultimately, this study emphasizes the crucial role of traffic signals in managing congestion and highlights the need for innovative, evidence-based strategies in urban planning to create more efficient and sustainable transportation systems. Policymakers and transportation authorities can use the insights gained to develop improved traffic management solutions to improve traffic safety at signalized intersections.

Keywords: LiDAR Sensor Technology, Signalized Intersections, Traffic Signal Failure, Traffic Safety, Congestion, Environmental pollution

INTRODUCTION

Traffic signals are fundamental components of urban transportation systems, crucial for regulating traffic flow and ensuring the safe and efficient movement of vehicles, pedestrians, and cyclists. However, like any mechanical or electrical system, traffic signals can experience malfunctions or stop working altogether due to various factors such as power outages, technical failures, or crashes. The effect of traffic signal failures on managing traffic volume and reducing congestion and delays at signalized intersections is a matter of paramount concern for traffic engineers, city planners, and commuters alike.

When a traffic signal ceases to function, the intersection reverts to an uncontrolled or "all-way stop" situation, where vehicles from all directions are required to yield the right-of-way. This sudden change in traffic control dynamics can have significant consequences, including increased congestion, delays, and potential safety hazards. Understanding the challenges posed by traffic signal failures and exploring strategies to mitigate their impact is essential for maintaining the smooth operation of urban transportation networks.

This study delves into the critical issue of how traffic signal failures affect traffic volume management, congestion levels, and delay mitigation at signalized intersections. It seeks to comprehensively analyze the dynamics of traffic behavior when signals malfunction and to

identify potential solutions and best practices for minimizing disruptions and ensuring safe traffic flow under such circumstances. The research not only examines the immediate consequences of signal failures but also considers the broader implications for urban mobility and transportation resilience. It is essential to recognize that traffic signal failures can occur due to various reasons, including natural disasters, crashes, equipment malfunctions, and power outages. Consequently, addressing the challenges associated with signal breakdowns is not only about responding to isolated crashes but also about developing proactive strategies to enhance the reliability and robustness of signalized intersections.

As cities continue to grow and transportation demands increase, the potential impact of traffic signal failures becomes even more significant [1]. This study aims to shed light on the multifaceted issues surrounding traffic signal failures and to provide insights into how urban planners, traffic engineers, and policymakers can better prepare for and respond to such events. By doing so, it contributes to the broader goal of creating resilient, efficient, and safe urban transportation systems that can withstand disruptions and provide consistent service to commuters, even when the traffic signal stops working. The following sections will delve into the methodology, data analysis, findings, conclusion, and references.

METHODOLOGY

Assessing the impact of traffic signal failures on traffic volume, congestion, and delays at signalized intersections using LiDAR sensor technology can be taken into account as the key aspect of intelligent mobility. The LiDAR sensor at E Cold Spring Ln – Hillen Rd intersection was installed in April 2022 to collect real-time traffic data with an efficient accuracy. LiDAR sensors can capture high-resolution, three-dimensional data of the intersection and its surroundings. The machine learning algorithms were utilized as explained in author's previous studies [2, 3, 5, 7, 9, 11, 14] to clean and preprocess LiDAR data to remove noise and outliers, and align it with signal failure timestamps.

The study focused on the traffic signal failures and investigating the driver's behavior at one of signalized intersections with significant frequency of crashes by using the recorded data by the LiDAR sensor technology specifically focusing on the absence of signal-related light patterns (e.g., red, green, yellow). By analyzing vehicle trajectories, speeds, and lane changes during signal failure periods to understand how traffic adapts to the absence of signal control, the findings of this research provide valuable insights regarding the importance of traffic signals. Additionally, the LiDAR data was analyzed to identify congestion patterns, such as stop-and-go traffic, queuing, and lane blockages during signal failures.

The research provides a baseline comparison. Hereupon, the traffic flow, congestion levels, and delays are compared during signal failure periods with data from the same intersection under normal signal operation. In order to examine the safety considerations, the LiDAR data was analyzed to identify potential safety incidents during signal failures, such as near-miss events or crashes. Also, the impact of signal failures was assessed on pedestrian and cyclist safety and behavior at signalized intersections. Finally, the key findings of the study are summarized, emphasizing the importance of understanding and mitigating the effects of traffic signal failures on traffic volume, congestion, and delays at signalized intersections using LiDAR sensor technology.

DATA ANALYSIS

This study concentrates on E Cold Spring Ln – Hillen Rd intersection in Baltimore City, Maryland. The LiDAR data from September, 12th to September, 13th 2023 was analyzed to provide a practical accurate understanding of the intersection safety conditions [2, 3]. Figure 1 shows the intersection's location. The installation location of the LiDAR sensor is shown by the black circle.

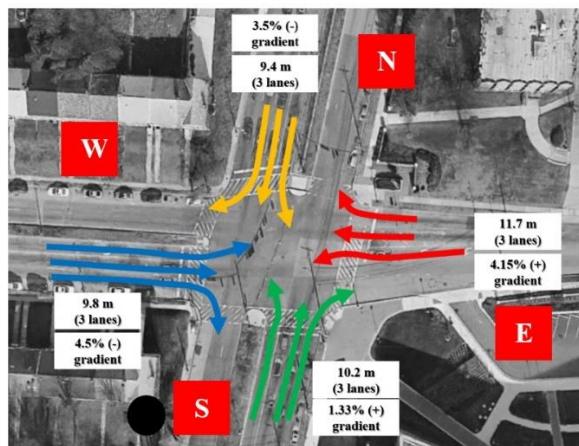


Figure 1. E Cold Spring Ln – Hillen Rd Intersection

The Research analyzed the traffic data from 00:00 AM in September 12th to 11:00 AM in September 13th. The signal controller stopped working from 05:30 AM to 10:23 AM on September 13th. As a result of improper traffic signal performance for four hours and fifty-three minutes, the research examined the traffic condition at the intersection in a chaos condition (traffic signal failure). Table 1 shows the vehicle counts before, during, and after the chaos condition. During the chaos, 10612 vehicles passed through the intersection. Also, Table 2 illustrates the pedestrian counts before, during, and after the chaos condition.

Table 1. Vehicle Counts Collected Data by The LiDAR Sensor

<u>Date</u>	<u>Hour</u>	<u>NE</u>	<u>NS</u>	<u>NW</u>	<u>ES</u>	<u>EW</u>	<u>EN</u>	<u>SW</u>	<u>SN</u>	<u>SE</u>	<u>WN</u>	<u>WE</u>	<u>WS</u>	<u>SUM</u>
	00:00 - 01:00 AM	22	59	20	8	43	31	6	82	9	55	63	1	399
	01:00 - 02:00 AM	13	36	10	7	34	20	3	72	5	22	32	0	254
	02:00 - 03:00 AM	7	52	9	3	16	5	1	40	2	11	16	0	162
	03:00 - 04:00 AM	8	67	3	3	18	7	4	45	4	18	22	0	199
	04:00 - 05:00 AM	9	119	25	12	32	9	0	47	1	17	28	0	299
	05:00 - 05:30 AM	19	219	20	11	43	12	3	62	3	26	38	0	457
September 13th	05:30 - 06:00 AM	12	146	13	8	29	8	2	42	2	18	25	0	304
	06:00 - 07:00 AM	75	720	83	42	177	67	16	213	19	75	135	7	1629
	07:00 - 08:00 AM	76	1129	214	101	412	166	25	498	34	118	271	4	3048
	08:00 - 09:00 AM	114	991	213	109	309	165	30	516	83	147	286	8	2971
	09:00 - 10:00 AM	113	655	170	67	209	114	29	471	46	167	222	1	2264
	10:00 - 10:23 AM	30	100	19	20	36	22	6	92	17	29	23	2	396
	10:23 - 11:00 AM	44	150	28	31	55	34	8	137	26	44	35	3	595

Table 2. Pedestrian Counts Collected Data by The LiDAR Sensor

<u>Date</u>	<u>Hour</u>	<u>N</u>	<u>E</u>	<u>S</u>	<u>W</u>	<u>SUM</u>
September 13th	00:00 - 01:00 AM	2	2	2	1	7
	01:00 - 02:00 AM	3	3	0	0	6
	02:00 - 03:00 AM	2	0	0	0	2
	03:00 - 04:00 AM	1	0	0	0	1
	04:00 - 05:00 AM	0	0	0	1	1
	05:00 - 05:30 AM	0	0	2	1	3
	05:30 - 06:00 AM	0	0	1	1	2
	06:00 - 07:00 AM	2	1	6	3	12
	07:00 - 08:00 AM	15	7	5	8	35
	08:00 - 09:00 AM	40	16	32	6	94
	09:00 - 10:00 AM	55	22	36	11	124
	10:00 - 10:23 AM	11	7	8	12	38
	10:23 - 11:00 AM	16	11	12	19	58

In terms of “**vehicle-vehicle conflicts**”, Figure 2 illustrated the frequency of conflicts and Figure 3 illustrates the severity of conflicts ($\sum \frac{1}{PET}$) [5, 7, 9, 11, 14]. The LiDAR sensor collects the Post Encroachment Time (PET) as one of key surrogate safety measures. PET is a critical parameter at signalized intersections that is essential for ensuring safety and minimizing the risk of crashes. It takes into account various factors, including vehicle speed, intersection design, and driver behavior, to determine the time vehicles have to safely clear the intersection after the signal changes to yellow. PET refers to the duration of time that a vehicle has to safely clear an intersection after the traffic signal changes from green (go) to yellow (caution) and before it turns red (stop). PET is a crucial safety parameter because it determines whether a vehicle can safely complete its passage through the intersection or whether it is at risk of entering the intersection when the opposing traffic receives a green signal [9-16]. PET is a crucial safety measure in intersection management, allowing vehicles to safely clear the intersection during signal transitions, considering factors like vehicle speed, intersection dimensions, driver behavior, and signal timing. PET determination involves a complex calculation from the onset of the yellow signal to a vehicle's complete passage beyond the stop line, with factors like speed, intersection dimensions, driver vigilance, and signal timing components playing pivotal roles.

As shown in Figures 2 and 3, the vehicle-vehicle conflicts during the chaos condition are more frequent and severe. A total of 632 conflicts with severity of 222.8 were collected during the chaos condition. The LiDAR captured 157 conflicts with an average severity of 56.7 during the same interval as the chaos condition on September 12th (the normal condition). The results highlighted that the frequency and severity of vehicle-vehicle conflicts increased by 302.5% and 292.9%, respectively. In terms of “**vehicle-pedestrian conflicts**”, Figure 4 and 5 illustrate the frequency and severity of conflicts, respectively.

As can be seen in Figures 4 and 5, 65 vehicle-pedestrian conflicts with severity 19.6 collected during the chaos condition. The LiDAR captured 188 conflicts with an average severity of 66.3 during the same interval as the chaos condition on September 12th (the normal condition). The results highlighted that the frequency and severity of vehicle-pedestrian conflicts decreased by 65.4% and 70.4%, respectively. The vehicle-pedestrian conflicts results confirmed

that pedestrians tend to exhibit a more cautious behavior during chaotic situations, where they can observe motorized vehicles moving freely from various directions at the intersection. Recognizing the significance of pedestrians as Vulnerable Road Users (VRUs), the research investigation interval on September 13th was extended until 19:00 PM. In terms of “jaywalking events conflicts”, Figure 6 shows the frequency of jaywalking events during the chaos condition.

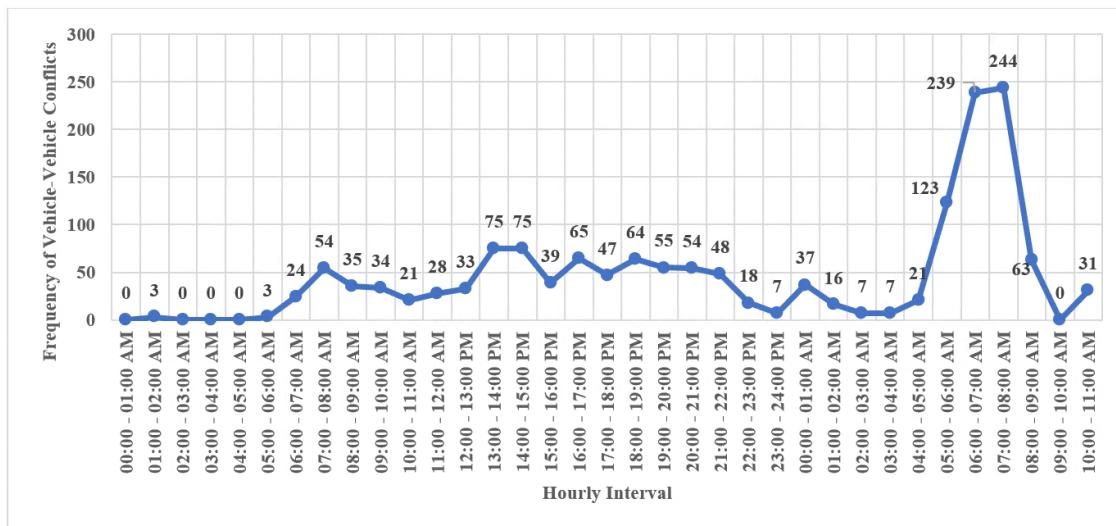


Figure 2. Frequency of Vehicle-Vehicle Conflicts

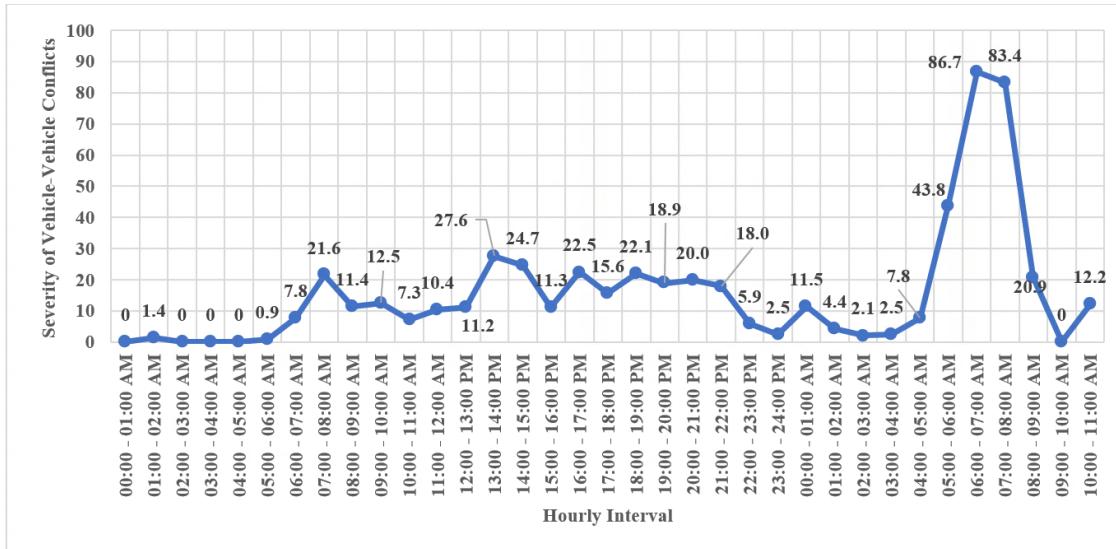


Figure 3. Severity of Vehicle-Vehicle Conflicts

Jaywalking conflicts at signalized intersections with a high rate of traffic crashes pose significant risks to pedestrian and driver safety [5]. Vehicle-pedestrian conflicts arise when pedestrian's cross streets outside designated areas, posing dangers. To address this issue, a multi-faceted approach involving education, infrastructure improvements, signal optimization, enforcement, and urban planning is necessary. Contributing factors include pedestrian behavior,

signal timing issues, deficient infrastructure, and safety perception, emphasizing the need to address these aspects for safer and more efficient traffic management at critical intersections. Legal consequences, including fines or liability, loom for both pedestrians and drivers entangled in jaywalking conflicts. Mitigation strategies encompass multifaceted approaches involving education campaigns, infrastructure improvements, optimized signal timing, enforcement measures, technological solutions, and considerations within urban planning to address and reduce jaywalking conflicts effectively. As can be seen in Figure 6, 35 jaywalking events were recorded by the LiDAR sensor during the chaos condition. In the same interval on September 12th (normal condition – from 05:30 AM to 10:23 AM), the LiDAR collected 25 events. As a result, the jaywalking events increased by 40% in comparison to the normal conditions observed on September 12th.

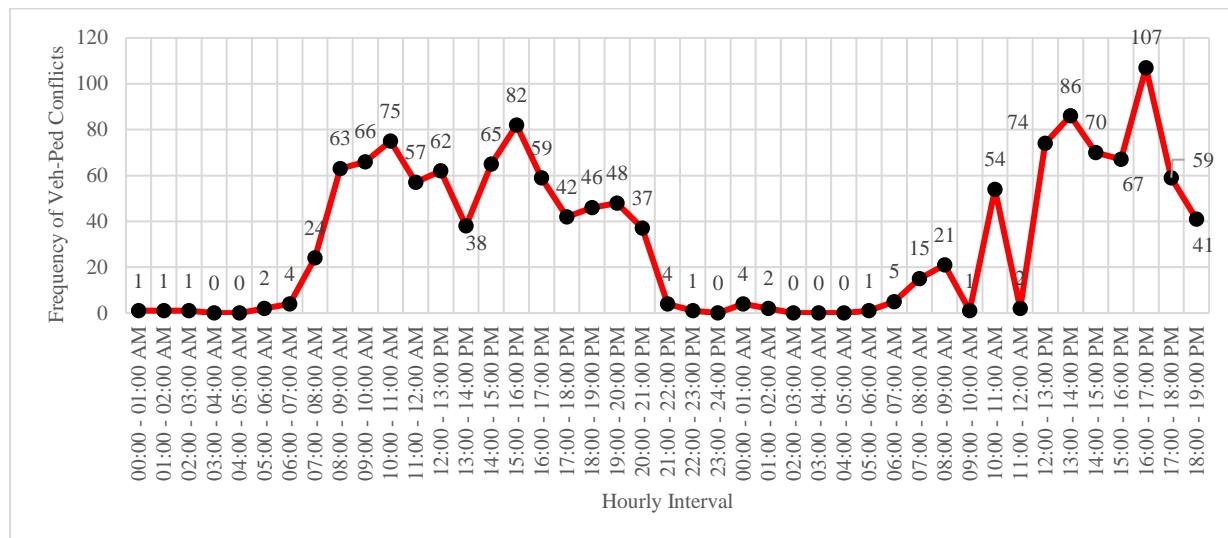


Figure 4. Frequency of Vehicle-Pedestrian Conflicts

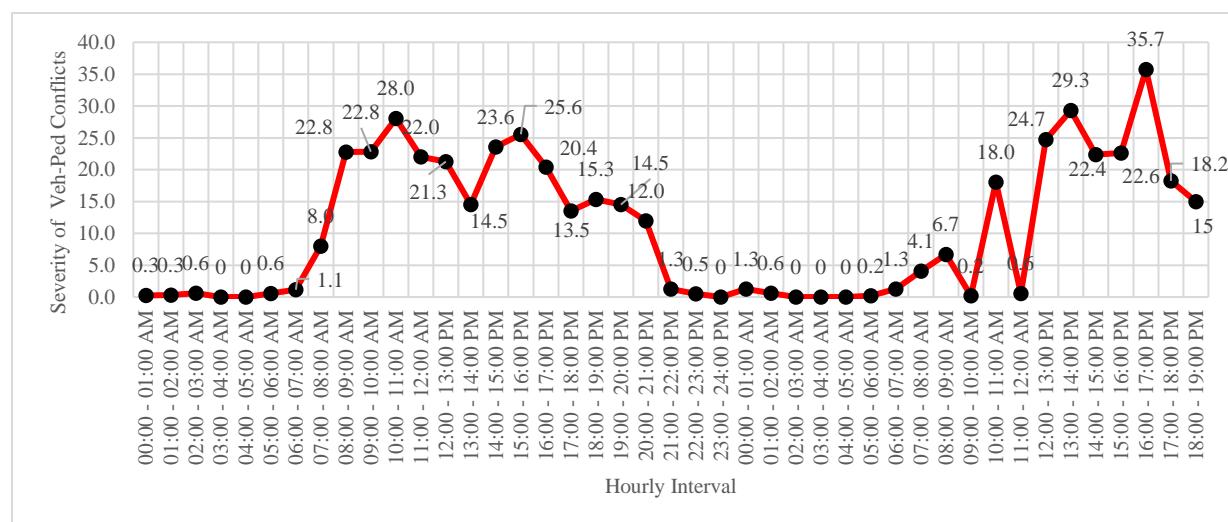


Figure 5. Severity of Vehicle-Pedestrian Conflicts

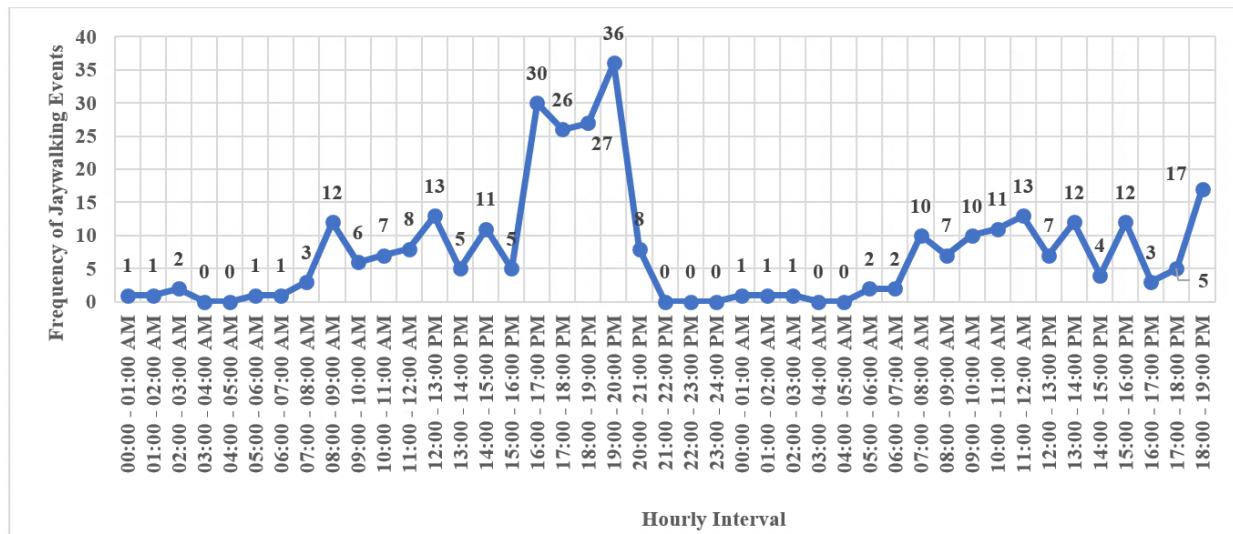


Figure 6. Jaywalking Events Analysis

FINDINGS

Basically, LiDAR sensors offer a novel and robust solution to overcome the aforementioned challenges. They operate based on the principle of emitting laser pulses and measuring their return time, allowing them to provide accurate and reliable traffic data regardless of environmental conditions. This inherent resilience makes LiDAR sensors especially well-suited for monitoring traffic dynamics at signalized intersections, where the consequences of data inaccuracies can be critical. Considering the accuracy of the installed LiDAR sensor at E Cold Spring Ln – Hillen Rd intersection [3, 7, 9, 11, 14], this section tries to compare the collected vehicle and pedestrian volumes (counts), vehicle-vehicle and vehicle-pedestrian conflicts frequency and severity, and the frequency of jaywalking events under chaos and normal conditions. Hereupon, results from 05:30 AM to 10:23 AM on Wednesday, September 13th (chaos condition) are compared with those from the same interval on Tuesday, September 12th.

- 1) **Vehicle Counts:** LiDAR sensors offer several key advantages when it comes to collecting vehicle count data during signal failures. Their ability to operate independently of traffic signal systems ensures the uninterrupted gathering of essential traffic information, enabling traffic managers to assess congestion levels, monitor traffic flow patterns, and make informed decisions even when signals are non-functional. LiDAR sensors are highly accurate in distinguishing between various types of vehicles and can provide real-time data with minimal downtime, allowing for rapid response and mitigation of congestion and safety risks. The changes in vehicle counts are shown in Figure 7. As can be seen in Figure 7, the vehicle counts in chaos condition is lower than the vehicle counts on September 12th.
- 2) **Pedestrian Counts:** LiDAR sensors offer high accuracy in distinguishing pedestrians from other objects and can provide real-time data with minimal disruption, allowing for rapid response to potential safety concerns. Moreover, their non-intrusive installation and low maintenance requirements reduce disruption to pedestrian flow, ensuring the uninterrupted movement of people even during signal system downtime. Traditional methods of pedestrian count data collection often rely on signal-controlled pedestrian

crossing systems, rendering them vulnerable to inaccuracies during signal outages. In contrast, LiDAR sensors offer a robust and dependable solution, ensuring the continuous availability of critical pedestrian data even in adverse conditions like signal failures. The changes in pedestrian counts are shown in Figure 8. As can be seen in Figure 8, the pedestrian counts in chaos condition is more than the pedestrian counts on September 12th.

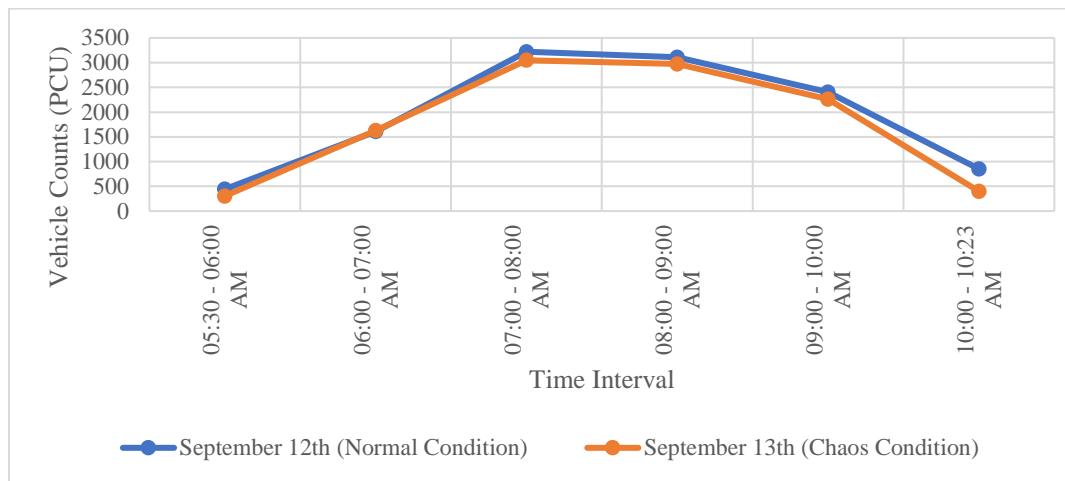


Figure 7. Vehicle Counts Changes

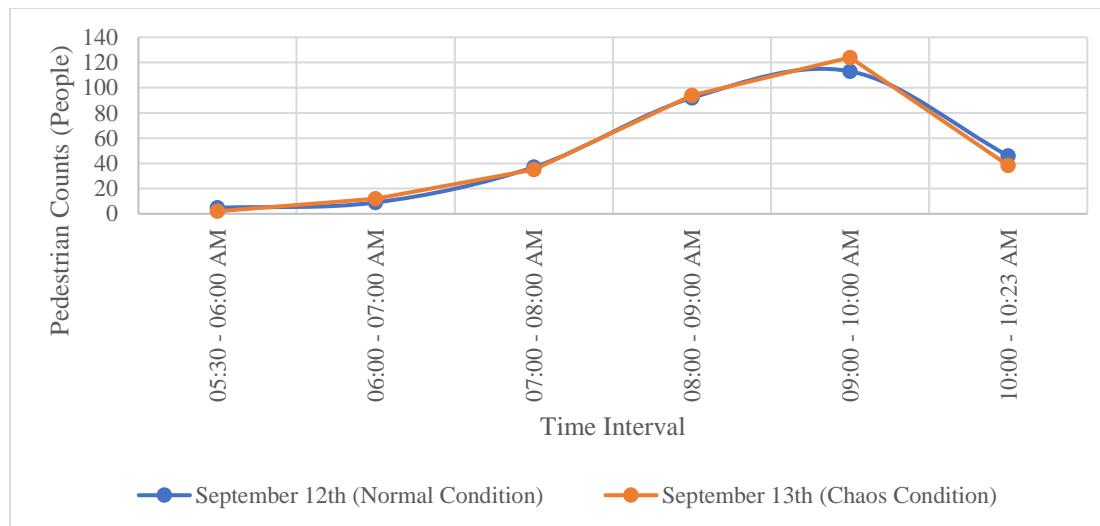


Figure 8. Pedestrian Counts Changes

- 3) **Vehicle-Vehicle Conflicts:** Traditional methods of monitoring vehicle-vehicle conflicts heavily rely on functioning traffic signals, making them vulnerable to inaccuracies and disruptions during signal outages. In contrast, LiDAR sensors offer a robust and reliable solution, ensuring the continual availability of critical traffic conflict data, even in adverse conditions. LiDAR sensors possess the capability to accurately distinguish between vehicles, measure their relative speeds, and assess potential conflict points,

allowing for real-time data collection with minimal disruption. Moreover, their non-intrusive installation and low maintenance requirements minimize the impact on traffic flow, ensuring the continual surveillance of vehicle interactions and conflicts, even during signal system downtime. Figure 9 compares vehicle-vehicle conflicts frequency (left figure) and vehicle-vehicle conflicts severity (right figure) in chaos and normal conditions.

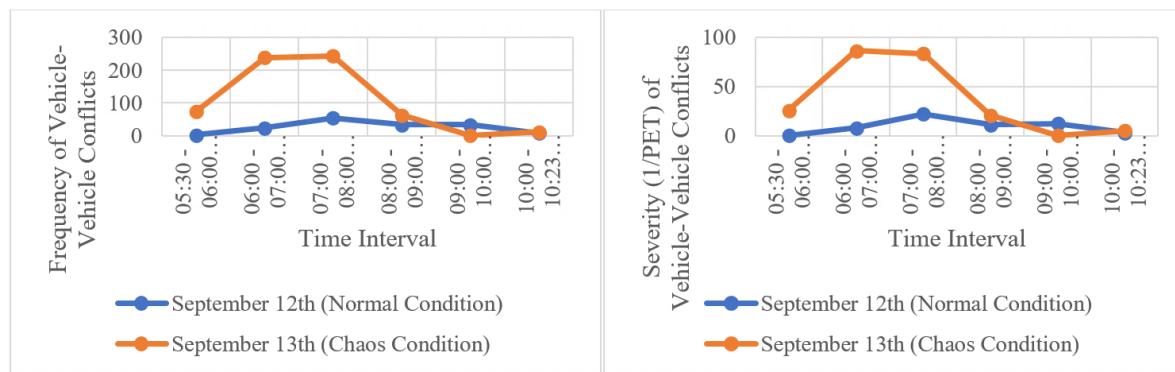


Figure 9. Frequency of Vehicle-Vehicle Conflicts Changes (left figure) and Severity of Vehicle-Vehicle Conflicts Changes (right figure)

Figure 10 illustrates the frequency and severity of vehicle-vehicle conflicts on September 12th and 13th.



Figure 10. Frequency and Severity of Vehicle-Vehicle Conflicts in Normal Condition (September 12th - left figure) and Chaos Condition (September 13th - right figure)

4) Vehicle-Pedestrian Conflicts: Traditional methods for monitoring vehicle-pedestrian conflicts at signalized intersections are susceptible to inaccuracies and safety risks during signal failures. Pedestrians tend to exhibit conservative behavior in crossing, emphasizing the need for effective traffic management and infrastructure maintenance to ensure safety and traffic flow in adverse conditions. Figure 11 compares vehicle-pedestrian conflicts frequency (left figure) and vehicle-pedestrian conflicts severity (right figure) in chaos and normal conditions.

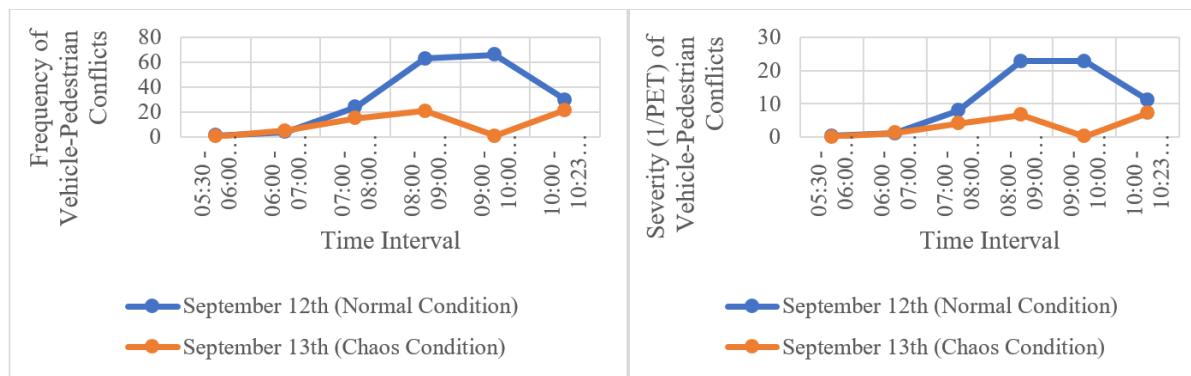
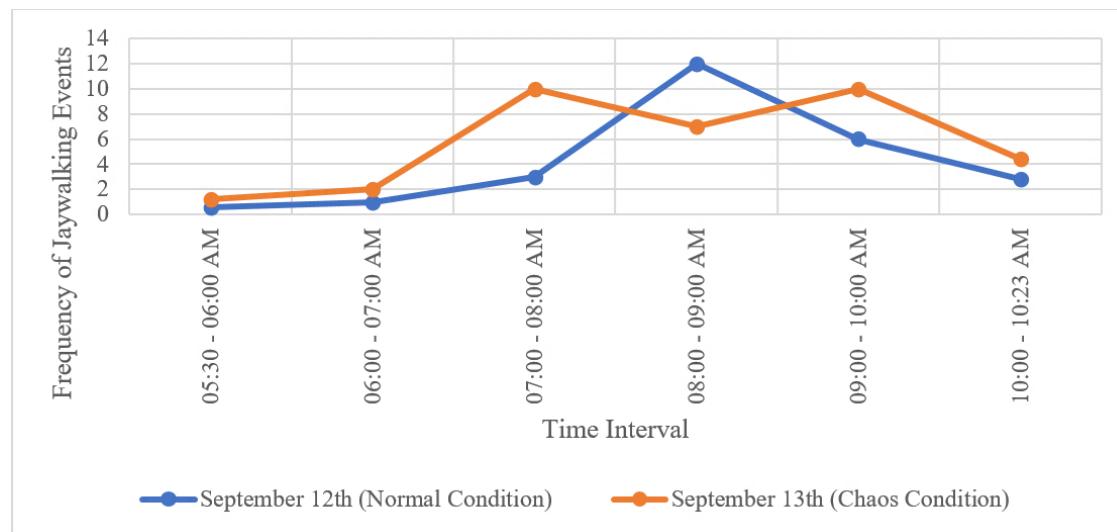


Figure 11. Frequency of Vehicle-Pedestrian Conflicts Changes (left figure) and Severity of Vehicle-Pedestrian Conflicts Changes (right figure)

5) **Jaywalking events Conflicts:** Pedestrians may face challenges when traffic signals malfunction, leading to uncertainty and potential dangers when crossing at signalized intersections. Factors such as impatience, convenience, habitual behavior, lack of alternative crossings, misunderstanding of traffic laws, and traffic flow disruption contribute to an increased likelihood of jaywalking in such situations. To mitigate jaywalking conflicts during signal malfunctions, municipalities and transportation authorities should focus on improved pedestrian education and awareness campaigns, installation of clear signage, and infrastructure enhancements that encourage safe and legal pedestrian crossings [17-20]. Additionally, efforts to promptly repair malfunctioning signals can reduce the instances where pedestrians are forced to make the choice to jaywalk due to signal failures [21-23]. Figure 12 shows the frequency of jaywalking events in chaos and normal conditions. According to Figure 12, jaywalking was more frequent on September 13th under chaos conditions than under normal conditions on September 12th.



