Results in Engineering 23 (2024) 102737

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks

Yetay Berhanu [a,b,*], Dietrich Schröder [b], Bikila Teklu Wodajo [c], Esayas Alemayehu [a]

[a] Faculty of Civil and Environmental Engineering, Jimma Institute of Technology, JIT, Jimma University, P.O. Box 378, Jimma, Ethiopia
[b] Stuttgart University of Applied Sciences, HFT, D-70174, Schellingstr. 24., Stuttgart, Germany
[c] School of Civil and Environmental Engineering, Addis Ababa Institute of Technology, AAiT, Addis Ababa University, P.O. Box 385, Addis Ababa, Ethiopia

ARTICLE INFO

Keywords:
Network analysis
Accident prediction
Safe route
Road traffic accident and traffic congestion
Machine learning- Random Forest

ABSTRACT

Road traffic accidents (RTAs) and the resulting traffic congestion are global concerns mainly in metropolitan environments. The need for road safety is directly correlated with the rapidly increasing impact of urbanization on infrastructure and day-to-day living. In this study, we have introduced an innovative approach integrating a Random Forest (RF) model, crash rates, and spatial network analysis to provide safe route recommendations for drivers aiming to reduce RTAs and congestion. Based on historical accident data from 2014–2019, the analysis of the RF model and crash rate served as a prediction of the likelihood and occurrence of RTAs. In applying the spatial network analysis, lower predicted crash counts from spatial joining were taken into consideration, as well as areas with lower crash rates that have had fewer incidents in the past. An alternative safe route and an optimum route that covers 32.27 km in 50.78 mins of travel time and 28.6 km in 41.58 mins of travel time were successfully identified, respectively. Having demonstrated 78 % predictive capability on the target variable, the RF model has proved its worth. Analyzing historically lower accident counts on segments leading to minimal crash rates validates the accuracy of the identified safe route. This advanced method significantly aids in improving traffic safety by making drivers and travelers aware of potentially high rates of accidents and traffic congestion on road segments. It assists travelers with their trip planning to anticipate potential risks and suggest safer alternate routes, making it a valuable contribution to the field.

1. Introduction

In the modern-day landscape, traffic accidents and road congestion emerge as pressing global challenges in urban planning and transportation management. In large cities and metropolitan areas, routing extensive distances for both regular and occasional trips raises crucial concerns about safety and route complexity, prompting safety management efforts to reduce the risk of accidents and injuries [1]. The rising demand for road transportation in recent decades has brought about numerous challenges, one of the most critical being traffic congestion. This congestion greatly increases the likelihood of accidents, often leading to severe injuries and fatalities [2]. Global studies emphasize the critical nature of traffic congestion, creating significant challenges in city traffic networks [3,4]. Moreover, traffic congestion is a major issue in many cities, particularly in large urban areas, causing people to waste significant amounts of time [5]. These challenges necessitate innovative solutions to improve safe urban mobility and alleviate pressure on transportation systems, affecting daily lives and placing significant stress on worldwide infrastructure systems [6].

Traffic safety has increasingly become a global concern particularly in low- and middle-income countries (LMIC), where nearly 90 % of traffic casualties occur. In contrast, high-income countries (HIC) report casualty rates below 9 %, while in LMICs average around 20 %, with the African region registering the highest rate at 26.6 % [7]. In 2021 alone, an estimated 1.19 million fatalities resulted from road traffic accidents, with 92 % of these incidents occurring in LMICs as per the World Health Organization's [8] report. The challenge facing transportation networks is to promote energy-efficient, diverse transportation modes while ensuring sustainability, reducing traffic congestion, and enhancing road safety measures.

* Corresponding author. Faculty of Civil and Environmental Engineering, Jimma Institute of Technology, JIT, Jimma University, P.O. Box 378, Jimma, Ethiopia.
E-mail address: yetadu26@gmail.com (Y. Berhanu).

https://doi.org/10.1016/j.rineng.2024.102737
Received 25 May 2024; Received in revised form 14 July 2024; Accepted 12 August 2024
Available online 16 August 2024

Safety experts in LMICs encounter difficulties in developing effective tools and strategies due to limited organizational resources [9] particularly in low-income countries like those across Africa, where road traffic accidents and resulting injuries present substantial health, economic, and developmental challenges [10]. According to United Nations Economic Commission for Africa and Economic Commission for Europe [11] similar to many African nations, Ethiopia failed to achieve the United Nations Decade of Action for Road Safety goal to halve road crash fatalities by 2020, as fatalities more than doubled from 2,161 in 2007 to 4,597 in 2018. The annual statistics reveal 5,180 deaths and over 20,000 disabilities, with Addis Ababa experiencing daily fatalities due to road traffic accidents [12].

Urbanization has intensified stress on the city populations, and exacerbated traffic congestion, highlighting the critical nature of road safety for sustainable urban development and the public welfare. Plenty of cities are struggling with these traffic mobility challenges, along with a sort of intensified traffic congestion owing to the vast effects of urbanization on easy and safe ways of travel [13]. The rapid traffic growth and unsafe road user behaviors on Addis Ababa's underdeveloped roads, combined by poorly planned road network and inefficient traffic operations contribute to frequent accidents [14]. The persistence of road traffic accidents and stagnant capacity has reduced traffic flow, affecting accessibility to economic centers, which impacts on road users' daily lives in Addis Ababa.

The connection between rapid urban development and increased traffic accidents, emphasizing the wider consequences beyond fatalities and financial damages. These accidents disrupt traffic flow, worsen congestion, and pose significant threats to overall traffic safety [15,16]. Gonzalez et al. [17] proven a strong correlation coefficient of 0.81 between congestion and accidents, clearly indicating congestion's significant impact on accident rates. Traffic congestion tends to gravely influences road safety, increasing crash risks due to erratic driving behaviors and speed variations within congested lanes. The capacity reduction brought on by the closure of lanes during accidents causes the majority of congestion, underscoring the critical need for a profound understanding of road crashes and their underlying causes [18]. Insufficient spatial road safety analysis strategies and poor traffic management factors have contributed to an increase in accidents and congestion, imposing economic losses and human casualties [19]. Hence, safety evaluation includes assessing the vehicle's "crash risk potential" during travel, particularly in areas with limited crash data, to determine the safest and most efficient routes [20].

Considering the uncertain demand and supply attributes, in conjunction with limited resources, can lead to congestion [16]. Severe primary traffic accidents lead to upstream traffic congestion and significant consequences, leading to more rise in the probability of subsequent occurrences [21]. And yet, Congestion does not only elevate the likelihood of road traffic accidents (RTAs) but it can also be a result due to several factors beyond accidents, such as escalating traffic volumes [10]. According to a correlation by the Institute of Transportation Engineers [22], the reduction of speed has been identified as a necessary measure during times of congestion as an essential way to mitigating the severity of crashes. This provides a robust rationale for investing in proper road safety interventions, especially in areas where allocation of constraint [23,24].

Higher traffic volumes and congested road conditions intensify interactions between vehicles, which makes accidents more likely to occur [25]. Addressing congestion in Addis Ababa is essential, as it affects residents citywide, causing extended travel times, work hours lost, and adverse economic impacts [26]. This persistent situation not only affects economic activities but is also increasing adverse environmental effects, including amplified noise pollution [27].

Academic literature examines the sophisticated relationship between congestion and accident rates, with González et al. [17] suggesting that less congestion could lead to a reduction in traffic accidents. The combined challenges of traffic accidents and congestion impact urban environments significantly, affecting the well-being of road users and urban populations [28]. Also, previous studies have extensively examined methods to identify locations with unusual crash rates, yet they often don't really consider how safety differences compared to the average accident frequency in the reference population create challenges when there are insufficient similar risky parts on the roads [29]. Addressing this issue requires the development of crash prediction models, as highlighted by World Road Association [30], to accurately estimate the number of crashes based on independent variables. The primary objectives of prediction studies have typically been to achieve accurate forecasts, such as the number of traffic accidents [31]. Accurate prediction of traffic conditions, as emphasized by Kai and Mo [27], enables the timely implementation of effective countermeasures and empowers drivers and pedestrians to avoid proactively accident-prone areas, as emphasized by Fan et al. [32] and Berhanu et al. [10].

Enhancing real-time safety requires a comprehensive understanding of the locations and temporal patterns associated with traffic accidents, emphasizing the need to identify high-risk segments along road networks [29]. Spatial analyses play a crucial role in detecting hotspots and identifying problematic regions on a larger scale, facilitating the prediction of high-risk areas and alerting drivers [33,34]. While conventional methods focus on altering physical road infrastructure, modern-day approaches rooted in Intelligent Transportation Systems (ITS) have lots of potential for reducing the risk of accidents [35].

Understanding the significance of traffic congestion in identifying optimal routes is crucial due to its substantial psychological and economic impacts on road users navigating complex road networks. These impacts cannot be emphasized enough. A widely accepted idea of optimality in stochastic routing pertains to reliability, it truly acknowledges that that the fastest route or policy, might involve big risks, making it not attractive for a risk-averse traveler [36]. Addressing this challenge involves providing route guidance to individual users to avoid congested or hazardous areas [37]. Such tools benefit not only individuals with diminished vision or perception abilities, including older drivers, but also help vulnerable road users such as bicyclists steer clear of congested areas and reduce the risk of collisions or accidents [20].

Interms of route planning strategies, researchers have introduced various methods to determine cost-effective routes within a network. For instance, Goczyła and Cielatkowski [38] proposed a method for identifying optimal paths based on minimizing travel time. Another approach, as presented by Geisberger et al. [39], involves the identification of key nodes within a network to plan routes effectively. The main consideration for route selection is to determine if the safest and quickest routes align. However, meeting this criterion may lead to vehicular traffic passing through residential areas [40].

