Research Article

Evaluating the Effect of Roadway and Development Factors on the Rural Road Safety Risk Index

Shahriar Afandizadeh and Shahab Hassanpour

Civil Engineering Department, Iran University Science and Technology, Tehran, Iran

Correspondence should be addressed to Shahriar Afandizadeh; zargari@iust.ac.ir

Received 8 September 2019; Revised 29 June 2020; Accepted 14 July 2020; Published 5 August 2020

Academic Editor: Heap-Yih (John) Chong

Copyright © 2020 Shahriar Afandizadeh and Shahab Hassanpour. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As roadway and development factors are identified as the most effective factors contributing to road traffic accidents, investigating these factors could lead to reducing the accident frequency rate. However, previous works focused on investigating the effect of roadway factors on the accident frequency rate using statistical analysis. The present study aimed to evaluate the effect of roadway and development factors on the accident frequency rate using ANOVA and Chi-square tests on a rural road. Secondly, it aimed to develop a rural road safety risk index based on K-means clustering and Gaussian models. The findings indicated that the operating speed and the differences between posted speed limits and the operating speed are the pivotal influencing factors on the accident frequency rate. Moreover, clustering analysis of the roadway and development factors on the two-lane, two-way road of Borujerd-Khorramabad indicated six clusters which were identified as highly, relatively highly, moderately, relatively lowly, lowly risky, and not risky (safe) clusters. Regarding clusters, the accident frequency rate increased by decreasing the difference between the posted speed limits and the operating speed from the safe cluster. In addition, the risky index model based on the Gaussian model showed that the average reducing factor of accident frequency rate reached 0.99 by increasing per km/hr in the difference between the posted speed limits and the operating speed among low risky and safe clusters, while it was equal to 1.17 in risky and unsafe clusters. The comparison of the clusters revealed that accident occurrence probability in risky clusters was more than the ones in low risky or safe clusters. Therefore, the maximum and minimum values of the safety risk index were observed in the sixth and the third clusters, respectively.

1. Introduction

Road traffic accidents cost most countries 3% of their gross domestic product [1], and traffic safety has become one of the most challenging issues in the recent decades. According to the report by the WHO, 1.52 million people are killed in traffic accidents every year [2]. Particularly, the cost of road fatalities and injuries is 2.19% of the gross national product in Iran, which is higher than the global average [3]. Location is considered as a crucial parameter in crash analyses since it closely relies on the identification of the traffic and geometric conditions that are related to an accident [4]. Anderson and Krammes [5] indicated that curves with a degree of curvature greater than four had higher accident rates. These curves required speed reductions while there was no need for

such a decrease in the curves with lower values. In addition, Caliendo and Lamberti [6] studied the influence of radius on accident rates and found a decrease in accident rates by increasing the radius between 200 and 500 m. Similarly, Cenek et al. [7] investigated the relationship for a wider range of radii, while Hauer [8] confirmed such a relationship for all radii. Hauer [9] also reported that curves with large deflection angles are more risky than the smaller ones.

Other research studies focused on evaluating geometric variables such as the lane and shoulder width, pavement type, skid resistance, annual average daily traffic, spiral transitions, and passing behavior [7, 10]. Furthermore, some other studies delved into the relationship between the speed and curvature [11, 12]. According to Tate and Turner [13], the difference between the negotiation speed and design

speed on curves has a significant effect on the injury crash rate. Studying the relationship between the operating speed and accident frequency rate, Bird and Hashim [14] indicated that higher operating speeds generally cause fewer accidents. Likewise, Wang et al. [15] investigated the relationship between average operating speed and accident severity and found that the operating speed with a 1% increase in the average operating speed results in a 0.074% decrease in the number of minor injuries with a 0.095% increase in the number of fatalities. Other studies evaluated the relationship between higher speed limits and the probability of accident severities and reported that higher speed limits increased the probability of a more severe accident and that accident severity increased outside level and straight roadways [16, 17]. Furthermore, Thomas [18] examined the influence of the segment length on crash analysis outside intersection-related sites and concluded that there is no definitive length which performs better than any other and that the length of the used segment solely depends on the type of the research. Few studies indicated that based on geometric and environmental features, variable-length segments perform better in the crash analysis compared to the fixed segments [19, 20]. Moreover, Caliendo and Lamberti [6], in a study, focused on the relationship between roadway factors and crash rates and demonstrated that segment types, access control, sight distance, and design consistency were highly correlated with crash rates.

Therefore, this study aims to evaluate the effect of roadway and development factors on accident frequency using ANOVA test and Chi-square tests on a rural road. Moreover, it develops a rural road safety risk index based on K-means clustering and Gaussian models to produce a technique for supporting the road safety analysis.

The organization of the remaining parts of the study is as follows. In Section 2, the literature review is presented, together with a discussion of previous studies related to the importance of factors contributing to road accidents and previous methods for accident data analysis. Additionally, Section 3 involves a description of data collection, followed by explaining the method of the present study about significance and clustering analyses and proposing the safety risk index. The obtained results regarding the proposed method are presented in Section 4. In Section 5, a sensitivity analysis is conducted by comparing the proposed safety risk index of the current study and that of the other studies. Finally, Section 6 contains the conclusion about the obtained results.

2. Literature Review

Several studies focused on driving safety affected by various factors and investigated the relationship between these factors and road accidents. Road accident data are classified as big data and include many attributes belonging to the accident such as driver attributes, environmental causes, as well as traffic, vehicle, and geometric characteristics and the location nature and the time of the day. In addition, data related to road accidents are taken for a long period of time and available as datasets, statistical tables and reports, or even Global Positioning System data. According to several studies, statistical and data mining techniques are proper for

analyzing the road accident data [21–24]. Lee et al. [25] designed a statistical framework as a fine choice for analyzing the road accidents with geometric factors including driver characteristics and road layout, along with the design of the car and weather condition. However, most road accidents are attributed to the “human factor,” most especially to road safety violations [26].

Some researchers investigated the effect of roadway factors on the number of road accidents on urban highways. They applied different techniques to establish a relationship between these factors and the accident frequency rate [27–29]. In addition, others reported that not only roadway factors but also development factors including land use and accessibility number are the main factors influencing the number of traffic accidents on multilane highways. They found a robust relationship between the accident frequency rate and development factors. In order to reduce accident frequency rates, it is vital to apply development factors in accident analysis in order to promote safety on roads [28–34].

