TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–16

DOI: 10.1177/03611981241270161
journals.sagepub.com/home/trr
Sage

Association Patterns of Work Zone Crashes using Bayesian Network

Subasish Das[1], M. Ashifur Rahman[2], Jinli Liu[1], Xinyue Ye[3] and Boniphace Kutela[4]

Abstract

Ensuring the safety of work zones is a top priority for transportation agencies because of the dangers posed by vehicles changing lanes and paths within these areas. Recent statistics highlight the seriousness of this issue, showing a shocking 46% increase in fatal collisions within work zones in 2019 compared with 2011. Therefore, this study examined crashes related to intrusions or encroachments in work zones to uncover the underlying mechanisms. Analyzing four years of crash data (2016–2019) from the Texas Department of Transportation, this research utilized Bayesian network to identify crucial factors, their relationships, and potential alternative scenarios. The severity of injuries in work zone intrusion accidents was significantly influenced by male drivers, curved roads, and specific patterns of driver distraction and condition. The study revealed three distinct scenarios with complete probability of specific attributes: (1) crashes on rural non-principal arterial roads; (2) collisions with non-barrier fixed objects; and (3) non-injury crashes involving non-barrier fixed objects and driving violations. The detailed findings from this study can provide valuable insights for safety engineers, enabling them to reduce work zone crashes caused by encroachments. By comprehending the key factors and their effects, transportation agencies can implement effective measures to lessen the risks associated with work zone encroachments, ultimately creating a safer environment for both drivers and road workers.

Keywords

work zones, safety, Bayesian network, worker safety, crash data

A work zone refers to a critical section of a road network where construction activities occur, leading to lane closures, detours, and the use of additional signs and warning devices. Ensuring the safety of work zones is a primary concern for transportation agencies because of the potential hazards posed to both drivers and workers. Data from the Fatality Analysis Reporting System show a troubling increase in fatal incidents within work zones over the past decade, rising from 521 in 2010 to 762 in 2019 (*1*). Work zone crashes constituted 1.7% of all fatal collisions in 2010, which escalated to 2.3% in 2019. In Texas alone, there were more than 25,000 reported work zone collisions in 2019, averaging one collision every 20 min (*1*). The proximity of work zone equipment, traffic control devices, and workers to travel lanes significantly heightens the risk of collisions (*2*). Over the last 10 years, more than 1,300 workers lost their lives in work zone crashes (*3*). The dynamic nature of work zones, with abrupt changes in travel paths and lanes, further increases crash risks, particularly if drivers become distracted and fail to notice warning signals and markings. Intrusion into the work zone area and road departures are among the major types of crashes occurring in work zones, affecting both drivers and workers. An urgent and comprehensive approach to addressing work zone safety is required to safeguard the public and work zone personnel from these hazards.

Although prior studies have explored factors implicated in rear-end collisions at work zones (*4–6*), there has not been extensive research on work zone encroachment crashes. Roadway work zone encroachment refers to a

[1]Ingram School of Engineering, Texas State University, San Marcos, TX
[2]Louisiana Transportation Research Center, Baton Rouge, LA
[3]School of Architecture, Texas A&M University, College Station, TX
[4]Texas A&M Transportation Institute, Houston, TX

Corresponding Author:
Subasish Das, subasish@txstate.edu

vehicle traveling outside the designated lane(s) of travel, and within work zones, it indicates a vehicle unintentionally crossing into the boundaries of the work zone. A few studies have attempted to investigate crashes involving vehicle encroachments in work zone areas (*7–9*). A vehicle's potential trajectory during an encroachment is influenced by several factors, including speed limits, roadway, roadside characteristics (such as alignment and width), weather conditions, and driver behavior. The frequency of roadway departures in work zones, the higher risk of fatalities, and the limited availability of information on encroachment conditions within work zones all highlight the immediate need for a comprehensive study in this field.

A Bayesian network (BN), a probabilistic graphical method, is highly suitable for improving work zone safety because of the complex nature of work zone encroachment crashes and the research gaps in this area. A BN is well suited for this task as it can handle multiple factors simultaneously and identify their interdependencies, providing a comprehensive understanding of work zone encroachment crashes. Moreover, the scarcity of information on encroachment conditions in work zones and the higher fatality risks associated with such crashes emphasize the need for a thorough study. Overall, this study aims to answer the following key research questions:

- **RQ1:** What factors contribute to work zone encroachment crashes and their increased risk within work zones?
- **RQ2:** How can a BN be utilized to analyze the complex interactions among factors influencing work zone encroachment crashes, providing valuable insights for safety improvement measures?

The choice to utilize BN in this study stems from its suitability for analyzing complex relationships and uncertainties within the dataset. Work zone safety is influenced by numerous interrelated factors, including driver behavior, road conditions, environmental factors, and traffic patterns. BN provides a robust framework for modeling these complex relationships by capturing both direct and indirect dependencies among variables. One key advantage of BN is its ability to handle probabilistic reasoning, which aligns with the inherently uncertain nature of crash data and the comprehensive nature of work zone intrusions. By incorporating probabilistic models, BN enables the assessment of various potential scenarios and their likelihoods, allowing for a comprehensive understanding of the underlying mechanisms contributing to work zone crashes. Furthermore, BN facilitates the identification of influential factors and their interactions, which may not be readily apparent through traditional statistical methods. This approach allows for the exploration of complex causal pathways and the identification of high-risk scenarios that may not be evident from descriptive analyses alone.

The paper's organization involves a literature review, data preparation, encroachment determination, exploratory data analysis, and the methodology used in this study. Subsequent sections present the results of the BN analysis, followed by a discussion of the findings and concluding remarks. This study sheds light on the intricacies of work zone encroachment crashes, providing valuable insights for safety improvement measures and strategies to protect drivers and work zone personnel from the inherent risks in work zones.

Literature Review

The transportation infrastructure of the United States heavily relies on roadways, yet over 43% of the U.S. roadway system is currently rated as poor or mediocre (*10*). Consequently, there is a growing demand for maintenance and rehabilitation projects, which involve repairing, resurfacing, or replacing sections of roadways and roadway structures. These projects often need to be executed while ensuring the flow of traffic nearby as the complete roadway closure is often impractical. Work zones have evolved from basic layouts, with varying levels of attention to safety, to modern designs that prioritize safety whenever feasible. Despite substantial efforts to improve work zone safety, collisions remain an ongoing concern.

Various factors, including driver behaviors and work zone characteristics, influence work zone-related crashes. Deviations from appropriate driving behaviors, such as speeding, following too closely, or making sudden lane changes, increased the crash risk and chances of injury by 10% (*11*). Speeding was reported in 71.4% of fatal work zone crashes compared with 30% in non-work zone fatal crashes (*12*). Some studies show that crashes were found to occur not only within the active work zone but also in areas like the taper and approach zones (*13*). The National Cooperative Highway Research Program Report 869 emphasized the impact of work zone attributes on crash severity and called for further research (*2*). Garber and Woo observed a significant increase in crash rates in work zones, with multilane highway crashes rising by 57% and crashes on two-lane urban highways skyrocketing by 168% (*14*). Work zone configurations, roadway facility types, access control, number of lanes, and other factors contributed to crash severity (*15*). Work zone areas were classified into distinct zones, and rear-end collisions were predominant in all work zones, particularly in warning zones (*16–18*). Several studies identified work zone activity areas as hotspots for crashes (*19–21*).

Understanding crash-contributing factors is crucial for developing effective safety measures and policies to reduce work zone-related crashes. Some studies showed that various crash types were associated with work zones, with rear-end collisions being the most frequent (*6*, *19*, *22*, *23*). Roadway departure, a form of single-vehicle crash (SVC), was another significant crash type, and these crashes have been analyzed in studies (*24*). Roadway departure crashes accounted for 54% of traffic fatalities in the United States, with multiple factors influencing their occurrence (*25*). Studies examined factors like driver behavior, road alignment, driver distraction, visibility, and roadway conditions concerning roadway departure crashes (*26–28*). Studies also examined work zone intrusion crashes by considering various factors such as roadway conditions, vehicle failure, speeding, driver inattention, and impairment (*7*, *29*). Research on work zone encroachment crashes during daytime and nighttime revealed that factors like driver inattention and abrupt changes in traffic flow were primary contributing factors (*8*, *9*).

The state of roadways in the United States demands urgent attention, with a significant portion classified as poor or mediocre. As work zone activities increase, so do collisions, making it crucial to understand the contributing factors. Driver behavior, including inaccurate maneuvers and speeding, plays a major role in work zone crashes, alongside work zone characteristics like signage and visibility. Rear-end collisions prevail in work zone advance warning areas, and work zone encroachments and single-vehicle crashes are also notable crash types. In-depth studies highlight factors influencing roadway departure crashes. The deliberate actions of drivers also contribute to work zone encroachments.

Several prior studies suggest that work zone encroachment crashes are especially concerning because of the risks they pose to both workers within the work zones and the drivers involved. Although some research has offered a simplistic quantitative description of work zone encroachment crashes, no earlier study has thoroughly integrated the multifaceted factors of driver behavior, roadway conditions, and specific crash characteristics. This research gap is addressed in this study, as work zone encroachment crashes are distinctively analyzed considering the multifaceted factors involved.

The BN-based probabilistic graphical method was selected for its inherent capacity to model intricate interdependencies between variables. In the context of work zone encroachment crashes, a myriad of factors is observed to interact in non-linear and interconnected ways. Through the BN approach, these relationships can be visualized and quantified in ways some traditional statistical techniques might not comprehensively capture. A comprehensive approach to gain more insights into work zone-related encroachment crashes is essential for improving work zone safety, minimizing collisions, and protecting workers and drivers alike.

Methodology

Probabilistic Graphical Method

A BN, a probabilistic graphical model, is constructed to explore the relationships among potential factors influencing work zone crash severity. Utilizing Bayes' rule on posterior simulation, the BN is a directed graphical model that incorporates prior probability distribution before observing evidence and posterior probability distribution after observing the evidence. Its nodes and links represent random variables, their associations, and conditional probability distributions for each variable's states. By applying Bayes' rule, the BN can predict the likelihood of future events based on the occurrence of previous events (*30*). BN provides an intuitive representation of interactions between components and their effects on the dependent variable when modifying a variable. Its visual depiction serves as a causal map, enhancing the understanding of the system compared with conventional analysis tools. Moreover, the training dataset used for the BN's learning algorithm is derived from actual events, making the model more grounded in real-world scenarios, and ensuring that it captures the complexities of work zone crash severity. This reliance on real data allows a more accurate representation of the system's behavior and better predictions of future events. In recent years, many safety studies have utilized BN (*31–34*).