Traditional statistical models may lack suitability for comprehensive spatial analysis, spatio-temporal analysis, and prediction highlighting the need for adaptive multivariate models strong generalization abilities and robust data utilization. Machine learning (ML) models excel in capturing complex associations between input and output variables, making them ideal for such tasks [33,41]. ML has emerged as a powerful tool in crash modeling in the past 2 decades, and offering superior in predicting performance over traditional methods (Y. [42]) and ensemble methods are also utilized due to their effectiveness in handling imbalanced data efficiently [43]. Many reviews cover ML in civil engineering, but few focus on physics-based ML or provide a roadmap for its application. Additionally, fundamental physics-based ML models in this field are rarely explored too [44]. Berhanu et al. [10] focus on the direct association between road network congestion and traffic accidents. They suggest that implementing traffic management systems with real-time data analysis and spatial predictive methodologies can decrease the impact of both factors through accident prevention and congestion mitigation. Functional models provides the benefit of easier understanding, spatial analysis studies have displayed weak transferability in both these scenarios [33].

The lack of a direct link between accident variables and injury severity makes developing road safety models a tough task [45]. Hence, crash prediction, which is crucial for road safety, heavily relies on modeling techniques to identify factors leading to increased crash risks and guide safety measures [46]. Similarly, studying behavioral features in RTAs is challenging due to the accidents' unpredictability and spontaneous nature, making direct observation difficult. Achieving 100 % accurate data is nearly impossible in this context, underscoring the need for advanced methods to enhance analysis. ML algorithms provide a practical and promising solution to address these challenges and achieve better results [33,47–50].

In Atumo et al.'s [51] study, ML appeared as a powerful tool for detecting trends and patterns in crash data, offering advantages over traditional methods. Supervised ML techniques, particularly Random Forest (RF), show high classification performance based on attributes and datasets provided [34] and higher accuracy compared to statistical models [52,53]. RF, known for its resistance to overfitting and enhanced accuracy in both classification and regression tasks, stands out in identifying variables strongly related to the target variable ([54]; Y. [42]). The study by Al-Mistarehi et al. [55] validates the superior performance of RF in predicting injuries, while Bokaba et al. [56] emphasize its excellence in accuracy metrics with the lowest root-mean-square error (RMSE), indicating a better fit compared to other classifiers. Pourroostaei Ardakani et al. [57] and Y. Ali et al. [42] back up RF's effectiveness, especially in complex scenarios, emphasizing its higher accuracy in predicting RTAs.

Routing is a key player in making transportation networks efficient and also, ensuring movement, especially when multiple issues like fires, plumes, and floods are sprouting up on the roadways during major disasters [58]. These roadblocks can trigger road congestion and restrict rescue vehicle mobility, emphasizing the importance of selecting the safest alternative among potential routes in urban transportation planning [1]. Determining optimal routes involves considering factors such as traffic conditions, road closures, congestion, and safety, as underlined by Hoseinzadeh et al. [59]. Notably, Sohrabi and Dominique Lord [60] emphasize that route safety recommendations from navigation apps may differ from the shortest route, underscoring the need for comprehensive safety considerations in route planning.

Previous researches in the field have explored accident prediction models and route optimization using various datasets, often with a focus on identifying accident hotspots and optimizing travel routes. However, these studies have typically operated focusing narrowly on specific aspects that provide limited insights. We identified a gap in previous research where they independently tackled either accident likelihood prediction or route optimization without integrating comprehensive spatial analysis and machine learning methodologies to address both aspects simultaneously.

However, to the best of our knowledge, no similar research has been done combining ML RF methods and spatial crash rate calculation to identify safe routes using spatial network analysis particularly in the study location of Addis Ababa. This segmented approach has left gaps in understanding the interconnected dynamics of traffic crashes and congestion, particularly in urban settings like Addis Ababa. Hence, this study aims to fill these gaps by integrating spatial analysis with RF machine learning models.

By combining these approaches, the research aims to enhance the accuracy of accident prediction and the identification of safe routes in congested urban environments. The novelty of this approach lies in its holistic integration of spatial analysis techniques with RF machine learning methods, which are adept at handling complex, multidimensional data sets. This integration leads to a richer analysis interpretation and understanding of crash likelihood. It provides actionable insights into safe route recommendations, thereby offering a comprehensive framework to effectively mitigate traffic crashes and congestion. The significance of this study extends beyond traditional statistical models by offering a refined methodology to enhance road safety and reduce traffic congestion.

This study aims to introduce a new approach leveraging spatial ML techniques, particularly the RF model, for accident prediction, crash rate calculation, and spatial network analysis. The goal is to provide safe routing solutions in accident- and congestion-prone road networks in Addis Ababa city. Ultimately, the study contributes to prioritizing road safety and traffic flow efficiency in urban transportation systems in the city. It aiming to reduce delays, improve overall travel experiences, and minimize the risk of accidents, rather than solely focusing on minimizing anticipated trip time. The paper is structured the subsequent sections as follows: Section 2 offers insights into the methodology for ML accident prediction, crash rate calculation, and spatial safe route identification. Following that, Section 3 presents the analysis and results of the applied methods. In Section 4, we explain our discussions tailored for safe routing to address the findings, considering crash risk and congestion problems. Finally, Section 5 functions as the conclusion of this study.

2. Methodology

Several datasets were used in this study; they were cleaned and spatial and data analysis techniques were applied to analyze crash rates, predict likelihood of road accidents, optimize streets, and find the safest and fastest route between a source and destination in a road network.

2.1. Data preprocessing

The preprocessing of data is essential in preparing data for spatial analysis and machine learning algorithms. According to many researchers, crash data is the most objective and trustworthy safety data commonly used for road safety analysis, as it is the most reliable indicator of road safety. This study makes use of a historical dataset of traffic accidents that took place in Addis Ababa between 2014 and 2019. The dataset totals 64,878 accident records obtained from the Addis Ababa Traffic office and the Addis Police Ababa Traffic Management Agency underwent thorough cleaning, transformation, and preprocessing. The accident records originally documented by the police, offers details, including the precise date and location of each incident. Geocoding is employed to address instances of unknown locations, while data with missing information is excluded from the analysis. Additionally, road network data and shapefiles outlining the city's boundaries were obtained from the Addis Ababa City administration and the Addis Ababa Traffic Management Agency (AATMA).

OpenStreetMap (OSM), an open-data initiative, offers a volunteer-driven repository of geographic information, including road networks. OSM is particularly valuable in low-income countries where official geographic data is lacking [61]. OSM has the capacity to generate a vast source of global geographic data, including transportation systems such as road networks [62]. The road networks of Addis Ababa City obtained from OSM were validated for their spatial correctness using Addis Ababa city road networks and boundary shapefiles collected from Addis Ababa City Roads Authority (AACRA) and Addis Ababa City Administration (AACA). To enable the utilization of OSM in spatial analysis, the QGIS tool was used to generate a compatible road segment shapefile data formats within the system. This preprocessing and data organization simplify the training of predictive models, conduct spatial crash rate analysis, and perform network analysis which finally leads to safe routing identification.

2.2. Forest based classification and regression

In this study, forest-based classification and regression methods, commonly known as Random Forest (RF) are used. RF, known for retaining all features, exhibits two key strengths. RF employs a bagging learning mechanism, creating a large ensemble of uncorrelated trees, effectively addressing overfitting. Moreover, each tree within the RF ensemble is constructed using randomly selected subsets of features and sample data, guaranteeing that only a subset of features plays a role in decision-making [63]. The use of Random Forest (RF) in ArcGIS Pro for accident prediction offers a broad perspective on potential accident-prone areas. We collected the road features and applied them for road accident prediction including road characteristics attributes like type, width, number of lanes, speed limits, and surface conditions, along with historical accident data providing crucial insights into past incidents. Additionally, environmental factors such as weather conditions and lighting conditions are considered. Through the calculation of crash rates and the creation of hotspot Probability Maps, it identifies regions with a higher likelihood of accidents, designating them as potential hotspots. By analyzing diverse predictors such as road features and historical accident data, RF aids in the identification of high-risk areas. RF analysis employed metrics like root mean square error (RMSE) and the coefficient of determination (R^2) when assessing model performance [64]. Thematic maps visually depict risk levels, and validation measures ensure the reliability of RF predictions, making this method crucial for proactive spatial analysis in understanding and visualizing concerns on road networks.

2.3. Calculate crash rate

Segment rating and prioritizing can be made possible by mapping the locations of safety priorities according to the degree to which the segment crash rate exceeds its critical rate [65]. The ArcGIS Crash Analysis Tool, which determines the crash rate, makes it comparatively simple to create a high-crash network [66]. This tool has the benefit of requiring the locations of intersections and streets to determine the frequency of crashes [67].