Shirmohammadi et al. [35] highlighted the clustering drivers regarding driving behaviors and skills as important factors which contribute to road accidents using the clustering analysis. Shen et al. [36] used clustering analyses to identify accident blackspots on rural roads. In addition, Alotaibi [37] employed data mining techniques to simplify road accident data since such methods are novel and superior to classical statistical techniques and help the researchers to discover the relationship between the hidden data. Several data mining methods in the transportation field are broadly utilized for road accident data analysis, including clustering algorithms, as well as classification and association rule mining [38, 39], although accident data are heterogeneous (different variables).

Among accident data analysis methods, clustering analysis is the best way to find several between-data correlations which probably remain unknown [40]. Moreover, data mining techniques are useful for overcoming the accident data [41]. Ma and Kockelman [42] classified road segments which have similar characteristics. The results of this study were based on a linear regression model to estimate crash frequency within each cluster. Other studies employed clustering analysis for roadway crashes and safety projects [43–45]. Similarly, Sekuła et al. [46] proposed a clustering approach to predict the probability of a collision occurring in the proximity of planned road maintenance operations (i.e., work zones). Different other studies also concluded data mining techniques are more advanced and better than traditional statistical techniques [47–51].

To our best knowledge, no study has investigated the effect of roadway and development factors, especially the difference between posted speed limits and operating speed and operating speed on accident frequency rate on rural roads. Furthermore, we did not find any previous study on developing a rural safety risk index using roadway and development factors. Furthermore, previous studies only used clustering analysis for drivers' behavioral characteristics concerning the accidents. Given this, the novelty of the present study is, firstly, investigating the effects of roadway

and development factors on the accident frequency rate. Secondly, it applies clustering analysis and the Gaussian model for developing a rural risk index of the clusters regarding roadway and development factors. Moreover, finding the contributing factors to accidents plays an important role in collision statistics, which is considered as another reason for developing the subjective and driver-based evaluation of road safety risk. Finally, SPSS 17.0 and MATLAB R2013a software were employed to obtain the results.

3. Research Method

The process of evaluating the effect of roadway and development factors on the accident frequency rate for the development of a rural road safety risk index is performed as follows (see Figure 1).

3.1. Case Study Area. Lorestan Province has an area of 29308 km² and a population of about 1.76 million. The capital city Khorramabad is located in the southern part of Lorestan. The province is widely known as a popular tourist destination. Since the Boroujerd-Khorramabad road is located throughout the transit road of the North to the South of Iran, it is the most densely populated part of the Lorestan roads, and the number of motor vehicles accidents had been steadily rising during 2013 to 2016. A comparison of the motor vehicle accidents from 2013–2016 along the Boroujerd-Khorramabad road revealed that the mortality rate reached up to 67% and the injury rate was up to 30%. During this period in total, there were 1409 accidents.

3.2. Data Collection. The accident frequency rate, normalized by the segment length, was used for this study and belongs to the accidents that occurred during three years (2013–2016). Regarding roadway and development factors in previous studies [28, 30–34] and data availability from the local police accident reports from 2013 to 2016 in the Boroujerd-Khorramabad rural road, evaluation of these factors and development of the rural risk index was based on such data. Using roadway and development factors not only makes the risk index more practical for rural roads but also reduces fatal and injury rates from accidents in future. Likewise, the roadway variables were average operating speed (km/hr), the difference between posted speed limits and operating speed (km/hr), annual average daily traffic (veh/day), segment length (km), the presence or absence of a speed control camera, homogeneous sections, and gradient (%). Moreover, development factors included dominant land uses along the roadways and the number of accessibility (Table 1). In this study, the two-lane, two-way rural highway of the Boroujerd-Khorramabad road in Lorestan province, Iran (Figure 2(a)), was considered as a case study, and the location map of the study area is shown in Figure 2(b). The geometric and traffic characteristics were classified into homogeneous sections, and based on the available information, some independent variables were used to divide the road network into homogeneous sections as well.

The Boroujerd-Khorramabad road is a two-lane, two-way road where the width of each lane and shoulder is constant

FIGURE 1: The proposed model.

and is equal to 3.65 and 1.85 meters, respectively, along the whole road and with no changes in lane or shoulder widths. Road pavement is in a relatively good condition along with road sections whose performance serviceability index (PSI) equals 3. The road sections are away from the zone of the influence of intersections, towns and so on. In addition, the value of side friction is considered 0.35 for the road sections according to AASHTO [52]. The value of the speed limit ranges from 40 km/hr to 95 km/hr with an average of 63 km/hr for road sections. Other geometric characteristics of the rural road including the characteristics of curvature and gradient sections are described in Table 1.

Therefore, based on the output of this approach, each road section was assigned a number of accidents varied from 0 to 13 per section. Considering the dynamic nature of traffic variables (i.e., operating speed and volume), traffic conditions were expressed by annual averages while road geometry was represented by categorical variables. The final dataset included 106 road sections (total length = 172 km) after the exclusion of sections applying missing traffic or geometry data.

3.3. Significance Analysis. The ANOVA test is one of the most applicable methods in transportation data analysis [53–56]. This method is used to evaluate whether the contributing factors have a significant impact on the accident frequency rate at the level of 0.05. Thus, the study examined the significance of the association between roadway and development factors and the accident frequency rate. The hypothesis was assumed as follows.

H_0 = there are no associations between roadway and development factors and the accident frequency rate

H_1 = there are associations between roadway and development factors and the accident frequency rate

TABLE 1: Definition of variables included in the link-based dataset and descriptive statistics analysis.

