COMPUTATIONAL RESEARCH PROGRESS IN APPLIED SCIENCE & ENGINEERING (CRPASE)

CRPASE: TRANSACTIONS OF CIVIL AND ENVIRONMENTAL ENGINEERING

Journal homepage: http://www.crpase.com

CRPASE: Transactions of Civil and Environmental Engineering 7 (1) Article ID: 2311, 1–9, March 2021 ISSN 2423-4591

Research Article

Calibration of Highway Safety Manual's Crash Prediction Model for Rural Two-Lane Two-Way Roads in a Developing Country: A Case Study

Mohamadreza Haghani[1*], Reza Jalalkamali[2], Hamidreza Haghani[3]

[1] *Civil and Environmental Engineering Department, University of Louisville, Louisville, KY, USA*
[2] *Faculty of Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*
[3] *Sirjan University of Technology, Sirjan, Iran*

Keywords

Rural Road Safety, Two-Lane Two-way Roads, Highway Safety Manual, Crash Prediction Models, Crash Models Calibration.

Abstract

Road safety has attracted much attention in the last decades owing to the high fatality rate of road crashes. The high speed of drivers on rural roads results in more severe crashes occurring on this type of roads. Therefore, the safety of rural roads was analyzed in this study. Roads with different topographic types were chosen for data collection in Kerman province, Iran. Roadway characteristics were collected from field surveys and documentations. Also, crash records of the studied roads were gathered for a three-year period. Then, the Highway Safety Manual's (HSM) crash prediction model was calibrated for the study area. The performance of the HSM' model was compared to the jurisdictional crash prediction model created in a previous research using the data from the same region. Also, the Empirical Bayes (EB) method was utilized to estimate expected crash frequencies. The calibration factor of 1.17 was calculated for the HSM's crash prediction model using EB. The evaluation results showed that the jurisdictional model had higher precision and lower bias compared to the HSM's model. However, utilizing EB, the performance of the HSM's model became better than the jurisdictional model. At the end, this research verified the transferability of the HSM's crash prediction model to a developing country.

1. Introduction

According to the World Health Organization, more than 1.35 million people died, and more than 50 million people were injured in road crashes that occurred in 2016. Road crash is now the eighth leading cause of death higher than HIV and Tuberculosis [1]. In 2010, governments of the world marked 2011 to 2020 as "the decade of action for road safety" [2]. This indicates the significance of road safety worldwide. Therefore, much research has been performed with the objective of improving the safety of roads in the last decade [3-10]. The Highway Safety Manual (HSM) is a useful tool to make decisions about road safety improvement with models which have the ability to estimate the frequency and severity of crashes. Since its models were developed based on the data collected only from a selection of the US states and various factors, including drivers' behavior, climate conditions, traffic conditions, and the like varies from region to region, HSM recommends the calibration of its models in accordance with each area's jurisdiction [11]. Also, developing new crash prediction models requires road, crash and traffic data collection, and much effort to process the data afterward [12-16]. Owing to this fact, jurisdictional agencies prefer to calibrate the crash prediction models presented in HSM instead of creating specific models for their localities [17]. Thus, after the publication of HSM, many researchers have performed studies on calibrating HSM's crash prediction models for their jurisdictions [18-20]. A summary of them is presented in the following section.

Since several draft versions of HSM had been developed and made available to researchers before the final version

* Corresponding Author: Mohamadreza Haghani
E-mail address: Mohamadreza.haghani@louisville.edu

Received: 6 February 2021; Revised: 18 March 2021; Accepted: 28 March 2021

Please cite this article as: M. Haghani, R. Jalalkamali, H. Haghani, Calibration of Highway Safety Manual's Crash Prediction Model for Rural Two-Lane Two-Way Roads in a Developing Country: A Case Study, Computational Research Progress in Applied Science & Engineering, CRPASE: Transactions of Civil and Environmental Engineering 7 (2021) 1–9, Article ID: 2311.

was published in 2010, researchers have conducted studies on the calibration and evaluation of crash prediction models of the aforementioned manual even before its publication [21]. One of the first studies on the calibration of the HSM's crash prediction model was performed by Sun et al. In this research, the crash prediction model of HSM's draft version was calibrated for rural two-lane two-way roads in Louisiana. A calibration factor of 1.63 was obtained [22]. The draft version model was also calibrated utilizing the data of Italia's rural roads. In addition, the calibration process was reformed, and twelve different methods were presented for calculating the calibration factor. The calibration factor of 0.37 was calculated for rural two-lane two-way roads [23]. Xie et al. aimed at calibrating the crash prediction models of all three types of highways available in HSM in a comprehensive research using highway data in Oregon. A calibration factor of 0.74 was obtained for the HSM's crash prediction model of rural two-lane two-way roads using the three-years crash data from 2004 to 2006 [24]. In another study, Banihashemi calibrated the HSM's crash prediction model for Washington States' roads. A calibration factor of 1.5 was calculated for rural two-lane two-way roads' model based on the three-years crash data of 2002 to 2004 [25]. Williamson and Zhou calibrated the HSM's crash prediction model and the jurisdictional crash prediction model of Illinois State based on the data from rural two-lane two-way roads in Illinois. The Illinois State's crash prediction model had been developed in a previous study [26]. Utilizing a three-years period of crash data from 2007 to 2009, the calibration factors of 1.4 and 1.58 were calculated, respectively, for HSM's and Illinois's crash prediction models [27].