Accordingly, we calculated the crash rate by considering the unit of measurement of the crash rate computation. These done by multiplying

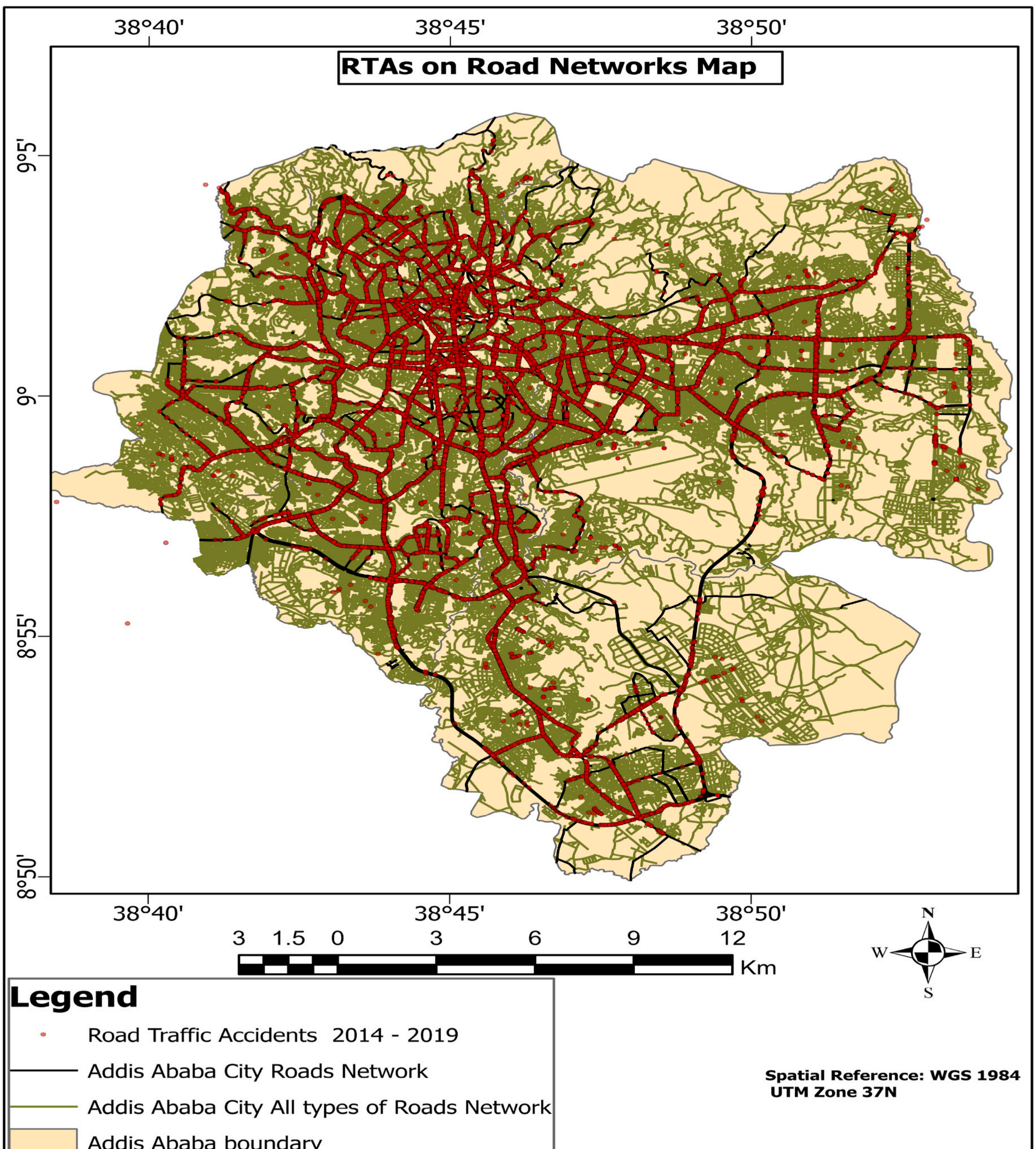

Fig. 1. Addis Ababa road networks with RTAs locations from the year 2014–2019.

the length of the road segment by the number of crashes related to each road segment over the study period. A shapefile comprising all crash locations between the years 2014–2019, a shapefile showing the city's roads and intersections from OSM, the time range (in years) during which the study is conducted, and the snap-to-street distance (in metres) are the inputs depicted in Fig. 1. The outcome creates for Addis Ababa a shapefile with crash rates color-coded to indicate how frequently crashes have occurred on each road.

2.4. Network analysis

Utilizing a spatial network analysis tool, a methodology is implemented to derive a safe route, strategically bypassing regions with heightened risks of road traffic accidents and congestion. This systematic process encompasses various stages. It commences with data preparation and the integration of relevant layers into ArcGIS Pro. Subsequently, a network dataset is constructed, facilitating the establishment of starting and destination points for route determination. Through route analysis, considerations for crash risk and associated restrictions are factored in, ensuring the identification of optimal paths while prioritizing safety. This approach not only equips road users, particularly drivers, with alternative routes that are both efficient and secure, but also it aids in directing enforcement efforts towards areas exhibiting higher crash risks and traffic congestion.

2.4.1. Creation of a network dataset

The methodology begins with the creation of a network dataset using the Network Dataset tool. This is coupled with the preparation of the road network dataset, which includes accident-related information. The road network database contains details regarding the road network, including nodes or vertices along the roads. Further steps involve configuring the analysis parameters, including the start and end points. Additionally, these steps specify safety constraints, such as avoiding of accident-prone segments and crash rates with higher concentrations of road traffic accident records.

2.5. Spatial safe routing

In transportation planning, trip generation involves the assignment of origins and destinations, is typically regarded as the initial step [68]. Although road navigation systems aim to find the shortest route between origin and destination locations, but there are circumstances where opting the fastest route raises the possibility of getting into an accident [69].

The optimal route precisely defines the safest and most efficient path between two points within the road network, ensuring a smooth journey without unnecessary delays [62]. Therefore, an inclusive procedure involves using a GIS-based network analysis tool for safe routing.

Initial steps include data preparation, ensuring the inclusion of accident data and pertinent attributes within the road network dataset. Routing analysis focuses on finding safe routes by assigning constraints

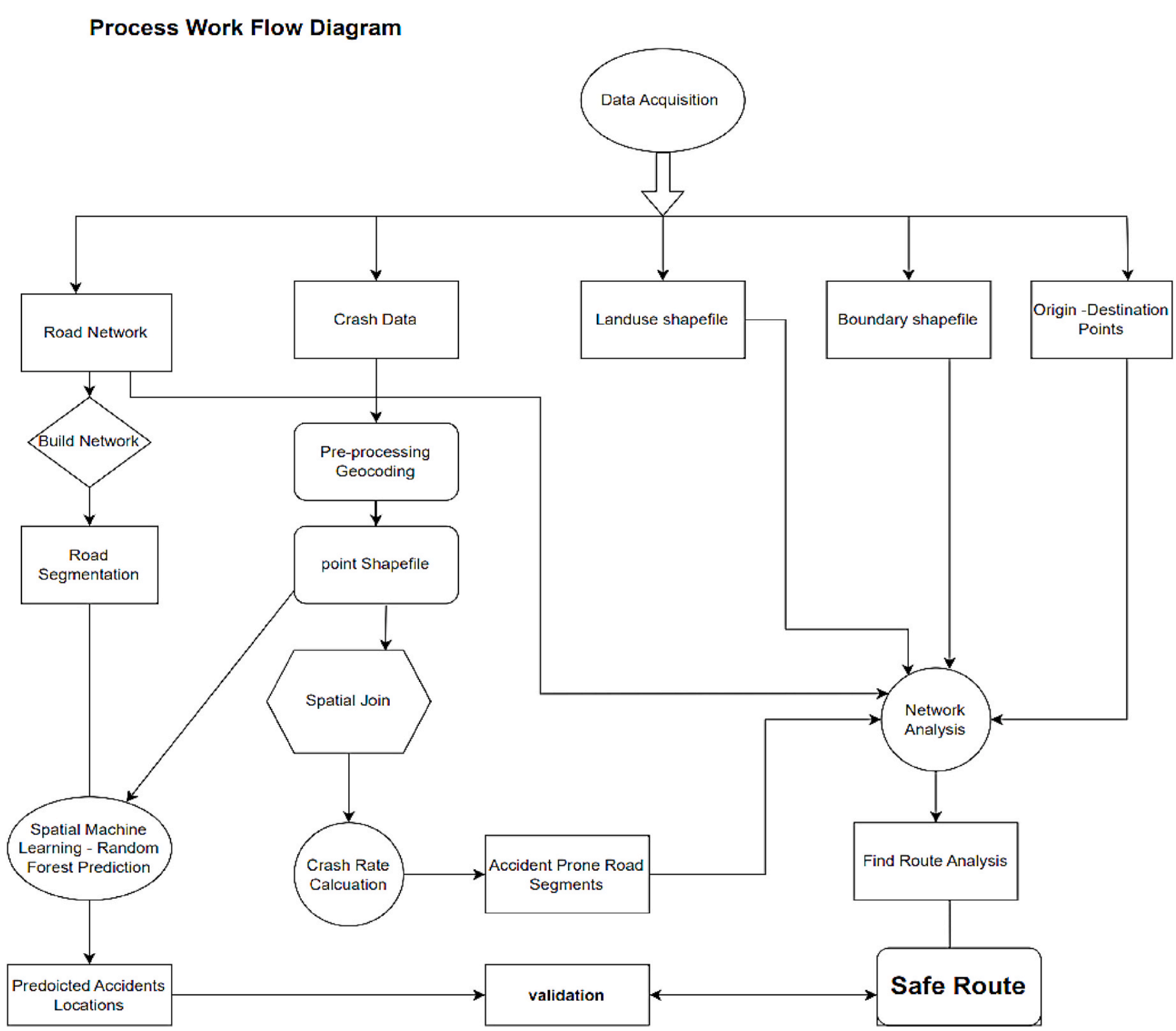

Fig. 2. Work flow diagram.

to accident-prone road segments, and weighted routing guides users away from these areas. The process establishes a route analysis layer, which facilitates finding the optimal route between specified starting and destination places of concern.