Variable	Link-based dataset	No. of samples	Minimum	Maximum	Mean	Std. deviation
Accident frequency rate	No. of accidents/segment length	106	0.00	12.66	0.40	1.44
Operating speed	—	106	50.10	106.00	83.87	12.39
Difference between posted speed and operating speed	—	106	-29.40	59.20	23.48	16.81
Segment length	—	106	0.10	5.20	1.62	0.88
Volume	(AADT)*1000	106	10.84	22.27	17.58	4.52
Presence or absence of a speed control camera	Binary variables: “1,” presence “0,” absence	106	0.00	1.00	0.03	0.17
Homogeneous sections	Binary variables: “1,” straight section	106	1.00	6.00	4.08	1.97
	“2,” straight uphill section					
	“3,” uphill section					
	“4,” straight downhill section					
	“5,” downhill section					
	“6,” horizontal curve section					
Gradient	G1: links with a median gradient below -0.3% (downhill)	106	0.00	2.00	1.17	0.93
	G2: links with median gradient between $\pm 0.3\%$ (Even)					
	G3: links with a median gradient above +0.3% (uphill)					
Dominant land uses along the roadways	Binary variables: “0,” residential “1,” commercial “2,” for other land uses	106	0.00	2.00	0.94	0.92
Number of accessibility	—	106	0.00	10.00	2.31	2.38

Note. AADT: annual average daily traffic.

FIGURE 2: Schematic map of the case study (source: Google map). (a) Plan of the Borujerd-Khorramabad road. (b) Location map of the study area.

Therefore, the hypothesis H_0 was rejected, while the hypothesis H_1 was accepted when the P value was less than 0.000.

3.4. Clustering Analysis. Clustering technique is one of the most commonly used data mining methods, and there are many clustering algorithms such as K-means and K-modes

[21, 57]. K-means algorithm is based on a centroid technique, while K-modes algorithm is based on the nominal data. The K-means algorithm is considered as one of the most popular data mining techniques for identifying the clusters based on accident frequencies [58, 59].

Using clustering techniques causes the problem of determining the best number of expected clusters. To solve this issue, the K-means algorithm is recommended to enter the

number of K clusters. According to the framework of this method, the best and optimal number of clusters is determined by the Elbow method [60]. This method is one of the optimal methods that depend on both the measure of similarities within a cluster and the parameters that are used for partitioning. Therefore, the steps of identifying the optimal number of clusters are summarized as follows [61].

- (1) Computing the clustering algorithm (i.e., K-means) for different values of K , $k = 2$ to $k = 15$
- (2) Calculating the total within-cluster sum of the square (wss) for each K cluster
- (3) Plotting the curve of wss according to the number of K clusters
- (4) Considering the location of a bend (knee) in the plot as a general indicator of the appropriate number of the clusters

3.5. Development of the Road Safety Risk Index. By the development of a risk index, it is vital to consider the fundamental elements that can contribute to road safety [62]. Ahmadinejad et al. [63] proposed a suitable index for road safety regarding deceleration numbers and safety parameters (e.g., crash rate and crash frequency rate). The results indicated that there is a significant correlation between safety parameters and deceleration numbers. Many studies defined safety risk by considering three variables including exposure, probability, and consequence [64, 65], which is shown in the following equation:

$$\text{Safety risk index} = \text{function} \quad (1)$$

(i.e., exposure, probability, and consequence),

where Exposure = measure to quantify the “exposure” of road users to potential roadway hazards. Probability = measure to quantify the chance of a vehicle being involved in a collision. Consequence = measure to quantify the severity level resulting from potential collisions.

4. Results and Discussion

4.1. Significance Analysis. To examine the effect of roadway and development factors on the accident frequency rate, the ANOVA test was run, the results of which are presented in Table 2. As shown in Table 2, operating speed and the difference between posted speed limits and operating speed have significant effects on the accident frequency rate due to Sig. (0.000) < 0.05. However, no significance is observed between the other factors and the accident frequency rate.

4.2. K-Means Clustering. The average linkage hierarchical clustering was used to determine the number of clusters although identifying the most optimal heterogeneous clusters has occasionally some limitations and deficiencies. Based on these limitations, the K-means cluster is applicable after determining the number of clusters. In this clustering method, using the centroids (i.e., the cluster center means)

generated from the average linkage hierarchical clustering is a starting point [66, 67].

Cluster analysis applies algorithms to collate individual variables with similar scores [68]. Based on the squared Euclidean distance measure, the cluster analysis utilizes the scores derived from the grouping variables. In the current study, the grouped variables included the accident frequency rate, operating speed, the difference between posted speed limits and the operating speed, segment length, annual average daily traffic, the number of accessibility, and dominant land uses along the roadways, as well as the presence or absence of a speed control camera, curvature, and gradient.

The standardized scores (Z-scores) of variables are used to avoid the problem of comparing Euclidean distances based on different measurement scales [69]. Based on Figure 3, the optimal number of a cluster is determined as six clusters based on the distinctive break (elbow) selected according to the squared Euclidean distance in comparison with agglomeration coefficients. Table 3 demonstrates the results of final cluster centers for independent and dependent variables.

Evaluating the ANOVA test of variables in the clusters for finding the most effective factors that play a role in the accident frequency rate, only the difference between posted speed limits and operating speed is specified as the most effective variable among the roadway and development factors due to the maximum statistical value or F-statistic observed in Tables 4 and 5. Regarding the accident frequency rate, clusters are arranged in a specific order as highly risky, relatively high risky, moderately risky, relatively low risky, low risky, and not risky (safe) clusters (Figure 4(a)).

The F tests should be used only for descriptive purposes because the clusters are chosen to maximize the differences among the cases in different clusters. However, the observed significance levels are not corrected for this and, thus, cannot be interpreted as the tests of the hypothesis that the cluster means are equal.

Similarly, based on the results of the Chi-square (X^2) test (Table 5), the maximum X^2 shows a difference between posted speed limits and the operating speed. Accordingly, the maximum X^2 indicates how much this factor (i.e., the difference between posted speed limits and operating speed) affects the accident frequency rate. Hence, the maximum X^2 was employed in the proposed model to discover the relationship between this variable and the accident frequency rate (Figure 4(b)). Additionally, the Chi-square distribution probability function was utilized to obtain the probability of each cluster (Figure 4(c)). As displayed, the maximum and minimum probability is determined for the fifth and the second clusters.

