



# Accident and hazard prediction models for highway–rail grade crossings: a state-of-the-practice review for the USA

Olumide F. Abioye<sup>1</sup> · Maxim A. Dulebenets<sup>1</sup> · Junayed Pasha<sup>1</sup> · Masoud Kavooosi<sup>1</sup> · Ren Moses<sup>1</sup> · John Sobanjo<sup>1</sup> · Eren E. Ozguven<sup>1</sup>

Received: 20 April 2020 / Revised: 15 July 2020 / Accepted: 20 July 2020 / Published online: 25 August 2020  
© The Author(s) 2020

**Abstract** Highway–rail grade crossings (HRGCs) are one of the most dangerous segments of the transportation network. Every year numerous accidents are recorded at HRGCs between highway users and trains, between highway users and traffic control devices, and solely between highway users. These accidents cause fatalities, severe injuries, property damage, and release of hazardous materials. Researchers and state Departments of Transportation (DOTs) have addressed safety concerns at HRGCs in the USA by investigating the factors that may cause accidents at HRGCs and developed certain accident and hazard prediction models to forecast the occurrence of accidents and crossing vulnerability. The accident and hazard prediction models are used to identify the most hazardous HRGCs that require safety improvements. This study provides an extensive review of the state-of-the-practice to

identify the existing accident and hazard prediction formulae that have been used over the years by different state DOTs. Furthermore, this study analyzes the common factors that have been considered in the existing accident and hazard prediction formulae. The reported performance and implementation challenges of the identified accident and hazard prediction formulae are discussed in this study as well. Based on the review results, the US DOT Accident Prediction Formula was found to be the most commonly used formula due to its accuracy in predicting the number of accidents at HRGCs. However, certain states still prefer customized models due to some practical considerations. Data availability and data accuracy were identified as some of the key model implementation challenges in many states across the country.

**Keywords** Highway–rail grade crossings · Accident prediction methods · Hazard prediction methods · Resource allocation · Critical review

---

✉ Maxim A. Dulebenets  
mdulebenets@eng.famu.fsu.edu

Olumide F. Abioye  
olumide1.abioye@famu.edu

Junayed Pasha  
jp17j@my.fsu.edu

Masoud Kavooosi  
mk17@my.fsu.edu

Ren Moses  
moses@eng.famu.fsu.edu

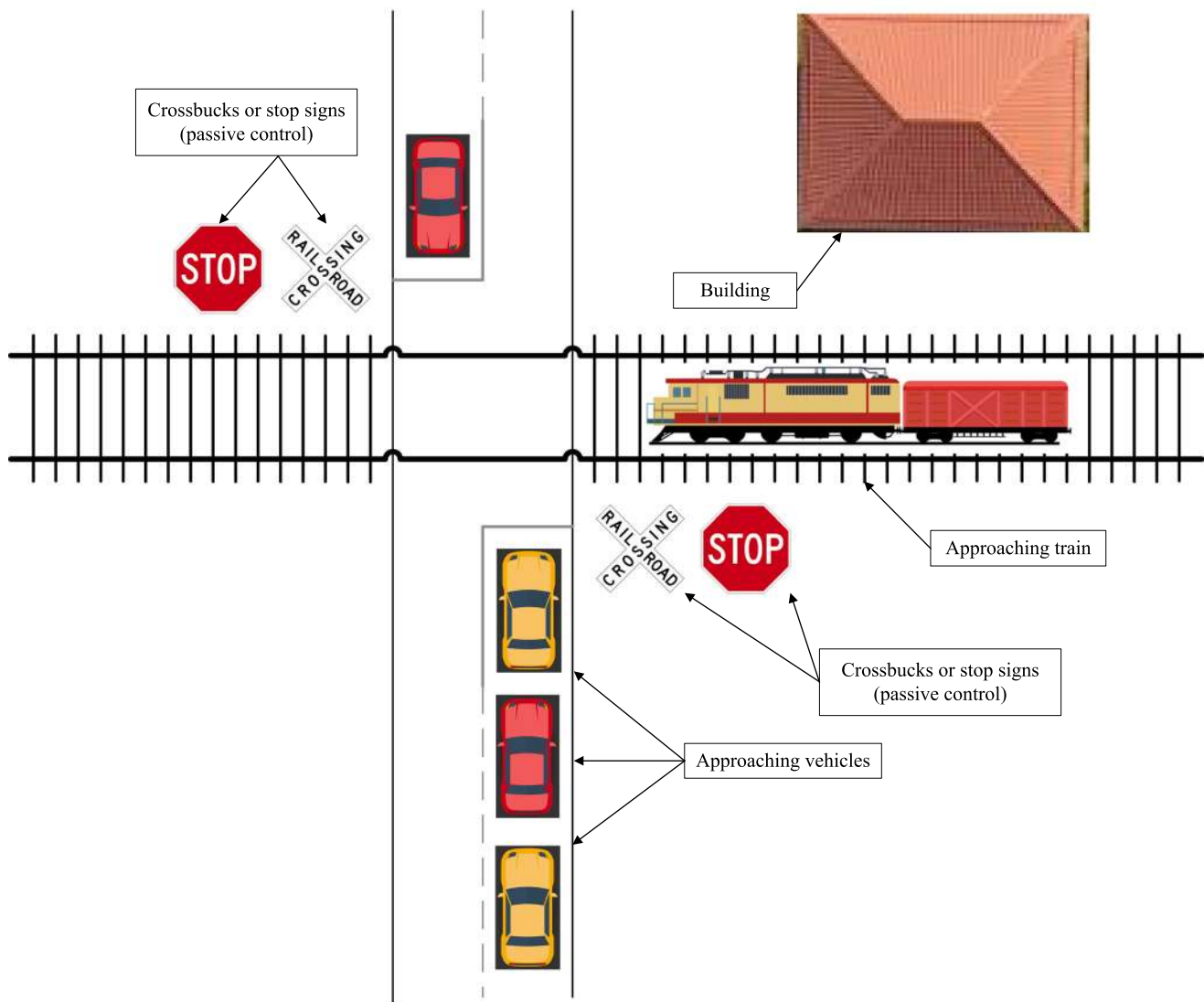
John Sobanjo  
sobanjo@eng.famu.fsu.edu

Eren E. Ozguven  
eozguven@eng.famu.fsu.edu

<sup>1</sup> Department of Civil and Environmental Engineering, Florida A&M University-Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310-6046, USA

## 1 Introduction

A highway–rail grade crossing (HRGC) is a segment of the transportation network, where the highway and the rail tracks intersect at the same elevation (Fig. 1). HRGCs can be found on both public and private highways in the USA. Over the past decades, several factors, including an increase in the demand for freight and passenger transportation, urbanization, and growing road and rail network, have led to an increase in the number of HRGCs in the USA. The public and private HRGCs in the USA were in



**Fig. 1** A schematic illustration of an HRGC

excess of 250,000 in the year of 2015 [1]. In the meantime, safety at HRGCs has been posing a major challenge over the past years [2]. At every HRGC, there is a likelihood of train–vehicle, vehicle–vehicle, and vehicle–warning device accidents. The Federal Railroad Administration (FRA) defines an HRGC accident as “any impact between a rail and highway user (both motor vehicles and other users) of the crossing at a designated crossing site, including walkways, sidewalks, etc., associated with the crossing” [3]. HRGC accidents may result in severe injuries, fatalities, and cause substantial property damage. A total of 6242 HRGC accidents involving trains and vehicles were reported in the USA from 2015 to 2017, which resulted in 895 deaths and 2745 injuries [4]. Therefore, quite a significant portion of the HRGC accidents in the USA cause

fatalities. Such a pattern has been observed in the previous years as well (Table 1).

HRGC accidents have received a lot of attention from the research community and state Departments of Transportation (DOTs) over the last decades [5–10]. Although accidents at HRGCs have been extensively investigated in the state-of-the-art, recent trends have shown that safety at HRGCs remains a major source of concern for the state DOTs. Severe injuries, fatalities, and substantial property damages are reported in each state of the USA every year [4]. Researchers and state DOTs have used various methods to predict accident occurrence and severity at HRGCs. Several studies have investigated different factors that may influence the occurrence of accidents at HRGCs and developed the accident and hazard prediction models for HRGCs [11, 12]. The accident prediction models rely on

**Table 1** Number of train–vehicle accidents, injuries, and fatalities in the USA between 2008 and 2017

Year	Accidents	Injuries	Fatalities
2008	2429	993	291
2009	1933	744	248
2010	2052	888	263
2011	2064	1051	271
2012	1988	982	273
2013	2104	983	290
2014	2296	882	295
2015	2078	1046	287
2016	2047	858	300
2017	2117	841	308

certain operational and physical HRGC characteristics (e.g., average daily traffic volume, train type, average daily train volume, highway vehicular speed, train speed, number of rail tracks, and type of warning devices used) as well as the accident history in order to estimate the predicted number of accidents for a specific period of time (generally, a 1-year period is considered). Unlike the accident prediction models, the hazard prediction models do not estimate the actual predicted number of accidents. The hazard prediction models assess vulnerability of HRGCs to potential accidents based on certain operational and physical HRGC characteristics as well as the accident history.

Some of the accident and hazard prediction models are nationally recognized and adopted by a number of states. On the other hand, certain states have developed their own accident and hazard prediction models for HRGCs [11, 12]. The models are used to identify the most dangerous HRGCs in each state and to prioritize them for safety improvement projects based on the available budget. Certain studies in the state-of-the-art have used customized statistical models (such as the Poisson model, the negative binomial model, the gamma model, the Conway–Maxwell–Poisson model, the Bernoulli model, and the zero-inflated Poisson model among others) to examine the relationship between accidents at HRGCs and contributing factors [6, 10, 13]. Considering a strong emphasis of public agencies on the safety aspects of transportation networks [14–18], the present study aims to conduct a comprehensive review of the existing accident and hazard prediction formulae. The formulae that have been used by different state DOTs in the USA to estimate the expected number of accidents or safety hazard at HRGCs and rank HRGCs for safety improvements will be of a primary interest. Furthermore, a detailed analysis will be performed to identify the common factors (accident or hazard predictors) that are

considered in the existing accident and hazard prediction formulae. Last but not least, this study will discuss the reported performance of the accident and hazard prediction models as well as the implementation challenges of these models.

The rest of this paper is structured as follows: The next section provides a concise overview of the state-of-the-art and the state-of-the-practice with a primary emphasis on the state DOT studies that evaluated and/or developed certain accident and hazard prediction formulae. Section 3 categorizes the identified models into two major groups (i.e., accident and hazard prediction formulae). Section 4 analyzes the common factors (i.e., predictors) that have been considered in the existing accident and hazard prediction formulae. Section 5 focuses on the reported performance of the identified accident and hazard prediction models. Moreover, the major implementation challenges of the models are discussed in the fifth section as well. The final section provides the necessary conclusions and outlines possible future research directions.

## 2 Literature review

This section provides a detailed review of the research efforts, undertaken by the state DOTs to predict accident occurrence and improve safety at HRGCs. Also, a concise review of the relevant state-of-the-art studies, which evaluated various accident and hazard prediction models and/or proposed certain statistical models for identifying the factors that may cause accidents at HRGCs, is presented in this section as well.

### 2.1 Previous research efforts by the state DOTs

Different state DOTs have made a number of research attempts to predict the occurrence of accidents at HRGCs. Throughout the literature search, ten state DOT reports were identified. The reports were prepared by the States of Virginia (1986), Alabama (1994), Illinois (2000), Missouri (2003), Tennessee (2012), Texas (2013), Iowa (2015), Nevada (2017), Ohio (2017), and Florida (2020). This section summarizes findings from the previous research efforts, undertaken by the state DOTs.

Faghri and Demetsky [19] conducted a study for the Virginia Highway & Transportation Research Council to evaluate the existing accident and hazard prediction models based on their ability to forecast accident occurrence and estimate hazard indexes for HRGCs. A survey was conducted among 45 state DOTs to determine the strategies used for identifying hazardous HRGCs. Throughout the study, a total of 13 accident and hazard prediction models were discovered. The analysis of the survey results

indicated that 32% of the states developed their own models, and 30% of the states used the US DOT Accident Prediction Formula. A set of statistical analyses were performed in the study to evaluate the top five most frequently used formulae based on their ability to rank the most hazardous HRGCs in the State of Virginia. The results confirmed that the US DOT Accident Prediction Formula performed better than other four formulae [19]. Bowman [20] performed a study to enhance the rail-highway safety program in the State of Alabama. The survey that was conducted in the study identified that the states had been primarily using the US DOT Accident Prediction Formula and the custom accident and hazard prediction formulae. Additionally, the survey responses revealed that four states were not using any accident and hazard prediction formulae. These states established a priority schedule using public complaints, feedback from railroad companies, accident history, and field inspections of HRGCs.