The analysis offers safe routing alternatives, ensuring relevance through the regular data updates. The primary objective is to find and provide a clear and safe routing option for road users and drivers, steering them away from high-risk road traffic crash areas and resulting congestion. To achieve this, we generate alternative routes and compare them for the selection of the safest route based on historical and

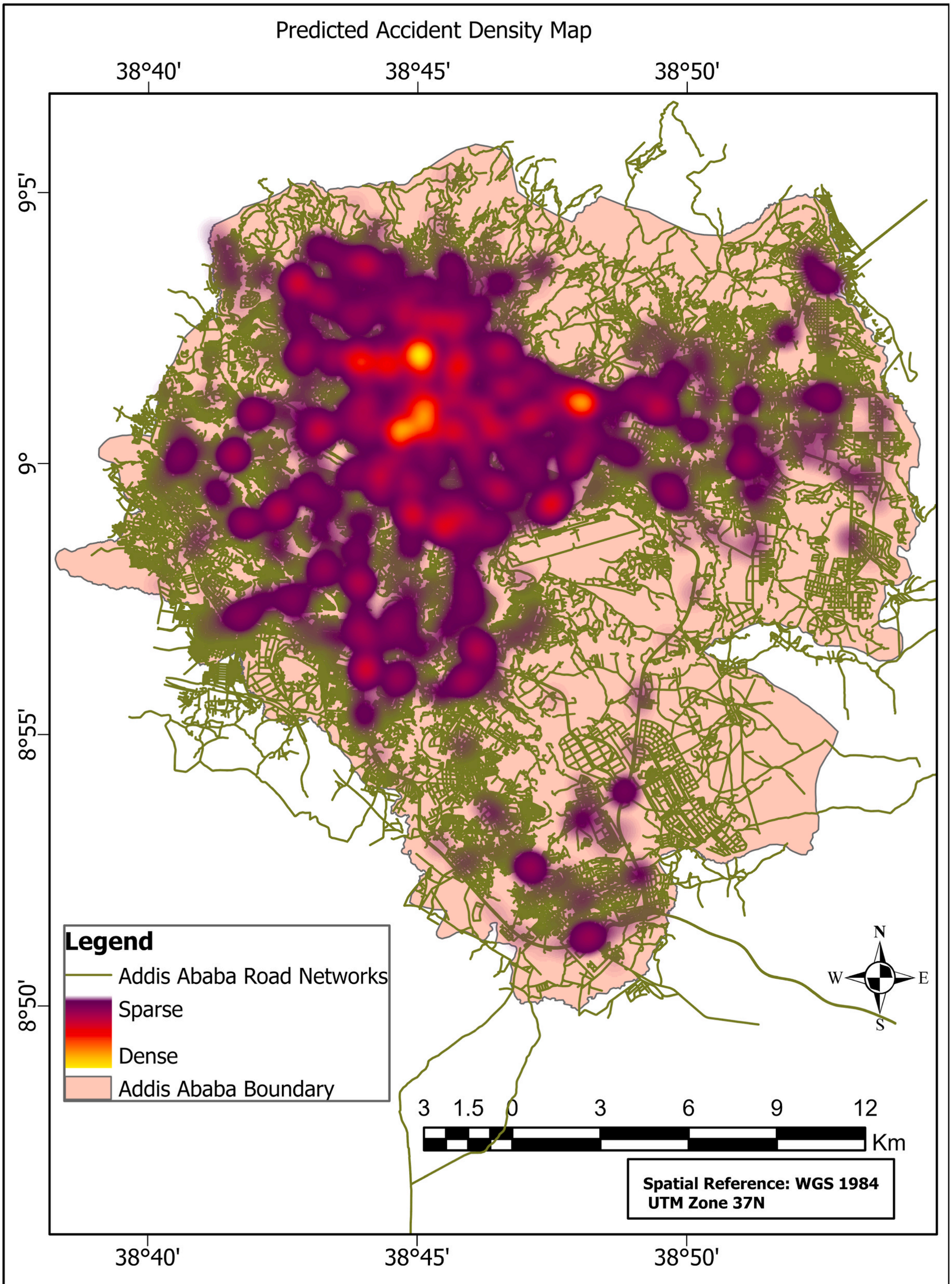

Fig. 3. Predicted accidents density map using RF on Addis Ababa's road networks.

predicted crash record data of each route.

Fig. 2 illustrates the general process of identifying accident-prone road segments and determining safe routing. The safest route is chosen based on spatial join of accident record data and crash rate calculation validating with the data from RF predicted likelihood of RTAs. These data serve as barriers for segments with high concentrations of accidents. After separately generating the alternative routes for comparison, the safest option is identified.

3. Analyses and results

While the objective of the study is twofold: first, to integrate spatial

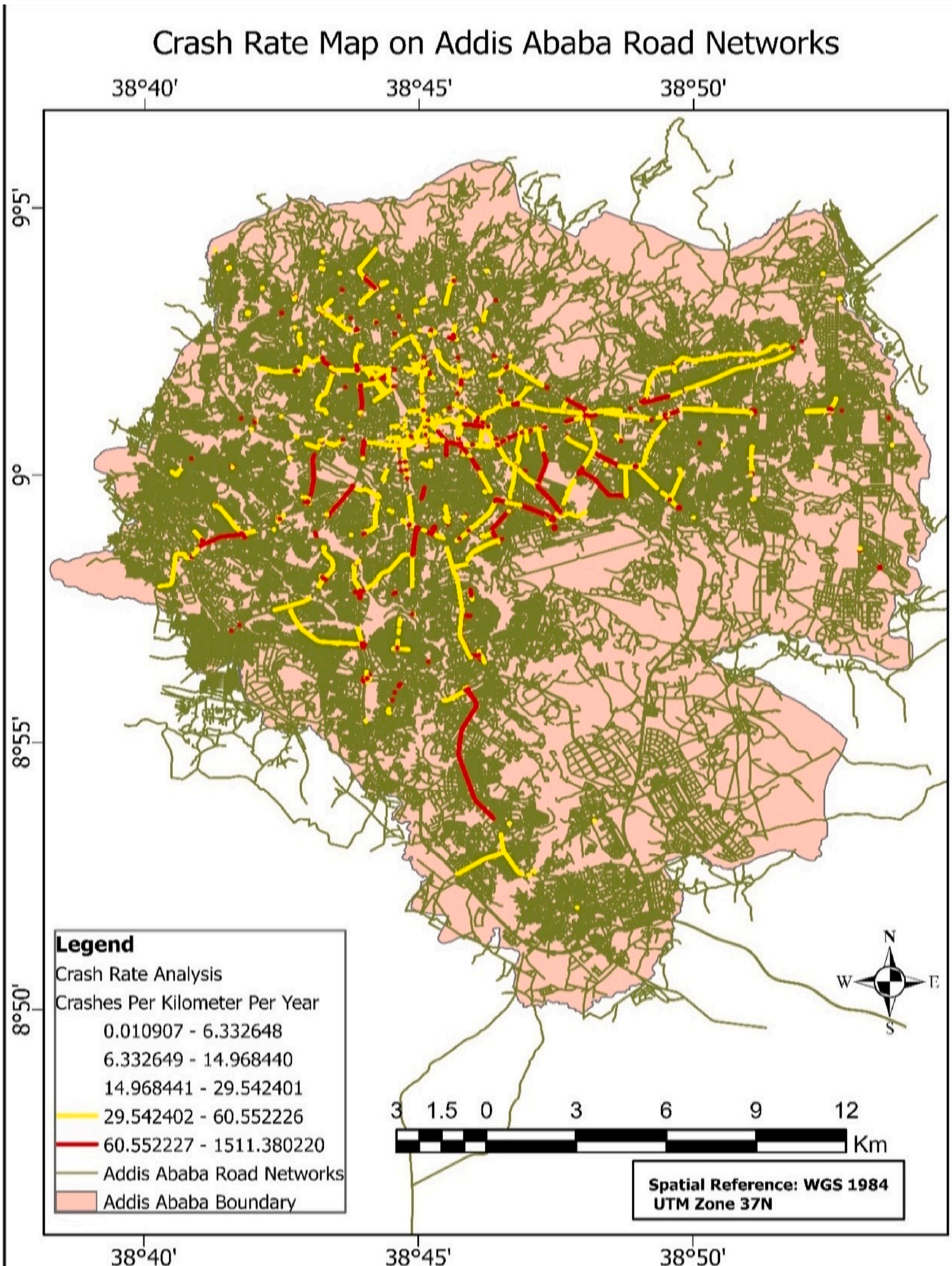

Fig. 4. Crashes rate map on Addis Ababa's road networks.

analysis with RF machine learning models to enhance accident prediction accuracy and identify safe routes in congested urban environments. This approach combines spatial analysis techniques with RF machine learning methods, which are adept at handling complex, multidimensional datasets. This integration enriches crash likelihood analysis. Second, the study aims to make an initial attempt to optimize traffic routes for safe navigation on road networks. It considers both historical accident data and crash rates of occurrences. Predicting the likelihood of road accidents using the Random Forest spatial machine learning method yields valuable insights for road safety, including crash rate maps, spatial patterns, and trends in predicted accidents. Based on the analysis and results described in the following sections, the identification of the safest route is determined after comparing alternatives, contributing to improved road safety.

3.1. Forest based classification and regression accident prediction

The application of highly sophisticated ML strategies is still in a minor stage of their execution (Y. [42]). As reviewed by Olugbade et al [49], the substantial involvement of AI and machine learning is crystal clear in the accurate prediction of accidents that occur on road systems. The analysis of the supervised spatial machine learning RF models is employed and was done using different spatial and non-spatial predictors to take into account the contributing risk factors that lead to accidents. The predictors used in the RF analysis include road features, historical data on accidents, weather conditions, and even road character, day of the week, time, and land use. The characteristics of the drivers (males and females), the type of roads, the surface of the roads, the type of junctions, and the lighting condition are all used. The output of the RF analysis plays a huge part in identifying specific locations or segments within the road network which the likelihood of an accident is highest.

