

Article

Not peer-reviewed version

A New Approach for Identifying Safety Improvement Sites on Rural Highways: Validation Study

[Bishal Dhakal](#) and [Ahmed Al-Kaisy](#) *

Posted Date: 20 December 2023

doi: 10.20944/preprints202312.1497.v1

Keywords: Rural Roads Safety; Network Screening; Highway Safety Improvement Programs; Ranking; Regression Analysis



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: 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.

Article

A New Approach for Identifying Safety Improvement Sites on Rural Highways: Validation Study

Bishal Dhakal ¹ and Ahmed Al-Kaisy ^{2,*}

¹ Department of Civil Engineering, Montana State University, Bozeman, Montana, USA. Email: bishaldhakal@montana.edu

² Department of Civil Engineering, Montana State University, Bozeman, Montana, USA. Email: alkaisy@montana.edu

* Correspondence: alkaisy@montana.edu

Abstract: The research presented in this paper examines a new proposed approach for identifying safety improvement sites on rural highways. Unlike conventional approaches, the proposed approach does not require crash history, but rather utilizes classified variables for traffic volume, geometric features, and roadside characteristics that do not require access to exact data or extensive technical expertise. The research validates the performance of the proposed approach using field data from a large sample of rural two-lane highway segments in the state of Oregon including traffic, roadway, and crash data. Using observed crash history as a reference, the performance of the proposed approach was compared to two of the well-established methods in practice, namely: the Empirical Bayes (EB) and the Potential for Safety Improvement (PSI) methods. The study results suggest that, using crash density for highway segments, the performance of the proposed method was lower than that of the EB and PSI methods. This is despite the high R-square value of the predictive model used in the proposed method. However, when using crash frequencies for highway segments, the performance of the proposed method was found comparable to the well-established EB and PSI methods.

Keywords: rural roads safety; network screening; highway safety improvement programs; ranking; regression analysis

1. Introduction

Rural roads constitute an integral part of the roadway network providing vital access to rural towns and communities (including farms and ranches). According to the fatality facts 2021, around 40% of traffic fatalities took place in rural areas although only 20% of people in the U.S. live there, and 32% of all vehicle miles traveled (VMT) occurred in rural areas [1]. These statistics demonstrate the need for improving rural road safety and the requirement to effectively include it in the ongoing state safety improvement programs.

Network screening, the first task in road safety management, involves a low-cost examination of the entire road network to identify sites with high potential for safety improvement. The identified sites, also known as blackspots or hotspots, are prioritized for detailed engineering studies that can identify crash patterns, contributing factors, and potential countermeasures [2,3]. Errors in ranking the sites may produce a large number of false negatives (i.e., sites needing safety improvement mistakenly considered as safe) and a large number of false positives (i.e., truly safe sites identified as needing safety improvement). These errors result in an inefficient use of the resources dedicated to safety improvements and eventually reduce the overall effectiveness of the safety management programs. Therefore, the correct identification of safety improvement sites is essential for the successful implementation of any highway safety program.

Most traditional network screening methods rely on the crash frequency, crash rate, crash severity, or any combination of these measures [4,5]. By relying on criteria derived from historical crash data only [6], most existing ranking techniques are intrinsically reactive [7], often requiring crashes to occur before hotspots can be identified. On the other hand, several proactive approaches have been proposed and/or used that require detailed data on roadway geometry and roadside characteristics, besides crash and traffic data [8]. A major challenge for network screening on rural roads is lack of access to the required crash, roadway, and traffic data at the network level especially for roadways that are owned and operated by local agencies.

This research seeks to evaluate the performance of a new, yet simplistic approach for network screening that was tailored for use on rural roadways [9]. The proposed method requires minimal data for implementation which could be valuable for agencies having limited access to detailed databases or extensive technical expertise.

2. Background

Numerous network screening techniques have been used in practice or proposed in the literature with each having its own strengths and limitations. Compared with the large number of studies focused on the development of various hotspot identifications methods, considerably less research has been dedicated to evaluating the performance of various methods [10]. In this section, the major studies that evaluated the performance of different network screening techniques are summarized and presented.

Kwon et al. [11] evaluated the performance of three network screening methods for detecting high collision concentration locations on highways: Sliding Moving Window (SMW), Peak Searching (PS), and Continuous Risk Profile (CRP). The study assessed the performance of the three methods for segmenting freeway sites to identify high-collision concentration locations. The traffic collision data from sites segmented in each method were used to estimate excess expected average crash frequency using Empirical Bayes adjustment with respect to two different sets of Safety Performance Functions (SPFs). The estimates from each method were then used to prioritize sites for safety investigation, comparing them with confirmed high collision concentration locations. The findings indicated that the CRP method exhibited the lowest false positive rate, accurately identifying sites requiring safety investigation.

A study by Ambrose et al. [12] investigated the difference between network screening results based on multivariate and simple crash prediction models. This study first compared the list of segments using Spearman's rank correlation coefficient between the two models. Secondly, this study assessed the equality of statistical distributions of potential for safety improvement (PSI) value, which was obtained as a difference between predicted crash frequency and EB estimate. The third comparison was the percentage of segments identified in both lists. The findings suggested that the results from both methods were comparable.

Another study in Brazil [13] presented a comparative analysis of safety performance measures, considering its applicability limitations using a sample of signalized intersections from Fortaleza City, Brazil. The safety performance measures assessed in this study are average crash frequency (ACF), crash rate (CR), Equivalent Property Damage Only (EPDO), Level of Service of Safety (LOSS), Excess Predicted Average Crash Frequency using Safety Performance Functions (SPFs), Expected Average Crash Frequency with EB Adjustments (EB), and Excess Expected Average Crash Frequency with EB Adjustment (EEB). The rank difference between each safety performance measure and the Excess Expected Average Crash Frequency with EB Adjustment (EEB) was used to determine how well each measure ranked the sample intersection. The outcome indicated that the most comprehensive safety performance measure EEB and the basic safety performance measures like crash frequency and crash rate reasonably matched.

The effectiveness of several network screening techniques was examined in an Italian study [14] using a set of criteria. These included reliability in detecting hazardous sites over a period, efficiency in finding sites with poor safety performance, and consistency in ranking. The potential for improvements (PFI) was examined along with the crash frequency, equivalent property damage only

(EPDO), crash rate, proportion method, and EB estimates for total and severe crashes. This analysis used crash data from an Italian highway and found that the EB method outperformed all other approaches investigated in this study.

A Federal Highway Administration (FHWA) study [15] examined four network screening techniques to identify places with the most potential for safety improvements. The four performance measures included crash frequency, crash rate, EB expected crashes, and EB excess expected crashes – all for the fatal and non-fatal injury severity levels. Utilizing intersection data from New Hampshire, the study tracked the safety management process from network screening to economic analysis. This study analyzed the overall economic benefit and benefit-cost ratio for each of the four techniques. The list of sites with the most significant overall economic benefit and the highest return on investment was generated using the EB excess expected measure and the EB expected measure, respectively.