BNs are graphical models that represent the probabilistic relationships between variables through directed links, forming a directed acyclic graph (DAG). The directed links indicate the conditional dependencies between variables, showing how changes in one variable may affect the probabilities of other variables. Feedback loops are not allowed in BNs, ensuring a clear and acyclic representation of the causal relationships between the variables (*35*). Nielsen and Jensen define BNs as a statistical multivariate model for a set of variables, $X = \{X_1, \ldots\ldots X_a\}$ represented by two components (*36*). In the qualitative component, a DAG represents the variables with each vertex representing a variable and the links between them indicating their statistical dependency. The graphical model then factorizes the joint probability distribution for this DAG. The quantitative component is made up of the conditional distribution $p(X_i \mid pa(X_i))$ for each $X_i, \ldots = 1\ldots\ldots n$ given its parents in the graph, denoted as $pa(x_i)$. Each node can have a Conditional Probability Table (CPT) associated with it, representing the strength of causal connections. By populating these CPTs using various information sources like model simulation, measured data, economic analysis, expert opinion, and stakeholder surveys, the BN allows for network quantification. BN models commonly utilize discrete nodes with a limited number of states, such as categories or intervals. BNs provide a systematic and concise representation of a joint probability distribution.

$$P(V) = \prod_{nV} P(X|pa(X))$$

where V is the set of discrete nodes X_1 to X_n, and $pa(X)$ are the parents of X.

One key feature that sets BNs apart from other modeling techniques, such as neural networks or decision trees, is their ability to incorporate prior knowledge about specific nodes in the network effectively. Using structural links in BNs enables the modeling of conditional probabilities and dependencies between nodes based on available data and domain expertise.

In the context of BNs, one key approach is to analyze counterfactual scenarios. A counterfactual scenario is a hypothetical situation where one or more variables in the network are set to specific values to explore potential outcomes under different conditions. This approach allows for examining "what-if" questions, such as the probability of an event if a certain factor had been different. Network quantifications in counterfactual scenarios are beneficial in practice as they enable decision-makers to understand the potential impact of interventions or changes in the system. By comparing outcomes of different scenarios, practitioners can identify influential factors and make informed decisions to optimize their efforts.

Data Preparation

The data preparation for the model development includes the following sub-tasks:

- **Data Source:** The main data source for this study was the Texas Department of Transportation crash database, covering four years (2016–2019) and containing information on crash specifications, vehicle units, and involved persons.
- **Data Filtering:** A unique crash identification number was used to merge the crash, unit, and person files, and vehicle and driver data were used for analysis. Work zone and work zone-related crashes were chosen for further investigation, whereas intersection and junction-related crashes were excluded because of complexity.
- **Focus on Specific Crashes:** The analysis specifically targeted encroachment and roadway departure crashes, regardless of the specifications of the intersection or junction location. Encroachment crashes were identified by filtering SVCs and harmful events associated with encroachment. For multiple vehicle crashes, only opposite sideswipe, head-on, and overturn crashes were considered if they involved or resulted in vehicle encroachment. The direction of encroachment (left-side or right-side) was determined based on available data on harmful event location and crash type.
- **Data Cleaning:** During the data preparation, outliers and redundant features were eliminated from the dataset. After filtering, 10,676 encroachment cases were identified in the Texas data.
- **Attribute Selection:** The final work zone crash database consisted of fifty attributes representing possible crash-contributing factors.
- **Exploratory Data Analysis:** The subsequent section presents the exploratory data analysis of the prepared dataset, aiming to gain insights into the characteristics and patterns of work zone crashes in Texas.

Data Description

Table 1 presents the description of encroachment-related work zone crash data, including different injury categories (fatal, serious, moderate, minor, and no injury). In Texas, there were 10,676 encroachment-related work zone crashes from 2016 to 2019, of which 3,678 resulted in fatality or injury (34%), whereas 6,998 resulted in no injuries (property damage only [PDO]). Among the 121 Fatal crashes, the majority occurred on straight roadways (69.4%), during daylight (41.3%), and involved male drivers (79.3%). On the other hand, 30.6% of Fatal crashes happened on curved roadways. For Serious Injury crashes ($N = 390$), 77.9% occurred on straight roadways, and 50.3% occurred during daylight. As regards driver gender, 31.5% of Serious Injury crashes involved female drivers. In Moderate Injury crashes ($N = 1,440$), 77.8% occurred on straight roadways, 47.8% happened during daylight, and 37.9% involved female drivers. For Minor Injury crashes ($N = 1,727$), 79.8% occurred on straight roadways, and 52.6% occurred during daylight. As regards driver gender, 42.0% of Minor Injury crashes involved female drivers. Lastly, for No Injury crashes ($N = 6,998$), 80.4% occurred on straight roadways, and 46.2% happened during daylight. With regard to driver gender, 63.6% of No Injury crashes involved male drivers. The severity-based distribution emphasizes the importance of a multi-faceted approach to examining work zone crashes, considering factors such as roadway design, lighting conditions, driver behavior, and gender dynamics.

Table 1. Percentage Distribution of Key Attributes by Crash Severity

Attribute	Fatal *N* = 121	Serious injury *N* = 390	Moderate injury *N* = 1,440	Minor injury *N* = 1,727	No injury *N* = 6,998
RDWY_ALIGN (roadway alignment)					
Curve	37 (30.6%)	83 (21.3%)	312 (21.7%)	334 (19.3%)	1,326 (18.9%)
Other	0 (0.00%)	3 (0.77%)	7 (0.49%)	15 (0.87%)	46 (0.66%)
Straight	84 (69.4%)	304 (77.9%)	1,121 (77.8%)	1,378 (79.8%)	5,626 (80.4%)
LIGHT_COND (lighting condition)					
Dark-lighted	26 (21.5%)	74 (19.0%)	327 (22.7%)	387 (22.4%)	1,743 (24.9%)
Dark-not lighted	39 (32.2%)	111 (28.5%)	368 (25.6%)	382 (22.1%)	1,724 (24.6%)
Daylight	50 (41.3%)	196 (50.3%)	688 (47.8%)	908 (52.6%)	3,233 (46.2%)
Other	6 (4.96%)	9 (2.31%)	57 (3.96%)	50 (2.90%)	298 (4.26%)
RDWY_GRADE (roadway grade)					
Grade	27 (22.3%)	82 (21.0%)	314 (21.8%)	302 (17.5%)	1,312 (18.7%)
Hillcrest/uphill/downhill	7 (5.79%)	13 (3.33%)	73 (5.07%)	101 (5.85%)	358 (5.12%)
Level	87 (71.9%)	292 (74.9%)	1,046 (72.6%)	1,309 (75.8%)	5,282 (75.5%)
Other	0 (0.00%)	3 (0.77%)	7 (0.49%)	15 (0.87%)	46 (0.66%)
WRK_PRSNT (worker presence)					
No	106 (87.6%)	310 (79.5%)	1,120 (77.8%)	1,309 (75.8%)	5,407 (77.3%)
Yes	15 (12.4%)	80 (20.5%)	320 (22.2%)	418 (24.2%)	1,591 (22.7%)
EVNT_WTHR_COND (weather condition)					
Clear	87 (71.9%)	280 (71.8%)	966 (67.1%)	1,093 (63.3%)	4,418 (63.1%)
Other	26 (21.5%)	72 (18.5%)	264 (18.3%)	336 (19.5%)	1,194 (17.1%)
Rain/snow	8 (6.61%)	37 (9.49%)	207 (14.4%)	298 (17.3%)	1,345 (19.2%)
Unknown	0 (0.00%)	1 (0.26%)	3 (0.21%)	0 (0.00%)	41 (0.59%)
D1_DR_COND (driver 1 condition)					
Apparently normal	93 (76.9%)	291 (74.6%)	1,167 (81.0%)	1,421 (82.3%)	5,832 (83.3%)
Not-normal	24 (19.8%)	87 (22.3%)	223 (15.5%)	228 (13.2%)	788 (11.3%)
Unknown	4 (3.31%)	12 (3.08%)	50 (3.47%)	78 (4.52%)	378 (5.40%)
TRAF_CTRL (traffic control)					
Controlled	12 (9.92%)	47 (12.1%)	141 (9.79%)	124 (7.18%)	635 (9.07%)
No controls	16 (13.2%)	46 (11.8%)	223 (15.5%)	330 (19.1%)	1,320 (18.9%)
Other	93 (76.9%)	297 (76.2%)	1,076 (74.7%)	1,273 (73.7%)	5,043 (72.1%)
D1_GEN (driver 1 gender)					
F	25 (20.7%)	123 (31.5%)	546 (37.9%)	726 (42.0%)	2,019 (28.9%)
M	96 (79.3%)	267 (68.5%)	888 (61.7%)	995 (57.6%)	4,448 (63.6%)
Unknown	0 (0.00%)	0 (0.00%)	6 (0.42%)	6 (0.35%)	531 (7.59%)
FRST_HARM_EVNT (first harmful event)					
Collision—barrier	42 (34.7%)	161 (41.3%)	732 (50.8%)	911 (52.8%)	3,822 (54.6%)
Collision—other-fix	33 (27.3%)	119 (30.5%)	350 (24.3%)	456 (26.4%)	1,804 (25.8%)
Collision—WZ-equip	16 (13.2%)	35 (8.97%)	151 (10.5%)	191 (11.1%)	967 (13.8%)
Non-collision	30 (24.8%)	75 (19.2%)	207 (14.4%)	169 (9.79%)	405 (5.79%)
D1_DR_DSTR (driver 1 distraction)					
Distracted	2 (1.65%)	32 (8.21%)	172 (11.9%)	202 (11.7%)	850 (12.1%)
Not distracted	115 (95.0%)	347 (89.0%)	1,220 (84.7%)	1,449 (83.9%)	5,784 (82.7%)
Unknown	4 (3.31%)	11 (2.82%)	48 (3.33%)	76 (4.40%)	364 (5.20%)
V1_FRST_DR_ACTN (vehicle 1 driver action)					
Improper driving	17 (14.0%)	61 (15.6%)	236 (16.4%)	328 (19.0%)	1,148 (16.4%)
No contributing action	39 (32.2%)	163 (41.8%)	600 (41.7%)	719 (41.6%)	2,963 (42.3%)
Other	7 (5.79%)	10 (2.56%)	30 (2.08%)	30 (1.74%)	201 (2.87%)
Unknown	4 (3.31%)	11 (2.82%)	48 (3.33%)	76 (4.40%)	364 (5.20%)
Violation	54 (44.6%)	145 (37.2%)	526 (36.5%)	574 (33.2%)	2,322 (33.2%)
FUN_CLASS (functional class)					
Rural other	18 (14.9%)	54 (13.8%)	128 (8.89%)	96 (5.56%)	403 (5.76%)
Rural principal arterial	26 (21.5%)	64 (16.4%)	182 (12.6%)	142 (8.22%)	731 (10.4%)
Unknown	10 (8.26%)	83 (21.3%)	276 (19.2%)	398 (23.0%)	1,706 (24.4%)
Urban other	3 (2.48%)	15 (3.85%)	36 (2.50%)	39 (2.26%)	148 (2.11%)
Urban principal arterial	64 (52.9%)	174 (44.6%)	818 (56.8%)	1,052 (60.9%)	4,010 (57.3%)
TRAF_WAY (traffic way)					
Two-way divided	77 (63.6%)	210 (53.8%)	903 (62.7%)	1,109 (64.2%)	4,319 (61.7%)
Two-way TWLTL	2 (1.65%)	16 (4.10%)	34 (2.36%)	58 (3.36%)	258 (3.69%)
Two-way undivided	32 (26.4%)	81 (20.8%)	226 (15.7%)	161 (9.32%)	704 (10.1%)
Unknown	10 (8.26%)	83 (21.3%)	277 (19.2%)	399 (23.1%)	1,717 (24.5%)