Brimley et al. calibrated the HSM's crash prediction model for rural two-lane two-way roads in Utah. A calibration factor of 1.16 was obtained based on the crash data of three years from 2005 to 2007 [28]. In another study, Lubliner and Schrock not only calibrated the HSM's crash prediction model but also modified the HSM's calibration method. Utilizing the crash data of 2005-2007, a calibration factor of 1.48 was calculated for rural two-lane two-way roads in Kansas. Eventually, for evaluating and comparing different calibration methods, statistical parameters including Pearson correlation, Mean Prediction Bias (MPB) and Mean Absolute Deviation (MAD) were used [29]. Bornheimer et al. utilized the same database of Lubliner and Schrock's study and developed jurisdiction specific crash prediction models for Kansas's rural two-lane two-way roads. The evaluation results showed that the HSM's calibrated prediction model performed better than the Kansas's prediction model. The calibrated models had a higher Pearson's correlation coefficient and lower MPB and MAD values [30].

Sun et al. estimated calibration factors for a variety of crash prediction models for urban and rural roads provided in HSM. A calibration factor of 0.82 was obtained for the rural two-lane two-way roads in Missouri [31]. Mehta and Lou conducted a research on calibrating the crash prediction models of HSM for rural roads in Alabama. Also, a new calibration method was proposed in this study. The calibration factors of 1.39 and 1.52 were calculated, respectively, for the HSM's and the new calibration method in rural two-lane two-way roads [32]. Qin et al. calibrated the HSM's crash prediction model and modified its calibration method according to the data collected from rural two-lane two-way roads in South Dakota. Using the crash data of 2009 to 2011, a calibration factor of 1.54 was subsequently calculated [33]. Russo et al. performed research on calibrating the HSM's crash prediction model for Italian rural two-lane two-way roads. Also, new Safety Performance Functions (SPF) and Accident Modification Factors (AMF) were developed for predicting three levels of crash frequencies, including only injuries, only deaths, and injuries plus deaths. The calibration factors were calculated as 0.606, 0.424, and 0.581, respectively, for the three levels [34].

HSM recommends that the minimum required database for calibration should include 30 to 50 road segments with at least 100 crash records in a year [11]. Based on rural two-lane two-way roads data from Arizona State, it was proposed to utilize a calibration function when a proper crash estimation cannot be performed using the calibration factor. Road segments were classified based on different characteristics including traffic volume, segment length, and alignment. Then, calibration factors were calculated for each group. The overall calibration factor was estimated as 1.08 [35]. The results of another study showed that the HSM's generalized recommendation does not fit all situations since various parameters including crash records, reference population, etc. influence the sample size needed for a reliable calibration process. The calibration factor of 0.47 was calculated for rural two-lane two-way roads in Florida State [20].

Rajabi et al. developed two calibration factors and compared the results with the results of the HSM's calibration method. Also, the calibration factor defined in the study of Mehta and Lou was utilized in this research. The calibration factors were calculated for eight different types of roads and ten different types of intersections based on four calibration methods with the data from South Carolina State's roadway network. The HSM's calibration factor for rural two-lane two-way roads was calculated as 0.99 [18, 32, 36]. Llopis-Castello and Findley also calibrated the HSM's crash prediction model for rural two-lane two-way roads of North Carolina. Different calibration factors were calculated for the different types of road elements. The aggregated calibration factor and disaggregated calibration factors for road sections, including only horizontal curves and only tangents, were calculated, respectively, as 1.34, 1.57, and 1.15 [19]. Using data from Missouri State, calibration factor functions were developed based on segment length and Average Annual Daily Traffic (AADT). Also, new crash prediction models specific to Missouri's roads were created. The results showed that the calibration factor based on AADT had the best performance. It also showed that the accuracy of prediction was approximately the same for the HSM's calibrated model and the jurisdictional crash prediction models while there was a difference in the effort it took to collect and process the data [37].

It was presented in this section that many studies calibrated the HSM's crash prediction models for their jurisdictions. Some investigated the influence of different variables, including sample size, road elements, and road characteristics, on the precision of the calibration process. Others developed calibration functions and more complicated calibration methods instead of the HSM's single calibration factor methodology. And many researchers created jurisdictional crash prediction models and compared their performance with the calibrated HSM's crash prediction models. Despite various research performed in the United States and Europe on investigating the transferability of the HSM's crash prediction models to different regions, few researchers evaluated the applicability of this manual for Iran's roads. Using the safety analysis methods provided in this manual, road safety can be significantly improved, and many lives can be saved in Iran. Therefore, this study was intended to fill this gap in the literature and to calibrate and evaluate the performance of the HSM's crash prediction model for rural two-lane two-way roads in Kerman province, Iran. Rural two-lane two-way roads were selected due to the high rate of casualties in this type of roads in Iran [38-42]. The calibration was performed in order to identify the efficiency of HSM when it is used to predict crashes in a developing country. Also, the performance of the calibrated HSM model was compared to the performance of a jurisdictional specific crash prediction model developed for this region in an earlier study [39]. The results of this study are of importance since it tests the transferability of a model created in the USA to a developing country with different driving behavior characteristics and traffic conditions.