Elvik [16] has presented an extensive review for the state-of-the-art approaches of road crash hot spot identification and concluded that the Empirical Bayes approach (EB) was the most reliable method. Moreover, using simulation experiments [10] and innovative robust evaluation criteria [17], Cheng and Washington, have also proved that the EB approach is the most consistent and reliable method for identifying sites with promise.

Although the EB method was found to better identify sites requiring safety improvement, it may be subject to several limitations under certain conditions, especially for local agencies. The EB method requires detailed data for roadway geometry and roadside characteristics besides technical expertise that is often inaccessible by local agencies (counties, towns, tribal governments, etc.).

3. Proposed Network Screening Approach: An Overview

This study validates a new network screening approach that was recently proposed for use on rural two-lane highways [9]. The main merit of the proposed approach is that it can easily be implemented by local agencies lacking access to detailed databases and technical expertise. Specifically, the more sophisticated network screening methods (e.g., Empirical Bayes method) use exact values for various roadway characteristics (sometimes referred to as risk factors). While the use of exact values may improve the accuracy of the screening process, it requires access to extensive databases or on-site detailed measurements which are typically beyond the resources available for small local agencies. Hence, implementing such sophisticated methods might not be feasible for most local agencies. To minimize data requirements, the proposed approach employs classified variables that can easily be compiled by local agencies without the need to access detailed and extensive databases. Using the classified variables, the proposed method consists of regression models that are developed using the EB expected number of crashes to predict the level of risk (or safety) of roadway segments that are part of the roadway network. Two models were proposed, one with and one without traffic data. The response (dependent) variable in both models was the EB expected number of crashes, which is a function of the HSM predicted number of crashes and the observed number of crashes. The explanatory (independent) variables included roadway and roadside characteristics besides traffic exposure (AADT). Table 1 shows the explanatory variables for the proposed models.

Table 1. Explanatory Variables of Proposed Model [9].

Risk Factors	Approximate Ranges of Variables	Categories	Terms
Segment Length (SL)	Exact Length		
Lane Width (LW)	LW < 11	1	Narrower
	LW ≥ 11	2	Wider

Shoulder Width (SW)	SW < 1.8	1	Narrower
	SW ≥ 1.8	2	Wider
Degree of Horizontal Curvature (DC)	DC = 0	0	Straight
	DC < 10	1	Mild
	10 ≤ DC < 27	2	Moderate
	DC ≥ 27	3	Sharp
Grade (G)	G < 4	0	Mild
	G ≥ 4	1	Steep
Driveway Density (DD) (driveways per mile)	Exact Number		
Side Slope (SS)	Steep	1	Steep
	Moderate	2	Moderate
	Flat	3	Flat
Fixed Objects (FO)	Many	1	Many
	Some	2	Some
	Few	3	Few
Volume (V)	Exact Volume		

4. Study Area

The two-lane rural road network in the state of Oregon was investigated as a case study, to validate the proposed method. A total roadway segment sample of around 1495 miles from eastern and western region was used in this investigation to ensure adequate geographic coverage. The study sample included roadways from different parts of the state as shown in Figure 1. In general, roadways in the eastern part of the state run in flatter terrain with less restrictive alignment while those in the western part of the state run in mountainous terrain with more restrictive alignment (winding routes, sharp radii, etc.).

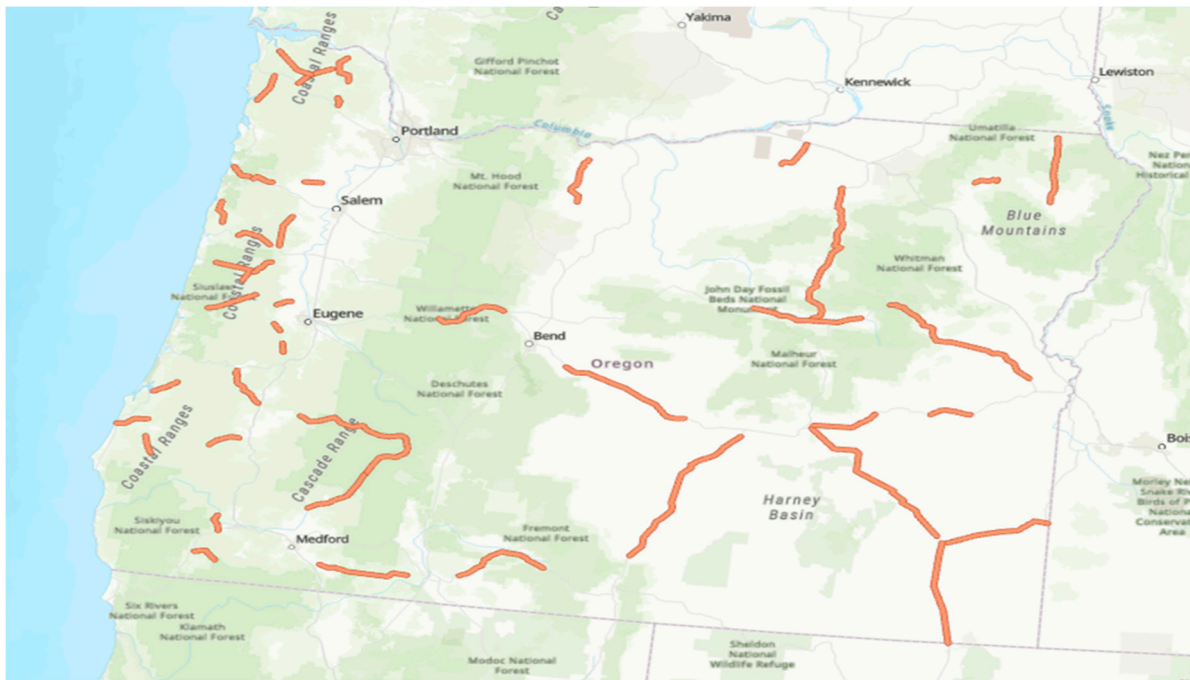


Figure 1. Study Area Showing Sampled Highway Segments.

5. Data Description

All state-owned rural two-lane roads were identified using online geographic information system (GIS) data [18]. Then selections were made from the results to get sample data considering geographical coverage. Data were collected for roadway segments using 0.05-miles increments to ensure that data would capture all changes in the physical characteristics of the roadway, thus eliminating the possibility of missing significant differences between consecutive observation points. All segments in the study sample had a posted speed limit of 55mph. Intersections and 0.05-mile segments along upstream and downstream approaches were excluded from the dataset.

The Oregon Department of Transportation (ODOT) online databases and video logs were used to identify and compile traffic, roadway geometry, roadside, and crash data [19,20]. The roadway geometric data consisted of lane width, shoulder type and width, horizontal curve presence degree of curvature, length of the horizontal curve, spiral curve presence and length, vertical curve presence and type, grade, and length of vertical curve taken from the online database. Roadside characteristics comprised driveway density, side slope ratings, and fixed object ratings taken from ODOT video logs at 0.05-mile resolution.

