Research Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
2022, Vol. 2676(3) 371–384

DOI: 10.1177/03611981211051351
journals.sagepub.com/home/trr
SAGE

Prediction of Occurrence and Severity of Run-off-Roadway Crashes on Rural Two-Lane Roadways Using Bayesian Networks

Boniphace Kutela[1], Raul E. Avelar[1], Srinivas R. Geedipally[2], and Ankit Jhamb[3]

Abstract

Run-off-roadway (ROR) crashes are among the most common crash types on rural two-lane roadways. Current methodologies to predict their occurrence and severity by considering conditional nature and interactions between independent variables require complex mathematical procedures. This study employs Bayesian networks (BNs), a non-functional form graphical model, to determine factors associated with the occurrence and severity of ROR crashes. The study used five-year (2014–2018) crash data collected from 397 randomly selected road segments within Texas. Out of 397 segments, 279 did not experience ROR crashes. The first BN model used all 397 segments and explored factors associated with occurrences of ROR crashes. The second BN model used the remaining 118 segments that involved ROR crashes and focused on factors associated with different crash types (guardrail [GR], overturning [OT], and fixed object [FO] crashes) and their associated severity levels. Study results revealed that the presence of horizontal curves and utility poles within the clear zone on the road individually increased the chance of ROR crashes by about 35%. Moreover, FO crashes resulted in 36% more fatal and injury crashes than GR crashes, which showed the effectiveness of guardrails in reducing severity. This study also explored the combined influence of variables on ROR crash occurrence and severity, as well as the interrelation between several independent variables. The proposed methodology can be used to evaluate the effectiveness of countermeasures.

Keywords

infrastructure, highway maintenance, roadside maintenance operations, roadside safety, safety, transportation safety management systems, rural road safety

Run-off-Roadway (ROR) crashes usually involve a single vehicle leaving a travel-way and either hitting a fixed object or overturning. Previous studies showed that ROR crashes are more likely to be severe compared with other single-vehicle crashes as they account for up to 70% of all fatal single-vehicle crashes and over 40% of all fatal crashes in the United States (*1–3*). Many researchers also investigated several factors associated with ROR crashes using different sets of methodologies and datasets. These can be broadly divided into human-related, environmental-related, and roadway-related factors.

Human-related factors contributed greatly to the occurrence and severity of ROR crashes (*2*, *4–7*). A study by McLaughlin et al., which utilized descriptive analysis, found that ROR crashes and ROR near-crashes occurred more frequently when drivers were distracted (*4*). Another study by Liu and Subramanian evaluated factors associated with fatal single-vehicle ROR crashes using Fatality Analysis Reporting System data (*2*). Their study revealed that sleeping while driving, driving under the influence of alcohol, and speeding were the main contributing factors. Similarly, Roy and Dissanayake compared ROR and non-ROR crashes using the Bayesian

[1]Texas A&M Transportation Institute, Bryan, TX
[2]Texas A&M Transportation Institute, Arlington, TX
[3]Texas A&M University, College Station, TX

Corresponding Author:
Boniphace Kutela, b-kutela@tti.tamu.edu

statistical approach (*5*). They reported a greater likelihood of ROR crashes when drivers sleep while driving, have an illness or medical condition, and drive under the influence. Peng and Boyle found that distraction, inattention, speeding, seat belt usage, drowsiness, and fatigue were associated with the occurrence and severity of commercial vehicle ROR crashes (*6*). Additionally, Das and Sun's study, which applied multiple correspondence analysis, revealed that male drivers, older female drivers, and impaired drivers were more likely to be associated with fatal ROR crashes (*7*).

Many studies also evaluated the association of several environmental-related and roadway-related factors to ROR crash occurrences and severity. Most commonly cited factors for ROR crashes include horizontal curves, lanes and shoulder characteristics, side slopes, lighting conditions, and roadway surface conditions, among others (*2*, *8–12*). Liu and Subramanian applied logistic regression and reported that the probability of fatal ROR crashes increased in the presence of horizontal curves, rural highways, high posted speed limits, and adverse weather conditions (*2*). Al-Bdairi et al. used two separate mixed logit models to investigate the severity of ROR crashes involving large trucks at different lighting conditions (*8*). Results showed that a complex interaction between driver characteristics, traffic flow, and roadway geometric features explain differences in injury severity under light and dark conditions. Lee and Mannering used zero-inflated count models and nested logit models to assess the influence of roadside features on ROR crashes' frequency and severity (*9*). Their study found that more ROR crashes were attributable to wider shoulders, the number of isolated trees, and the steep side slope. The severity of the ROR crash was associated with horizontal curves and utility poles, among other factors (*10*).

As a result of previous research findings, different treatments were developed to reduce the occurrence and severity of ROR crashes. These treatments can be divided into three categories based on their intent: (1) keep vehicles on the road, (2) provide a safe recovery to vehicles, and (3) reduce crash severity. Pavement friction, markings, shoulder rumbles, and curve warning signs were a few among many strategies applied to warn drivers and keep vehicles on the road. These treatments significantly reduced ROR crash frequency (*13–16*). Other methods such as providing a clear zone on sides of roadway and widening of shoulders aided vehicles in recovering safely and avoiding severe accidents (*17*, *18*). Another set of treatments such as guardrails, barriers, and cushions can be employed to reduce crash severity. However, they might lead to an increment in crash frequency (*11*, *12*, *19*).

Past studies have investigated several factors associated with occurrences and severity of ROR crashes using either traditional regression methods (*2*, *8*), Bayesian approaches (*5*), or machine learning approaches (*7*). These methods assumed a direct relationship between predictors and response variables. However, there exists an interrelationship between multiple predictors and response variables. For instance, it is very common to find that locations with high annual average daily traffic (AADT) have wider lanes and paved shoulders (*20*). Also, previous methodologies did not consider conditional dependencies between safety countermeasures and other existing features. For instance, the same safety countermeasure might have different magnitudes of effects at different posted speed limits. Further, there are situations where two or more countermeasures, such as guardrails and chevrons, can exist at the same curved segment. To evaluate the individual effect of each countermeasure at this location requires conditioning of other features. Implementing conditional effectiveness using traditional regression approaches is not trivial.

This study applies Bayesian network (BN) to evaluate various factors associated with occurrences and severity of ROR crashes. The approach considers not only the conditional nature of the outcome of variables but also interactions between predictors. BN has been utilized by several researchers in traffic safety (*21–24*). Zou and Yue developed a BN model to compute the probability of an accident occurring in particular traffic conditions in Adelaide's Central Business District in South Australia and find the combination of factors that cause accidents (*22*). Hossain and Muromachi employed BN to build a model which can predict real-time crashes on basic freeway segments (*23*). The model predicts two-thirds of future crashes, with less than one-fifth of the false alarm rate, that can occur within the following 4–9 min over a 250 m section of road. Similarly, the BN model created by Deublein et al. correctly predicted the number of accidents at 86% of road segments on Swiss highways with 25% tolerance (*24*). The next section describes the methodology adopted, followed by the data description where probabilistic sampling of data is discussed. Results and discussions are then provided, followed by conclusions and future study direction.

Methodology

This study developed two separate BNs, one for ROR crash occurrence and the other for injury severity. The next sections provide an overview of BNs, individual and combined inference analysis, and variable impact analysis.