CONCLUSION

The application of LiDAR sensor technology in evaluating the consequences of traffic signal failures at intersections marks a significant advancement in understanding and managing challenges associated with unexpected malfunctions. This study, employing LiDAR sensors, has provided detailed, three-dimensional data capturing traffic dynamics during signal failures, offering valuable insights and potential solutions to mitigate disruptions and enhance urban transportation resilience. In conclusion, LiDAR sensor technology proves to be a powerful tool for comprehensively analyzing the impact of traffic signal failures, providing statistical insights to strengthen urban transportation system resilience. The findings from this research empower city planners and policymakers to make informed decisions for improving intersection safety, reducing congestion, and minimizing delays during unexpected signal malfunctions, contributing to the broader goal of creating efficient, reliable, and safe urban transportation networks. Looking ahead, the future of studying traffic signal management during failures using LiDAR sensor technology holds great promise for revolutionizing urban transportation. Addressing specific research areas will contribute to safer, more efficient, and resilient traffic systems, ultimately enhancing the quality of life in smart cities and reducing the societal and economic impact of signal failures. These advancements are poised to play a crucial role in shaping the future of transportation in urban environments.

REFERENCES

- 1) Rahman, M. S., M. Abdel-Aty, S. Hasan, and Q. Cai. (2019). "Applying machine learning approaches to analyze the vulnerable road-users' crashes at statewide traffic analysis zones." *Journal of safety research* 70: 275-288.
- 2) Ansariyar, A. (2022). "Investigating the Car-Pedestrian Conflicts Based on an Innovative Post Encroachment Time Threshold (PET) Classification." Available at SSRN 4377745, Available at SSRN: <https://ssrn.com/abstract=4377745>.
- 3) Ansariyar, A., and A. Taherpour. (2023). Statistical analysis of vehicle-vehicle conflicts with a LIDAR sensor in a signalized intersection. *Advances in Transportation Studies*, 60.
- 4) Arun, A., M. M. Haque, A. Bhaskar, S. Washington, and T. Sayed. (2021). "A systematic mapping review of surrogate safety assessment using traffic conflict techniques." *Accident Analysis & Prevention* 153: 106016.
- 5) Ansariyar, A., and M. Jeihani. (2023). "Statistical Analysis of Jaywalking Conflicts By a LiDAR Sensor." *Scientific Journal of Silesian University of Technology*. Series Transport 120: 17-36, https://sjsutst.polsl.pl/sjsutst_vol120.html.
- 6) Astarita, V., et al. Surrogate safety measures from traffic simulation models a comparison of different models for intersection safety evaluation. *Transportation research procedia*, 2019. 37: p. 219-226.
- 7) Ansariyar, A., and M. Jeihani. (2023, June). Investigating the Vehicle-Bicyclists Conflicts using LiDAR sensor technology at signalized intersections. In *ICTTE 2023: International Conference on Transportation and Traffic Engineering* <https://publications.waset.org/abstracts/166804/investigating-the-vehicle-bicyclists-conflicts-using-lidar-sensor-technology-at-signalized-intersections>, Available at SSRN: <https://ssrn.com/abstract=4473399>.

- 8) Soo, H. Y., D. Teodorovic, and J. Collura. (2006). "A DSS framework for advanced traffic signal control system investment planning." *Journal of Public Transportation* 9(4): 87-106.
- 9) Ansariyar, A., A. Ardestiri, E. Vaziri, and M. Jeihani. (2022). Investigating the Traffic Behavior of Bicyclists in Interaction with Car Users on Shared Bike Lanes Without Physical Barriers. In *Transportation Research Board 102nd Annual Meeting* Location: Washington DC, United States, Available at SSRN: <https://ssrn.com/abstract=4377734>.
- 10) Stylianou, K., and L. Dimitriou. (2018). "Analysis of rear-end conflicts in urban networks using Bayesian networks." *Transportation research record* 2672(38): 302-312.
- 11) Ansariyar, A., and A. Taherpour. (2023). Investigating the accuracy rate of vehicle-vehicle conflicts by LIDAR technology and microsimulation in VISSIM and AIMSUN. *Advances in Transportation Studies*, 61.
- 12) Sun, Y., H. Xu, J. Wu, E. Y. Hajj, and X. Geng. (2017). "Data processing framework for development of driving cycles with data from SHRP 2 naturalistic driving study." *Transportation Research Record* 2645(1): 50-56.
- 13) Suzuki, K., and H. Ito. (2017). "Empirical analysis on risky behaviors and pedestrian-vehicle conflicts at large-size signalized intersections." *Transportation research procedia* 25: 2139-2152.
- 14) Ansariyar, A., A. Ardestiri, and M. Jeihani. (2023). "Investigating the collected vehicle-pedestrian conflicts by a LIDAR sensor based on a new Post Encroachment Time Threshold (PET) classification at signalized intersections." *Advances in Transportation Studies* 61: 103-118. Available at: <https://www.atsinternationaljournal.com/index.php/2023-issues/Ixi-november-2023/1442-investigating-the-collected-vehicle-pedestrian-conflicts-by-a-lidar-sensor-based-on-a-new-post-encroachment-time-threshold-pet-classification-at-signalized-intersections>.
- 15) Suzuki, S., P. Raksincharoensak, I. Shimizu, M. Nagai, and R. Adomat. (2010). Sensor fusion-based pedestrian collision warning system with crosswalk detection. *2010 IEEE Intelligent Vehicles Symposium*, IEEE.
- 16) Thakur, R. (2016). "Scanning LIDAR in Advanced Driver Assistance Systems and Beyond: Building a road map for next-generation LIDAR technology." *IEEE Consumer Electronics Magazine* 5(3): 48-54.
- 17) Uzondu, C., S. Jamson, and F. Lai. (2018). "Exploratory study involving observation of traffic behaviour and conflicts in Nigeria using the Traffic Conflict Technique." *Safety science* 110: 273-284.
- 18) Van der Horst, A. R. A. (1991). *A time-based analysis of road user behaviour in normal and crucial encounters*.
- 19) Vanlaar, W., M. M. Hing, S. Brown, H. McAteer, J. Crain, and S. McFaull. (2016). "Fatal and serious injuries related to vulnerable road users in Canada." *Journal of safety research* 58: 67-77.
- 20) Vasconcelos, L., L. Neto, Á. M. Seco, and A. B. Silva. (2014). "Validation of the surrogate safety assessment model for assessment of intersection safety." *Transportation Research Record* 2432(1): 1-9.
- 21) Vogel, K. (2003). "A comparison of headway and time to collision as safety indicators." *Accident analysis & prevention* 35(3): 427-433.

- 22) Wang, D. Z., I. Posner, and P. Newman. (2012). What could move? finding cars, pedestrians and bicyclists in 3d laser data. *2012 IEEE International Conference on Robotics and Automation*, IEEE.
- 23) Wang, Y., X. Yang, H. Liang, and Y. Liu. (2018). “A review of the self-adaptive traffic signal control system based on future traffic environment.” *Journal of Advanced Transportation* 2018.

A Comparative Analysis of Probe Data Sources for Provision of Work Zone Congestion Information to the Public

Dorcas Okaidjah¹; Skylar Knickerbocker²; and Christoher Day, Ph.D.³

¹Dept. of Civil, Construction, and Environmental Engineering, Iowa State Univ., Ames, IA.
Email: okaidjah@iastate.edu

²Dept. of Civil, Construction, and Environmental Engineering, Iowa State Univ., Ames, IA.
Email: sknick@iastate.edu

³Dept. of Civil, Construction, and Environmental Engineering, Iowa State Univ., Ames, IA.
Email: cmday@iastate.edu

ABSTRACT

With the rise in highway maintenance demands, work zones have proliferated across the United States, posing traffic and safety challenges. Dissemination of work zone information to the public is essential to improve traffic flow and efficiency in work zones. Central to the success of smart work zones (SWZ) public information is real-time data availability. This study compares two alternative probe data types to ground truth data from field sensors. The two emerging probe vehicle datasets tested in this study are connected vehicle (CV) data supplied by a third-party data vendor and segment average speeds (for each minute) obtained from a different vendor. The mean and minimum speeds of the probe data types were analyzed using frequency distributions and descriptive statistics. The CV data displayed a high amount of variability in minimum and mean speeds in both travel directions compared to segment average speeds, making them more similar to the ground truth data. This finding is useful in providing precise real-time information to the driving public and will improve the accuracy with which work zone information is relayed.

INTRODUCTION

Work zones are necessary for road maintenance but can be dangerous for workers and drivers. Despite speed limit reductions in work zones, motorists often do not comply, leading to increased work zone crashes. Work zone crashes in the United States are estimated to cost the construction sector \$3.5 billion annually (1). Work zones aim to create a secure space for both road users and workers, but they are often regarded as crash-prone areas due to alterations in road layouts and posted speed limits. Work zone safety remains a crucial topic for transportation agencies, highlighting the importance of this study.

Dynamic message signs (DMSs) allow for the efficient communication of information to the public, such as notifications of slowed or stopped traffic, lane closures, travel times, road construction, traffic detours, and other information. DMSs have been used to improve work zone safety by reducing vehicle speeds and managing congestion. Previously, SWZ equipped with real-time traffic information would necessitate multiple sensors to fully cover a work zone. However, with the advent of probe data, which is widespread, it's now possible to equip SWZ with real-time traffic updates without this extensive infrastructure. Therefore, evaluating the accuracy of these probe data types for relaying real time information is vital.

Several studies have evaluated the effectiveness of SWZ to optimize traffic flow, enhance congestion, and convey information to the public (2–4). Varying data types, such as sensor data,

naturalistic driving data, and probe data, have been used for smart work zone studies. Segment average speeds have been employed in previous studies to develop performance measures, identify shockwave boundary, assess the extent of congestion, and examine signal performance measures (5–8).

Unlike sensor data, which have a relatively low level of geographic coverage and a high sample rate, CV data has a high level of coverage and a lower sample rate. Segment average speeds are provided on a minute-by-minute basis. They also have wide coverage, but the data is provided as an aggregate value for all the vehicles observed within the minute. Studies have noted delays in the identification of traffic congestion with the use of segment average speeds, with reported lags varying between 3 minutes and up to 6 minutes when benchmarked against sensor data (9,10). Another study reported speed bias between sensor and segment average speeds (11). Therefore, this study is essential to understand the frequency and temporal distributions of speed metrics of the multiple probe data types to enhance the delivery of accurate information to the public, which would ultimately improve work zone safety.

LITERATURE REVIEW

Dynamic Message Signs (DMS) effectively enhance public understanding and operational efficiency, particularly on rural highways, leading to significant speed reductions (3,12,13). The effectiveness of Speed Monitoring Displays (SMD) was highlighted in the Midwest States Smart Work Zone Deployment Initiative, with three SMDs installed along I-80 near Lincoln, Nebraska, showing significant improvements in speed compliance, although one SMD showed varied results (2). The introduction of dynamic and graphical Variable Message Sign (VMS) systems has positively influenced driver behavior, encouraging safer driving practices (14).

The Utah Department of Transportation's trials with Portable Variable Speed Limit (PVSL) systems reported speed reductions of 15–25 mph in active work zones, improving compliance with speed limits (15). Online algorithms for variable speed limits have also been developed, enhancing traffic throughput and reducing delays (16). A Virginia study evidenced the efficacy of radar-activated changeable message signs (CMS) in reducing speeds across all levels on interstate highways, particularly impacting high-speed vehicles (17). Additionally, lane ends, speed limits, and active signage were identified as significant influencers of driver behavior, particularly in advance warning areas of work zones (18). Specific signage types, such as speed photo signs and dynamic speed display signs, have also been noted for their substantial impact on reducing speeds (19).

Traffic simulation models have been crucial in developing strategies for managing congestion due to lane reductions in work zones (20). An intelligent work zone study on Interstate 80 in Nebraska evaluated four automatic queue detection (AQD) systems, demonstrating low error rates (0.7–2.3%) and a 47% greater reduction in space mean speed (SMS) with warning messages compared to those without (3). Wang et al. (21) proposed enhancements to DMS algorithms using machine learning for more accurate congestion messaging.

The use of unmanned aerial vehicles for real-time monitoring of work zone traffic, particularly for detecting end-of-queue, has shown the potential to enhance safety through dynamic warnings (22). A developed queue-end warning system, integrating traffic sensors and artificial neural network models, has proven effective in predicting queue-end locations and issuing warnings (4). Lastly, Li et al. (23) assessed different sensor networks to estimate work

zone traffic. Results revealed that nonlinear Kalman filter algorithms were the most accurate, reducing errors by up to 30%. Another study used SWZ data in an algorithm to enhance real-time travel time and delay predictions, showing significant improvements, especially in areas with fluctuating queue conditions (24).

METHODOLOGY

To compare various probe data types, we collected data from multiple sources:

- *Segment average speed* – This data was obtained from INRIX. The data consists of an average speed for each minute of the day. The vendor includes a confidence score that indicates whether the data consists of entirely real-world data, some amount of historic/imputed data, or entirely historical or imputed data. Because this study focuses on a real-time application, we included all data in our analyses.
- *Connected vehicle data* - CV trajectory data was obtained from Wejo. This data is obtained from CVs outfitted with equipment that permits them to broadcast their locations. Vehicle manufacturers have recently been introducing this capability into new vehicles and have been collecting the data and reselling it through data vendors. CV data includes the full geographic record of the movement of individual vehicles with a 3-second time resolution. Consequently, there are more concerns regarding the privacy of CV data than segment average speeds.
- *Radar speed sensors*: These sensors are deployed to detect speeds locally and typically are limited to observing vehicles within a short distance (up to hundreds of feet). Data from Wavetronix sensors deployed in the field were obtained from the Iowa Department of Transportation (IADOT). The data consists of average speed per lane and vehicle count for each minute.

The Iowa Department of Transportation (IADOT) procured segment average speed data and CV trajectory data continuously throughout the year 2022. IADOT has also used radar speed sensors to measure traffic conditions in work zones. This paper examines June 2022 data from the three datasets listed above. To support this analysis, we selected a work zone encompassing diverse traffic scenarios and varying levels of congestion. The work zone was located in Johnson County, Iowa, and Figure 1 shows the work zone selected for the study. The annual average daily traffic (AADT) for this road is 52,300 vehicles per day.

We selected two sensors within the work zone, creating a defined study area with a 0.5-mile radius surrounding each sensor. The 0.5-mile radius is employed because the distance between a sensor and a DMS is typically 0.5 to one mile. The sensors and study locations ("Location1" and "Location2") within the work zone are shown in Figure 1. The 0.5-mile radius was applied in both the Northbound and Southbound directions to capture a comprehensive range of traffic conditions within the study area. Considering both directions of travel enabled us to determine whether directional factors significantly influence the data or whether the characteristics are consistent. We aggregated metrics into one-minute time bins for each probe data type to maintain consistency in temporal analysis. This aggregation approach allowed for a detailed yet manageable dataset to assess performance metrics over time.

The sensor data was defined as our ground truth data. The CV data provided granular insights into individual vehicle movements, while the sensor data offered a general perspective on traffic flow. A key focus of our study was on the speed metrics, specifically the average and minimum speeds. For the CV data, we aggregated this information across all waypoints for each location.

We further filtered the data to exclude instances where no waypoint was recorded within a given one-minute timestamp, ensuring the integrity and relevance of our data set. Similarly, for the Inrix data, we aggregated the speed metrics across the segments in each location. In the case of traditional sensor data, we aggregated the data across each sensor point to obtain the average and minimum speed. This approach provided an understanding of traffic patterns within our study area, enabling us to compare and contrast the probe data obtained from different sources.

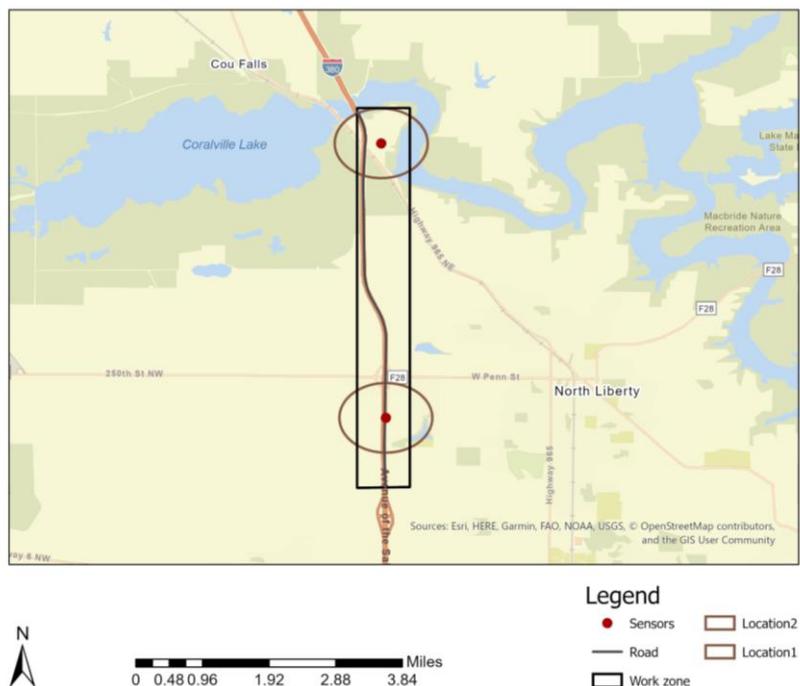


Figure 1: Work zone area with study locations in the northbound direction

RESULTS

The mean speed frequency distribution for the various probe data types for the northbound and southbound directions is shown in Figures 2 and 3, respectively. The minimum speed frequency distribution is shown in Figures 4 and 5, respectively.

The figures demonstrate a comprehensive view of speed patterns for the three data types.

Table 1 displays the descriptive statistics of the mean and minimum speed of all three probe data types in northbound and southbound directions. The temporal speed distribution of mean speeds captured by different probe data sources is displayed in Figures 6 and 7.

The CV data demonstrates a wider spread of both mean and minimum speeds in both directions (Figures 2–5). Descriptive statistics shown in Table 1 show that the CV data has much more variability than the segment average speed data. This greater variability in both directions demonstrates that the CV data captures a broader range of speed and is consequently more sensitive to a range of traffic conditions, including both the flow and interruptions of traffic. Despite this more significant variability, the primary concentration of data points for mean speed is within the 60–70 range in both directions, while the minimum speeds show occurrences of very low speeds, which could indicate traffic stops or heavy congestion.

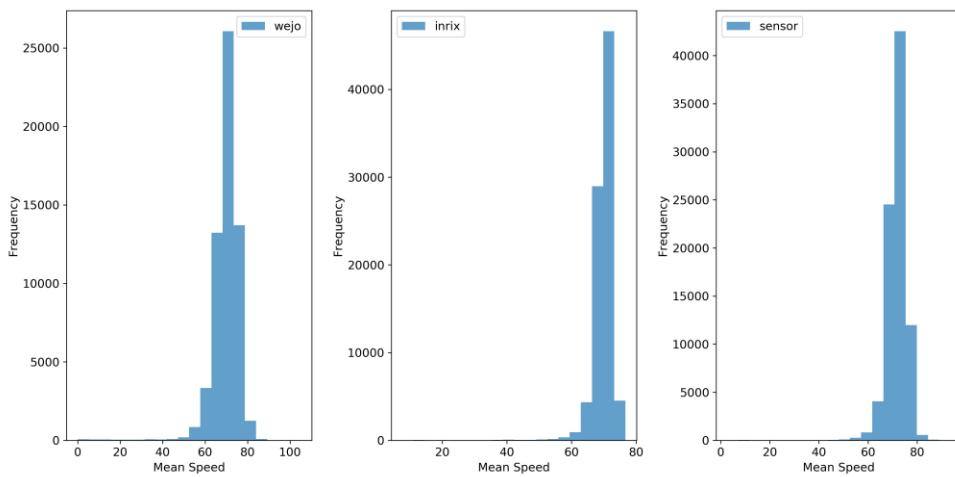


Figure 2: Mean speed distribution of probe data types (Northbound)

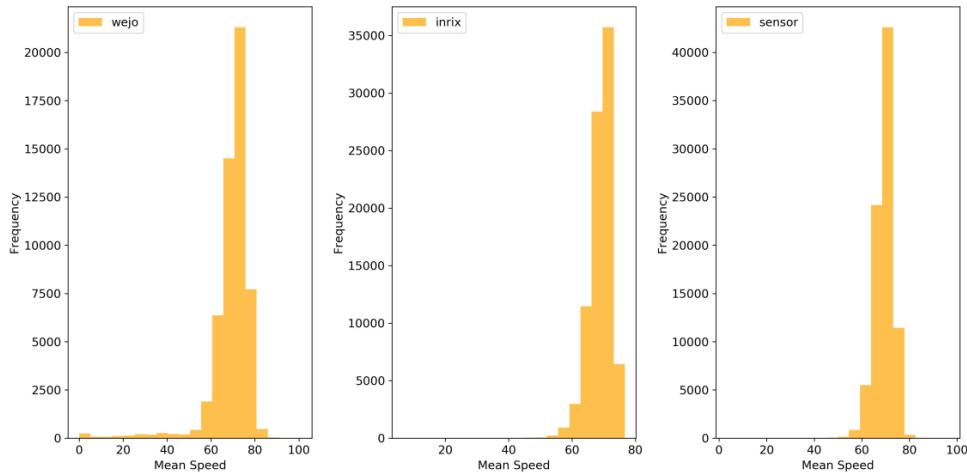


Figure 3: Mean speed distribution of probe data types (Southbound)

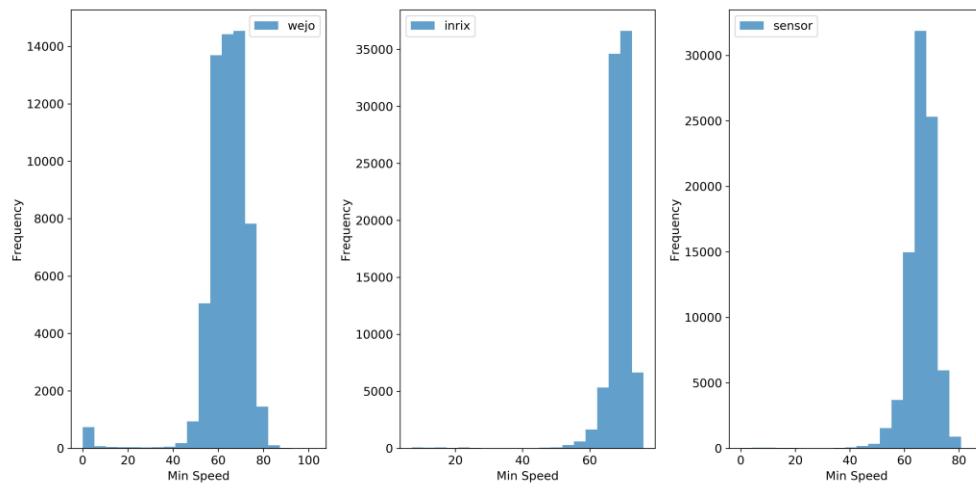


Figure 4: Minimum speed distribution of probe data types (Northbound)

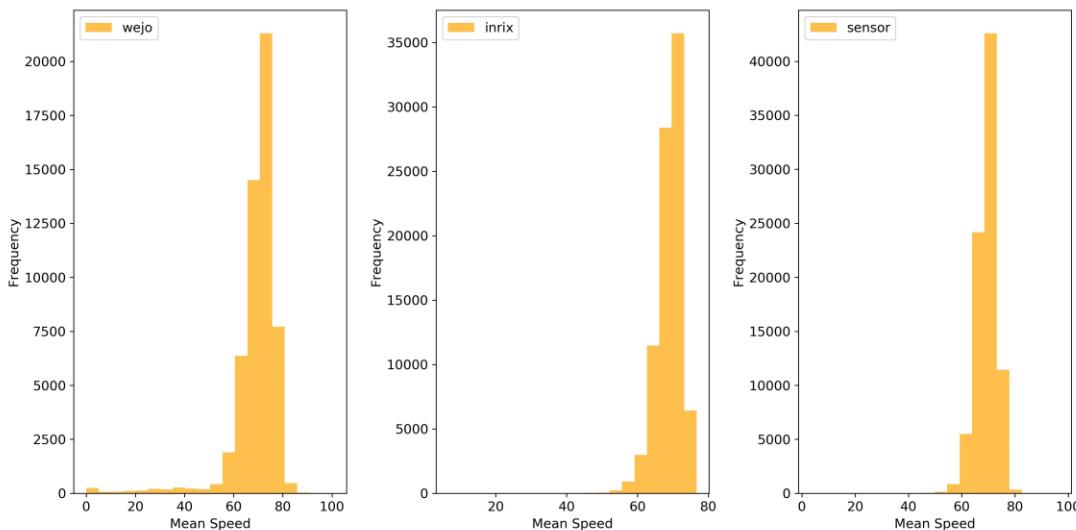


Figure 5: Minimum speed distribution of probe data types (Southbound)

Table 1. Descriptive Statistics of Mean and Minimum Speed

Statistics	CV	INRIX	Sensor	CV	INRIX	Sensor
	Mean Speed (NB)			Min Speed (NB)		
Count	59301	86348	85203	59301	86348	85203
Mean	69.825	69.801	71.619	63.890	68.800	65.970
Standard	6.711	3.619	4.677	10.273	4.665	5.668
Minimum	0.000	7.667	2.667	0.000	7.000	0.000
25%	67.356	68.667	69.833	59.409	68.000	64.000
50%	70.615	70.333	72.000	65.139	70.000	66.000
75%	73.547	71.667	74.1667	70.147	71.000	69.000
Maximum	104.987	76.667	93.50	102.601	76.000	85.000
	Mean Speed (SB)			Min Speed (SB)		
Count	54400	86223	85207	54400	86223	85207
Mean	69.401	68.974	69.439	63.326	67.372	64.105
Standard	9.446	3.692	4.055	13.873	4.605	4.718
Minimum	0.000	6.667	3.333	0.000	4.000	0.000
25%	67.046	67.000	67.000	59.409	65.000	62.000
50%	71.273	69.333	69.333	65.853	68.000	64.000
75%	74.387	71.667	71.667	70.861	71.000	67.000
Maximum	100.929	76.667	76.667	100.929	76.000	84.000

The segment average speed data exhibits the highest consistency and the lowest standard deviation among the datasets for both mean and minimum speeds in both directions, as shown in Table 1. The narrower distribution in both directions indicates that this data is highly influenced by the aggregation process. The numbers suggest more uniform speeds, but this is probably mainly because the outliers are lost in the aggregation process. In addition, Figure 7 illustrates

that the segment speed data ("Inrix") does not report congestion until a few minutes after the sensor data begins showing congestion. Segment average speed data has been extensively used to identify congestion for comparative purposes. The similar National Performance Measures Research Dataset (NPMRDS) is used by USDOT to evaluate travel time reliability on an annual basis. However, the aggregate nature of the data and its latency may reduce its usefulness for timely detection of slowed traffic for public information.

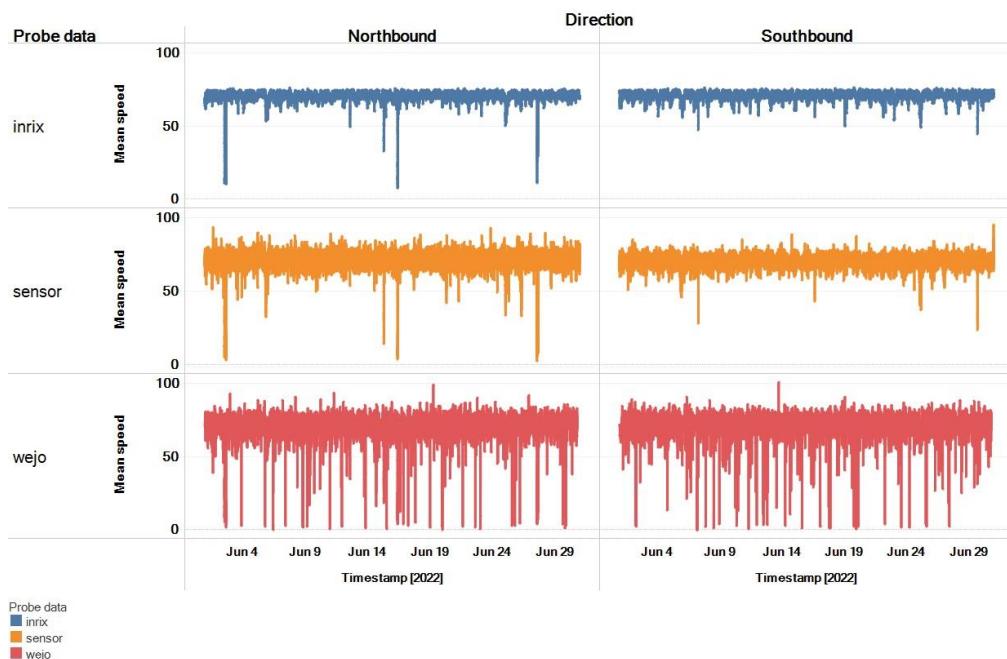


Figure 6: Temporal distribution of mean speeds in June

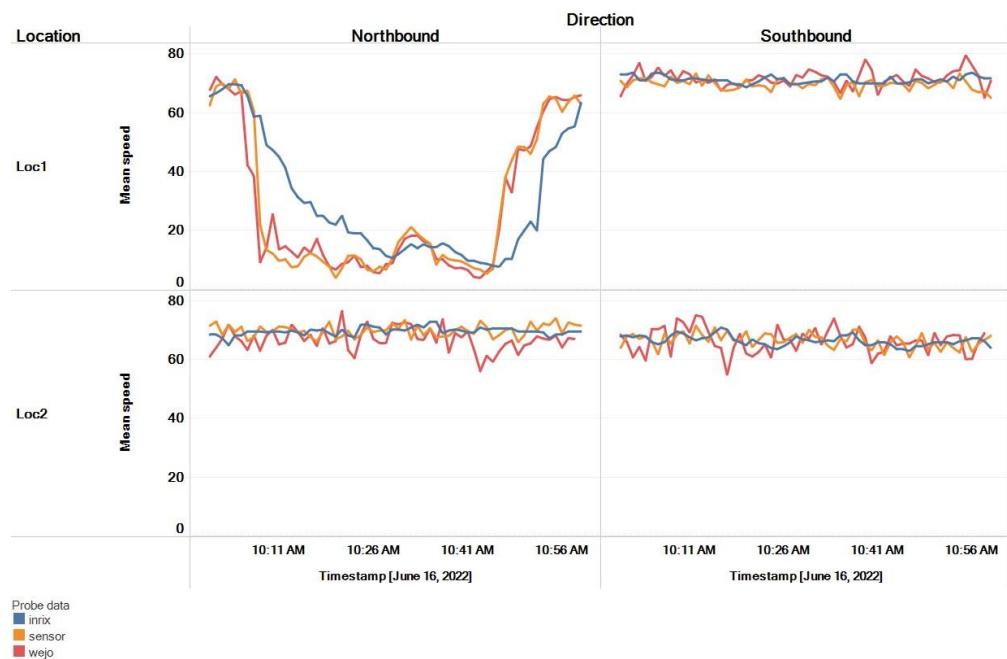


Figure 7: Temporal distribution of mean speeds for an hour

The temporal distribution of mean speeds in Figure 6 indicates INRIX has a consistent pattern with minimal deviations over the observed period. This is primarily due to the aggregation across multiple vehicles and along segments. Meanwhile, the CV data exhibits the most significant amount of variability. Unlike the sensor data, the CV data can be obtained across the entire work zone and includes individual vehicles, so it is not surprising that it is more variable. The comparison between Northbound and Southbound directions does not show a significant change in mean speed patterns, implying that the observed trends are likely symmetric with respect to directional flow.

CONCLUSION

This study sought to compare various probe data types to determine their effectiveness in relaying work zone information to the driving public. The study employed frequency distributions, descriptive statistics, and temporal distribution of mean speeds for its analysis, and several visualizations were presented, along with a discussion of the key differences in the compared data sets. CV data and segment average speed data were compared against sensor data, which served as ground truth. CV data, with its wide range and significant variability, is well suited for analyzing diverse traffic conditions for public information and is able to capture varying traffic dynamics in both directions of the work zone. Segment average speed data is slower to detect incidents and exhibits less variability because of the nature of the aggregation process. Compared to sensor data, the CV data is quicker to identify congestion than segment average speed data. So, whereas segment average speed data is convenient for evaluating performance over long time periods, such as for congestion reports or travel time reliability analysis, these results show that CV data may be better suited for real-time applications such as providing public work zone information. One limitation of this study is that it did not examine the integration of CV data with segment average speed data, which could potentially offer a robust approach to real-time work zone information dissemination. Future research will explore opportunities for data fusion in application to work zone monitoring and provision of public information. Ongoing research will expand the analysis to more work zones to address transferability.

CV data is relatively new, and the results indicate a potential for a data-driven method to generate real-time work zone public information messages. The findings from this study present significant insights for transportation agencies to refine their work zone safety strategies and enhance public communication mechanisms. As the research expands to include more work zones and diverse traffic scenarios, the insights gathered from CV data will be central in informing future traffic management and public safety initiatives.

REFERENCES

1. Cunningham, I., Jack, R., Anderson, S. M., Fitzsimmons, E. J., and Nye, B. G. *Analysis of Speed Profiles and Evaluation of Dynamic Signs in Kansas Work Zones*. Kansas Department of Transportation. Bureau of Research; 2021.
2. Pesti, G., and McCoy, P. T. Long-term effectiveness of speed monitoring displays in work zones on rural interstate highways. *Transportation Research Record*. 2001;1754(1):21–30.
3. Zhao, L., and Rilett, L. R. Verification and Efficacy of Automated Queue Detection Systems on Interstate Intelligent Work Zones. *Transportation Research Record*. 2022;03611981221108155.