Crash data from 2011-2020 were collected for the study sample using ODOT online database. Data on more than 20 individual crash characteristics- including crash location, road character, impact location, traffic control, crash type, crash severity, vehicle type, and weather condition- were collected and combined with the geometric and roadside database forming an integrated dataset for analysis.

Traffic data for 10 years (2011-2020) were also collected separately and added to the study integrated dataset described earlier.

The study sample was shared with the Oregon DOT personnel to learn about any changes that may have taken place on any of the sample segments. The information received from the agency confirmed that there were no changes for the study period.

5.1. Segmentation

As mentioned earlier, the data was collected using 0.05-mile increments. Afterwards, the total sample was compiled into homogeneous segments concerning the following variables: annual average daily traffic (AADT), speed limit, lane and shoulder type and width, and roadside

characteristics. Any change in these variables marked the end of a segment and the beginning of another segment.

Upon completing the segmentation, a total of 377 segments consisting of 1495-miles were compiled. Crash frequency data for two different periods (2011-2015, and 2016-2020) were compiled for roadway segments.

6. Methodology

The proposed approach was implemented using the study data and evaluated using observed crash history collected for the study sample. Specifically, the crash data collected over ten years was divided into two datasets: a training dataset spanning five years (2011-2015) to develop the regression model, and a testing dataset covering the subsequent five-year period (2016-2020). The approach that was used in evaluating the proposed network screening method is to compare site rankings using the proposed method with the respective rankings using crash history data. Both crash frequency and crash density (number of crashes per mile) were used in ranking sites. Further, rankings from the Empirical Bayes (EB) method and the potential for safety improvement (PSI) method were included in the analyses to see how the performance of the proposed method compares to that of the two well-established techniques. Using the EB and PSI in this study is based on the fact that these techniques has become the gold standard in performing quantitative safety analyses, with merits and effectiveness well established in the literatures [21-27].

6.1. Model Development

Multivariate linear regression analysis was used in developing the model for road network screening. The open-source statistical software R was used for running the analysis. The model with different explanatory variables is shown in Equation 1.

$$\ln \left(\frac{\text{crashes}}{\text{year}} \right) = -8.27049 + 0.72962 * \frac{1}{FO} + 0.02717 * DD + 0.98309 * \ln(V) + 0.10126 * DC + 0.94290 * \ln(SL) \quad (1)$$

Where, exp = EB expected number of crashes, SL = Segment length in miles, V = Traffic volume (AADT), SW = Shoulder width, DD = driveway density, DC = Degree of curvature, SS = Side slope

This model has a coefficient of determination (adjusted R-squared) of 0.932. This indicates that the model can explain about 93 percent of the variability of the EB expected number of crashes, which is relatively high given the classified format used for most of the variables in this model. All variable coefficients are found significant at the 95% confidence level. The regression output for the first model is shown in Figure 2.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.270493   0.195408 -42.324 < 2e-16 ***
`1/FO`       0.729629   0.186246   3.918 0.000106 ***
DD           0.027170   0.003241   8.384 1.07e-15 ***
`Ln (v)`     0.983096   0.021492  45.743 < 2e-16 ***
DC           0.101260   0.019409   5.217 3.03e-07 ***
`Ln (SL)`    0.942908   0.017038  55.340 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3178 on 371 degrees of freedom
Multiple R-squared:  0.9325,    Adjusted R-squared:  0.9316
F-statistic: 1025 on 5 and 371 DF,  p-value: < 2.2e-16