2. Material and Method

2.1. Data Collection

In this section, the data collection procedure is described.

2.1.1. Study Area

The intended study area in the present research consists of three rural two-lane two-way roads with a total length of 83 miles (133 km) in Kerman province, Iran. The topography was level, rolling and mountainous, respectively, for the first road with the length of 14 miles (22 km), the second road with the length of 34 miles (54 km) and the third road with the length of 35 miles (57 km).

2.1.2. Collected Road Characteristics

In the present study, the collected data included the required variables for the HSM's crash prediction model according to the following: AADT, lane width, shoulder width, the characteristics of the roadside area including clear zone width and side slope, the characteristics of horizontal curvatures including radius, length and super-elevation, vertical grade, the number of driveways, the existence of passing lane, and the existence of lighting [11, 43].

2.1.3. Data Sources

The data was collected from different sources. Some parameters were collected from field data surveys, and some from documentations and schematics. Some characteristics of the road including the radius and the length of horizontal curves and longitudinal slope were gathered from the documented plan and the profile of roads. Also, other field surveys were performed to measure the super-elevation of horizontal curves.

2.1.4. Crash and Traffic Data

Normally, a period of 2 to 5 years is considered to collect crash data [44]. According to the literature review and the changes that have occurred in the study road due to improvement projects, a period of three years was chosen for crash data collection [29, 33]. Crash records of a three-year period from 2010 to 2012 were obtained from Road Maintenance and Transportation Organization for the intended roads. Furthermore, the traffic volume data was collected from the same organization for the chosen period. This data was recorded by traffic detectors.

2.2. Method

In the following section, the methodology is presented. Firstly, the procedure developed by HSM to create homogenous segments is described. Then, the crash prediction model of HSM is presented. This model is utilized to predict crashes for the created homogenous segments. In addition, the HSM's calibration method is described. At the end, the statistical methods utilized for evaluation are mentioned with their formula.

2.2.1. Homogeneous Road Segments

Once the data was collected, the study roads were divided into homogeneous segments based on the HSM's segmentation method. A homogeneous segment is a segment through which all road characteristics are the same. According to HSM, the start of the road is considered as the start of the first segment. This segment continues until one of the following conditions are met:

Start or end of horizontal curve, Point of Vertical Intersection (PVI), start or end of passing lane, start or end of the two-way left-turn lane. In this situation, the last segment ends, and a new segment begins. Also, a new segment starts when a change in any of the following characteristics is observed:

AADT, lane width, shoulder width, shoulder type, Roadside Hazard Rating (RHR), centerline rumble strip, lighting, and automated speed enforcement [11]. All the variables used in this study are presented in Table A.1 in Appendix A.

2.2.2. The HSM's Crash Prediction Model

Afterwards, crash frequencies were predicted for the obtained homogeneous segments using the HSM's crash prediction model for rural two-lane two-way roads. The crash prediction model of HSM utilizes an 18-stage procedure and a base equation for the segments of rural two-lane two-way roads in order to estimate the frequency of crashes. The base formula is shown in Eq. (1)

$$N_{predicted} = N_{spf} \times (CMF_1 \times CMF_2 \times \ldots \times CMF_{12}) \times C \tag{1}$$

Where:

$N_{predicted} =$ The predicted crash frequency by the HSM' model,

$N_{spf} =$ The predicted crash frequency for base conditions,

$CMF_i =$ The ith crash modification factor,

$C =$ The calibration factor.

Eq. (2) shows the SPF, and Eq. (3) presents the calculation of the over-dispersion parameter, which is used in the Empirical Bayes (EB) method.

$$N_{spf} = AADT \times L \times 365 \times 10^{-6} \times e^{(-0.312)} \quad (2)$$

$$k = \frac{0.236}{L} \quad (3)$$

Where:
N_{spf} = The predicted crash frequency for base conditions,
$AADT$ = Average annual daily traffic (vehicles/day),
L = The segment's length (kilometers),
k = The over-dispersion parameter.

The HSM's SPF was developed based on some base conditions that were prevailing in the roads of the states whose data was used in the HSM's crash prediction model development. The base conditions for rural two-lane two-way roads are presented in Table 1.

Table 1. The HSM's base conditions for rural two-lane two-way roads

Number	Parameter	Base condition
1	Lane width	3.65 m (12 ft)
2	Shoulder width	1.85 m (6 ft)
3	RHR	3
4	Driveway density	3 driveways per km (5 driveways per mile)
5	Horizontal curvature	None
6	Vertical curvature	None
7	Longitudinal slope	Zero percent
8	Centerline rumble strip	None
9	Passing lane	None
10	Two-way left-turn lane	None
11	Lighting	None
12	Automated speed enforcement	None

Considering the geometric conditions of studied roads, Crash Modification Factors (CMF) are used to modify the frequency of crashes predicted by the SPF. There are twelve CMFs defined in HSM for rural two-lane two-way roads.