4. Khan, A. M. Intelligent infrastructure-based queue-end warning system for avoiding rear impacts. *IET intelligent transport systems*. 2007;1(2):138–43.
5. Remias, S. M., Brennan, T. M., Day, C. M., Summers, H. T., Horton, D. K., and Cox, E. D. Spatially referenced probe data performance measures for infrastructure investment decision makers. *Transportation Research Record*. 2014;2420(1):33–44.
6. Li, H., Remias, S. M., Day, C. M., Mekker, M. M., Sturdevant, J. R., and Bullock, D. M. Shock wave boundary identification using cloud-based probe data. *Transportation Research Record*. 2015;2526(1):51–60.
7. O'Brien, P., Stevanovic, A., Day, C., Hale, D., and Matout, N. *A Methodology and Case Study: Evaluating the Benefits and Costs of Implementing Automated Traffic Signal Performance*. United States. Federal Highway Administration; 2020.
8. Day, C., McNamara, M., Li, H., Sakhare, R., Desai, J., Cox, E., Horton, D., and Bullock, D. (2016). *2015 Indiana Mobility Report and Performance Measure Dashboards*.
9. Adu-Gyamfi, Y., Sharma, A., Knickerbocker, S., Hawkins, N., and Jackson, M. Reliability of probe speed data for detecting congestion trends. In IEEE; 2015. p. 2243–9.
10. Kim, S., and Coifman, B. Comparing INRIX speed data against concurrent loop detector stations over several months. *Transportation Research Part C: Emerging Technologies*. 2014;49:59–72.
11. Ahsani, V., Amin-Naseri, M., Knickerbocker, S., and Sharma, A. Quantitative analysis of probe data characteristics: Coverage, speed bias and congestion detection precision. *Journal of Intelligent Transportation Systems*. 2019;23(2):103–19.
12. Kwigizile, V., Oh, J. S., Van Houten, R., Lee, K., Kwayu, K. M., and Abushattal, M. *Quantifying Effectiveness and Impacts of Digital Message Signs on Traffic Flow*. Michigan. Dept. of Transportation. Research Administration; 2022.
13. Firman, U., Li, Y., and Bai, Y. Determining the Effectiveness of Portable Changeable Message Signs in Work Zones. In: *2009 Mid-Continent Transportation Research Symposium* Iowa Department of Transportation Iowa State University, Ames University of Northern Iowa, Cedar Falls National Center for Freight and Infrastructure Research and Education (CFIRE) Wisconsin Department of Transportation. 2009.
14. Almallah, M., Hussain, Q., Alhajyaseen, W. K. M., Pirdavani, A., Brijs, K., and Dias, C. Improved traffic safety at work zones through animation-based variable message signs. *Accident Analysis and Prevention*. 2021 Sep;159.
15. Van Jura, J., Haines, D., and Gemperline, A. Use of portable and dynamic variable speed limits in construction zones. *Transportation research record*. 2018;2672(16):35–45.
16. Lin, P. W., Kang, K. P., and Chang, G. L. *Exploring the effectiveness of variable speed limit controls on highway work-zone operations*. In Taylor & Francis; 2004. p. 155–68.
17. Garber, N. J., and Patel, S. T. *Effectiveness of changeable message signs in controlling vehicle speeds in work zones*. Virginia Transportation Research Council; 1994.
18. Thapa, R., Hallmark, S., Smadi, O., and Goswamy, A. Assessing driving behavior upstream of work zones by detecting response points in speed profile: A naturalistic driving study. *Traffic injury prevention*. 2019;20(8):854–9.
19. Banerjee, S., Jeihani, M., and Morris, D. Impact of Work Zone Signage on Driver Speeding Behavior: A Driving Simulator Study. In: *Transportation Research Board 98th Annual Meeting*. 2019.

20. Kamyab, A., Maze, T. H., Nelson, M., and Schrock, S. D. Using simulation to evaluate the performance of smart work zone technologies and other strategies to reduce congestion. In: *Proceedings Of 6th World Congress On Intelligent Transport Systems (Its)*, Held Toronto, Canada, November 8-12, 1999. 1999.
21. Wang, S., Sharma, A., and Knickerbocker, S. Analyzing and improving the performance of dynamic message sign reporting work zone-related congestion. *Transportation Research Record*. 2017;2617(1):71–7.
22. Messenger, R., Islam, M. Z., Whitlock, M., Spong, E., Morton, N., and Claggett, L. Real-Time Traffic End-of-Queue Detection and Tracking in UAV Video. arXiv preprint arXiv:230201923. 2023.
23. Li, Y., Martinez Mori, J. C., and Work, D. B. Improving the effectiveness of smart work zone technologies. Illinois Center for Transportation Series No 16-023. 2016.
24. Songchitruksa, P., and Middleton, D. Enhancing work zone travel time estimation using smart work zone data. *Presented at the Transportation Research Board 92nd Annual Meeting Transportation Research Board*, 2013.

Investigating the Role of Dimension Reduction in Counteracting Machine Learning Performance Bias in TSMO Strategies

Mustafa Attallah¹ and Jalil Kianfar²

¹Doctoral Candidate, Dept. of Civil, Computer, and Electrical Engineering, Saint Louis Univ., Saint Louis, MO. Email: mustafa.attallah@slu.edu

²Associate Professor, Dept. of Civil, Computer, and Electrical Engineering, Saint Louis Univ., Saint Louis, MO (corresponding author). Email: jalil.kianfar@slu.edu

ABSTRACT

Machine learning models are increasingly being utilized to develop predictive models for Transportation Systems Management and Operations (TSMO) applications. These models are often assessed based on a global performance metric that evaluates the model's performance when the entire testing dataset is presented to the model. A TSMO application is expected to perform reliably and consistently in various situations and roadway conditions. The reliability and consistency of the model predictions for various scenarios are critical to the success of transportation agencies' efforts to address mobility and safety issues. Performance bias might be influencing the model when a model's performance is inconsistent for different scenarios. This paper investigates the performance bias that the traffic management center may face when applying machine learning methods to predict incident clearance time. Additionally, dimension-reduction techniques are employed as mitigation techniques in the model development process. This paper investigates the impact of two common dimension-reduction methods, important feature selection and principal component analysis, on performance bias. In a case study, this paper investigates the performance bias of RF, BRNN, SVR, KNN, XGB, GP, and NNET in incident clearance time prediction. Incident data from three interstate corridors in Missouri, USA, were utilized to develop and evaluate the models. Repeated k-fold cross-validation was used to prepare 20 training and testing sets to demonstrate and assess the learners' performance variations due to data splits. The results indicated that the seven learners suffered from performance bias. The analysis of the impact of dimension-reduction models revealed that the important feature selection method did not significantly mitigate the performance bias. On the other hand, the principal component analysis method significantly mitigated this bias for all learners, with poor-performing learners gaining the most improvements. In addition to contributing to reducing the performance bias, the principal component analysis significantly reduced the learners' global (i.e., overall) error metrics.

INTRODUCTION

Traffic congestion in major urban areas results in air pollution, economic loss, and decreased quality of life. Traffic congestion is classified into two forms: (i) recurrent congestion, which primarily occurs at active bottlenecks due to increased demand during peak hours, and (ii) non-recurrent congestion as a result of events such as incidents, severe weather, special events, and work zones (Grigorev et al. 2022). Traffic incidents are considered a major factor in disrupting traffic flow (Ghosh et al. 2018) and cause traffic congestion and delays (Jiang et al. 2014). Previous studies suggested that these non-recurrent events contribute to 60% of traffic

congestion in the USA (Hojati et al. 2011). The Australian government also announced in 2017 that the annual economic cost of road incidents in Australia was estimated at \$27 billion annually. A European study suggests that reducing the incident clearance time by one minute can save €57 per incident (Adler et al. 2013).

Traffic Incident Management Systems (TIMS) aim to respond promptly to incidents and mitigate their impact (Grigorev et al. 2022). TIMS collects incident data such as duration, incident characteristics, and other important factors. Incident response process is described in following phases: (1) incident detection time, which is the time between the occurrence of an incident and its reporting; (2) incident response time represents the time interval between the reporting and the arrival of the first respondent to the location of the incident, (3) incident clearance time which is the time between the arrival of the first respondent and the clearance of the incident, (4) incident recovery time indicates the time interval between the complete clearance of an incident and the time of the normal traffic returns. This paper focuses on incident clearance time as estimating the incident duration promptly could save operational costs and end-user time (Grigorev et al. 2022), improving the overall TIMS process. Considering the importance of the estimated incident clearance time in TIMS strategies, developing and utilizing a reliable and trustworthy clearance time estimation model is essential for mitigating incident impacts. Models that produce inaccurate predictions may lead to the adoption of less-than-optimal response strategies and may result in undesirable outcomes. Over the past decades, researchers have worked to develop and improve incident clearance time prediction models. However, the previous research in incident clearance duration primarily focused on the global prediction performance metric of the models in training, evaluating, testing, and selection of the models (Abdel-Aty and Radwan 2000; Grigorev et al. 2022; Ozbay and Noyan 2006; Salum et al. 2023).

In all previous research on incident clearance time prediction, the entire testing (i.e., unseen) data are presented to the model and a performance metric such as root-mean-square-error is calculated. This global performance metric is reported for all the learners, and typically, the model with the best performance metric is selected as the preferred model. Such an approach, however, assumes that the model performs reliably and consistently for all incident scenarios. For example, this assumption implies that the model produces a similar error when predicting incident clearance time for one blocked and three blocked lanes. However, the attributes, characteristics, and clearance time distribution of incidents with one-blocked lane might differ from incidents with two-, three- or more than three-blocked lanes. Given the differences in incident attributes, different machine learning algorithms (i.e., learners) might not produce similar error metrics for all incident scenarios. The inconsistent performance of different learners towards one or some incident scenarios is called performance bias. Therefore, reducing the error variation for a learner across different scenarios reduces the performance bias of that learner across different scenarios. This paper investigates the performance bias in RF, BRNN, SVR, KNN, XGB, GP, and NNET for incident clearance time prediction. Considering that these learners follow different training algorithms, they may not exhibit the same bias towards the same incident scenarios even though all these learners are provided with the same training and testing data. Such a trend is evidence of performance bias, where learners perform differently across different scenarios.

This paper proposes a scenario-aware perspective to analyze and assess the performance of common machine learning models to predict incident clearance time. Investigating the performance of the conventional models to predict the incident clearance time for different

incident scenarios might reveal weaknesses and biases of learners towards some scenarios, which impacts the overall reliability of a model. In addition to investigating the performance bias of the common learners when all the relevant independent variables are utilized for incident clearance time prediction, this paper investigates the impact of two-dimension-reduction approaches to mitigate the performance bias of conventional learners. The first dimension-reduction approach, feature selection (Chand et al. 2022), utilizes only the important original features as independent variables to predict the incident clearance time. The second dimension-reduction approach, Principal Component Analysis (Wold et al. 1987), converts the original independent features into new components while preserving the variation of these features.

The remainder of this paper is organized as follows. First, the relevant literature on incident clearance time and dimension-reduction is reviewed. Next, the method sections discuss performance bias, machine learning, dimension-reduction, performance metrics, and Repeated k-folds cross-validation. The following section describes the case study and data. Then, the results obtained from the base, feature selection, and Principal Component Analysis (PCA) methods are presented. The following section provides a summary and conclusion of the study.

LITERATURE REVIEW

This section reviews the relevant literature on incident clearance time prediction and discusses applications of dimension-reduction methods to incident clearance time prediction.

Li et al. (2015) introduced a competing risk mixture hazard-based model that leverages multinomial logistic and parametric hazard-based models to analyze and predict the duration of traffic incidents. By incorporating text analysis techniques and utilizing these models, the performance of the proposed approach improved compared to the non-mixture model in sequential incident duration prediction. The study dataset included 11,568 incidents reported in Singapore with incident durations ranging from 2 to 180 minutes, and Mean Absolute Percentage Error (*MAPE*) was utilized as the metric to evaluate the models.

Ghosh et al. (2018) proposed Treebagger, an adaptive ensemble model, to predict the duration of incidents. They demonstrated that the Treebagger model outperformed other common regression methods. The *MAPE* of incidents with clearance time of 36 to 200 minutes ranged from 25% to 55%. Notably, the accuracy of predictions improved significantly over time for longer incidents, with *MAPE* values ranging from 76% to 50% for incidents exceeding 65 minutes. Moreover, averaging all incidents, the overall *MAPE* value showed a 50% improvement with elapsed time for both reported and effective duration prediction. The effectiveness of the adaptive ensemble model, particularly the Treebagger model, provides reasonable and increasingly accurate forecasts for the incident duration, even with limited initial information, and supports its value in mitigating traffic congestion. The study collected data from 11,278 expressway incidents over six months in Singapore, including 1,745 incidents reported at 197 exits. The authors noted that *MAPE* values tended to be higher for shorter incidents than longer ones due to the utilization of more features in predicting the duration of longer incidents.

Lee (2011) proposed an Artificial Neural Network-based model to predict the accident duration. The Principal Component Analysis (PCA) method was utilized to reduce the number of features to predict the clearance time while preserving sufficient information from traffic and accident components. A model was trained with 24 accidents and tested with 15 unseen accidents. The proposed model could reduce the *MAPE* from 12.37% to 10.760% compared to the base method.

Shang et al. (2019) developed a Neighborhood Components Analysis (NCA) and Bayesian Optimization Algorithm (BOA)- optimized Random Forest (RF) (NCA-BOA-RF) to model incident clearance time. The NCA method selected the relevant features for traffic incident duration prediction. A total of 440 traffic incidents were used in this study, with a split of 70% and 30% for training and testing. The proposed method achieved an accuracy of 85%, outperforming SVM with an accuracy of 80% and CART with an accuracy of 75.8%.

Zhu et al. (2021) proposed a multi-layer perceptron (MLP) and long short-term memory (LSTM) model to predict the incident clearance duration. The models integrated the traffic incident factors and real-time traffic flow components to predict the residual incident duration. A total of 4041 incidents were collected from Shanghai Zhoughuan Expressway in China to develop the models. The models reported an accuracy of 80.7% and 85.7% for the residual periods of more than 5 minutes and more than 10 minutes, respectively.

Pereira et al. (Pereira et al. 2013) proposed a topic modeling approach where real-time incident reports were utilized to predict the incident clearance duration. Data collected in Singapore over two years were used to develop the models. The longest reported incident duration was 180 minutes. The Radial Basis Function (RBF) reported the least MAE of 15.9 minutes when only the text was used to predict the incident duration, indicating the richness of the report provided at the incident time.

Bokaba et al. (2022) applied five modeling techniques to predict incidents on highways in South Africa. A total of 46,692 incidents over four years from Gauteng province were collected to conduct the study. The Multiple Imputations by Chained Equations (MICE) were utilized to estimate the missing data. PCA technique was applied to the data to reduce the dimensions of the explanatory variables. The results showed that the Random Forest classifier outperformed the other models.

Vlahogianni and Karlaftis (2013) proposed a neural network approach to model incident clearance time. Feature selection using fuzzy entropy similarity was applied to the data to select the important features. A total of 1449 incidents from the urban highway Attiki Odos Tollway in Greece were utilized to train and test the model. Of the 1449 incident records, 65% of data were utilized for training, 15% were utilized for cross-validation, and 20% were utilized for testing. Traffic flow, lane blockage, and vehicles involved were reported to be among the top important features.

The studies reviewed in this section focused on assessing the proposed incident clearance time estimation models on a global level without assessing the performance of these models concerning different incident scenarios, such as the number of blocked lanes or locations. The evaluation metrics were only calculated for the whole testing set, and only the overall performance measure of the models was reported. In other words, the performance bias of the incident clearance time prediction models has not been previously studied, and it is not known if these models performed consistently for all incidents. This paper evaluates the common algorithms for incident clearance time prediction for their performance bias. Such evaluation complements the global metrics often reported for incident clearance time models. This research proposes PCA as a potential solution to the performance bias. To the author's knowledge, all previous studies used one split for the training and testing data. This research applies repeated k-fold cross-validation to present the results of the predictions from twenty different iterations (i.e., splits) for training and testing sets.

METHODS

Various factors influence incident clearance time, making its prediction a complex task. Moreover, the heterogeneity observed in incident clearance time data (Pereira et al. 2013) poses a challenge for the prediction algorithms. In other words, incidents with different clearance time durations (e.g., 15-min versus 90-min clearance time) exhibit distinct characteristics. In essence, the information required to predict incidents lasting over ninety minutes significantly differs from that needed for incidents with durations under fifteen minutes (Li et al. 2015). Unfortunately, a learner is usually not presented with an equal amount of information from the whole distribution of the incident clearance time; thus, these learners encounter limitations that result in biases and poor performance, particularly for severe incidents. Given the limitations of the available incident data, this issue becomes prominent when severe incidents share similar traits with common and frequent incidents. Consequently, the learner becomes distracted, producing higher error levels and biased prediction. In addition to the traditional method, where all relevant independent variables are used to train and test the learners, two dimension-reduction approaches are proposed to reduce the performance bias. The first approach is feature selection (Chand et al. 2022), and the second technique is PCA (Washington et al. 2020).

Performance Bias

The main goal of this paper is to investigate the performance bias in incident clearance time prediction models. In incident clearance time prediction models, performance bias is diagnosed by reviewing the error metrics of the learners across different scenarios. A learner might produce a low error when predicting the clearance time of an incident scenario (e.g., one blocked lane); however, the same learner produces undesirable error metrics for some other incident scenarios (e.g. three blocked lanes). In this case, a learner exhibits bias and favoritism toward a scenario(s). This bias is caused by unfair representation of scenarios in the data. It can also be caused by other resources such as the sampling process and the data collection process. It is worth mentioning that different learners do not exhibit the same bias towards the same scenarios even if they are trained and tested with identical training and testing data sets. For example, learner A may produce the lowest error when predicting the clearance time for incidents on roadway alpha but produces unacceptable errors when predicting clearance time on roadway beta. Learner B, conversely, may perform differently by producing acceptable error metrics for roadway beta incidents clearance time and producing undesirable errors for incidents on roadway alpha. This bias is solely attributed to the algorithms of these learners; in other words, it is attributed to the decision-making process of each learner, especially in complex data such as incident clearance time.

Machine Learning

This paper investigates seven commonly used machine learning algorithms for developing incident clearance time prediction models. These learners are Random Forest (RF) (Breiman 2001), Bayesian Regularization for Feed-Forward Neural Networks (BRNN) (Nguyen and Widrow 1990), K-Nearest Neighbor (KNN) (James et al. 2013), Support Vector Regression (SVR) (Schölkopf et al. 1998), Feed-Forward Neural Networks (NNET) (Anderson 1995), Gaussian Process Regression with Radial Basis Function Kernel (GP) (Rasmussen and Williams

2006), and Extreme Gradient Boosting (XGB) (Chen and Guestrin 2016). To maintain conciseness, the details of these learners are not included.

Dimension Reduction

This research investigates two dimension-reduction methods as a potential solution to reduce the performance bias in clearance time prediction. These two methods are important feature selection and PCA. Performance bias could exist due to the learner's process to learn from the data and make predictions; it could result from the interaction between the learner's algorithm and the data features, or both the learner and interactions between the learner's algorithm and the data could contribute to the bias. This paper suggests changing the composition of the data to reduce the performance bias. Important Feature selection and PCA methods reduce the number of features used to model the clearance time; however, the two methods differ in how they reduce the data dimension. Feature selection eliminates some of the less important features of the data, while PCA replaces the original features with new components that retain the variation in the original features.

Feature Selection

The feature selection technique aims to select the important features with a prominent impact on the model output quality compared to the ground truth incident clearance time. This technique improves the learning performance of the model, reduces computational complexity, builds generalizable models, and avoids overfitting or multicollinearity problems (Chand et al. 2022). Feature selection preserves the important features and their full variation while eliminating less important features. This paper uses a tree-based technique (Kuhn 2008) to identify the important features by ranking the original incident features according to their contribution to the model's performance.

Principal Component Analysis

PCA (Wold et al. 1987) combines the highly correlated variables into one component. It is challenging to interpret the relationship between the independent and dependent variables in high-dimensional data such as incident clearance time. It is also not trivial to determine which independent variables impact the incident clearance time and in which way. The PCA technique provides a solution for this problem by identifying and clustering the independent variables that contain similar information into new components centered around zero. The first component usually combines all the important variables; therefore, this component holds high variation from the original data. Component two holds the second highest variation. Therefore, converting the many independent variables into two new components makes the interpretability of the data easier by reducing the dimensions of the data, which is the motivation for exploring PCA in this paper.

Performance Measure and Evaluation

Root-mean-square-error (*RMSE*) is utilized to quantify the prediction errors of the seven learners, as shown in Eq. (1).

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n}} \quad (1)$$

where, y_i is the ground truth incident clearance time, \hat{y}_i is the predicted clearance time, and n is the total number of incidents in the testing dataset.

Two methods are used to validate and test the models. K-fold cross-validation (Fushiki 2011) with ten folds is used to validate the models and select the values of the hyperparameters for each model, while Repeated k-fold Cross-Validation (Kim 2009) is used to test the models. Twenty distinct models are built using twenty distinct data splits. For each model, 10-fold cross-validation is applied to select the set of hyperparameters. Therefore, twenty RMSE values are reported for each model to account for partitioning bias.

INCIDENT DATA AND STUDY AREA

A total of 1151 incident records that occurred on sections of Interstate 64, Interstate 70, and Interstate 270 in St. Louis, Missouri, USA, are utilized to develop the models (Attallah et al. 2022). Ten incident scenarios are defined for the study area. The scenarios are defined based on the number of blocked lanes, the freeway, and the incident type. The number of blocked lanes included four scenarios: one, two, three, and more than three blocked lanes. Interstate scenarios are I-64, I-70, I-270 incidents. Lastly, incidents are classified as type 1, type 2, and type 3. Type 3 incident is characterized as an accident that involves three or more vehicles. Type 2 incident refers to an accident involving fewer than three vehicles, resulting in the closure of 50% or more of the roadway lanes. Incidents not falling into other categories are designated as type 1 incidents. Traffic data were obtained from roadway sensors.

RESULTS

This paper aimed to investigate the performance bias in incident clearance time prediction models and evaluate two dimension-reduction methods as potential solutions for mitigating the performance bias. This section provides the global and scenario-specific error metrics of learners developed on the basis of FS and PCA methods. The outcomes of FS and PCA with regard to data composition are discussed, and finally, the impact of dimension reduction methods on performance bias is discussed. All the models are compared with base models.

In the base models, all the available independent variables were utilized to develop the learners to predict incident clearance time. Repeated k-fold was applied to create 20 training and testing sets to train and test 20 models for each learner. Figure 1 shows the global and scenario-specific RMSE of the models. Figure 1(a), Figure 1(b), and Figure 1(c) illustrate the global RMSE of these learners. These figures reveal that the PCA offers two advantages compared to the base models and FS models. First, the models developed on the basis of the PCA method produced a lower global RMSE for all learners. Second, the range of the RMSE during the twenty partitioning iterations is narrower, suggesting an improvement in the models' generalization ability. Figure 1(d), Figure 1(e), and Figure 1(f) depict the scenario-specific RMSE for each learner for the three approaches. These plots reveal that: 1) Majority of the models developed on the basis of PCA reduced the RMSE of severe incident scenarios, such as incidents blocking more than three lanes. 2) The seven learners (e.g., RF and NNET) exhibited a comparable error when they were developed on the basis of the PCA method.

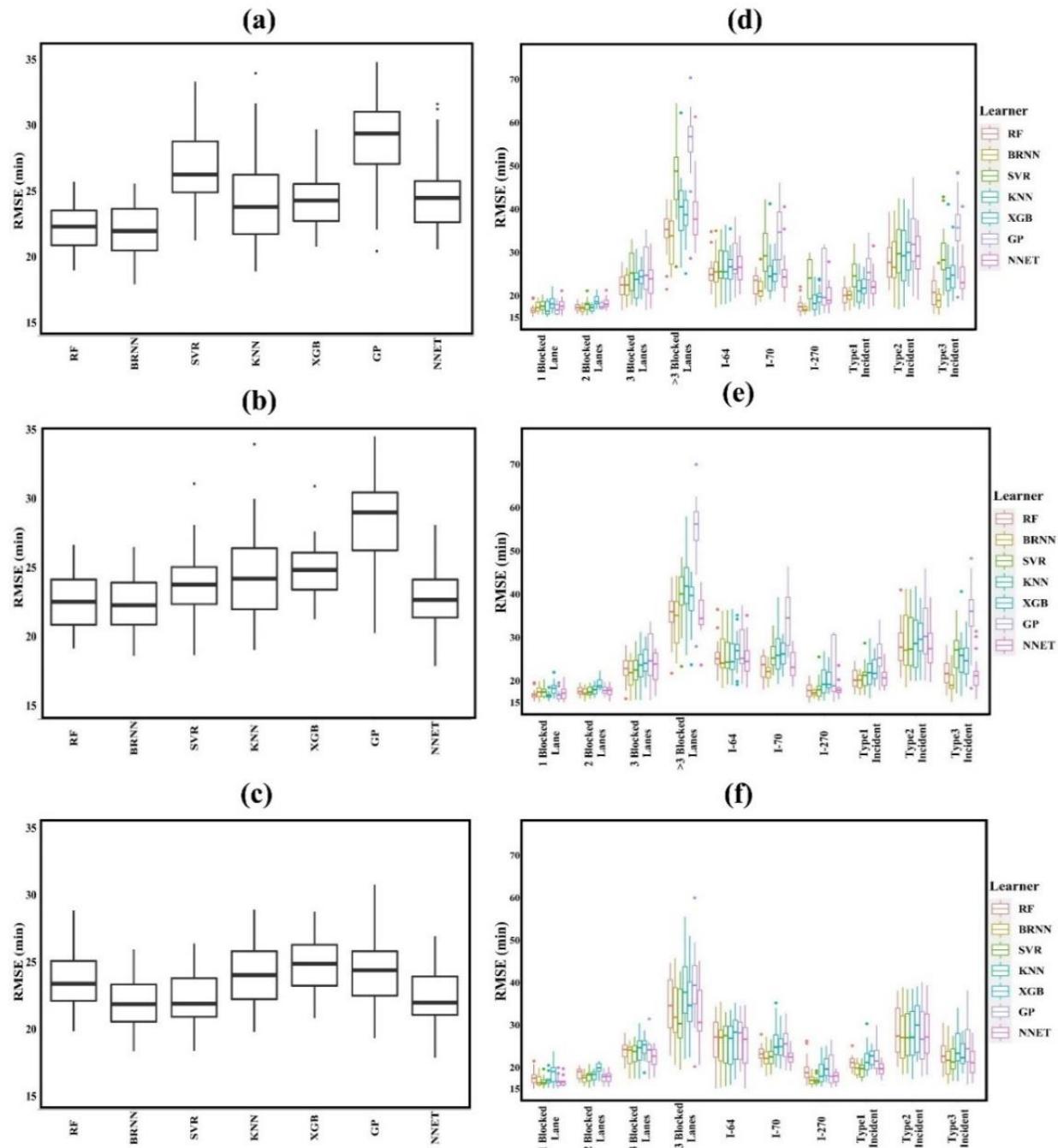


Figure 1. (a) Global RMSE for Base Model. (b) Global RMSE for Feature Selection Model. (c) Global RMSE for PCA Model. (d) Scenario-Specific RMSE for Base Model. (e) Scenario-Specific RMSE for Feature Selection Model. (f) Scenario-Specific RMSE for PCA Model.

Figure 2(a) reports the results from the FS method and illustrates the importance of the features according to their contribution to the prediction process. The independent variables were ranked according to their contribution to the clearance time prediction error. Then, a breakpoint in the variable importance chart was identified, and the variables before this breakpoint were

selected for the model development. In the important feature selection approach, only the first seven features were used as predictor variables to predict incident clearance time.

Figure 2(b) displays the relationship between PC1, PC2, blocked lanes, and incident clearance time. PC1 and PC2 are combinations of independent variables. PC1 ranges from -2 to 6, while PC2 ranges from -6 to 0. The wider range of PC1 indicated that PC1 captures more information. Both components have an average of zero, representing values centered at zero. Despite the skewed nature of the clearance time, Figure 2 (b) shows a clear pattern: smaller PC2 values correlate with longer clearance times. More blocked lanes are associated with lower PC2 values. PC1 also positively correlates with clearance time; as PC1 increases, so does incident clearance time. These two components only captured 37% of the incident clearance time data variation. Therefore, original features, blocked lanes, and location alongside the two principal components were utilized to develop the models.

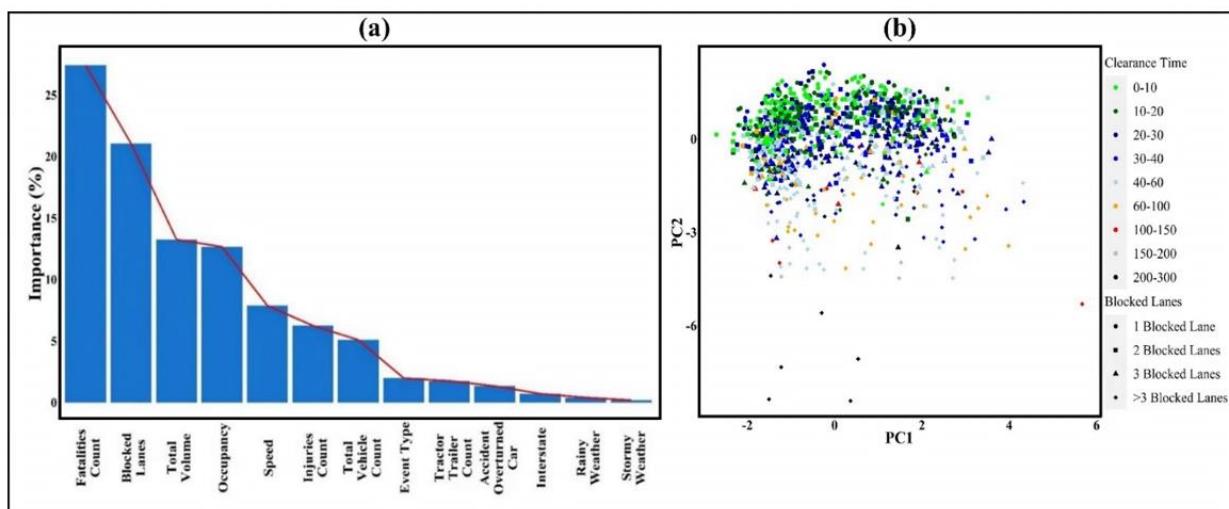


Figure 2. Dimension Reduction Methods Results (a) Feature Importance. (b) Principal components analysis.

Figure 3(a) shows the performance bias in the seven learners, depicting the median RMSE for each learner in the base method across ten scenarios. The RMSE spectrum highlights the performance bias, with learners exhibiting varying biases towards different scenarios. For instance, XGB's RMSE in predicting I-270 incidents' clearance time was higher than SVR's, yet XGB's RMSE for incidents with over three blocked lanes was lower than SVR's. Figures 3(b) and Figure 3(c) depict results from the models developed on the basis of FS and PCA. The models developed on the basis of FS moderately mitigated the performance bias but particularly mitigated scenarios with inconsistent learner performance. On the other hand, models developed using the data reduced by PCA significantly improved global error and reduced the performance bias. For instance, applying PCA reduced GP's median RMSE for scenarios with more than three blocked lanes from 56.8 to 39.4 minutes, a 17.4-minute improvement. SVR's RMSE for the same scenario decreased from 48.8 to 30.3 minutes, a 38% enhancement. GP, SVR, and NNET models benefited most from dimension-reduction methods, with average RMSE reductions of 4.9, 4.7, and 2 minutes, respectively.

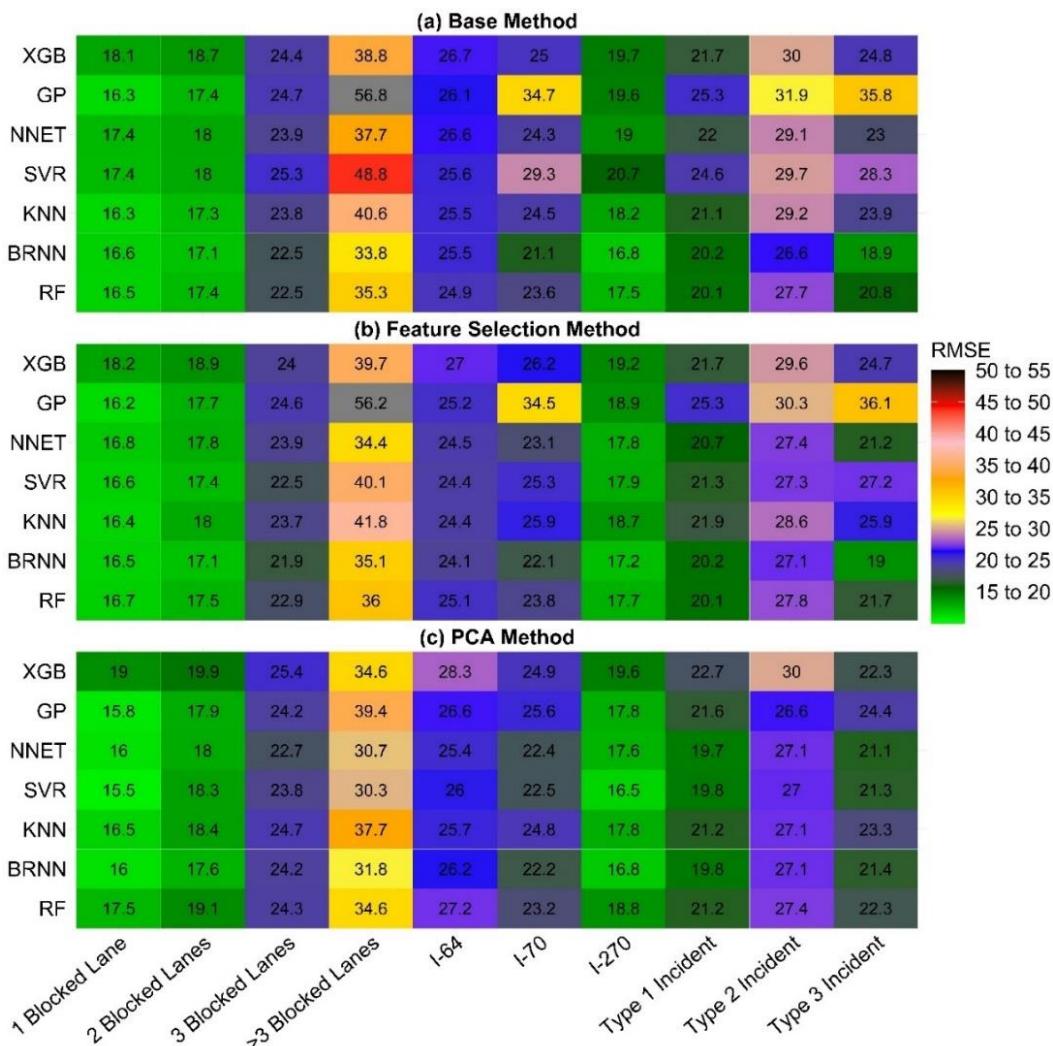


Figure 3. The performance bias heatmap across all scenarios and learners.

CONCLUSIONS

Incident clearance time prediction models inform transportation agencies' strategies for responding to and mitigating incident impacts. Accurate and reliable incident clearance time estimates lead to the success of the transportation agencies' efforts in responding to an incident. On the other hand, inaccurate estimates may compound the incident's impacts. A potential contributing factor to inaccurate clearance time estimates is the performance bias in the learners chosen to develop the prediction models. This paper investigated the performance bias in seven commonly applied learners for incident clearance time prediction: RF, BRNN, SVR, KNN, XGB, GP, and NNET. In addition to reviewing the global error measure obtained from the models for clearance time prediction, this paper reviewed the scenario-specific performance measure. The results demonstrated that the global error measures do not represent the model's actual performance for all scenarios. For example, the BRNN learner produced the minimum global error for the incident clearance time but did not produce the minimum error for all scenarios. RF outperformed the BRNN in three out of ten incident scenarios. Additionally, it was

demonstrated that learners do not exhibit the same biases toward the same scenarios despite being presented with the same data to learn from.

The primary objective of this paper was to highlight and analyze performance bias in incident clearance time prediction models. Two dimension-reduction methods, important feature selection, and principal component analysis, were investigated as potential approaches for mitigating the performance bias in incident clearance time prediction models. The results obtained from a case study revealed that both methods contributed to mitigating the performance bias of the learners in general; however, the learners developed on the basis of the PCA method were more effective in mitigating the performance bias. The learners developed using outputs of the PCA method improved the error when predicting clearance time of severe scenarios, such as incidents with more than three blocked lanes, compared to base learners. In addition to mitigating the performance bias, the learners developed on the basis of the PCA method improved the models' global error metric. Furthermore, these models harmonized the error metric for all learners across all scenarios.