> coef(regression_models)
(Intercept)      `1/FO`          DD      `Ln (v)`          DC      `Ln (SL)`
-8.27049313   0.72962895   0.02717009   0.98309595   0.10126046   0.94290773

```

Figure 2. R-output of Regression Model.

7. Study Analysis and Results

This section presents the discussion of the metrics used for the evaluation of the performance of the proposed method. The performance of the EB method and the PSI method is also calculated and compared with the proposed method.

7.1. Spearman Correlation Coefficient

Initially, the spearman rank correlation coefficient was calculated between the proposed method's ranking and the ranking using observed crash data. Spearman rank correlation coefficient is a nonparametric technique that is usually applied to evaluate the degree of linear association between two independent variables [28]. Here, this coefficient is used to measure the degree of association between two lists of hazardous sites ordered based on two ranking criteria. The simple expression for ' ρ ' based on the difference between the two ranked variables is as shown in Equation 2:

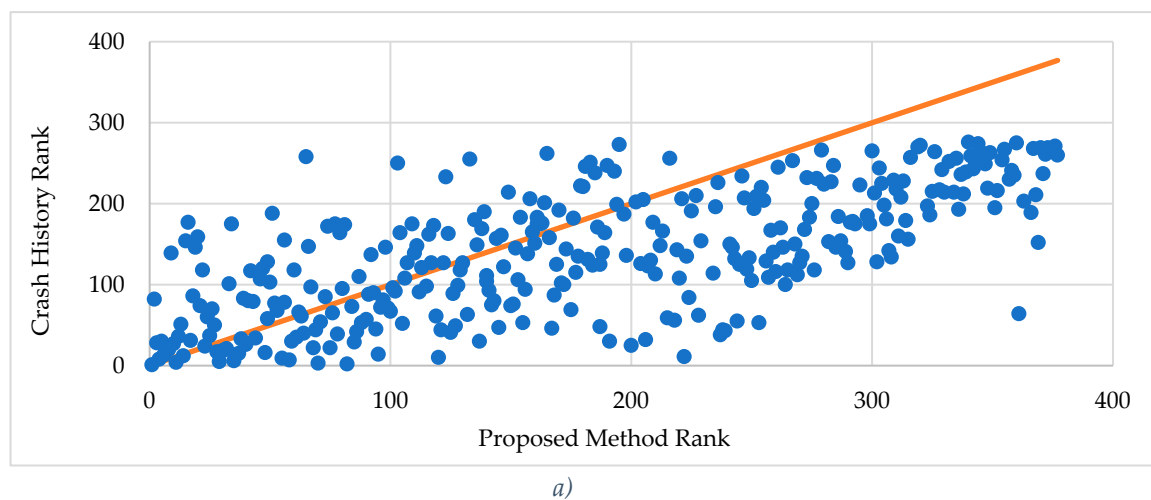
$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} \quad (2)$$

where, ρ = spearman's rank correlation coefficient; $d_i = R(X_i) - R(Y_i)$ is the difference between the two ranks of each observation; n = number of observations.

Figure 3 shows scatterplot with site rankings using the proposed method, EB method, and PSI method versus crash history. An examination of Figure 3a, which is proposed method versus crash history rank, reveals that the data points are spread around the diagonal line. It shows only some segments are sharing same rank by two methods. The discrepancy between the ranks increases with the increase in ranks. The tightness of the data points around the line and the correlation coefficient of 0.686 indicates a moderate correlation between the ranks by the two methods.

The EB method ranking was also compared with the ranking using the observed number of crashes and the results are shown in Figure 3b. In this figure, the proximity of data points to the line passing through the origin (diagonal line) indicates that it is more compact compared to that of the proposed method. This observation is confirmed by a higher correlation coefficient ($r=0.816$) between the two rankings.

Additionally, Figure 3c shows the comparison of ranking using PSI method versus the ranking using the observed number of crashes. The data points clustered around the diagonal line and the correlation coefficient of 0.879 indicate the very strong association between the ranks.



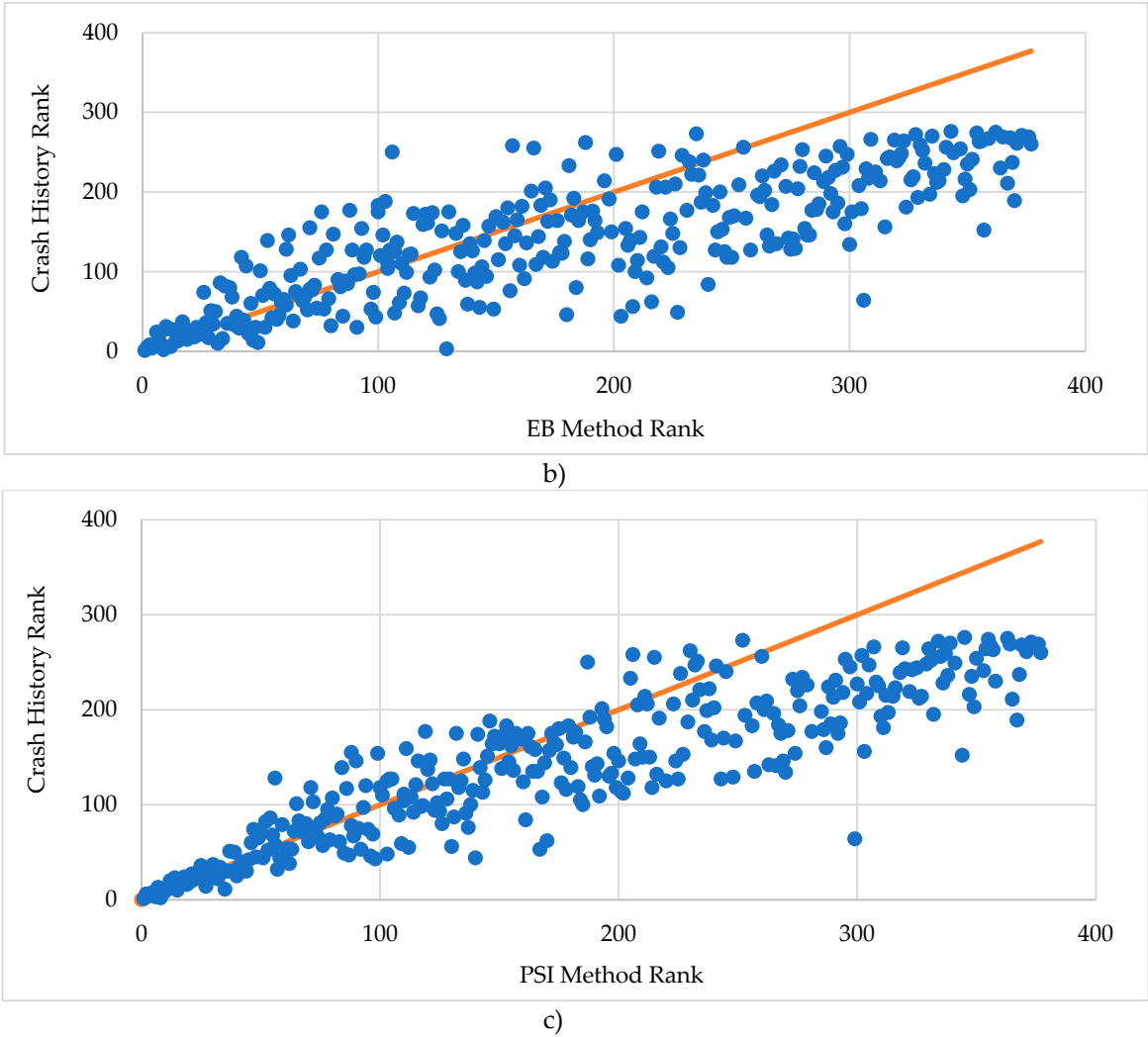


Figure 3. Rank comparison using proposed method, EB method, and PSI method versus crash history.

Table 2 presents the Spearman's Rank Correlation Coefficient values for three different methods (Proposed Method, Empirical Bayes Method, and PSI Method) based on crash density and crash frequency in different upper tail segments and for the total sample size of 377 segments. In analyzing crash density for the upper tail segments (20, 40, 60, 80, 100), the proposed method shows moderate to strong positive correlations with crash history, with values ranging from 0.587 to 0.744. The Empirical Bayes method and PSI method consistently exhibit higher correlations, ranging from 0.786 to 0.828 and 0.954 to 0.959, respectively. For the total sample size of 377 segments, the proposed method exhibits a moderate positive correlation (0.683), while the Empirical Bayes method (0.816) and PSI method (0.879) display stronger positive correlations. This shows that, in analyzing crash densities, the Empirical Bayes and PSI methods outperform the proposed method at identifying safety improvement sites.

In the case of crash frequency, for the upper tail segments (20, 40, 60, 80, 100), the proposed method shows weak to moderate positive correlations, ranging from 0.502 to 0.703. The Empirical Bayes Method exhibits overall stronger positive correlations, varying in a broad range from 0.406 to 0.840, while the PSI Method consistently demonstrates the highest correlations, ranging from 0.738 to 0.893. In the total sample, the proposed method shows a relatively strong positive correlation (0.815), whereas the Empirical Bayes method (0.928) and PSI method (0.942) display even stronger positive correlations with crash history. This indicates a higher association between all three methods and the crash history when the total sample is analyzed.