To understand the effect of the difference between posted speed limits and the operating speed on the accident frequency rate, the probability of the occurrence was obtained for each cluster. Based on Figure 4(b), when the difference between posted speed limits and the operating speed reduces from the safe cluster, the probability of accident occurrence risk in each cluster increases (Figure 4(c)). Therefore, the following results are obtained by comparing the difference

TABLE 2: ANOVA test for the examination of the effect of roadway and development factors on the accident frequency rate.

Results of significance analysis of roadway factors					
Results of significance analysis (operating speed)					
	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	218.24	93	2.35	625.77	0.000*
Within groups	0.05	12	0.00		
Total	218.28	105			
Results of significance analysis (the difference between posted speed limits and operating speed)					
	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	218.24	96	2.27	454.66	0.000*
Within groups	0.05	9	0.01		
Total	218.28	105			
Results of significance analysis (segment length)					
Segment length	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	35.41	33	1.07	0.42	0.996
Within groups	182.87	72	2.54		
Total	218.28	105			
Results of significance analysis (volume)					
Volume	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	5.900	3	1.967	0.945	0.422
Within groups	212.38	102	2.082		
Total	218.28	105			
Results of significance analysis (the presence or absence of a speed control camera)					
Presence or absence of a speed control camera	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	0.49	1	0.49	0.23	0.629
Within groups	217.79	104	2.09		
Total	218.28	105			
Results of significance analysis (homogeneous sections)					
Homogeneous sections	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	5.71	5	1.14	0.53	0.751
Within groups	212.41	99	2.15		
Results of significance analysis (gradient)					
Grade	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	0.69	2	0.35	0.17	0.849
Within groups	217.59	103	2.11		
Results of significance analysis of development factors					
Results of significance analysis (dominant land uses along the roadways)					
Dominant land uses along the roadways	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	4.45	2	2.22	1.07	0.346
Within groups	213.84	103	2.08		
Total	218.28	105			
Results of significance analysis (the number of accessibility)					
Number of accessibility	Sum of squares	df	Mean square	<i>F</i>	Sig.
Between groups	9.77	9	1.086	0.50	0.871
Within groups	208.51	96	2.172		
Total	218.28	105			

Note. All the italic coefficients are not statistically significant. *It is significant at the 0.05 level.

between posted speed limits and the operating speed and the probability in each cluster (Figure 4).

As shown, the first cluster, namely, “relatively high risk,” is ranked the second based on the accident frequency rate,

and its probability risk value is less than 10%. Hence, the occurrence of an accident is relatively low in this cluster.

The second cluster is ranked the fourth, “relatively low risk,” based on the observed accident frequency rate, and its

FIGURE 3: Number of clusters.

TABLE 3: Final cluster centers for independent and dependent variables in this study.

Final cluster centers		Cluster					
		1	2	3	4	5	6
Independent variables	Operating speed	56.66	62.65	91.79	81.80	97.53	88.00
	Difference between posted speed limits and operating speed	-17.34	8.10	41.02	27.77	18.69	-0.75
	Segment length	1.26	1.56	1.90	1.42	1.84	1.56
	Volume	16.42	16.87	15.75	18.95	17.73	18.56
	Presence or absence of a speed control camera	0.00	0.09	0.04	0.03	0.00	0.00
	Homogeneous sections	3.00	3.64	4.43	4.26	3.86	3.67
	Gradient	1.20	1.09	1.07	1.23	1.21	1.22
	Dominant land uses along the roadways land use	1.20	1.27	0.75	0.69	1.29	1.56
	Number of accessibility	0.60	1.36	3.46	2.13	2.93	0.67
Dependent variable	Accident frequency rate	0.80	0.53	0.01	0.12	0.75	1.87

TABLE 4: Examination of the ANOVA test of variables in the clusters.

Variables	ANOVA					
	Cluster		Error		F	Sig.
	Mean square	df	Mean square	df		
Operating speed	2669.37	5	27.74	100	96.21	0.000*
Difference between posted speed and operating speed	5174.64	5	37.82	100	136.81	0.000*
Segment length	1.04	5	0.76	100	1.37	0.241
Volume	37.55	5	19.57	100	1.92	0.098
Presence or absence of a speed control camera	0.01	5	0.03	100	0.47	0.796
Homogeneous sections	2.97	5	3.93	100	0.76	0.583
Gradient	0.11	5	0.90	100	0.12	0.988
Dominant land uses along the roadways	2.01	5	0.80	100	2.52	0.034*
Number of accessibility	18.55	5	5.00	100	3.71	0.004*
Accident frequency rate	5.88	5	1.89	100	3.10	0.012*

Note. All the italic coefficients are not statistically significant. *It is significant at the 0.05 level.

TABLE 5: Results of the chi-square test for examining the most effective factors in relation to the accident frequency rate.

Variables	(Chi-square) χ^2	Sig. (alpha)
Operating speed	1051.77	0.003*
Difference between posted speed limits and operating speed	1060.00	0.043*
Segment length	218.32	1.000
Volume	34.85	0.248
Presence or absence of a speed control camera	0.95	1.000
Homogeneous sections	64.38	0.083
Gradient	21.38	0.375
Dominant land uses along the roadways	19.16	0.511
Number of accessibility	44.35	1.000

Note. All the italic coefficients are not statistically significant. *It is significant at the 0.05 level.

FIGURE 4: (a) Relationship between the cluster and accident frequency rate. (b) Relationship between the cluster and the difference between posted speed limits and the operating speed. (c) Relationship between the cluster and probability.

probability risk value is less than 5%; thus, the incidence of a high accident frequency rate is very low in this cluster.

Likewise, the third cluster is ranked the sixth, "safe cluster," based on the accident frequency rate. Identically, the probability risk value is less than 5%, which demonstrates that the accident occurrence is very low in this cluster.

The fourth cluster is ranked the fifth, "low risk," based on the accident frequency rate. By comparing the probability risk value in this cluster with safe clusters, it can be found

that the probability of accident occurrence in this cluster is 10% which might lead to a lower rate of accident.

In addition, the fifth cluster is put on the third, "moderately risk," place considering the accident frequency rate. Based on the evaluation of the accident occurrence probability of this risky cluster and its comparison with the other cluster, the probability is 85%, which is high, and thus, the accident frequency rate is expected to demonstrate a significant increase.

FIGURE 5: Examination of the effect of the difference between posted speed limits and the operating speed on the accident frequency rate.