Overview of BNs

BN is a non-functional form graphical model that consists of arcs and nodes. Nodes represent random variables, usually described by probability tables, whose values portray probabilistic interdependence. Arcs represent the conditional association between random variables. The relationship between variables should be directional, meaning that no circle should be introduced. This condition is referred to as directed acyclic graphs. The tail and head of arcs depict the directional relationship between variables. Tail stands as the "parent" variable, while arc's head is the "child" variable (*25*, *26*). In most cases, parent and child nodes are interpreted to represent a causal relationship (*26*).

The use of BNs aims to obtain the posterior conditional probability distribution of possible unobserved antecedents given observed evidence, that is, P [Antecedent | evidence]. However, in practice, researchers observe evidence given antecedent, P [Evidence|Antecedent]. Therefore, BNs use Bayes' theorem to express the conditional probability distribution of unobserved antecedents, given observed evidence (*26*). Bayes' theorem for this case can be written as

$$P[Antecedent|Evidence] = P[Evidence|Antecedent].\frac{P[Antecedent]}{P[Evidence]} \quad (1)$$

The relationship between random variables can be expressed using the Markov property, which states that every random variable X_i directly depends on its parents Π_{Xi} (*27*). For discrete variables, such a relationship can be written as

$$P(X_1 \ldots\ldots\ldots X_n) = \prod_{i=1}^{n} P(X_i|\Pi_{Xi}) \quad (2)$$

Using this framework, a researcher can fit models of anticipated or latent causal relationships among variables in a cross-sectional dataset.

Schwarz introduced the Bayesian information criterion (BIC) to select a best-fitting model based on the maximum likelihood function (*28*). BIC is defined as

$$BIC(M) = -2l(\hat{\theta}) + plog(n)$$

where M is a statistical model, $l(\hat{\theta})$ is log-likelihood for model M, p is the number of parameters, n is the number of independent observations. A lower BIC value represents a better model (*28*, *29*).

BN's Structure and Parameter Learning

Performing BN analysis involves structure and parameter learning. Structure learning is the process of understanding the graphical connection between random variables. A researcher can learn BN structure using either an analytical approach (learn from data), expert knowledge, or a combination of both (*30*, *31*). Learning BN structure from data is appropriate if a researcher has no background knowledge about variables' relationships. The disadvantage of this approach is that BN structure can be affected by the data-generation process; thus, a resulting relationship can be meaningless. Conversely, the expert knowledge-based approach can result in a biased network against the actual relationship presented in data. In this study, the authors used a composite approach. BN structure was first learned using an analytical approach by applying the hill-climbing method; then, expert knowledge was applied to alter arrows' directions, insert, or remove arcs. Parameter learning involves obtaining conditional probability distribution between a\the parent node and the child node. Both structure and parameter learning used maximum likelihood estimation.

Predictive Inference and Variable Impact Analysis

Predictive inference and variable impact analysis can be performed for either individual treatment/scenario or combined treatments/scenarios. This section provides details for predictive inference analysis and variable impact analysis for both approaches.

Individual Inference Analysis. Individual inference analysis involved obtaining posterior conditional probability distribution of possible unobserved antecedents given evidence (*26*, *32–34*). For instance, given that the segment is curved, the predicted probability of crash occurrence can be found using individual inference analysis. This can be attained by assigning certainty to one parent variable (see Equation 3).

$$\begin{gathered} Predicted\ probability_i = P(Event = i|Evidence_x = 1.0) \\ P(Crashoccurence = i|Curve = 1.0) \end{gathered} \quad (3)$$

where i is the probability of crash occurrence, and x represents a hypothesis variable, for example, the presence of a curve.

Combined Inference Analysis. Contrary to individual inferences, combined inferences consider more than one condition linked to the target variable. Thus, evaluating the impact of individual treatment using traditional regression approaches requires more effort. For instance, given a horizontal curve segment is treated by guardrail and chevrons, the predicted probability of crash occurrence can be found using combined inferences. This can be attained by assigning certainty to more than one parent variable (see Equation 4).

$$\textit{Predicted probability}_i = P(\textit{Event} = i|\textit{Evidence}_1 = 1.0\,\&\,\textit{Evidence}_2 = 1.0\,\&\,\ldots\ldots\ldots\ldots\,\&\,\textit{Evidence}_n = 1.0)$$

$$P(\textit{Crashoccurence} = i|\textit{Curved} = 1.0\,\&\,\text{Guardrail} = 1.0) \tag{4}$$

where i is the probability of crash occurrence, and $Evidence_1$ to $Evidence_n$ represent hypothesis variables, for instance, presence of curves, guardrail treatment, side slope, and so forth.

Variable Impact Analysis. After obtaining predicted conditional probabilities, finding the impact of the variable of interest on either occurrence or severity of ROR crashes becomes important. This can be analyzed by considering changes in predicted probabilities per variable category. For instance, the difference in predicted probability of crashes for a segment with and without a horizontal curve can be found by using the sensitivity equation below (*34*, *35*).

$$E_x^{P_i} = P_i[Evidence_1] - P_i[Evidence_0] \tag{5}$$

where $E_x^{P_i}$ is direct sensitivity of variable x, $Evidence_0$ is set as base condition, while $Evidence_1$ represents the category of interest.

Further, using combined inference analysis and variable impact analysis, it is possible to determine the impact of variable(s) of interest conditioning on one or more variables. For instance, one can determine the impact of a guardrail's presence on crash severity for a location with a guardrail, steep side slope, and high AADT. The impact of an individual factor out of several factors under various combinations of scenarios can easily be determined. In contrast, handling multiple conditions in more traditional statistical approaches escalates quickly in complexity, involving linear algebra with estimates of coefficients and covariance matrix from estimation.

Data Description

This study used crash data, traffic data, and roadway characteristics data collected from rural two-lane roadways located in Texas. The authors used five-year (2014–2018) crash data collected from Crash Records Information System database, roadway, and traffic data from Geospatial Roadway Inventory Database database. In addition, street view and satellite photographs were used to obtain roadway characteristics data that were not available in databases. The study considered a total of 397 road segments that cover over 300 mi. Data from all 397 road segments were used to develop a BN model to predict factors associated with the crash occurrence. On the other hand, a reduced dataset of 118 road segments that experienced crashes was used for developing the BN model to predict factors associated with crash severity.

Study Sample

As collecting additional data for the entire State can be time-consuming, the authors developed a probabilistic sampling method to capture a sample representative of two-lane rural roadway segments in the entire State. The authors divided the State into four regions (north, west, south, and east). The cube sampling procedure was used 100 times with replacement for different sample sizes to obtain the right sample size. Results of resampling procedures showed that a random selection of 600 highway segments would be a good starting point to remain within the desired accuracy (within 10% of state-wide parameters for various safety metrics). The total length of road segments was 353 mi.

After completing sample collection, 137 of 600 roadway segments were discarded, thus reducing the sample size to 463 segments. The discarded samples were too close to intersections, road segments shorter than 100 ft, and segments with uneven cross-sections because of merging lanes, decreasing median width, and changing lanes' width. Further investigation into the remaining segments disclosed segments with missing information. More specifically, about 66 segments did not have proper roadway imagery, which resulted in the unavailability of key variables such as posted speed limit, lane width, and presence of roadside objects. Therefore, the final number of segments used in this study was 397. The average segment length is 1.23 mi with minimum length of 0.1 mi and maximum of 6.34 mi. A TxDOT report by Avelar et al. provides more details on the sample design, target precision, data collection, and summary statistics (*36*).

Descriptive Analysis

Among 397 segments, 279 segments experienced no ROR crashes, whereas 118 segments experienced a varying number of crashes. As the number of crashes per segment varies, each crash was recorded as a single observation to avoid underrepresentation. Thus, a total of 590 observations were available for analysis. A total of 15 variables (Table 1) were investigated to determine their association to the occurrence, types (guardrail [GR], overturning [OT], and fixed object [FO] crashes), and severity (property damage only and fatal injury crashes) of ROR crashes.