REFERENCES

- Abdel-Aty, M. A., and Radwan, A. E. (2000). "Modeling traffic accident occurrence and involvement." *Accident Analysis & Prevention*, 32(5), 633-642.
- Adler, M. W., Van Ommeren, J., and Rietveld, P. (2013). "Road congestion and incident duration." *Economics of transportation*, 2(4), 109-118.
- Anderson, J. A. (1995). *An introduction to neural networks*, MIT press.
- Attallah, M., Kianfar, J., and Wang, Y. (2022). "Impact of High Resolution Radar-Obtained Weather Data on Spatio-Temporal Prediction of Freeway Speed." *Sustainability*, 14(22), 14932.
- Bokaba, T., Doorsamy, W., and Paul, B. S. (2022). "Comparative study of machine learning classifiers for modelling road traffic accidents." *Applied Sciences*, 12(2), 828.
- Breiman, L. (2001). "Random forests." *Machine learning*, 45, 5-32.
- Chand, S., Li, Z., Alsultan, A., and Dixit, V. V. (2022). "Comparing and Contrasting the Impacts of Macro-Level Factors on Crash Duration and Frequency." *International Journal of Environmental Research and Public Health*, 19(9), 5726.
- Chen, T., and Guestrin, C. "Xgboost: A scalable tree boosting system." *Proc., Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 785-794.
- Fushiki, T. (2011). "Estimation of prediction error by using K-fold cross-validation." *Statistics and Computing*, 21, 137-146.
- Ghosh, B., Asif, M. T., Dauwels, J., Fastenrath, U., and Guo, H. (2018). "Dynamic prediction of the incident duration using adaptive feature set." *IEEE Transactions on Intelligent Transportation Systems*, 20(11), 4019-4031.
- Grigorev, A., Mihaita, A.-S., Lee, S., and Chen, F. (2022). "Incident duration prediction using a bi-level machine learning framework with outlier removal and intra-extra joint optimisation." *Transportation research part C: emerging technologies*, 141, 103721.
- Hojati, A. T., Charles, P., Ferreira, L., and bin Kabit, M. R. (2011). *An analysis of traffic incidents on an Australian urban road network*, Universiti Malaysia Sarawak.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, Springer.

- Jiang, R., Qu, M., and Chung, E. (2014). "Traffic incident clearance time and arrival time prediction based on hazard models." *Mathematical Problems in Engineering*, 2014.
- Kim, J.-H. (2009). "Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap." *Computational statistics & data analysis*, 53(11), 3735-3745.
- Kuhn, M. (2008). "Building predictive models in R using the caret package." *Journal of statistical software*, 28, 1-26.
- Lee, Y. "A Computerized Feature Reduction Using Principal Component Analysis for Accident Duration Forecasting on Freeway." *Proc., Proceedings of the International Conference on Data Science (ICDATA)*, The Steering Committee of The World Congress in Computer Science, Computer ..., 1.
- Li, R., Pereira, F. C., and Ben-Akiva, M. E. (2015). "Competing risk mixture model and text analysis for sequential incident duration prediction." *Transportation Research Part C: Emerging Technologies*, 54, 74-85.
- Nguyen, D., and Widrow, B. "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights." *Proc., 1990 IJCNN international joint conference on neural networks*, IEEE, 21-26.
- Ozbay, K., and Noyan, N. (2006). "Estimation of incident clearance times using Bayesian Networks approach." *Accident Analysis & Prevention*, 38(3), 542-555.
- Pereira, F. C., Rodrigues, F., and Ben-Akiva, M. (2013). "Text analysis in incident duration prediction." *Transportation Research Part C: Emerging Technologies*, 37, 177-192.
- Rasmussen, C., and Williams, C. (2006). *Gaussian processes for machine learning*. (MIT press: Cambridge, MA). mit press: Cambridge ma.
- Salum, J. H., Kodi, J. H., Kidando, E., Alluri, P., and Sando, T. (2023). "Associating incident clearance duration with freeway segment types using hierarchical Bayesian survival model." *Journal of Transportation Engineering, Part A: Systems*, 149(1), 04022114.
- Schölkopf, B., Bartlett, P., Smola, A., and Williamson, R. C. (1998). "Shrinking the tube: a new support vector regression algorithm." *Advances in neural information processing systems*, 11.
- Shang, Q., Tan, D., Gao, S., and Feng, L. (2019). "A hybrid method for traffic incident duration prediction using BOA-optimized random forest combined with neighborhood components analysis." *Journal of Advanced Transportation*, 2019.
- Vlahogianni, E. I., and Karlaftis, M. G. (2013). "Fuzzy-entropy neural network freeway incident duration modeling with single and competing uncertainties." *Computer-Aided Civil and Infrastructure Engineering*, 28(6), 420-433.
- Washington, S., Karlaftis, M. G., Mannering, F., and Anastasopoulos, P. (2020). *Statistical and econometric methods for transportation data analysis*, CRC press.
- Wold, S., Esbensen, K., and Geladi, P. (1987). "Principal component analysis." *Chemometrics and intelligent laboratory systems*, 2(1-3), 37-52.
- Zhu, W., Wu, J., Fu, T., Wang, J., Zhang, J., and Shangguan, Q. (2021). "Dynamic prediction of traffic incident duration on urban expressways: A deep learning approach based on LSTM and MLP." *Journal of intelligent and connected vehicles*, 4(2), 80-91.

Hierarchical Classification—Regression (HiClassR) to Improve Incident Clearance Time Prediction

Mustafa Attalah¹ and Jalil Kianfar²

¹Ph.D. Candidate, Dept. of Civil, Computer, and Electrical Engineering, Saint Louis Univ., Saint Louis, MO. Email: mustafa.attallah@slu.edu

²Associate Professor, Dept. of Civil, Computer, and Electrical Engineering, Saint Louis Univ., Saint Louis, MO (corresponding author). Email: jalil.kianfar@slu.edu

ABSTRACT

Incident clearance time prediction is a key task in Advanced Transportation Management and Advanced Traveler Information Systems, as it informs mitigation and response strategies adopted by transportation agencies. Incident clearance time prediction is not a trivial task given that the duration of an incident is influenced by various factors, some of which are difficult to measure or not measured at all. Such constraints in the data might limit the accuracy of the models developed for predicting clearance time. Over the past decades, researchers have employed various statistical methods and machine learning approaches to overcome these limitations and improve the accuracy of the incident clearance time prediction models. While these efforts have successfully improved the overall accuracy of models, very few studies have focused on assessing and improving the generalization ability of incident clearance time prediction models. This paper proposes the hierarchical classification-regression (HiClassR) framework to improve the predictive performance of common machine-learning algorithms in predicting incident clearance time. HiClassR is a framework for developing a bi-level model in which the incidents are first assigned to an incident class; then, a class-specific model estimates the incident clearance time. While the HiClassR method is expected to improve error measures such as the RMSE over the entire dataset, it is also expected to improve the model's generalization ability across various incident scenarios. In a case study, the HiClassR method was applied to six conventional predictive methods to develop incident clearance time prediction models for St. Louis, MO, USA. The case study results indicate that the HiClassR model consistently outperforms the conventional base models and effectively generalizes the overall prediction power of the model to various incident categories of different conventional predictive methods. On average, the preferred HiClassR model reduced the RMSE from 22.7 to 17.6 min for the random forest model, from 22.5 to 19 min for Bayesian regularization neural networks, from 23.7 to 16.9 min for support vector machines, from 25.2 to 21.1 min for K-nearest neighbors, from 24.5 to 19.7 min for XGB, and from 23.9 to 18.3 min for neural networks.

INTRODUCTION

Transportation management centers consider the expected clearance time of roadway incidents in their day-to-day operations. Over the past decades, researchers have explored and proposed various models to improve incident clearance time prediction models (Garib et al. 1997; Grigorev et al. 2022; Khattak et al. 2016; Tang et al. 2020). Predicting incident clearance time is not trivial, given that incident clearance time data is often skewed and imbalanced with complex and unique incidents. This paper builds on the work of Weng et al. (2015), Kuang et al.

(2019), and Grigorev et al. (2022) to develop Category Aware Models and Hierarchical Classification-Regression models to estimate incident clearance time. However, this paper is distinct from the work of Kuang et al. (2019) and Grigorev et al. (2022), aiming not only to reduce the overall incident clearance prediction error but also to improve the model performance across different incident scenarios. An incident scenario refers to a specific set of circumstances characterizing a roadway incident. For example, incidents that result in three-blocked lanes, or incidents that occur on a specific roadway. Additionally, this paper aims to explore the partitioning bias in the development of incident clearance time prediction models and explores repeated k-fold cross-validation to mitigate this bias. The remainder of this paper is organized as follows: 1) A review of the state of the art in incident clearance time modeling and the application of bi-level models is provided. 2) Data and case study is introduced. 3) The methods applied in this work, including the machine learning algorithms, modeling techniques, and assessment metrics, are explained. 4) The results obtained from all models are presented in the next chapter. 5) Lastly, a conclusion is drawn from this work.

LITERATURE REVIEW

Over the past two decades, researchers have utilized various statistical and machine learning methods to analyze and estimate incident clearance time. This section provides an overview of state-of-the-art incident clearance time prediction models and the application of bi-level models in ITS incident management-related subsystems.

Li et al. (2018) conducted an in-depth review of the incident clearance time estimation methods. They noted the success of machine learning methods in predictive modeling across various disciplines predicting in various disciplines and suggested machine learning methods be explored for incident clearance estimation.

Kuang et al. (2019) proposed a bi-level incident clearance time estimation model. First, a cost-sensitive Bayesian network classification was utilized to classify incidents as short-duration or long-duration based. Then, the k-nearest neighbor regression was applied to predict the incident clearance time. Incidents with a duration of less than 30 min were considered short, and the rest of the incidents were considered long. Incident data collected in Xiamen City, China, were used to develop the models and included 27712 incident records for training and 10000 incidents for testing the models. The MAPE varied between 13.6% and 60.3%, depending on the tested dataset. The MAPE varied between 13.6% and 60.3%, depending on the tested dataset.

Weng et al. (2015) proposed a cluster-based lognormal distribution model to predict incident clearance time. A lognormal distribution model incorporates heavier tails, allowing for a greater likelihood of extreme events. Initially, incidents were classified into three groups using a decision tree. Then, a dummy variable representing the incident cluster was added to the regression model as a predictor variable. Introducing the cluster-based dummy variable reduced the prediction error and resulted in a MAPE of 42.3%.

Grigorev et al. (2022) developed a bi-level machine learning-based incident clearance time prediction model for arterial and motorways in Sydney, Australia, and San Francisco, California, USA. A classification model was first used to split incidents into short-term and long-term incidents. Various durations were explored for splitting the incidents into short-term and long-term, and it was determined that a 40-45 min threshold provides the best split. Then, a regression model was developed to predict clearance time for each class separately. The authors found that XGB is the best incident clearance time classification model. XGB was also the best overall

regression model. However, LGBM resulted in the lowest MAPE for San Francisco data. It is worth mentioning that San Francisco's sample size is significantly larger than the sample from Sydney. The total MAPE varied between 33.19% and 229.53% for the best and worst-performing models.

The previous research primarily considered the global performance of a model when evaluating models. This paper builds on previous research and introduces a scenario-specific analysis of incident clearance time prediction models. Previous research did not consider the impact of partitioning bias on splitting the data into training and testing sets. To mitigate the impact of partitioning bias, this paper suggests applying repeated k-fold cross-validation to build more than one training set and one testing set. This enables the researchers to understand the behavior of the machine learning algorithms in predicting incident clearance time and their generalization ability.

METHODOLOGY

This section provides an overview of Class Aware Model framework and Hierarchical Classification and Regression framework for developing incident clearance time predictions models.

Class Aware Model

The Class Aware Model (CAM) consists of a classification and a regression model. The classification model predicts the overall category of the incident duration. Then, the incident category and features are presented to a regression model to estimate the incident clearance time. Several sets of pre-established thresholds are used to define the incident duration categories. For example, set 1 labels incidents with clearance time shorter than 41 minutes as class 1, and labels incidents with clearance time longer than 40 minutes as class 2. Set 2 labels the incidents to three categories: incidents with clearance time shorter than 31 minutes, incidents with clearance time between 31 and 60 minutes, and incidents with clearance time longer than one hour. The TMC defines the thresholds in the subsets based on factors such as internal agency incident response performance metrics, distribution of the clearance time data, and motorists' perception of delays resulting from incidents. The CAM framework is developed based on work of Weng et al. (2015).

The incident clearance time is typically imbalanced and right-skewed. The motivation for CAM is to address the limitations of the imbalanced data for developing regression models by introducing a new variable that indicates the expected category of the incident clearance time. Additionally, an oversampling method is utilized to oversample the minority class with replacement to balance the training set size in each incident duration category such that all categories have the same number of data points (Mohammed et al. 2020).

Hierarchical Classification and Regression Model

Given the skewness of the incident clearance time, developing a dedicated regression model for each part of the distribution can improve the generalization ability of that regression model. Therefore, a Hierarchical Classification and regression (HiClassR) model is proposed to develop incident clearance time prediction models. The development of HiClassR begins with pre-

processing the incident data. In the pre-processing task, the incident data are split into m quantiles based on incident clearance time, and data in each quantile are labeled as class 1 to class m . In the pre-preprocessing task, the incident data are split into a Classification Training Set (CTrS) and a Classification Testing Set (CTeS). Then, the HiClassR model is developed in the following two tasks. First, a supervised classification algorithm is used to classify incident data into classes 1 to m . This model is trained with CTrS and tested with CTeS. Next, CTeS is split into m subsets according to predictions made by the classification algorithm. In the second task, each subset is further partitioned into Regression Training Set_i (RTrS_i) and Regression Testing Set_i (RTrS_i), which are used to train and test the regression model for class i . The regression models are unique for each class. The HiClassR model development process begins with exploring two quantiles. The number of quantiles is incrementally increased until increasing the number of quantiles does not improve the model error metric or until further splitting the data to additional quantiles would result in insufficient sample sizes for clusters. The HiClassR framework follows the approach proposed by Kuang et al. (2019) and Grigorev et al. (2022).

ML Algorithms

Six common machine learning algorithms are explored to develop the classification layer and regression layer of the HiClassR model. These algorithms include Random Forest (RF) (Breiman 2001), Bayesian Regularization for Feed-Forward Neural Networks (BRNN) (MacKay 1992), K-Nearest Neighbor (KNN) (James et al. 2013), Support Vector Regression (SVR) (Schölkopf et al. 1998), Feed-Forward Neural Networks (NNET) (Anderson 1995), and Extreme Gradient Boosting (XGB) (Chen and Guestrin 2016). For brevity, details of these methods are not provided.

Partitioning and Repeated K-fold Cross-Validation

The k-folds cross-validation method is utilized for training the classification and regression models. The data are uniformly and randomly split into k folds (Rodriguez et al. 2009) such that all the folds are approximately the same size. Then, the model is trained and validated k times with various combinations of folds. In the first step, the first fold is set aside to validate the model, and the remaining $k-1$ folds are used to train the model. The second fold is reserved for validation in the second step, and the remaining folds are utilized to train the model. This process continues until every fold is utilized to validate the model. The average error for the models developed in the k iterations is reported as the overall training error; the testing error is calculated similarly. This paper proposes ten-fold cross-validation for the development of the models. Additionally, repeated k-fold cross-validation (Burman 1989) is applied thirty times to reduce the likelihood of partitioning bias. Partitioning bias is a concern in small datasets, especially those with unbalanced data. A random seed is applied to split the entire dataset into training and testing sets. Then, k-fold cross-validation is applied to the training test. This process is repeated thirty different times (i.e., thirty iterations) with thirty different random seeds. Rather than the conventional one value of the error metric, the repeated k-fold cross-validation produces thirty values of the error metric, which helps avoid data partitioning bias (Bergmiller et al. 2017).

Error Metrics

The performance of the incident prediction models is evaluated at the global and scenario-specific levels. As the name implies, the global performance metrics were applied to the entire

testing dataset and represent a model's overall performance for predicting the entire incident clearance time. The scenario-specific performance metrics are used to assess the performance of a model for various incident scenarios. Scenario-specific performance metrics are applied to investigate the consistency of model predictive performance across various incident scenarios. The global and scenario-specific performance of the models is assessed based on Root-Mean-Squared-Error (RMSE) and Mean Absolute Percentage Error (MAPE) as provided in Eq. (1) and Eq. (2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (2)$$

where, y_i is the observed (i.e., ground truth) clearance time for incident i , \hat{y}_i is the predicted clearance time for incident i , and n is the number of observations. The model performance is reported based on a dataset that includes RTeS₁ to RTeS_m.

In addition to these error metrics, Pearson's correlation coefficient (Cohen et al. 2009), also known as product-moment correlation coefficient, is used to calculate the correlation between the predicted incident clearance time and the ground truth incident clearance time. When a model provides similar correlation values between prediction and ground truth data across various scenarios, it indicates that the model is able to generalize the knowledge fairly and does not favor a group of scenarios over other scenarios.

Error metrics of the CAM and HiClassR models are compared with base models. Base models are machine learning regression models that predict incident clearance time based on incident features. These models are not bi-level and are developed with a standard training and testing split. In other words, the base models are the conventional machine learning models developed to predict incident clearance time.

CASE STUDY AND DATA

Incident and traffic flow data from three freeway corridors in St. Louis, Missouri, United States, were used to develop the models. These corridors include a 63-kilometer section of Interstate 64, a 60-kilometer section of Interstate 70, and a 31-kilometer section of Interstate 270. Figure 1 illustrates freeway corridors and the spatial distribution of the incidents. In total, 1,151 incidents were reported over two years. The average clearance time of incidents was 32 minutes, and 75% were cleared within 41 minutes, with less than 1% requiring more than three hours to be cleared.

The incident logs included information about the incident type, location and time of the incident, number of blocked lanes, number of vehicles involved, number of tractor-trailers involved, weather conditions (i.e., clear or inclement weather), number of injuries, if the injuries included a fatality, and if an over-turned vehicle was involved. The traffic data included speed, volume, and occupancy (Attallah et al. 2022). The TMC classified the incidents into incident type 1, incident type 2, and incident type 3. A collision that involved three or more vehicles was classified as incident type 3. A collision that did not meet the definition of incident type 3 was classified as either incident type 1 or 2. Incident type 2 involved less than three vehicles and

resulted in the closure of 50% or more of the roadway lanes. Every other type of collision was classified as incident type 1.



Figure 1. Saint Louis, Missouri, USA Study Area.

RESULTS

This section discusses the results obtained from the CAM and HiClassR models. First, the global error metrics for a base model, CAM, and HiClassR model are evaluated, and a preferred model is selected. Then, the scenario-specific error metrics for the preferred models are investigated. Finally, the performance of the models during the repeated k-fold process is presented.

Global Performance

In this section, the global performance for the base, CAM-2, CAM-3, HiClassR-2, and HiClassR-3 models in terms of global MAPE obtained from 30-repeated k-folds cross-validation is reported. In CAM-2 and HiClassR-2 models, incident data are split into two classes; and in CAM-3 and HiClassR-3 models, incident data are split into three classes.

The MAPE for these models is depicted in Figure 2 and includes RF, BRNN, SVR, KNN, XGB, and NNET learners. Among the base models, SVR and RF learners outperformed the other learners in terms of both median error and lowest error. The median errors represent the median of error obtained from the 30-repeated k-folds, and the lowest error represents the results obtained from the best iteration of the repeated k-folds iteration. The base models served as a benchmark for evaluating the CAM and HiClassR models and are shown in both Figure 2 (a) and Figure 2 (b) to facilitate comparison.

Figure 2 (a) illustrates the global MAPE of CAM-2 and CAM-3 models. CAM-2 and CAM-3 models utilized RF as a classifier to assign a class to each incident. CAM-2 and CAM-3 slightly improved the global MAPE. Additionally, there were no significant improvements from increasing the number of splits in the CAM from two to three for all six learners. With regards to learners' performance, the lowest MAPE was produced by the CAM-2 random forest model.

HiClassR-2 places the incidents in two classes, and HiClassR-3 places the incidents in three classes. HiClassR-2 and HiClassR-3 models utilized RF as a classifier to assign a class to each incident. Figure 2 (b) depicts the MAPE of the base, HiClassR-2, and HiClassR-3 models developed with the six learners to serve as the second-level regression model. The results reported in Figure 2 (b) were obtained from thirty repeated k-fold cross-validation for each

learner. Overall, HiClassR-2 and HiClassR-3 outperformed the base models. For example, the MAPE of the best HiClassR-2 KNN was 60.01%, a 40.55% improvement compared to the best base KNN model with a MAPE of 100.65%. Among the HiClassR-2 models, the SVR learner produced the best results.

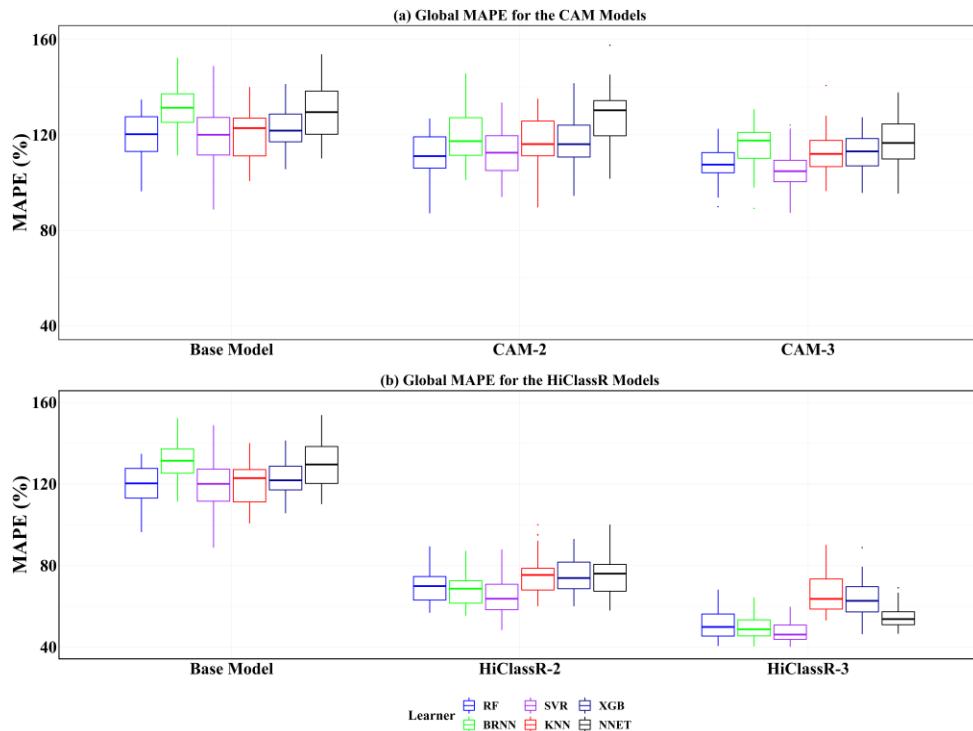


Figure 2. Global Performance of the Base, CAM, and HiClassR Models.

Furthermore, HiClassR-3 learners outperformed HiClassR-2 learners. For example, HiClassR-3 SVR learner reduced the MAPE by 63.85% compared to the base SVR learner, while the HiClassR-2 SVR reduced the MAPE by 52.31% compared to the base SVR learner. In summary, the HiClassR-2 and HiClassR-3 learners outperformed the base, CAM-2, and CAM-3 models, with the HiClassR-3 learners providing the best results. On average, the HiClassR-2 reduced the median global MAPE of the 30-repeated k folds for the six learners by 54.47% compared to the base learners. The HiClassR-3 reduced the median global MAPE of the 30-repeated k folds across the six learners by 71.55% in comparison to the base learners. Consequently, the HiClassR-3 models were selected as preferred models, and their scenario-specific performance is further investigated.

Scenario-Specific Performance

In this section, the performance of HiClassR-3 learners is compared with the performance of base learners for 1-, 2-, 3-, and more than three blocked-lane(s) scenarios, roadway scenarios, and incident-type scenarios.

Figure 3 shows the MAPE for the six learners across blocked lanes and location scenarios. Overall, the HiClassR models outperformed the base model across all scenarios and all learners. In terms of the lowest error of the 30 repeated k-folds, HiClassR-3 outperformed the base model

in all scenarios across the six learners. With regards to the median MAPE of the 30 repeated k-folds models, the HiClassR-3 model also outperformed the base model in all the 42 scenarios. The black circular, and triangular markers on the base and HiClassR-3 MAPE plots indicate the performance metric corresponding to the learner with the lowest global MAPE. In HiClassR-3 learners, the scenario-specific performance metrics of the learner with the lowest global MAPE are within the first two quartiles for 81% of the scenarios.

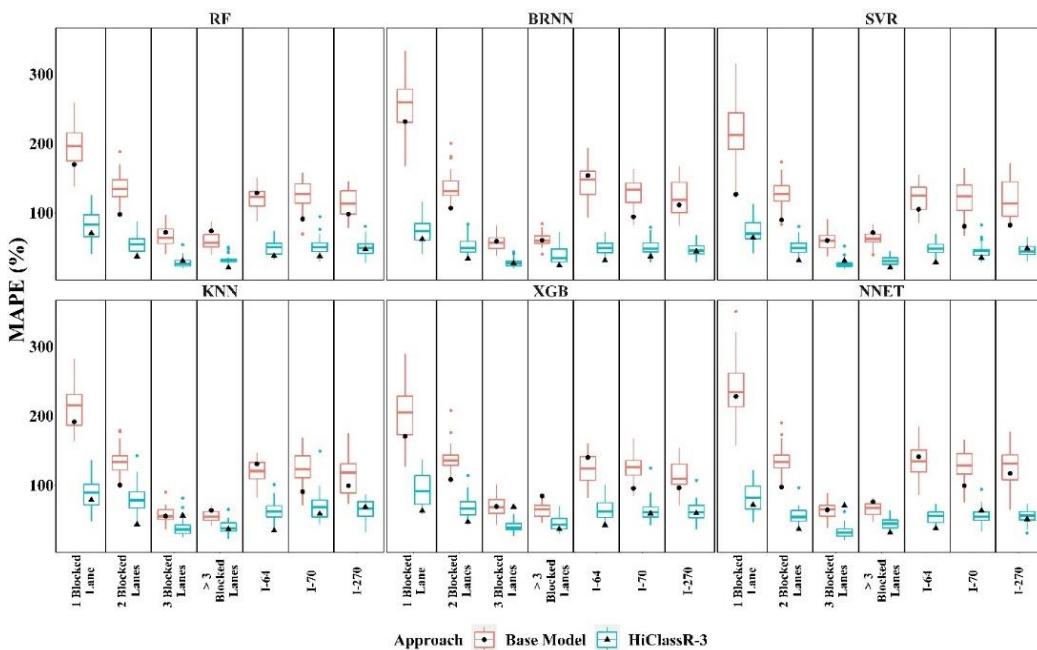


Figure 3. Scenario-Specific Performance of the Base and HiClassR-3 Models.

Figure 4 compares the HiClassR-3 model with the base model across ten scenarios and six learners. Two metrics are reported in Figure 4: RMSE and the product-moment correlation coefficient between the predicted and ground truth clearance time. A smaller square marker indicates lower RMSE, and a marker with a lighter color indicates higher Pearson's correlation coefficient. The HiClassR-3 learners consistently outperformed the base learners in terms of RMSE. Generally, the one blocked-lane scenario and the I-64 scenario gained the most improvements across all the learners. In terms of learners' performance, the SVR outperformed the other five learners.

With regard to product-moment correlation coefficient, the base learners exhibited lower correlation values in comparison to the HiClassR-3 model. Additionally, a narrower spectrum of the correlation coefficient was observed for the HiClassR-3 learners, which indicates the learners have a similar generalization ability across the ten scenarios. In other words, the learners in the HiClassR-3 model did not favor one scenario over another. Furthermore, the SVR demonstrated the correlation coefficient with the slightest variance across different scenarios.

Analysis of the Repeated K-fold cross-validation

As mentioned in the literature review section, the previous research on developing incident clearance time prediction models primarily used one random seed to partition the available data

into training and testing sets. Few studies have utilized k-fold cross-validation to train the models; however, the k-fold cross-validation has been applied to a single split of training data. Using a single random seed to split data into training and testing sets and the k-fold cross-validation might mitigate the risk of overfitting for the data split. However, it does not address the biases that might be introduced to the models as a result of a specific split.

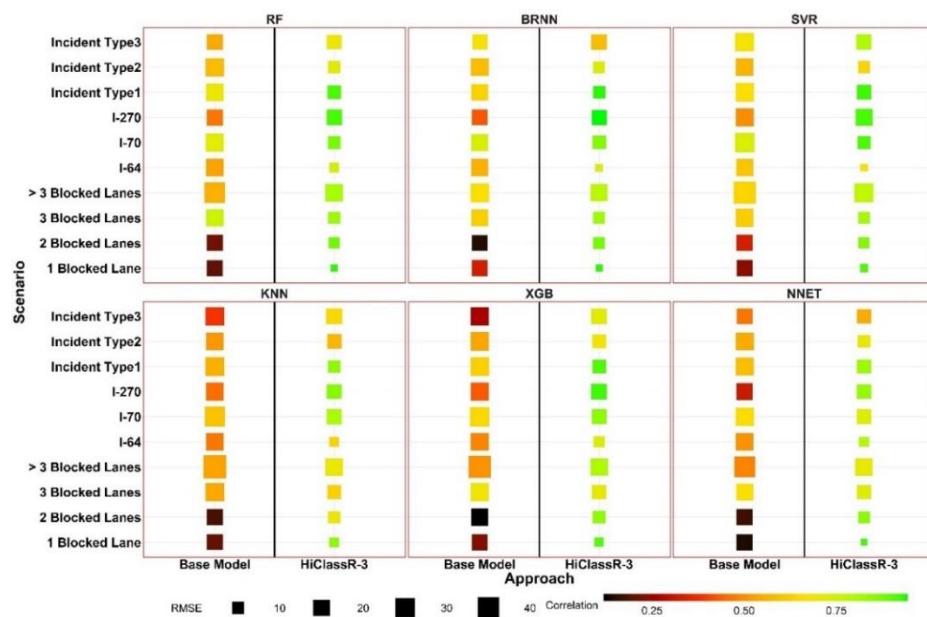


Figure 4. RMSE and product-moment correlation coefficient of the Base and HiClassR-3 Models.

To better assess the generalization ability of a model for a specific dataset, repeated k-fold cross-validation is proposed in this paper. Figure 5 provides the global MAPE for six learners when 30 distinct randomly generated seeds are used to split the data into training and testing sets.

Several findings are reported in Figure 5. First, exploring only one data split for creating training and testing datasets might not yield the best model. Second, the six learners exhibited similar trends across the 30 splits, confirming that the data split influenced the learners' performance. For the same reason, similar trends are observed across the base model and the HiClassR-3 models. It is worth noting that the lowest global MAPE for the base model was obtained from iteration 17, while the lowest global MAPE for the HiClassR-3 model was obtained from iteration 20. In both base and HiClassR-3 models, the SVR learner outperformed the other learners. It is also noted that a learner's performance depends on the splitting of the data. For example, even though SVR produced the lowest error among the base learners in iteration 17, it produced the highest error among the base learners in iteration 19. Thus, the development of the model with the lowest error requires selecting a suitable learner and a suitable data split.

Relative Performance

The violins in Figure 6 exhibit the variation of the ten scenario-specific MAPE for each learner, while the red circular marker reports the global MAPE of a learner. Figure 6 is

developed from results obtained from iteration 17 for the base model and iteration 20 for the HiClassR-3 model.

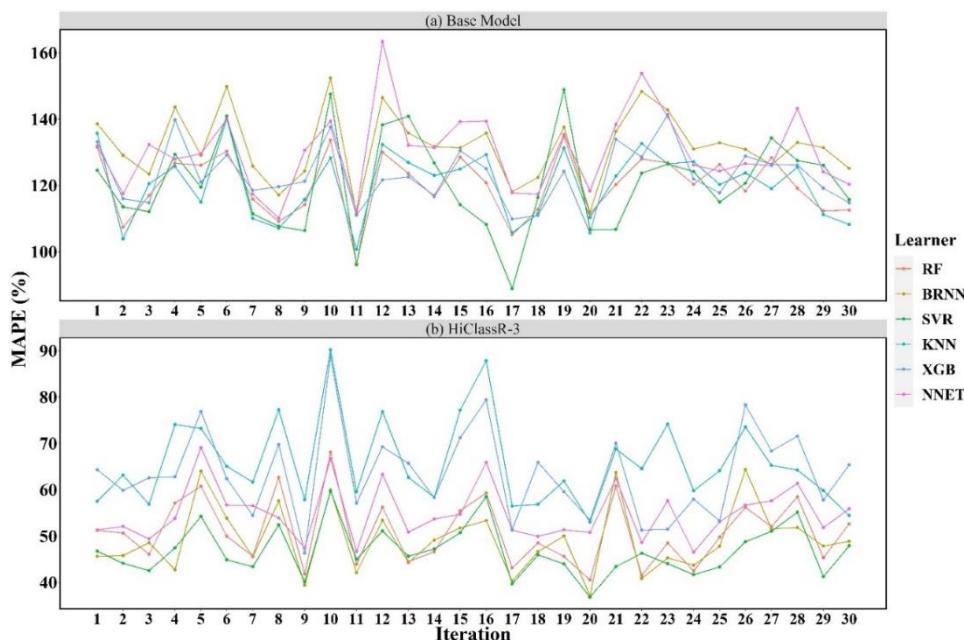


Figure 5. MAPE of Base and HiClassR-3 models over thirty iterations.

There are several trends observed in Figure 6. First, the HiClassR-3 violins are shorter than the base model violins, which indicates that the HiClassR-3 learner performs more consistently across the scenarios. Second, the global MAPE produced by the HiClassR-3 learner is more representative of a learner's performance across all scenarios in comparison to the base learners.

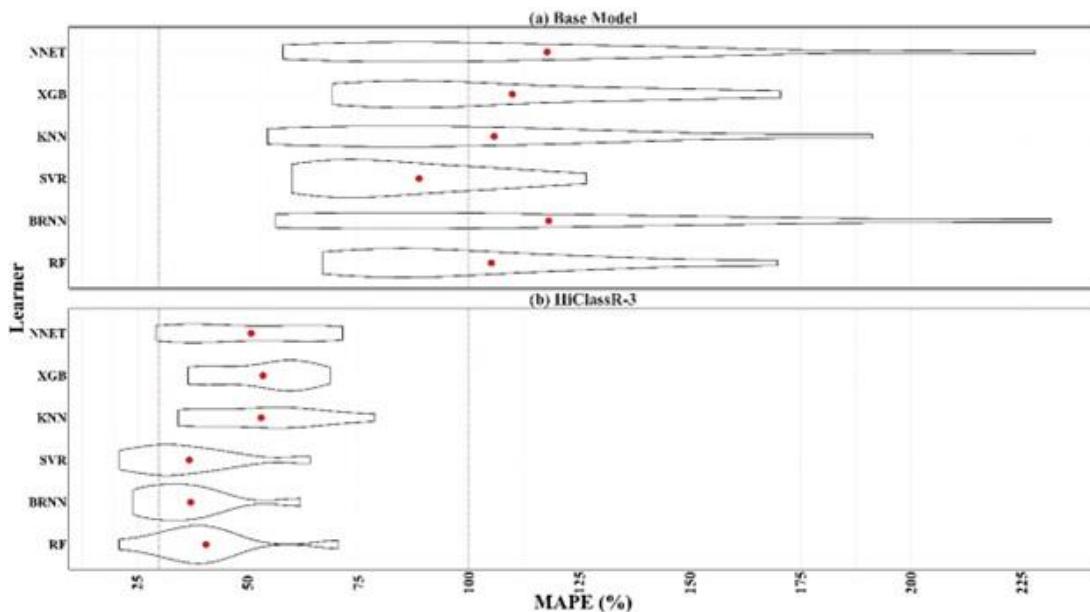


Figure 6. Variation in the Scenario-Specific MAPE and Global MAPE.