In this study, the analysis resulted in a total of 3,096 accidents that were predicted by the RF model on the road network of Addis Ababa City as clearly depicted in Fig. 3 of the density map. The predicted likelihood of accidents falls below the observed distribution of accidents and patterns. Performance indicators, such as Mean Squared Error (MSE), calculate the average of the squared differences between predicted and actual values, therefore emphasize larger errors in regression tasks. Root Mean Squared Error (RMSE) works hand in hand with MSE and takes the square root of the MSE, offering a measure of the average magnitude of errors in regression tasks [70]. The study produced an RMSE of 78 % as well as performance metrics that are very beneficial in evaluating the RF model's accuracy, precision, and robustness. These metrics provide some extremely important insights into the model's effectiveness.

3.2. Crash rate

When assessing the safety of a road network, the crucial aspect of the safety review involves analyzing collisions occurring at junctions and along roads to find areas where a high incidence of fatalities and serious injuries occur [71]. The resulting map, depicted in Fig. 4, highlights the vulnerable roads in Addis Ababa a shapefile of crash rates color-coded based on concentration of crashes on each road. The road segments with red and yellow color, show high concentration of accidents.

The analysis's reveals which streets have the highest crash frequencies. As a result the road segments marked by the red and yellow symbols, respectively, signify the top percentage of collisions per kilometer annually.

Table 1
Generated alternative routes data.

Alternative Route Names	Historical RTAs count (2014–2019)	Predicted RTAs Count	Total Route Length (m)	Travel Time (minutes)
Route 1	4091	108	28.6	41.58
Route 2	3803	123	30.38	48.20
Route 3	3375	129	32.27	50.78

3.3. Safe route determination

Given limited budgets and time constraints, it is crucial to start by identifying hotspot locations and prioritizing areas with the highest crash risk in efforts to improve traffic safety [72]. To find a safe route using network analysis in ArcGIS Pro, we have used road network data and accident crash frequent segments identified by crash rate analysis that probably result in traffic congestion, and two points as the start-end of public facility locations.

We employed the Network Analyst tool to set analysis parameters, including defining the starting point, destination, and safety-related constraints of avoiding high concentrations of accident areas and barrier locations. Using the network route analysis tool, we generated three alternative routes while considering the analysis's safety requirements. Layers for route network analysis were created, and configurations were customized to examine routes in accordance with predetermined safety parameters.

We considered both predictable and historical numbers of RTAs along with other parameters such as travel time and total route length, to identify the safest route among the three alternatives. Prioritizing the route with the fewest historical and predicted RTA counts along with a manageable overall route length and travel time is one preferred approach. The methodology's results verified its efficiency in identifying safer routes, highlighting potential enhancements through iterative analysis and regular data updates.

When considering overall length, travel time, and crash counts (as shown in Table 1 and Fig. 5 below), Route 1 has the lowest expected number of RTAs (108), followed by Route 2 (123) and Route 3 (129). Route 1 also boasts the quickest travel time (41.58 min) and the shortest overall route length (28.6 km).

Identifying the high-risk and safe roads segments with less congestion enables users to make informed choices about their route selection. In route analysis, the identified routes traverse along the resident areas in some sections, avoiding the primary road with accident-prone sections and segments prone to congestion due to these accidents [25]. Heavy traffic congestion can lead to a greater number of crashes and interactions.

Route selection primarily focuses on aligning safest and quickest routes, but this may result in vehicular traffic passing through residential areas [40], optimizing travel time and minimizing the impact of congestion on the road. The safe route map shown in Fig. 6 takes into account the road traffic crash rates, emphasizing areas with historically lower incidents. By navigating clear of segments associated with higher crash rates, the route prioritizes safety.

The analysis delivers safe routing suggestions, and regular data updates maintain applicability. Therefore, in terms of safety, Route 3 would probably be the safest option based on the information available, as shown in Table 1.

4. Discussion

The use of RF for accident prediction provides a comprehensive insight into potential accident-prone areas. By analyzing various predictors, including road features and historical accident data, it assesses the likelihood of accidents across various regions, identifying higher-risk locations. This assessment helps in identifying critical segments within the road network, facilitating targeted safety measures. Visual representation through thematic maps facilitates the identification of risk levels, while validation and accuracy processes ensure the reliability of RF predictions. This helps in determining the reliability of the model's predictions.

In the analysis, the R-squared statistic, ranging from 0 to 1, measures

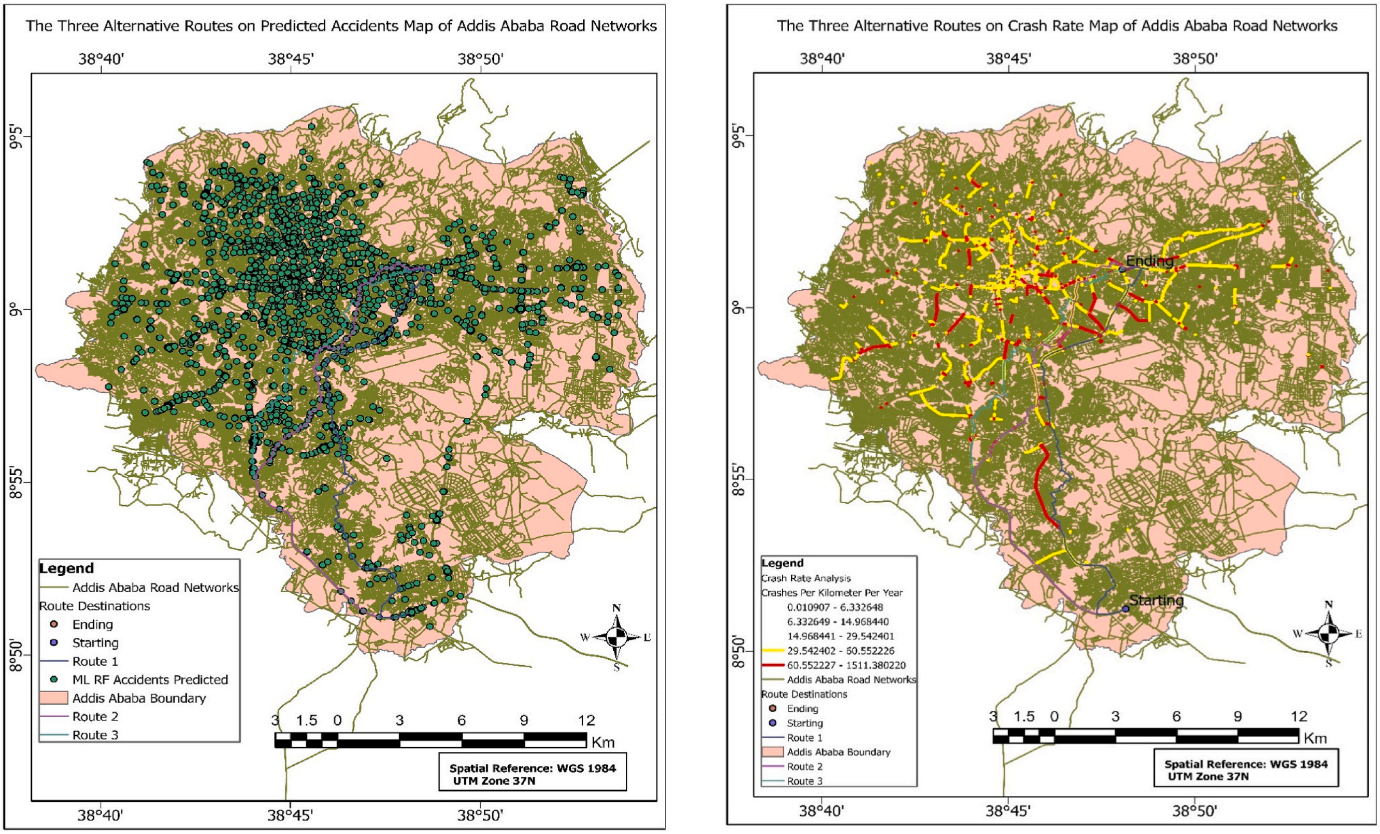

Fig. 5. Comparison of the three alternatives routes on predicted accidents and the crash rate maps.

how well the model aligns with the actual data. A value of 0 indicates poor explanatory power, while 1 signifies a perfect fit [70].

In our study, an R-squared of 0.78 reveals that the model explains around 78 % of the variance in the target variable, demonstrating strong predictive capability. The accuracy of the RF model's predictions is typically assessed and validated using statistical measures or by comparing predicted results against actual accident occurrences, which we were unable to do due to real-time data unavailability. Since the advent of the Internet of Things (IoT), real-time applications have become integral to daily life. Consequently, IoT requires a more dependable infrastructure to fulfill real-time demands that addresses the limitations of cloud computing [73].

The study by Jamal et al. [74] validates this approach. Their findings indicated that ensemble machine learning methods could predict injury severity accurately. Among these methods the Random Forest model demonstrating the highest performance, whereas the statistical approach performed the least effectively. Overall, this method offers a predictive spatial analysis crucial for understanding, visualizing, and proactively addressing safety concerns on the road network. The resulting safe route map integrates the predictive power of RF models with consideration of road crash rates and barriers. This integrated approach aims to deliver a safe and optimized travel experience, aligning with the primary goal of improving road safety and reducing congestion.

To establish a safe route, we implemented comprehensive strategy involving the analysis of the supervised machine-learning models, RF, crash rate calculation, and spatial network analysis functions within ArcGIS Pro. These models analyze historical accident data, incorporating various predictors like road geometry, weather conditions, land use, and time of day. Crash rates for road segments were calculated to determine accident frequency per unit length. Identifying accident-frequent segments involved establishing a basis for high crash rates and prioritizing areas prone to accidents.