Elzohairy and Benekohal [11] performed a survey among 49 states in the USA to identify the techniques used by the state DOTs to rank HRGCs for safety improvement projects. The following criteria were identified as important for ranking HRGCs [11]: (i) higher hazard indexes and higher number of predicted accidents; (ii) site review of vehicle types (mass transit, school buses); (iii) benefit-cost analysis; (iv) public concerns and complaints; (v) engineering judgment and HRGC geometry; (vi) sight distance; and (vii) service condition. A new model (referred to as the Illinois Hazard Index Formula) was developed for predicting potential hazard of the HRGCs in the State of Illinois. The Missouri DOT conducted a study jointly with the University of Missouri—Columbia/Rolla [21]. A number of different accident and hazard prediction models were evaluated for the HRGCs in the State of Missouri. The Illinois Hazard Index Formula was found to be the most accurate for active HRGCs. On the other hand, the California Hazard Rating Formula outperformed the other models for passive HRGCs. The Tennessee DOT and the University of Memphis conducted a study to improve the HRGC safety in the state [22]. A total of two optimization models were designed as a part of the study to assist with efficiently allocating monetary resources for safety improvement projects based on the US DOT resource allocation procedure [23].

Weissmann et al. [24] proposed a Revised Texas Priority Index for ranking public HRGCs in the State of Texas for safety improvement projects. The canonical Texas Priority Index Formula and the Revised Texas Priority Index Formula were evaluated for 9108 HRGCs, located in the State of Texas, using the 2011 accident data. The results revealed that the Revised Texas Priority Index Formula was able to identify more hazardous HRGCs as compared to the Texas

Priority Index Formula. The Iowa State University and the Iowa DOT conducted a study to design a methodology for ranking the HRGCs in the State of Iowa and to effectively conduct resource allocation [25, 26]. The benefit-cost analysis has been used in the State of Iowa for allocation of the available monetary resources among the existing HRGCs. Hans et al. [26] suggested a weighted index method and a supplementary Microsoft Excel spreadsheet-based tool for ranking the HRGCs, which can serve as an addition to the existing procedures in the State of Iowa. Ryan and Mielke [27] discussed the conventional methods that are frequently used to prioritize HRGCs and proposed a new model for ranking the HRGCs in Nevada. The formula accounted for the daily train volume, average daily highway traffic, near misses within the past 3 years, number of accidents within the past 5 years, existing protection, train speed, and highway speed.

Sperry et al. [28] reported that 110 accidents were recorded on average every year between 2005 and 2010 at the HRGCs in the State of Ohio. The analysis of the accident data showed that the HRGCs where accidents occurred had higher train speeds, higher train volumes, and more tracks. The study recommended that the State of Ohio should keep using the US DOT Accident Prediction Formula for ranking the HRGCs for safety improvement projects. The North Carolina Investigative Index and the Missouri Exposure Index were recommended for ranking passive HRGCs upon completion of the initial prioritization. The Florida DOT conducted a study in 2020, aiming to improve the HRGC safety in the State of Florida [12]. The Modified Texas Priority Index Formula, which unlike the canonical Texas Priority Index Formula accounted for the warning device upgrades in the accident history estimations, was found to be the most accurate in ranking the state HRGCs. Furthermore, two optimization models were developed as a part of the study to minimize the overall HRGC hazard and to minimize the overall HRGC hazard severity. A set of heuristic-based algorithms were developed to solve the optimization models. Moreover, a standalone decision support tool was designed to assist the FDOT personnel by assessing the overall hazard of the HRGCs in the State of Florida as well as resource allocation among the state HRGCs. A set of illustrative examples were presented to demonstrate applicability of the developed decision support tool.

A summary of the state DOT efforts in terms of the conducted activities is presented in Table 2. It can be observed that most of the previous state DOT studies conducted interviews or surveys among the appropriate representatives regarding the safety aspects at HRGCs, provided a detailed review of some of the existing accident and hazard prediction models, as well as evaluated some of the existing accident and hazard prediction models for the

**Table 2** Summary of the state DOT efforts

Study/conducted activities	Interviews or surveys	Detailed review of the existing models	Evaluation of the existing models	New model development	Decision support tools
Virginia (1986) [19]	✓	✓	✓		
Alabama (1994) [20]	✓	✓	✓		
Illinois (2000) [11]	✓	✓	✓	✓	
Missouri (2003) [21]	✓	✓	✓	✓	
Tennessee (2012) [22]		✓	✓		✓
Texas (2013) [24]	✓	✓	✓	✓	
Iowa (2015) [25, 26]	✓	✓			✓
Nevada (2017) [27]	✓	✓	✓	✓	
Ohio (2017) [28]	✓	✓	✓		
Florida (2020) [12]		✓	✓	✓	✓

HRGCs in a particular state. However, only several studies developed new models for accident and hazard prediction at HRGCs as well as designed decision support tools that could be used to improve safety at HRGCs.

## 2.2 Other relevant studies

Similar to the state DOT efforts, some state-of-the-art studies have also evaluated various accident and hazard prediction models for HRGCs (including the US DOT Accident Prediction Formula, the Peabody–Dimmick Formula, the Coleman–Stewart Model, the NCHRP Report 50 Accident Prediction Formula, and others). Austin and Carson [5] underlined that some of the accident and hazard prediction models, including the Peabody–Dimmick Formula, the NCHRP Report 50 Accident Prediction Formula, and the New Hampshire Hazard Index Formula, lack descriptive capabilities because they consider a limited number of explanatory variables. Furthermore, the authors stated that the accuracy of the US DOT Accident Prediction Formula in predicting accidents might significantly reduce over time due to the model complexity. The study developed an alternative HRGC accident prediction model that was based on the negative binomial regression.

Oh et al. [13] used various statistical models to examine some of the factors that cause accidents at HRGCs. A total of four statistical models, including the gamma probability model, the Poisson model, the negative binomial model, and the zero-inflated Poisson model, were considered. Also, four nationally recognized HRGC accident and hazard prediction models were evaluated, namely the Peabody–Dimmick Formula, the New Hampshire Hazard Index Formula, the NCHRP Report 50 Accident Prediction Formula, and the US DOT Accident Prediction Formula. The HRGC accident data, obtained for 162 HRGCs in

Korea, were used to evaluate the models. The results from the analysis of the models revealed that the gamma probability model fit the data best. However, the authors pointed out that the US DOT Accident Prediction Formula was more comprehensive and descriptive in terms of the number of predictors considered as compared to the other accident and hazard prediction models. Chadwick et al. [8] performed a comprehensive review of the literature to identify the challenges of shared high-speed passenger and freight railroads. A number of accident and hazard prediction models were described, including the Peabody–Dimmick Formula, the New Hampshire Hazard Index Formula, the NCHRP Report 50 Accident Prediction Formula, the US DOT Accident Prediction Formula, and the Transport Canada Accident Model. Other accident and hazard prediction models were mentioned in the study as well, including the Poisson regression model, the negative binomial regression model, the gamma probability model, and the Bayesian model. The US DOT Accident Prediction Formula was reported to be the most commonly used model by the state DOTs to rank HRGCs.

Some studies specifically focused on the development of custom statistical models for safety improvement at HRGCs. For example, Hu et al. [6] proposed a generalized logit model with stepwise variable selection to analyze accident severity at the HRGCs in Taiwan. Highway separation, number of daily trains, obstacle detection device, number of daily trucks, and presence of HRGC markings were found to be statistically significant factors that influenced accident severity. Yan et al. [7] applied a hierarchical tree-based regression model for predicting the train-vehicle accidents at passive HRGCs. The FRA accident history database was used in the analysis for a time period between 1980 and 2006. The study results highlighted that stop signs could be considered as an effective



countermeasure for passive HRGCs. Hao and Daniel [29] developed the ordered probit model to study the driver injury severity at the HRGCs with different types of warning devices (active vs. passive) in the USA. It was found that visibility, train speed, motor vehicle speed, driver's age, traffic volume, area type, and highway pavement could significantly affect the injury severity for both active and passive HRGC types in the USA. A similar study was conducted by Hao et al. [30], which aimed to investigate the driver injury severity levels for the truck and non-truck involved accidents at the HRGCs in the USA. The study concluded that speed control for truck drivers could substantially decrease the injury severity of truck drivers.

Lu and Tolliver [10] evaluated a wide range of different statistical models for accident prediction at HRGCs, including the Poisson model, the negative binomial model, the Conway–Maxwell–Poisson model, the gamma model, the Bernoulli model, the Poisson–Hurdle model, and zero-inflated Poisson model. The analysis, conducted for the HRGCs in North Dakota, demonstrated that the Poisson model was the most promising one. Hao et al. [31] investigated the effects of foggy conditions on injury levels for the HRGCs in the USA. It was found that older drivers typically suffered severe injuries in foggy conditions primarily due to longer reaction times and a decline in vision. Furthermore, the study highlighted that some drivers were still driving at fairly high speeds when approaching to the HRGCs despite foggy conditions.

Khan et al. [32] conducted a study for the HRGCs in North Dakota. A binary logistic regression model was developed to determine the factors that influence accident occurrence at the HRGCs. The study results showed that the accident likelihood was substantially affected by the number of daily trains, maximum speed of trains, number of through railroad tracks, and number of highway lanes. Moreover, the presence of pavement markings and the presence of populations within five miles of the HRGCs also could influence accident occurrence at the considered HRGCs. Some other studies focused on improving safety of roadway travelers at HRGCs in other countries as well, including Canada [33, 34], Great Britain [35, 36], Hungary [37], Finland [38], France [39, 40], and Australia [41–43].

### 2.3 Literature summary and contribution

A detailed review of the literature revealed that safety issues at HRGCs have received significant attention from the research community over the last decades. A number of state DOTs conducted various research projects, which aimed to reduce accident occurrence at HRGCs. Several studies have developed a variety of formulae to predict the

number of accidents or estimate a hazard index that was further used in ranking HRGCs for safety improvement projects. It was found that many state DOTs have been using the following accident and hazard prediction formulae [11, 12, 20]: (a) the US DOT Accident Prediction Formula; (b) the New Hampshire Hazard Index Formula; (c) the NCHRP Report 50 Accident Prediction Formula; and (d) the Peabody–Dimmick Formula. Certain state DOTs developed accident and hazard prediction models that are customized for their state needs and compared the performance of those models with the existing models that are nationally recognized.

Furthermore, some scientific studies developed alternative statistical models (such as the Poisson model, the negative binomial model, the Conway–Maxwell–Poisson model, the gamma model, the Bernoulli model, and the zero-inflated Poisson model among others) to evaluate the relationship between accidents at HRGCs and contributing factors [6, 10, 13, 29–32]. Typically, these statistical models are based on the same predictors as the ones that are adopted for the accident and hazard prediction formulae that have been used by the state DOTs. However, such statistical models are not generally used by the state DOTs for accident and hazard prediction at HRGCs due to their complexity. The main contributions of this study to the state-of-the-art and the state-of-the-practice include the following:

- Conduct a detailed up-to-date review of the nationally recognized accident and hazard prediction formulae that have been used by the state DOTs to rank the HRGCs in the USA for safety improvement projects.
- Conduct a detailed up-to-date review of the state-specific accident and hazard prediction formulae that have been used by the state DOTs to rank the HRGCs in the USA for safety improvement projects.
- Perform a comprehensive analysis of the major factors that are being considered in the existing accident and hazard prediction formulae.
- Discuss the reported performance of the identified accident and hazard prediction formulae as well as the reported implementation challenges.
- Suggest future research directions based on the existing challenges that are associated with implementation of the existing accident and hazard prediction formulae.