FIGURE 6: The relationship between the difference of posted speed limits and the operating speed and the accident frequency rate based on the proposed model.

Finally, the sixth cluster is ranked the first, “high risk,” based on the increasing accident frequency rate. Regarding the probability of accident occurrence in the cluster, the obtained probability is less than 5%, indicating that the frequency related to this kind of the cluster of accident might happen less than the other risky clusters.

Therefore, the probability of the occurrence of a moderate risky cluster is higher as compared to the other clusters, and more accident frequency rates occur in this cluster. Furthermore, the difference between the posted speed limits and operating speed in this cluster is nearly 18.69 km/hr which is near to the mean of the difference between the posted speed limits and operating speed. As a result, the accident frequency rate significantly increases by decreasing

the difference between the posted speed limits and operating speed from the safe cluster (Figure 5).

4.3. Assessment of the Association of Posted Speed Limits and the Operating Speed on the Accident Frequency Rate. The relationship between difference posted speed limits and the operating speed and the accident frequency rate, as well as the behavior of the frequency of risky and unrisky clusters was evaluated using the Gaussian function. The findings (Figure 6) indicated that this function shows a better performance based on the considering coefficients (with 95% confidence bounds) and the goodness of fit parameters including the sum of the squared errors, R-square, adjusted R-square, and root mean square error presented in Table 6. According to the Gaussian function, the difference between posted speed limits and the operating speed can cause an increase and decrease trend in the accident frequency rate in each cluster. Therefore, the average reducing factor of the accident frequency rate is 0.99 by increasing per km/hr in the difference between posted speed limits and the operating speed among the safe clusters. This means that drivers in safe clusters maintain an operating speed lower than the posted speed limits. Hence, by increasing the difference between the posted speed limits and the operating speed, maximum difference is obtained, thereby decreasing the number of accidents per length by the rate of 0.99. From this finding, one can infer that drivers in safe clusters do not exceed the speed limits. However, in risky and unsafe clusters, drivers exceed the speed limits, and their operating speed is more than the speed limits, which, in turn, could lead to 1.17 rise, on average, the in accident frequency rate. In other words, a minimum difference is obtained, and the number of accidents per length went up by the rate of 1.17. Therefore, the growth factor in risky and unsafe clusters is 1.18 times and is as often as the accident frequency rate in low risky and safe clusters. These results are consistent with the findings of the probability of accident occurrence risk when the difference between posted speed limits and the operating speed reduces from the safe cluster in which drivers keep the minimum difference, and therefore, the probability of accident occurrence risk in each cluster increases.

As an example in Figure 6 and Table 7, when the difference between posted speed limits and the operating speed is 0, the accident frequency rate is 1.7. In such cases, drivers are categorized in the high risk cluster based on the proposed risk index. Based on Leur and Sayed’s study [62], when accident frequency rate is 11.1, drivers are categorized as the high-risk cluster. However, when the difference between posted speed limits and the operating speed is -20 km/hr, the accident frequency is 0.6, and drivers are categorized in the relatively high-risk cluster according to the proposed the risk index. Based on Leur and Sayed’s study [62], when the accident frequency rate is 12.12, drivers are categorized in the relatively high-risk cluster. Thus, by comparing the results of the present study with those of Leur and Sayed [62], it can be shown that the proposed method has categorized clusters appropriately similar to Leur and Sayed’s study [62] as the high-risk cluster and relatively high-risk cluster, while the accident frequency rate and risk index are different.

TABLE 6: Proposed model for the analysis of the accident frequency rate for the clusters.

General model Gaussian	Coefficients (with 95% confidence bounds)			Goodness of fit			
$F(x) = a_1 * \exp^{-((x-b_1)^2/c_1)}$	a_1	b_1	c_1	SSE	R-square	Adjusted R-square	RMSE
	1.93	-4.82	13.70	0.51	0.77	0.61	0.41

Note. SSE: the sum of the squared errors; RMSE: root mean square error.

TABLE 7: Examination of the obtained risk value based on the proposed risk index and its comparison with that of the other studies for each cluster.

Cluster	Operating speed	Proposed study based on collected data				Leur and Sayed's study [62]			
		Difference between posted speed limits and the operating speed	Risk degree	Accident frequency ratio	Risky index (predicted value) (Eq. 2)	Accident frequency rate (consequences)	Probability	Exposure	Final risk
1	56.66	-17.34	Relatively high risk	0.80	1.62	12.12	0.013	3	0.476
2	62.65	8.10	Relatively low risk	0.53	2.02	10.12	0.027	1	0.275
3	91.79	41.02	Not risk (safe)	0.01	0.14	7.83	0.058	0	0.00
4	81.80	27.77	Low risk	0.12	0.59	12.4	0.012	1	0.146
5	97.53	18.69	Moderately risk	0.75	1.98	9.06	0.039	2	0.708
6	88.00	-0.75	High risk	1.87	14.81	11.1	0.019	3	0.636

TABLE 8: ANOVA test for the proposed safety risk index.

Parameter estimates				
Parameter	Estimate	Std. error	95% confidence interval	
b1	1.01	0.07	Lower bound 0.84	Upper bound 1.18
ANOVA ^a				
Source	Sum of squares		df	Mean squares
Regression	231.74		1	231.74
Residual	1.26		5	0.25
Uncorrected total	233.00		6	
Corrected total	150.20		5	

Dependent variable: risk value

a. R squared = $1 - (\text{residual sum of squares})/(\text{corrected sum of squares}) = 0.99$

4.4. Safety Risk Index Model. Based on the findings of ANOVA and Chi-square tests, among the roadway and development factors and their effects on frequency accident rate, only operating speed and the difference between posted speed limits and the operating speed were employed to the safety risk index model in equation (2).

The results of the safety risk index for each cluster are displayed in Table 8. Based on the obtained data, the third cluster with the lowest risky index is regarded as the safest cluster, while the sixth cluster is considered as an unsafe cluster with the maximum risk index among the six clusters.

$$\text{Safety Risk index} = \sqrt{\frac{\text{Operating speed} \times \text{accident frequency rate}}{\text{Difference between posted speed limits and operating speed}}}. \quad (2)$$

Moreover, the Chi-Square distribution probability function was used as a probability generator for obtaining

the probability of each cluster, the results of which are presented in Table 7. The final risk for the study by Leur and

FIGURE 7: Sensitivity analysis between the two studies.