In ArcGIS Pro, we created a road network dataset for road segmentation, integrating road network data with attributes of restrictions and crash rates. The route analysis, configured with the Network Analyst extension, found routes based on crash rates to influence route selection. Subsequently, the safe route was generated using spatial network analysis algorithms, visually represented on a map (Fig. 6), alongside predicted accident occurrences on the road network.

The safe route's accuracy was validated by considering predicted accidents on those road segments that have a minimal rate of accidents, especially for safe route segments that completely avoid the existing main road. Compared to Routes 2 and 3, Route 1 has the shortest total route length, the shortest travel time, and the lowest predicted RTA count, despite having the highest historical RTA count. Route 2 shows moderate values for overall route length, travel time, and both historical and predicted RTA counts. Route 3 has the longest total route length, the longest travel duration, and the lowest historical RTA count, but it has the highest predicted RTA count, as shown in Table 1.

Hence, Route 1 appears to be the optimal route choice out because it has the least expected RTA count, the shortest route length, and the shortest travel time. Specifically, Route 1 spans 28.6 km between the two points and takes 41.58 min, making it the optimal route, as shown in Table 1.

The difficulties remain complex after the time dimension was introduced, as people typically budget a large amount of time for travel when planning for important events. These "reliable" paths assist travelers in better scheduling their routes and preparing for the risk of arriving late [36]. While Route 1 appears to be the fastest option, other considerations such as safety, congestion, road conditions, and individual traveler preferences, may also influence the route chosen. This promotes safety while simultaneously lowering the risk of being late, rather than just focusing on minimizing anticipated trip time.

It can be computationally difficult to find the optimal route in a large network that is both the safest route and the shortest in travel time

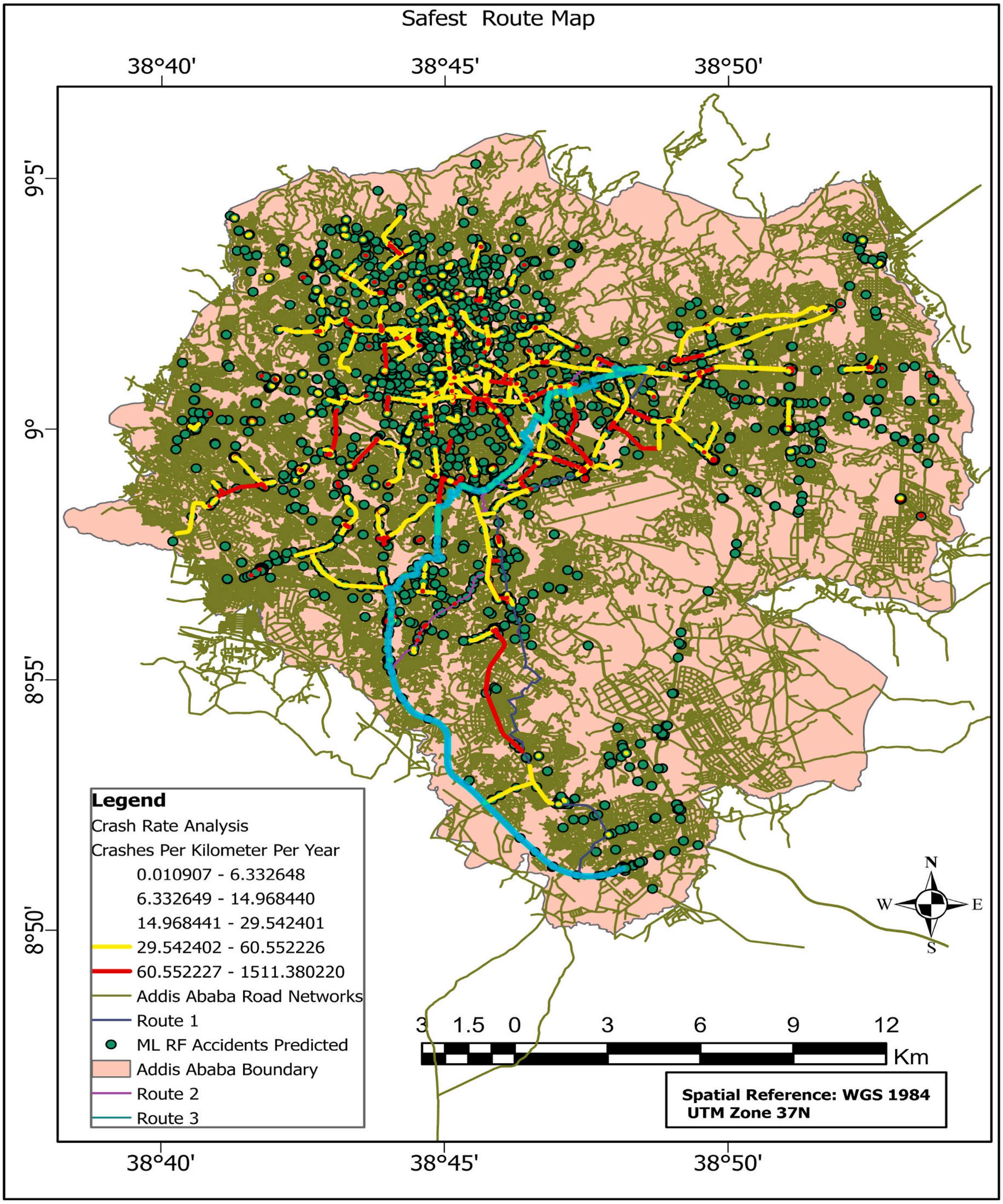

Fig. 6. Identified safest route map.

simultaneously due to the many street connections and intersections [20]. The goal of safety-based routing algorithms is to determine the safest route between any two road network points [35]. Despite being more time-efficient approach, it may not be sufficient to generate the most efficient or optimal route schedules [75]. On top of this, prioritizing travel duration over safety in route selection can have negative consequences for other road users. Incorporating safety considerations and adopting centralized algorithms can mitigate these issues [69].

In planning transportation routes, factors such as inadequate information awareness can also play an important role [1]. Better awareness of accident hotspots and timing assists drivers and pedestrians in choosing safer routes [19]. Therefore, due to the above scenarios, Route 3 is being selected as the safest route of the three alternative routes because, historically, it has recorded the lowest accident counts. This is generally considered safer because they signify actual past incidents that have occurred on the road network, providing tangible evidence of safety levels over time. This route is visually depicted on the safe route map in Fig. 6, providing a tangible representation of the navigation route that shows traversing in some parts of residential areas for safety reasons. The safe route map is a visually intuitive tool that provides information awareness and clear guidance to drivers and road users.

5. Conclusions

Road accidents and resulting traffic congestion are significant global public-health concern that burden society immensely. Addis Ababa, Ethiopia, experiences daily fatalities due to RTAs, which further worsen congestion. Addressing these challenges involves effectively utilizing and analysis crash, spatial and non-spatial data to predict accidents and identify crash-prone road segments and inform road users about alternate safe routes. This study introduces three novel methodologies: RF machine learning, crash rate analysis, and spatial network analysis. These methods are integrated to improve road safety as well as mitigate traffic congestion to deliver route suggestions that prioritize safety while optimizing traffic flow.

Applying the RF model for spatial analysis resulted in predicting 3,096 accidents based on five years of historical accident data and road characteristic attributes from 2015 to 2019 in Addis Ababa city. The model's R-squared value resulted with 0.78 which signifies robust performance, and accounts for roughly 78 % of the variance in the target variable. After a thorough investigation into crash rates analysis, it was discovered that there is a notable concentration of traffic accidents on the road segments connecting any two urban locations. These segments were used as barriers for safe routing analysis to avoid. Therefore, Route 3, which covers a distance of 32.27 km and takes 50.78 min to travel, has been identified as the safest route due to its historically lower accident rate. However, out of three alternative routes, an optimal Route 2 of 28.6 km with a travel time duration of 41.58 min was chosen. The validity of the safer route was confirmed by assessing the outcomes of predicted incidents on specific road segments with low crash rates.

This methodical approach significantly enhances understanding of accident-prone areas and traffic congestions, directly contributing to improved traffic safety and efficient flow control. It alerts travelers and drivers to potential accident risks on specific road segments offering alternate safe routes to ensure safer journeys. This research contributes by integrating spatial analysis and ML techniques to address road accident issues and enhance road safety. The study's findings also advance road safety approaches and mitigate traffic congestion, thereby supporting global efforts to improve urban mobility and infrastructure resilience, significantly contributing to the field. For future, we suggest refining model complexity, validate in diverse environments, update data regularly, integrate real-time information, use traffic volume and user input parameters, and improve crash data recording systems.

CRediT authorship contribution statement

Yetay Berhanu: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Dietrich Schröder:** Writing – review & editing, Supervision, Resources, Conceptualization. **Bikila Teklu Wodajo:** Writing – review & editing, Supervision. **Esayas Alemayehu:** Writing – review & editing, Supervision, Resources, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used *Quillbot*, *Grammarly* and *chatGPT-3.5* in order to grammar check, paraphrase and summarize. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

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 research article.

Data availability

The authors do not have permission to share data.

Acknowledgements

The authors would like to express their gratitude to the Jimma Institute of Technology (JIT), Jimma University (Jimma, Ethiopia), Addis Ababa Science and Technology University (AASTU) (Addis Ababa, Ethiopia) and Stuttgart Technology University of Applied Sciences (HFT Stuttgart, Germany). We would also like to thank to the Addis Ababa Traffic Police Office, Traffic Management Agency, and Land management bureau for kindly providing the necessary data for our research.