### 3 Categorization of the identified models

A detailed review of the relevant studies identified a total of 21 accident and hazard prediction formulae. The discovered formulae were classified into the following two major groups: (i) accident prediction formulae that can be

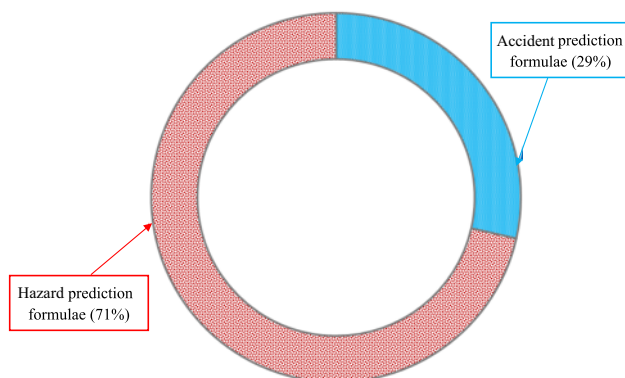
used to calculate the expected number of accidents at HRGCs for a particular time period and (ii) hazard prediction formulae that can be used to assess hazard or a safety index value for a given HRGC, which can be further considered throughout the safety improvement project selection. A distribution of the identified accident and hazard prediction formulae is presented in Fig. 2. Approximately 71% of the formulae (15 formulae out of 21 formulae) are the hazard prediction formulae, while the remaining 29% of the formulae (6 formulae out of 21 formulae) can be classified as the accident prediction formulae. A higher percentage of the hazard prediction formulae among the identified formulae can be explained by the fact that it is more difficult to forecast the expected number of accidents at HRGCs with a high degree of accuracy due to a large variety of factors (including human behavior that is difficult to emulate). Thus, the majority of the identified formulae aim to calculate the hazard index for HRGCs instead of the predicted number of accidents.

The following accident prediction formulae were discovered after a detailed state-of-the-practice review:

- (1) The Coleman–Stewart Model
- (2) The Iowa Accident Prediction Formula
- (3) The Jaqua Formula (developed by the State of Oregon)
- (4) The NCHRP Report 50 Accident Prediction Formula
- (5) The Peabody–Dimmick Formula
- (6) The US DOT Accident Prediction Formula

The following hazard prediction formulae were discovered after a detailed state-of-the-practice review:

- (1) The Arkansas Hazard Rating Formula
- (2) The California Hazard Rating Formula
- (3) The Connecticut Hazard Rating Formula
- (4) The Florida Accident Prediction and Safety Index Formula



**Fig. 2** A distribution of the identified accident and hazard prediction formulae

- (5) The Illinois Hazard Index Formula
- (6) The Kansas Design Hazard Rating Formula
- (7) The Michigan Hazard Index Formula
- (8) The Missouri Exposure Index Formula
- (9) The Nevada Hazard Index Formula
- (10) The New Hampshire Hazard Index Formula
- (11) The New Mexico Hazard Index Formula
- (12) The North Carolina Investigative Index Formula
- (13) The Revised Texas Priority Index Formula
- (14) The South Dakota Hazard Index Formula
- (15) The Texas Priority Index Formula

Certain states, including Alaska, North Dakota, and Washington, have been mostly utilizing customized spreadsheets without application of any particular accident or hazard prediction formulae to rank HRGCs for safety improvement projects. A number of hazard prediction formulae, such as the Connecticut Hazard Rating Formula, the Texas Priority Index Formula, and the New Mexico Hazard Index Formula, are inspired by the nationally recognized New Hampshire Hazard Index Formula. Certain state DOTs modified the canonical New Hampshire Hazard Index Formula and included some additional factors in the hazard prediction to improve the formula accuracy (such as train speed, type of train, sight distance, vehicle speed, highway–rail grade crossing angle, number of school buses, presence of the nearby intersection, and number of accidents among others). Moreover, the US DOT Accident Prediction Formula has been used by many states over the years (e.g., Alabama, Idaho, Indiana, Maine, Maryland, Ohio, South Carolina, Utah, Virginia, and Wisconsin).

A detailed description and formulation of the identified nationally recognized accident and hazard prediction models is presented in the following sections of this paper. The nationally recognized accident and hazard prediction models include the Coleman–Stewart Model, the NCHRP Report 50 Accident Prediction Formula, the New Hampshire Hazard Index Formula, the Peabody–Dimmick Formula, and the US DOT Accident Prediction Formula [8, 27]. Moreover, the other accident and hazard prediction models and resource allocation procedures used by the state DOTs are comprehensively described in “Appendix” of this paper. Note that the description of the accident and hazard prediction formulae was prepared using the available HRGC safety literature and the reviewed studies [11, 12, 19–28, 44–46].

### 3.1 Coleman–Stewart Model

The Coleman–Stewart Accident Prediction Model assesses the relationship between the observed accident rates and some factors that may cause accidents at HRGCs (e.g., daily train movements, daily vehicular movements, and

existing protection). Moreover, the model assumes that HRGCs are identical if they have similar characteristics, including the warning device type, location, number of tracks, and highway and traffic volumes. The model can be expressed as follows [11, 45]:

$$\lg A = B_0 + B_1 \lg C + B_2 \lg T_{cs} + B_3 (\lg T_{cs})^2, \quad (1)$$

where  $A$  is the average number of accidents per HRGC per year;  $C$  are the average daily vehicular movements (if  $C = 0$ , use 0.5 instead);  $T_{cs}$  are the average daily train movements (if  $T_{cs} = 0$ , use 0.5 instead);  $B_0$ ,  $B_1$ ,  $B_2$ , and  $B_3$  are the coefficients of the accident prediction equation (Table 3).

Although the variation among a collection of HRGCs in a given location may affect accident prediction at individual HRGCs, the Coleman–Stewart Model does not account for such a variation.

### 3.2 NCHRP Report 50 Accident Prediction Formula

The National Cooperative Highway Research Program (NCHRP) Report 50 proposed a formula for predicting accidents at HRGCs. The formula considers a number of factors, including the number of trains per day, existing warning devices, number of highway vehicles per day, and location (urban/rural). The NCHRP Report 50 Accident Prediction Formula can be expressed as follows [8, 11, 23, 27, 45]:

$$N = A \times B \times T_{nc}, \quad (2)$$

where  $N$  is the number of accidents per year;  $A$  is the factor based on the number of highway vehicles per day;  $B$  is the factor based on the existing warning devices and urban/rural classification;  $T_{nc}$  is the current train volume per day.

The values of  $A$  and  $B$  can be found in Table 4 and Fig. 3 [23, 45].

### 3.3 New Hampshire Hazard Index Formula

The New Hampshire Hazard Index Formula presents a fairly simple approach for estimating the hazard indexes at HRGCs. The hazard indexes are further used to rank HRGCs based on the probability of accidents. The formula states that the hazard index is dependent on the average daily volume of trains, average daily volume of vehicles, and type of warning devices installed at a given HRGC. The New Hampshire Hazard Index Formula can be expressed as follows [8, 27, 45]:

$$NHHI = V \times T_{nh} \times PF, \quad (3)$$

where  $NHHI$  denotes the New Hampshire Hazard Index;  $V$  is the annual average daily traffic;  $T_{nh}$  is the average daily volume of trains;  $PF$  is the protection factor (1.00 for stop signs, 0.60 for flashing lights, and 0.10 for gates).

Although a number of state DOTs have used the canonical New Hampshire Hazard Index Formula, certain states have customized the formula with the introduction of some variables in order to improve the accuracy of the formula in predicting safety hazard at HRGCs. Some of the additional variables include the train speed, sight distance, vehicle speed, HRGC width, HRGC angle, type of train,

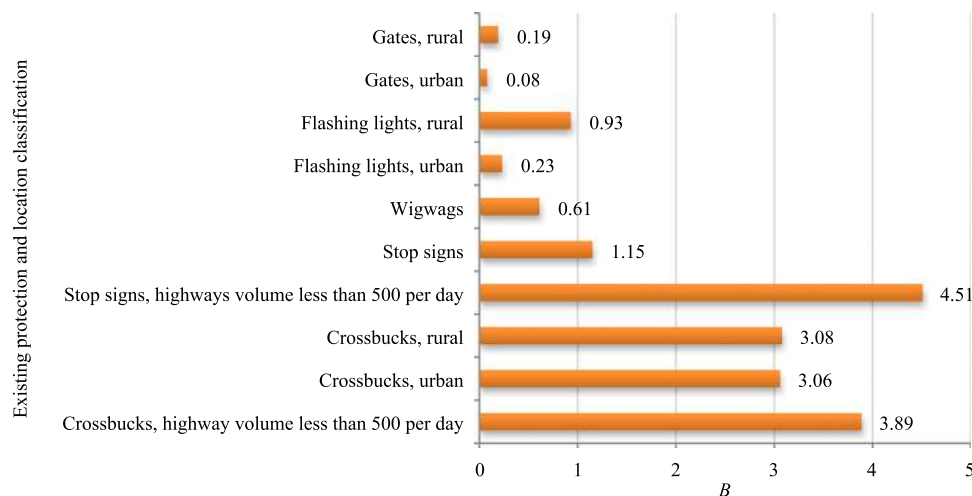
**Table 3** Coleman–Stewart Model coefficients

Item	$B_0$	$B_1$	$B_2$	$B_3$	Item	$B_0$	$B_1$	$B_2$	$B_3$
<i>Single-track urban</i>					<i>Multiple-track urban</i>				
Automatic gates	− 2.17	0.16	0.96	− 0.35	Automatic gates	− 2.58	0.23	1.30	− 0.42
Flashing lights	− 2.85	0.37	1.16	− 0.42	Flashing lights	− 2.50	0.36	0.68	− 0.09
Crossbucks	− 2.38	0.26	0.78	− 0.18	Crossbucks	− 2.49	0.32	0.63	− 0.02
Other active	− 2.13	0.30	0.72	− 0.30	Other active	− 2.16	0.36	0.19	0.08
Stop signs	− 2.98	0.42	1.96	− 1.13	Stop signs	− 1.43	0.09	0.18	0.16
None	− 2.46	0.16	1.24	− 0.56	None	− 3.00	0.41	0.63	− 0.02
Item	$B_0$	$B_1$	$B_2$	$B_3$	Item	$B_0$	$B_1$	$B_2$	$B_3$
<i>Single-track rural</i>					<i>Multiple-track rural</i>				
Automatic gates	− 1.42	0.08	− 0.15	0.25	Automatic gates	− 1.63	0.22	− 0.17	0.05
Flashing lights	− 3.56	0.62	0.92	− 0.38	Flashing lights	− 2.75	0.38	1.02	− 0.36
Crossbucks	− 2.77	0.40	0.89	− 0.29	Crossbucks	− 2.39	0.46	− 0.50	0.53
Other active	− 2.25	0.34	0.34	− 0.01	Other active	− 2.32	0.33	0.80	− 0.35
Stop signs	− 2.97	0.61	− 0.02	0.29	Stop signs	− 1.87	0.18	0.67	− 0.34
None	− 3.62	0.67	0.22	0.26	None	−	−	−	−



**Table 4** A values based on 10-year annual average daily traffic (AADT)

Vehicles per day (10-year AADT)	A	Vehicles per day (10-year AADT)	A
250	0.000347	9000	0.011435
500	0.000694	10,000	0.012674
1000	0.001377	12,000	0.015012
2000	0.002627	14,000	0.017315
3000	0.003981	16,000	0.019549
4000	0.005208	18,000	0.021736
5000	0.006516	20,000	0.023877
6000	0.007720	25,000	0.029051
7000	0.009005	30,000	0.034757
8000	0.010278		


**Fig. 3** B values based on the existing protection and location classification

population, surface type, number of school buses, number of buses, surface condition, number of tracks, presence of a nearby intersection, vertical alignment, functional class of highway, horizontal alignment, number of hazardous material trucks, number of accidents, number of passengers, etc. [23].