Third, in the HiClassR-3 model, the highest scenario-specific MAPE was approximately 75%. In contrast, the highest scenario-specific MAPE for base learners was approximately 225%, demonstrating that HiClassR-3 successfully reduced the scenario-specific error.

CONCLUSION

Intelligent Transportation Systems rely on predicted incident clearance times to implement strategies to mitigate incidents' impact and communicate the expected roadway delays to travelers. Roadway incidents include a wide range of severity and collision types, resulting in a broad range of incident clearance time. Additionally, the various incident types do not occur with similar frequency, yielding a skewed clearance time distribution and imbalanced data. To address the potential impact of the imbalanced data on clearance time prediction, this paper proposed the Category Aware Model (CAM) and the Hierarchical Classification-Regression (HiClassR) model. The HiClassR model aims to improve the incident clearance time prediction in two aspects. First, this approach aims to improve the overall incident clearance time prediction errors in comparison to the conventional incident clearance time prediction model by building quantile-specific prediction models. Second, the HiClassR model improves the quality of incident clearance time predictions across various incident scenarios. Assessment of the incident clearance time predictions model at the scenario level is crucial to developing a reliable model that can predict clearance time for different incident scenarios with comparable error metrics.

CAM and HiClassR models were developed in a case study to predict incident clearance time in St. Louis, Missouri, USA. The results obtained from the models were compared to base conventional machine learning models. The models were evaluated in terms of their global performance as well as their scenario-specific performance. Additionally, the paper applied repeated k-fold cross-validation in the model development process to demonstrate the impact of partitioning bias on the models' performance. This validation approach enables the selection of a prediction algorithm in conjunction with a training-testing split those results in the best model. The results indicate that HiClassR models outperform both CAM and base models in terms of global and scenario-specific performance measures. Additionally, the repeated k-fold cross-validation revealed that one split of the data to train and test a learner might not yield the best model. Furthermore, a learner's performance might vary depending on the training-testing splits.

ACKNOWLEDGMENTS

Authors would like to thank Missouri Department of Transportation Gateway Guide Transportation Management Staff for providing the incident and traffic sensors data.

REFERENCES

- Anderson, J. A. (1995). *An introduction to neural networks*, MIT press.
- Attallah, M., Kianfar, J., and Wang, Y. (2022). "Impact of High Resolution Radar-Obtained Weather Data on Spatio-Temporal Prediction of Freeway Speed." *Sustainability*, 14(22), 14932.
- Bergmiller, T., Andersson, A. M., Tomasek, K., Balleza, E., Kiviet, D. J., Hauschild, R., Tkačik, G., and Guet, C. C. (2017). "Biased partitioning of the multidrug efflux pump AcrAB-TolC underlies long-lived phenotypic heterogeneity." *Science*, 356(6335), 311-315.

- Breiman, L. (2001). "Random forests." *Machine learning*, 45, 5-32.
- Burman, P. (1989). "A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods." *Biometrika*, 76(3), 503-514.
- Chen, T., and Guestrin, C. "Xgboost: A scalable tree boosting system." *Proc., Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 785-794.
- Cohen, I., Huang, Y., Chen, J., Benesty, J., Benesty, J., Chen, J., Huang, Y., and Cohen, I. (2009). "Pearson correlation coefficient." *Noise reduction in speech processing*, 1-4.
- Garib, A., Radwan, A., and Al-Deek, H. (1997). "Estimating magnitude and duration of incident delays." *Journal of Transportation Engineering*, 123(6), 459-466.
- Grigorev, A., Mihaita, A.-S., Lee, S., and Chen, F. (2022). "Incident duration prediction using a bi-level machine learning framework with outlier removal and intra-extra joint optimisation." *Transportation research part C: emerging technologies*, 141, 103721.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, Springer.
- Khattak, A. J., Liu, J., Wali, B., Li, X., and Ng, M. (2016). "Modeling traffic incident duration using quantile regression." *Transportation Research Record*, 2554(1), 139-148.
- Kuang, L., Yan, H., Zhu, Y., Tu, S., and Fan, X. (2019). "Predicting duration of traffic accidents based on cost-sensitive Bayesian network and weighted K-nearest neighbor." *Journal of Intelligent Transportation Systems*, 23(2), 161-174.
- Li, R., Pereira, F. C., and Ben-Akiva, M. E. (2018). "Overview of traffic incident duration analysis and prediction." *European transport research review*, 10(2), 1-13.
- MacKay, D. J. (1992). "Bayesian interpolation." *Neural computation*, 4(3), 415-447.
- Mohammed, R., Rawashdeh, J., and Abdullah, M. "Machine learning with oversampling and undersampling techniques: overview study and experimental results." *Proc., 2020 11th international conference on information and communication systems (ICICS)*, IEEE, 243-248.
- Rodriguez, J. D., Perez, A., and Lozano, J. A. (2009). "Sensitivity analysis of k-fold cross validation in prediction error estimation." *IEEE transactions on pattern analysis and machine intelligence*, 32(3), 569-575.
- Schölkopf, B., Bartlett, P., Smola, A., and Williamson, R. C. (1998). "Shrinking the tube: a new support vector regression algorithm." *Advances in neural information processing systems*, 11.
- Tang, J., Zheng, L., Han, C., Liu, F., and Cai, J. (2020). "Traffic incident clearance time prediction and influencing factor analysis using extreme gradient boosting model." *Journal of Advanced Transportation*, 2020.
- Weng, J., Qiao, W., Qu, X., and Yan, X. (2015). "Cluster-based lognormal distribution model for accident duration." *Transportmetrica A: Transport Science*, 11(4), 345-363.

An Exploratory Analysis of Factors Affecting Adoption of Electric Vehicles

Apurva Pamidimukkala, Ph.D.¹; Sharareh Kermanshachi, Ph.D., P.E., F.ASCE²;
Jay Michael Rosenberger, Ph.D.³; and Greg Hladik, Ph.D.⁴

¹Assistant Professor of Research, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: apurva.pamidimukkala@mavs.uta.edu

²Associate Vice Chancellor and Associate Dean of Research, Penn State Univ., Middletown, PA (corresponding author). Email: svk5464@psu.edu

³Professor, Dept. of Industrial, Manufacturing, and Systems Engineering, Univ. of Texas at Arlington, Arlington, TX. Email: jrosenbe@uta.edu

⁴Executive Director, Auxiliary Services, Univ. of Texas at Arlington, Arlington, TX. Email: hladik@uta.edu

ABSTRACT

The development of electric vehicles (EVs) has accelerated in recent years, in part because of significant governmental support. Despite this support, however, the market share of EVs in the United States remains relatively small, and a large majority of consumers exhibit reluctance toward adopting them. This study aims to facilitate the effective development of EVs by identifying the barriers that affect their adoption in the United States and analyzing responses to an online questionnaire designed to provide insight into various perspectives of the barriers. A literature search identified 17 barriers that served as the basis of the survey questionnaire. A total of 733 individuals provided complete responses, and cluster analysis was performed. The results revealed three cluster segments and indicated that middle-aged males who possess higher levels of education and income exhibit greater levels of interest in the adoption of electric vehicles. The outcomes of this study may offer direction to policymakers in formulating efficacious energy and transportation policies and guide designers of electric vehicles to align the requirements and preferences of prospective consumers with their designs.

Keywords: Electric vehicles; adoption intention; barriers, cluster analysis

INTRODUCTION

The transport sector makes a substantial contribution to both atmospheric pollution and emissions. Furthermore, it has been observed that approximately 25% of the world's fossil fuel resources are consumed by the transportation sector, with a significant portion being allocated to road transportation (Gnann et al., 2018; Moeletsi, 2021; Patel et al. 2023). According to the 2020 report by the International Energy Agency (IEA), the transportation sector is responsible for 24% of direct carbon dioxide (CO₂) emissions from fuel combustion, and based on IEA's 2020 projections, this figure is expected to surge nearly 70% by 2050 under a status quo scenario (IEA, 2020). The statistics mentioned above highlight the need to develop a technological solution to mitigate automobile CO₂ emissions(Adnan et al., 2018; Chanamallu et al. 2023; Pamidimukkala et al. 2023a).

Electric vehicles (EVs) are recognized for their capacity to significantly diminish dependence on non-renewable energy sources, like fossil fuels, and reduce CO₂ emissions and other

environmental issues. A diverse range of EVs are currently accessible in the market, encompassing battery electric vehicles (BEV), plug-in hybrid electric vehicles (PHEV), and hybrid electric vehicles (HEV) (Ghosh, 2020; Patel et al. 2021; Sanguesa et al., 2021; Stockkamp et al., 2021). EVs present considerable ecological and financial benefits over conventional vehicles by substituting petroleum and natural gas with grid-based electricity (Pamidimukkala et al. 2023b). According to Zhou et al. (2020), implementing these measures will result in reduced greenhouse gas emissions, improved energy dependability, and enhanced green energy utilization.

The auto sector has taken notice of the 14% rise in electric vehicle sales worldwide in 2019. The sales in Europe experienced a significant increase of 80%, and Canada's sales grew by 43%. Conversely, sales in China and the United States remained stable. Several countries, including Norway, with a rate of 39.5% and the United Kingdom, with a rate of 1.94%, have followed suit in adopting the trend of increasing their procurement of electric vehicles. According to Ackaah et al. (2022), there has been a substantial surge in the global sales of EVs, with a record high of 6.6 million units sold in 2021. This figure represents a twofold increase from the previous year's sales. In 2021, there was a noteworthy surge in the global market share of electrified automobile sales in the global market, with a near 10% increase observed, a figure four times greater than the market share documented in 2019. As per the findings of Patyal et al. (2021), the worldwide tally of electric vehicles has surged to around 16.5 million, denoting a threefold upsurge from 2018.

Despite the recent surge in the proportion of EVs in the United States, the country is still grappling with challenges that impede their widespread adoption. Ali & Naushad (2022) and She et al. (2017) correlated the widespread adoption of electric vehicles with consumer perceptions. Therefore, the aim of this study is to understand the consumer perceptions of electric vehicles and to identify the segments of the population who will be EV users, explore the cluster differences in terms of barriers towards EV adoption, and profile each segment-based on demographics. The outcomes of this research can provide direction to policymakers in formulating efficacious energy and transportation policies and will provide guidance to those who design electric vehicles so that their designs align with the requirements and expectations of prospective customers.

LITERATURE REVIEW

According to findings presented in Table 1, limited driving range is one of the important barriers for consumers adoption of EVs. Consequently, individuals who engage in activities that involve extensive travel are less inclined to embrace EVs (Etminani-Ghasrodashti et al. 2023). Furthermore, it is noteworthy that the typical warranty period for an electric vehicle battery is eight to ten years (Patel et al. 2022a; Khan et al. 2023) which, given their high cost, makes their replacement a substantial challenge. According to existing literature, prospective customers also exhibit concerns about the length of time required for charging EVs (Adhikari et al., 2020; Khan et al. 2022; Li et al., 2017), as well as the lack of public charging facilities (Noel et al., 2020).

As presented in Table 1, the paucity of information on both reliability and safety of EVs, which is significant concern for potential consumers (She et al., 2017; Etminani-Ghasrodashti et al. 2021). The limited number of available models (Kongklaew et al., 2021), is also a barrier to EV sales; however, while a greater variety of vehicle options would appeal a broader market segment (Linzenich et al., 2019), EV production is typically limited (Xue et al., 2014).

Table 1. List of Identified Barriers

#	Barrier	Previous Studies
1	Limited driving range	(Berkeley et al., 2018; Singh et al., 2020)
2	Limited battery life	(Noel et al., 2020)
3	High battery replacement cost	(Kongklaew et al., 2021)
4	Long charging times	(Li et al., 2017; Adhikari et al., 2020)
5	Insufficient public charging stations	(Kongklaew et al., 2021)
6	Concerns about reliability	(Adhikari et al., 2020)
7	Poor safety	(Kongklaew et al., 2021)
8	Few EV models	(She et al., 2017)
9	High purchase price	(Noel et al., 2020)
10	Low resale value	(Lim et al., 2015)
11	High electricity price for charging	(Kim et al., 2018)
12	Problems of battery disposal	(Berkeley et al., 2018)
13	Environmental impact of battery production	(Giansoldati et al., 2020)
14	Cost of adapting home electrical system	(Patt et al., 2019)
15	Charging problem in the absence of a garage	(Illmann & Kluge, 2020)
16	Insufficient maintenance and repair services	(Giansoldati et al., 2020)
17	Unreliable charging electricity grid performance	(Kumar & Alok, 2020)

Table 1 demonstrates that EVs' cost and resale value are significant barriers to consumer adoption, as the absence of a well-established secondary market means that they have lower residual values than conventional vehicles. EVs have higher manufacturing costs (Cherchi, 2017; Anastasiadou & Gavanas, 2022), which are ultimately transferred to customers. Additionally, EVs rely on electrical energy for operation. Consequently, an increase in power prices leads to a decline in the demand for EVs (Kim et al., 2018; Patel et al. 2022b; Pamidimukkala et al., 2024).

The production of batteries and electricity required for electric vehicles (EVs) generates considerable pollution, and the disposal of used batteries poses a challenge due to the lack of a sufficient recycling infrastructure (Stockkamp et al., 2021; Ali & Naushad, 2022). This has resulted in divergent viewpoints concerning the environmental benefits of EVs (Liu et al., 2020; Etminani-Ghasroddashti et al. 2021; Ramesan et al., 2022).

According to the literature, individuals who rent their homes must engage in discussions with landlords regarding the expenses associated with installing a residential electrical system. The need for a comprehensive and expensive update of the electrical infrastructure can be an important barrier to the purchase of an EV (Liu et al., 2021; Parker et al., 2021; Patt et al., 2019). Apartment dwellers who want to recharge their cars have a hurdle due to the absence of a garage (Illmann & Kluge, 2020).

The dearth of maintenance and repair support centres and facilities for electric vehicles (EVs) in comparison to their conventional counterparts has resulted in discontentment among current EV owners. Furthermore, the intricacy of the processes required for servicing and repairing EVs is a notable challenge, compounded by the limited availability of mechanics who possess the necessary proficiency in these fields (Giansoldati et al., 2020; Pamidimukkala et al., 2023c).

METHODOLOGY

A survey questionnaire was designed and distributed electronically in March 2023 to potential consumers in Texas, using the online application QuestionPro to determine their perceptions of EVs. The questionnaire comprised 27 questions, which were categorized into two sections and took around 7 minutes to finish. After sending two reminder emails, 733 complete responses were received that were used for further analysis.

The analysis of socio-demographics revealed that females comprised 52.9% of the responses, males 44.3%, those who identified as “Others” constituted 2.8%. Around 62% of the respondents possessed bachelor's degrees. Most of the respondents (52.2%) belonged to the age group of 35 years and older. A mere 17.6% of the participants belonged to households earning less than \$35,000 annually. Most of the respondents had more than one vehicle, and 77% of participants are experienced drivers.

RESULTS AND DISCUSSION

Identifying Homogenous Respondent Groups

Cluster analysis was adopted to identify the homogenous groups of respondents. A clustering technique was performed in two stages utilizing SPSS 29, with the 17 identified barriers as the clustering variables. The initial step involved the implementation of a hierarchical clustering methodology to generate viable clustering outcomes, utilizing Ward's linkage technique with agglomeration coefficient, as previously employed in a few scholarly investigations (Jaiswal et al., 2022; Saleem et al., 2018). The outcome of the hierarchical analysis led to an assessment of the variance percentage in the heterogeneity-stopping rule, ultimately concluding that a 3-cluster solution was the most suitable. Subsequently, a K-means algorithm was employed to examine the membership of the clusters. Upon completion of the cluster analysis and attainment of an optimal cluster solution, an ANOVA was executed to ascertain the presence of statistically significant differences among the groups. The results indicate significant differences among the three groups for all the variables. Table 2 provides an overview of the three cluster memberships, with 43.1%, 12.7%, and 44.2% of respondents belonging to Cluster 1, Cluster 2, and Cluster 3, respectively.

Identifying the clusters based on demographic characteristics and using cross-tabulations of various demographic factors and segment memberships further enhanced our understanding of the clusters. The socio-demographic attributes of the respondents in each cluster are presented in Table 3. The cluster profiles are discussed below, based on the results from Table 2 and Table 3.

The Indifferents

The first cluster group, labelled the “Indifferents,” was the second largest cluster, comprised of 316 respondents. Most of them indicated that they were neutral on all the barriers, except for

poor safety ($X= 2.10$), doubts about reliability ($X= 2.35$), high purchase price ($X=4.21$), and insufficient public charging infrastructure ($X= 4.16$). Safety and reliability barriers were perceived to be less important in this cluster; the purchase price and lack of enough public charging stations were perceived to be more important.

Table 2. Respondents Scores of Each Cluster

Barrier	Cluster 1 (n ₁ =316)	Cluster 2 (n ₂ =93)	Cluster 3 (n ₃ =324)	Significance
	Indifferents	Enthusiasts	Skeptics	
Limited driving range	3.79	2.14	4.45	<0.001
Long charging times	3.70	2.10	4.56	<0.001
Limited battery life	3.58	1.99	4.65	<0.001
Poor safety	2.10	1.57	2.99	<0.001
Doubts about reliability	2.35	1.74	3.60	<0.001
Fewer EV models	2.98	1.83	3.87	<0.001
Problems of Battery disposal	3.06	1.77	4.42	<0.001
Environmental impact of battery production	3.36	1.99	4.50	<0.001
High purchase price	4.21	2.99	4.74	<0.001
High Battery replacement cost	3.87	2.47	4.83	<0.001
High electricity price for charging	3.00	1.66	4.40	<0.001
Lower resale value	3.09	1.73	4.07	<0.001
Cost of adapting a residential electrical infrastructure	3.64	2.14	4.73	<0.001
Insufficient public charging stations	4.16	2.51	4.64	<0.001
Charging problem in the absence of a garage	3.68	2.12	4.56	<0.001
Insufficient maintenance and repair services	3.85	2.05	4.69	<0.001
Unreliable charging electricity grid performance	3.41	1.62	4.62	<0.001

Significant at p<0.05

As illustrated in Table 3, Analysis of the demographic characteristics of this cluster showed that the proportion of male (49%) and female respondents (47.5%) was almost same, with most of them (53.8%) below the age of 34. The majority (23.7%) of them held a bachelor's degree; 23.4% held a graduate degree. Approximately 35.8% had an annual household income greater than \$100,000, and 44.2% owned two vehicles. About 74.7% of them had more than five years of driving experience.

Table 3. Sociodemographic Characteristics Based on Individual Clusters

Demographic Item		Indifferents	Enthusiasts	Skeptics
		$n_1 = 43.1\%$	$n_2 = 12.7\%$	$n_3 = 44.2\%$
		%	%	%
Gender	Female	47.5%	37.6%	62.6%
	Male	49.0%	58.1%	35.8%
	Other	3.5%	4.3%	1.5%
Age	18-24	27.8%	21.6%	29.3%
	25-34	26.0%	21.6%	13.9%
	35-44	14.9%	13.9%	17.9%
	45-54	11.7%	19.3%	13.6%
	55-64	14.9%	19.3%	16.7%
	65+	4.7%	4.3%	8.6%
Education	High school/GED	8.6%	5.4%	6.5%
	Some college/technical school	15.6%	14.0%	23.1%
	Associate degree	10.7%	10.8%	15.4%
	Bachelor's degree	23.7%	14.0%	20.1%
	Graduate degree	23.4%	31.1%	21.3%
Household income	Ph.D. or other equivalent degree	18.0%	24.7%	13.6%
	Less than \$20,000	10.4%	8.6%	6.2%
	\$20,000 - \$34,999	10.1%	4.3%	9.9%
	\$35,000 - \$49,999	13.0%	7.5%	15.1%
	\$50,000 - \$74,999	15.8%	15.1%	16.7%
	\$75,000 - \$99,999	14.9%	18.3%	18.5%
Vehicle ownership	\$100,000 or more	35.8%	46.2%	33.7%
	1 vehicle	30.1%	23.7%	21.3%
	2 vehicles	44.3%	45.2%	40.7%
Driving experience	3 or more vehicles	25.6%	31.2%	38.0%
	0-2 years	12.3%	7.5%	10.5%
	3-5 years	13.0%	8.6%	12.3%
	Over 5 years	74.7%	83.9%	77.2%

The Enthusiasts

The “Enthusiasts” were the smallest cluster, representing 12.7% of the respondents. This group has the lowest score on all the barriers as compared to other two groups. In addition, all these values are less than sample means, and they did not find any of the 17 barriers to be of importance. This indicates that the respondents in this cluster perceive EVs positively, and they tend to adopt EVs in future.

As shown in Table 3, the majority (58.1%) of the Enthusiasts were males, and 38.6% of them were between the ages of 45 and 64, which is higher than those in the other two clusters. Most of them held either a graduate or postgraduate degree (surpassing individual contribution of each of the other two clusters, and the total sample average) with the highest percentage (24.7%) holding

a Ph.D. or other equivalent degree. Regarding income, the enthusiasts have the highest proportion of individuals in the highest income category (\$100,000 and more 46.2%). In addition, 45.2% of respondents own 2 vehicles, and 83.9% of them are experienced drivers.

The Skeptics

The third cluster group labelled as “Skeptics.” was the largest group, representing 44.2% of the sample. They ranked all the barriers high except poor safety, doubts about reliability, and fewer EV models. The mean values of this cluster were higher than those of the other two groups and higher than those of the sample means. This shows that the Skeptics considered almost all the identified barriers highly important and indicates their negative and unfavorable views towards adopting EVs.

Table 3 presents a sociodemographic profile of the Skeptics, which is primarily distinguished by the exceedingly larger percentage of females (62.6%) over males (35.8%) in comparison to other groups examined in the study. On the income scale, the respondents in this cluster have relatively low income. In addition, when compared with the other clusters, this group has lower proportion of graduates and post-graduates and has a greater number of beginners and intermediate drivers compared to other two clusters.

CONCLUSION

The present study delved into the barriers to adopting EVs as perceived by consumers. This was accomplished by scrutinizing 733 responses from a survey administered to potential consumers of EVs. The findings suggest that the enthusiasts are more inclined to purchase EVs, as manufacturers could readily target them as early adopters using extrinsic incentive measures. In addition, this segment should be informed about the fundamental advantages of EVs, including savings on fuel cost and environmental conservation, through increased social persuasion and a positive attitude into using such vehicles. The findings also revealed that the indifferents are the second largest cluster, emerging with a neutral attitude of adopting EVs. This homogeneous cluster might be converted into potential EV adopters with convincing evidence demonstrating the significant benefits of EV use, including its affordability and environmental friendliness. The findings of this study also suggest that the promotion of EVs would be most effective if targeted towards male, middle-aged drivers with prior experience, given their high level of interest and propensity to adopt this technology. The findings of this study will provide valuable guidance to EV makers, designers, marketers, and other groups in effectively educating the public about the benefits of eco-friendly transportation and in shaping their inclination to possess an EV.

REFERENCES

- Ackaah, W., Kanton, A. T., and Osei, K. K. 2022. Factors influencing consumers' intentions to purchase electric vehicles in Ghana. *Transportation Letters*, 14(9), 1031-1042.
- Adhikari, M., Ghimire, L. P., Kim, Y., Aryal, P., and Khadka, S. B. 2020. Identification and analysis of barriers against electric vehicle use. *Sustainability*, 12(12), 4850.
- Adnan, N., Nordin, S. M., Amini, M. H., and Langove, N. 2018. What make consumer sign up to PHEVs? Predicting Malaysian consumer behavior in adoption of PHEVs. *Transportation Research Part A: Policy and Practice*, 113, 259-278.

- Ali, I., and Naushad, M. 2022. A study to investigate what tempts consumers to adopt electric vehicles. *World Electric Vehicle Journal*, 13(2), 26.
- Berkeley, N., Jarvis, D., and Jones, A. 2018. Analysing the take up of battery electric vehicles: An investigation of barriers amongst drivers in the UK. *Transportation Research Part D: Transport and Environment*, 63, 466-481.
- Channamallu, S. S., Kermanshachi, S., and Pamidimukkala, A. 2023. Impact of Autonomous Vehicles on Traffic Crashes in Comparison with Conventional Vehicles. In *International Conference on Transportation and Development 2023* (pp. 39-50).
- Etmiani-Ghasroddashti, R., Patel, R. K., Kermanshachi, S., Rosenberger, J. M., Weinreich, D., and Foss, A. 2021. Integration of shared autonomous vehicles (SAVs) into existing transportation services: A focus group study. *Transportation Research Interdisciplinary Perspectives*, 12, p.100481.
- Etmiani-Ghasroddashti, R., Hladik, G., Kermanshachi, S., Rosenberger, J. M., Arif Khan, M., and Foss, A. 2022. Exploring shared travel behavior of university students. *Transportation Planning and Technology*, 1-23.
- Etmiani-Ghasroddashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2023. Exploring motivating factors and constraints of using and adoption of shared autonomous vehicles (SAVs). *Transportation Research Interdisciplinary Perspectives*, 18, p.100794.
- Ghosh, A. 2020. Possibilities and challenges for the inclusion of the electric vehicle (EV) to reduce the carbon footprint in the transport sector: A review. *Energies*, 13(10), 2602.
- Giansoldati, M., Monte, A., and Scorrano, M. 2020. Barriers to the adoption of electric cars: Evidence from an Italian survey. *Energy Policy*, 146, 111812.
- Gnann, T., Stephens, T. S., Lin, Z., Plötz, P., Liu, C., and Brokate, J. 2018. What drives the market for plug-in electric vehicles?-A review of international PEV market diffusion models. *Renewable and Sustainable Energy Reviews*, 93, 158-164.
- IEA. 2020. *Electric Vehicles*. In. Paris, France: IEA.
- Illmann, U., and Kluge, J. 2020. Public charging infrastructure and the market diffusion of electric vehicles. *Transportation Research Part D: Transport and Environment*, 86, 102413.
- Khan, M. A., Etmiani-Ghasroddashti, R., Kermanshachi, S., Rosenberger, M., Pan, Q., and Foss, A. 2022. Do ridesharing transportation services alleviate traffic crashes? A time series analysis. *Traffic injury prevention*.
- Khan, M. A., Patel, R. K., Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., Hladik, G., and Foss, A. 2023. Factors that determine a university community's satisfaction levels with public transit services. *Frontiers in Built Environment*, 9, 1125149.
- Kim, M.-K., Oh, J., Park, J.-H., and Joo, C. 2018. Perceived value and adoption intention for electric vehicles in Korea: Moderating effects of environmental traits and government supports. *Energy*, 159, 799-809.
- Kongklaew, C., Phoungthong, K., Prabpayak, C., Chowdhury, M. S., Khan, I., Yuangyai, N., and Techato, K. 2021. Barriers to electric vehicle adoption in Thailand. *Sustainability*, 13(22), 12839.
- Kumar, R. R., and Alok, K. 2020. Adoption of electric vehicle: A literature review and prospects for sustainability. *Journal of Cleaner Production*, 253, 119911.
- Li, W., Long, R., Chen, H., and Geng, J. 2017. A review of factors influencing consumer intentions to adopt battery electric vehicles. *Renewable and Sustainable Energy Reviews*, 78, 318-328.

- Liu, R., Ding, Z., Jiang, X., Sun, J., Jiang, Y., and Qiang, W. 2020. How does experience impact the adoption willingness of battery electric vehicles? The role of psychological factors. *Environmental Science and Pollution Research*, 27, 25230-25247.
- Liu, Z., Song, J., Kubal, J., Susrarla, N., Knehr, K. W., Islam, E., and Ahmed, S. 2021. Comparing total cost of ownership of battery electric vehicles and internal combustion engine vehicles. *Energy Policy*, 158, 112564.
- Moeletsi, M. E. 2021. Socio-economic barriers to adoption of electric vehicles in South Africa: Case study of the gauteng province. *World Electric Vehicle Journal*, 12(4), 167.
- Noel, L., de Rubens, G. Z., Kester, J., and Sovacool, B. K. 2020. Understanding the socio-technical nexus of Nordic electric vehicle (EV) barriers: A qualitative discussion of range, price, charging and knowledge. *Energy Policy*, 138, 111292.
- Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., and Hladik, G. 2023a. "Evaluation of barriers to electric vehicle adoption: A study of technological, environmental, financial, and infrastructure factors." *Transportation Research Interdisciplinary Perspectives*, 22, p.100962.
- Pamidimukkala, A., Patel, R. K., Kermanshachi, S., Rosenberger, J. M., and Tanvir, S. 2023b. A Review on Shared Mobility and Electric Vehicles. In *International Conference on Transportation and Development 2023* (pp. 333-342).
- Pamidimukkala, A., Kermanshachi, S., and Patel, R. 2023c. An Exploratory Analysis of Crashes Involving Autonomous Vehicles. In *International Conference on Transportation and Development 2023* (pp. 343-350).
- Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., and Hladik, G. 2024. "Barriers and Motivators to the Adoption of Electric Vehicles: A Global Review." *Journal of Green Energy and Intelligent Transportation*, p.100153.
- Parker, N., Breetz, H. L., Salon, D., Conway, M. W., Williams, J., and Patterson, M. 2021. Who saves money buying electric vehicles? Heterogeneity in total cost of ownership. *Transportation Research Part D: Transport and Environment*, 96, 102893.
- Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., Pamidimukkala, A., and Foss, A. 2023. Identifying individuals' perceptions, attitudes, preferences, and concerns of shared autonomous vehicles: During-and post-implementation evidence. *Transportation Research Interdisciplinary Perspectives*, 18, 100785.
- Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., and Foss, A. 2022a. Exploring willingness to use shared autonomous vehicles. *International Journal of Transportation Science and Technology*.
- Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J., and Foss, A. 2022b. How Riders Use Shared Autonomous Vehicles. In *Automated People Movers and Automated Transit Systems, ASCE International Conference on Transportation & Development*, pp. 81-93.
- Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., and Weinreich, D. 2021. Exploring preferences towards integrating the autonomous vehicles with the current microtransit services: A disability focus group study. In *International Conference on Transportation and Development 2021* (pp. 355-366).
- Patt, A., Aplyn, D., Weyrich, P., and van Vliet, O. 2019. Availability of private charging infrastructure influences readiness to buy electric cars. *Transportation Research Part A: Policy and Practice*, 125, 1-7.

- Patyal, V. S., Kumar, R., and Kushwah, S. 2021. Modeling barriers to the adoption of electric vehicles: An Indian perspective. *Energy*, 237, 121554.
- Ramesan, S., Kumar, P., and Garg, S. K. 2022. Analyzing the enablers to overcome the challenges in the adoption of electric vehicles in Delhi NCR. *Case Studies on Transport Policy*, 10(3), 1640-1650.
- Sanguesa, J. A., Torres-Sanz, V., Garrido, P., Martinez, F. J., and Marquez-Barja, J. M. 2021. A review on electric vehicles: Technologies and challenges. *Smart Cities*, 4(1), 372-404.
- She, Z.-Y., Sun, Q., Ma, J.-J., and Xie, B.-C. 2017. What are the barriers to widespread adoption of battery electric vehicles? A survey of public perception in Tianjin, China. *Transport Policy*, 56, 29-40.
- Singh, V., Singh, V., and Vaibhav, S. 2020. A review and simple meta-analysis of factors influencing adoption of electric vehicles. *Transportation Research Part D: Transport and Environment*, 86, 102436.
- Stockkamp, C., Schäfer, J., Millemann, J. A., and Heidenreich, S. 2021. Identifying factors associated with consumers' adoption of e-mobility—A systematic literature review. *Sustainability*, 13(19), 10975.
- Zhou, Y., Wen, R., Wang, H., and Cai, H. 2020. Optimal battery electric vehicles range: A study considering heterogeneous travel patterns, charging behaviors, and access to charging infrastructure. *Energy*, 197, 116945.

Smartwatches in Transportation: Unleashing Innovations and Advancements—A Comprehensive Systematic Review

Eazaz Sadeghvaziri, Ph.D.¹; Ramina Javid²; and Mohammadsoroush Tafazzoli, Ph.D.³

¹Assistant Professor, Dept. of Environmental and Civil Engineering, School of Engineering, Mercer Univ., Macon, GA (corresponding author). Email: sadeghvaziri_e@mercer.edu

²Ph.D. Candidate, Dept. of Transportation and Urban Infrastructure, Morgan State Univ. Email: rajah1@morgan.edu

³Assistant Professor, Dept. of Civil Engineering and Construction, Georgia Southern Univ., Statesboro, GA. Email: mtafazzoli@georgiasouthern.edu

ABSTRACT

This study explores the growing popularity of smartwatches since 2015, particularly their application in various transportation sectors. Utilizing the PRISMA guideline, the research systematically reviews 33 articles focusing on smart wearable devices, with an emphasis on smartwatches. The evaluation encompasses goals, methods, findings, advantages, and disadvantages, drawing comparisons with smartphones. Studies are categorized into safety and driver behavior, health and physical activities, equity, comparison of smartwatches and smartphones, and workers and construction. Results indicate a surge in research interest in safety, driver behavior, and health-related aspects of smartwatches, with a noticeable rise in popularity in recent years. The study highlights the potential for smartwatches, given advancements in smartphone health sensors, especially in examining the physical activity and health issues of cyclists, scooter riders, and pedestrians in transportation engineering. This comprehensive review aids researchers in selecting quantification methods aligned with their objectives and study constraints.

INTRODUCTION

This review of the literature focuses on the use of smartwatches in different fields of transportation engineering, which researchers have investigated during the past decade. The results summarized in this review encompass multiple domains of smartwatch use in transportation engineering. Understanding the existing status of research reveals the significance of this study's direction in terms of the advantages of smartwatches over other devices, as well as potential future research. Hence, this systematic review produces a varied knowledge base for review users to provide insight into the potential use of smartphones in transportation.

This study follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2011), which is an evidence-based minimum set of items used for reporting in systematic reviews and meta-analyses, to systematically review relevant publications. The PRISMA is a global initiative created by pertinent professionals to solve the persistent problem of a lack of transparent and well-documented review techniques presented in published review papers (Tepper 2022). Therefore, this review was conducted on academic publications that were found on ScienceDirect, Transportation Research Record (TRR), and Transport Research International Documentation (TRID). ScienceDirect is a popular platform for finding peer-reviewed articles from books and journals published by Elsevier, including

thousands of open-access papers. ScienceDirect combines reliable full-text articles in the areas of science, technology, and health, with smart, user-friendly functionality to help users keep updated in their disciplines and become more successful and effective in their profession (Elsevier 2023). The TRR, which is one of the most often cited transportation publications worldwide and has unparalleled depth and breadth in its coverage, is the Transportation Research Board's (TRB's) top publication (TRR 2023). The TRID is an integrated database that combines records from TRB's Transportation Research Information Services (TRIS) Database and the Organization for Economic Co-operation and Development's (OECD) Joint Transport Research Centre's International Transport Research Documentation (ITRD) Database. The TRID provides access to more than 1.3 million records of transportation research worldwide (TRID 2023). The databases were searched to identify published articles, reports, conference papers, and book chapters using any combination of keywords in their title, abstract, and keywords, including "smartwatch" or "smart-watch", or "smartwatch" or "wristwatch". Only published material in English was included due to the burdensome translating process. All material considered in the review was published between 2013 and July 2022. This period of time was chosen because most of the relevant literature was published in the last decade, and it is important to study the most recent literature. Many studies have used systematic reviews to investigate mobile phone applications in transportation, active transportation for underrepresentation populations, biking activities among low-income households, etc. (Rojas IV, Sadeghvaziri, and Jin 2016; Sadeghvaziri, Javid, and Jeihani 2023; Sadeghvaziri et al. 2023). However, to the best knowledge of the authors, no study has yet conducted a systematic review of smartwatch applications in transportation.