References

[1] K. Shubenkova, A. Boyko, P. Buyvol, The technique of choosing a safe route as an element of smart mobility, Transport. Res. Procedia 36 (2018) 718–724, https://doi.org/10.1016/j.trpro.2018.12.100.
[2] Fouzi Harrou, Abdelhafid Zeroual, Farid Kadri, Ying Sun, Enhancing road traffic flow prediction with improved deep learning using wavelet transforms, Results in Engineering 23 (2024) 102342, https://doi.org/10.1016/j.rineng.2024.102342. ISSN 2590-1230.
[3] Sajjad Shafiei, Ziyuan Gu, Meead Saberi, Calibration and validation of a simulation-based dynamic traffic assignment model for a large-scale congested network, Simulat. Model. Pract. Theor. 86 (2018) 169–186, https://doi.org/10.1016/j.simpat.2018.04.006. ISSN 1569-190X.
[4] M. Papageorgiou, C. Kiakaki, V. Dinopoulou, A. Kotsialos, Yibing Wang, Review of road traffic control strategies, Proc. IEEE 91 (12) (2003) 2043–2067, in: https://doi:10.1109/jproc.2003.819610.
[5] E. Ghaffari, A.M. Rahmani, M. Saberikamarposhti, A. Sahafi, An optimal path-finding algorithm in smart cities by considering traffic congestion and air pollution, IEEE Access 10 (2022) 55126–55135, https://doi.org/10.1109/ACCESS.2022.3174598.
[6] World Health Organization (WHO), Global Status Report on Road Safety 2018, World Health Organization, Geneva, Switzerland, 2018.
[7] F. Wegman, The future of road safety: a worldwide perspective, IATSS Res. 40 (2) (2017) 66–71, https://doi.org/10.1016/j.iatssr.2016.05.003.
[8] World Health Organization (WHO). Global Status Report on Road Safety 2023, World Health Organization, Geneva, 2023. Licence: CC BY-NC-SA 3.0 IGO, https://iris.who.int/bitstream/handle/10665/375016/9789240086517-eng.pdf?sequence=1.
[9] World Road Association, (PIARC), Strategic plan 2020-2023: PIARC world road association, Retrieved from, https://piarc.ch/wp-content/uploads/Strategic-Plan-2020-2023-PIARC-World-Road-Association-Update-October-2020.pdf, 2020.
[10] Yetay Berhanu, Esayas Alemayehu, Dietrich Schröder, Examining car accident prediction techniques and road traffic congestion: a comparative analysis of road safety and prevention of world challenges in low-income and high-income countries", J. Adv. Transport. 2023 (2023) 6643412, https://doi.org/10.1155/2023/6643412, 18.
[11] United Nations Economic Commission for Africa and United Nations Economic Commission for Europe (UNECE). Road safety performance review, UNECE, Ethiopia. Geneva, 2020. https://unece.org/sites/default/files/2021-01/RSPR%20Ethiopia%20Report_Final%20web_1.pdf.
[12] Addis Ababa Traffic Management Agency (AATMA), Report: Special Edition Volume 1 No. 1 March/2020, 2020.
[13] J. Docker, M. Johnson, Safe roads for cyclists: an investigation of Australian and Dutch approaches, Journal of the Australasian College of Road Safety 29 (3) (2018) 39–49.
[14] G. Berhanu, Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa, Accid. Anal. Prev. 36 (5) (2004) 697–704, https://doi.org/10.1016/j.aap.2003.05.002. ISSN 0001-4575.
[15] Y. He, Z. Liu, X. Zhou, B. Zhong, Analysis of urban traffic accidents features and corelation with traffic congestion in large-scale construction district, in: Paper Presented at the 2017 International Conference on Smart Grid and Electrical Automation, IEEE, 2017, https://doi.org/10.1109/ICSGEA.2017.110.
[16] N. Geroliminis, M.G. Karlaftis, A. Skabardonis, A spatial queuing model for the emergency vehicle districting and location problem, Transp. Res. Part B Methodol. 43 (7) (2009) 798–811, https://doi.org/10.1016/j.trb.2009.01.006.
[17] S. Sanchez Gonzalez, F. Bedoya-Maya, A. Calatayud, Understanding the effect of traffic congestion on accidents using big data, Sustainability 13 (13) (2021) 7500, https://doi.org/10.3390/su13137500.
[18] C. Pasquale, D. Anghinolfi, S. Sacone, S. Siri, M. Papageorgiou, A two-class traffic control scheme for reducing congestion and improving safety in freeway systems, in: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2016, pp. 1767–1772, https://doi.org/10.1109/ITSC.2016.7795797.
[19] Yetay Berhanu, Schroeder Dietrich, Bikila Teklu, Esayas Alemayehu, Spatial analysis of road traffic accidents: identifying hotspots for improved road safety in Addis Ababa, Ethiopia, Cogent Engineering 10 (2023) 2, https://doi.org/10.1080/23311916.2023.2269655.