Table 2. Comparative Analysis of Rank Correlation for Proposed Method versus the EB Method versus the PSI Method.

Groups/ Methods	Spearman's Rank Correlation Coefficient		
	Proposed Method	EB Method	PSI Method
Crash Density Based Analysis			
Upper Tail (20)	0.744	0.815	0.958
Upper Tail (40)	0.587	0.786	0.954
Upper Tail (60)	0.638	0.813	0.959
Upper Tail (80)	0.655	0.828	0.959
Upper Tail (100)	0.605	0.825	0.958
Total Sample (377)	0.683	0.816	0.879
Crash Frequency Based Analysis			
Upper Tail (20)	0.502	0.406	0.738
Upper Tail (40)	0.562	0.768	0.878
Upper Tail (60)	0.65	0.84	0.893
Upper Tail (80)	0.703	0.823	0.87
Upper Tail (100)	0.702	0.829	0.867
Total Sample (377)	0.815	0.928	0.942

7.2. True Positive Identification

Secondly, the identification of true positive segments is calculated. A true positive occurs when the proposed method correctly identifies a segment as a potential for safety improvement. This evaluation listed the number and percentage of common segments between two lists: the one compiled using ranks from the proposed method and the other using ranks from observed crash data. The higher the number and percentage of the common sites, the better the performance of the subject method.

Table 3 shows a comparison of the identification of true positive segments by three different methods—the proposed method, the Empirical Bayes (EB) method, and the PSI method—while using observed crash data as a reference in the context of crash density and crash frequency analysis for various upper tail segments.

When evaluating the identification of common segments using crash density, for upper tail segments (20), the proposed method was least consistent with crash history (30%), followed by the EB method (65%) and the PSI method (90%), respectively. With the increase in upper tail segments, the discrepancy between the proposed method and the EB and PSI methods decreases. However, the performance of the proposed method is overall less favorable compared to that of the other two

methods. For example, in upper tail segments (100), the proposed method identified 68 sites versus 78 sites for the EB method and 87 sites for the PSI method.

When using crash frequency in the analysis, show better performance for the proposed method and lower discrepancy between the three methods. For instance, the proposed method identified 78% of true positive segments compared to 86% by the EB method and 88% by the PSI method when considering 100 upper tail segments. Further, for every other upper tail considered, results from the proposed method shows lower discrepancy with those using the other two methods.

Table 3. True Positive Segments Identification Comparison for Proposed Method versus the EB Method versus the PSI Method.

Groups	% of segments	# of Common Segments			% of Common Segments		
		Proposed Method	EB Method	PSI Method	Proposed Method	EB Method	PSI Method
Crash Density Analysis							
Upper Tail (20)	5.31%	6	13	18	30	65	90
Upper Tail (40)	10.61%	22	31	37	55	77.5	92.5
Upper Tail (60)	15.92%	30	45	51	50	75	85
Upper Tail (80)	21.22%	46	60	64	57.5	75	80
Upper Tail (100)	26.53%	68	78	87	68	78	87
Crash Frequency Analysis							
Upper Tail (20)	5.31%	12	17	17	60	85	85
Upper Tail (40)	10.61%	30	34	34	75	85	85
Upper Tail (60)	15.92%	41	49	50	68.33	81.67	83.33
Upper Tail (80)	21.22%	61	68	70	76.25	85	87.5
Upper Tail (100)	26.53%	78	86	88	78	86	88

7.3. Precision and Percent Deviation

The precision and percent deviation from crash history are calculated for various upper tail segment groups (20, 40, 60, 80, and 100) for the proposed method and the performance is compared with that of the EB and the PSI methods. A higher percentage of precision and lower value of percent deviation indicates better performance.

$$\text{Precision} = \frac{TP}{TP+FP} * 100\% \quad (3)$$

$$\text{Percent Deviation} = (1 - TP/M) * 100\% \quad (4)$$

Where, TP = true positive is the number of sites correctly identified as safety improvement sites by the evaluating method.

FP = false positive is the number of sites falsely identified as safety improvement sites by the evaluating method (not listed in crash history).

M is the number of upper tail segments that are considered.