METHODOLOGY REVIEW

The implemented review process of this study is shown in Figure 1. This flow diagram shows how information moves through the various stages of a systematic review. It illustrates how many records were found, how many were included and excluded, and why they were excluded. Different templates were made available depending on the type of review (new or updated) and the sources utilized to find relevant research (Tepper 2022). As of July 2022, a total of 549 publications were identified using the developed search strategy. After checking for duplicates, screening the identified articles, and reviewing articles' full text, we excluded 581 articles: 55 duplicates, 226 after title screening, 257 after abstract screening, 3 after full-text retrieve, and 7 after full-text review. Finally, 33 articles met inclusion criteria and were included in this review.

Characteristics of Included Studies

Based on the literature, the previous study contributed significantly to the literature on the uses of smartwatches in transportation. These studies can be classified into five categories: Safety and Driver Behavior, Health and Physical Activities, Equity, Comparison of Smartwatches and Smartphones, and Workers and Construction.

Figure 2 shows that most of the literature on this subject is related to Safety and Driver Behavior, Health and Physical Activities. Figure 3 shows the publication year of the studies included in this review. This time-series analysis of publications indicated that smartwatch-related studies in transportation and their focus on safety and health began receiving more attention over time, with the number of publications increasing in 2019. Increases in the

availability of new features (such as fall detection and electrical heart sensor) may be one of the reasons behind this change.

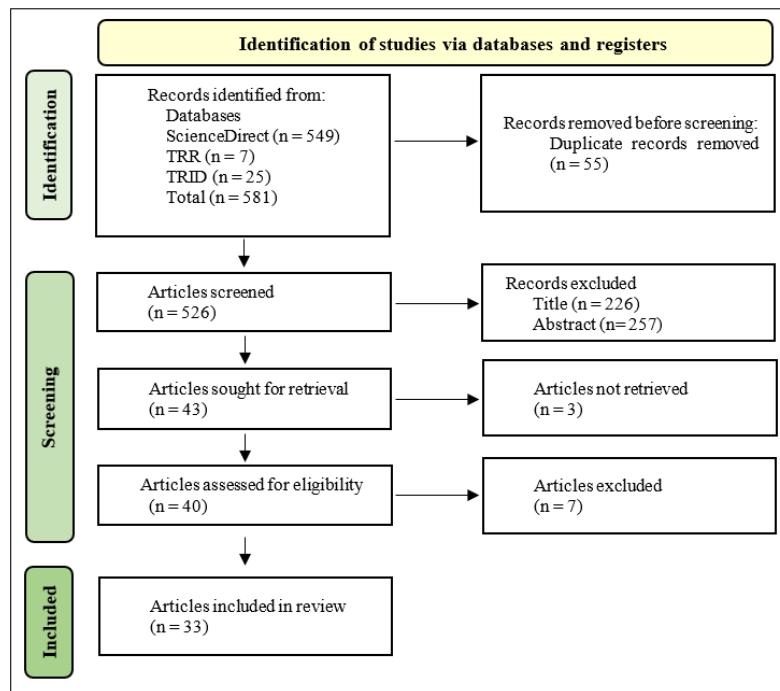


Figure 1. PRISMA Flowchart for reporting systematic review

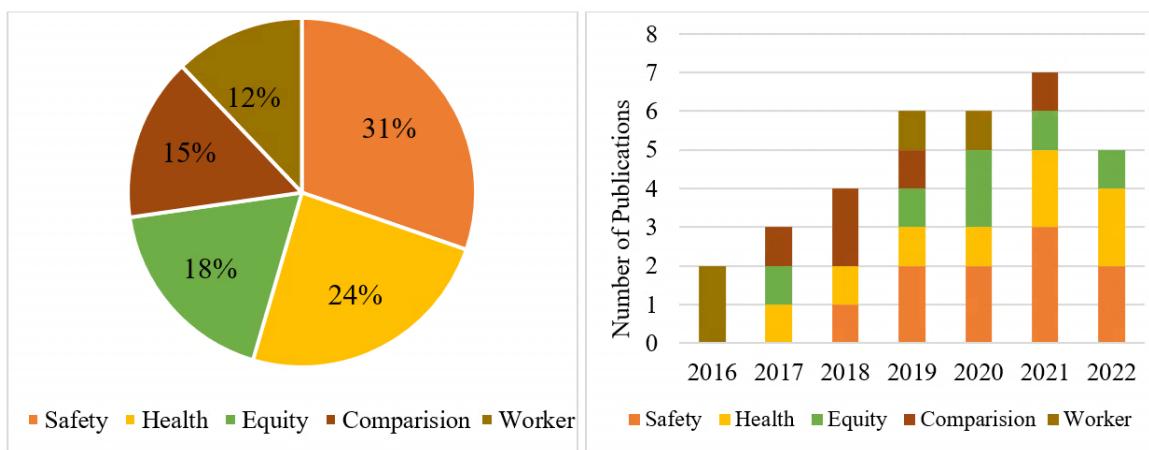


Figure 2. Distribution of the identified smartwatch study approaches; Figure 3. Publication date of the studies included in this review

LITERATURE REVIEW

Safety and Driver Behavior

Smartwatches have been used in traffic safety and driver behavior, including distracted driving, driver fatigue, etc. Researchers are more likely to use physiological and behavior-based

approaches to identify driving fatigue or drowsiness. Although smartwatches can be used to improve the safety of drivers, hand movement data are very rarely used by researchers to detect driver fatigue. For instance, a study used a Bluetooth-enabled EEG headband system and a wristwatch to detect driver drowsiness. Moreover, another study detected driver drowsiness by a built-in motion sensor in a smartwatch using hand movement data (Kamti and Iqbal 2022).

According to studies, environmental elements substantially impact emotions, which in turn incredibly impact drivers' emotional states and driving behavior. Furthermore, given that demand for autonomous vehicles is anticipated to grow dramatically over the next ten years, building trust with people will require a thorough understanding of drivers' and passengers' emotions, behaviors, and preferences. The results of a naturalistic driving research study by Balali et al. (Balali et al. 2020) suggested a novel semi-automated method for determining how ambient influences affect drivers' emotions and behavioral changes. A frontal head-mounted camera, a smartwatch for monitoring physiological data, and a Controller Area Network (CAN) serial data logger were all part of their configuration. Researchers found that the type of road and the weather conditions, which can alter driving behaviors, significantly impact the driver's effect. One of the most critical steps in improving traffic safety is improving data collection methods, data fusion, and edge computing. Edge computing is a distributed Information Technology (IT) architecture in which client data is processed at the periphery of the network, as close to the originating source as possible (Bigelow n.d.).

In another traffic safety research study, He et al. (C. He et al. 2021) used the sensor data of smartwatches to detect whether the driver's hand is on the steering wheel in real-time. Results showed a detection rate of 90% accuracy of whether or not the driver's hand stayed on the steering wheel. Furthermore, Bachmann et al. (Bachmann, Morold, and David 2021) investigated the accuracy of recognizing a pedestrian's movements by analyzing the "Car-to-Pedestrian Nearside" scenarios from the "Test Protocol for Autonomous Emergency Braking Vulnerable Road Users (VRU) Systems" of the European New Car Assessment Program (Euro NCAP). They provided a comprehensive analysis of how the accuracy of recognizing a pedestrian's movement influences the reliability of detecting collisions between the vehicle and the pedestrian in cooperative VRU collision avoidance systems. Ninjaed et al. (Niknejad et al. 2020) conducted a systematic review of studies on smart wearables in 2019. Ninjaed found an upward trend in the number of smart wearable studies from 2010 to 2019. Moreover, Ninjaed classified the studies into five main categories, including Technology-focused (54%), User Behavior (27%), Design (14%), Security and Privacy (4%), and Social Acceptability (1%). Additionally, Lee et al. (Lee et al. 2019) systematically reviewed wearable and IoT devices. They proposed several research directions in intelligent, positive computing systems research.

Studies on human factors and safety have shown that personality traits and emotional states, such as stress, weariness, and others, significantly impact people's decisions when it comes to driving. Paschalidis et al. (Paschalidis, Choudhury, and Hess 2019) aimed to build a car-following model that takes into account the driver's stress level and quantifies the effect it has on acceleration-deceleration decisions. They used a non-intrusive wristband to collect the physiological measurements of stress and regularly utilized variables, such as heart rate, blood volume pulse, and skin conductance. They suggested a framework extension of the GM stimulus-response model, where stress was viewed as a latent (unobserved) variable, and the specification also considers the effects of drivers' sociodemographic traits. In another study, Bej et al. (Bej et al. 2019) developed a methodical, affordable, and sensor-free solution that alerts distracted drivers to prevent crashes. They developed a distracted driving prevention system

whereby wearable devices sound an alarm when a driver is distracted. Their prototype device was installed on the dashboard of a moving car to determine real-world results. The results showed that drivers had an adequate alertness level of 76 percent. Overall, smartwatches have been widely used in traffic safety-related studies.

Health and Physical Activities

Investigating a traveler's health and physical activities is another transportation research area that shows how smartwatches play a significant and beneficial role. Smartwatches can significantly help identify cognitive load and impaired driving performance. A driver's cognitive load and impaired driving performance can be increased using in-vehicle information systems. He et al. explored the performance of several machine learning models in classifying three levels of cognitive load on drivers. The fusion of eye-tracking and physiological sensors, such as smartwatches, seems more promising for real-time assessment of a driver's cognitive load. Yet, research is lacking in this area (D. He et al. 2022).

Numerous health studies have shown through their research findings that regular Physical Activity (PA) improves health outcomes. People spend a considerable time traveling, especially when commuting to their jobs. Driving is generally the least physically taxing means of transportation, which adds to today's sedentary lives and health issues. The use of public transportation, along with walking and cycling, gives people a chance to boost their PA levels regularly. Harms et al. (Harms, Olaru, and Pattison 2019) aimed to use smartwatches to assess a novel mix of data-gathering instruments and techniques in order to develop a thorough understanding of travel-related PA. Their kit of instruments offered a precise spatial and temporal account of daily activities and supplemental contextual data (such as heart rate, the state of the transportation infrastructure and services, and degree of enjoyment) that are not typically collected in travel and health studies. They asked 50 participants in office-based (sedentary) occupations to wear a Garmin smartwatch and an EDESIX wearable camera (objective measures) for two days and complete a self-report time-use diary before participating in a reconstruction interview a few days later, which were self-report measures. They were able to cross-validate GPS and PA data with diary reports, as well as gain a deeper understanding of activity planning and travel mode preferences thanks to the combination of quantitative (GPS tracks, heart rate, PA, distances, and activity duration) and qualitative (contextual information from the video footage and reconstruction interviews) data.

Activity recognition technology is one of the most crucial technologies for life-logging and the care of older people. Elderly people prefer to reside in their own homes, and in their communities. And if they can do so, society and the economy stand to gain from several factors. However, living alone may come with significant risks. Wearable sensors have been created to avoid these hazards, and they should soon be suitable for medical applications. Wearable sensors can assist in discreetly observing an elderly person's everyday routines to monitor their well-being if they live alone. Kumari et al. (Kumari, Mathew, and Syal 2017) examined the rising popularity of wearable technology and the requirement for multimodal recognition for continuous or intermittent monitoring of human activity, and biological signals like electroencephalograms (EEG), electrooculograms (EOG), electromyograms (EMG), electrocardiograms (ECG), and parameters, and other symptoms. Moreover, gesture recognition is crucial in smart environments, such as autonomous vehicles. Travelers can specify gestures connected to commands for the smart environment thanks to the availability of machine-learning technologies.

Equity

Equity in transportation includes different areas, including the transportation of people with disabilities, low-income household families that mainly use walking or cycling as their primary mode of transportation, senior adults' mobility, etc. Furthermore, one of the main aspects of equity in transportation is to ensure that new transportation systems and technologies are available to everyone, especially for people with disabilities. Smartwatches have the potential to help people with disabilities by using Connected and Autonomous Vehicles (CAVs).

Ranjbar et al. (2022) investigated the potential ways that CAVs can be used by people with different types of disabilities, including blindness, deafblindness, and deafness. They investigated whether vibrotactile aid could enable persons with blindness, deafblindness, and deafness to use CAVs. They recruited 15 participants with hearing and vision impairments to compare trips with and without vibrotactile guidance in a simulated autonomous vehicle in real traffic. One of the survey's questions asked the participants whether they had any suggestions for how to make the communication aids better. To make it easy to conceal within clothing, together with bracelets and a watch, they intended the device to be the size of a wristwatch, preferably integrated with already existing devices, such as a smartwatch or a cell phone (Ranjbar et al. 2022). Moreover, Mascetti et al. (2020) presented SmartWheels, which detects urban features by analyzing inertial sensor data produced by wheelchair movements. They collected data from different sources, including smartwatches. Their smartwatch ran custom Android applications to collect data from built-in inertial sensors. Their experiments, which included 17 wheelchairs and smartwatches, showed that the proposed approach is indeed effective, especially when using the data collected in a controlled environment (Mascetti et al. 2020).

Additionally, it can be difficult for senior pedestrians to maintain their quality of life when navigating metropolitan settings. However, the maps that older pedestrians typically use might not be appropriate for their individual needs, and the current digital aids do not consider older persons' user experience or perceptual and cognitive deficits. Montuwy et al. (2019) provided a detailed description of the navigation experience of elderly pedestrians using wearable technology that was either visual (augmented reality glasses), auditory (bone conduction headphones), or both visual and haptic (smartwatch). These wearables were contrasted with the navigation tool that older adults typically use to navigate cities (their own digital or paper map). The study consisted of 18 participants, tested on navigation performance, and interviewed to record the in-depth comments of each user's experience. Overall, smartwatches have been used in different aspects of equity in transportation (Montuwy, Dommes, and Cahour 2019).

Comparison of Smartwatches and Smartphones

Although some wearable devices have disadvantages over smartphones, smartwatches have many advantages. For instance, due to the small size of smartwatches, their batteries need to be recharged frequently to provide continuous use. For instance, Apple Watch Series 7 needs to be recharged every 18 hours. Moreover, smartwatches should be well-fitting and light for consumers to feel at ease while using them. Therefore, these size and weight restrictions might force smartwatches to use low-end hardware, such as low-capacity batteries.

Moreover, wearable devices like Google Glass and smartwatches can be less distractible than smartphones. Head-up and wearable displays, like Google Glass, have been promoted as safe in-vehicle substitutes for phone-based displays because they allow users to receive messages

without taking their eyes off of the road. However, using a head-up display while driving can still be dangerous, as using a device while driving might be distracting depending on the user's multitasking techniques. He et al. (2018) investigated how drivers interacted with a worn head-up display and a head-down smartphone display. Participants used the head-mounted display (HMD) on Google Glass or a smartphone to receive and respond to text messages while they were engaged in a driving simulation. An audio alarm announced new communications, and responses were spoken aloud. They understood that participants responded more quickly when using Google Glass than when using a smartphone, and the amount of time needed to complete a task was unaffected by the difficulty of lane-keeping. The findings imply that the potential safety advantages of wearable technology may be negated by a propensity to participate in distracted activities more readily (J. He et al. 2018).

In another study, Perlman et al. (2019) compared the effects of smartwatches and smartphones on drivers' behaviors by collecting data from 36 participants in a driving simulator. The effects of smartphone visual-manual (VM) and auditory-vocal (AV) interface on driver workload, attention, and performance were compared to those of using a smartwatch to initiate phone calls. The remote detection task (R-DRT) responsiveness, mean single glance duration, percentage of long duration off-road glances, total off-road glance time, and percent of time looking off-road were similar, while task time and number of glances were higher for AV calling on the smartwatch compared to the smartphone. The metrics were significantly higher for the VM interface vs. the AV methods. When making phone calls, heart rate and skin conductance were higher than when "simply driving," although they did not reliably distinguish between calling methods. The comparison of the smartphone VM calling to both AV techniques showed that participants drove more erratically (lane position and severe steering wheel reversals). The workload ratings for AV calling on both devices were lower than VM calling (Perlman et al. 2019).

On the other hand, smartwatches can dramatically increase the quality of life of active transportation mode users, such as walkers, bicyclists, scooter riders, etc., since they interact with the human body without manual intervention (Jin et al. 2022). For instance, Apple Watch can collect motion data from the accelerometer sensor to detect whether the active transportation mode user experiences a hard fall. If the traveler does not move for about a minute after the fall is detected, an automatic call for help is placed. Furthermore, Fitbit Sense provides a robust set of tracking measures, including pace, body temperature when sleeping, sleep quality, menstrual cycle, food and drink intake, and more. It also includes an integrated electrodermal activity (EDA) scan, which could be another method of determining the degree of stress. Overall, the primary source of a smartwatch's unique advantage is its ability to be worn while used.

Smart user gadgets are becoming more commonplace and influential in determining a user's present behavior and context. San-Segundo et al. examined several methods for enhancing the stability of a Human Activity Recognition (HAR) system that leverages accelerometer readings from various smartphones and smartwatches. Their investigation highlights a few of the difficulties brought on by device heterogeneity and the differing functionality of smartwatches compared to smartphones. For instance, arm movements bring additional unpredictability when utilizing smartwatches to identify whole-body activities, leading to a considerable reduction in HAR. They identified numerous issues with HAR for smart devices in their study by analyzing the motion sensor outputs from a range of users and utilizing a variety of smartphones and smartwatches. Due to the greater unpredictability in the recordings from smartwatches, including arm and whole-body motions, these difficulties are noticeably more severe for smartwatches than

smartphones. From a traffic safety point of view, smartwatches can be more distractible than other gadgets while driving (San-Segundo et al. 2018). The use of smartwatches raises several concerns regarding their potential for distraction in circumstances where sustained attention is crucial, such as operating a motor vehicle. Brodeur et al. (2021) investigated whether smartwatches are more distractible than smartphones. Results of an experimental study using a driving simulator and eye-tracking goggles showed that drivers become more distracted in the smartwatch condition than in the mobile phone condition, and they become less distracted in the speaker condition than in the phone condition (Brodeur et al. 2021).

Workers and Road Construction

The advancement of wearable technology has led to a growing number of research projects that have examined the viability of utilizing wearable sensors to alert construction workers to potentially dangerous situations. Businesses widely use wearable technology to boost operational effectiveness and reinforce competitive advantage. When making such judgments, managers may find it challenging to choose the best technology for their business. Buyukozkan et al. (2020) introduced an evaluation framework built on a Hesitant Fuzzy Linguistic (HFL) Multi-Criteria Decision-Making technique to take into account all of the factors influencing the final choice in order to address the Smart Watch (SW) selection dilemma. During decision-making, managers can find it difficult to select the right device for their company. To address the Smart Watch (SW) selection problem, this article introduces an assessment framework established on a Hesitant Fuzzy Linguistic (HFL) Multi-Criteria Decision-Making technique to consider parameters affecting the eventual decision collectively. Furthermore, the high frequency of work-related accidents and fatalities makes the construction process and work zone areas hazardous. The gathering and analysis of safety data is a crucial component in developing measurement and improvement strategies. The implementation of wearable technology has the potential to provide a data collection and analysis method focused on results that can give workers real-time information (Buyukozkan and Güler 2020). Awolusi et al. (2018) thoroughly analyzed wearable technology applications for customized safety monitoring. Predicting safety performance and management practices is possible by looking at the characteristics of wearable devices and safety measurements. According to the review, various safety performance criteria can be monitored and measured using wearable technologies. For multi-parameter safety performance monitoring, the advantages of individual wearable sensors or systems can be merged based on their characteristics (Awolusi, Marks, and Hallowell 2018).

Smartwatches are unobtrusive, ordinary objects that can be used to interact with machines through gestures. Villani et al. (2016) suggested using a smartwatch to communicate with a sophisticated troubleshooting program that will be deployed in an industrial setting. The application is a hypermedia information system that aims to assist workers in performing preventive and corrective machine maintenance. The smartwatch's hands-free interaction feature makes this possible when using the entire system while wearing personal protective equipment, such as gloves or having fingers covered in oil or dust that hinder functioning touch screens.

SUMMARY

Smartwatches have been applied in a wide range of industries, and more recently, they entered the field of transportation engineering. This study uncovered prior studies that

concentrate on various smart wearable devices, particularly smartwatches. The PRISMA recommendation was followed to review pertinent publications thoroughly. A total of 33 articles were included in the review after the literature was screened and examined. The investigations were divided into five categories: workers and construction, equity, health and physical activity, comparison of smartwatches and smartphones, and safety and driving behavior. Given that newer smartphones are outfitted with more sophisticated health sensors, it appears that smartwatches can all the more be employed in the field of transportation engineering, particularly when examining the physical activity and health-related concerns of cyclists, scooter riders, and walkers. It is understood that there seems to be a significant gap between a traveler's needs and the potential of autonomous vehicles. Hence there is an excellent opportunity to design services based on understanding the trip from start to end. If society is to guarantee autonomous mobility for everyone, regardless of limitations, a holistic viewpoint is required. Moreover, in terms of limitations, although smartwatches can be very useful in navigating people with disabilities, it is important to avoid sending too much information. For instance, it was found that the haptic codes utilized in a particular experiment have limits on how much information can be conveyed. In other words, a user can only remember a certain number of codes. Although training can improve it, a limited number of comments can be used when starting to learn the tool. In future studies, the intensity of the haptics of smartwatches can be investigated for better navigation of people with disabilities. The design of haptic communication needs to be improved; for instance, the intensity of the vibrations should change based on the importance of the information. Researchers should be able to select the optimal quantification approach for their objectives and study constraints with the aid of this review.

REFERENCES

- Awolusi, I., E. Marks, and M. Hallowell. 2018. "Wearable Technology for Personalized Construction Safety Monitoring and Trending: Review of Applicable Devices." *Automation in Construction* 85 (January): 96–106.
- Bachmann, M., M. Morold, and K. David. 2021. "On the Required Movement Recognition Accuracy in Cooperative VRU Collision Avoidance Systems." *IEEE Transactions on Intelligent Transportation Systems* 22 (3): pp 1708-1717.
- Balali, V., Mineta Transportation Institute, State of California SB1 2017/2018, California State University Transportation Consortium, and Long Beach California State University. 2020. "A Multimodal Approach for Monitoring Driving Behavior and Emotions." Digital/other. Final Report. 01752854.
- Bej, D., S. Rakshit, A. K. Mal, and R. Mahapatra. 2019. "A Cost-Effective System for Triggering Alarm to Distracted Drivers/ Nurses." *Computers & Electrical Engineering* 76 (June): 24–39.
- Bigelow, S. n.d. "What Is Edge Computing? Everything You Need to Know." Tech Accelerator. Accessed July 30, 2022. <https://www.techtarget.com/searchdatacenter/definition/edge-computing>.
- Brodeur, M., P. Ruer, P.-M. Léger, and S. Séncal. 2021. "Smartwatches Are More Distracting than Mobile Phones While Driving: Results from an Experimental Study." *Accident Analysis & Prevention* 149 (January).
- Büyüközkan, G., and M. Güler. 2020. "Smart Watch Evaluation with Integrated Hesitant Fuzzy Linguistic SAW-ARAS Technique." *Measurement* 153 (March): 107353.

- Elsevier. 2023. "ScienceDirect | Peer Reviewed Literature | Elsevier." 2023. <https://www.elsevier.com/products/sciedirect>.
- Harms, T., D. Olaru, and C. Pattison. 2019. "Active Travel: Using Wearable Technology to Analyse Daily Travel Behaviour." In, 15 p. <https://www.australasiantransportresearchforum.org.au/papers/2019>.
- He, C., W. Liu, D. Sun, Z. Gao, and ASCE. 2021. *Hand off Steering Wheel Detection Using Sensor Data of Smartwatch and Smartphone*. In, pp 630-640.
- He, D., Z. Wang, E. B. Khalil, B. Donmez, G. Qiao, and S. Kumar. 2022. "Classification of Driver Cognitive Load: Exploring the Benefits of Fusing Eye-Tracking and Physiological Measures." *Transportation Research Record*, May, 03611981221090937.
- He, J., J. S. McCarley, K. Crager, M. Jadliwala, L. Hua, and S. Huang. 2018. "Does Wearable Device Bring Distraction Closer to Drivers? Comparing Smartphones and Google Glass." *Applied Ergonomics* 70 (July): 156–66.
- Jin, X., L. Li, F. Dang, X. Chen, and Y. Liu. 2022. "A Survey on Edge Computing for Wearable Technology." *Digital Signal Processing, Sensing, Signal Processing and Computing for the Era of Wearables*, 125 (June): 103146.
- Kamti, M. K., and R. Iqbal. 2022. "Evolution of Driver Fatigue Detection Techniques—A Review From 2007 to 2021." *Transportation Research Record*, July, 03611981221096118.
- Kumari, P., L. Mathew, and P. Syal. 2017. "Increasing Trend of Wearables and Multimodal Interface for Human Activity Monitoring: A Review." *Biosensors and Bioelectronics* 90 (April): 298–307.
- Lee, U., K. Han, H. Cho, K.-M. Chung, H. Hong, S.-J. Lee, Y. Noh, S. Park, and J. M. Carroll. 2019. "Intelligent Positive Computing with Mobile, Wearable, and IoT Devices: Literature Review and Research Directions." *Ad Hoc Networks* 83 (February): 8–24.
- Mascetti, S., G. Civitarese, O. El Malak, and C. Bettini. 2020. "SmartWheels: Detecting Urban Features for Wheelchair Users' Navigation." *Pervasive and Mobile Computing* 62 (February): 101115.
- Moher, D., D. G. Altman, A. Liberati, and J. Tetzlaff. 2011. "PRISMA Statement." *Epidemiology* 22 (1): 128.
- Montuwy, A., A. Dommes, and B. Cahour. 2019. "Using Sensory Wearable Devices to Navigate the City: Effectiveness and User Experience in Older Pedestrians." *Multimodal Technologies and Interaction* 3 (1): 24p.
- Niknejad, N., W. B. Ismail, A. Mardani, H. Liao, and I. Ghani. 2020. "A Comprehensive Overview of Smart Wearables: The State of the Art Literature, Recent Advances, and Future Challenges." *Engineering Applications of Artificial Intelligence* 90 (April): 103529.
- Paschalidis, E., C. F. Choudhury, and S. Hess. 2019. "Combining Driving Simulator and Physiological Sensor Data in a Latent Variable Model to Incorporate the Effect of Stress in Car-Following Behaviour." *Analytic Methods in Accident Research* 22 (June): 100089.
- Perlman, D., A. Samost, A. G. Domel, B. Mehler, J. Dobres, and B. Reimer. 2019. "The Relative Impact of Smartwatch and Smartphone Use While Driving on Workload, Attention, and Driving Performance." *Applied Ergonomics* 75 (February): 8–16.
- Ranjbar, P., P. K. Krishnakumari, J. Andersson, and M. Klingegård. 2022. "Vibrotactile Guidance for Trips with Autonomous Vehicles for Persons with Blindness, Deafblindness, and Deafness." *Transportation Research Interdisciplinary Perspectives* 15 (September): 100630.

- Rojas, M. B., IV, E. Sadeghvaziri, and X. Jin. 2016. "Comprehensive Review of Travel Behavior and Mobility Pattern Studies That Used Mobile Phone Data." *Transportation Research Record* 2563 (1): 71–79.
- Sadeghvaziri, E., R. Javid, and M. Jeihani. 2023. "Active Transportation for Underrepresented Populations in the United States: A Systematic Review of Literature." *Transportation Research Record*, 03611981231197659.
- Sadeghvaziri, E., R. Javid, M. Jeihani, and Equity Center. 2023. "Investigating Walking and Biking Activities Among Low-Income African Americans."
- San-Segundo, R., H. Blunck, J. Moreno-Pimentel, A. Stisen, and M. Gil-Martín. 2018. "Robust Human Activity Recognition Using Smartwatches and Smartphones." *Engineering Applications of Artificial Intelligence* 72 (June): 190–202.
- Tepper, K. 2022. "Subject Guides: Systematic Review: Home." 2022. <https://guides.lib.monash.edu/c.php?g=943295&p=6955838>.
- TRID (Transport Research International Documentation). 2023. "Home - Transport Research International Documentation - TRID." 2023. <https://trid.trb.org/>.
- TRR (Transportation Research Record). 2023. "Transportation Research Record (TRR)." 2023. <https://journals.sagepub.com/home/trr>.

Use of Synthetic Aperture Radar and Uncrewed Aerial Vehicle for Assessing Pavement Condition

Amit Gajurel, S.M.ASCE¹; Nripojyoti Biswas, Ph.D.²; Hiramani Chimauriya, S.M.ASCE³;
and Anand J. Puppala, Ph.D., P.E., F.ASCE⁴

¹Doctoral Candidate, Zachry Dept. of Civil and Environmental Engineering, Texas A&M Univ., College Station, TX. Email: amitgajurel@tamu.edu

²Senior Research Engineer, Zachry Dept. of Civil and Environmental Engineering, Texas A&M Univ., College Station, TX. Email: nripojyoti.biswas@tamu.edu

³Doctoral Candidate, Zachry Dept. of Civil and Environmental Engineering, Texas A&M Univ., College Station, TX. Email: hiramani12@tamu.edu

⁴A.P. and Florence Wiley Chair Professor, Zachry Dept. of Civil and Environmental Engineering, Texas A&M Univ., College Station, TX. Email: anandp@tamu.edu

ABSTRACT

The long-term monitoring of transportation infrastructure assets, such as pavements, highway embankments, and earth retaining systems, at a lower cost, and with short mobilization time, is of interest to owners and stakeholders. Due to the improvement in spatial and temporal resolution of synthetic aperture radar (SAR) remote sensing systems as well as the drastic reduction in the price of data acquisition, SAR has now become a viable method to provide an economical and rapid condition assessment of transportation assets. This paper presents an example of this through a case study on the inspection and characterization of a pavement surface based on the amplitude of backscattering from an X-band radar. In situ characterization of the test site was first performed using aerial photogrammetry techniques with uncrewed aerial vehicle (UAV). The pavement features, extracted from the digital twin model created using photogrammetry, were compared with the radar cross section (RCS) of the pavement. The results show that problematic areas evident during in situ characterization can be delineated and quantified based on the normalized radar cross-section of the pavement surface. Overall, the outcomes of this research demonstrate the potential of SAR for future transportation asset management undertakings and can be of significant interest to state engineers and practitioners.

Keywords: Remote Sensing, Synthetic Aperture Radar (SAR), Transportation Assets, Uncrewed Aerial Vehicle (UAV)

INTRODUCTION

Transportation assets are an indispensable and intricate system of infrastructures that provide an efficient and safe mode of transit for goods and people. Timely monitoring and maintenance of these assets is of major interest as they support and contribute abundantly to the day-to-day lives of billions of people. The average age of transportation assets has been increasing – indicating a deterioration in asset condition in the long run (USDOT 2021, 2023). The Transportation Asset Management (TAM) plan mandates state Department of Transportations (DOTs) to annually and/or biennially inspect certain assets – bridges and pavement, and encourages inspection of other assets – walls, slopes, embankments, and subgrades within the

pavement Right-of-Way (ROW). The primary goal of TAM is to maintain physical assets at an acceptable level over its lifecycle at minimum cost to the stakeholders. Among many assets, pavements and bridges have been the primary focus of transportation agencies including those in the south-central states. Such focus has indeed improved the condition of the higher functioning pavements and bridges since 2000 (USDOT 2021, 2023). Besides pavement and bridges, geotechnical assets - such as walls, slopes, embankments, and subgrades, within its vicinity have a significant effect on the lifecycle of a transportation system. However, the lack of data on inventory, condition data, and performance measures of assets other than pavement and bridges on NHS limits the scope for TAM applications.

Recent years have seen a significant boom in space-borne remote sensing platforms like Synthetic Aperture Radar (SAR) satellites in both public and private sectors. Such rapid development has resulted in free access to large archives of medium resolution data— with global coverage every 6 - 12 days, and significant drop in image acquisition cost for high-resolution data (ICEYE 2023; Ignatenko et al. 2020; Torres et al. 2012). Unlike traditional optical sensors, SAR can collect data irrespective of weather conditions, cloud cover, and the presence of sunlight. Such data have a potential to be used for rapid and on-demand condition assessment and performance monitoring of transportation assets. Data from SAR can be utilized to effortlessly extend the current TAM scope as well as increase the monitoring frequency of assets as needed. Recent studies have shown that these sensor can be used to obtain data during inclement weather, natural disasters, and other hazard events (Ardila et al. 2022; Kourkouli 2023). The continued development of SAR satellite constellation has made possible capabilities like monitoring an asset at a given area multiple times in a single day with up to 0.5 m ground resolution irrespective of illumination from sunlight and weather conditions (Paek et al. 2020). This advancement is also accompanied by a sharp decrease in data acquisition and processing costs. These developments make it possible to closely monitor transportation systems, quickly identify distress, and swiftly deploy countermeasures that aid in preserving existing assets and enhancing service life.

Advancement in digital sensors for taking images and storage technology enabled capturing numerous high resolution images affordable in mid to late 2000s. FM is a photogrammetric method used in UAVs to generate high resolution 3D structures for a series of overlapping images that is typically derived from moving sensor (Snavely et al. 2008). Scale and rotation invariant key features are identified in each of these overlapping image which are then used to estimate camera pose and scene geometry to extract point cloud in image-space (Snavely et al. 2008; Westoby et al. 2012). The 3D point clouds are transformed to real-world co-ordinate system using Ground Control Points (GCPs) during the post-processing step. Redundant network of these evenly distributed GCPs, which are high-contrast targets both in field and point cloud, is established to account to any potential issues with sparse data or errors in SfM reconstruction (Congress et al. 2018; Westoby et al. 2012).

In this study, the authors illustrate an application of high-resolution X-band SAR data for analyzing pavement surfaces, utilizing backscattering amplitude as the primary metric. The methodology involves a comprehensive in-situ assessment of the pavement utilizing advanced aerial photogrammetry using Uncrewed Aerial Vehicles (UAVs) platform. This process generated a digital twin model of the pavement, from which specific features are extracted and subsequently correlated with the radar cross-section data of the same surface. The findings of this research highlight the ability to identify and quantify problem areas on the pavement, as detected during the in-situ evaluation, by analyzing the normalized radar cross-section. The

implications of this study are significant, demonstrating the utility of Synthetic Aperture Radar (SAR) in the field of transportation asset management. This research holds considerable relevance for state engineers and professionals in the field, underscoring the potential benefits of SAR technology in future infrastructure maintenance and management projects.