Note: TWLTL = Two way left turn lane.

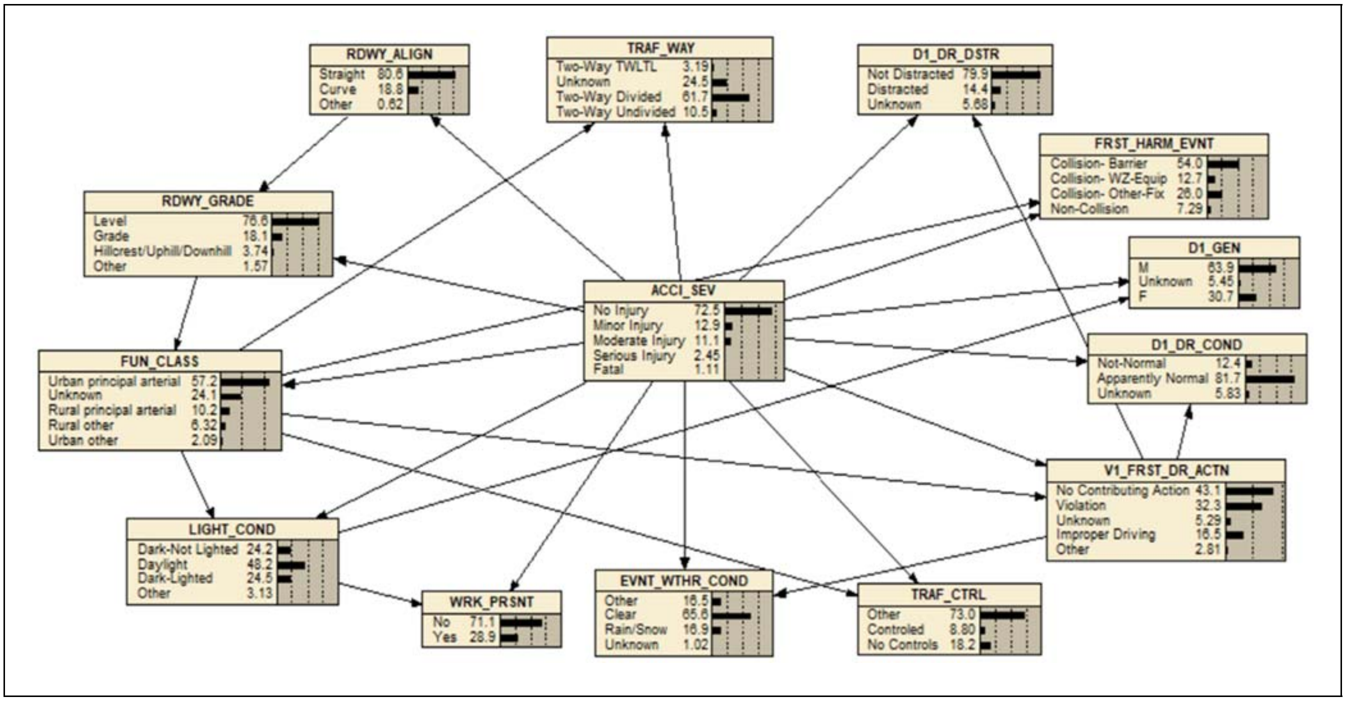

Figure 1. Bayesian network of full data.

Results and Discussions

The BN model for this study was constructed using a series of steps. Firstly, a DAG representing the model was created using Netica 6.04, user-friendly software tailored for BN applications, offering an efficient user interface (*37*). The constructed networks were then used to perform various types of inference using advanced and effective methods. To fit the network to multiple input–output combinations obtained from the Monte Carlo runs from the model, Netica's expectation maximization (EM) technique (*38*) was employed. This learning process resulted in estimates of the model parameters of the BN, known as conditional probability distributions. The final layout of the BN network is shown in Figure 1, where each node's CPTs are depicted as belief bars.

Figure 1 shows the initial data of work zone-related collisions modeled through a BN, centered around injury severities representing a real-world scenario. In this model, most crashes resulted in no injury (72.5%), followed by minor injury (12.9%) and moderate injury (11.1%). Only 2.45% resulted in serious injury, and only 1.11% resulted in fatality. The roadway was usually straight (80.6%), level (76.6%), two-way divided (61.7%), and an urban principal arterial road (57.2%). Traffic control was most reported as other (73.0%), followed by no controls (18.2%). Driver 1 was usually a male (63.9%), with a normal condition (81.7%), and not distracted (75.9%). There was usually no contributing action as vehicle 1 first action (43.1%), followed by a violation (32.3%). The first harmful event was usually a collision with a barrier (54.0%). The environment was usually clear weather (65.6%) and in the daylight (48.2%). Workers were only present in 28.9% of the crashes.

Tables 2 and 3 provide conditional probability scores for two variables by certain other categories: (1) lighting condition by crash severity and functional class, and (2) work zone protection type by crash severity and functional class. Higher scores are displayed in dark red, and lower in lighter red to almost white. Specifically, Table 2 shows the conditional probability of different crash severities given functional class and lighting conditions. The highest probability score occurred with a fatal crash occurring on a rural principal arterial class in the daylight (0.748), followed by a fatal crash on an urban "other" class (0.679). Overall, the highest probability scores occurred with a lighting condition of daylight.

Similarly, Table 3 displays the likelihood of different types of work zones being involved in a crash, given the severity of the crash and the functional class. The highest probability occurs with a no-injury crash occurring on an unknown functional class, with the crash type being a collision with a barrier (0.652). The second highest probability score occurs with a moderate injury on an urban principal arterial functional class with a collision with a barrier (0.640). Overall, the highest probabilities occurred with a collision with a barrier.

Table 2. Conditional Probability Table for Lighting Condition by Crash Severity and Functional Class

Crash severity	Functional class	Lighting condition			
		Daylight	Dark-not lighted	Dark-lighted	Other
Fatal	Urban principal arterial	0.304	0.322	0.331	0.043
Fatal	Rural principal arterial	0.748	0.143	0.037	0.072
Fatal	Rural other	0.480	0.515	0.002	0.002
Fatal	Urban other	0.679	0.311	0.005	0.005
Fatal	Unknown	0.171	0.217	0.483	0.130
Serious Injury	Urban principal arterial	0.405	0.237	0.333	0.026
Serious injury	Rural principal arterial	0.646	0.306	0.033	0.015
Serious injury	Rural other	0.447	0.546	0.004	0.003
Serious injury	Urban other	0.528	0.348	0.118	0.006
Serious injury	Unknown	0.606	0.213	0.170	0.010
Moderate injury	Urban principal arterial	0.442	0.222	0.313	0.023
Moderate injury	Rural principal arterial	0.551	0.361	0.059	0.029
Moderate injury	Rural other	0.569	0.354	0.015	0.062
Moderate injury	Urban other	0.461	0.473	0.051	0.016
Moderate injury	Unknown	0.484	0.196	0.277	0.042
Minor injury	Urban principal arterial	0.555	0.179	0.242	0.024
Minor injury	Rural principal arterial	0.552	0.423	0.009	0.016
Minor injury	Rural other	0.536	0.378	0.019	0.067
Minor injury	Urban other	0.351	0.513	0.118	0.018
Minor injury	Unknown	0.512	0.190	0.281	0.017
No injury	Urban principal arterial	0.466	0.228	0.279	0.027
No injury	Rural principal arterial	0.580	0.324	0.050	0.047
No injury	Rural other	0.452	0.509	0.010	0.028
No injury	Urban other	0.379	0.453	0.103	0.065
No injury	Unknown	0.461	0.161	0.338	0.039

Note: The shading in the cells represents the relative frequency or intensity of certain crash outcomes on different types of roadways. Darker shades indicate higher frequencies, while lighter shades indicate lower frequencies.