Figure 4 shows precision across various upper tail groups (20, 40, 60, 80, and 100). Precision is also calculated for the EB method, and the PSI method to see how the performance of the proposed method compares to that of the other two methods. The precision percentage based on crash density analysis is shown in Figure 4a. The figure shows that the precision value is less for the proposed method when compared with the EB and the PSI methods. The value for the proposed method varies from 30% to 68%, versus 62% to 76% for the EB method and 85% to 90% for the PSI methods. On the other hand, the precision percentage based on crash frequency shown in Figure 4b indicates that the precision value varies from 60% to 78% for the proposed method versus 81.6% to 86% for the EB method and 83.3% to 88% for the PSI method. This confirms that both the EB and PSI methods were more consistent with the crash history compared with the proposed method. Further, the proposed method exhibited better performance and higher consistency with crash history when analyzed using crash frequency over crash density.

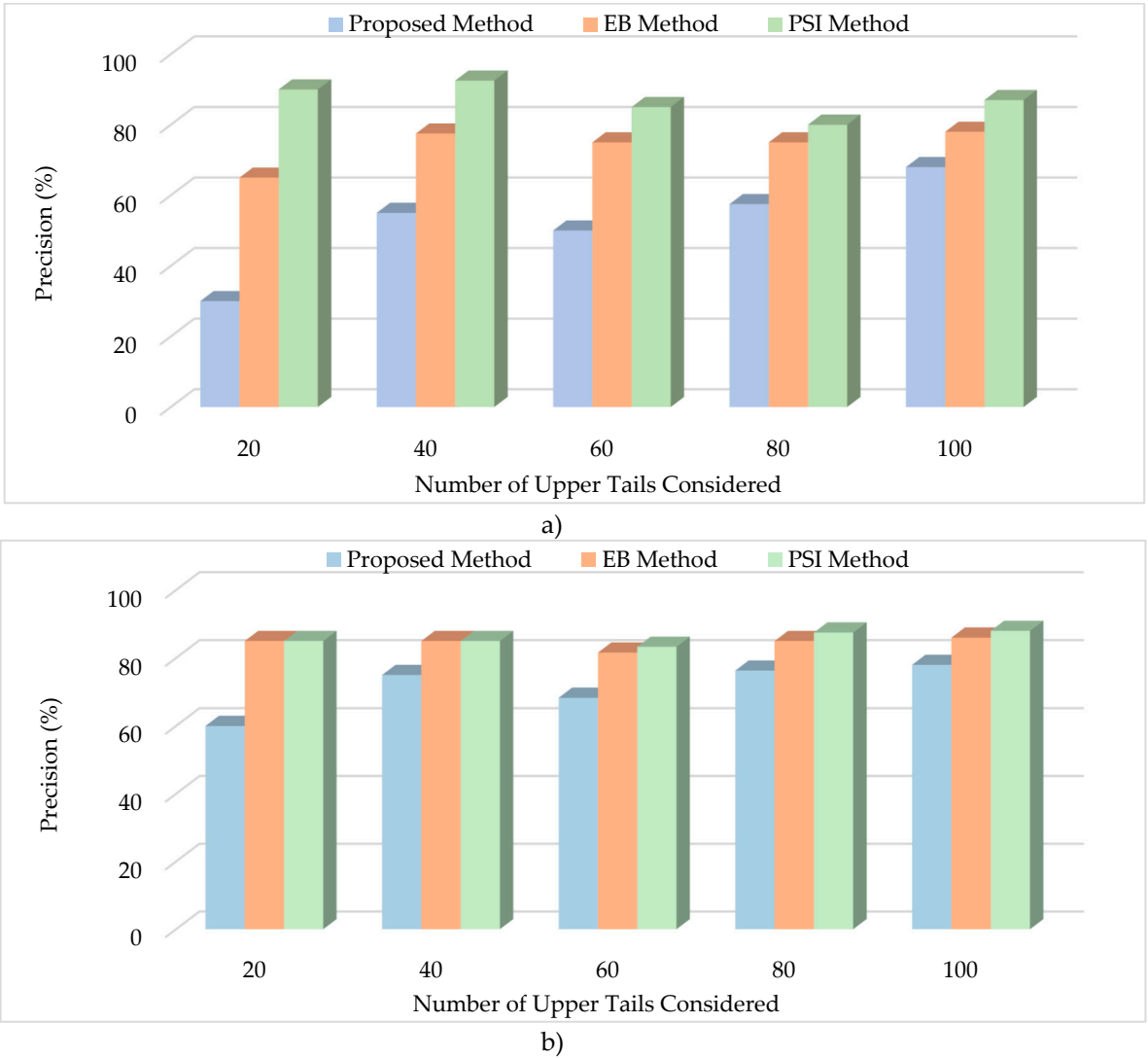


Figure 4. Precision comparison using proposed method, EB method, and PSI method versus crash history a) crash density-based analysis b) crash frequency-based analysis.

Figure 5 shows percent deviation from the crash history for the proposed method, the EB method, and the PSI method. A quick examination of Figure 5 clearly shows higher deviation for the proposed method compared to the EB and the PSI methods both for crash density and crash frequency analyses. The other trend clearly exhibited in this figure is that the percent deviation for the proposed method decreases with the increase in the number of upper tail segments. Specifically,

the percent deviation of 70% for density-based analysis and 40% for frequency-based analysis were observed for the 20 upper tail segments. However, the corresponding percent deviation for the 100 upper tails is 32% and 22% for density-based and frequency-based analyses respectively. The EB and the PSI methods, while generally showing less deviation and better performance, exhibited less variation in percent deviation with the increase in upper tail segments.

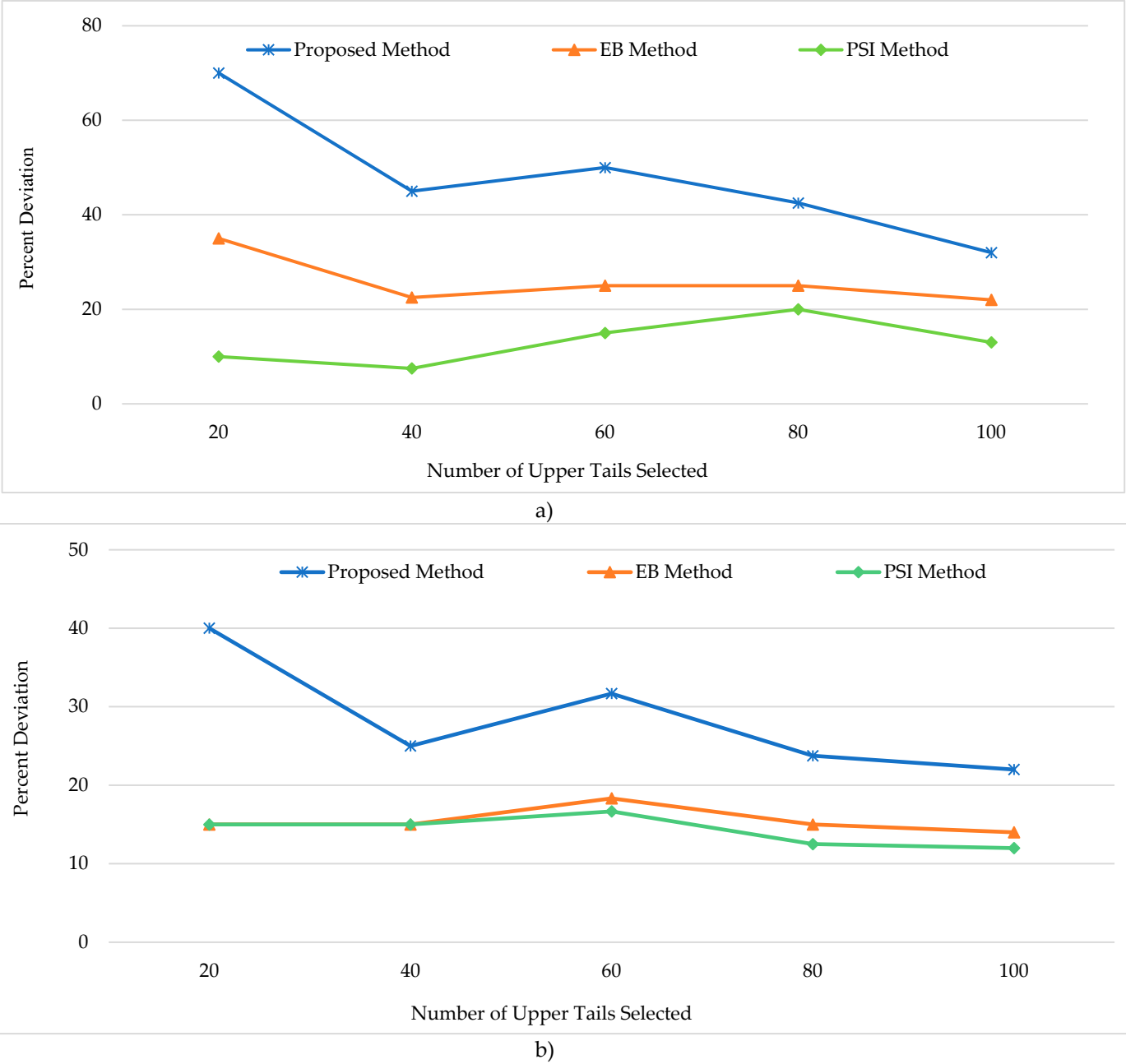


Figure 5. Percent deviation comparison using proposed method, EB method, and PSI method versus crash history a) crash density-based analysis b) crash frequency-based analysis.

7.4. Treatable Crashes

The last analysis in the evaluation of the proposed method involved the use of the number of treatable crashes (calculated for five years 2016-2020) as the performance measure. Treatable crashes are defined as those crashes that may be treated by engineering countermeasures such as lane widening. The remaining crashes constitute what might be expected due to traffic exposure and may not be treatable by engineering countermeasures. A reasonable estimate for the number of treatable crashes at a location is the number of crashes in excess of what would normally be expected at

locations with similar characteristics and traffic exposure [29]. To assess the degree of similarity between the segments identified by a method and those identified by the crash history, a segment overlap metric was used.

Table 4 shows the summary of the estimate of treatable crashes for the proposed method the EB method, and crash history. The table also provides the ratio and overlap for treatable crashes in the top 100 segments defined earlier using both crash density and crash frequency.

The table shows that, based on crash density, the proposed method identified a fewer number of treatable crashes for the top 100 segments compared to crash history (469 versus 726 respectively). The segment overlap was 66%. Using crash frequency, the segment overlap increased to 79% indicating a significant improvement in performance. Regarding the EB method, the number of treatable crashes and overlap indicate a higher level of consistency with crash history. However, the discrepancy in performance between the proposed and the EB methods is less evident when crash frequency is used in the analysis. Specifically, there were only 11.6% fewer treatable crashes on the segments identified with the proposed method than those identified with the crash history, compared to 11.3% fewer treatable crashes using the EB method. Further, 79% of all segments identified by the proposed method were also on the top 100 segments list of the observed crash.

Table 4. Treatable Crashes for the top 100 segments identified.

Methods	Total Length (mi)	Total Crashes (2016-2020)	Treatable Crashes (2016-2020)	
			Estimate	Overlap
Crash Density Based Analysis				
Proposed Method	260.55	565	469	66%
EB	265	614	522	81%
Crash History	241.1	805	726	-
Crash Frequency Based Analysis				
Proposed Method	807.45	2220	1147	79%
EB	787.05	2144	1152	85%