[20] Shailesh Chandra, Safety-based path finding in urban areas for older drivers and bicyclists, Transport. Res. Part C 48 (2014) (2014) 143–157, https://doi.org/10.1016/j.trc.2014.08.018.
[21] H. Park, A. Haghani, S. Samuel, M.A. Knodler, Real-time prediction and avoidance of secondary crashes under unexpected traffic congestion, Accid. Anal. Prev. 112 (2018) 39–49, https://doi.org/10.1016/j.aap.2017.11.025.
[22] Institute of Transportation Engineers (ITE), Traffic calming: reducing speed to save lives, Retrieved from, https://www.ite.org/wp-content/uploads/traffic-calming-reducing-speed-to-save-lives.pdf, 2021.
[23] ITF, Road Safety Annual Report, OECD Publishing, Paris, France, 2020, 2020. https://www.itf-oecd.org/road-safety-annual-report-2020.
[24] Sulaiman Yunus, Ishaq A. Abdulkarim, Road traffic crashes and emergency response optimization: a geo-spatial analysis using closest facility and location-allocation methods, Geomatics, Nat. Hazards Risk 13 (1) (2022) 1535–1555, https://doi.org/10.1080/19475705.2022.2086829.
[25] Robert B. Noland, Mohammed A. Quddus, Congestion and safety: a spatial analysis of London, Transport. Res. Pol. Pract. 39 (Issues 7–9) (2005) 737–754, https://doi.org/10.1016/j.tra.2005.02.022. ISSN 0965-8564, https://www.sciencedirect.com/science/article/pii/S0965856405000467.
[26] M.G. Alebe, B.W. Hussen, A.W. Kitila, The challenging the supply side approach: the example of road traffic congestion in Addis Ababa: the case of Torhayloch – Mexico street corridor, Journal of Urban Development Studies 3 (1) (2023) 57–69, https://doi.org/10.58891/ecsujuds.v3i1.155.
[27] C. Kai, Z. Mo, Design of Real-Time Traffic Information Prediction and Simulation System Based on AOSVR and On-Line Learning, IEEE International Conference on Vehicular Electronics and Safety, Shanghai, China, 2006, pp. 189–193, https://doi.org/10.1109/ICVES.2006.371580.
[28] Li Shen, Guofa Li, Cheng Yang, Bin Ran, Urban arterial traffic status detection using cellular data without cellphone GPS information, Transport. Res. C Emerg. Technol. 114 (2020) 446–462, https://doi.org/10.1016/j.trc.2020.02.006. ISSN 0968-090X,, https://www.sciencedirect.com/science/article/pii/S0968090X19306370.
[29] H. Harirforoush, An Integrated GIS-Based and Spatiotemporal Analysis of Traffic Accidents: A Case Study in Sherbrooke. PhD Diss, University of Sherbrooke, 2017, https://doi.org/10.3990/1.9789073946088.
[30] World Road Association, (PIARC), Road safety manual: recommendations from the world road association (PIARC). https://www.piarc.org/en/PIARC-knowledge-base-Roads-and-Road-Transportation/Road-Safety-Sustainability/Road-Safety/safety-manual, 2003.
[31] Gorzelanczyk Piotr, Using neural networks to forecast the number of road accidents in Poland taking into account weather conditions, Results in Engineering 17 (2023) 100981, https://doi.org/10.1016/j.rineng.2023.100981. ISSN 2590-1230.
[32] Y. Fan, X. Zhu, B. She, W. Guo, T. Guo, Network-constrained spatio-temporal clustering analysis of traffic collisions in Jianghan District of Wuhan, China, PLoS One 13 (4) (2018) e0195093, https://doi.org/10.1371/journal.pone.0195093.
[33] A. Ziakopoulos, G. Yannis, A review of spatial approaches in road safety, Accid. Anal. Prev. 135 (2020) 105323, https://doi.org/10.1016/j.aap.2019.105323.
[34] B.D.S.P. Amorim, A.A. Firmino, C.D.S. Baptista, G.B. Júnior, A.C.D. Paiva, F.E.D. A. Júnior, A machine learning approach for classifying road accident hotspots, ISPRS Int. J. Geo-Inf. 12 (6) (2023) 227, https://doi.org/10.3390/ijgi12060227.
[35] Tarek Ghoul, Sayed Tarek, Chuanyun Fu, Real-time safest route identification: examining the trade-off between safest and fastest routes, Analytic Methods in Accident Research 39 (2023) 100277, https://doi.org/10.1016/j.amar.2023.100277. ISSN 2213-6657.
[36] Y.M. Nie, X. Wu, Shortest path problem considering on-time arrival probability, Transp. Res. Part B Methodol. 43 (6) (2009) 597–613, https://doi.org/10.1016/j.trb.2009.01.008.
[37] A. Wahid, M.T.U. Haider, M.M. Ahmad, A.K. Singh, Adaptive Route Planning in Spatial Road Networks, International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2016, pp. 1281–1285, https://doi.org/10.1109/WiSPNET.2016.7566343, 2016,.
[38] K. Goczyłla, J. Cielatkowski, Optimal routing in a transportation network, Eur. J. Oper. Res. 87 (2) (1995) 214–222, https://doi.org/10.1016/0377-2217(95)00177-R.
[39] R. Geisberger, P. Sanders, D. Schultes, D. Delling, Contraction hierarchies: faster and simpler hierarchical routing in road networks, in: Experimental Algorithms: 7th International Workshop, WEA 2008 Provincetown, MA, USA, May 30-June 1, 2008 Proceedings 7 (Pp. 319-333), Springer Berlin Heidelberg, 2008, https://doi.org/10.1007/978-3-540-68552-4_24.
[40] A. Dijkstra, En Route to Safer Roads. How Road Structure and Road Classification Can Affect Road Safety, University of Twente]. SWOV, 2011, https://doi.org/10.3990/1.9789073946088 dissertation [PhD Thesis - Research external, graduation UT.
[41] Yangsong Gu, Diyi Liu, Ramin Arvin, Asad J. Khattak, Lee D. Han, Predicting intersection crash frequency using connected vehicle data: a framework for geographical random forest, Accid. Anal. Prev. 179 (2023) 106880, https://doi.org/10.1016/j.aap.2022.106880. ISSN 0001-4575.
[42] Y. Ali, F. Hussain, M.M. Haque, Advances, challenges, and future research needs in machine learning-based crash prediction models: a systematic review, Accid. Anal. Prev. 194 (2024) 107378, https://doi.org/10.1016/j.aap.2023.107378.
[43] M. Abed Ahmed, AlArjani Ali, Laila f. Seddek, Samia ElAttar, Reduce the delivery time and relevant costs in a chaotic requests system via lean-Heijunka model to enhance the logistic Hamiltonian route, Results in Engineering 21 (2024) 101745, https://doi.org/10.1016/j.rineng.2023.101745. ISSN 2590-1230.
[44] S.R. Vadyala, S.N. Betgeri, J.C. Matthews, E. Matthews, A review of physics-based machine learning in civil engineering, Results in Engineering 13 (2022) 100316, https://doi.org/10.1016/j.rineng.2021.100316. ISSN 2590-1230.
[45] Shakil Ahmed, Md Akbar Hossain, Sayan Kumar Ray, Md Mafijul Islam Bhuiyan, Saifur Rahman Sabuj, A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance, Transp. Res. Interdiscip. Perspect. 19 (2023) 100814, https://doi.org/10.1016/j.trip.2023.100814. ISSN 2590-1982, https://www.sciencedirect.com/science/article/pii/S2590198223000611.
[46] J. Wang, T. Luo, T. Fu, Crash prediction based on traffic platoon characteristics using floating car trajectory data and the machine learning approach, Accid. Anal. Prev. 133 (2019) 105320, https://doi.org/10.1016/j.aap.2019.105320.
[47] M. Kanevski, A. Pozdnukhov, V. Timonin, Machine learning algorithms for GeoSpatial data, Applications and Software Tools" International Congress on Environmental Modelling and Software 53 (2008). https://scholarsarchive.byu.edu/iemssconference/2008/all/53.
[48] M.F. Labib, A.S. Rifat, M.M. Hossain, A.K. Das, F. Nawrine, Road accident analysis and prediction of accident severity by using machine learning in Bangladesh, in: 2019 7th International Conference on Smart Computing & Communications (ICSCC) (Pp. 1-5). 978-1-7281-1557-3/19/$31.00 ©2019 IEEE, 2019, https://doi.org/10.1109/ICSCC.2019.8843640.
[49] S. Olugbade, S. Ojo, A.L. Imoize, J. Isabona, M.O. Alaba, A review of artificial intelligence and machine learning for incident detectors in road transport systems, Math. Comput. Appl. 27 (5) (2022) 77, https://doi.org/10.3390/mca27050077.
[50] K. Kopczewska, Spatial machine learning: new opportunities for regional science, Ann. Reg. Sci. 68 (2022) 713–755, https://doi.org/10.1007/s00168-021-01101-x.
[51] E.A. Atumo, T. Fang, X. Jiang, Spatial statistics and random forest approaches for traffic crash hot spot identification and prediction, Int. J. Inj. Control Saf. Promot. 29 (2) (2022) 207–216, https://doi.org/10.1080/17457300.2021.1983844.
[52] Z. Pu, Z. Li, R. Ke, X. Hua, Y. Wang, Evaluating the nonlinear correlation between vertical curve features and crash frequency on highways using random forests, J. Transport. Eng., Part A: Syst. 146 (10) (2020) 2020.
[53] M. Abdel-Aty, A. Pande, A. Das, W.J. Knibbe, Assessing safety on Dutch freeways with data from infrastructure-based intelligent transportation systems, Transport. Res. Rec. 2083 (1) (2008) 153–161, https://doi.org/10.3141/2083-18.
[54] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[55] B.W. Al-Mistarehi, A.H. Alomari, R. Imam, M. Mashaqba, Using machine learning models to forecast severity level of traffic crashes by R studio and ArcGIS, Front. Built Environ. 8 (2022) 860805, https://doi.org/10.3389/fbuil.2022.860805.
[56] T. Bokaba, W. Doorsamy, B.S. Paul, Comparative study of machine learning classifiers for modelling road traffic accidents, Appl. Sci. 12 (2022) 828, https://doi.org/10.3390/app12020828, 2022.
[57] Pourroostaei Ardakani, Saeid Xiangning Liang, Kal Tenna Mengistu, Richard Sugianto So, Xuhui Wei, Baojie He, Cheshmehzangi Ali, Road car accident prediction using a machine-learning-enabled data analysis, Sustainability 15 (7) (2023) 5939, https://doi.org/10.3390/su15075939.
[58] Z. Wang, S. Zlatanova, Safe route determination for first responders in the presence of moving obstacles, IEEE Trans. Intell. Transport. Syst. 21 (3) (2019) 1044–1053, https://doi.org/10.1109/TITS.2019.2900858.
[59] Nima Hoseinzadeh, Ramin Arvin, Asad J. Khattak, Lee D. Han, Integrating safety and mobility for pathfinding using big data generated by connected vehicles, Journal of Intelligent Transportation Systems 24 (4) (2020) 404–420, https://doi.org/10.1080/15472450.2019.1699077.
[60] Soheil Sohrabi, Lord Dominique, Navigating to safety: necessity, requirements, and barriers to considering safety in route finding, Transport. Res. C Emerg. Technol. 137 (2022) 103542, https://doi.org/10.1016/j.trc.2021.103542. ISSN 0968-090X.
[61] C. Barrington-Leigh, A. Millard-Ball, The world's user-generated road map is more than 80% complete, PLoS One 12 (8) (2017) e0180698, https://doi.org/10.1371/journal.pone.0180698.
[62] K.Z. Phyo, Optimal route identification on complex road network by using advanced Dijkstra's algorithm, Dissertation (Doctoral dissertation, MERAL Portal) (2019).
[63] T. Hastie, R. Tibshirani, J.H. Friedman, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2, springer, New York, 2009, pp. 1–758.
[64] S. Das, I. Tsapakis, Interpretable machine learning approach in estimating traffic volume on low-volume roadways, International Journal of Transportation Science and Technology (2019), https://doi.org/10.1016/j.ijtst.2019.09.004.
[65] Oregon Department of Transportation (ODOT), Highway safety manual analysis: procedure manual version 2. https://www.oregon.gov/ODOT/Engineering/Pages/HSM.aspx, 2018.
[66] Manushi Majumdar, Spatial data science solution for vision zero, Retrieved from, https://www.esri.com/arcgis-blog/products/constituent-engagement/constituent-engagement/spatial-data-science-solution-for-vision-zero/, 2018.
[67] ArcGIS Solutions, Use traffic crash analysis, Retrieved from, https://doc.arcgis.com/en/arcgis-solutions/11.2/reference/use-traffic-crash-analysis.htm, 2023.
[68] J.M. Vorster, M.H.P. Zuidgeest, Cycle route network development and evaluation using spatial multi-criteria analysis and shortest path analysis, Southern African Transport Conference (2019). http://hdl.handle.net/11427/30136.
[69] S. Sohrabi, Y. Weng, S. Das, S.G. Paal, Safe route-finding: a review of literature and future directions, Accid. Anal. Prev. 177 (2022) 106816, https://doi.org/10.1016/j.aap.2022.106816.
[70] Esri, How forest-based and boosted classification and regression works, Retrieved from, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-forest-works.htm, 2023.

[71] Esri, ArcGIS solutions: introduction to traffic crash analysis, Retrieved from, https://doc.arcgis.com/en/arcgis-solutions/latest/reference/introduction-to-traffic-crash-analysis.htm, 2021.
[72] H. Harirforoush, L. Bellalite, A new integrated GIS-based analysis to detect hotspots: a case study of the city of Sherbrooke, Accid. Anal. Prev. 130 (2016) 62–74, https://doi.org/10.1016/j.aap.2016.08.015.
[73] G. Javadzadeh, A.M. Rahmani, M.S. Kamarposhti, Mathematical model for the scheduling of real-time applications in IoT using Dew computing, J. Supercomput. (2022) 7464–7488, https://doi.org/10.1007/s11227-021-04170-z.
[74] M. Ijaz, L. lan, M. Zahid, A. Jamal, A comparative study of machine learning classifiers for injury severity prediction of crashes involving three-wheeled motorized rickshaw, Accid. Anal. Prev. 154 (July 2020) (2021) 106094, https://doi.org/10.1016/j.aap.2021.106094.
[75] Krishna Veer Tiwari, Satyendra Kumar Sharma, An optimization model for vehicle routing problem in last-mile delivery, Expert Syst. Appl. 222 (2023) 119789, https://doi.org/10.1016/j.eswa.2023.119789. ISSN 0957-4174.