Table 3. Conditional Probability Table for Work Zone Type by Crash Severity and Functional Class

Crash severity	Functional class	Work zone type			
		Collision—barrier	Collision—WZ-equip	Collision—other-fix	Non-collision
Fatal	Urban principal arterial	0.474	0.186	0.194	0.146
Fatal	Rural principal arterial	0.154	0.029	0.321	0.497
Fatal	Rural other	0.180	0.003	0.387	0.429
Fatal	Urban other	0.351	0.260	0.385	0.003
Fatal	Unknown	0.191	0.119	0.345	0.345
Serious injury	Urban principal arterial	0.462	0.129	0.276	0.133
Serious injury	Rural principal arterial	0.352	0.081	0.215	0.351
Serious injury	Rural other	0.229	0.005	0.414	0.352
Serious injury	Urban other	0.472	0.007	0.347	0.174
Serious injury	Unknown	0.396	0.096	0.418	0.090
Moderate injury	Urban principal arterial	0.640	0.097	0.179	0.084
Moderate injury	Rural principal arterial	0.413	0.062	0.250	0.275
Moderate injury	Rural other	0.256	0.064	0.392	0.287
Moderate injury	Urban other	0.194	0.182	0.282	0.343
Moderate injury	Unknown	0.402	0.107	0.406	0.086
Minor injury	Urban principal arterial	0.591	0.093	0.237	0.079
Minor injury	Rural principal arterial	0.486	0.158	0.160	0.196
Minor injury	Rural other	0.237	0.104	0.413	0.245
Minor injury	Urban other	0.308	0.196	0.278	0.218
Minor injury	Unknown	0.446	0.104	0.369	0.081
No injury	Urban principal arterial	0.652	0.120	0.198	0.030
No injury	Rural principal arterial	0.473	0.100	0.315	0.111
No injury	Rural other	0.344	0.069	0.389	0.198
No injury	Urban other	0.449	0.155	0.289	0.107
No injury	Unknown	0.420	0.202	0.337	0.041

Note: The shading in the cells represents the relative frequency or intensity of certain crash outcomes on different types of roadways. Darker shades indicate higher frequencies, while lighter shades indicate lower frequencies.

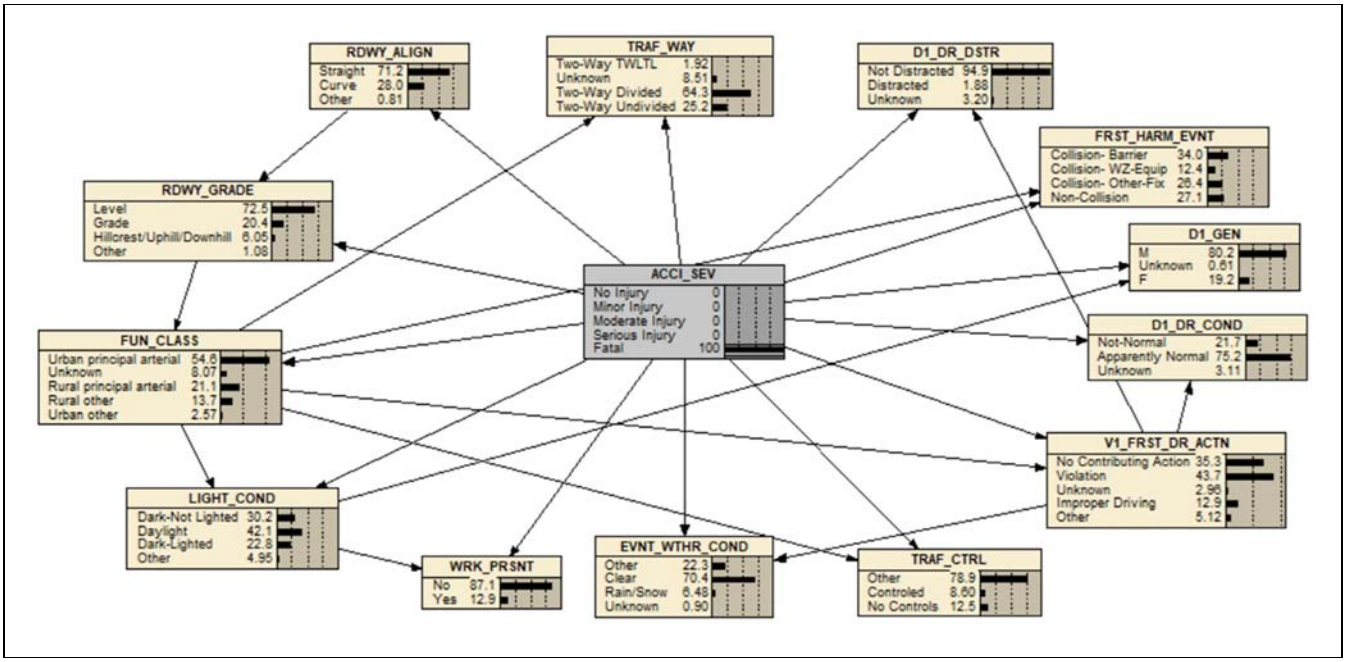

Figure 2. Counterfactual Bayesian network considering all injuries are fatal.

Counterfactual Scenarios

Statistically significant counterfactual analysis is a powerful technique used to assess the impact of variable changes on outcomes. In the context of Netica Software, it involves conducting parameter learning of CPTs using a counting-learning technique based on the Tree-Augmented Naive Bayes (TAN) model. This approach allows the software to learn from examples and construct accurate CPTs, representing the conditional probabilities of variables given to their parents in the graphical model. Once the CPTs have been generated, Netica Software can compute posterior probabilities for each variable. This means it can estimate the probability of a particular event occurring given the evidence observed in the data. By analyzing these probabilities, one can identify how different variables influence the likelihood of specific outcomes, providing insights into the factors contributing to safety hazards and crash occurrences. The results of this statistical examination can offer intriguing preliminary findings that can be utilized to develop safety precautions and crash prevention strategies to improve work zone safety.

Figure 2 shows the BN for work zone-related crashes by considering the counterfactual scenario of all work zone injuries as fatal crashes. In the fatal crash model, a larger proportion of roadways were two-ways divided (25.2%; 2.4 times the full data) and rural principle arterial roads (21.1%; 2.07 times the full data). More roadways were curved (28.0%; 1.49 times the full data). The findings align with previous studies on work zone-related rear-end crashes (*39*, *40*), which suggest that curved sections tend to result in high levels of injury severity. More of the drivers were male (80.2%; 1.26 times the full data), were not distracted (94.9%; 1.19 times the full data) and were not in a normal condition (21.7%; 1.75 times the full data) than in the real-world scenario. According to the study conducted by Wang (40), male drivers traveling through work zones tend to be more involved in fatal rear-end crashes than female drivers. Non-collision was a more frequent first harmful event than in the real-world model (27.1%; 3.72 times the full data), and vehicle 1's first action was more often a violation (43.7%; 1.35 times the full data). More crashes occurred in clear weather (70.4%; 1.07 times the full data), and fewer occurred with workers present (12.9%; a reduction of 55.36%). Roadway grade, lighting condition, and traffic control had only minimal changes.

Figure 3 shows the BN for work zone-related crashes by considering the counterfactual scenario of all work zone injuries as serious injury crashes. Fewer collisions occurred on urban principal arterial roads (42.3%; 1.35 times the full data). More of the drivers were not distracted (88.7%; 1.11 times the full data) and not in a normal condition (23.4%; 1.89 times the full data). More first harmful events were collisions with other fixed objects (32.4%; 1.25 times the full data). More crashes involved traffic controls (13.2%; 1.5 times the full) and occurred in clear weather (73.1%; 1.11 times the full data). Like the fatal scenario, fewer crashes occurred with workers present (18.2%; 0.37 times less than the full

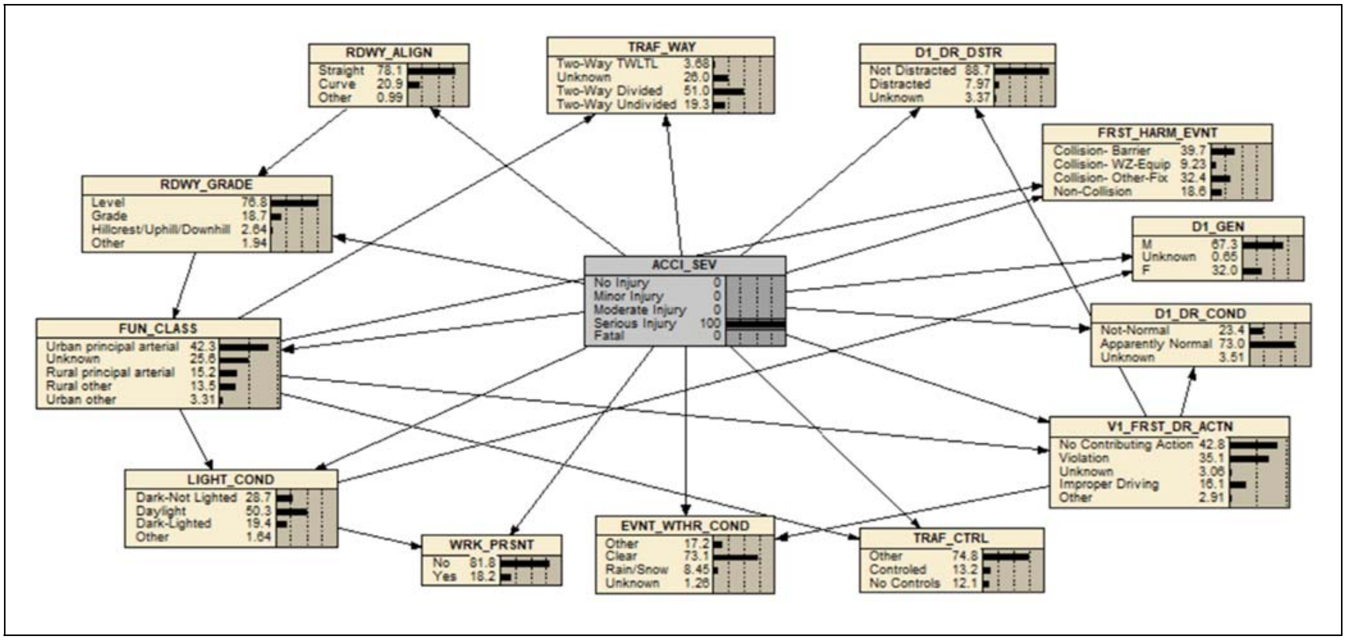

Figure 3. Counterfactual Bayesian network considering all injuries are serious injury.