The crash prediction model of HSM has been developed based on the data of selected few states in the USA. Also, several parameters including driving behavior, weather conditions, traffic circumstances, etc. differ from one region to another. Therefore, HSM proposes that this model should be calibrated for the intended region before usage. Based on HSM, there are five steps in the calibration process. The first step is defining the type of road (for instance rural two-lane two-way) on which the calibration will be done. The second step is choosing the intended section of road for calibration. The selected section should have at least 30 to 50 road segments with a total annual crash frequency of 100 or more. The next step is summing the observed crash frequency for the chosen road segments in the selected period. The fourth step is calculating the total predicted crash frequency for the study road section using the HSM's crash prediction model with the calibration factor of one and without the EB method. Eventually, the last step is calculating calibration factor utilizing Eq. (4)

$$C = \frac{\sum_{i=1}^{n} N_{observed}}{\sum_{i=1}^{n} N_{predicted}} \quad (4)$$

Where:
C = The calibration factor,
$N_{observed}$ = The observed crash frequency,
$N_{predicted}$ = The predicted crash frequency by the HSM' model.

EB method was also implemented in the present study. In the case of the presence of crash history, the observed and the predicted number of crashes are combined to create the expected frequency of crashes. This method is utilized to counteract the bias of regression to the mean. The expected frequency of crashes is calculated through Eq. (5) and Eq. (6) [11].

$$N_{expected} = w \times N_{predicted} + (1 - w) \times N_{observed} \quad (5)$$

$$w = \frac{1}{1 + k \times (\sum_{all\ study\ years} N_{predicted})} \quad (6)$$

Where:
$N_{expected}$ = The expected crash frequency,
w = The weighted adjustment EB coefficient,
$N_{predicted}$ = The predicted crash frequency by the HSM' model,
$N_{observed}$ = The observed crash frequency,
k = The over-dispersion parameter.

In an earlier study, crash prediction models were developed for estimating frequency of crashes in rural two-lane two-way roads based on the data from the same region [39]. In the current study, the performance of these jurisdiction specific crash prediction models was compared to the performance of the HSM's crash prediction model in order to find out the best model.

2.2.3. Evaluation Methods

The statistical parameters of Spearman correlation coefficient, MPB and MAD were used in order to evaluate the crash prediction models. The statistical measure of Spearman correlation is a coefficient which shows the dependence between two variables. As much as this coefficient is closer to 1 or -1, the correlation between the observed and the predicted frequency of crashes is higher. Eq. (7) is used to estimate this statistical parameter.

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n(n^2 - 1)} \quad (7)$$

Where:
r_s = Spearman correlation,
d_i = The difference between the predicted and the observed crash frequency of the ith road segment,
n = The number of road segments.

The parameter of MPB is utilized to measure the mean difference between the predicted and the observed frequency of crashes. Low values of this parameter show the little difference between predicted and observed frequencies. Positive values show that predicted frequencies are greater than observed frequencies and vice versa. The value of this parameter is estimated using Eq. (8)

$$MPB = \frac{\sum_{i=1}^{n} d_i}{n} \tag{8}$$

Where:
$MPB =$ Mean prediction bias,
$d_i =$ The difference between the predicted and the observed crash frequency of the ith road segment,
$n =$ The number of road segments.

The value of MAD shows the magnitude of the mean difference between the model results and the observed frequency of crashes. The difference between MPB and MAD is that, in MAD, positive and negative values cannot zero out each other. It is calculated through Eq. (9)

$$MAD = \frac{\sum_{i=1}^{n} |d_i|}{n} \tag{9}$$

Where:
$MAD =$ Mean absolute deviation,
$d_i =$ The difference between the predicted and the observed crash frequency of the ith road segment,
$n =$ The number of road segments.

3. Results and Discussion

In this section, firstly, the results of homogenous segment creation are presented. Then, the crash prediction results for the created segments are offered. Afterwards, the results of the evaluation and the comparison of the jurisdictional and the HSM's model are presented. The discussion is presented at the end.

3.1. Crash Prediction Results

After performing the segmentation process on the study road based on the HSM's method and the elimination of intersections, 880 homogeneous segments were produced. Then, crash frequencies were predicted for the obtained homogeneous segments. Also, the calibration factor was calculated using the crash prediction model of HSM. The statistical summary of homogeneous segments is presented in Table 2.

Table 2. The summary statistics of homogeneous segments

Variable	Min	Max	Mean	Standard Deviation
L [mile (km)]	0.016 (0.025)	0.766 (1.233)	0.087 (0.140)	0.083 (0.133)
AADT [vehicles/day]	1496	2330	1972.47	269.44
LW [ft (m)]	8.530 (2.6)	12.467 (3.8)	10.545 (3.214)	0.915 (0.279)
SW [ft (m)]	0 (0)	10.171 (3.1)	2.635 (0.803)	1.831 (0.558)
CL [mile (km)]	0.029 (0.046)	0.337 (0.543)	0.133 (0.214)	0.067 (0.107)
CR [ft (m)]	285.433 (87)	32808.4 (10000)	4181.5 (1274.52)	5894.49 (1796.64)
SV [%]	0	0.05	0.012	0.015
G [%]	0	9	3.763	2.625
DD [1/mile (1/km)]	0 (0)	82.51 (51.28)	3.673 (2.283)	10.539 (6.550)
CRS	0	0	0	0
PL	0	1	0.025	0.156
TWLTL	0	0	0	0
RHR	1	7	4.6	1.44
Li	0	1	0.115	0.319
ASE	0	0	0	0

According to Table 2, there were not any Centerline Rumble Strips, Two-Way Left-Turn Lanes, and Automated Speed Enforcements in the study roads. The aggregated prediction results of the HSM's model and the calculated calibration factors for each year are shown in Table 3.