Sayed [70] was obtained according to the values of the accident frequency rate, probability values, and exposure or scores. As shown in Table 8, the findings of the ANOVA test also approved that the proposed model has a high prediction power of risk for clusters.

4.5. Future Research Works. Future works might consider investigating the effect of geometric factors such as road width, weather, and lightening conditions on accident frequency, and development of the rural risk index. In addition, data mining and multicriteria decision making approaches including decision tree techniques, fuzzy AHP, and fuzzy COPRAS could be noteworthy to expand this risk index for rural roads for drivers based on database and experts' opinion in the field.

5. Sensitivity Analysis

To examine the reliability of the safety risk index for clusters, a sensitivity analysis was performed between the results of the proposed model and the findings of Leur and Sayed [70], as shown in Figure 7. Based on Figure 7, it is evident that the value of risk index from cluster 1 to 5 is close to the value of risk index in clusters in Leur and Sayed's study [70] except the sixth cluster. In addition, the maximum risk index of the proposed study is observed in the sixth cluster which is a high-risk one. However, in the sixth cluster of Leur and Sayed [70], the risk index is 0.636 which is lower than the proposed study which makes it different. This difference is due to the use of the difference between posted speed limits and operating speed in the development of the rural risk index in the present study. This discrepancy can include more high risk drivers in the sixth cluster for the proposed study.

6. Conclusions

Given the fact that roadway and development factors are known as the most effective parameters contributing to road traffic accidents on roads, applying these factors in safety analysis could be instrumental in reducing the accident frequency rate and preventing the growth fatality and injury rate on rural roads. Therefore, this study evaluated the effect

of roadway and development factors on accident frequency in order to develop a rural road safety risk index using the K-means clustering and Gaussian model. Relying on the obtained data and the results of the analysis, the main findings of the study and the evaluation of the rural accident risk index among roadway and development factors are summarized based on the ANOVA test, as well as clustering and risk analyses as follows.

- (1) Based on the results of the ANOVA test, among roadway and development factors, only operating speed and the difference between posted speed limits and the operating speed had significant effects on the accident frequency rate. Furthermore, the results of the Chi-square test demonstrated that the maximum chi-square of the operating speed in the risky index has a lower effect on the accident frequency rate compared to the difference between posted speed limits and the operating speed.
- (2) Based on the K-means clustering analysis of roadway and development factors respecting the accident frequency rate, six easily understandable clusters were investigated as high risky, relatively high risky, moderately risky, relatively low risky, low risky, and not risky (safe) drivers for each cluster. The comparison of the clusters regarding the accident frequency rate revealed that the sixth cluster was categorized as the high risky cluster, whereas the third cluster was considered as a safe cluster.
- (3) The risky index model was proposed based on the Gaussian model to analyze the behavior of the accident frequency rate for clusters and to obtain the risk value. Therefore, the average reducing factor of the accident frequency rate was achieved by 0.99 through increasing (per km/hr) the difference between the posted speed limits and the operating speed among the safe clusters. However, in unsafe clusters, the average increasing factor of the accident frequency rate was obtained as 1.17. Therefore, the growth factor in risky and unsafe clusters was 1.18 times the accident frequency rate in low risky and safe clusters.
- (4) Based on the comparison of the difference between posted speed limits and the operating speed and the probability of accident occurrence, it is concluded that, by decreasing the difference of posted speed limits and the operating speed from the safe cluster, the probability of accident occurrence risk in each cluster increases, followed by an increase in the accident frequency rate. As a result, the maximum probability of the accident occurrence was observed in the fifth cluster, which was achieved by 85%. The probability of accidents in the fifth cluster increased as well.
- (5) Sensitivity analysis showed that the proposed safety risk index has a better performance regarding predicting the risk values for the clusters when compared to the other study.