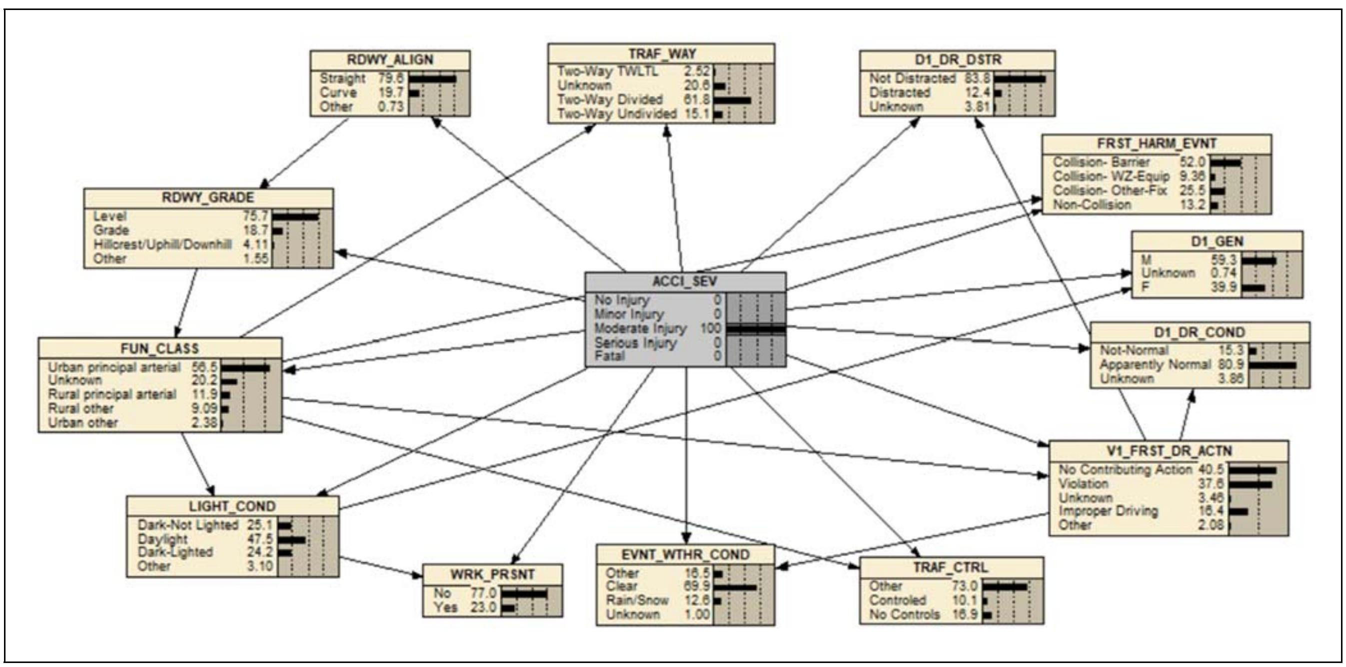

Figure 4. Counterfactual Bayesian network considering all injuries are moderate injury.

data). Traffic way, driver 1's gender, roadway grade, lighting condition, and vehicle 1's first action had only minimal changes.

Figure 4 shows the BN for work zone-related crashes by considering the counterfactual scenario of all work zone injuries as moderate injury crashes. More females than males were in the moderate injury crashes (39.9%; 1.30 times the original data). Similar results were found in Koilada et al. (*41*), where female drivers were found to be 1.18 times more likely to be involved in a moderate injury crash in the activity area compared with male drivers. The odds of female drivers being involved in a severe injury crash in the activity area are 0.23 times lower than those for male drivers. This indicates that males are more

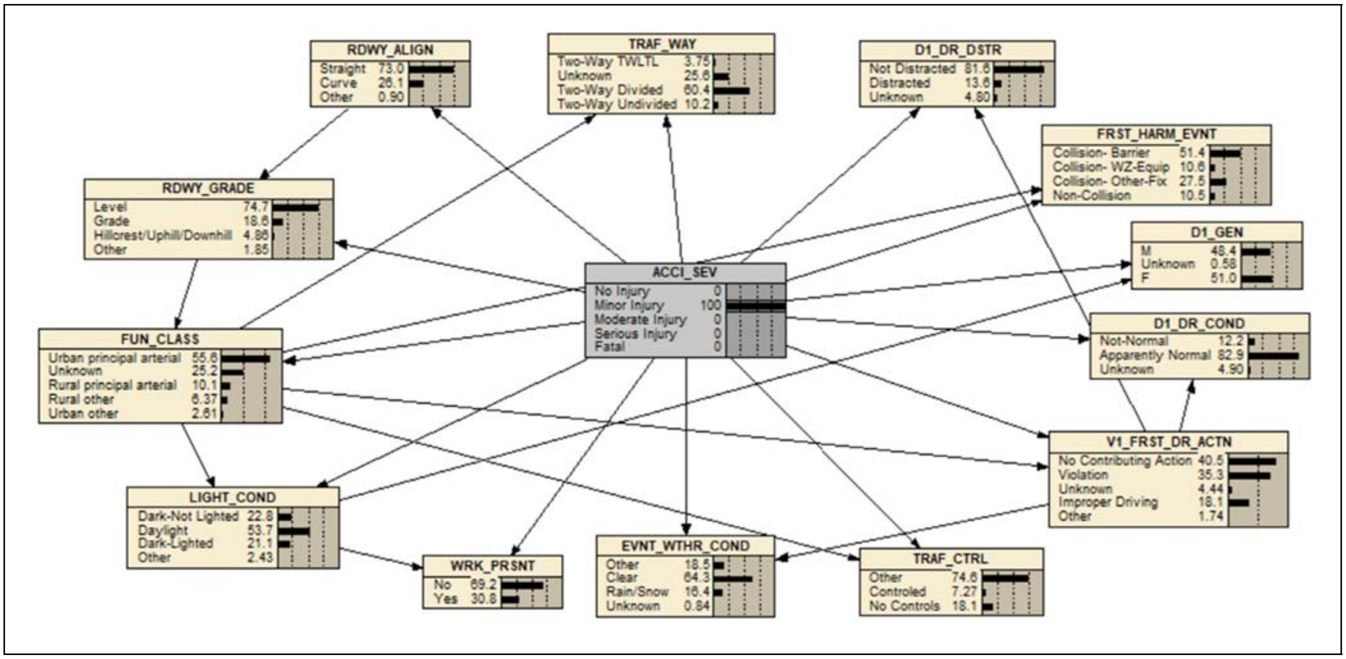

Figure 5. Counterfactual Bayesian network considering all injuries are minor injury.

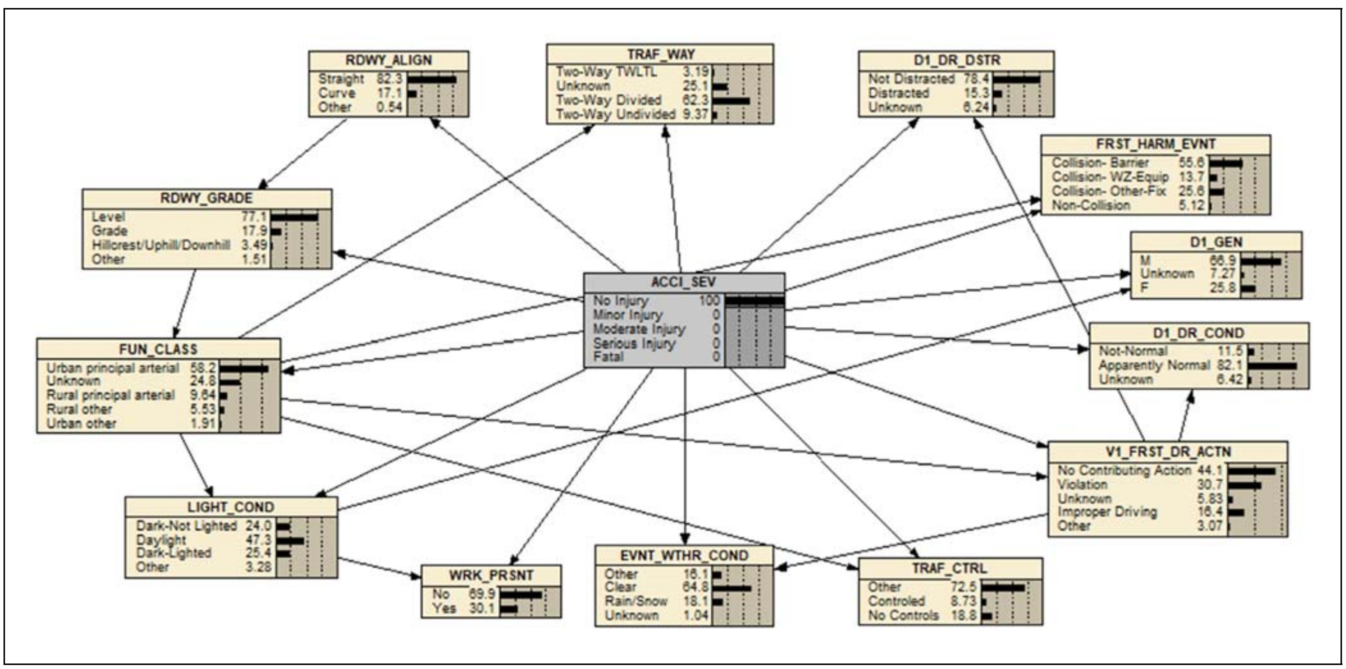

Figure 6. Counterfactual Bayesian network considering all injuries are no injury.

likely to experience a severe crash. Interestingly, none of the other attributes had major changes from the original real-world data.

Figure 5 shows the BN for work zone-related crashes by considering the counterfactual scenario of all work zone injuries as minor injury crashes. Like the moderate injury scenario, more drivers were female (51.0%; 1.66 times the full data). In addition, more crashes occurred on roads with a curved alignment (26.1%, which is 1.39 times the full data). None of the other attributes had any major changes from the full data.