Table 1 summarizes the number and percentage distribution of 311 crashes that occurred and their severity in relation to property damage only (PDO), and fatal and

Table 1. Descriptive Statistics of Variables

Variable	No ROR crash		ROR crash		ROR crash severity			
	Count	Percent	Count	Percent	PDO count	PDO percent	FI count	FI percent
Guardrail								
None	212	55.8	168	44.2	94	56.0	74	44.0
Present	67	31.9	143	68.1	98	68.5	45	31.5
Curved section								
None	193	73.1	71	26.9	40	56.3	31	43.7
Present	86	26.4	240	73.6	152	63.3	88	36.7
VMT								
<1,000	246	78.6	67	21.4	39	58.2	28	41.8
≥1,000	33	11.9	244	88.1	153	62.7	91	37.3
AADT								
<1,000	147	65.6	77	34.4	45	58.4	32	41.6
≥1,000 to <3000	78	45.1	95	54.9	50	52.6	45	47.4
≥3000	54	28.0	139	72.0	97	69.8	42	30.2
Posted speed limit								
<60 mph	136	53.1	120	46.9	78	65.0	42	35.0
≥60 mph	143	42.8	191	57.2	114	59.7	77	40.3
Side slope								
≥4:1	97	55.7	77	44.3	51	66.2	26	33.8
>4:1	182	43.8	234	56.3	141	60.3	93	39.7
Lane width								
<12 ft	227	50.3	224	49.7	138	61.6	86	38.4
≥12 ft	52	37.4	87	62.6	54	62.1	33	37.9
Paved shoulder width								
None	73	57.5	54	42.5	28	51.9	26	48.1
1 ft to 6 ft	125	45.6	149	54.4	95	63.8	54	36.2
>6 ft	81	42.9	108	57.1	69	63.9	39	36.1
Poles								
None	147	68.4	68	31.6	36	52.9	32	47.1
Present	132	35.2	243	64.8	156	64.2	87	35.8
Lone trees								
None	227	56.6	174	43.4	104	59.8	70	40.2
Present	52	27.5	137	72.5	88	64.2	49	35.8
Clusters of trees								
None	243	52.3	222	47.7	140	63.1	82	36.9
Present	36	28.8	89	71.2	52	58.4	37	41.6
TCD								
None	235	57.0	177	43.0	100	56.5	77	43.5
Present	44	24.7	134	75.3	92	68.7	42	31.3
Minor intersections density (intersection/mile)								
None	169	71.9	66	28.1	41	62.1	25	37.9
1 to 2	36	15.2	201	84.8	124	61.7	77	38.3
>2	74	62.7	44	37.3	27	61.4	17	38.6
Driveways								
None	114	76.5	35	23.5	17	48.6	18	51.4
Present	165	37.4	276	62.6	175	63.4	101	36.6
Rumble strips								
None	260	49.1	269	50.9	171	63.6	98	36.4
Present	19	31.1	42	68.9	21	50.0	21	50.0

Note: AADT = annual average daily traffic; FI = fatal and injury; PDO = property damage only; ROR = run-off-roadway; TCD = traffic control devices (chevrons and delineators); VMT = vehicle miles traveled.

injury (FI) crashes for different categories in each variable. According to results in Table 1, most crashes occurred on segments with 1–2 intersections per mile (84.8%), higher AADT (88.1%), locations with traffic control devices (TCDs) which include chevrons and delineators (75.3%), and at curved locations (73.6%). Further, Table 1 presents the distribution of crashes and their associated severities for different categories of

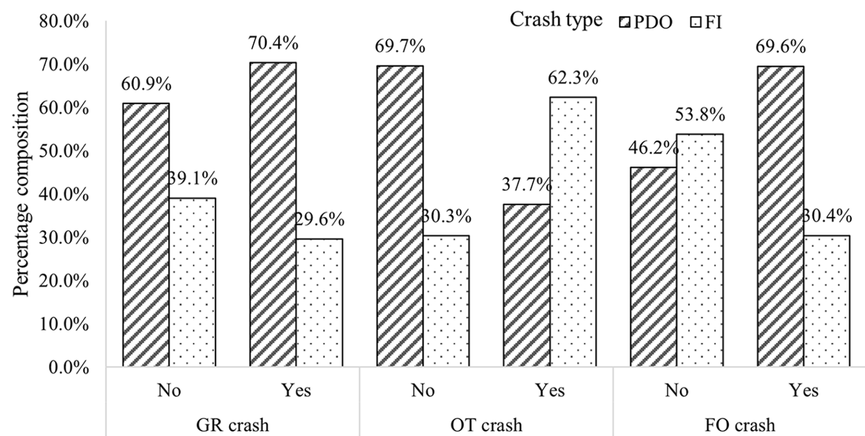

Figure 1. Crash severity statistics for different types of crashes.
Note: FI = fatal and injury; FO = fixed object; GR = guardrail; OT = overturning; PDO = property damage only.

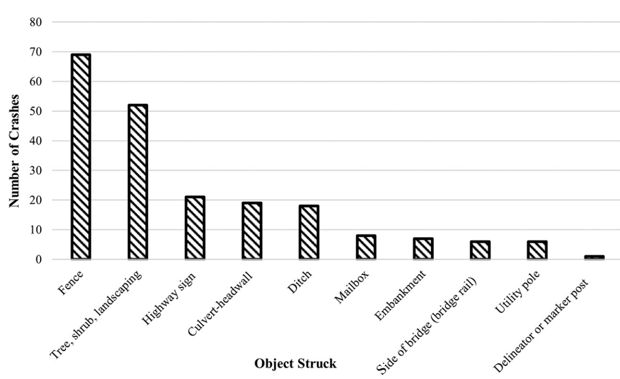

Figure 2. Composition of objects struck for fixed object crashes.

predictors, including driveways, poles, TCD, lone trees, guardrails, clusters of trees, and rumble strips.

Figure 1 presents the percentage composition of PDO, FI crash severity for three types of crashes, GR crashes, OT crashes, and FO crashes. Among crashes analyzed, none had two outcomes at once; for example, no crash involved overturning after hitting a fixed object. It can be observed that GR crashes and FO crashes have a higher percentage of PDO crashes, whereas OT crashes resulted in higher severity crashes (FI crashes).

Additionally, Figure 2 shows the composition of objects struck in FO crashes. It can be observed that most FO crashes involved either fence or trees/shrubs/landscaping features. Apart from that, a significant portion of FO crashes involved either highway signs, culvert headwalls, or ditches. Injury severity of crashes involving different objects may differ significantly.

Results and Discussions

This section is divided into two major parts; the first part presents a discussion on optimal BN structures and factors associated with ROR crashes. The second part focuses on parameters linked to crash types and severities.

The Optimal BN Structures

Figures 3 and 4 show optimal BN structures for crash occurrences and crash types and severities, respectively. Both structures were learned using the composite approach; that is, structures were first learned using data; then, expert knowledge was used to add, remove, and reverse links. For example, BN's initial structure for crash occurrence (Figure 3) had four variables: guardrail, minor intersections, vehicle miles traveled (VMT), and horizontal curves connected to ROR crash. These variables had a strong relationship with the occurrence of crashes. However, prior research showed an association between ROR crash occurrence and posted speed limit and side slope (*37*). Therefore, four other variables that had a relatively weaker relationship were connected to the ROR crash variable. The connections improved the BIC score from −5734 to −6889. The difference of −1155 BIC score is deemed significant (*38*). Furthermore, the out of sample prediction accuracy of the BN model in Figure 3 was 82.6% compared with 80.4% of a comparable logistic regression. A similar approach was used to attain optimal BN structure for crash type and severity (Figure 4).