Table 3. The calculated calibration factors in different years

Year	Total observed crashes	Total predicted crashes	Calibration factor
2011	89	74	1.21
2012	117	77	1.52
2013	119	80	1.49
Overall	325	230	1.41

As shown in Table 3, the overall three-year calibration factor was calculated as 1.4. This value is greater than 1 i.e., the predicted frequencies of the HSM's model are less than the observed frequencies of crashes in general. In a previous research, several crash prediction models were created for rural two-lane two-way roads based on the data from the same area. Based on evaluation results, the best performance was observed in the following model [39].

$$N_{Pre-jur} = L^{0.9938} \times AADT^{2.5897} \times EXP\begin{pmatrix} -0.5071 \times SW + 0.2370 \times LU + 49.3516 \times 1/CR \\ +12.5746 \times SV + 0.0193 \times DD - 18.6762 \end{pmatrix} \tag{10}$$

Where:
$N_{Pre-jur} =$ The predicted crash frequency by the jurisdictional model,
$L =$ The segment's length (kilometers),
$AADT =$ Average annual daily traffic (vehicles/day),
$SW =$ The shoulder's width (meters),
$LU =$ Land use (1 for the presence of residential land use on roadside and 0 for absence),
$CR =$ The curvature radius (meters),
$SV =$ The variance between the super-elevation recommended by the highway design code and the actual super-elevation of the curve,
$DD =$ Driveway density calculated by dividing the number of driveways on both sides of highway by the length of segment (driveways/kilometer).

It is observed that, in the developed model, influencing variables include segment length, traffic volume, shoulder width, the presence of residential land use on roadside, curvature radius, the variance between recommended super-elevation and observed super-elevation, and driveway density.

3.2. Evaluation Results

In the current study, the performance of the jurisdiction specific crash prediction model was compared with the HSM's model. Also, the EB method was utilized as an additive to crash prediction models in order to enhance their performance. The results of evaluation are presented in Table 4.

Table 4. Evaluation results

Model	r_s	MPB	MAD	Calibration Factor
HSM's model without EB	0.48	-0.88	2.00	1.41
HSM's model with EB	0.77	-0.36	0.86	1.17
Jurisdictional Model without EB	0.56	0.00	1.90	1.00
Jurisdictional Model with EB	0.71	0.008	1.42	0.997

The addition of the EB method improved results as expected. According to Table 4, in both the HSM's and the jurisdictional models, the value of Spearman correlation coefficient using EB was higher and closer to 1 compared to without EB. This indicates that the addition of EB leads to higher correlation between the predicted and the observed frequency of crashes. Also, using EB with the HSM's model, the value of calibration factor became smaller and closer to 1 showing more precision and more similarity of predicted and observed values. The calculated values for MPB and MAD were smaller and closer to 0, utilizing EB, indicating lower deviation and better performance.

By comparing Spearman correlation values in different models, it is observed that, without EB, the jurisdictional model had a higher correlation than the HSM's model. However, using EB, the HSM's model showed better performance than the jurisdictional model. It may be due to the reason that the utilized EB method was developed in HSM and is more compatible with the HSM's model. The same trend was observed in MAD values. The value of MPB for the jurisdictional model was around 0. However, this model had the MAD value of more than 1 which means errors canceled each other out. The calculated calibration factors were equal to 1 for the jurisdictional model. It may be owing to the fact that the jurisdictional model was developed based on the data from the same region and did not require calibration. All in all, it is concluded that the addition of the EB method improved the performance of crash prediction models. Also, consistent with literature review, significant difference was not observed between the performance of the jurisdictional and the HSM's model with respect to the effort that has been put to develop the jurisdictional model [30, 37]. Utilizing the EB method, the HSM's model even outperformed the jurisdictional model. A summary of the conducted research on the calibration of the HSM's crash prediction model for rural two-lane two-way roads is presented in Table 5.