- (6) The proposed risk index model is considered as a useful tool for obtaining the safety risk value for studies concerning the accident rate and clustering analysis of drivers on rural roads. Finally, this study can be useful for safety research organizations such as governmental institutes and police centers to consider the maximum risk value in order to accurately present their plans and strategies toward minimizing accidents.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] World Health Organization (WHO), "Road traffic injuries," 2018, <http://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- [2] M. Ahadi, M. Hassanpour, P. Bashiri, and P. Bashiri, "Strategies to promote safety to prevent pedestrian accidents in the city of Qazvin," *Safety Promotion and Injury Prevention*, vol. 4, no. 3, pp. 143–150, 2017.
- [3] World Health Organization, *Violence, Injury Prevention. Global Status Report on Road Safety 2013: Supporting a Decade of Action*, World Health Organization, Geneva, Switzerland, 2013.
- [4] M.-I. M. Imprialou, M. Quddus, D. E. Pitfield, and D. Lord, "Re-visiting crash-speed relationships: A new perspective in crash modelling," *Accident Analysis & Prevention*, vol. 86, pp. 173–185, 2016.
- [5] I. B. Anderson and R. A. Krammes, "Speed reduction as a surrogate for accident experience at horizontal curves on rural two-lane highways," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1701, no. 1, pp. 86–94, 2000.
- [6] C. Caliendo and R. Lamberti, "Relationships between accidents and geometric characteristics for four lanes median separated roads," in *Proceeding of the Road Safety on Three Continents, Moscow, Russia*, Moscow, Russia, September 2001.
- [7] P. D. Cenek, R. B. Davies, and R. J. Henderson, "Crash risk relationships for improved road safety management (no. 488)," 2012.
- [8] E. Hauer, "Traffic conflicts and exposure," *Accident Analysis & Prevention*, vol. 14, no. 5, pp. 359–364, 1982.
- [9] E. Hauer, "Safety and the choice of degree of curve," *Transportation Research Record*, vol. 1665, no. 1, pp. 22–27, 1999.
- [10] M. G. Karlaftis and I. Golias, "Effects of road geometry and traffic volumes on rural roadway accident rates," *Accident Analysis & Prevention*, vol. 34, no. 3, pp. 357–365, 2002.
- [11] V. Andjus and M. Maletin, "Speeds of cars on horizontal curves," *Transportation Research Record*, vol. 1612, no. 1, pp. 42–47, 1998.
- [12] J. Collins, K. Fitzpatrick, K. M. Bauer, and D. W. Harwood, "Speed variability on rural two-lane highways," *Transportation Research Record*, vol. 1658, no. 1, pp. 60–69, 1999.
- [13] F. Tate and S. Turner, "Road geometry and drivers' speed choice," *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*, vol. 16, no. 4, p. 53, 2007.
- [14] R. Bird and I. Hashim, "Exploring relationship between safety and consistency of geometry and speed on british roads (no. 06-1509)," 2006.
- [15] X. Wang, T. Fan, M. Chen, B. Deng, B. Wu, and P. Tremont, "Safety modeling of urban arterials in Shanghai, China," *Accident Analysis & Prevention*, vol. 83, pp. 57–66, 2015.
- [16] S. Dissanayake and I. Ratnayake, "Identification of factors leading to high severity of crashes in rural areas using ordered probit modeling," *Journal of the Transportation Research Forum*, vol. 45, no. 2, pp. 87–101, 2006.
- [17] N. V. Malyskhina, F. L. Mannering, and S. A. Labi, *Influence of Speed Limits on Roadway Safety in Indiana*, Joint Transportation Research Program, West Lafayette, IN, USA, 2007.
- [18] I. Thomas, "Spatial data aggregation: exploratory analysis of road accidents," *Accident Analysis & Prevention*, vol. 28, no. 2, pp. 251–264, 1996.
- [19] G. Koorey, "Road data aggregation and sectioning considerations for crash analysis," *Transportation Research Record*, vol. 2103, no. 1, pp. 61–68, 2009.
- [20] J. M. P. Mayora, "Relevant variables for crash-rate prediction on Spain's two-lane rural roads," in *Proceedings of 82nd Annual Meeting*, Transportation Research Board, Washington, DC, USA, January 2003.
- [21] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques: The Morgan Kaufmann Series in Data Management Systems*, Elsevier, Amsterdam, Netherlands, 3rd edition, 2011.
- [22] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1666–1676, 2011.
- [23] F. L. Mannering, V. Shankar, and C. R. Bhat, "Unobserved heterogeneity and the statistical analysis of highway accident data," *Analytic Methods in Accident Research*, vol. 11, pp. 1–16, 2016.
- [24] G. Janani and N. R. Devi, "Road traffic accidents analysis using data mining techniques," *JITA-Journal of Information Technology and Applications*, vol. 14, no. 2, 2016.
- [25] C. Lee, F. Saccomanno, and B. Hellinga, "Analysis of crash precursors on instrumented freeways," *Transportation Research Record*, vol. 1784, no. 1, pp. 1–8, 2002.
- [26] M. J. Sullman, M. L. Meadows, and K. B. Pajo, "Aberrant driving behaviours amongst New Zealand truck drivers," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 5, no. 3, pp. 217–232, 2002.
- [27] C. Wang, M. A. Quddus, and S. G. Ison, "The effect of traffic and road characteristics on road safety: a review and future research direction," *Safety Science*, vol. 57, pp. 264–275, 2013.
- [28] M. Mohanty and A. Gupta, "Factors affecting road crash modeling," *Journal of Transport Literature*, vol. 9, no. 2, pp. 15–19, 2015.
- [29] H. Shirmohammadi, A. S. Najib, and F. Hadadi, "Identification of road critical segments using wavelet theory and multi-criteria decision-making method," *European Transport-trasporti Europei*, vol. 68, no. 2, 2018.
- [30] T. Litman, "Measuring transportation: traffic, mobility and accessibility," *ITE Journal*, vol. 73, no. 10, p. 28, 2003.
- [31] J. Withanaarachchi, S. Setunge, and S. Bajwa, "Traffic impact assessment and land use development and decision making," in *Proceedings of International Conference on Disaster Management*, pp. 256–273, Kumamoto, Japan, August 2012.