According to Figure 3, eight variables, posted speed limit (PSL), driveway density, guardrail, curve density, minor intersection density, side slope, clusters of trees, and AADT directly associate with crash occurrences. On

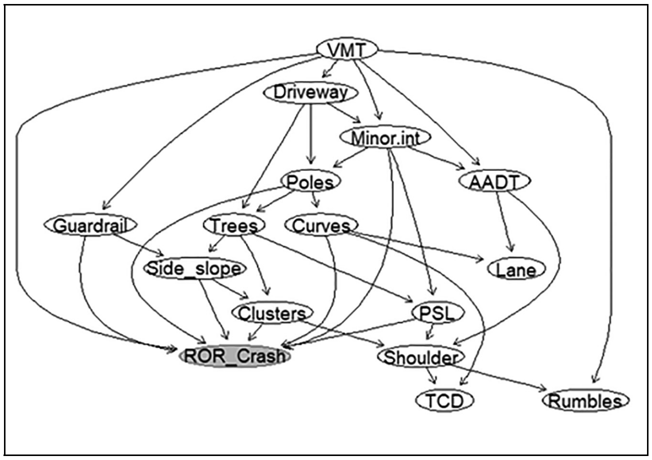

Figure 3. Optimal Bayesian network for ROR single-vehicle crash occurrence.
Note: AADT = annual average daily traffic; PSL = posted speed limit; ROR = run-off-roadway; TCD = traffic control device.

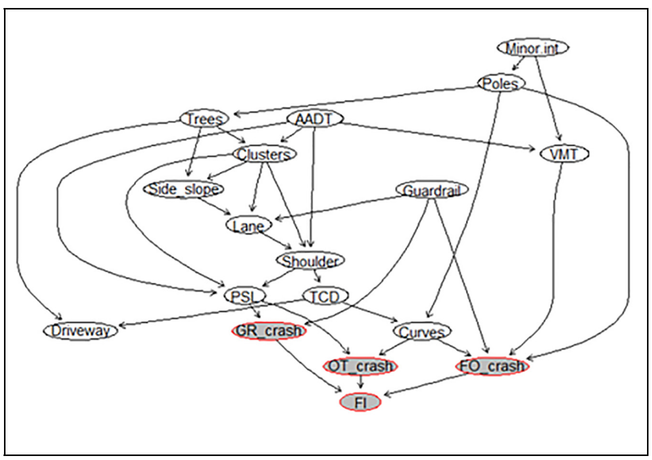

Figure 4. Optimal Bayesian network for crash type and severity.
Note: AADT = annual average daily traffic; FI = fatal and injury; FO = fixed object; GR = guardrail; OT = overturning; PSL = posted speed limit; ROR = run-off-roadway; TCD = traffic control device.

the other hand, Figure 4 shows the relationship between several predictor variables and crash types and severity. According to Figure 4, crash severity (FI) is associated with crash types only, and each crash type is linked to its predictor variables. FO crash type is connected to four predictor variables. On the other hand, both OT and GR crashes are associated with two variables. Again, the representation of the structure of relationships in Figure 3 requires multiple higher-order interaction terms in a more traditional modeling approach whose individual interpretation is limited, and that requires linear combinations of estimates and model information matrix. BN results can produce estimates directly from trees and estimates, as explained in the earlier section.

Figures 3 and 4 also show interrelationships between predictor variables. These interrelationships can represent either a safety relationship or merely a data-generation process. Most data-generation process interrelationships do not have a logical explanation. For example, Figure 3 shows that there exists an association between VMT, guardrail, and AADT. A possible explanation could be that guardrails are more likely to be installed on roadways with higher VMT, given that all other conditions are kept constant. The relationship between VMT and AADT is a direct relationship obtained during data processing, as VMT is a product of segment length and AADT. Another set of interrelationships of variables shown in Figure 3 is between roadside features, including lone trees, poles, guardrails, clusters of trees, and side slopes. It is very common to find a location with both steep side slopes and guardrails. Also, locations with numerous lone trees can also have clusters of trees or poles. Figure 4 shows several interrelationships that are likely to represent real-life scenarios at sites. For instance, locations with higher AADT likely have wider shoulders (*20*). Similarly, higher AADT locations can also have higher PSL. Moreover, TCDs such as chevrons and delineators are used where horizontal curves are present, and therefore they have a relationship.

Using optimal BNs, predicted conditional probabilities of crash occurrence, type, and severity are determined by querying evidence from variables directly linked to the outcome of interest. As only three crash types are presented in the network, severity comparison is performed across three crash types.

Crash Occurrences

Table 2 shows predictive probabilities and sensitivities for nodes that are directly linked to the ROR crash node. Sensitivity represents a change in probability of crash occurrence with respect to change in conditions of the variable in question. These variables are PSL, VMT, GR, curved section, minor intersections, side slope, poles per mile, and the cluster of trees. The next section discusses the influence of each variable on ROR crash occurrence.

Individual Variable Impact on Crash Occurrence. According to predictive probabilities results (Table 2), the presence of a guardrail is associated with increased ROR crash frequency. The probability of observing a ROR crash is higher when there is a guardrail (0.663) than when there is no guardrail (0.400), making a difference of 0.263. A similar conclusion was drawn by one of the previous studies (*11*) using logistic regression. This does not necessarily mean that guardrails increase crash occurrence, but they are generally installed at high-risk locations for crashes to reduce their severity. This phenomenon is

Table 2. Crash Occurrence Prediction Inferences

Variable and observed evidence	Predicted probability	Sensitivity
Guardrail		
No guardrail	0.400	—
Guardrail present	0.663	0.263
Curved section		
No	0.277	—
Yes	0.632	0.354
Vehicle miles traveled		
<1,000	0.218	—
≥1,000	0.815	0.596
Posted speed limit		
< 60	0.389	—
≥60	0.538	0.149
Side slope		
≤1:4	0.465	—
>1:4	0.498	0.033
Poles per mile		
No pole	0.256	—
Poles present	0.605	0.348
Cluster of trees		
No clusters	0.456	—
Clusters present	0.563	0.107
Minor intersections		
No intersection	0.220	—
1 or more intersection/mile	0.808	0.588

Note: — = Base category.

called endogeneity, where a variable is capturing the effect of an unobserved variable; for more details, the reader is referred to Kim and Washington (*39*).

The results in Table 2 show that curved segments with at least one minor intersection per mile increase the chance of ROR crashes by 35.4% and 58.8%, respectively. Additionally, higher PSL and VMT road segments are associated with higher probabilities of crash occurrences. Locations with 60 mph or higher PSL have about a 14.9% increased chance of ROR crashes. For locations with 1,000 or more VMT, the chance of ROR crashes increases by about 60% compared with locations with VMT less than 1,000. Further, side slopes are associated with an increased likelihood of ROR crashes; however, the magnitude of difference in predicted crash occurrences is relatively small. The probability of ROR crash for locations with side slopes greater than 4:1 is about 3.3% more than that of flatter side slopes (less or equal to 4:1). The presence of utility poles and clusters of trees within the clear zone also increases the chance of ROR crashes by about 35% and 10.7%, respectively.