Figure 6 shows the BN for work zone-related crashes by considering the counterfactual scenario of all work zone injuries as no-injury crashes. Interestingly, none of

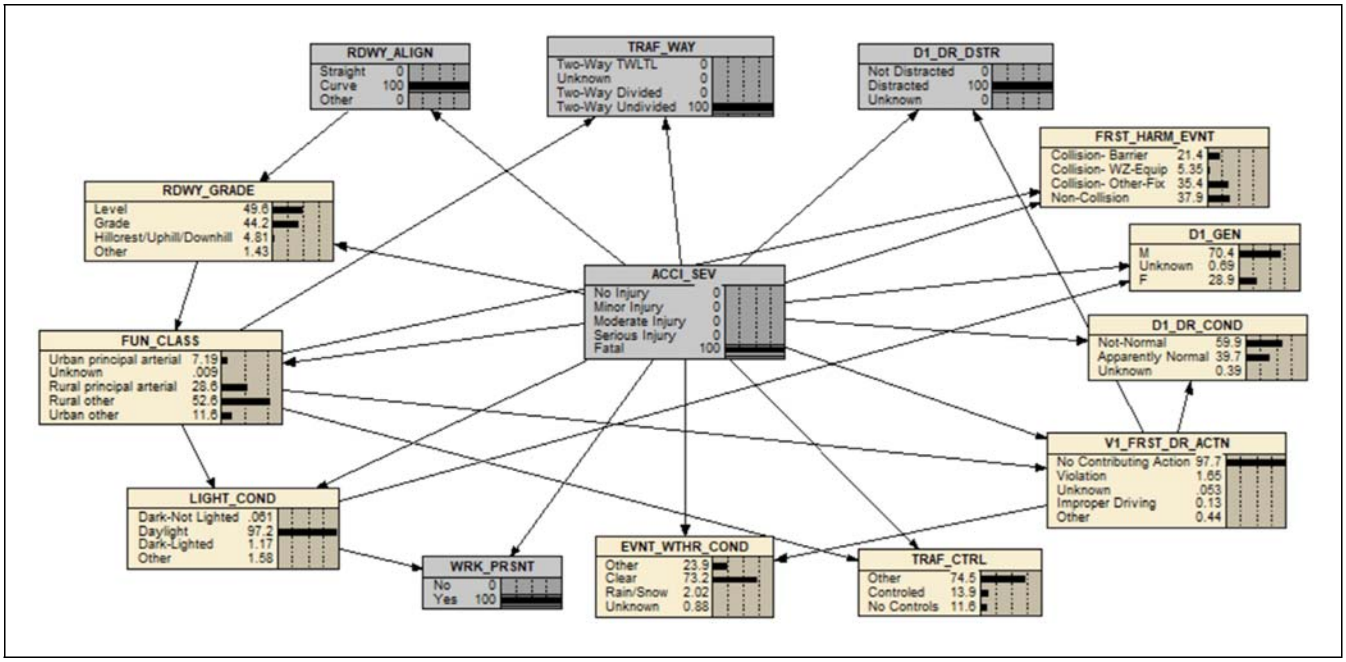

Figure 7. Special case 1.

the attributes had major changes from the real-world data. This is most likely because of the representativeness of no-injury data in real-world data (around 73% of crash data are no-injury crashes).

Special Case Scenarios

A significant benefit of BN lies in its capacity for joint strategies, enabling the integration of multiple hypotheses to ascertain the probability of another variable. Figure 7 shows the BN for work zone-related crashes with fatal crashes, curved roadway alignments, two-way undivided traffic ways, driver 1 distraction, and worker presence all at 100% (special case 1). Fewer crashes involved a collision with a barrier as the first harmful event (21.4%; a reduction of 60.37%), and more involved non-collision as the first harmful event (37.9%; 5.20 times the full data). More drivers were male (70.4%; 1.10 times the full data) and in a not-normal condition (59.9%; 4.83 times the full data). No contributing action was vehicle 1's first action in 97.7% of the special case 1 model (2.27 times the full data). Fewer crashes occurred on roads with no traffic controls (11.6%; a reduction of 36.26%), and fewer crashes occurred in the rain/snow (2.02%; a reduction of 88.05%). More crashes occurred in the daylight (97.2%; 2.02 times the full data). More crashes occurred on rural other roads (52.6%; 8.32 times the full data) and on graded roads (44.2%; 2.44 times the full data).

Figure 8 shows the BN for work zone-related crashes with serious injury crashes, dark-not-lighted lighting conditions, worker presence, and no traffic controls at 100% (special case 2). Fewer roads were two-way divided (50.1%; a reduction of 18.95%). More of the drivers were not distracted (88.5%; 1.11 times the full data), were not in a normal condition (23.4%; 1.89 times the full data) and were female (36.3%; 1.18 times the full data). More collisions involved collisions with other fixed objects as the first harmful event (32.9%; 1.27 times the full data). Fewer collisions involved rain/snow (8.49%; a reduction of 49.76%). Fewer collisions involved urban principal arterial as the function class (44.5%; a reduction of 22.20%). Roadway alignment, roadway grade, roadway alignment, and Vehicle 1 first action had only minor changes.

Figure 9 shows the BN for work zone-related crashes with a curved roadway alignment, two-way undivided traffic way, and worker presence all at 100% (special case 3). No injury was the most prevalent crash severity (62.4%). More crashes involved collisions with other fixed objects (34.1%; 1.31 times the full data) or non-collision (19.2%; 2.63 times the full data) as the first harmful event, and more crashes involved violations as the vehicle 1 first action (42.3%; 1.31 times the full data). More drivers were female (36.4%; 1.19 times the full data). More crashes involved traffic control (23.5%; 2.67 times the full data). More collisions involved daylight (80.2%; 1.66 times the full data) for the lighting conditions and rural other roads as the functional class (56.0%; 8.86 times the full data). More collisions occurred on graded roadways (40.6%; 2.24 times the full data). Driver 1 distraction, driver 1 condition, and weather had minimal changes.

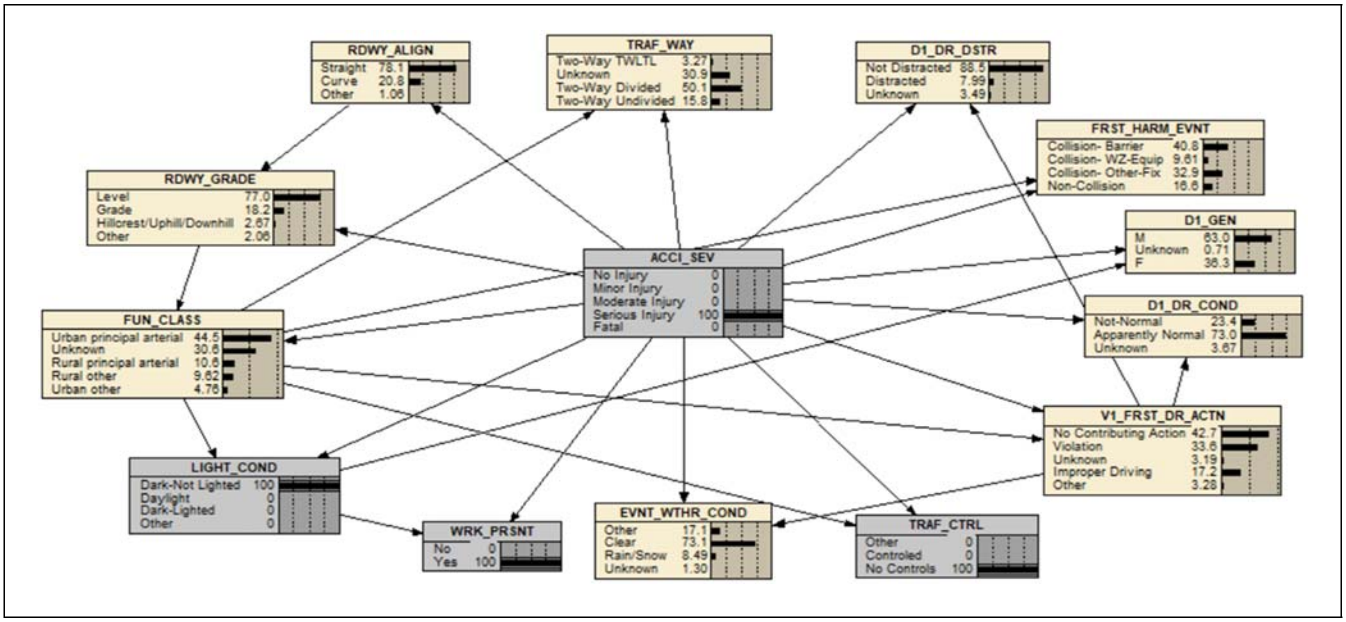

Figure 8. Special case 2.

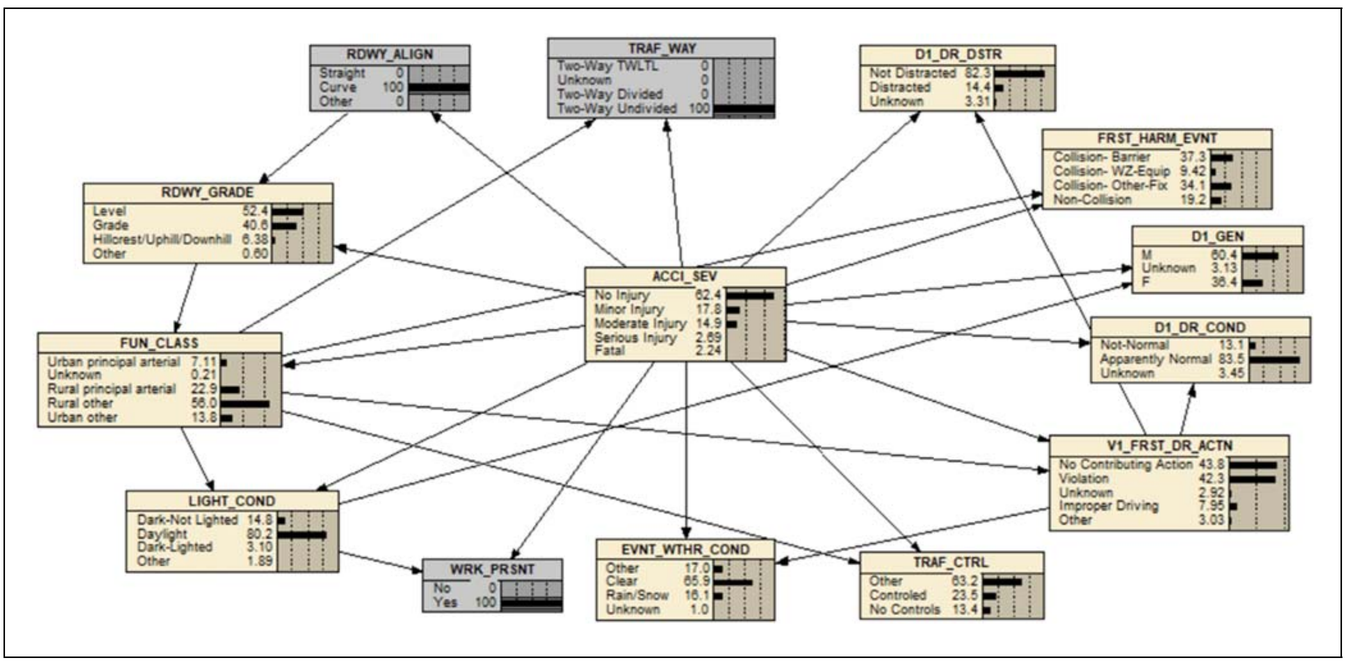

Figure 9. Special case 3.