Table 5. The calculated calibration factors for the HSM's crash prediction model of rural two-lane two-way roads

Number	Researcher	Year	Calibration factor	Study area
1	Sun et al.	2006	1.63	Louisiana (USA)
2	Martinelli et al.	2009	0.37	Italy
3	Xie et al.	2011	0.74	Oregon (USA)
4	Banihashemi	2011	1.50	Washington (USA)
5	Williamson and Zhou	2012	1.40	Illinois (USA)
6	Brimley et al.	2012	1.16	Utah (USA)
7	Lubliner and Schrock	2012	1.48	Kansas (USA)
8	Sun et al.	2013	0.82	Missouri (USA)
9	Mehta and Lou	2013	1.39	Alabama (USA)
10	Qin et al.	2014	1.54	South Dakota (USA)
11	Russo et al.	2014	0.58	Italy
12	Srinivasan et al.	2016	1.08	Arizona (USA)
13	Alluri et al.	2016	0.47	Florida (USA)
14	Rajabi et al.	2018	0.99	South Carolina (USA)
15	Llopis Castello and Findley	2019	1.34	North Carolina
16	The present study	2020	1.17	Iran

It is observed in Table 5 that, considering calibration factor, the HSM's crash prediction model had an acceptable performance in the study area of the present research comparing to other studies.

4. Conclusion

In this study, the HSM's crash prediction model was calibrated for the rural two-lane two-way roads of Kerman province, Iran. In addition, the performance of a jurisdictional crash prediction model that was developed in a previous research was compared to the performance of the HSM's model. It was concluded that the HSM's model had an acceptable performance in comparison with the jurisdictional model.

Owing to the importance of rural roads safety especially rural two-lane two-way roads in Iran, the safety of this type of roads was investigated in this research. Several rural two-lane two-way roads were chosen for data collection in Kerman province, Iran. The selected roads had different topographic characteristics including level, rolling, and mountainous. A three-year period of 2011 to 2013 was selected for collecting crash data. Some of road features were gathered through field surveys and some from maps and documents. Then, using the HSM's segmentation method, the studied roads were divided into homogeneous segments. The crash prediction model of HSM was employed to predict crash frequencies for the created road segments. Also, a jurisdictional crash prediction model developed in a previous research was utilized to estimate the safety of the studied roads. Furthermore, the EB method was used with respect to the accessibility of crash history. Afterwards, according to the HSM's calibration method, the calibration factor of 1.17 was obtained indicating the satisfying performance of the HSM's crash prediction model on evaluating the safety of the studied roads. The value of 0.77 for Spearman correlation coefficient between the output of the HSM's model using EB and the observed frequency of crashes, and the small values of MPB and MAD were indicative of low bias and accurate prediction of this model. With respect to correlation coefficient and MAD values, it is observed that, without EB, the jurisdictional model showed better performance than the HSM's model. However, utilizing EB, the HSM's model represented higher correlation and lower bias. It may be due to the compatibility of the EB method with the HSM' model. Consequently, the HSM's crash prediction model showed the same safety evaluation performance compared to the jurisdictional model. However, much more time and effort were spent on data preparation and modeling to develop the jurisdictional specific model.

The limitations of this study were as follows. The first limitation was using a small dataset since the data collection required performing field surveys needing much time and high costs. Another limitation of this research was the absence of some parameters that were considered in the HSM's crash prediction model but were not present in the studied roads. These parameters included centerline rumble strips, two-way left-turn lanes, and automated speed enforcements. In future studies, the HSM's crash prediction model can be calibrated for other road types in this area. Also, jurisdictional crash prediction models can be developed for other road types. In addition, specific crash modification factors can be created for the road features that influence the safety of roads in developing countries. The results of this paper can be utilized by road designers, policy makers, and road safety administrators to better understand the contributing factors to road safety. Through using the calibrated HSM's crash prediction model, the road segments with high crash risks can be identified and solutions can be given to improve their safety. Also, the safety of future roads can be evaluated in order to find the deficiencies of road designs before moving into the construction stage. Therefore, many lives could be saved using the calibrated models.

5. Author contributions

The contribution of authors in this study is according to the following. The original idea was from **Mohamadreza Haghani.**

Mohamadreza Haghani: Conceptualization, Methodology, Writing - Original Draft. **Reza Jalalkamali**: Methodology, Writing- Reviewing and Editing. **Hamidreza Haghani**: Formal analysis, Writing- Original Draft.

Conflict of Interest Statement

The authors declare no conflict of interest.

Appendix

Table A.1. The variables implemented in this study

Number	Variable	Definition	Source
1	$AADT$	Average annual daily traffic	Road Maintenance and Transportation Organization Database
2	L	Length of segment	Field surveys, plans
3	LW	Lane width	Field surveys
4	SW	Shoulder width	Field surveys
5	RHR	Roadside hazard rating[14]	Field surveys
6	CL	Length of horizontal curvature	Plans
7	CR	Radius of horizontal curvature	Plans
8	SV	The variance between the super-elevation recommended by AASHTO Green Book[13] and the actual superelevation of horizontal curve[14]	Field surveys, plans
9	G	Vertical grade	Profiles
10	DD	Driveway density that is calculated by dividing the number of driveways on both sides of the road by the length of segment[14]	Field surveys, plans
11	Li	Lighting (1 for the presence and 0 for the absence of it)	Field surveys
12	PL	Passing Lane (1 for the presence and 0 for the absence of it)	Field surveys
13	CRS	Centerline rumble strip (1 for the presence and 0 for the absence of it)	Field surveys
14	$TWLTL$	Two-way left-turn lane (1 for the presence and 0 for the absence of it)	Field surveys
15	ASE	Automated speed enforcement (1 for the presence and 0 for the absence of it)	Field surveys
16	$N_{observed}$	Observed crash frequency	Road Maintenance and Transportation Organization Database
17	$N_{predicted}$	Predicted crash frequency by the HSM' model	Equation 1
18	$N_{pre-jur}$	Predicted crash frequency by the jurisdictional model	Equation 10
19	$N_{expected}$	Expected crash frequency	Equation 5
20	N_{spf}	Predicted crash frequency for base conditions	Equation 2
21	CMF	Crash modification factor	Equation 1
22	C	Calibration factor	Equation 4
23	k	Overdispersion parameter	Equation 3
24	w	Weighted adjustment EB coefficient	Equation 6
25	r_s	Spearman correlation	Equation 7
26	MPB	Mean prediction bias	Equation 8
27	MAD	Mean absolute deviation	Equation 9
28	d_i	The difference between the predicted and the observed crash frequency of the i^{th} segment	Equations 7, 8 and 9
29	n	Number of road segments	Equations 7, 8 and 9