Crash History	805.75	2439	1299	-

In addition to evaluating the top 100 segments identified by the methods, similar calculations were performed for other possible levels of selection up to upper tails having positive treatable crashes. Plots of the treatable crashes identified at various levels are shown in Figure 6. For the range of top segments studied, the chart shows relatively close lines for the proposed and the EB methods, with the EB method consistently identifying a slightly higher number of treatable crashes. The figure also suggests that the larger the total segments selected, the closer the proposed method is in terms of selection efficiency compared to that of EB method. These patterns are similar for the crash density and crash frequency analyses.

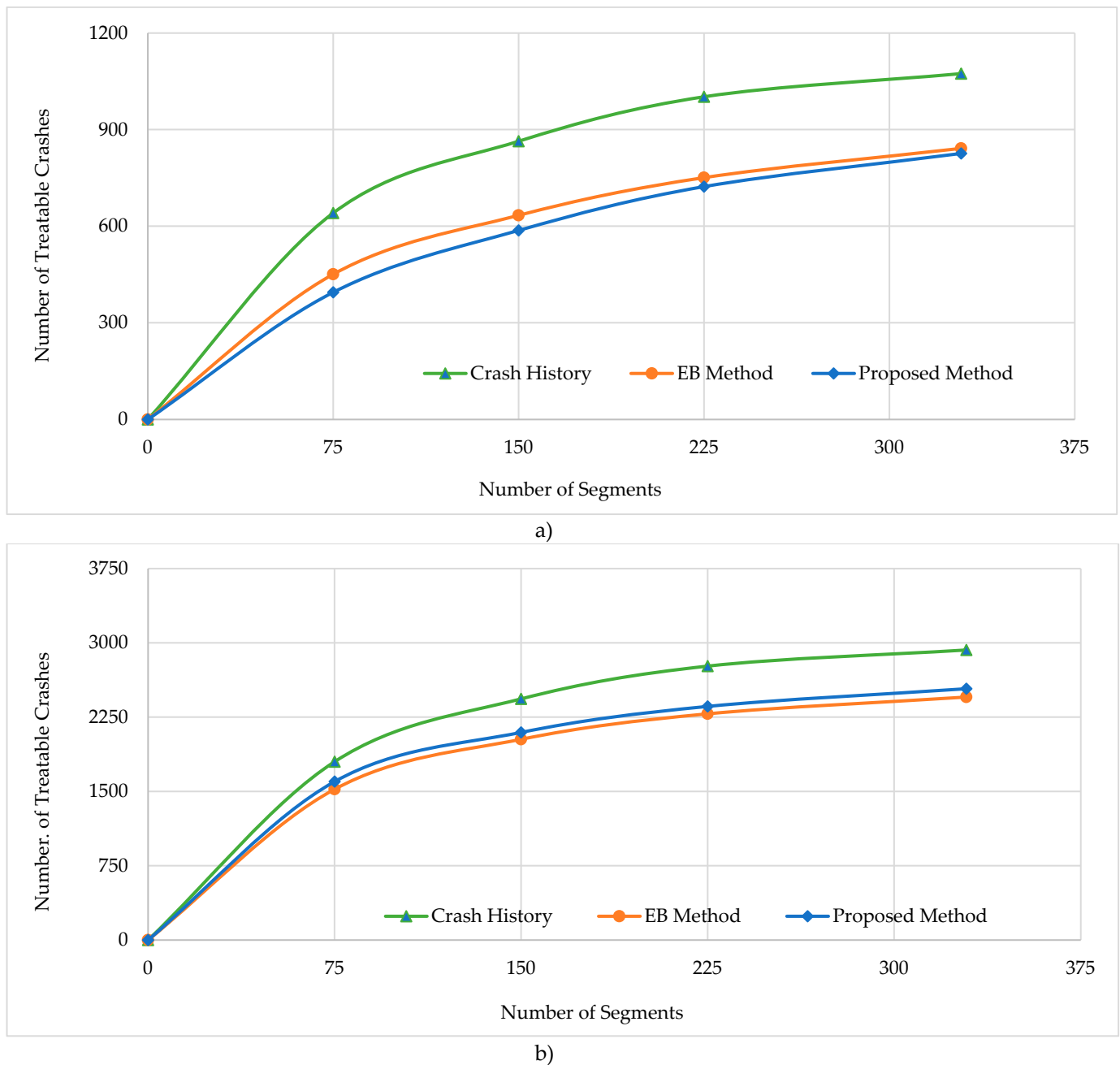


Figure 6. Treatable Crash Count by Different Methods and Total Number of Segments Selected a) crash density-based analysis b) crash frequency-based analysis.

8. Summary and Conclusions

Network screening is a critical initial step of the Highway Safety Improvement Programs (HSIP) which aims to identify sites with potential for safety improvement. Ineffective network screening can lead to a significant number of false positives and/or false negatives resulting in an inefficient allocation of agency resources and consequently impacting the overall efficacy of safety management programs. Hence, accurate identification of sites that are in need of safety improvements is crucial for the success of any highway safety program.

The research presented in this paper examined a new proposed approach for identifying safety improvement sites on rural highways. The proposed approach involves the use of EB prediction models using traffic and roadway variables only. Specifically, the proposed approach utilizes classified variables for traffic volume, geometric, and roadside characteristics that neither requires exact data nor extensive technical expertise. The evaluation was performed using a dataset

comprising 1495 miles of rural two-lane highway segments in the state of Oregon. The data used in the evaluation included roadway geometry, roadside features, traffic conditions, and ten years of crash records. A training dataset spanning five years (2011-2015) was employed to develop the model, while the subsequent five-year period (2016-2020) served as the basis for evaluating the model's performance in network screening.

The study findings suggest that using crash density for highway segments, the performance of the proposed method was fair and not as effective as the well-established EB and PSI methods. This is despite the high R-square value of the predictive model used by the proposed approach. However, when using crash frequencies for highway segments, the performance of the proposed method was found to be comparable to the EB and PSI methods. Moreover, the results suggest that, in identifying segments most likely to have treatable crashes, the proposed method outperformed the EB method for higher upper tail segments considered in the analysis.

Overall, despite the practical merits of the proposed approach, its effectiveness in identifying safety improvement sites is modest given that many highway agencies use crash densities for highway segments in safety analyses.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Ahmed Al-Kaisy; data collection: Bishal Dhakal; analysis and interpretation of results: Bishal Dhakal and Ahmed Al-Kaisy; draft manuscript preparation: Bishal Dhakal and Ahmed Al-Kaisy. All authors reviewed the results and approved the final version of the manuscript.

Funding: This work was supported by the US Department of Transportation (USDOT) Small Urban, Rural, and Tribal Center on Mobility (SURTCOM) at the Western Transportation Institute of Montana State University (grant # 4W9679).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fatality Facts 2021: Urban/Rural Comparison Available online: <https://www.ihs.org/topics/fatality-statistics/detail/urban-rural-comparison> (accessed on 20 July 2023).
2. Hauer, E.; Kononov, J.; Allery, B.; Griffith, M.S. Screening the Road Network for Sites with Promise. *Transp. Res. Rec.* **2002**, *1784*, 27–32, doi:10.3141/1784-04.
3. Schmelkin, E.J.P., Liora Pedhazur *Measurement, Design, and Analysis: An Integrated Approach*; Psychology Press: New York, 1991; ISBN 978-0-203-72638-9.
4. Park, P.Y.; Sahaji, R. Safety Network Screening for Municipalities with Incomplete Traffic Volume Data. *Accid. Anal. Prev.* **2013**, *50*, 1062–1072, doi:10.1016/j.aap.2012.08.012.