Summary Findings and Countermeasure Implications

The initial data analysis using a BN showed that most crashes resulted in no injury (72.5%), followed by minor injury (12.9%) and moderate injury (11.1%). Workers were present in only 28.9% of the crashes. CPTs for lighting conditions and work zone protection types were analyzed, revealing the influence of different variables on crash severity and functional class. Furthermore, the study explored counterfactual scenarios to understand the impact of variable changes on outcomes.

One scenario considered all work zone injuries as fatal crashes, which resulted in a higher proportion of roadways being two-way divided and rural principal arterial roads. Male drivers, not in normal condition and not distracted, were more involved in fatal crashes. Non-collision was a more frequent first harmful event, and violations were more common as vehicle 1 first action. Another scenario considered all injuries as serious injuries, leading to changes in road types and first harmful events. Unlike the fatal crash scenarios, female drivers had higher representation in moderate and minor injury scenarios. Similar analyses were conducted for moderate, minor, and no-injury scenarios. The moderate injury model indicated a heightened probability of curved roads being involved and a greater prevalence of female drivers in moderate injury crashes. Interestingly, the model for no-injury crashes did not show significant changes from the real-world data in relation to the attributes studied.

Special cases were investigated by setting multiple variables at 100% probability. These scenarios provided insights into the influence of other variables. For example, in special case 1, more crashes occurred on rural other roads, and collisions with barriers were less frequent. Special case 2 showed more collisions involving other fixed objects and less involvement of rain/snow. Special case 3 revealed more no-injury crashes, collisions with other fixed objects, and violations as the first vehicle action.

Effective countermeasures to reduce work zone encroachment crashes should be implemented to improve overall roadway safety and mitigate the risks associated with these incidents. Based on the findings from the study, the following countermeasures can be considered:

- **Enhanced Warning Signs and Traffic Control:** Improved and conspicuous warning signs and clear and visible traffic control measures can help alert drivers about upcoming work zones in advance. This includes providing sufficient distance and warning time to allow drivers to safely adjust their speed and lane positions. A significant number of crashes in special cases occurred on roads with no traffic controls, underscoring the importance of conspicuous warning signs.
- **Temporary Traffic Barriers:** Installing temporary traffic barriers, such as crash cushions and water-filled barriers, can prevent vehicles from encroaching into work zones. Special case 1 indicated a reduction in crashes involving a collision with barriers, highlighting the potential effectiveness of temporary traffic barriers. However, when installing temporary traffic barriers, it is important to consider factors such as the duration and extent of the work zone, the speed limit on the roadway, visibility conditions, and the presence of roadside hazards such as drop-offs.
- **Speed Reduction and Enforcement:** Implementing reduced speed limits in work zones and ensuring strict enforcement of these limits can discourage speeding and give drivers more time to react to changing conditions. The prominence of curved roadways in fatal crashes underscores the need for reduced speed limits in challenging alignments.
- **Public Awareness Campaigns:** Public awareness campaigns to educate drivers about the potential dangers and risks associated with work zones can promote responsible driving behavior and increase compliance with safety measures. The overrepresentation of male drivers in fatal scenarios suggests that targeted campaigns for this demographic might be beneficial.
- **Proper Work Zone Design:** Ensuring proper work zone design with clear lane markings, adequate signage, and well-defined traffic patterns can reduce confusion for drivers and minimize the likelihood of encroachments. Various injury scenarios have emphasized specific roadway types and alignments, highlighting the importance of well-designed work zones. The finding of fatal crash involvement in work zones on curved roads potentially implies that extra precautions may be necessary, along with supplementary facilities such as advanced warning signs on curved roads.
- **Real-Time Traffic Information:** With many crashes occurring in clear weather, real-time traffic updates can offer additional insights for drivers. Providing real-time traffic information through Variable Message Signs can inform drivers about current work zone conditions, helping them make informed driving decisions.
- **Driver Training and Education:** Offering specialized training for drivers on the subject of safe driving practices in work zones can enhance their awareness and preparedness for encountering such situations. The high percentage of drivers not in a normal condition in some scenarios suggests an area that driver training can address.
- **Innovative Technologies:** Implementing advanced technologies, such as automated warning systems, collision avoidance systems, and intelligent transportation systems, can enhance safety and reduce the risk of work zone encroachment crashes. The variability in the first harmful events across different scenarios highlights the potential for innovative technologies to address specific crash types.

Conclusions

Despite substantial improvements in overall roadway safety, there are still a concerning number of fatalities at work zones, with 752 fatal crashes reported in the U.S. in 2019 alone. Previous research indicates that many work zone crashes result from vehicle encroachments, where vehicles briefly leave their travel lanes, and even minor mistakes can lead to severe consequences. However, there needs to be more information on work zone encroachment conditions, with the lack of such data hindering safety researchers from identifying crash-contributing factors and understanding the circumstances leading to these crashes. This study utilizes Texas work zone crash data to analyze driver, roadway, and environmental conditions that could be associated with encroachment-related crashes, including crash conditions, driver information, and vehicle details. The findings from the figures and tables provide valuable insights into work zone encroachment-related crashes. Most work zone crashes resulted in no injury, with serious injury and fatality accounting for a small percentage. Lighting conditions and functional class influenced crash severity. Male drivers were more likely to be involved in fatal and serious injury crashes, whereas female drivers had higher representation in moderate and minor injury scenarios. Collisions with other fixed objects and non-collision events were more common in specific counterfactual scenarios. Curved roadways showed correlations with higher crash probabilities in certain cases.

The study makes several unique contributions and policy implications for enhancing work zone safety. Firstly, it applies a BN model to analyze work zone-related crashes, revealing complex interactions among factors influencing crash severity. This allows policymakers to prioritize resources effectively and implement targeted safety measures. Secondly, the study identifies crash patterns related to road classifications and lighting conditions through empirical analysis, providing evidence-based insights for formulating strategies to mitigate work zone crashes. Thirdly, the study proposes joint strategies within the BN model to address specific scenarios contributing to severe crashes, enabling customized interventions for maximum impact. Lastly, by quantifying conditional probabilities, the study guides policymakers in setting priorities and tailoring approaches based on the unique characteristics of different work zones, leading to precise, effective, and adaptable safety policies.

This study on work zone encroachment crashes provides valuable insights into factors influencing crash severity, but it has some limitations that suggest potential future research directions. The lack of detailed crash sequence information and the need to consider external factors and human behavior could be addressed in future studies to enhance the model's accuracy and predictive capabilities. Temporal and spatial analyses could reveal trends and regional variations in work zone encroachment crashes. In contrast, comparative analyses with other crash types within work zones would offer a more comprehensive understanding of work zone safety. Incorporating factors such as flow, composition, work zone layout and safety warning types into the analysis offers a deeper insight into traffic volume and patterns, important precursors to work zone encroachment crashes. This inclusion can also shed light on potential shortcomings in specific warning mechanisms. Additionally, research on the effectiveness of safety measures and strategies to improve work zone worker safety would further enhance work zone safety practices and policies. Validating the BN model using additional datasets and continuous improvement would ensure its reliability and applicability in diverse work zone contexts.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Subasish Das; data collection: Subasish Das; analysis and interpretation of results: Subasish Das and M. Ashifur Rahman; draft manuscript preparation: Subasish Das, M. Ashifur Rahman, Jinli Liu, Xinyue Ye, and Boniphace Kutela. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Subasish Das https://orcid.org/0000-0002-1671-2753
M. Ashifur Rahman https://orcid.org/0000-0001-6940-1599
Jinli Liu https://orcid.org/0000-0002-6152-8808
Boniphace Kutela https://orcid.org/0000-0002-5450-1623

References

1. *Fatality Analysis Reporting System (FARS)*. National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington, D.C.
2. The National Work Zone Safety Information Clearinghouse. Library of Resources to Improve Roadway Work Zone Safety for All Roadway Users. https://www.workzonesafety.org/. Accessed July 26, 2021.
3. Ullman, G. L., M. Pratt, S. Geedipally, B. Dadashova, R. J. Porter, J. Medina, and M. D. Fontaine. *Analysis of Work Zone Crash Characteristics and Countermeasures*. Transportation Research Board, Washington, D.C., Publication 0309471281. 2018.
4. Meng, Q., and J. Weng. Evaluation of Rear-End Crash Risk at Work Zone Using Work Zone Traffic Data. *Accident Analysis & Prevention*, Vol. 43, No. 4, 2011, pp. 1291–1300. https://doi.org/10.1016/j.aap.2011.01.011.
5. Weng, J., Q. Meng, and X. Yan. Analysis of Work Zone Rear-End Crash Risk for Different Vehicle-Following

Patterns. *Accident Analysis & Prevention*, Vol. 72, 2014, pp. 449–457.