Impact of Combined Scenarios on Crash Occurrence. Given the number of variables in this study, various combinations of scenarios are possible. However, it is difficult to present the results of all combinations. Further, some combinations of scenarios might not be practical. Because it is known that VMT and PSL are key players in ROR crashes, these variables were selected to construct combined scenarios. Therefore, the impact of guardrails, curves, poles, and clusters of trees is evaluated under varying VMT and PSL conditions. These scenarios are chosen for illustration purposes because they are realistic and relevant. Results show the predicted probability of combined factors and sensitivity values of guardrails, curves, poles, and tree clusters. Sensitivity values represent predicted probabilities of individual factors conditioned on other factors. The sensitivity values were computed for the presence of a guardrail, curves, poles, and clusters against their absence. Thus, the base categories constitute of the absence of the object of interest. In general, predicted probabilities of variables in combined scenarios either increased or decreased compared with individual variables predictions. This is because of the multiplicative nature of the combined impact.

The results in Table 3 show that given the presence of guardrails, the chance of ROR crash is highest (0.947) if PSL is above 60 mph and VMT is greater than 1,000. On the other hand, the chance of ROR crash is lowest (0.177) if PSL is above 60 mph, VMT is less than 1,000, and there is no guardrail. Sensitivity values revealed that the presence of guardrails is associated with a 7.4% to 21.6% increased chance of ROR crashes. Magnitudes of increased chance of ROR crashes with presence of guardrails for combined scenarios are smaller than estimates of individual guardrail variable impact (26.3%) in Table 2.

Also, with respect to horizontal curves, the highest chance of ROR crash (0.931) is observed on horizontal curves where PSL is above 60 mph, and VMT is greater than 1,000. Conversely, the lowest chance of ROR crash (0.102) is observed on straight segments with no curves where PSL is above 60 mph, and VMT is less than 1,000. Sensitivity analysis revealed that the presence of a curve is associated with an increased chance of ROR crashes of between 11.5% and 55.5%. The largest magnitude of increased chance of ROR crashes associated with horizontal curves (55.5%) for combined scenarios is greater than that of individual variable impact analysis (35.4%) obtained from Table 2.

Additionally, the influence of poles and tree clusters on ROR crashes is slightly different when combined scenarios and individual variables analysis are compared. For instance, the largest sensitivity values for poles and clusters of trees are 23.1% and 10.2%, which are smaller than that obtained in individual variable impact (34.8%) and (10.7%), respectively, in Table 2.

The results in Table 3 show that it is important to consider combined scenarios to evaluate the variable's impact.

Table 3. Combined Scenarios for Crash Occurrence Prediction Inferences

		Predictions	
ROR crash occurrence	PSL and VMT	Probability	Sensitivity
Clusters			
No	<60 mph		
	<1,000	0.192	Base for PSL<60 & VMT<1,000
	≥1,000	0.676	Base for PSL<60 & VMT≥1,000
	≥60 mph		
	<1,000	0.177	Base for PSL≥60 & VMT<1,000
	≥1,000	0.843	Base for PSL≥60 & VMT≥1,000
Yes	<60 mph		
	<1,000	0.267	0.074
	≥1,000	0.735	0.059
	≥60 mph		
	<1,000	0.393	0.216
	≥1,000	0.947	0.104
Guardrail			
No	<60 mph		
	<1,000	0.136	Base for PSL<60 & VMT<1,000
	≥1,000	0.297	Base for PSL<60 & VMT≥1,000
	≥60 mph		
	<1,000	0.102	Base for PSL≥60 & VMT<1,000
	≥1,000	0.815	Base for PSL≥60 & VMT≥1,000
Yes	<60 mph		
	<1,000	0.280	0.143
	≥1,000	0.853	0.555
	≥60 mph		
	<1,000	0.364	0.262
	≥1,000	0.931	0.115
Curves			
No poles	<60 mph		
	<1,000	0.233	Base for PSL<60 & VMT<1,000
	≥1,000	0.526	Base for PSL<60 & VMT≥1,000
	≥60 mph		
	<1,000	0.110	Base for PSL≥60 & VMT<1,000
	≥1,000	0.762	Base for PSL≥60 & VMT≥1,000
Poles present	<60 mph		
	<1,000	0.178	–0.055
	≥1,000	0.719	0.193
	≥60 mph		
	<1,000	0.341	0.231
	≥1,000	0.902	0.140
Poles			
No clusters	<60 mph		
	<1,000	0.220	Base for PSL<60 & VMT<1,000
	≥1,000	0.677	Base for PSL<60 & VMT≥1,000
	≥60 mph		
	<1,000	0.212	Base for PSL≥60 & VMT<1,000
	≥1,000	0.861	Base for PSL≥60 & VMT≥1,000
Clusters present	<60 mph		
	<1,000	0.174	–0.045
	≥1,000	0.752	0.075
	≥60 mph		
	<1,000	0.243	0.031
	≥1,000	0.963	0.102

Note: PSL = posted speed limit; ROR = run-off-roadway; VMT = vehicle miles traveled.

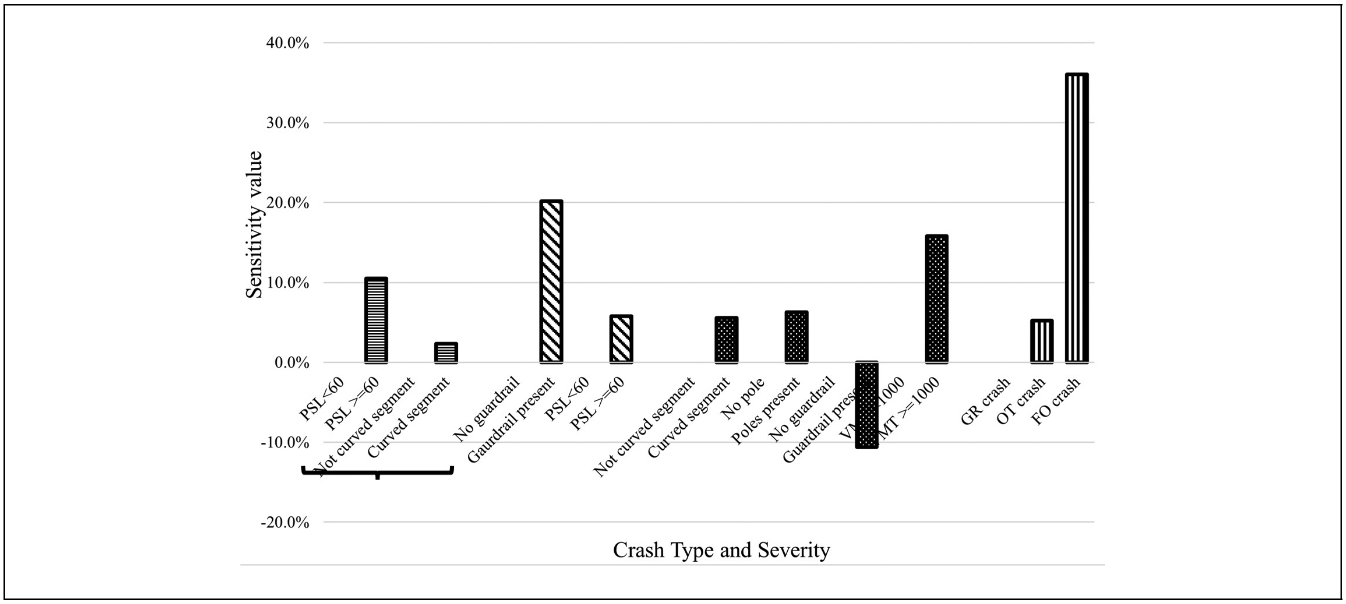

Figure 5. Individual variables sensitivity values for crash type and severity.
Note: FO = fxed object; GR = guardrail; OT = overturning; PSL = posred speed limit; VMT = vehicle miles traveled.