5. Anastasopoulos, P.Ch.; Mannering, F.L. An Empirical Assessment of Fixed and Random Parameter Logit Models Using Crash- and Non-Crash-Specific Injury Data. *Accid. Anal. Prev.* **2011**, *43*, 1140–1147, doi:10.1016/j.aap.2010.12.024.
6. Huang, H.; Chin, H.C.; Haque, Md.M. Empirical Evaluation of Alternative Approaches in Identifying Crash Hot Spots: Naive Ranking, Empirical Bayes, Full Bayes Methods. *Transp. Res. Rec.* **2009**, *2103*, 32–41, doi:10.3141/2103-05.
7. Agerholm, N.; Lahrmann, H. Identification of Hazardous Road Locations on the Basis of Floating Car Data: Method and First Results.; December 15 2012.
8. Lee, C.; Hellinga, B.; Ozbay, K. Quantifying Effects of Ramp Metering on Freeway Safety. *Accid. Anal. Prev.* **2006**, *38*, 279–288, doi:10.1016/j.aap.2005.09.011.
9. Al-Kaisy, A. and Huda, K. T. Network Screening on Low-Volume Roads: A New Proposed Method. Presented at the 100th Transportation Research Board Annual Meeting, Washington, DC, 21-29 January 2021.
10. Cheng, W.; Washington, S.P. Experimental Evaluation of Hotspot Identification Methods. *Accid. Anal. Prev.* **2005**, *37*, 870–881, doi:10.1016/j.aap.2005.04.015.
11. Kwon, O.H.; Park, M.J.; Yeo, H.; Chung, K. Evaluating the Performance of Network Screening Methods for Detecting High Collision Concentration Locations on Highways. *Accid. Anal. Prev.* **2013**, *51*, 141–149, doi:10.1016/j.aap.2012.10.019.

12. Ambros, J.; Valentová, V.; Janoška, Z. Investigation of Difference between Network Screening Results Based on Multivariate and Simple Crash Prediction Models.; Transportation Research Board, January 13 2015.
13. Mesquita Xavier, V.J.; Craveiro Cunto, F.J. Comparative Analysis of Performance Measures for Network Screening: A Case Study of Brazilian Urban Areas. *J. Adv. Transp.* **2017**, *2017*, e4360414, doi:10.1155/2017/4360414.
14. Montella, A. A Comparative Analysis of Hotspot Identification Methods. *Accid. Anal. Prev.* **2010**, *42*, 571–581, doi:10.1016/j.aap.2009.09.025.
15. Gross, F.B.; Harmon, T.; Albee, M.; Himes, S.; Srinivasan, R.; Carter, D.; Dugas, M. *Evaluation of Four Network Screening Performance Measures*; United States. Federal Highway Administration. Office of Safety, 2016;
16. Elvik, R. State-of-the-Art Approaches to Road Accident Black Spot Management and Safety Analysis of Road Networks.
17. Cheng, W.; Washington, S. New Criteria for Evaluating Methods of Identifying Hot Spots. *Transp. Res. Rec.* **2008**, *2083*, 76–85, doi:10.3141/2083-09.
18. ODOT TransGIS Available online: <https://gis.odot.state.or.us/transgis> (accessed on 5 June 2023).
19. Oregon Department of Transportation : Road Assets and Mileage : Data & Maps : State of Oregon Available online: <https://www.oregon.gov/odot/Data/pages/road-assets-mileage.aspx> (accessed on 5 June 2023).
20. TDS - Crash Reports Available online: <https://tvc.odot.state.or.us/tvc/> (accessed on 5 June 2023).
21. Cafiso, S.; Di Silvestro, G.; Persaud, B.; Begum, M.A. Revisiting Variability of Dispersion Parameter of Safety Performance for Two-Lane Rural Roads. *Transp. Res. Rec.* **2010**, *2148*, 38–46, doi:10.3141/2148-05.
22. Manepalli, U.R.R.; Bham, G.H. An Evaluation of Performance Measures for Hotspot Identification. *J. Transp. Saf. Secur.* **2016**, *8*, 327–345, doi:10.1080/19439962.2015.1048015.
23. Manepalli, U.R.R.; Bham, G.H. CRASH PREDICTION: EVALUATION OF EMPIRICAL BAYES AND KRIGING METHODS.
24. Elvik, R. The Predictive Validity of Empirical Bayes Estimates of Road Safety. *Accid. Anal. Prev.* **2008**, *40*, 1964–1969, doi:10.1016/j.aap.2008.07.007.
25. Hauer, E. *Observational Before–after Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*; Pergamon: Amsterdam, 2002; ISBN 978-0-08-043053-9.
26. Persaud, B.; Lyon, C. Empirical Bayes Before-after Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. *Accid. Anal. Prev.* **2007**, *39*, 546–555, doi:10.1016/j.aap.2006.09.009.
27. Hauer, E.; Persaud, B.N.; Smiley, A.; Duncan, D. Estimating the Accident Potential of an Ontario Driver. *Accid. Anal. Prev.* **1991**, *23*, 133–152, doi:10.1016/0001-4575(91)90044-6.
28. Hauke, J.; Kossowski, T. Comparison of Values of Pearson’s and Spearman’s Correlation Coefficients on the Same Sets of Data. *QUAGEO* **2011**, *30*, 87–93, doi:10.2478/v10117-011-0021-1.
29. Persaud, B.; Lyon, C.; Nguyen, T. Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement. *Transp. Res. Rec.* **1999**, *1665*, 7–12, doi:10.3141/1665-02.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.