SYNTHETIC APERTURE RADAR (SAR)

SAR sensors are imaging radars that operate at microwave range, i.e., the wavelength of 0.75 cm to 120 cm, and provide unique information about the incident surface which is distinct from conventional optical and infrared images (Moreira et al. 2013). Unlike passive sensors, which rely on external sources of radiation like sunlight, SAR systems actively transmit electromagnetic waves and measure the backscatter from various surface features. The transmitter generates successive bursts of the chirped pulse at regular intervals which are focused by the antenna into a beam. The beam illuminates the surface obliquely in the direction right angle to the flight line (also known as range direction) and the SAR sensor system receives a portion of the transmitted energy. This is used to obtain the Radar Cross Section (RCS) – the inferred target area based on the ratio of the received response signal intensity and transmitted signal intensity. For a given wavelength and observation angle, RCS is dependent on the physical properties, such as geometry and roughness, and electrical properties, such as the dielectric constant of the incident surface. Depending on the type/orientation of the area used to normalize, a number of RCS measures are estimated. Radar brightness (β^0) is the quantity measured by a SAR sensor which is equal to RCS normalized with a solid rectangle defined in the slant range plane representing the ground surface. The Normalized Radar Cross Section (NRCS), also known as sigma nought (σ^0) is calculated by normalizing RCS with an actual geometrical area on the ground surface. σ^0 is the primary value of interest to the image analyst as it quantifies the ground surface response and minimizes the radiometric difference with change in incidence angles, acquisition geometry, and look direction (Small 2011; Woodhouse 2006). For this study, the equation used for calculating σ^0 is given by Equation 1 – calibration equation from provided by Capella Space (Capella Space 2022).

$$\sigma_{\text{dB}}^0 = 20 \log_{10}(\text{SC} * \text{DN}_{\text{geo}}) \quad (1)$$

Where,

σ_{dB}^0 = Normalized Radar Cross section (NRCS) (m^2/m^2) in dB scale

SC = Scale factor provided in the metadata for each scene

DN_{geo} = UInt16 raster value of the image

The radar response of surfaces like pavement, low-vegetation fields, and bare soils is dominated by surface scattering – which is the main contributor to the NRCS. Surface scattering is primarily the function of the roughness of the surface, the wavelength of the SAR sensor, and the incidence angle. Therefore, for a given wavelength and incidence angle of the SAR sensor, the roughness of the surface becomes the main contributing factor for the radar response of pavements and alike surfaces.

One of the criteria to classify the smoothness or roughness of a surface with respect to the incident electromagnetic wave is the Fraunhofer roughness criterion given by Equation 2 (Jensen 2014; Ulaby and Long 2014; Woodhouse 2006). This criterion is useful in modeling the scattering and emission behavior of natural surfaces in the microwave region where the

wavelength (λ) is of the same order as that of the root mean square (rms) height of the surface. It is calculated based on the formula shown in Equation 3. In this study, the SAR system is a X-band ($\lambda = 3.1$ cm) radar with incidence angle (θ) of 30.6°. Substituting these values in Equation 3, shows that for $s < 0.11$ cm, the surface can be considered smooth.

$$s < \frac{\lambda}{32 \cos(\theta)} \quad (2)$$

$$S = \left[\frac{1}{N-1} \sum_{i=1}^N z_i^2 - N \bar{z}^2 \right]^{\frac{1}{2}} \quad (3)$$

Where,

$s = \text{rms height}$

λ = wavelength of electromagnetic wave

θ = incidence angle

N = number of samples

$$\bar{z}_l = \frac{1}{N} \sum_{i=1}^N z_i$$

z_i = Height profile of the surface in consideration

The rms heights (roughness) from UAV platform as well as mean NRCS data (radar response) from SAR satellite were extracted for analysis. The analysis was performed for 0.5 m segment length.

CASE STUDY: PAVEMENT IN RELLIS CAMPUS, BRYAN, TX

The site for this case study was the Proving Grounds Research Facilities located at the 2,000-acre RELLIS Campus of Texas A&M University System in Bryan, Texas. This facility contains multiple runways, aprons, and transportation-related pavements. This study includes RTA Zone 2 35C Sect 3 as shown in Figure 1.



Figure 1: Location of the study site at RELLIS Campus in Texas A&M University

Digital twin model. UAV was used to collect high-resolution aerial images of the site as shown in Figure 2(a). A reconnaissance survey of the site was used to finalize the flight plan and locations of the ground control points. Data acquisition was performed using the UAV platform

equipped with an optical sensor, Real Time Kinematics (RTK) navigation system, and Post Processing Kinematics (PPK) geotagging system. The GPS information of each image from the UAV platform was corrected using the data from the PPK system after flight. Similar processes have been used by several researchers in correcting the GPS coordinate post-flight (Congress and Puppala 2022; Thomas et al. 2020; Zhang et al. 2019). The images with corrected GPS information and Ground Control Points (GCPs) were ingested into a SfM photogrammetry software to generate dense point cloud, Digital Surface Model (DSM), and ortho-mosaic. DSMs and ortho-mosaics were exported to GIS environments for creating 3D models and extracting elevation data. DSM overlain by the ortho-mosaic facilitated the creation of a digital twin model of the site with a spatial resolution of 0.49 cm Figure 2(b).

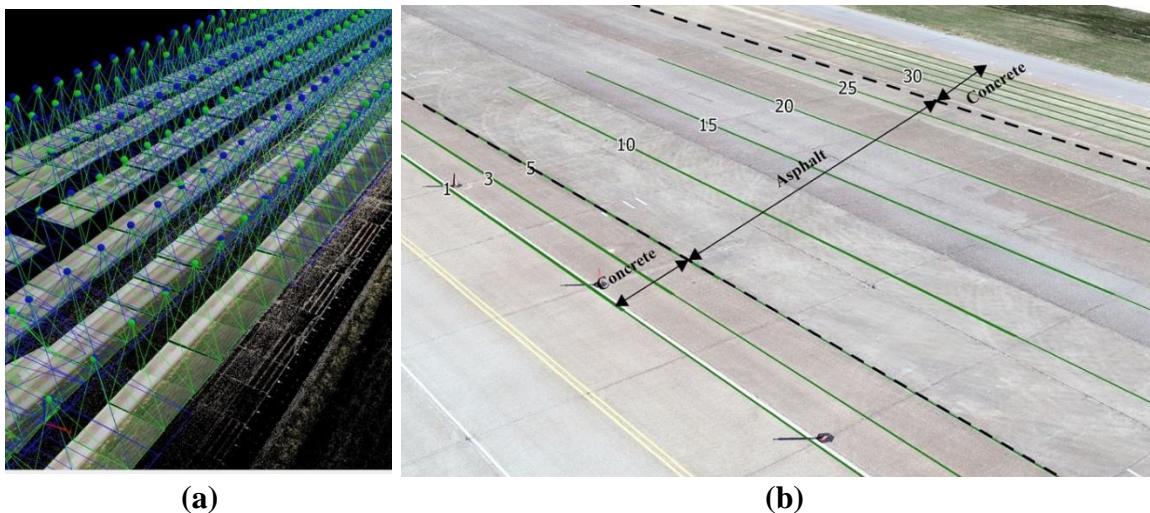


Figure 2: UAV data of the site (a) dense point could generation (b) digital twin

The pavement elevation profiles along a 200 m long stretch of the sections labelled in Figure 2(b) were extracted using ENVI. The rms value along these sections are computed using Equation 3 for segment length of 0.5 m. for both concrete and asphalt pavement sections.

High resolution SAR. Satellite tasking, data acquisition, and download of high-resolution SAR data was performed through Capella Space® for the RELLIS campus (Figure 3). Spotlight Geocoded Terrain Corrected (GEO) SAR data product downloaded for this site is a one dimensional raster, where each pixels has been calibrated for radar internal subsystems and corrected for location (Capella Space 2022). The imagery was imported into a GIS environment and calibrated to obtain a raster with NRCS in decibels (dB) scale. Geo-referencing of the SAR data was performed using 1st-order affine transformation to correctly align the SAR data. NRCS along the profile sections corresponding to the digital twin section were extracted.

RESULTS AND DISCUSSION

The NRCS value in dB is plotted against the rms (in cm.) value along the profile (both concrete and asphalt surfaces) for segment length of 0.5 m to access relationship between them. Figure 4, Figure 6, and Figure 7 shows the comparison between NRCS value and rms value for the concrete pavements on either side of the central asphalt pavement. Figure 5 and Figure 6 shows similar comparison between profile sections on the asphalt pavement. It is evident that the

rms value, calculated from the digital twin model do not show variation among different pavement types whereas the NRCS of the pavement surface shows variation. This suggests that SAR data can capture the subtle surface texture difference between asphalt and concrete pavements. This is clearly visible in the case of comparison of NRCS value along Line 10, 15, 20, and 25 (which are all asphalt pavements) with other sections. In case of Line 25, the pavement surface looks aged and has early signs of alligator cracks along its length in comparison to other asphalt pavement (shown in Figure 8) – which in turn has higher value of NRCS. Although the NRCS is higher, the rms value of the surface remains unaffected by these distresses. This suggests that initiation of pavement failure with distress like alligator cracking are visible with SAR data at its early stages than with measuring surface profiles. The concrete pavements, as visible in the orthomosaics, do not show any visible deterioration and seem to be in good condition. Therefore, the rms value of the surface is consistently low in all the profiles. For a concrete pavement without any visible surface distresses, the NRCS value of the surface had a mean value of -18.6 dB with standard deviation of 1.4 dB. This suggests that using an X-band radar a good concrete pavement can be identified. Similarly, for an asphalt pavement without any visible surface distresses, the mean NRCS value of -15.8 dB with standard deviation of 1.5 dB. Likewise, for line 25 (shown in Figure 8), the average NRCS value if -13.5 dB with standard deviation of 1.6 dB.

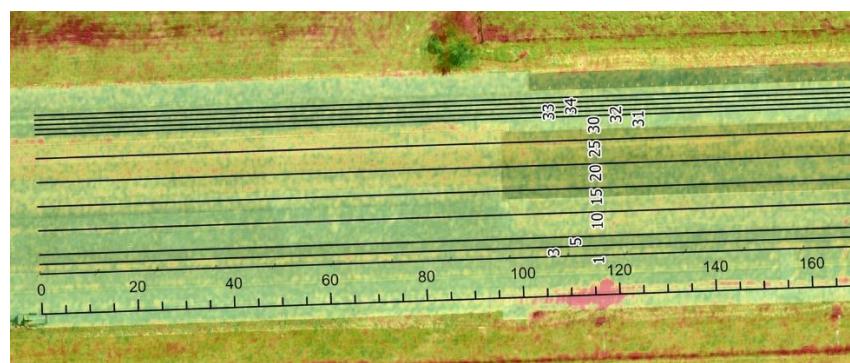


Figure 3: Location of profile sections overlain by SAR amplitude data

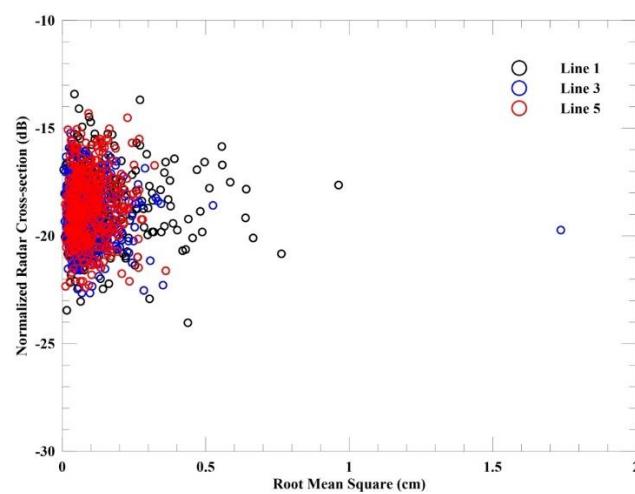


Figure 4. NRCS vs rms value for 0.5 m segment length for Line 01, 03, and 05

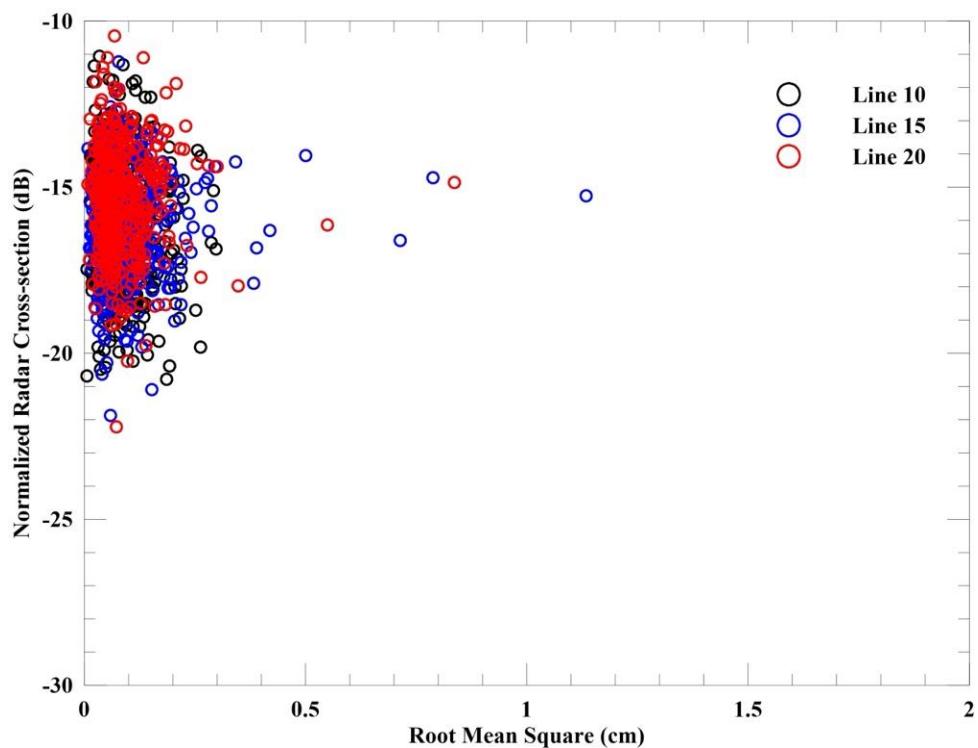


Figure 5. NRCS vs rms value for 0.5 m segment length for Line 10, 15, and 20

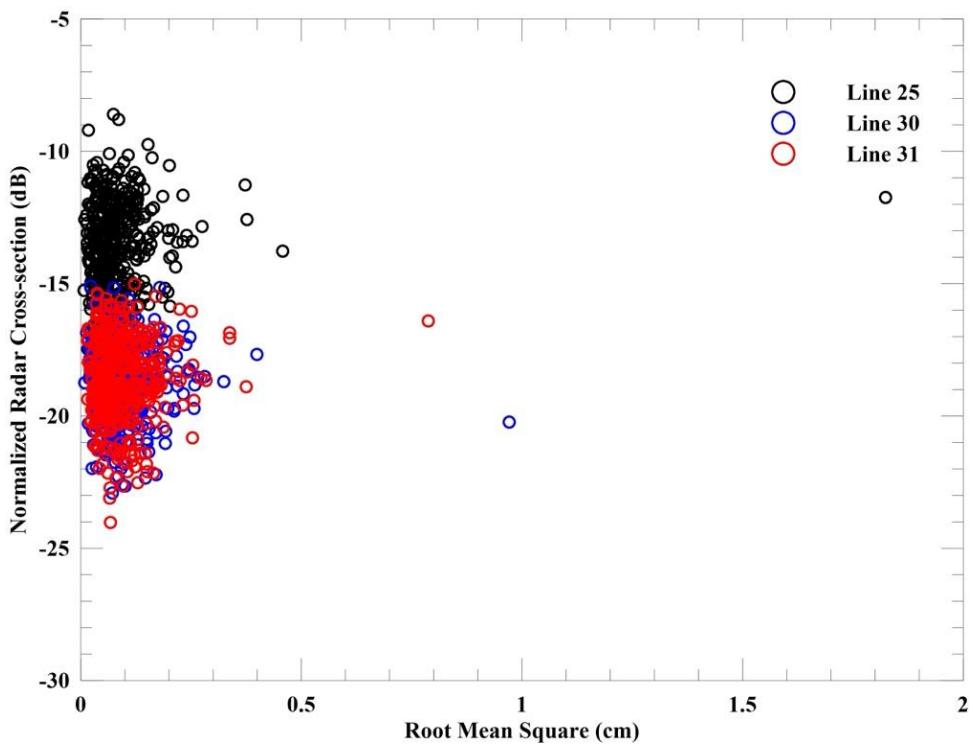


Figure 6. NRCS vs rms value for 0.5 m segment length for Line 25, 30, and 31

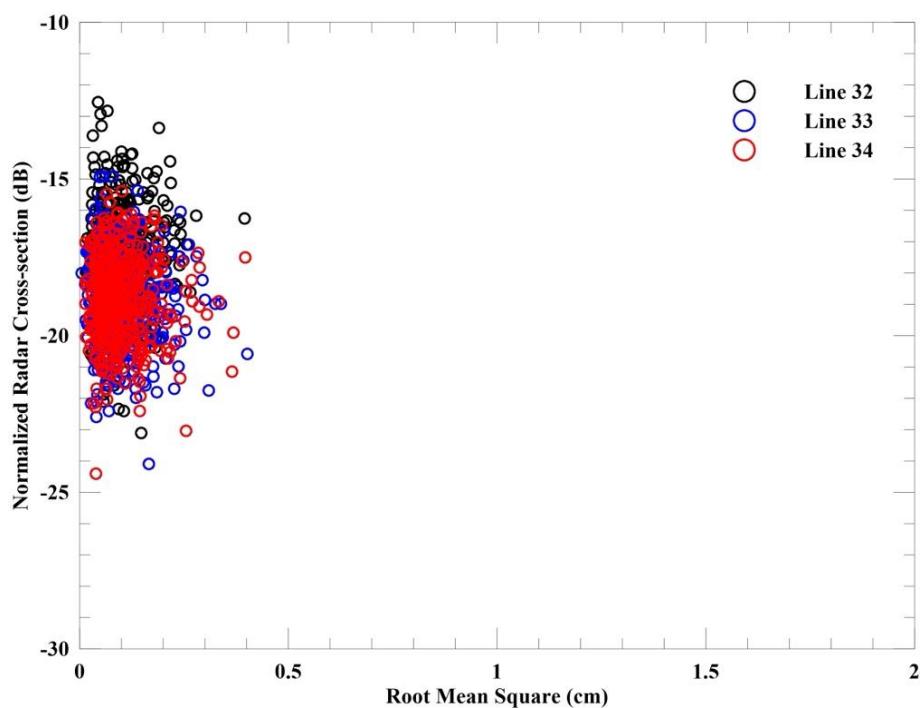


Figure 7. NRCS vs rms value for 0.5 m segment length for Line 32, 33, and 34



Figure 8. Early-stage alligator cracks present along Line 25

CONCLUSIONS

The study showcased a noteworthy relationship between the NRCS values and the surface profiles of various pavement types. Even though rms values derived from the digital twin model didn't display a significant distinction between different pavements, the NRCS distinctly captured the subtle differences in surface texture between concrete and asphalt pavements. Key findings include:

- The analysis distinctly highlighted how SAR data can recognize early signs of pavement failures, such as alligator cracking, even when there aren't noticeable variations in the rms values.
- The concrete and asphalt pavements appeared to be in good condition with consistent but different rms values, underlining the robustness of X-band SAR data.
- By employing an X-band radar, it is possible to identify the spatial location and condition of good concrete and asphalt pavements based on their respective mean NRCS values.
- High resolution X-band SAR data can be a reliable tool for detecting early signs of pavement distress and for characterizing the integrity of different pavement materials.

REFERENCES

- Ardila, J., P. Laurila, P. Kourkouli, and S. Strong. 2022. "Persistent Monitoring and Mapping of Floods Globally Based on the Iceye Sar Imaging Constellation." *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, 6296–6299. Kuala Lumpur, Malaysia: IEEE.
- Capella Space. 2022. "Capella Space Synthetic Aperture Radar (SAR) Open Dataset." Accessed November 22, 2022. https://registry.opendata.aws/capella_opendata.
- Congress, S., and A. Puppala. 2022. "Eye in the sky: Condition monitoring of transportation infrastructure using drones." *Proceedings of the Institution of Civil Engineers: Civil Engineering*, 1–9.
- Congress, S. S. C., A. J. Puppala, and C. L. Lundberg. 2018. "Total system error analysis of UAV-CRP technology for monitoring transportation infrastructure assets." *Engineering Geology*, 247: 104–116.
- ICEYE. 2023. "Satellite Data - How to use satellite data for better decision making." Accessed April 30, 2023. <https://www.iceye.com/satellite-data>.
- Ignatenko, V., P. Laurila, A. Radius, L. Lamentowski, O. Antropov, and D. Muff. 2020. "ICEYE Microsatellite SAR Constellation Status Update: Evaluation of First Commercial Imaging Modes." *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, 3581–3584. Waikoloa, HI, USA: IEEE.
- Jensen, J. R. 2014. *Remote sensing of the environment : an earth resource perspective*. Pearson.
- Kourkouli, P. 2023. "Natural disaster monitoring using ICEYE SAR data." *Geoinformatics for Geosciences*, 163–170. Elsevier.
- Moreira, A., P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou. 2013. "A tutorial on synthetic aperture radar." *IEEE Geosci. Remote Sens. Mag.*, 1 (1): 6–43.
- Paek, S. W., S. Balasubramanian, S. Kim, and O. De Weck. 2020. "Small-Satellite Synthetic Aperture Radar for Continuous Global Biospheric Monitoring: A Review." *Remote Sensing*, 12 (16): 2546.

- Small, D. 2011. "Flattening gamma: Radiometric terrain correction for SAR imagery." *IEEE Transactions on Geoscience and Remote Sensing*, 49 (8): 3081–3093.
- Snavely, N., S. M. Seitz, and R. Szeliski. 2008. "Modeling the World from Internet Photo Collections." *Int J Comput Vis*, 80 (2): 189–210.
- Thomas, O., C. Stallings, and B. Wilkinson. 2020. "Unmanned aerial vehicles can accurately, reliably, and economically compete with terrestrial mapping methods." *Journal of Unmanned Vehicle Systems*, 8 (1): 57–74.
- Torres, R., et al. 2012. "GMES Sentinel-1 mission." *Remote Sensing of Environment*, 120: 9–24.
- Ulaby, F. T., and D. G. Long. 2014. *Microwave Radar and Radiometric Remote Sensing*. Ann Arbor: The University of Michigan Press.
- USDOT. 2021. *National Transportation Statistics 2021*. Washington, DC: US Department of Transportation, Bureau of Transportation Statistics.
- USDOT. 2023. *Transportation Statistics Annual Report 2022*. United States. Department of Transportation. Bureau of Transportation Statistics. Not Available.
- Westoby, M. J., J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds. 2012. "'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications." *Geomorphology*, 179: 300–314.
- Woodhouse, I. H. 2006. *Introduction to Microwave Remote Sensing*. CRC Press. Boca Raton: Taylor & Francis.
- Zhang, H., E. Aldana-Jague, F. Clapuyt, F. Wilken, V. Vanacker, and K. Van Oost. 2019. "Evaluating the potential of post-processing kinematic (PPK) georeferencing for UAV-based structure-from-motion (SfM) photogrammetry and surface change detection." *Earth Surface Dynamics*, 7 (3): 807–827.

Using AI and Change Detection in Geospatial UAS Airfield Pavement Inspection for Pavement Management

Michael T. McNerney, Ph.D., P.E., M.ASCE¹; Grant Bishop²; and Valerie Saur³

¹Dept. of Civil Engineering, Univ. of Texas at Arlington, Fort Worth, TX
(corresponding author). Email: michaelmcnerney@gmail.com

²President and CEO, Silent Falcon UAS Technologies, Oakton, VA.
Email: gbishop@silentfalconUAS.com

³Vice President of Operations, Silent Falcon UAS Technologies, Front Royal, VA.
Email: vsaur@silentfalconUAS.com

ABSTRACT

The authors have conducted detailed pavement inspections at over 100 airports using uncrewed aerial systems with high resolution photogrammetry to identify pavement distresses and calculate pavement condition index using the formula of ASTM D-5340. New aerial vehicles with new 100-megapixel cameras allow the team to collect the same 1.5 mm pixel GSD imagery three times faster than the previously reported data collection system. The team uses artificial intelligence coupled with deep machine learning models to automatically identify both distress type and distress severity levels. The large library of distresses from 140 airports that have previously been flown and verified has greatly improved the accuracy of the distress identity and distress severity determination. Although the AI automatically generated distress determinations are reviewed by technicians, the quality of the AI generated distresses and severities have greatly improved from what we reported two years ago. The result is that distresses that may have gone unnoticed with a visual scan are being discovered and correctly identified by the AI. Although an FAA report has claimed that the UAS technology is not progressed enough to identify several low severity distresses (alkali-silica reaction, corner spalling, joint spalling, joint seal damage, depression, raveling, swell and weathering), the authors' analyses using AI and deep machine learning have identified each of these low severity distresses without additional imagery. Now that the team is scanning many airports for a second time, we are using change detection at the pixel level to see changes that are happening to distresses and distress severities. Using AI and change detection software together has yielded important results. It provides the ability to visually see in what geospatial area specific distresses are increasing. A change of 1–3 PCI points is normally expected in year-to-year inspections. However, determining what types of distress changes are taking place is much more meaningful to the pavement management program (PMP).

BACKGROUND

The Federal Aviation Administration Office of Airports requires that airports receiving federal funds must have a Pavement Management Program (PMP) which is described in Advisory Circular 150/5380-7B (FAA 2014). This requirement comes from Public Law 103-305 passed in Aug 23, 1994, that all airports have implemented an effective pavement maintenance-management program as a condition of assistance. The consequences of not having a pavement maintenance management plan are that airports not documenting good maintenance could be denied new pavement grants.

AC 150/5380-7B has added some definition to what constitutes a Pavement Management Plan. Federally obligated airports must perform a detailed inspection of airfield pavements annually. If a pavement condition index (PCI) survey is performed, as set forth in ASTM D5340, Standard Test Method for Airport Pavement Condition Index Surveys, the **frequency** of detailed inspections by PCI surveys may be extended to three years.

In 2020 and 2021 the William J Hughes Technical Center conducted a research project on using UAS for pavement inspections using a variety of UAS platforms and sensors. The report is not yet published, and the authors took part in this research project flying UAS at three airports and the FAA National Pavement and Materials Research Center on pavements that had been rutted using the Heavy Vehicle Simulator - Airfields (HVS-A). The Goal of the FAA technical is to develop a specification for UAS pavement inspection in 2025 (FAA 2022). However, the airport industry and especially the general aviation airports want to use this UAS pavement inspection technology now.

The authors have the most experience in UAS pavement inspections and PCI calculation by a large margin having scanned over one hundred airports in the US, Canada and Mexico. Also, while others have used UAS for PCI data collection, the authors have the distinction of using Artificial Intelligence Deep Learning Neural Network (AI-DLNN) Process. This patented process was presented in a paper in the T&DI International Conference on Transportation and Development in Seattle in 2022 (McNerney 2022).

In 2023 the authors presented in Austin at the Airfield and Highway Pavement Conference the lessons learned for inspecting over eighty airports (McNerney 2023). This third paper in the series details the amazing results from using change detection software at airports that Silent Falcon previously inspected to identify distress changes in only one year. The authors have also used AI to identify low severity depression with Digital Elevation Models (DEM) on an airport runway.

BENEFITS OF A DETAILED PAVEMENT INSPECTION PROCESS USING UAS

All airports can benefit from UAS scans because they reduce risk, provide 100% data and imagery for all airfield pavements quickly and very economically. The FAA is aware of the potential cost savings and has embarked on a research project at the William J. Hughes Technical Center in Atlantic City to study the capabilities required. The authors have participated in the data collection for the research project with two years of firsthand experience in providing to small and medium airports imagery and pavement condition surveys in geospatial format.

There is a benefit to having 100 percent coverage of distresses on the airfield pavements rather than a statistical sample as suggested in the ASTM D5340 standard. There is also a benefit to having all the distresses and severities quantified and geolocated for the planning of maintenance and rehabilitation actions. Having all this data in a GIS database such as ArcGIS Pro that the airport clients can review the imagery of all the distressed on their desktop is a tremendous benefit for visualizing the data and making informed decisions for pavement management.

The FAA requirement for PMP is for annual detailed pavement inspections beyond routine maintenance inspections. The FAA requirement for PMP also states that if a pavement inspection is conducted using the procedure of ASTM D5340, the inspection frequency can be delayed until every three years.

However, because of the simpler and less expensive UAS inspection process, we are finding that there is benefit in annual inspections. Using the change detection evaluation of airfield

pavements, we are finding amazing results of identifying new distresses or expansion of existing distress in only one year. Silent Falcon has conducted over fifty airport second year UAS inspections in Utah. The results have exceeded expectations. The results have been amazingly accurate, and we found more changes than were expected.

Second Year Data Collection Process. Second year data collection process is the same as the first year with lessons learned that makes the planning and execution go faster. In Utah the UAS used for the first year required operations at 33 feet above ground (AGL) level to achieve a 1.5 mm pixel resolution. In the second year the UAS used permits flying at 133 feet AGL to achieve the same 1.5 mm pixel resolution which saves time and can be completed with less personnel.

The airport is divided into branches (runways, taxiways, and aprons) and that is further subdivided into smaller sections that are based on total flight time of the UAS being used. Each section is then assigned a mission number. All flight operations are pre-planned before arriving at the airfield and each pilot is tasked with specific missions and times to ensure time, space and altitude deconfliction. Pilots do not “fly” the aircraft manually. They manage a pre-programmed flight that adheres to the FAA AC standards for remote sensing. The flight team is located around the airfield to minimize flight time to/from the tasking area and to also avoid conflicts with air traffic. Clear, concise communications are maintained throughout the mission profile between ATC, Airport Operations, Aircraft and Airport stakeholders. All support systems for the UAS are managed and self-contained with the SF team. After all missions are completed, the data is transferred to a secure cloud network. This allows for virtual processing of the data by team members regardless of physical location.

CHANGE DETECTION ANALYSIS FOR AIRFIELD PAVEMENTS

Silent Falcon has entered into an exclusive licensing agreement with Collins Aerospace to use their change detection software for airfield pavement evaluation of changes in distresses. The process is a pixel-to-pixel imagery change detection evaluation. The results have been stunning.

Examples of the change detection are shown in Figure 1 and Figure 2 of the pavement evaluation of Dutch John Airport, (33U), Utah. The 2022 images for Dutch John Airport were taken by Silent Falcon UAS on 9 Aug 2022 with an Autel Evo 2 quadcopter UAS. The second scan was taken on 8 June 2023 with an Inspired Flight IF1200A quadcopter UAS using a Phase One 100mp camera.

In Figure 1 the low severity cracking detected in 2022 is shown in the yellow box which may or not be visible based on the print resolution. In 2023 the crack has extended all the way to sealed paving joint on the right side of the image. The change detection shows in white exactly what new cracking has been detected as different. The brightness in the change detection indicates the depth of the lower elevation, with less bright indicated less elevation change.

In Figure 2 in 2022 there is an existing left to right crack that looks as if it has been painted over by striping the centerline. The crack is well sealed and there are no depressions in the sealant. There is a second partial crack on the left side below the main crack. There is no top to bottom cracking visible in 2022.

In 2023 the top to bottom crack is highly visible. The main left to right crack is visible through the paint stripe and there are considerable depressions in the crack sealant. The second crack is visible, but it is not deep. However, the second crack has spalled where it crosses the top to bottom crack.

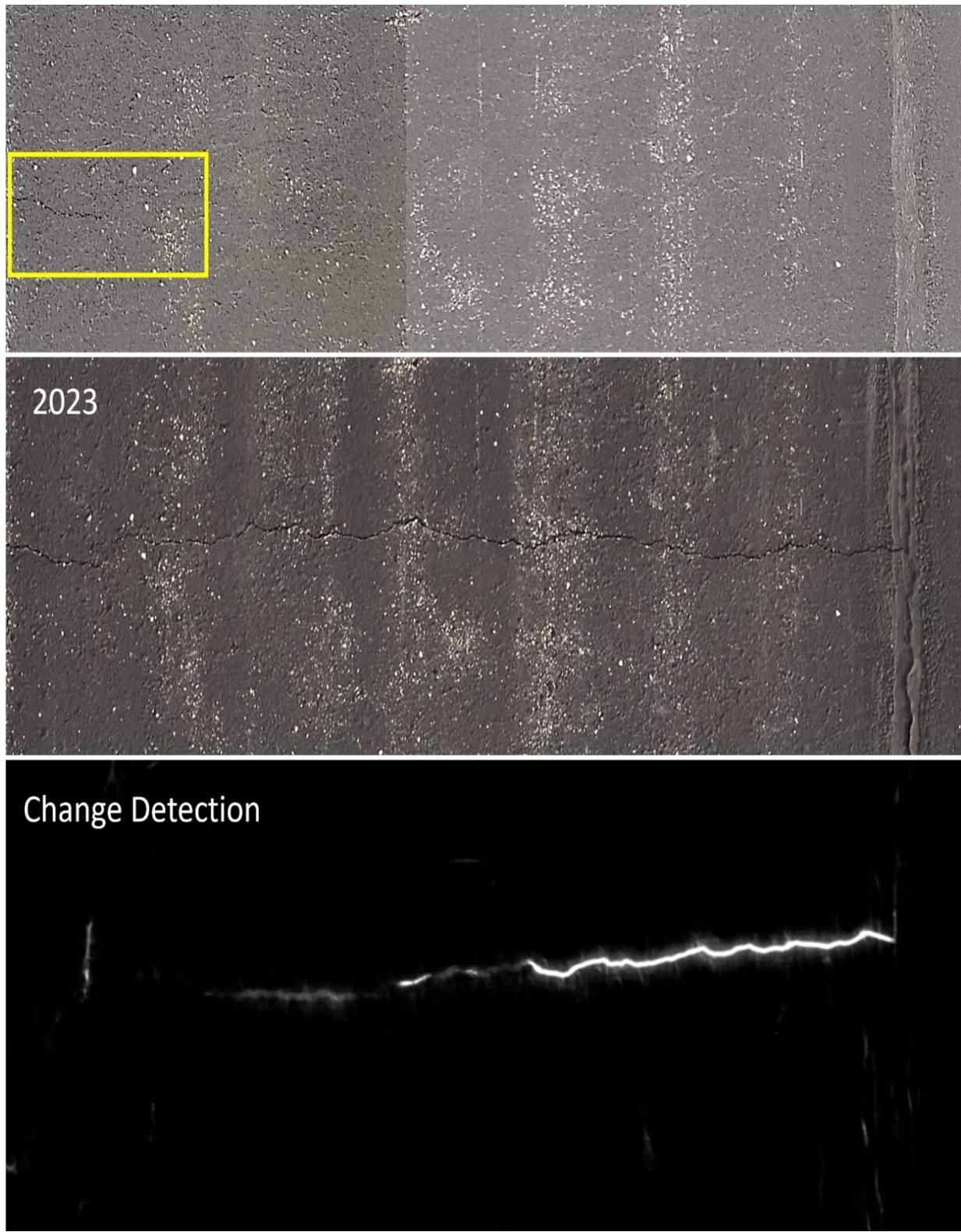


Figure 1. Example 1 of Change Detection of Linear Cracking at Dutch John Airport, Utah

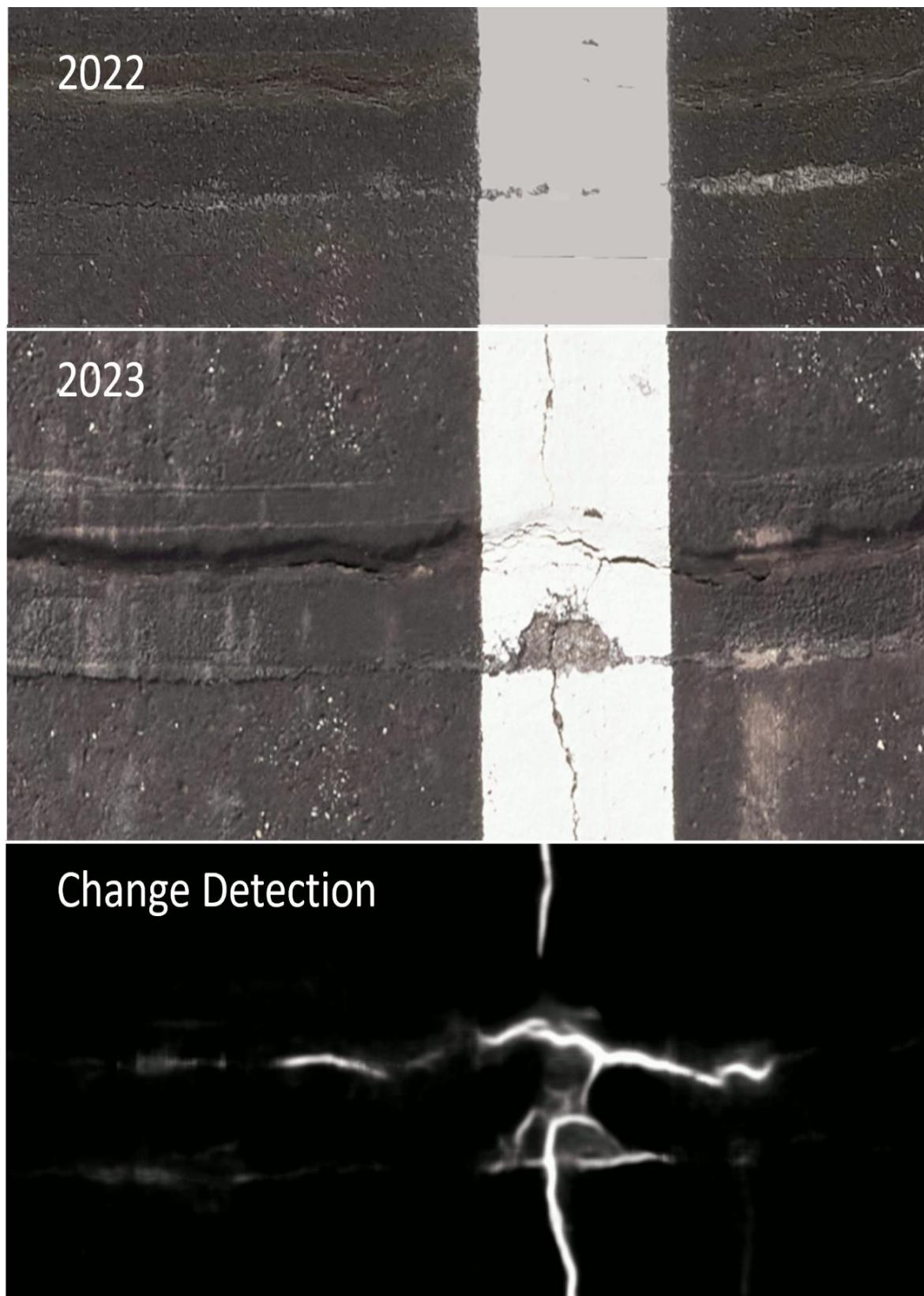


Figure 2. Example 2 of Change Detection of Linear Cracking at Dutch John Airport, Utah

The change detection portion of Figure 2 shows in white, only that cracking that has changed and the elevation difference is indicated by the brightness. The new vertical cracking displays as very bright and bold. The main left to right crack has some areas which are brighter than others based upon what is new and what has changed little. The spalled area of the cracking is shown lightly in brightness because there is little or no depth to the spall.