Table 4. Crash Type and Severity Prediction Inferences

Variable and observed evidence	Predicted probability
Overturning (OT) crash type	
Posted speed limit	
<60 mph	0.179
≥60 mph	0.284
Curved section	
No	0.223
Yes	0.246
Guardrail (GR) crash type	
Guardrail	
No guardrail	0.000
Guardrail present	0.202
Posted speed limit	
<60 mph	0.056
≥60 mph	0.114
Fixed object (FO) crash type	
Curved section	
No	0.621
Yes	0.677
Poles	
No pole	0.613
Poles present	0.676
Guardrail	
No guardrail	0.702
Guardrail present	0.596
Vehicle miles traveled	
<1,000	0.584
≥1,000	0.741
Crash severity (fatal or injury)	
GR crash	0.258
OT crash	0.310
FO crash	0.618

Variations of the impact of a variable using the combined scenarios approach provide more information to policy-makers and practitioners at a minimal computation cost than that obtained from individual variable impact.

Crash Type and Severity

This section presents the individual variable impact on crash type and severity. The section is divided into two major parts: crash type and crash severity.

Individual Variable Impact on Crash Type and Severity. Table 4 and Figure 5 present inferences of predicted probabilities and sensitivity values, respectively, for three types of ROR crashes: OT crashes, GR crashes, and FO crashes. OT crashes are directly linked to two variables, which are PSL and VMT. GR crashes are associated with the presence of guardrails and PSL. On the other hand, FO crashes are associated with four variables, which are VMT, poles, horizontal curve, and guardrail.

The higher PSL (60 mph and above) is associated with about 10.5% increased probability of OT crashes. The likelihood of OT crashes increases by about 2.3% at curved segments compared with straight segments (Figure 5). This result is similar to findings from one of the previous studies (*2*), which reported an increase in ROR crashes on curved sections.

The presence of the guardrail is associated with about a 20.2% chance of occurrence of GR crashes. This is

Table 5. Combined Scenarios for Crash Type and Severity Prediction Inferences

Variables	Probability
Guardrail (GR) crash type	
Guardrail and posted speed limit	
No	
< 60 mph	0.000
≥60 mph	0.000
Yes	
< 60 mph	0.122
≥60 mph	0.237
Overturning (OT) crash type	
Curves and Vehicle miles traveled (VMT)	
No	
<1,000 miles	0.204
≥1,000 miles	0.208
Yes	
<1,000 miles	0.254
≥1,000 miles	0.244
Fixed object (FO) crash type	
Curves and Vehicle miles traveled (VMT)	
No	
<1,000 miles	0.231
≥1,000 miles	0.224
Yes	
<1,000 miles	0.245
≥1,000 miles	0.260

Note: PSL = posted speed limit.

revealed by the probability of GR crashes for locations with guardrails (0.202). Further, high PSL (60 mph and above) is associated with about a 5.8% increase in GR crashes compared with PSL of up to 60 mph for locations with guardrails (Figure 5).

The guardrail has a relatively large magnitude of effects on FO crashes. The presence of the guardrail is associated with a decrease of FO crashes by about 10.6% (Figure 5). Also, results in Table 4 show that locations with VMT of 1,000 or more have a 74.1% chance of FO crashes, whereas locations with VMT less than 1,000 have a 58.4% chance of FO crashes. Thus, 1,000 or more VMT is associated with an increase in FO crashes by about 16% (Figure 5). Curved sections are associated with an increase of about 5.6% in FO crashes. Lastly, presence utility poles is associated with an increased chance of FO crashes by about 6.3% (Figure 5).

Table 4 also presents predicted probabilities of crash severities. The severity of crashes was grouped into two levels: FI crashes and PDO crashes. Results in Table 4 show that the chance of FI crashes increases for OT crashes but decreases for either GR or FO crashes. The chance that a crash will be either fatal or injury increases by 5.2% when it is an OT crash (0.643) compared with a GR crash. Further, the chance of FI crash severity increases by about 36% for FO crash compared with GR crash. In other words, the findings imply that the presence of the guardrail reduces the chance of severe crashes by about 36%.

Impact of Combined Scenarios on Crash Types and Severity. In this section, different combined scenarios associated with crash occurrences are evaluated. For GR and OT crashes, two variables that are directly connected to them are evaluated together. On the other hand, selected combinations of variables were performed for FO crashes. Table 5 presents prediction inferences for combined scenarios for GR crashes, OT crashes, and FO crashes.

The probability value represents the chance of either GR crash, OT crash, or FO crash to occur given combined scenarios. Results in Table 5 show that GR crashes are higher with the guardrail, and when the roadway has PSL of 60 mph or above. As it is impossible to observe GR crashes without guardrails, GR crashes' sensitivity for guardrails is the same as the probability of observing GR crashes (Table 5 and Figure 6). On the other hand, results show that curved segments with similar VMT can experience an increase of 3.6% to 5% OT crashes compared with straight segments. Lastly, given segments with similar VMTs, curved locations are associated with an increase of about 1.4% to 3.7% of FO crashes.

A comparison of sensitivity analysis values for the same variable for individual impact analysis (Figure 5) and combined scenarios (Figure 6) provide more information on the impact of the variable on ROR crash occurrences, crash type, and severity. For instance, whereas individual impact analysis shows that guardrail is associated with a 20.2% increase in GR crashes, such increase is between 12.2% and 23.7% when combined scenario analysis with PSL is used (Figure 6). Furthermore, horizontal curves are associated with a 2.3% increase of OT crashes using individual variable analysis, whereas for combined scenarios with VMT, the chance of OT crashes varies between 3.6% and 5%. Lastly, although the impact of horizontal curves on FO crashes varies between 1.4% and 3.7% under combined scenarios with VMT, individual impact analysis shows a relatively larger value (5.6%). These observations underscore the importance of conditional analysis when evaluating the influence of roadway treatments.

Conclusions and Future Works

ROR crashes are common on two-lane roadways and can be severe. Several past studies explored factors associated with ROR crashes. Cited factors are human related, such as sleeping while driving, driving under the influence of alcohol, and speeding. Other factors are environmental related, such as rainy/wet pavement, snow conditions, and roadway-related factors, such as horizontal curves, grades, narrow lanes, and shoulders. However,