References

[1] W.H. Organization, Global status report on road safety 2018. Geneva, Switzerland, 2019.

[2] W. WHO, Global status report on road safety 2013: supporting a decade of action. Geneva, Switzerland, 2013.

[3] A. Sheykhfard, F.R. Haghighi, M. Soltaninejad, A. Karji, Analyzing Drivers' Mental Patterns Using Q-Methodology, Journal of Transportation Technologies 10 (2020) 169.

[4] A.M. Amiri, A. Sadri, N. Nadimi, M. Shams, A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers, Accident Analysis & Prevention 138 (2020) 105468.

[5] A. Hosseinzadeh, A. Karimpour, R. Kluger, R. Orthober, A Framework to Link Crashes to Emergency Medical Service Runs and Trauma Admissions: For Improved Highway Safety Monitoring and Crash Outcome Assessment. In: 99th Annual Meeting of Transportation Research Board, Washington DC, January (2020).

[6] A. Hosseinzadeh, R. Kluger, Do EMS Times Associate with Injury Severity?, Accident Analysis & Prevention 153 (2021) 106053.

[7] I. Bargegol, V.N.M. Gilani, M. Ghasedi, M. Ghorbanzadeh, Delay modeling of un-signalized roundabouts using neural network and regression, Computational Research Progress in Applied Science & Engineering (CRPASE) 2 (2016) 28─34.

[8] M. Koloushani, E.E. Ozguven, A. Fatemi, M. Tabibi, Mobile Mapping System-based Methodology to Perform Automated Road Safety Audits to Improve Horizontal Curve Safety on Rural Roadways, Computational Research Progress in Applied Science & Engineering (CRPASE) 6 (2020) 263─275.

[9] N. Nadimi, S.S.N. Alavi, A.H.Z. Mirhosseini, A.M. Amiri, Proposing a Model for Ranking Hotspots in Rural Roads Using a Multi-criteria Decision-making Method, Computational Research Progress in Applied Science & Engineering (CRPASE) 3 (2017) 8─14.

[10] A. Hosseinzadeh, M. Haghani, R. Kluger, Exploring Influencing Factors on Crash-related Emergency Response Time: A Machine Learning Approach. In: 100th Annual Meeting of Transportation Research Board, Washington DC, January 5─29 (2021).

[11] H.S. Manual, Aashto, Washington DC, 529 (2010).

[12] N. Soltani, A. Naderan, N. Soltani, Modeling crash frequencies by transportation mode using micro/macro level variables, AUT Journal of Civil Engineering (2020).

[13] A. Foley, Highway design and maintenance, road safety and traffic management. In: Proceedings of the Institution of Civil Engineers-Transport, Thomas Telford Ltd (2019) 247─248.

[14] A.M. Amiri, N. Nadimi, A. Yousefian, Comparing the efficiency of different computation intelligence techniques in predicting accident frequency, IATSS Research 44 (2020) 285─292.

[15] A. Hosseinzadeh, M. Moeinaddini, A. Ghasemzadeh, Investigating Factors Affecting Severity of Large Truck-involved Crashes: Comparison of the SVM and Random Parameter Logit Model, Journal of Safety Research (2021).

[16] M.B. Ulak, E.E. Ozguven, H.H. Karabag, M. Ghorbanzadeh, R. Moses, M. Dulebenets, Development of safety performance functions for restricted crossing U-turn intersections, Journal of transportation engineering, Part A: Systems 146 (2020) 04020038.

[17] A. Farid, M. Abdel-Aty, J. Lee, A new approach for calibrating safety performance functions, Accident Analysis & Prevention 119 (2018) 188─194.

[18] M. Rajabi, J.H. Ogle, P. Gerard, Highway Safety Manual calibration: assessing alternate definitions of the calibration factor, Transportation research record 2672 (2018) 99─108.

[19] D. Llopis-Castelló, D.J. Findley, Influence of calibration factors on crash prediction on rural two-lane two-way roadway segments, Journal of Transportation Engineering, Part A: Systems 145 (2019) 04019024.

[20] P. Alluri, D. Saha, A. Gan, Minimum sample sizes for estimating reliable Highway Safety Manual (HSM) calibration factors, Journal of Transportation Safety & Security 8 (2016) 56─74.