- [32] A. Bako and I. Musa, "Effect of land use on road traffic accidents in urban zaria area, Nigeria," *BEST: International Journal of Humanities, Arts, Medicine and Sciences (BEST: IJHAMS)*, vol. 2, no. 1, pp. 35–42, 2014.
- [33] C. Berthod, "Land use planning measures promoting road safety," in *Proceedings of TAC 2016: Efficient Transportation-Managing the Demand-2016 Conference and Exhibition of the Transportation Association of Canada*, Transportation Association of Canada (TAC), Ottawa, Ontario, Canada, November 2016.
- [34] S. B. Kusselson, *Investigating How Land Use Patterns Affect Traffic Accident Rates Near Frontage Road Cross-Sections: A Case Study on Interstate 610 in Houston, Texas*, Oklahoma State University, Stillwater, OK, USA, 2013.
- [35] H. Shirmohammadi, F. Hadadi, and M. Saeedian, "Clustering analysis of drivers based on behavioral characteristics regarding road safety," *International Journal of Civil Engineering*, vol. 17, no. 8, pp. 1–14, 2019.
- [36] L. Shen, J. Lu, M. Long, and T. Chen, "Identification of accident blackspots on rural roads using grid clustering and principal component clustering," *Mathematical Problems in Engineering*, vol. 4, pp. 1–12, 2019.
- [37] A. S. Alotaibi, "Density-based clustering for road accident data analysis," *International Journal of Advanced and Applied Sciences*, vol. 5, no. 8, pp. 113–121, 2018.
- [38] S. K. Barai, "Data mining applications in transportation engineering," *Transport*, vol. 18, no. 5, pp. 216–223, 2003.
- [39] P. C. Srividhya, "A comparative analysis of clustering approach for predicting road traffic accident dataset," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, no. 6, pp. 468–473, 2017.
- [40] B. Depaire, G. Wets, and K. Vanhoof, "Traffic accident segmentation by means of latent class clustering," *Accident Analysis & Prevention*, vol. 40, no. 4, pp. 1257–1266, 2008.
- [41] S. Kumar and D. Toshniwal, "A data mining framework to analyze road accident data," *Journal of Big Data*, vol. 2, no. 1, p. 26, 2015.
- [42] J. Ma and K. Kockelman, "Crash frequency and severity modeling using clustered data from Washington state," in *Proceedings of 2006 IEEE Intelligent Transportation Systems Conference*, pp. 1621–1626, IEEE, Toronto, Canada, 2006 September.
- [43] S. Y. Sohn, "Quality function deployment applied to local traffic accident reduction," *Accident Analysis & Prevention*, vol. 31, no. 6, pp. 751–761, 1999.
- [44] T. F. Golob and W. W. Recker, "A method for relating type of crash to traffic flow characteristics on urban freeways," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 1, pp. 53–80, 2004.
- [45] S. C. Wong, B. S. Y. Leung, B. P. Loo, W. T. Hung, and H. K. Lo, "A qualitative assessment methodology for road safety policy strategies," *Accident Analysis & Prevention*, vol. 36, no. 2, pp. 281–293, 2004.
- [46] P. Sekula, Z. Vander Laan, K. Farokhi Sadabadi, and M. J. Skibniewski, "Predicting work zone collision probabilities via clustering: application in optimal deployment of highway response teams," *Journal of Advanced Transportation*, vol. 1, no. 1529, pp. 1–16, 2018.
- [47] L. Y. Chang and W. C. Chen, "Data mining of tree-based models to analyze freeway accident frequency," *Journal of Safety Research*, vol. 36, no. 4, pp. 365–375, 2005.
- [48] S. Kumar and D. Toshniwal, "Analysing road accident data using association rule mining," in *Proceedings of 2015 International Conference on Computing, Communication and Security (ICCCS)*, pp. 1–6, IEEE, Pamplemousses, Mauritius, 2015, December.
- [49] A. Tavakoli Kashani, A. Shariat-Mohaymany, and A. Ranjbari, "A data mining approach to identify key factors of traffic injury severity," *Promet-Traffic&Transportation*, vol. 23, no. 1, pp. 11–17, 2011.
- [50] J. Abellán, G. López, and J. De Oña, "Analysis of traffic accident severity using decision rules via decision trees," *Expert Systems with Applications*, vol. 40, no. 15, pp. 6047–6054, 2013.
- [51] S. Kumar and D. Toshniwal, "A data mining approach to characterize road accident locations," *Journal of Modern Transportation*, vol. 24, no. 1, pp. 62–72, 2016.
- [52] AASHTO, *A Policy on Geometric Design of Highways and Streets*, AASHTO, Washington, DC, USA, 1984.
- [53] X. Qu, Q. Meng, and S. Li, "Analyses and implications of accidents in Singapore Strait," *Transportation Research Record*, vol. 2273, no. 1, pp. 106–111, 2012.
- [54] X. Qu, Y. Yang, Z. Liu, S. Jin, and J. Weng, "Potential crash risks of expressway on-ramps and off-ramps: a case study in Beijing, China," *Safety Science*, vol. 70, pp. 58–62, 2014.
- [55] S. Jin, X. Qu, and D. Wang, "Assessment of expressway traffic safety using Gaussian mixture model based on time to collision," *International Journal of Computational Intelligence Systems*, vol. 4, no. 6, pp. 1122–1130, 2011.
- [56] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transportation Research Part C: Emerging Technologies*, vol. 35, pp. 46–56, 2013.
- [57] P. N. Tan, M. Steinbach, and V. Kumar, "Cluster analysis: basic concepts and algorithms," *Introduction to Data Mining*, vol. 8, pp. 487–568, 2006.
- [58] A. M. Aljofey and K. Alwagih, "Analysis of accident times for highway locations using K-means clustering and decision rules extracted from decision trees," *International Journal of Computer Applications Technology and Research*, vol. 7, no. 01, pp. 001–011, 2018.
- [59] W. Budiawan, S. Saptadi, A. Arvianto, and P. Andarani, "Implementation K-means clustering analysis of traffic accident in semarang city using weka interface," *International Journal of Science and Engineering Investigations*, vol. 7, no. 81, pp. 83–86, 2018.
- [60] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: an Introduction to Cluster Analysis*, John Wiley & Sons, Hoboken, NY, USA, 2009.
- [61] A. Kassamara, "Determining the optimal number of clusters: 3 must known methods – unsupervised machine learning," 2015, <http://www.sthda.com/english/wiki/determining-the-optimal-number-of-clusters-3-must-known-methods-unsupervised-machine-learning>.
- [62] U. D. Hasmukhrai, K. V. Ganeshbabu, and P. J. Gundaliya, "Identification of crash risk index for urban road: a case study of ahmedabad city," *International Journal of Innovative Research in Technology*, vol. 2, no. 12, pp. 134–140, 2016.
- [63] M. Ahmadinejad, S. Afandizadeh Zargari, and R. Jalalkamali, "Are deceleration numbers a suitable index for road safety?" *Proceedings of the Institution of Civil Engineers-Transport*, vol. 171, no. 5, pp. 247–252, 2017.
- [64] W. Haddon, "Advances in the epidemiology of injuries as a basis for public policy," *Public Health Reports*, vol. 95, no. 5, p. 411, 1980.
- [65] M. J. Koornstra, "The evolution of road safety and mobility," *IATSS Research*, vol. 16, no. 2, pp. 129–148, 1992.
- [66] M. Sarstedt and E. Mooi, *Cluster Analysis in: A Concise Guide to Market Research*, pp. 273–324, Springer, Berlin, Germany, 2014.

- [67] R. C. De Amorim, "Constrained clustering with minkowski weighted k-means," in *Proceedings of 2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI)*, pp. 13–17, IEEE, Budapest, Hungary, 2012 November.
- [68] B. S. Everitt, *Cluster Analysis*, Halsted Press, New York, NY, USA, 3rd edition, 1993.
- [69] G. W. Milligan and L. M. Sokol, "A two-stage clustering algorithm with robust recovery characteristics," *Educational and Psychological Measurement*, vol. 40, no. 3, pp. 755–759, 1980.
- [70] P. D. Leur and T. Sayed, "Development of a road safety risk index," *Transportation Research Record*, vol. 1784, no. 1, pp. 33–42, 2002.