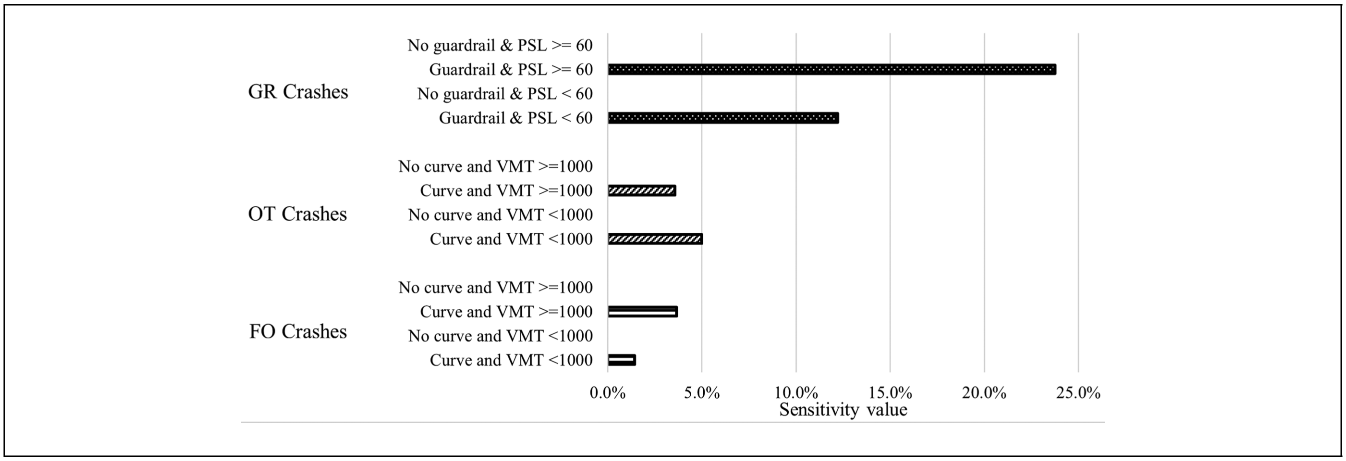

Figure 6. Combined variables sensitivity values for crash type.
Note: FO = fixed object; GR = guardrail; OT = overturning; PSL = posted speed limit; VMT = vehicle miles traveled.

methodologies used in previous studies need more computational effort to explain the conditional nature and interactions between predictor variables. Also, it is very common to find road segments having more than one treatment at the same location. Evaluating the influence of one treatment on ROR crashes using traditional regression methods requires significant efforts.

This study introduces BNs, a graphical model having a non-functional form, to determine factors associated with the occurrence, types, and severity of ROR crashes. It uses data collected from two-lane roadways located at various locations throughout the State of Texas. A total of 600 segments of roadways were systematically selected using a probability sampling procedure to represent the entire State of Texas. A total of 397 segments were used for analysis. ROR crashes were classified as GR crashes, OT crashes, and FO crashes. To perform the BN analysis, both BN structures and parameters were learned using data and expert knowledge.

The study found that sections with horizontal curves and utility poles within clear zone limit increase the chance of ROR crashes by about 35%. Moreover, the presence of guardrails increases the chance of GR crashes by about 20% but decreases the chance of FO crashes by 11%. The severity of the crash decreases by about 36% when it is GR crash compared with FO crash. The combined influence of variables revealed a noticeable variation of variable's impact when individual impact analysis and combined scenarios analysis are used. Further, the interrelation between several predictors and the influence of other roadside objects on ROR crash occurrence and severity was revealed.

This paper also makes BN's case over traditional regression approaches in its relative simplicity of estimating the effects of interrelated variables. Whereas traditional regression requires multiple interaction terms and significant mathematical statistics applications to obtain marginal estimates, BN allows for expedited calculations and graphic representation of relationships.

The approach in this study can be used by practitioners and researchers when evaluating impacts of various treatments. Evaluations of individual variable impacts are appropriate when prevailing site conditions such as PSL, VMT, and so forth, are similar. On the other hand, a combined scenarios approach should be preferred when evaluating the impact of treatment across uneven prevailing site conditions. More specifically, this methodology is more appropriate where two or more countermeasures exist. For instance, it is common to have chevrons, delineators, and reduced speed limit (advisory speed) on the curve. In such conditions, it becomes difficult to evaluate the influence of individual countermeasures; thus, the need for conditional assessment arises. Assessment of the effectiveness of treatments can be done by evaluating the increase/decrease in odds of ROR crashes or any other type of crashes conditioning on the presence/absence of treatment. Changes in odds of occurrence of the crash because of the presence of a countermeasure can be considered as the impact of the installed countermeasure.

This study was able to show the ability of BNs to capture the effectiveness of roadway treatment under various conditional situations. For instance, it has been shown that guardrails are associated with a decline of FO crashes but increase of GR-related crashes. However, such an increase in GR crashes is not uniform for all conditions. That is, the increase in GR crashes is higher for locations with PSL of 60 mph or higher, as compared with locations with less than 60 mph PSL. Therefore, if a study considers the presence of guardrail only without conditioning on other factors, the results might be misleading as they can overestimate or underestimate the influence of guardrail in GR crashes.

Although the presented methodology offers more flexibility in analyzing ROR crashes and potentially suggesting causal relationships, it does not offer several metrics available in traditional regression that would be desirable to ease the interpretation of results. For instance, the BN approach does not offer a statistical test of coefficient that would enable a researcher to determine the confidence level typically expected to recommend treatment. However, it offers a graphical way to represent complex interactions and a simpler computational approach to explore such interactions. Moreover, future work should show the prediction capability of the proposed approach compared with the traditional approach using a dataset different from data used in modeling. Further, as ROR crashes are likely associated with human behaviors, future research may include person-level variables.

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.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Boniphace Kutela and Raul Avelar; data collection: Raul Avelar, Ankit Jhamb; analysis and interpretation of results: Boniphace Kutela, Ankit Jhamb, Raul Avelar, and Srinivas Geedipally; draft manuscript preparation: Boniphace Kutela, Ankit Jhamb, Srinivas Geedipally. All authors reviewed the results and approved the final version of the manuscript.

Funding

The preliminary work and database development were funded by TxDOT project 0-6991.

ORCID iDs

Boniphace Kutela https://orcid.org/0000-0002-5450-1623
Raul E. Avelar https://orcid.org/0000-0002-3962-1758
Srinivas R. Geedipally https://orcid.org/0000-0002-6842-8172
Ankit Jhamb https://orcid.org/0000-0002-6123-9352

References

1. Liu, X., B. J. Schroeder, T. Thomson, Y. Wang, N. M. Rouphail, and Y. Yin. Analysis of Operational Interactions Between Freeway Managed Lanes and Parallel, General Purpose Lanes. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2262: 62–73.
2. Liu, C., and R. Subramanian. *Factors Related to Fatal Single-Vehicle Run-Off-Road Crashes*. Report DOT HS 811 232. URC Enterprises, Inc., Olney, MD, 2009.
3. Siddiqui, S. *Run-Off-Road (ROR) Crashes: A Description of Conventional and Contemporary Countermeasures*. Technical Report. Montana State University, Bozeman, 2015.
4. McLaughlin, S. B., J. M. Hankey, S. G. Klauer, and T. A. Dingus. *Contributing Factors to Run-Off-Road Crashes and Near-Crashes*. Report DOT HS 811 079. Virginia Tech Transportation Institute, Blacksburg, 2009.
5. Roy, U., and S. Dissanayake. Comparison of Factors Associated with Run-Off-Road and Non-Run-Off Road Crashes in Kansas. *Journal of the Transportation Research Forum*, Vol. 50, No. 2, 2011, pp. 69–86.
6. Peng, Y., and L. N. Boyle. Commercial Driver Factors in Run-off-Road Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2012. 2281: 128–132.
7. Das, S., and X. Sun. Association Knowledge for Fatal Run-off-Road Crashes by Multiple Correspondence Analysis. *IATSS Research*, Vol. 39, No. 2, 2016, pp. 146–155. https://doi.org/10.1016/j.iatssr.2015.07.001.
8. Al-Bdairi, N. S. S., S. Hernandez, and J. Anderson. Contributing Factors to Run-Off-Road Crashes Involving Large Trucks under Lighted and Dark Conditions. *Journal of Transportation Engineering Part A Systems*, Vol. 144, No. 1, 2018, pp. 1–9. https://doi.org/10.1061/JTEPBS.0000104.
9. Lee, J., and F. Mannering. Impact of Roadside Features on the Frequency and Severity of Run-off-Roadway Accidents: An Empirical Analysis. *Accident Analysis and Prevention*, Vol. 34, No. 2, 2002, pp. 149–161. https://doi.org/10.1016/S0001-4575(01)00009-4.
10. Lord, D., M. A. Brewer, K. Fitzpatrick, S. R. Geedipally, and Y. Peng. *Analysis of Roadway Departure Crashes on Two-Lane Rural Roads in Texas*. Report FHWA/TX-11/0-6031-1. Texas Transportation Institute, College Station, 2011.
11. Li, N., B. Park, and J. H. Lambert. Effect of Guardrail on Reducing Fatal and Severe Injuries on Freeways: Real-World Crash Data Analysis and Performance Assessment. *Journal of Transportation Safety & Security*, Vol. 10, No. 5, 2018, pp. 455–470. https://doi.org/10.1080/19439962.2017.1297970.
12. Zou, Y., A. P. Tarko, E. Chen, and M. A. Romero. Effectiveness of Cable Barriers, Guardrails, and Concrete Barrier Walls in Reducing the Risk of Injury. *Accident Analysis and Prevention*, Vol. 72, 2014, pp. 55–65. https://doi.org/10.1016/j.aap.2014.06.013.
13. Brimley, B. K., M. Saito, and G. G. Schultz. Calibration of Highway Safety Manual Safety Performance Function. *Transportation Research Record: Journal of the Transportation Research Board*, 2012. 2279: 82–89.
14. Gan, A., J. Shen, and A. Rodriguez. *Update of Florida Crash Reduction Factors and Countermeasures to Improve the Development of District Safety Improvement Projects*. Report FDOT 99700-3596-119. Lehman Center for Transportation Research, Miami, FL, 2005.
15. Perrillo, K. *The Effectiveness and Use of Continuous Shoulder Rumble Strips*. Albany, New York, 1998.
16. Khan, M., A. Abdel-Rahim, and C. J. Williams. Potential Crash Reduction Benefits of Shoulder Rumble Strips in