### 3.4 Peabody–Dimmick Formula

The Peabody–Dimmick Formula was developed by the US Bureau of Public Roads using the data obtained from 3563 rural HRGCs in 29 states. The formula computes the expected number of accidents for 5 years based on some factors, including the average daily train traffic, annual average daily traffic, and protection coefficient (dependent on the warning device type) using the following equation [8, 23, 27, 45]:

$$A_5 = K + \frac{1.28V^{0.170} \times T_{pd}^{0.151}}{P^{0.171}}, \quad (4)$$

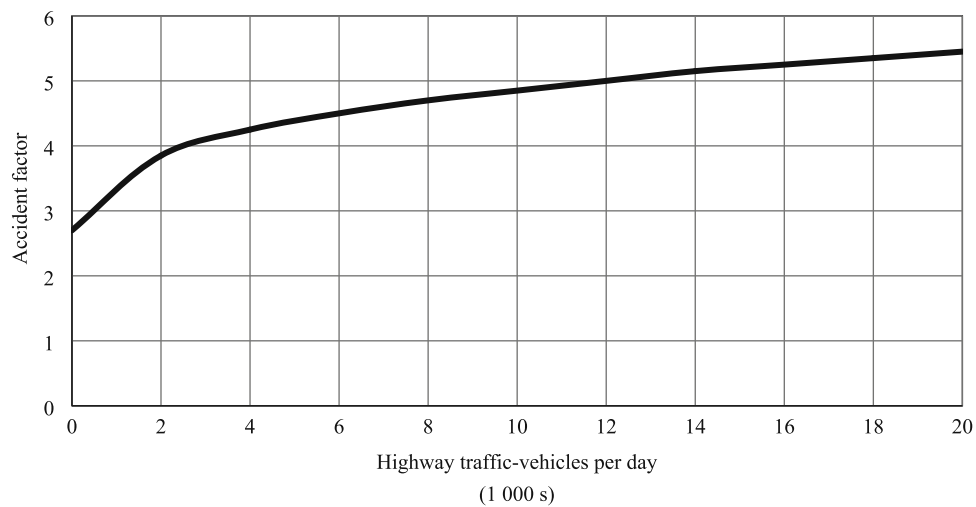
where  $A_5$  is the expected number of accidents in 5 years;  $V$  is the annual average daily traffic factor;  $T_{pd}$  is the average daily train traffic factor;  $P$  is the protection coefficient;  $K$  is the additional parameter.

In order to estimate the additional parameter  $K$ , an unbalanced accident factor ( $l_u$ ) is required that can be calculated as follows [23]:

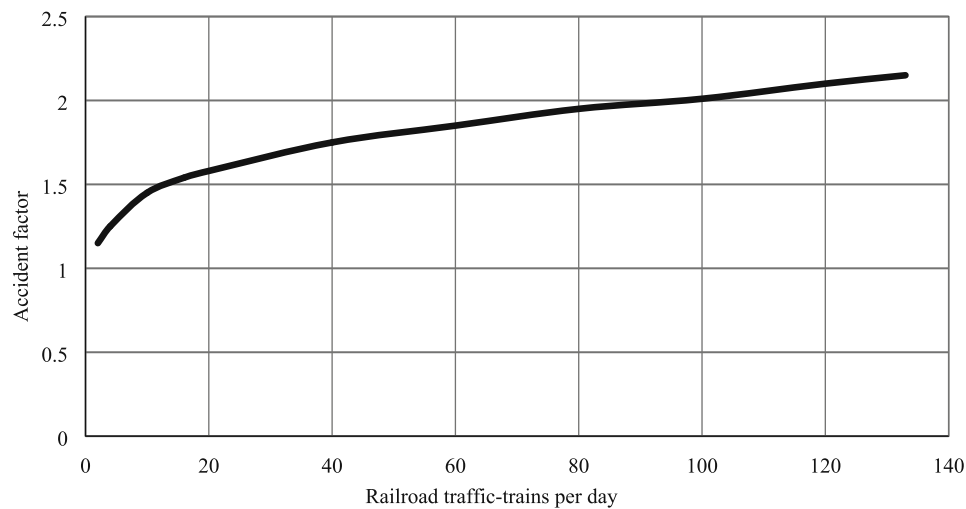
$$l_u = 1.28 \frac{V^a \times T^b}{P^c}, \quad (5)$$

where  $V^a$ ,  $T^b$  and  $P^c$  are the accident factors (i.e., the annual average daily traffic factor, the average daily train traffic factor, and the protection coefficient).

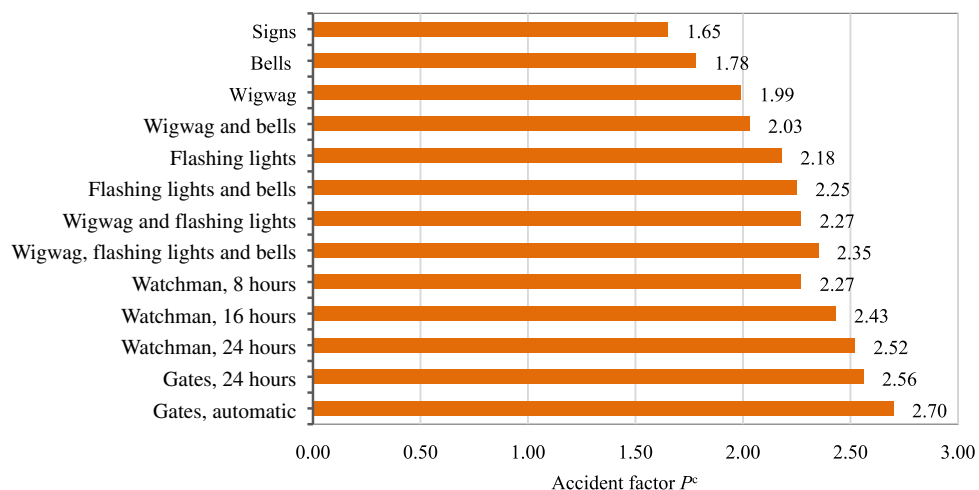
The expected number of accidents in 5 years ( $A_5$ ) can be identified using a set of charts, which are presented in Figs. 4, 5, 6 and 7.



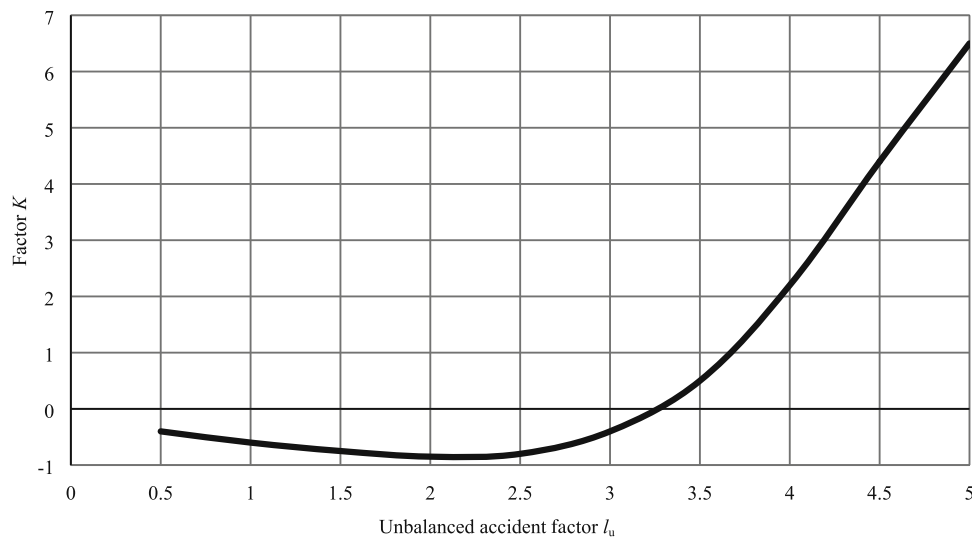
**Fig. 4** Relationship between highway traffic and accident factor  $V^a$



**Fig. 5** Relationship between railroad traffic and accident factor  $T^b$



**Fig. 6** Relationship between warning device type and accident factor  $P^c$



**Fig. 7** Relationship between  $K$ -factor and unbalanced accident factor  $I_u$

### 3.5 US DOT procedure for accident prediction and resource allocation

The expected number of accidents at an HRGC over a defined time period can be forecasted using an accident prediction model. Generally, accident prediction models take into account operational and physical characteristics of HRGCs to forecast accident occurrence. The US DOT developed an accident prediction model that would be able to assist the states with satisfying certain requirements under the Federal-Aid Policy Guidelines (FAPG) [44]. The US DOT Model uses a total of three equations to predict accidents at HRGCs. The first equation, which is referred to as the initial accident prediction formula, predicts the accident frequency per year at HRGCs using their operational and physical characteristics. The initial accident prediction formula is presented as follows [8, 21, 23, 27, 45]:

$$a = K \times EI \times MT \times DT \times HP \times MS \times HT \times HL, \quad (6)$$

where  $a$  denotes the initial accident prediction, accidents per year at an HRGC;  $K$  is the formula constant;  $EI$  is the factor for exposure index based on the product of highway and train traffic;  $MT$  is the factor for the number of main tracks;  $DT$  is the factor for the number of through trains per day during daylight;  $HP$  is the factor for highway paved (yes or no);  $MS$  is the factor for maximum timetable speed;  $HT$  is the factor for highway type;  $HL$  is the factor for the number of highway lanes.

The values for the factors of the initial accident prediction formula can be found in Table 5. The second accident prediction formula captures the accident history recorded for a given HRGC over a defined time period. A major assumption in the formula is that the future accidents

are expected to occur at the same rate as the past accidents. Hence, the formula is expected to give the most accurate results when the available accident history of up to 5 years is considered. However, the accident data, collected for the time period that exceeds 5 years, may produce misleading results, as substantial changes in operational and physical HRGC characteristics may occur over such a time period. The second accident prediction can be estimated as follows [8, 21, 23, 27, 45]:

$$B = \frac{T_0}{T_0 + T_y} (a) + \frac{T_0}{T_0 + T_y} \left( \frac{N}{T_y} \right), \quad (7)$$

where  $B$  denotes the second accident prediction (measured in accidents per year at an HRGC);  $a$  is the initial accident prediction (measured in accidents per year at an HRGC);  $\frac{N}{T_y}$  is the accident history prediction (measured in accidents per year at an HRGC), and  $N$  is the number of accidents that were observed in  $T_y$  years at an HRGC;  $T_0$  is the formula weighting factor,  $T_0 = \frac{1}{0.05 + a}$ .

The final accident prediction is further estimated by applying a normalizing constant ( $A$ ) to the second accident prediction for each HRGC category. The HRGC categories are differentiated based on the protection type into the following groups: (a) passive HRGCs, (b) HRGCs with flashing lights, and (c) HRGCs with gates. Application of normalizing constants ensures that the procedure is transformed from the past accident trends to the current accident trends. Thus, the normalizing constants have to be updated regularly. The normalizing constant can be estimated by equating the sum of the predicted accidents multiplied by the corresponding normalizing constant to the most recent number of accidents for each HRGC category over a specific time period [23, 46]. The final accident prediction

**Table 5** Highway–rail grade crossing characteristic factors for the Initial US DOT Accident Prediction Formula

Crossing category	Formula constant $K$	Exposure index factor $EI$	Main tracks factor $MT$	Day thru trains factor $DT$	Highway paved factor $HP$	Maximum speed factor $MS$	Highway type factor $HT$	Highway lanes factor $HL$
Passive	0.002268	$\left(\frac{ct+0.2}{0.2}\right)^{0.3334}$	$e^{0.2094mt}$	$\left(\frac{d+0.2}{0.2}\right)^{0.1336}$	$e^{-0.6160(hp-1)}$	$e^{0.0077ms}$	$e^{-0.1000(ht-1)}$	1.0
Flashing lights	0.003646	$\left(\frac{ct+0.2}{0.2}\right)^{0.2953}$	$e^{0.1088mt}$	$\left(\frac{d+0.2}{0.2}\right)^{0.0470}$	1.0	1.0	1.0	$e^{0.1380(hl-1)}$
Gates	0.001088	$\left(\frac{ct+0.2}{0.2}\right)^{0.3116}$	$e^{0.2912mt}$	1.0	1.0	1.0	1.0	$e^{0.1036(hl-1)}$

$c$  denotes the annual average number of highway vehicles per day (total in both directions);  $t$  is the average total train movements per day;  $mt$  is the number of main tracks;  $d$  is the average number of thru trains per day during daylight;  $hp$  is the highway paved (1.0 for paved and 2.0 for unpaved);  $ms$  is the maximum timetable speed in mph;  $ht$  is the highway type factor value;  $hl$  is the number of highway lanes

Highway type	Inventory code	$ht$
Rural		
Interstate	01	1
Other principal arterial	02	2
Minor arterial	06	3
Major collector	07	4
Minor collector	08	5
Local	09	6
Urban		
Interstate	11	1
Other freeway and expressway	12	2
Other principal arterial	14	3
Minor arterial	16	4
Collector	17	5
Local	19	6

can be used afterward to rank HRGCs for resource allocation. The periodic updates used for accident prediction as well as the normalizing constants for resource allocation procedure are presented in Table 6.