6. Zhang, K., and M. Hassan. Identifying the Factors Contributing to Injury Severity in Work Zone Rear-End Crashes. *Journal of Advanced Transportation*, Vol. 2019, 2019, p. 4126102. https://doi.org/10.1155/2019/4126102.
7. Bryden, J. E., L. B. Andrew, and J. S. Fortuniewicz. Intrusion Accidents on Highway Construction Projects. *Transportation Research Record: Journal of the Transportation Research Board*, 2000. 1715: 30–35.
8. Ullman, G. L., M. D. Finley, J. E. Bryden, R. Srinivasan, and F. M. Council. *Traffic Safety Evaluation of Nighttime and Daytime Work Zones*. NCHRP Report. Transportation Research Board, Washington, D.C., 2008.
9. Ullman, G. L., M. D. Finley, and L. Theiss. Categorization of Work Zone Intrusion Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2258: 57–63.
10. ASCE. *A Comprehensive Assessment of America's Infrastructure: 2021 Infrastructure Report Card*. American Society of Civil Engineers, Washington, D.C., 2021.
11. Liu, T., Y. Yang, G.-B. Huang, Y. K. Yeo, and Z. Lin. Driver Distraction Detection Using Semi-Supervised Machine Learning. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 4, 2016, pp. 1108–1120. https://doi.org/10.1109/TITS.2015.2496157.
12. Fatality Analysis Reporting System (FARS). NHTSA. https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars. Accessed July 19, 2021.
13. Raub, R. A., O. B. Sawaya, J. L. Schofer, and A. Ziliaskopoulos. Enhanced Crash Reporting to Explore Work Zone Crash Patterns. Paper No. 01, Northwestern University, Evanston, IL,Vol. 166, 2001.
14. Garber, N. J., and T.-S. H. Woo. *Accident Characteristics at Construction and Maintenance Zones in Urban Areas*. Virginia Transportation Research Council, Charlottesville, VA, 1990.
15. Osman, M., R. Paleti, and S. Mishra. Analysis of Passenger-Car Crash Injury Severity in Different Work Zone Configurations. *Accident Analysis & Prevention*, Vol. 111, 2018, pp. 161–172. https://doi.org/10.1016/j.aap.2017.11.026.
16. Garber, N. J., and M. Zhao. Distribution and Characteristics of Crashes at Different Work Zone Locations in Virginia. *Transportation Research Record: Journal of the Transportation Research Board*, 2002. 1794: 19–25.
17. Das, S., A. Dutta, R. Tamakloe, and M. N. Khan. Analyzing the Time-Varying Patterns of Contributing Factors in Work Zone-Related Crashes. *Journal of Transportation Safety & Security*, Vol. 16, No. 6, 2023, pp. 655–682.
18. Dobrovolny, C. S., B. Dadashova, M. Tabesh, S. Das, H. Kwon, R. Bligh, L. E. Riexinger, C. P. Smith, H. C. Gabler, and S. Hallmark. *Addressing Encroachment-Related Safety Issues in Work Zones: A Guide*. Transportation Research Board, Washington, D.C., 2023.
19. Hargroves, B. T. Vehicle Accidents in Highway Work Zones. *Transportation Engineering Journal of ASCE*, Vol. 107, No. 5, 1981, pp. 525–539. https://doi.org/10.1061/TPEJAN.0000946.
20. Nemeth, Z., and A. Rathi. Freeway Work Zone Accident Characteristics. *Transportation Quarterly*, Vol. 37, No. 1, 1983, pp. 145–159.
21. Das, S., M. Tabesh, B. Dadashova, and C. Dobrovolny. Diagnosis of Encroachment-Related Work-Zone Crashes by Applying Pattern Recognition. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 222–236.
22. Abdel-Aty, M. Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models. *Journal of Safety Research*, Vol. 34, No. 5, 2003, pp. 597–603. https://doi.org/10.1016/j.jsr.2003.05.009.
23. Anastasopoulos, P. C., and F. L. Mannering. An Empirical Assessment of Fixed and Random Parameter Logit Models Using Crash-and Non-Crash-Specific Injury Data. *Accident Analysis & Prevention*, Vol. 43, No. 3, 2011, pp. 1140–1147.
24. Akepati, S. R., and S. Dissanayake. Characteristics of the Work Zone Crashes. Presented at the First Congress of Transportation and Development Institute (TDI), Chicago, IL, 2011.
25. McGee, H. W. Transportation Research Board, and National Academies of Sciences, Engineering, and Medicine. *Practices for Preventing Roadway Departures*. Transportation Research Board, Washington, D.C., 2018.
26. Alshatti, D. A. *Examining Driver Risk Factors in Road Departure Conflicts Using SHRP2 Data*. Master's thesis. University of Dayton, OH, 2018.
27. McLaughlin, S. B., J. M. Hankey, C. Klauer, and T. A. Dingus. *Contributing Factors to Run-off-Road Crashes and Near-Crashes*. United States. National Highway Traffic Safety Administration, Washington, D.C., 2009.
28. Das, S. *Artificial Intelligence in Highway Safety*. CRC Press, Boca Raton, FL, 2022.
29. Hasan, A. S., M. A. B. Kabir, M. Jalayer, and S. Das. Severity Modeling of Work Zone Crashes in New Jersey Using Machine Learning Models. *Journal of Transportation Safety & Security*, Vol. 15, No. 6, 2023, pp. 604–635.
30. Koski, T., and J. M. Noble. *Bayesian Networks: An Introduction*. John Wiley & Sons, Ltd, Chichester, 2009.
31. Das, S., R. Chakraborty, and M. M. Mimi. Unraveling Crash Causation: A Deep Dive into Non-Motorists on Personal Conveyance. *Proc., International Conference on Transportation and Development (ICTD)*, Conference Proceeding Paper, 47-58, Atlanta, GA, American Society of Civil Engineers, Reston, VA, 2024.
32. Das, S., A. Hossain, S. Barua, S. Kavianpour, and A. Sheykhfard. Causal Insights into Speeding Crashes. *Proc., International Conference on Transportation and Development (ICTD)*, Conference Proceeding Paper, 348-359, Atlanta, GA, American Society of Civil Engineers, Reston, VA, 2024.
33. Sun, M., X. Sun, D. Shan, D. Armstrong, and S. Das. Louisiana Pedestrian Crash Analysis with Multinomial Logit Model and Bayesian Network. Presented at 98th Annual Meeting of the Transportation Research Board, Washington, D.C., 2 2019.
34. Das, S., V. Vierkant, J. Cruz Gonzalez, B. Kutela, and A. Sheykhfard. Bayesian Network for Motorcycle Crash

Severity Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 51–63.
35. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann, San Mateo, CA, 1988.
36. Nielsen, T. D., and F. V. Jensen. *Bayesian Networks and Decision Graphs*. Springer Science & Business Media, Berlin, Germany, 2009.
37. Norsys. Netica 6.04. 2020. https://www.norsys.com/netica.html. Accessed July 2022.
38. Lauritzen, S. L. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, Vol. 19, 1995, pp. 191–201.
39. Li, Y., and Y. Bai. Highway Work Zone Risk Factors and Their Impact on Crash Severity. *Journal of Transportation Engineering*, Vol. 135, No. 10, 2009, pp. 694–701.
40. Wang, Q. Study on Crash Characteristics and Injury Severity at Roadway Work Zones. M.S. Thesis. *Civil Engineering*, University of South Florida, Tampa, FL, 2009.
41. Koilada, K., A. S. Mane, and S. S. Pulugurtha. Odds of Work Zone Crash Occurrence and Getting Involved in Advance Warning, Transition, and Activity Areas by Injury Severity. *IATSS Research*, Vol. 44, No. 1, 2020, pp. 75–83.
42. Sun, M., X. Sun, and D. Shan. Pedestrian Crash Analysis with Latent Class Clustering Method. *Accident Analysis & Prevention*, Vol. 124, 2019, pp. 50–57. https://doi.org/10.1016/j.aap.2018.12.016.
43. Chang, F., P. Xu, H. Zhou, A. H. S. Chan, and H. Huang. Investigating Injury Severities of Motorcycle Riders: A Two-Step Method Integrating Latent Class Cluster Analysis and Random Parameters Logit Model. *Accident Analysis & Prevention*, Vol. 131, 2019, pp. 316–326. https://doi.org/10.1016/j.aap.2019.07.012.
44. Nilsson, D., M. Lindman, T. Victor, and M. Dozza. Definition of Run-off-Road Crash Clusters—For Safety Benefit Estimation and Driver Assistance Development. *Accident Analysis & Prevention*, Vol. 113, 2018, pp. 97–105. https://doi.org/10.1016/j.aap.2018.01.011.
45. Li, Z., C. Chen, Y. Ci, G. Zhang, Q. Wu, C. Liu, and Z. Qian. Examining Driver Injury Severity in Intersection-Related Crashes Using Cluster Analysis and Hierarchical Bayesian Models. *Accident Analysis & Prevention*, Vol. 120, 2018, pp. 139–151. https://doi.org/10.1016/j.aap.2018.08.009.
46. van de Velden, M., A. I. D'Enza, and F. Palumbo. Cluster Correspondence Analysis. *Psychometrika*, Vol. 82, No. 1, 2017, pp. 158–185. https://doi.org/10.1007/s11336-016-9514-0.
47. Feigenbaum, B., M. G. Fields, and S. Purnell. 24th Annual Highway Report. Reason Foundation, Los Angeles, CA, 2019.
48. Schrock, S. D., G. L. Ullman, A. S. Cothron, E. Kraus, and A. P. Voigt. *An Analysis of Fatal Work Zone Crashes in Texas*. Report FHW A/TX-05/0-4028, Vol. 1. Texas Transportation Institute, College Station, 2004.
49. Whitmire, I. J., J. F. Morgan, T. Oron-Gilad, and P. A. Hancock. The Effect of In-Vehicle Warning Systems on Speed Compliance in Work Zones. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 14, No. 5, 2011, pp. 331–340.
50. Debnath, A., R. Blackman, and N. Haworth. A Review of the Effectiveness of Speed Control Measures in Roadwork Zones. *Proc., 2012 Occupational Safety in Transport (OSIT) Conference*, CARRS-Q, Queensland University of Technology, Australia, 2012.