Using Change detection at Dutch John Airport (33U) we were able to quantify the increase in cracking that happened in the 10-month period from August to June. Table 1 shows that on Runway 11, there was a 37% increase in low severity cracking and 32% increase in medium severity cracking from 2022 to 2023.

In the branch Apron 1, sixty-three feet of medium severity longitudinal and transverse cracking increased to 121 feet. However, the longitudinal and transverse low severity cracking decreased but medium severity alligator cracking was observed that was previously only low severity.

On branch Taxiway 1, both low and medium severity linear cracking was reduced. However, fifty-three square feet of medium severity alligator cracking was observed in 2023.

Table 1. Changes in Longitudinal Cracking Identified from Change Detection

Branch	Distress Type and Severity	2022	2023	
Apron 1	Longitudinal and Transverse Cracking Low Severity	1,973	1,340	-32%
	Longitudinal and Transverse Cracking Medium Severity	63	121	92%
	Alligator Cracking Low Severity	152	0	
	Alligator Cracking Medium Severity	0	491	
Runway 11	Longitudinal and Transverse Cracking Low Severity	33,495	45,997	37%
	Longitudinal and Transverse Cracking Medium Severity	524	709	35%
Taxiway 1	Longitudinal and Transverse Cracking Low Severity	1,498	693	-54%
	Longitudinal and Transverse Cracking Medium Severity	169	132	-22%
	Alligator Cracking Medium Severity	0	53	

USING AI WITH DIGITAL ELEVATION MODEL TO IDENTIFY DEPRESSIONS

A digital elevation model (DEM) is a 3D computer graphic representation of elevation data to represent terrain or surface. When a photogrammetric image is orthorectified a DEM is created in the rectification process. DEM is often used in GIS to create digital terrain models (DTM) of the surface. Normally, this process is used in mapping large areas. However, using the high resolution of our imagery and precise control points, we have developed DEM or DTM of the runway service with sufficient accuracy for our AI and digital library to identify low severity depressions.

ASTM D5340 defines depressions as localized pavement surface areas having elevations slightly lower than those of the surrounding pavement. The severity of a depression is measured by its mean depth. A low severity depression on a runway has a mean depth of less than or equal to 6 mm. Medium severity depression on a runway is greater than 6 mm and less than 13 mm. High severity depression has a mean depth greater than 13mm. Taxiways and aprons have double the depth of allowable depression. Having a depression depth of less than 6 mm to be identified by UAS imagery was thought to be beyond the capability of UAS technology in FAA research published this year (FAA 2023). Although the research and report lags in current UAS technology and current Silent Falcon data proves that it is possible.

Silent Falcon conducted an Airport Scan of a commercial service airport in Mexico and the AI identified significant depressions on one runway. Initially it was thought that this was too much depression and after field verification it was determined that the AI was correct. The low severity depression does really exist on the runway. Figure 3 shows the polygon on the runway outline the areas of low severity depression.



Figure 3. Green Polygons of Low Severity Depression on Runway as Identified by AI

CONCLUSIONS

Using Artificial intelligence with high quality imagery and a large library of distresses gathered from many airports that can be used for distress identification, is a valuable option for UAS pavement inspection of airfield pavements. Currently, it is necessary for quality control purposes to have the human pavement engineer to review AI output and make corrections.

Using AI and digital elevation models it is possible to identify low severity depressions on a runway, which was previously reported as not yet technically achievable.

Using pixel-to-pixel change detection software can easily show new linear cracking in asphalt. Research is still in progress to use the change detection software to identify other visually observable distresses. By the time of this conference, we expect to be able to show examples of change detection software with other distresses in both asphalt and concrete pavements with a 3D resolution of 2mm.

REFERENCES

- ASTM. ASTM D5340-20, *Standard Test Method for Airport Pavement Condition Index Surveys*, ASTM International, Philadelphia, PA, 2020.

- FAA. Advisory Circular 150/5380-7B, Airport Pavement Management Program (PMP), Washington D.C., Nov10, 2014.
- FAA. DOTIF/FAA/TC-23-50, Small Unmanned Aircraft Systems for Pavement Inspection, William J. Hughes Technical Center, August 2023.
- McNerney, M. T., G. Bishop, and V. Saur. "Detailed Pavement Inspection of Airports using Remote Sensing UAS and Machine Learning of Distress Imagery" *International Conference on Transportation and Development 2022*, May 31–June 3, 2022, |Seattle, Washington, International Conference on Transportation and Development 2022: Other Modes—Rail, Transit, and Aviation, Edited by Heng Wei, Ph.D., P.E., Book set: ICTD 2022ISBN (PDF): 978078448437.
- McNerney, M. T., G. Bishop, and V. Saur. "Experiences Gained and Benefits from Using Uncrewed Aerial Systems to Calculate Pavement Condition Index at over 80 Airports in the United States. *Airfield and Highway Pavements 2023: Innovation and Sustainability in Airfield and Highway Pavements Technology*. pp 244-253, Edited by Navneet Garg, Ph.D.; Amit Bhasin, Ph.D., P.E.; and Julie Vandenbossche, Ph.D., P.E.

Integration of Small Unmanned Aircraft Systems and Deep Learning for Efficient Airfield Pavement Crack Detection and Assessment

Md. Abdullah All Sourav, Ph.D.¹; Halil Ceylan, Ph.D., Dist.ASCE²;
Sunghwan Kim, Ph.D., P.E.³; and Matthew Brynick⁴

¹Postdoc Research Associate, Dept. of Civil, Construction, and Environmental Engineering, Iowa State Univ., Ames, IA. Email: sourav.abe@gmail.com

²Professor and Director, Dept. of Civil, Construction, and Environmental Engineering and Program for Sustainable Pavement Engineering and Research, Iowa State Univ., Ames, IA. Email: hceylan@iastate.edu

³Associate Director, Program for Sustainable Pavement Engineering and Research, Institute for Transportation, Iowa State Univ., Ames, IA. Email: sunghwan@iastate.edu

⁴Civil Engineer, Federal Aviation Administration, Airport Pavement R&D Section, William J. Hughes Technical Center, NJ. Email: matthew.t.brynick@faa.gov

ABSTRACT

Airfield pavement inspection and maintenance are critical aspects of aviation infrastructure, representing a substantial portion of life-cycle costs. Longitudinal, transverse, and diagonal (LTD) cracks; corner breaks; shattered slabs in Portland cement concrete (PCC) pavement; and longitudinal and transverse (L&T) cracks of asphalt concrete (AC) pavement consist of most of the airfield pavement distresses. Traditional airfield pavement inspection methods are manual, time-consuming, laborious, and reliant on the inspector's experience, leading to increased expenses and safety risks. This research explores the potential to automatically identify those distresses in red-green-blue images using four variants of deep learning (DL) model YOLOv8, ranging from nano to large. YOLOv8 is a widely used off-the-shelf DL object detection model that allows rapid training and easy execution. A DL training dataset of 5,273 small uncrewed aircraft systems (sUAS) collected images was developed. The transfer learning technique was used, and the dataset passed through each model 100 times for adequate training. The model exhibits mean average precision values exceeding 0.65, with varying processing times. Such accuracy showed that crack-related distress detection using DL models could enhance airfield pavement inspection efficiency.

Keywords: Airfield pavement inspection, small Uncrewed Aircraft Systems (sUAS), cracks detection, pavement distress detection, deep learning models, drone for cracks detection, drone image, orthophoto, Deep Learning (DL)

INTRODUCTION

The traditional process of airfield pavement inspection involves manual inspection of some selected pavement areas, yielding numerical values crucial for subsequent maintenance and safety measures (Pietersen et al., 2022). Airfield pavement safety and maintenance contribute significantly to life-cycle costs, with a substantial portion of the U.S. government's 2019 airfield improvement program funds, 68.31% of the 3.32 billion USD, of airfield improvement program funds allocated to the maintenance and rehabilitation of apron, runway, and taxiway pavements

(FAA, 2022). The inspection is time-consuming, labor-intensive, dependent on inspector experience, and poses safety risks (Ellenberg et al., 2016; Hubbard & Hubbard, 2020; Pietersen et al., 2022).

The rising expenses associated with trained professional airfield pavement inspection underscore the need for innovative solutions. The adoption of small Uncrewed (Unpersonned/Unmanned) Aircraft Systems (sUAS), commonly known as a drone, has gained significant popularity in recent years for non-destructive inspection of transportation infrastructure. Especially in the U.S., different parts of the bridges are being inspected by sUAS and the need for limiting traffic flow during bridge inspections has been eliminated while ensuring the safety of the inspector, reducing total cost, and achieving high-details of different parts of the bridge (Hubbard & Hubbard, 2020; Seo et al., 2018; Wells & Lovelace, 2018). sUAS have found various applications in transportation infrastructure, including construction planning, work monitoring, mast light pole inspection, accident scenario recording, traffic flow monitoring, and vehicle counting (FHWA, 2018).

Despite the increasing use of sUAS, there has been limited research on airfield inspection and inventory management using this technology. Proof-of-concept studies conducted by airport authorities and industries in the U.S. and Europe demonstrate the potential of sUAS in collecting airfield pavement data (Airport Improvement, 2020; Airsight, 2020; Banks et al., 2018). Congress and Puppala (2022) conducted an inventory study using sUASs, collecting data on airfield pavements through DJI Matrice 210 RTK quadrotor. The study focused on creating orthophotos of 0.6 mm/pix to note pavement distresses, calculate pavement condition index values, and categorize good, fair, and poor pavement systems. Sourav, Ceylan, Brooks, et al. (2023), Sourav, Ceylan, Kim, et al. (2022), and Sourav, Mahedi, et al. (2022) conducted comprehensive studies on the usefulness of sUAS data for airfield pavement distress detection. Their research involved collecting red-green-blue (RGB) images and processing them to create orthophotos and digital elevation models (DEM). Following airfield pavement inspection procedures, image analysis manually detected distresses, classified their severity, and calculated pavement condition index values. The result demonstrated the usefulness of RGB orthophotos in detecting and rating at least one severity of 13 out of 14 available Portland Cement Concrete (PCC) pavement distresses and 6 out of 9 available Asphalt Concrete (AC) pavement in the test sites (Sourav, Ceylan, Brooks, et al., 2022). Additionally, DEM showed very promising performance in detecting vertical displacement distresses (faulting, shoving) (Sourav, Ceylan, Brooks, et al., 2023; Sourav, Ceylan, Kim, et al., 2022, 2023). It was shown that sUAS-collected images are useful in distress detection and rating. Thus, it can be said that sUAS data could be successfully used for airfield pavement distress detection.

In recent times, machine learning and deep learning (DL) object detection models have been applied to sUAS-collected data for pavement damage detection, focusing primarily on road pavement cracks and potholes (Y. Li et al., 2022; Zhu et al., 2022). With proper training, object detection models can deal with different lighting conditions and complex scenarios while localizing distresses with higher accuracy. Previous studies (R. Li et al., 2023; Y. Li et al., 2022; Zhu et al., 2022) often focused on detecting distress on road pavement cracks and potholes; however, there is a notable gap in research regarding crack detection in airfield pavement. This study addresses this gap by evaluating the performance of DL models in airfield pavement crack detection. RGB images are collected and processed to train object detection models, with the evaluation based on standard object detection metrics. This research aims to train and evaluate different variants of YOLO for crack-based distress detection in AC and PCC pavements.

DATA COLLECTION AND METHODS

Data Collection and Preparation

The first step to initiate the training and development of the DL model involves preparing the data. The study utilized data collected from Grosse Ile Municipal Airport, Grosse Ile Township, Michigan (ONZ) and Custer Airport, Monroe, Michigan (TTF) (Figure 1). The Apron and Runway 17/35 of ONZ airport consist of PCC pavement system. In May 2021, two DJI Mavic 2 Enterprise Advanced sUASs equipped with 48 mp Quad-Bayer RGB optical cameras and 640x512 pixel thermal sensors were flown at 15.1 m altitude to collect RGB optical data from Runway 17/35 and apron. The same sUAS was employed for AC pavement data collection from Runway 3/21 at an altitude of 15.1 m. Each collected image measured 8000 pixels x 6000 pixels, covering approximately 20 meters in length and 15 meters in width. For the DL model's training and validation, a block size representing approximately 5 m x 3.75 m was used, leading to the division of each collected image into 16 equal-sized blocks with Python programming.

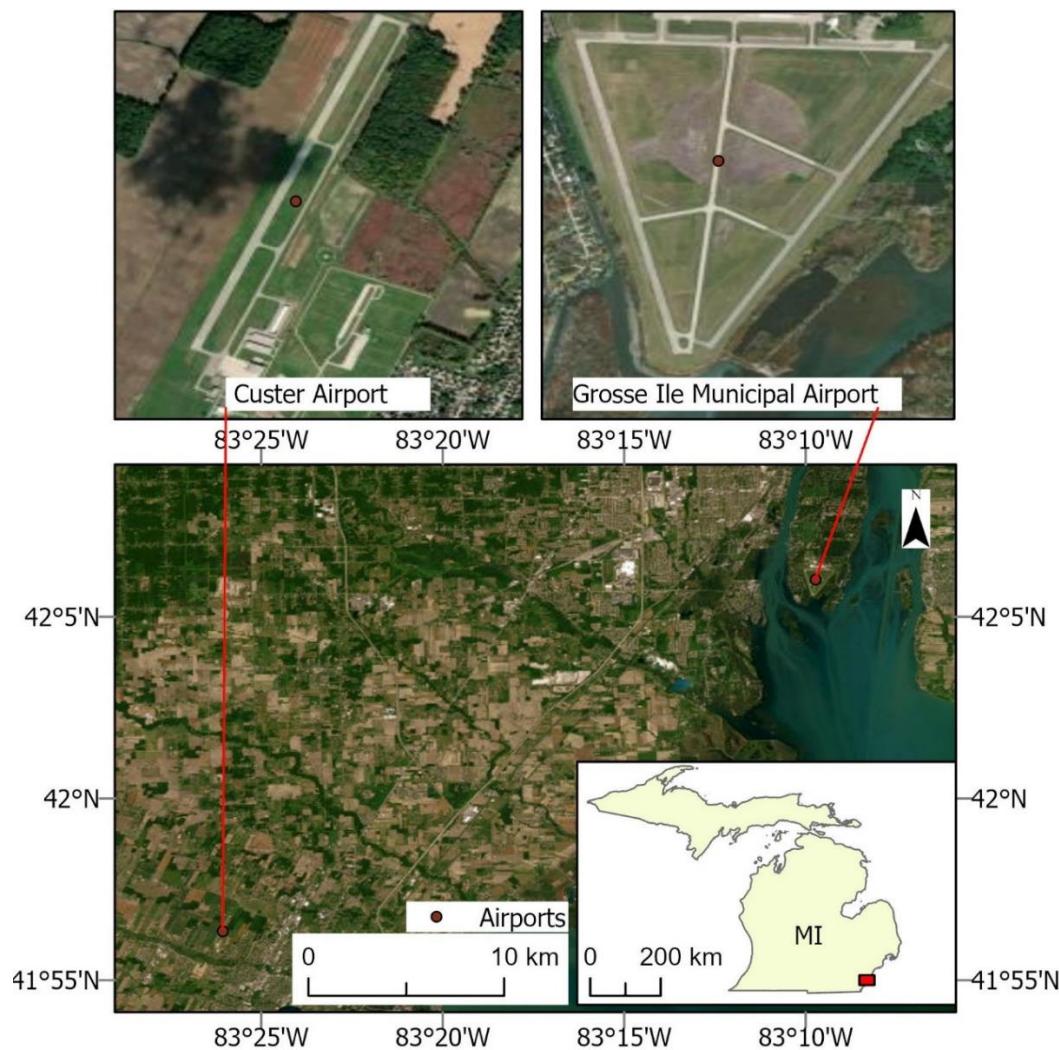


Figure 1: Data Collection Site

Training dataset creation

Creating the training dataset involves annotating the split images using LabelImg, an open-source image annotation software (Tzutalin, 2015). The primary objective of annotation was to draw bounding boxes around each crack in the image. The number and shape of cracks varied in each image, and a bounding box was drawn when a crack was completely visible. The keyword "D" was provided for the bounding box drawing, and LabelImg automatically recorded the user input in an XML file. The spatial location of the bounding box, along with its predefined single-class cracks (keyword "D"), was recorded in Pascal VOC format, noting the coordinates of the upper-left and bottom-right corners of the bounding box (Tzutalin, 2015). A total of 3,118 images are processed to create the dataset, which was then uploaded to Roboflow.com, an online platform facilitating image annotation and computer vision model building (Roboflow, 2023). Data augmentation, including rotation and cropping, was applied to randomly selected images, resulting in 5,273 images. These images were randomly split into three groups: the training set with 4,310 images, the testing set with 583 images, and the validation set with 572 images.

Training DL Object Detection Model

Training DL object detection models involves utilizing computationally expensive architectures. Models, such as Efficientnet, Faster R-CNN, YOLO, and Single Shot Detector, are known for providing highly accurate object detection and localization results after adequate training. The speed and accuracy of these models vary based on their architecture and working principles. With proper software and a high-performance Graphics Processing Unit (GPU), these DL models can detect multiple objects in a video in real-time (30 frames per second).

The dataset was smaller than widely used object detection training datasets like Microsoft COCO, PASCAL VOC, Open Images, KITTI, and ImageNet (Everingham et al., 2014; Geiger et al., 2013; Kuznetsova et al., 2020; Lin et al., 2015; Russakovsky et al., 2015). However, transfer learning, leveraging knowledge acquired for one task to solve a related problem (Kim et al., 2018; Shin et al., 2016), was employed in this research, where four variants of YOLO models (nano, small, medium, and large) were trained and evaluated, each having different architectures and yielding different data processing and accuracy times. Table 1 shows the properties of each YOLO variant. Larger models offer higher accuracy for object detection but contain more parameters, resulting image/video processing.

Table 1: Properties of YOLO models

Model	size (pixels)	Mean Average Precision ^{val} 50-95	Parameters (Millions)
YOLOv8-nano	640x640	37.3	3.2
YOLOv8-small	640x640	44.9	11.2
YOLOv8-medium	640x640	50.2	25.9
YOLOv8-large	640x640	52.9	43.7

The complete dataset with 5,273 images and annotation files was downloaded from the Roboflow hub. It had three separate folders for training, testing, and validation data, along with a YAML file containing information on folder structure and object class. The training was initiated

for 100 epochs and the trained model was saved locally after every five epochs. Epochs represent the passing of the complete training dataset through the model. A Dell 5820 workstation computer with a 12-core CPU, 128 GB DDR4 RAM, and RTX-5000 16GB VRAM was used for model training. The batch size for each model was different, with nano, small, medium, and large YOLO models having batch sizes of 64, 56, 32, and 22, respectively.

Once the DL model training starts, random weight and bias values are assigned or acquired from the pre-trained model. However, as the training proceeds, the weights and biases are adjusted over time. The errors in the predictions made by the model after the adjustments are known as losses. The loss indicates a penalty for bad predictions, showing how bad the model's predictions are. The lower the loss, the better the prediction. The training dataset was used to train the model and calculate the training loss. However, the validation dataset was only used for validation loss calculation. The losses decreased drastically at the beginning of the training of all models, but after a certain number of epochs, the decrease rate was slow (Figure 2). For all four models, the losses are less than 2 after 20 epochs, with very minor to no improvement in the model performance after 60 epochs. Figure 2 also shows that the losses were between 1 and 2 after 100 epochs of training, indicating the models were adequately trained for crack detection in AC and PCC pavement. The trained models were saved into hardware for further use.

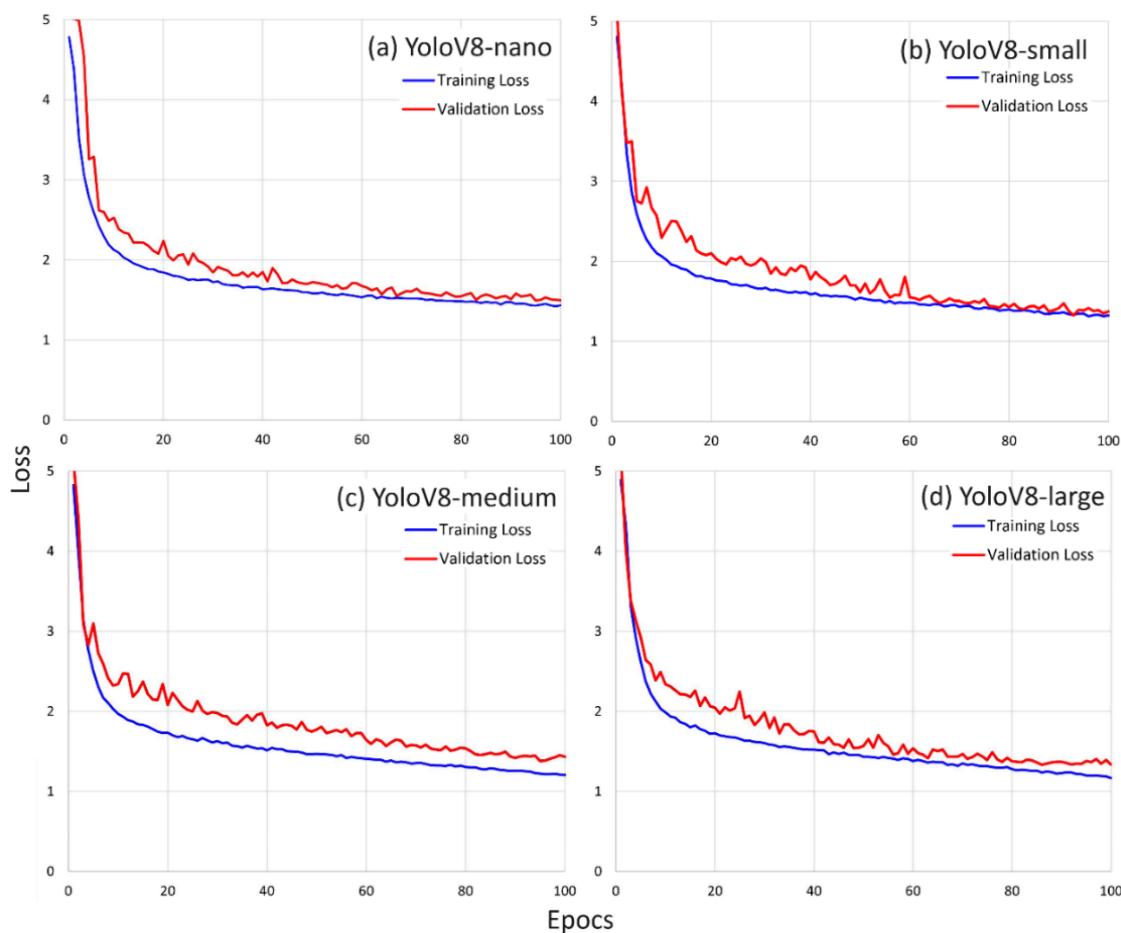


Figure 2: Training and validation loss for (a) YOLOv8-nano, (b) YOLOv8-small, (c) YOLOv8-medium, and (d) YOLOv8-large

RESULTS AND ANALYSIS

Since crack detection is a typical object tracking problem, the accuracy and efficiency of the models were evaluated using mean average precision. Additional standard evaluation metrics are described in Table 2. Here, \uparrow indicates that a higher number denotes better performance while \downarrow indicates that a lower number denotes better performance. The detection result provides the location of a bounding box located in the same place or a little further from the ground-truthing bounding box. The intersection over union (IOU) parameter was used to determine whether the predicted bounding box was a true positive (tp) or false positive (fp). For our study, the threshold value was selected to be 0.5. A detection was considered true positive if the ground truth bounding box and predicted bounding box had an intersecting area that was more than 50% of the union of those bounding boxes. Additionally, examples of some detections from YOLOv8-nano are provided below in Figure 3.

Table 2: Evaluation metrics to assess accuracies.

Metric	Description
True positive (TP) \uparrow	An object is detected and also present in the image
False positive (FP) \downarrow	An object is detected but is not present in the image
False negative (FN) \downarrow	An object is not detected but is present in the image
True negative (TN) \uparrow	An object is not detected and is not present in the image
Recall \uparrow (TP / TP + FN)	Percentage of correctly matched detections to ground-truth detections
Precision \uparrow (TP / TP + FP)	Percentage of correctly matched detections to total detections
mAP50 \uparrow	Mean average precision calculated at an IoU threshold of 0.50

YOLOv8 models automatically calculated precision, recall, and mAP50 after each epoch. The results for each model and their speed for processing individual images are provided in Table 3 below. Table 3 shows that with adequate training, the lightweight YOLO model can perform with a mean average precision of more than 0.65. Each model can detect cracks with high accuracy in AC and PCC pavement. The speed also demonstrates that nano and small variants are the fastest among all four models, whereas the large version of the YOLO model takes 22.7 milliseconds to process each image. However, each of the models has the capability of real-time crack detection (30 image processing per second).

Table 3: Model results and speeds

Models	Precision	Recall	mAP50	Individual Image Processing time (millisecond)
YOLOv8- nano	0.69055	0.6404	0.65082	7.9
YOLOv8-small	0.7408	0.6404	0.69054	8.1
YOLOv8-medium	0.69186	0.69853	0.66505	14.4
YOLOv8-large	0.66921	0.7193	0.67755	22.7



Figure 3: Sample detection result

FUTURE WORK

This research successfully demonstrated the use of a DL object detection model for crack detection in airfield pavement. The authors envision deploying multiple DL models with varying accuracy and speed, aiming for higher precision in distress identification. Moreover, The study recommends future work to extend the scope by incorporating additional distress detections, such as, corner spalling, joint spalling, joint seal damage, patching, and raveling. The ultimate goal is to achieve higher accuracy with lightweight object detection DL models that can be deployed using minicomputers, further optimizing the practicality of the proposed methodology.

SUMMARY AND CONCLUSIONS

The research employed two DJI Mavic 2 Enterprise Advanced sUASs equipped with RGB optical cameras to collect RGB data from two airports in Michigan. The collected images

underwent annotation and augmentation, forming a robust dataset for training and evaluating DL models. Specifically, the study investigates four variants of YOLO models, ranging from nano to large, for their performance in crack detection in both AC and PCC pavements. This research represents a significant advancement in airfield pavement inspection methodologies by integrating state-of-the-art technologies. Timely airfield distress identification is necessary for airfield pavement rehabilitation and maintenance. The current approach to airfield pavement inspection is time-consuming. Thus, this research attempted to use DL object detection models for crack detection. The successful application of DL models, particularly the YOLO variants, highlights their potential to transform the inspection landscape. The findings affirm that these models can effectively detect and categorize cracks in AC and PCC pavements, demonstrating their utility in enhancing inspection accuracy and efficiency. The results showcase promising outcomes, with the YOLO models exhibiting mean average precision values exceeding 0.65. The promising performance of the DL object detection model will be extended with training and deployment of additional object detection models and incorporating multiple pavement distresses.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Federal Aviation Administration for supporting this study. The authors would like to thank the project technical point of contact (TPOC), Mike DiPilato, for his invaluable guidance, support, and direction during this study. Special thanks are expressed to Grosse Ile Municipal Airport manager Michael Duker and Custer Airport manager Dan Diesing, for their full support during data collection. The authors would like to express their sincere gratitude to other team members from Iowa State University, Michigan Tech Research Institute, and Applied Pavement Technology, Inc., who assisted with data collection and organization. Although the FAA has sponsored this project, it neither endorses nor rejects the findings of this research. The presentation of this information is in the interest of evoking comments by the technical community with respect to the results and conclusions of the research. This paper does not constitute a standard, specification, or regulation.

AUTHOR CONTRIBUTIONS

The authors contribute to the paper as follows: Dr. Halil Ceylan, Dr. Sunghwan Kim, Dr. Md Abdullah All Sourav, and Matthew Brynick conceived and designed the study; Dr. Md Abdullah All Sourav processed data, trained models, and conducted evaluations; Dr. Md Abdullah All Sourav, Dr. Halil Ceylan, Dr. Sunghwan Kim, and Matthew Brynick analyzed and interpreted results; Dr. Md Abdullah All Sourav, Dr. Halil Ceylan, Dr. Sunghwan Kim, and Matthew Brynick prepared the draft manuscript. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Airport Improvement. (2020). *Savannah/Hilton Head Int'l Prepares to Integrate Drones Into Airport & Airfield Operations / Airport Improvement Magazine*.
- Airsight. (2020). Runway pavement inspections using airsight drone.
<https://www.airsight.de/projects/item/runway-pavement-inspections-using-airsight-drone/>.

- Banks, E., et al. (2018). Successful Approaches for the Use of Unmanned Aerial System by Surface Transportation Agencies. <https://domesticscan.org/download/4364/>.
- Congress, S. S. C., and Puppala, A. J. (2022). Lessons Learned in Airport Asset Inspection Using Unmanned Aerial Vehicle (UAV) Based Close-Range Photogrammetry. *International Conference on Transportation and Development 2022*, 212–222.
- Ellenberg, A., Kontsos, A., Moon, F., and Bartoli, I. (2016). Bridge related damage quantification using unmanned aerial vehicle imagery. *Structural Control and Health Monitoring*, 23(9), 1168–1179.
- Everingham, M., Eslami, S. M. A., Van Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2014). The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision*.
- FAA. (2022). Report to Congress on the Airport Improvement Program for FY 2019. https://www.faa.gov/sites/faa.gov/files/2022-08/Report_Congress_AIP_accomplishments_FY2019.pdf.
- FHWA. (2018). Use of Unmanned Aerial Systems (UAS) by State DOTs. <https://rosap.ntl.bts.gov/view/dot/43679>.
- Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*.
- Hubbard, B., and Hubbard, S. (2020). Unmanned Aircraft Systems (UAS) for Bridge Inspection Safety. *Drones*, 4(3), 40.
- Kim, H., Kim, H., Hong, Y. W., and Byun, H. (2018). Detecting Construction Equipment Using a Region-Based Fully Convolutional Network and Transfer Learning. *Journal of Computing in Civil Engineering*.
- Kuznetsova, A., et al. (2020). The Open Images Dataset V4: Unified Image Classification, Object Detection, and Visual Relationship Detection at Scale. *International Journal of Computer Vision*.
- Li, R., Yu, J., Li, F., Yang, R., Wang, Y., and Peng, Z. (2023). Automatic bridge crack detection using Unmanned aerial vehicle and Faster R-CNN. *Construction and Building Materials*, 362, 129659.
- Li, Y., Ma, J., Zhao, Z., and Shi, G. (2022). A Novel Approach for UAV Image Crack Detection. *Sensors*, 22(9), 3305.
- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollár, P. (2015). Microsoft COCO: Common Objects in Context. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Pietersen, R. A., Beauregard, M. S., and Einstein, H. H. (2022). Automated method for airfield pavement condition index evaluations. *Automation in Construction*, 141, 104408.
- roboflow. (2023). <https://roboflow.com/>.
- Russakovsky, O., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*.
- Seo, J., Duque, L., and Wacker, J. (2018). Drone-enabled bridge inspection methodology and application. *Automation in Construction*, 94, 112–126.
- Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., and Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*.

- Sourav, M. A. A., Ceylan, H., Brooks, C., Peshkin, D., Kim, S., Dobson, R., Cook, C., Mahedi, M., and Brouillette, O. (2022). *Small Unmanned Aircraft System for Pavement Inspection: Task 4—Execute the Field Demonstration Plan and Analyze the Collected Data*. United States. Department of Transportation. Federal Aviation Administration.
- Sourav, M. A. A., Ceylan, H., Brooks, C., Peshkin, D., Kim, S., Dobson, R., Cook, C., Mahedi, M., and Jenkins, A. (2023). *Small Unmanned Aircraft System for Pavement Inspection*. United States. Department of Transportation. Federal Aviation Administration. Airport Engineering Division.
- Sourav, M. A. A., Ceylan, H., Kim, S., Brooks, C., Peshkin, D., Dobson, R., and Brynick, M. (2023). Use of Digital Elevation Model for Detecting Airfield Pavement Distress. In *Airfield and Highway Pavements 2023* (pp. 254–265).
- Sourav, M. A. A., Ceylan, H., Kim, S., Brooks, C., Peshkin, D., Dobson, R., Brynick, M., and DiPilato, M. (2022). Small Uncrewed Aircraft Systems-Based Orthophoto and Digital Elevation Model Creation and Accuracy Evaluation for Airfield Portland Cement Concrete Pavement Distress Detection and Rating. *International Conference on Transportation and Development 2022*, 168–180.
- Tzutalin. (2015). LabelImg. <https://github.com/tzutalin/labelImg>.
- Wells, J., and Lovelace, B. (2018). *Improving the quality of bridge inspections using unmanned aircraft systems (UAS)*.
- YOLOV8. (2023). <https://github.com/ultralytics/ultralytics>.
- Zhu, J., Zhong, J., Ma, T., Huang, X., Zhang, W., and Zhou, Y. (2022). Pavement distress detection using convolutional neural networks with images captured via UAV. *Automation in Construction*, 133, 103991.