Two-Lane Rural Highways. *Accident Analysis and Prevention*, Vol. 75, 2015, pp. 35–42. https://doi.org/10.1016/j.aap.2014.11.007.

17. Ornek, M., and A. Drakopoulos. Analysis of Run-Off-Road Crashes in Relation to Roadway Features and Driver Behavior. *Proc., 2007 Mid-Continent Transportation Research Symposium*, Ames, IA, 2007.
18. Dixon, K., K. Fitzpatrick, R. Avelar, and S. Das. *Analysis of the Shoulder Widening Need on the State Highway System*. College Station, TX, 2017.
19. Elvik, R. Meta-Analysis of Evaluations of Public Lighting as Accident Countermeasure. *Transportation Research Record, National Research Council*, No. 1485, 1995, pp. 112–123.
20. Geedipally, S., M. Ko, and W. Lingtao. *A Systemic Approach to Project Selection for Highway Widening*. Texas Department of Transportation, 2016.
21. Kitali, A. E., S. M. Asce, E. Kidando, M. Asce, Boniphace Kutela, C. Kadeha, P. Alluri, and T. Sando. Safety Evaluation of High-Occupancy Toll Facilities Using Bayesian Networks. *Journal of Transportation Engineering, Part A: Systems*, Vol. 147, No. 5, 2021. https://doi.org/10.1061/JTEPBS.0000523.
22. Zou, X., and W. L. Yue. A Bayesian Network Approach to Causation Analysis of Road Accidents Using Netica. *Journal of Advanced Transportation*, Vol. 2017, 2017. https://doi.org/10.1155/2017/2525481.
23. Hossain, M., and Y. Muromachi. A Bayesian Network Based Framework for Real-Time Crash Prediction on the Basic Freeway Segments of Urban Expressways. *Accident Analysis & Prevention*, Vol. 45, 2012, pp. 373–381. https://doi.org/10.1016/J.AAP.2011.08.004.
24. Deublein, M., M. Schubert, Adey, B. T., and B. García de Soto. A Bayesian Network Model to Predict Accidents on Swiss Highways. *Infrastructure Asset Management*, Vol. 2, No. 4, 2015, pp. 145–158. https://doi.org/10.1680/JINAM.15.00008.
25. Markov, Z., and I. Russell. *Probabilistic Reasoning with Naïve Bayes and Bayesian Networks*. New Britain, 2007.
26. Horný, M. *Bayesian Networks*. Report 5. Boston Univeristy School of Public Health, MA, 2014.
27. Korb, K. B., and A. E. Nicholson. *Bayesian Artificial Intelligence*. Chapman & Hall/CRC, Boca Raton, FL, 2004.
28. Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics*, Vol. 6, No. 2, 1978, pp. 461–464. https://doi.org/10.2307/2958889.
29. Wit, E., E. van den Heuvel, and J.-W. Romeijn. 'All Models Are Wrong...': An Introduction to Model Uncertainty. *Statistica Neerlandica*, Vol. 66, No. 3, 2012, pp. 217–236. https://doi.org/10.1111/J.1467-9574.2012.00530.X.
30. Demiroluk, S., and K. Ozbay. Adaptive Learning in Bayesian Networks for Incident Duration Prediction. *Transportation Research Record: Journal of the Transportation Research Board*, 2014. 2460: 77–85.
31. Kim, J., and G. Wang. Diagnosis and Prediction of Traffic Congestion on Urban Road Networks Using Bayesian Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2595: 108–118.
32. Xie, C., and S. Waller. Estimation and Application of a Bayesian Network Model for Discrete Travel Choice Analysis. *Transportation Letters*, Vol. 2, No. 2, 2010, pp. 125–144. https://doi.org/10.3328/TL.2010.02.02.125-144.
33. Kidando, E., R. Moses, T. Sando, and E. E. Ozguven. Probabilistic Inference and Prediction of Travel Time Reliability: Empirical Analysis of Associated Factors. Presented at 97th Annual Meeting of the Transportation Research Board, Washington, D.C., 2018.
34. Kutela, B., and H. Teng. Prediction of Drivers and Pedestrians' Behaviors at Signalized Mid-Block Danish Offset Crosswalks Using Bayesian Networks. *Journal of Safety Research*, Vol. 69, 2019, pp. 75–83. https://doi.org/10.1016/j.jsr.2019.02.008.
35. Cong, H., C. Chen, P. S. Lin, G. Zhang, J. Milton, and Y. Zhi. Traffic Incident Duration Estimation Based on a Dual-Learning Bayesian Network Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672: 196–209.
36. Avelar, R., S. Geedipally, S. Das, L. Wu, B. Kutela, D. Lord, and I. Tsapakis. *Evaluation of Roadside Treatments to Mitigate Roadway Departure Crashes*. Report FHWA/TX-20/0-6991-R1. Texas A&M Transportation Institute, College Station, 2020.
37. Smith, E. B., and J. N. Ivan. Evaluation of Safety Benefits and Potential Crash Migration Due to Shoulder Rumble Strip Installation on Connecticut Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2005. 1908: 104–113.
38. Raftery, A. E. Bayesian Model Selection in Social Research. *Sociological Methodology*, Vol. 25, 1995, pp. 111–163. https://doi.org/10.2307/271063.
39. Kim, D. G., and S. Washington. The Significance of Endogeneity Problems in Crash Models: An Examination of Left-Turn Lanes in Intersection Crash Models. *Accident Analysis and Prevention*, Vol. 38, No. 6, 2006, pp. 1094–1100. https://doi.org/10.1016/j.aap.2006.04.017.