#### 4 Key factors considered in the identified models

Table 7 presents the list of the factors (or predictors) considered by the identified accident and hazard prediction formulae. A detailed review of the 21 accident and hazard prediction formulae, which were identified in the literature, revealed a total of 20 unique predictors that may directly influence the expected number of accidents at HRGCs. A distribution of the factors considered by the identified accident and hazard prediction formulae is presented in

Fig. 8. The analysis results indicate that all the identified accident and hazard prediction formulae directly account for the number of vehicles per day and the number of trains per day. Such a finding can be supported by the fact that the number of vehicles per day and the number of trains per day are the key factors that determine exposure of a given HRGC to highway and railway traffic. A total of 19 accident and hazard prediction formulae consider the protection type at HRGCs (i.e., type of warning devices used). Some other predictors that are commonly used by the identified accident and hazard prediction formulae include: (i) accident history—captured in 12 formulae; (ii) train speed—captured in 11 formulae; (iii) number of tracks—captured in 10 formulae; (iv) sight distance—captured in 8 formulae; and (v) number of traffic lanes—captured in 6 formulae.

**Table 6** Values of the normalizing constants for the accident prediction and resource allocation procedure

Warning device groups	New	Prior year constants							
	2010	2007	2005	2003	1998	1992	1990	1988	1986
Passive	0.4613	0.6768	0.6407	0.6500	0.7159	0.8239	0.9417	0.8778	0.8644
Flashing lights	0.2918	0.4605	0.5233	0.5001	0.5292	0.6935	0.8345	0.8013	0.8887
Gates	0.4614	0.6039	0.6513	0.5725	0.4921	0.6714	0.8901	0.8911	0.8131

This table was prepared using the data reported by FRA [46]

**Table 7** Factors considered by the identified accident and hazard prediction formulae

Accident/hazard prediction formulae	Factors considered
Arkansas Hazard Rating Formula	Trains per day; vehicles per day; accident history; number of tracks
California Hazard Rating Formula	Trains per day; vehicles per day; accident history; existing protection
Coleman–Stewart Model	Trains per day; vehicles per day; number of tracks; location; existing protection
Connecticut Hazard Rating Formula	Trains per day; vehicles per day; accident history; existing protection
Florida Accident Prediction and Safety Index Formula	Trains per day; vehicles per day; number of traffic lanes; train speed; sight distance; highway vehicular speed; school buses; accident history; existing protection
Illinois Hazard Index Formula	Trains per day; vehicles per day; train speed; number of traffic lanes; number of tracks; existing protection; accident history
Iowa Accident Prediction Formula	Trains per day; vehicles per day; time of day; daylight thru trains per day; existing protection; train speed; number of tracks; highway pavement type; accident history; number of traffic lanes
Jaqua Formula	Trains per day; number of cars in a train; type of train; train speed; number of tracks; vehicles per day; highway vehicular speed; sight distance; number of traffic lanes; approach gradient; angle of crossing; other roadway geometrics; location; existing protection
Kansas Design Hazard Rating Formula	Trains per day; vehicles per day; type of train; sight distance; angle of crossing; number of tracks
Michigan Hazard Index Formula	Trains per day; vehicles per day; existing protection
Missouri Exposure Index Formula	Existing protection; vehicles per day; sight distance; highway vehicular speed; trains per day; train speed; type of train
NCHRP Report 50 Accident Prediction Formula	Trains per day; vehicles per day; location; existing protection
Nevada Hazard Index Model	Trains per day; vehicles per day; existing protection; accident history; highway vehicular speed; number of tracks; train speed; angle of crossing
New Hampshire Formula	Trains per day; vehicles per day; existing protection
New Mexico Hazard Index Formula	Trains per day; vehicles per day; existing protection; sight distance; accident history; train speed
North Carolina Investigative Index Formula	Existing protection; school buses; vehicles per day; trains per day; type of train; number of tracks; train speed; sight distance; accident history
Peabody–Dimmick Formula	Trains per day; vehicles per day; existing protection
South Dakota Hazard Index Formula	Trains per day; vehicles per day; sight distance; existing protection
Texas Priority Index Formula	Trains per day; vehicles per day; existing protection; accident history; train speed
Revised Texas Priority Index Formula	Existing protection; location; highway pavement type; number of traffic lanes; sight distance; number of tracks; train speed; trains per day; vehicles per day; highway vehicular speed; presence of a nearby highway intersection; accident history
US DOT Accident Prediction Formula	Existing protection; trains per day; vehicles per day; daylight thru trains per day; highway pavement type; number of tracks; train speed; number of traffic lanes; location; accident history

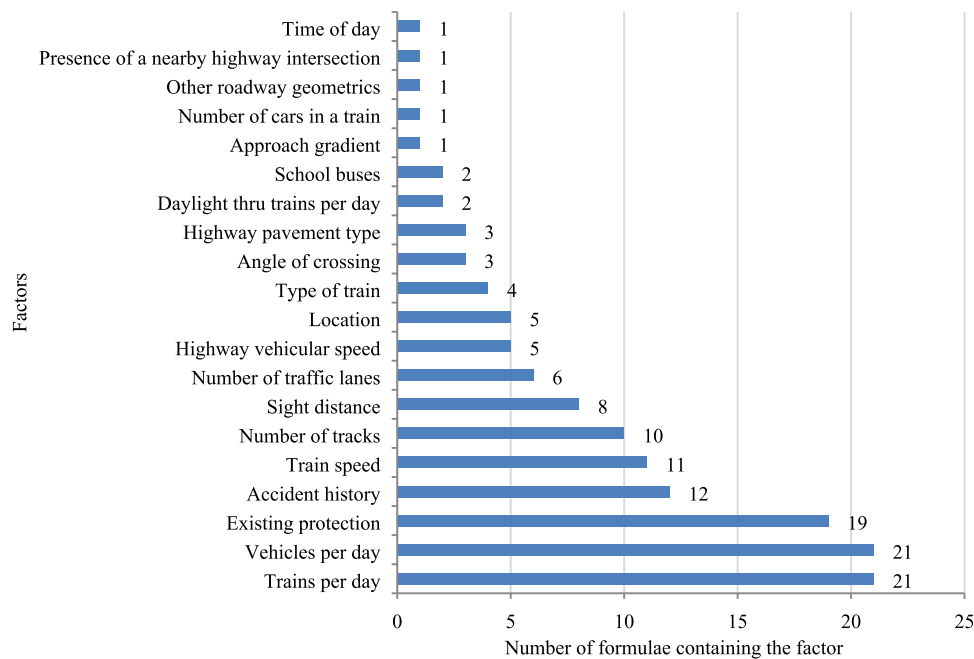
Certain predictors have been captured only by a very limited number of accident and hazard prediction formulae. For instance, only the Jaqua Formula, which has been deployed by the Oregon DOT, directly accounts for the number of cars in a train, approach gradient, and different roadway geometric features (e.g., existence of entrances and exits to the streets as well as the street intersections near an HRGC, and curvature of the roadway). Furthermore, the Florida Accident Prediction and Safety Index Formula and the North Carolina Investigative Index Formula are the only hazard index formulae that capture the number of school buses, which traverse a given HRGC on a daily basis. The Revised Texas Priority Index Formula is the only hazard index formula that explicitly considers presence of a nearby highway intersection throughout estimation of a hazard index for a given HRGC. Consideration of the additional predictors (i.e., number of cars in a

train, approach gradient, different roadway geometric features, number of school buses, presence of a nearby highway intersection) is expected to improve the performance of the accident and hazard prediction formulae. However, the existing HRGC inventory databases may not have any information regarding the additional predictors, and supplementary field reviews will have to be conducted in order to collect the required data.

## 5 Performance of the identified models and implementation challenges reported

Many of the reviewed studies investigated the performance of the existing accident and hazard prediction models. A review of some accident and hazard prediction models, which was conducted by Chadwick et al. [8], revealed that





**Fig. 8** A distribution of the factors considered by the identified accident and hazard prediction formulae

the US DOT Accident Prediction Formula is the most commonly used model by the state DOTs for prioritizing HRGCs for safety improvements. Moreover, the research efforts that were conducted by the States of Alabama, Virginia, and Ohio concluded that the US DOT Accident Prediction Formula outperformed the alternative accident and hazard prediction formulae for the HRGCs in the respective states and returned more accurate ranking of the HRGCs [19, 20, 28]. The Illinois Hazard Index Formula was found to be the most accurate formula for ranking the HRGCs in the State of Illinois [11]. Moreover, the Illinois Hazard Index Formula was found to be the most accurate for active HRGCs in the State of Missouri, while the California Hazard Rating Formula outperformed the other models for passive HRGCs [21]. The original Texas Priority Index Formula was outperformed by the Revised Texas Priority Index Formula when ranking the HRGCs in the State of Texas [24].

Austin and Carson [5] highlighted that some accident and hazard prediction formulae (including the New Hampshire Hazard Index Formula, the NCHRP Report 50 Accident Prediction Formula, and the Peabody–Dimmick Formula) lack descriptive capabilities because they consider only a limited number of explanatory variables. For instance, the Texas Priority Index, which is an extension of the New Hampshire Hazard Index Formula, does not capture certain important predictors, such as transport of hazardous materials, school bus usage, train speeds, and urban/rural distinction [27]. Furthermore, another issue with the Texas Priority Index consists in the fact that it

does not allow differentiating between two HRGCs with the same operational and physical features, which have one accident and no accidents over the past 5 years, respectively. In particular, both HRGCs will have exactly the same values of the Texas Priority Index despite substantial differences in the past accident history [27].

On the other hand, some other formulae (e.g., the US DOT Accident Prediction Formula) have a significant number of explanatory variables. However, there is a need to update the parameters in the US DOT Accident Prediction Formula to prevent a decline in accuracy of the model [5]. The normalizing constant values that are used in the US DOT Accident Prediction Formula significantly affect its accuracy and have to be adjusted over time. Moreover, the survey conducted by Bowman [20], which involved the state highway–rail program coordinators across the USA, revealed that the US DOT Accident Prediction Formula still fails to account for some important factors (such as roadway approach characteristics and quadrant sight distance) and is significantly dependent on the accident history. Several studies pointed out that the data availability is one of the major challenges that affect a successful implementation of accident and hazard prediction formulae [11, 20, 21, 28]. Certain predictors may not be available in the existing FRA and state HRGC databases (e.g., daily average train movements by type of train, number of school bus passengers, number of blind quadrants, and speed of each type of train). Field reviews will be required in order to collect the necessary data for implementation of certain accident and hazard prediction

formulae. Moreover, the state DOT representatives may have to contact railroad companies directly to obtain specific data.

Some accident and hazard prediction formulae (e.g., the US DOT Accident Prediction Formula, the California Hazard Rating Formula, and the New Hampshire Hazard Index Formula) generally give a higher priority value to the HRGCs with high exposure and may not accurately capture vulnerability of the HRGCs with low exposure [24, 27]. Aiming to address the latter challenges, Ryan and Mielke [27] proposed a new hazard index formula (named as “the Nevada Hazard Index Formula”) that accounted for several factors, including the daily train volume, average daily highway traffic, number of accidents within the past 5 years, protection factor, highway speed factor, number of near misses within the past 3 years, rail speed factor, HRGC angle factor, and track configuration factor. The Nevada Hazard Index Formula uses the square root of exposure to prevent the assignment of “too much weight” to the train and roadway traffic volumes. Another important issue that was pointed out by some of the studies consists in the fact that AADT and train counts may not be updated regularly by certain state DOTs [20, 26, 28]. The latter can further result in errors throughout accident and hazard prediction for HRGCs as well as misleading resource allocation.

Table 8 summarizes the implementation challenges that were identified throughout a detailed review of the state DOT efforts and proposes potential solutions to address these challenges. The issue of insufficient descriptive capabilities can be addressed by introducing additional factors in the accident and hazard prediction formulae, including train speed, sight distance, vehicle speed, HRGC width, HRGC angle, type of train, surface type, and

number of school buses [23, 27]. A periodic calibration of the model parameters is required to prevent a decline in their accuracy that may occur over time [5]. The issue of overreliance of certain models on the accident history can be addressed by considering the number of near misses in the models along with the number of past accidents [27]. Furthermore, the number of warrants met should be accounted for when ranking passive HRGCs [24]. Periodic field reviews are required in order to resolve the issue of the data unavailability. In certain cases, the state DOT representatives may have to contact railroad companies or other relevant stakeholders directly to obtain specific data. The use of the square root of exposure rather than the base exposure value can effectively resolve the issue of assigning “too much weight” to the train and roadway traffic volumes [27]. Periodic reviews of the HRGC databases based on the feedback from the appropriate stakeholders are required to resolve the data inconsistency issues. Furthermore, the information regarding operational and physical HRGC features should be updated in the HRGC databases with a sufficient frequency to ensure accurate accident and hazard prediction.