[21] H. Lubliner, Evaluation of the highway safety manual crash prediction model for rural two-lane highway segments. In Kansas, in, University of Kansas (2011).

[22] X. Sun, Y. Li, D. Magri, H.H. Shirazi, Application of highway safety manual draft chapter: Louisiana experience, Transportation research record 1950 (2006) 55─64.

[23] F. Martinelli, F. La Torre, P. Vadi, Calibration of the Highway Safety Manual's accident prediction model for Italian secondary road network, Transportation research record 2103 (2009) 1─9.

[24] F. Xie, K. Gladhill, K.K. Dixon, C.M. Monsere, Calibration of highway safety manual predictive models for Oregon state highways, Transportation research record 2241 (2011) 19─28.

[25] M. Banihashemi, Highway Safety Manual, new model parameters vs. calibration of crash prediction models. In: Moving Toward Zero. 2011 ITE Technical Conference and Exhibit Institute of Transportation Engineers (ITE), April 3─6 (2011).

[26] R.A. Tegge, J.-H. Jo, Y. Ouyang, Development and application of safety performance functions for Illinois, Illinois, 2010.

[27] M. Williamson, H. Zhou, Develop calibration factors for crash prediction models for rural two-lane roadways in Illinois, Procedia-social and behavioral sciences 43 (2012) 330─338.

[28] B.K. Brimley, M. Saito, G.G. Schultz, Calibration of Highway Safety Manual safety performance function: development of new models for rural two-lane two-way highways, Transportation research record 2279 (2012) 82─89.

[29] H. Lubliner, S.D. Schrock, Calibration of the Highway Safety Manual prediction method for rural Kansas highways. In: 91st Annual Meeting of Transportation Research Board, Washington DC, January 22─26 (2012).

[30] C. Bornheimer, S.D. Schrock, M.-H. Wang, H. Lubliner, Developing a regional safety performance function for rural two-lane highways. In: 91st Annual Meeting of Transportation Research Board, Washington DC, January 22─26 (2012).

[31] C. Sun, H. Brown, P. Edara, B. Claros, K. Nam, Calibration of the highway safety manual for Missouri. Mid-America Transportation Center, 2013.

[32] G. Mehta, Y. Lou, Calibration and development of safety performance functions for Alabama: Two-lane, two-way rural roads and four-lane divided highways, Transportation research record 2398 (2013) 75─82.

[33] X. Qin, C. Zhi, K. Vachal, Calibration of Highway Safety Manual predictive methods for rural local roads. In: 93rd Annual Meeting of Transportation Research Board, Washington DC, January 12─16 (2014).

[34] F. Russo, S.A. Biancardo, M. Busiello, G. Dell'Acqua, G. Coraggio, Crash Severity Prediction Functions on Italian Rural Roads. In: 93rd Annual Meeting of Transportation Research Board, Washington DC, January 12─16 (2014).

[35] R. Srinivasan, M. Colety, G. Bahar, B. Crowther, M. Farmen, Estimation of calibration functions for predicting crashes on rural two-lane roads in Arizona, Transportation research record 2583 (2016) 17─24.

[36] J.H. Ogle, M. Rajabi, Development of South Carolina Databases and Calibration Factors for the Highway Safety Manual, Federal Highway Administration, South Carolina, 2018.

[37] B. Claros, C. Sun, P. Edara, HSM calibration factor, calibration function, or jurisdiction-specific safety model–A comparative analysis, Journal of Transportation Safety & Security 12 (2020) 309─328.

[38] M. Ahmadinejad, S. Afandizadeh Zargari, R. Jalalkamali, Are deceleration numbers a suitable index for road safety?. In: Proceedings of the Institution of Civil Engineers-Transport, Thomas Telford Ltd 171 (2018) 247─252.

[39] M. Haghani, R. Jalalkamali, M. Berangi, Assigning crashes to road segments in developing countries. In: Proceedings of the Institution of Civil Engineers-Transport, Thomas Telford Ltd 172 (2019) 299-307.

[40] A.A. Kordani, B. Shirini, M. Yazdani, The Effect of Road Shoulder and Weather Conditions on the Occurrence of Rollover Crashes in Two-lane Highways, Periodica Polytechnica Civil Engineering 63 (2019) 1235─1241.

[41] W. Ackaah, M. Salifu, Crash prediction model for two-lane rural highways in the Ashanti region of Ghana, IATSS research 35 (2011) 34─40.

[42] D.W. Harwood, F. Council, E. Hauer, W. Hughes, A. Vogt, Prediction of the expected safety performance of rural two-lane highways, Federal Highway Administration, Virginia, 2000.

[43] A.A.o.S. Highway, T. Officials, A policy on geometric design of highways and streets, American Association of State Highway and Transportation Officials, Washington DC, 2011.

[44] S. Cafiso, A. Di Graziano, G. Di Silvestro, G. La Cava, Safety performance indicators for local rural roads: comprehensive procedure from low-cost data survey to accident prediction model. In: 87th Annual Meeting of Transportation Research Board, Washington DC, January 13─17 (2008).