## 6 Concluding remarks and future research needs

Highway–rail grade crossings (HRGCs) are considered as one of the most dangerous segments of the transportation network. At every HRGC, there is a risk of an accident between a highway vehicle and a train. Other types of accidents, such as rear-end accidents between highway vehicles and collisions of vehicles with warning devices, may also occur at HRGCs. Although there has been a decline in the number of accidents at HRGCs over the past

**Table 8** Implementation challenges identified and potential solutions proposed

Implementation challenges	Potential solutions
Lack of descriptive capabilities	Introduce additional factors in the model (e.g., train speed, sight distance, vehicle speed, HRGC width, HRGC angle, type of train, surface type, number of school buses)
Decline in accuracy of the model constants	Conduct periodic calibration of the model parameters with a sufficient frequency depending on the changes in operational and physical HRGC features
Overreliance of certain models on the accident history	Consider the number of near misses in the model along with the number of past accidents; account for the number of warrants met when ranking passive HRGCs
Unavailability of the data required to apply the models	Perform periodic field reviews to collect the necessary data; contact railroad companies or other relevant stakeholders directly as needed to obtain specific data
Assignment of “too much weight” to the train and roadway traffic volumes	Apply the appropriate modifications in the model (e.g., use the square root of exposure rather than the base exposure value)
Inconsistencies and outdated information in the HRGC databases	Conduct periodic reviews of the HRGC databases to resolve the inconsistency issues based on the feedback from the appropriate stakeholders and update the information regarding operational and physical HRGC features with a sufficient frequency

years, thousands of accidents are still recorded annually in the USA, which result in fatalities, injuries, and staggering costs in property damages. Numerous studies have investigated the factors that may influence accident occurrence and developed various models to estimate the expected number of accidents at HRGCs as well as to assess the HRGC hazard. This paper presented a comprehensive review of the existing accident and hazard prediction formulae that have been used by the state Departments of Transportation (DOTs) over the years to prioritize the HRGCs for safety improvement projects in the USA. From an extensive literature review, a total of 21 accident and hazard prediction formulae were identified, five of which are nationally recognized accident and hazard prediction formulae. Furthermore, a number of customized accident and hazard prediction models, developed by certain state DOTs, were identified and evaluated as well.

The identified accident and hazard prediction formulae were divided into two categories based on the measure used in ranking HRGCs, which include: (i) accident prediction formulae (6 out of 21 formulae) that can be used to calculate the expected number of accidents at HRGCs for a particular time period; and (ii) hazard prediction formulae (15 out of 21 formulae) that can be used to assess hazard or a safety index value for a given HRGC, which can be further considered throughout the safety improvement project selection. A higher percentage of the hazard prediction formulae among the identified formulae can be explained by the fact that it is more difficult to forecast the expected number of accidents at HRGCs with a high degree of accuracy. A detailed analysis of the identified accident and hazard prediction formulae indicated that all the formulae directly accounted for the number of vehicles per day and the number of trains per day that determine exposure of a given HRGC to highway and railway traffic. A total of 19 accident and hazard prediction formulae considered the protection type at HRGCs (i.e., type of warning devices used).

Furthermore, this study discussed the reported performance and implementation issues, associated with the identified accident and hazard prediction formulae. The US DOT Accident Prediction Formula was found to be the most preferential among the state DOTs for prioritizing HRGCs for safety improvements. The key issues, associated with implementation of the identified accident and hazard prediction formulae, include: (a) limited number of explanatory variables in certain formulae (e.g., the NCHRP Report 50 Accident Prediction Formula, the New Hampshire Hazard Index Formula, and the Peabody–Dimmick Formula) causing a lack of descriptive capabilities; (b) decline in accuracy of the model constants (e.g., the US DOT Accident Prediction Formula normalizing constants have to be adjusted over time); (c) overreliance of certain accident

and hazard prediction formulae on the accident history; (d) unavailability of the data required to apply the models (such as sight distance, number of blind quadrants, angle of the intersection between the roadway and the track); (e) assignment of “too much weight” to the train and roadway traffic volumes (exposure value), which results in higher priority values for the HRGCs with higher traffic volumes as compared to the lower-volume HRGCs; and (f) inconsistencies and outdated information in the HRGC databases maintained by some state DOTs.

Future research can be extended in the following directions: (i) there is a need to evaluate the effects of improvement in technology on the performance of existing accident and hazard prediction models; (ii) there is a need for future studies to investigate the effects of driver behavior on accident occurrence at HRGCs; (iii) future studies may include additional driver and roadway geometric characteristics into the existing accident and hazard prediction models (for instance, the effects of driver fatigue, level of frustration, time of accident, age of driver, population distribution, etc. may directly influence accident occurrence at HRGCs); (iv) there is a need for practitioners to develop new procedures for creating HRGC databases that will address the existing data availability and inconsistency issues; (v) conduct a survey among the state DOT representatives to identify any changes in the accident and hazard prediction procedures that have occurred over the years; and (vi) compare the identified accident and hazard prediction formulae against customized statistical models (e.g., the Poisson model, the negative binomial model, the gamma model) for a group of selected HRGCs in the US and other countries as well.

**Acknowledgements** This study was supported by the Florida Department of Transportation (Grant Number BDV30-977-26). The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Florida Department of Transportation or the US Department of Transportation. The authors would like to thank Mr. Rickey Fitzgerald, Freight & Multimodal Operations Office Manager, for his involvement and valuable feedback throughout this study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix: Other accident and hazard prediction models for highway–rail grade crossings and resource allocation procedures used by the state DOTs

A number state DOTs have developed their own accident and hazard prediction formulae to achieve the following objectives: (i) estimate the number of accidents at HRGCs; (ii) assess potential hazard at HRGCs; and (iii) prioritize HRGCs for safety improvement projects. The models used by the state DOTs that were identified throughout the review of the literature include: (1) the Arkansas Hazard Rating Formula; (2) the California Hazard Rating Formula; (3) the Connecticut Hazard Rating Formula; (4) the Florida Accident Prediction and Safety Index Formula; (5) the Illinois Hazard Index Formula; (6) the Iowa Accident Prediction Formula; (7) the Kansas Design Hazard Rating Formula; (8) the Michigan Hazard Index Formula; (9) the Missouri Exposure Index Formula; (10) the Nevada Hazard Index Formula; (11) the New Mexico Hazard Index Formula; (12) the North Carolina Investigative Index Formula; (13) the Jaqua Formula (used by the State of Oregon); (14) the South Dakota Hazard Index Formula; (15) the Texas Priority Index Formula; and (16) the Revised Texas Priority Index Formula. The state-specific accident and hazard prediction formulae that were identified in the reviewed literature are presented in the following sections of the manuscript.

### Arkansas Hazard Rating Formula

A hazard rating index and field diagnostic team reviews have been used by the Arkansas Highway and Transportation Department (AHTD) to identify the HRGCs that require safety improvements. Moreover, the AHTD does not set any specific thresholds throughout resource allocation. HRGCs are selected for safety improvements based on the hazard rating index and available budget. The Arkansas Hazard Rating Formula is expressed as follows [11]:

$$AHR = HTP \times RTP \times ARP, \quad (8)$$

where *AHR* is the Arkansas Hazard Rating of an HRGC; *HTP* denotes the highway traffic points, 5 points maximum (depending on ADT). *RTP* denotes the railway traffic points, 5 points maximum. Up to 75% of the railway traffic points are dependent on the number of trains. The rest of the railway traffic points depend on the number of side and main tracks at an HRGC. *ARP* denotes accident record points, 4 points maximum (depending on the number of accidents over the past 15 years).

### California Hazard Rating Formula

The California Hazard Rating Formula estimates the hazard index of a given HRGC, which is a representative of the number of accidents. The formula requires the 10-year accident history for the HRGCs. The California Hazard Rating Formula is expressed as follows [11, 21]:

$$CaHI = \frac{V \times T \times PF}{1000} + AH, \quad (9)$$

where *CaHI* is the California Hazard Index; *V* is the number of vehicles; *T* is the number of trains; *PF* is the protection factor (1.00 for stop signs or crossbucks, 0.67 for wigwags, 0.33 for flashing lights, and 0.13 for gates); *AH* is the accident history (the total number of accidents in the last 10 years multiplied by a factor of “3”).

### Connecticut Hazard Rating Formula

The rating formula that has been used by the State of Connecticut is similar to the one used by the State of California as both formulae estimate the hazard index, not the number of accidents. However, the formulae differ based on the fact that the Connecticut Hazard Rating Formula requires the 5-year accident history, while the California Hazard Rating Formula considers the 10-year accident history. The Connecticut Hazard Rating Formula is expressed as follows [11, 21, 45]:

$$CoHI = \frac{(T + 1)(A + 1) \times AADT \times PF}{100}, \quad (10)$$

where *CoHI* is the Connecticut Hazard Index; *AADT* is the annual average daily traffic; *T* is the number of trains per day; *PF* is the protection factor (1.25 for passive warning devices, 1.00 for stop sign control, 0.75 for stop signs and protect control, 0.75 for manually activated traffic signals, 0.25 for railroad flashing lights, 0.25 for traffic signal control with preemption, 0.01 for gates with railroad flashing lights, and 0.001 for inactive rail line); *A* is the accident history (the total number of accidents in the last 5 years).

### Florida Accident Prediction and Safety Index Formula

The Florida DOT developed an accident prediction formula, which has been used to prioritize the HRGCs for safety improvement projects and estimate safety index at the HRGCs in the State of Florida. The Florida Accident Prediction and Safety Index Formula is expressed with the following equations [23]:

$$t_p = -8.075 + 0.318 \ln S_t + 0.484 \ln T + 0.437 \ln A + 0.387 \ln V_V \\ + \left(0.28 - 0.28 \frac{MASD}{RSSD}\right) + \left(0.33 - 1.23 \frac{MCSD}{RSSD}\right) + 0.15 N_{xbucks}, \quad (11)$$

$$y = e^{(0.968t_p + 1.109)} / 4, \quad (12)$$

$$t_a = -8.075 + 0.318 \ln S_t + 0.166 \ln T + 0.293 \ln A + 0.387 \ln V_V \\ + \left(0.28 - 0.28 \frac{MASD}{RSSD}\right) + 0.225(L - 2) - 0.233P_g, \quad (13)$$

$$y = e^{(0.938t_a + 1.109)} / 4, \quad (14)$$

where  $A$  denotes the vehicles per day or annual average daily traffic;  $L$  is the number of lanes;  $MASD$  is the actual minimum stopping sight distance along a highway;  $MCSD$  is the clear sight distance (ability to see an approaching train along a highway, recorded for the four quadrants established by the intersection of the railroad tracks and that highway);  $RSSD$  is the required stopping sight distance on wet pavement;  $S_t$  is the maximum speed of a train;  $T$  is the yearly average of the number of trains per day;  $t_a$  is  $\ln$  of predicted number of accidents in a 4-year period at HRGCs with active traffic control devices;  $t_p$  is  $\ln$  of predicted number of accidents in a 4-year period at HRGCs with passive traffic control devices;  $V_V$  is the posted vehicle speed limit unless geometrics dictates a lower speed;  $N_{xbucks}$  is the total number of crossbucks at an HRGC;  $P_g$  is the gate presence indicator (1 if gated, and 0 if not);  $y$  is the predicted number of accidents per year at an HRGC.

To account for the accident history, the number of accidents predicted at an HRGC per year ( $y$ ) is adjusted as follows:

$$Y = y \sqrt{\frac{H}{yP}}, \quad (15)$$

where  $Y$  is the accident prediction adjusted for the accident history;  $y$  is the accident prediction based on the regression model;  $H$  is the number of accidents for the 6-year history or since the year of last improvement;  $P$  is the number of years of the accident history period.

Using the accident prediction formula, a safety/hazard index approach was developed to rank the HRGCs in the State of Florida. An HRGC with a safety index value of 70

is considered as safe and does not require any safety improvements. Also, a safety index value of 60, which shows that one accident occurred in 9 years, is considered as marginal. The safety index is calculated using the predicted number of accidents per year that has been adjusted for the accident history as follows [23]:

$$R = X(1 - \sqrt{Y}), \quad (16)$$

where  $R$  is the safety index;  $Y$  is the adjusted accident prediction value;  $X = 90$  when less than 10 school buses per day traverse an HRGC,  $X = 85$  when 10 or more school buses per day traverse an HRGC with active traffic control devices without gates, and  $X = 80$  when 10 or more school buses per day traverse an HRGC with passive traffic control devices.

### Illinois Hazard Index Formula

The Illinois Hazard Index Formula was developed from the study that was performed by the State of Illinois, which aimed to identify various factors that influence accident occurrence at the HRGCs in the state. The Illinois Hazard Index Formula is expressed as follows [11, 21, 45]:

$$IHI = 10^{-6} A^{2.59088} B^{0.09673} C^{0.40227} D^{0.59262} (15.59N^{5.60977} + PF), \quad (17)$$

where  $IHI$  is the Illinois Hazard Index;  $A = \ln(ADT \times NTT)$ ;  $ADT$  is the average daily traffic;  $NTT$  is the number of total trains per day;  $B$  is the maximum timetable speed, mph;  $C$  is the number of main and other tracks;  $D$  is the number of highway lanes;  $N$  is the average number of accidents per year (generally, a 5-year period is considered);  $PF$  is the protection factor (86.39 for crossbucks, 68.97 for flashing lights, and 37.57 for gates).

### Iowa Accident Prediction Formula

The accident prediction formula that has been used by the Iowa DOT was developed based on the US DOT Accident Prediction Formula [25]. The accident prediction and severity calculations are conducted in the following order: (1) estimation of the exposure factor; (2) estimation of the predicted number of accidents; and (3) accident severity assessment. The exposure factor ( $EF$ ) can be estimated as follows [25]:



$$\begin{aligned}
 EF = & AADT^{12 \text{ a.m.} - 6 \text{ a.m.}} \times T^{12 \text{ a.m.} - 6 \text{ a.m.}} \\
 & + AADT^{6 \text{ a.m.} - 12 \text{ p.m.}} \times T^{6 \text{ a.m.} - 12 \text{ p.m.}} \\
 & + AADT^{12 \text{ p.m.} - 6 \text{ p.m.}} \times T^{12 \text{ p.m.} - 6 \text{ p.m.}} \\
 & + AADT^{6 \text{ p.m.} - 12 \text{ a.m.}} \times T^{6 \text{ p.m.} - 12 \text{ a.m.}}, \\
 & \text{divided by the GREATER of} \\
 & \left[ (AADT^{12 \text{ a.m.} - 6 \text{ a.m.}})^2 + (AADT^{6 \text{ a.m.} - 12 \text{ p.m.}})^2 \right. \\
 & \left. + (AADT^{12 \text{ p.m.} - 6 \text{ p.m.}})^2 + (AADT^{6 \text{ p.m.} - 12 \text{ a.m.}})^2 \right], \\
 & \text{or} \\
 & \left[ (T^{12 \text{ a.m.} - 6 \text{ a.m.}})^2 + (T^{6 \text{ a.m.} - 12 \text{ p.m.}})^2 \right. \\
 & \left. + (T^{12 \text{ p.m.} - 6 \text{ p.m.}})^2 + (T^{6 \text{ p.m.} - 12 \text{ a.m.}})^2 \right], \\
 E = & 1.35EF \times AADT \times T, \quad (19)
 \end{aligned}$$

where  $AADT^{12 \text{ a.m.} - 6 \text{ a.m.}}$  is the percentage of AADT between 12:00 a.m. and 6:00 a.m.;  $AADT^{6 \text{ a.m.} - 12 \text{ p.m.}}$  is the percentage of AADT between 6:00 a.m. and 12:00 p.m.;  $AADT^{12 \text{ p.m.} - 6 \text{ p.m.}}$  is the percentage of AADT between 12:00 p.m. and 6:00 p.m.;  $AADT^{6 \text{ p.m.} - 12 \text{ a.m.}}$  is the percentage of AADT between 6:00 p.m. and 12:00 a.m.;  $T^{12 \text{ a.m.} - 6 \text{ a.m.}}$  is the percentage of trains between 12:00 a.m. and 6:00 a.m.;  $T^{6 \text{ a.m.} - 12 \text{ p.m.}}$  is the percentage of trains between 6:00 a.m. and 12:00 p.m.;  $T^{12 \text{ p.m.} - 6 \text{ p.m.}}$  is the percentage of trains between 12:00 p.m. and 6:00 p.m.;  $T^{6 \text{ p.m.} - 12 \text{ a.m.}}$  is the percentage of trains between 6:00 p.m. and 12:00 a.m.;  $AADT$  is the annual average daily traffic;  $T$  is the number of trains per day;  $E$  is the exposure of a given HRGC.

The predicted number of accidents is dependent on the existing warning devices at an HRGC. Moreover, the initial number of predicted accidents is adjusted based on the 5-year accident history. The number of predicted accidents at HRGCs with passive warning devices can be estimated as follows [25]:

$$PA = 0.0006938[(E + 0.2)/0.2]^{0.37}[(d + 0.2)/0.2]^{0.1781} e^{(0.0077ms)} e^{[-0.5966(hp-1)]}, \quad (20)$$

$$PA^{\text{adj}} = 0.65 \frac{PA[1/(0.05 + PA)] + N_5}{[1/(0.05 + PA)] + 5}, \quad (21)$$

where  $PA$  denotes the predicted accidents;  $d$  is the daylight thru trains per day;  $ms$  is the maximum timetable speed;  $hp = 1$  if an HRGC is on a paved road, and  $hp = 2$  if an HRGC is on a dirt or gravel road;  $PA^{\text{adj}}$  is the adjustment of predicted accidents;  $N_5$  is the number of accidents in 5 years.

The number of predicted accidents at HRGCs with flashing lights can be estimated as follows [25]:

$$PA = 0.0003351[(E + 0.2)/0.2]^{0.4106}[(d + 0.2)/0.2]^{0.1131} e^{(0.1917mt)} e^{[0.1826(hl-1)]}, \quad (22)$$

$$PA^{\text{adj}} = 0.5001 \frac{PA[1/(0.05 + PA)] + N_5}{[1/(0.05 + PA)] + 5}, \quad (23)$$

where  $mt$  is the number of main tracks, and  $hl$  is the number of highway lanes.

The number of predicted accidents at HRGCs with flashing lights and gates can be estimated as follows [25]:

$$PA = 0.0005745[(E + 0.2)/0.2]^{0.2942}[(d + 0.2)/0.2]^{0.1781} e^{(0.1512mt)} e^{[0.142(hl-1)]}, \quad (24)$$

$$PA^{\text{adj}} = 0.5725 \frac{PA[1/(0.05 + PA)] + N_5}{[1/(0.05 + PA)] + 5}. \quad (25)$$

HRGC accidents can be classified in terms of severity as fatal, casualty, injury, and property-damage-only accidents. A number of factors are used in evaluating the severity of accidents at HRGCs, such as the number of tracks, train speed, number of switching trains, number of through trains, and type of location (rural or urban) [25]. The predicted number of accidents by severity category can be estimated as follows:

$$PA^{\text{fat}} = \frac{PA^{\text{adj}}}{1 + 440.9ms^{-0.9931}(T + 1)^{-0.0873}(ST + 1)^{0.0872} e^{(0.3571ht)}}, \quad (26)$$

$$PA^{\text{cas}} = \frac{PA^{\text{adj}}}{1 + [4.481ms^{-0.343} e^{0.1153mt} e^{0.2960ht}]}, \quad (27)$$

$$PA^{\text{inj}} = PA^{\text{cas}} - PA^{\text{fat}}, \quad (28)$$

$$PA^{\text{prop}} = PA^{\text{adj}} - PA^{\text{cas}}, \quad (29)$$

where  $PA^{\text{fat}}$  denotes the predicted fatal accidents;  $PA^{\text{cas}}$  is the predicted casualty accidents;  $PA^{\text{inj}}$  is the predicted injury accidents;  $PA^{\text{prop}}$  is the predicted property accidents;  $ST$  is the number of switch trains per day;  $ht = 1$  if an HRGC is located in urban settings;  $ht = 0$  if an HRGC is located in rural settings.

### Kansas Design Hazard Rating Formula

The Kansas Design Hazard Rating Formula computes a hazard index as a surrogate of the number of accidents at HRGCs. The factors, considered in the formula, include the number of highway vehicles, number of slow trains, number of fast trains, angle of the intersection between the roadway and the track, number of main tracks, and sight distances for all four quadrants. Note that any negative

**Table 9** Protection factor values used by the Michigan DOT

Protection type	Factor
Flashing light signals with cantilever arms, half-roadway gates, and traffic signal interconnection	0.05
Flashing light signals with cantilever arms and half-roadway gates	0.08
Flashing light signals with half-roadway gates	0.11
Flashing light signals with cantilever arms and traffic signal interconnect	0.24
Flashing light signals with cantilever arms	0.27
Flashing light signals	0.30
Stop and flag procedures	0.75
Stop sign	0.80
Reflectorized crossbuck with or without a yield sign	1.00

hazard index value is reset to zero. The Kansas Design Hazard Rating Formula is expressed as follows [11, 21]:

$$KDHR = \frac{A(B + C + D)}{4}, \quad (30)$$

where  $KDHR$  is the Kansas Design Hazard Rating;  $A = \frac{HT(2NFT + NST)}{400}$ ;  $HT$  is the highway traffic;  $NFT$  is the number of fast trains;  $NST$  is the number of slow trains (switch trains are not included);  $B = 2\sqrt[3]{\frac{8000}{SMSD}}$ ;  $SMSD$  is the sum of maximum sight distance 4 ways;  $C = \sqrt{\frac{90}{AI}}$ ;  $AI$  is the angle of intersection;  $D$  = main track factor (1.00 for 1 track, 1.50 for 2 tracks, 1.80 for 3 tracks, and 2.00 for 4 tracks).

### Michigan Hazard Index Formula

The Michigan Hazard Index Formula is based on the New Hampshire Hazard Index Formula (described earlier in the manuscript), which estimates an index to be used for ranking HRGCs [11, 45]. However, the State of Michigan has modified the protection factor ( $PF$ ) values that have been used in the canonical New Hampshire Hazard Index Formula. The values of the protection factor, used by the Michigan DOT, are shown in Table 9.

### Missouri Exposure Index Formula

An exposure index has been used by the Missouri DOT, taking into account the existing protection type at HRGCs. A number of factors are considered in estimating the exposure index, including the number and speed of vehicles, number of freight and passenger trains, speed of freight and passenger trains, number of switching movements, required and actual sight distance. The Missouri Exposure Index Formula is expressed as follows [11, 21]:

For passive to active upgrade:

$$MEI = TI + SDO \times TI. \quad (31)$$

For active upgrade:

$$MEI = TI, \quad (32)$$

where  $MEI$  is the Missouri Exposure Index;  $SDO$  is the sight distance obstruction factor,  $SDO = \frac{RSD - ASD}{RSD}$ ;  $RSD$  is the required sight distance;  $ASD$  is the actual sight distance;  $TI$  is the traffic index,  $TI = \frac{(VM \times VS)(FM \times FS + PM \times PS + 10SM)}{10,000}$ ;  $VM$  are the vehicle movements;  $VS$  is the vehicle speed;  $FM$  are the freight train movements;  $FS$  is the freight train speed;  $PM$  are the passenger train movements;  $PS$  is the passenger train speed;  $SM$  are the switching movements.

### Nevada Hazard Index Formula

Ryan and Mielke [27] developed a revised hazard index model to rank the HRGCs for safety improvement projects in the State of Nevada. Some of the factors, considered in the model, include the daily train volume, average daily highway traffic, number of accidents within the past 5 years, protection factor, highway speed factor, number of near misses within the past 3 years, rail speed factor, HRGC angle factor, and track configuration factor. The Nevada Hazard Index Formula is expressed as follows [27]:

$$NHI = \sqrt{EI} \times ANMF \times PF \times HSF \times RSF \times TCF \times CAF, \quad (33)$$

where  $NHI$  is the Nevada Hazard Index;  $EI$  is the exposure index, written as the product of average daily highway traffic and daily train volume, and  $ANMF$  is the accident and near miss factor,  $ANMF = 1.3^{A+\frac{N}{3}}$ ;  $A$  is the number of the accidents within the past 5 years;  $N$  is the number of near misses within the past 3 years;  $PF$  is the protection factor, 0.15 for 4 quad gate or gates with medians, 0.30 for gates only, and 1.00 for flashing lights or passive;  $HSF$  is the highway speed factor; 0.50 for 0 to 15 mph, 1.00 for

20 to 35 mph, 1.50 for 40 to 65 mph, and 2.00 for 70 mph or above;  $RSF$  is the rail speed factor, 1.00 for 0 to 59 mph, and 1.50 for 60 mph and above;  $TCF$  is the track configuration factor; 1.25 for 1 siding/other track; 1.50 for 2 siding/other tracks; and 2.00 for 3 or more siding/other tracks;  $CAF$  is the HRGC angle factor, 2.00 for  $0^\circ$  to  $30^\circ$ , 1.50 for  $30^\circ$  to  $60^\circ$ , and 1.00 for  $60^\circ$  to  $90^\circ$ .

### New Mexico Hazard Index Formula

A hazard index formula has been used by the New Mexico State Highway and Transportation Department. The formula was developed from the Modified New Hampshire Hazard Index Formula. The formula estimates a hazard index, which is used to prioritize the HRGCs for safety improvements in the State of New Mexico. The New Mexico Hazard Index Formula is expressed as follows [11]:

$$NMHI = \frac{TADT \times VADT \times PF}{100} SD_f \times T_s \times AH_f, \quad (34)$$

where  $NMHI$  is the New Mexico Hazard Index;  $TADT$  is the train ADT;  $VADT$  is the highway vehicle ADT;  $PF$  is the protection factor, 0.11 for gates, 0.20 for lights, 0.34 for wigwags, 0.58 for signs, 1.00 for crossbucks, and 2.00 for no protection;  $SD_f$  is the sight distance factor, 1.0 for no restrictions, 1.2 for restrictions at one quadrant, and 1.5 for restrictions at more than one quadrant;  $T_s$  is the train speed in mph;  $AH_f$  is the accident history factor,  $AH_f = A + B + C$ ;  $A = 0.10$  for each property damage only accident;  $B = 0.20$  for each injury accident;  $C = 0.30$  for each fatal accident.

### North Carolina Investigative Index Formula

The Investigative Index Formula has been used by the North Carolina DOT to rank the HRGCs for safety improvement projects in the State of North Carolina. The formula considers three main factors, which include the accident history, exposure, and sight distance. The North Carolina Investigative Index Formula is expressed as follows [11]:

$$NCII = \frac{PF \times ADT \times TV \times TSF \times TF}{160} + (70A/Y)^2 + SDF, \quad (35)$$

where  $NCII$  is the North Carolina Investigative Index;  $PF$  is the protection factor, 1.0 for no warning devices or crossbucks, 0.50 for traffic signals, 0.20 for flashing lights, and 0.10 for gates;  $ADT$  is the average daily traffic. When school buses use an HRGC, add  $N_p/1.2$  ( $N_p$  denotes No. of school bus passengers) to  $ADT$ . When passenger trains use an HRGC, multiply  $ADT$  by the average vehicle

occupancy, which is 1.2;  $TV$  is the train volume;  $TSF$  is the train speed factor,  $TSF = V_{ma}/50 + 0.8$ , and  $V_{ma}$  is the maximum allowable train speed;  $TF$  is the track factor, which depends on the number of through tracks and the number of total tracks;  $A/Y$  is the number of train-vehicle accidents per year, and a 10-year accident history is required for the model;  $SDF$  is the sight distance factor,  $SDF = 16 \sum (SDF_n/4)$ ;  $SDF_n$  is the sight distance factor for quadrant  $n$  (0 for clear sight, 2 for average sight, and 4 for poor sight).

### Oregon

The Jaqua Formula has been used by the Oregon DOT to predict the number of accidents at the HRGCs in the State of Oregon. The Jaqua Formula is expressed as follows [11]:

$$ACC5 = \frac{ABC}{1610}, \quad (36)$$

$$A = \sum_{i=1}^n T_i \left( \frac{C_i V}{3S_i} + V \right), \quad (37)$$

where  $ACC5$  denotes accident prediction for the next 5 years;  $A$  is the exposure factor;  $n$  is the number of train types;  $T_i$  is the number of trains of type  $i$ ;  $C_i$  is the number of cars in a train of type  $i$ ;  $S_i$  is the speed of a train of type  $i$ ;  $V$  is the AADT;  $B$  is the hazard rating, which depends on the number of tracks, speed of vehicles and trains, number of blind quadrants, number of lanes, angle of intersection, approach grade, curvature of the roadway, existence of entrances and exits to the streets as well as the street intersections near an HRGC;  $C$  is the protection factor, which depends on the type of area (urban vs. rural) and the type of existing warning devices at the HRGCs.

### South Dakota Hazard Index Formula

The State of South Dakota developed a hazard index formula to prioritize the HRGCs for safety improvement projects. The factors, considered in estimating the hazard index of candidate HRGCs, include the average daily highway traffic, train traffic, obstruction factor, and HRGC protection factor. The South Dakota Hazard Index Formula is expressed as follows [11]:

$$SDHI = \frac{TV \times ADT \times PF \times OF}{5}, \quad (38)$$

where  $SDHI$  is the South Dakota Hazard Index;  $TV$  is the number of trains per day;  $ADT$  is the average daily highway traffic;  $PF$  is the HRGC protection factor;  $OF$  is the obstruction factor.

### Texas Priority Index Formula

Several states have adopted the Texas Priority Index. The formula is similar to the New Hampshire Hazard Index Formula. However, unlike the New Hampshire Hazard Index Formula, the Texas Priority Index Formula considers certain additional factors, such as accident history and train speed. Moreover, the formula distinguishes between the cantilever and mast-mounted flashing lights. Despite the fact that the 5-year accident history is considered, the priority index remains the same if no accidents or one accident are recorded for an HRGC. The Texas Priority Index Formula is expressed as follows [11, 27, 45]:

$$TPI = V \times T \times 0.1S \times PF \times 0.01A^{1.15}, \quad (39)$$

where  $TPI$  is the Texas Priority Index;  $V$  is the average daily traffic volume;  $T$  is the average daily train volume;  $S$  is the train speed;  $PF$  is the protection factor (1.00 for passive, 0.70 for mast-mounted flashing lights, 0.15 for cantilever flashing lights, and 0.10 for gates);  $A$  are the train accidents in the past 5 years (default = 1).

### Revised Texas Priority Index Formula

Researchers and practitioners found that the original Texas Priority Index Formula gives a higher priority ranking to the HRGCs with high traffic volume based on the accident history. In 2013, the Texas DOT performed a study in collaboration with the Texas A&M Transportation Institute and the University of Texas at San Antonio, aiming to address the existing drawbacks of the original Texas Priority Index Formula. The study proposed a new formula, called “the Revised Texas Priority Index Formula.” The Revised Texas Priority Index Formula is expressed as follows [24]:

$$TPI_{rev} = 1000\hat{\mu}(A_5 + 0.1), \quad (40)$$

$$\begin{aligned} \hat{\mu} = \exp[ & -6.9240 + PF + 0.2587hp - 0.3722ht \\ & + 0.0706hl + 0.0656tt + 0.0022ASD + 0.0143mst^{\max} \\ & + 0.0126mst^{\min} + 1.0024\lg(T + 0.5) \\ & + 0.4653\lg(AADT) - 0.2160NIP + 0.0092msv], \end{aligned} \quad (41)$$

where  $TPI_{rev}$  is the Revised Texas Priority Index;  $\hat{\mu}$  is the predicted number of accidents per year at an HRGC;  $PF$  is the protection factor (0.5061 for flashing lights,  $-0.2006$  for gates, and 0.0000 for passive);  $hp$  is the highway pavement factor (1 for paved; 2 for unpaved);  $ht$  is the urban/rural designation factor (1 for urban and 2 for rural);  $hl$  is the number of traffic lanes;  $tt$  is the number of the main and other tracks;  $ASD$  is the actual sight distance, approach 1;  $mst^{\max}$  is the maximum train speed (through

trains);  $mst^{\min}$  is the minimum train speed (switching trains);  $T$  is the daily train volume;  $AADT$  is the vehicular AADT;  $NIP$  is the nearby roadway intersection presence (1 if present and 2 if not present);  $msv$  is the higher roadway speed limit between approach 1 and approach 2;  $A_5$  is the number of accidents in the last 5 years at an HRGC.

In order to reduce the bias in ranking both passive and active HRGCs, an adjustment factor for the Revised Texas Priority Index was developed (since active HRGCs are likely to receive a higher priority ranking due to a higher accident record over a 5-year time period). The adjustment factor for a given warranted passive HRGC can be estimated using the following equation [24]:

$$AF_{pas} = 1.5(nw + c), \quad (42)$$

where  $AF_{pas}$  is the adjustment factor for warranted passive HRGCs;  $nw$  is the number of warrants met;  $c$  is the number of accidents in the most recent 5-year period.

### Other states

A review of the literature revealed that the States of Alaska, North Dakota, and Washington developed their state-specific procedures for accident and hazard prediction at HRGCs as well as resource allocation. Specifically, the Alaska Department of Transportation and Public Facilities has been using the Accident Prediction Value (APV) computational methods, described in the Railroad-Highway Grade Crossing Handbook [11]. A sufficiency rating system, which is called “Performance Appearance Rating (PAR),” has been used by the North Dakota DOT to prioritize the HRGCs for safety improvement projects. The PAR value is estimated based on visibility, exposure, railroad conditions, and highway conditions. The HRGCs with the lowest PAR rating are selected for safety improvements [11]. A priority matrix and a field review matrix have been used by the State of Washington to rank the HRGCs for safety improvement projects [11]. The matrices account for a number of factors, including the accident history, sight distance, number of vehicles per day, crossing angle, number of tracks, heavy truck percentage, bus presence, and train speed.

The States of Alabama, Idaho, Indiana, Maine, Maryland, Ohio, South Carolina, Utah, Virginia, and Wisconsin have been using the US DOT Accident Prediction Formula [11, 28]. The US DOT Accident Prediction Formula has been used by the State of Indiana. Nevertheless, the state does not apply any threshold values for prioritization of the HRGCs for safety improvement projects. A benefit–cost analysis has been also used by the State of Indiana in order to accurately select the HRGCs for upgrading [11]. Ryan and Mielke [27] highlighted that the State of Arizona has been using the Texas Priority Index for ranking the existing



HRGCs. A Modified New Hampshire Hazard Index Formula has been used for the HRGCs in the State of Louisiana [11]. The New Jersey DOT does not use any particular formula for predicting the number of accidents or the hazard index at the HRGCs. The state DOT primarily relies on the accident history throughout recourse allocation among the HRGCs [11].

The South Carolina DOT has been using the US DOT Accident Prediction Formula along with various criteria (including school bus crossings, hazardous material hauling on the roadway, sight distance, passenger rail service, and implementation feasibility) for prioritization of the HRGCs [11]. The US DOT Accident Prediction Formula has been used by the Virginia DOT along with some additional factors (i.e., sight distance, vehicle type, roadway geometric characteristics, and adjacent land use development) in order to identify the HRGCs that require safety improvements [11]. The Wisconsin DOT has been also using the US DOT Accident Prediction Formula for prediction of the expected number of accidents at the HRGCs in the State of Wisconsin. Both maintenance and installation costs for the selected countermeasures are generally considered throughout recourse allocation among the HRGCs. The Wisconsin DOT typically gives a serious consideration to the HRGCs that experienced more than one accident in past 10 years [11].

## References

1. Federal Motor Carrier Safety Administration (2015) Highway–rail grade crossing safety. <https://www.fmcsa.dot.gov/safety/rail-crossing/highway-rail-grade-crossing-safety>. Accessed 06 June 2019
2. Kavoosi M, Dulebenets MA, Pasha J et al (2020) Development of algorithms for effective resource allocation among highway–rail grade crossings: a case study for the State of Florida. *Energies* 13(6):1419
3. FDOT (2010) The Florida rail system plan: investment element. <http://www.fdot.gov/rail/PlanDevel/Documents/FinalInvestmentElement/A-2010FLRailPlan-InvestmentElement.pdf>. Accessed 24 Sept 2018
4. FRA (2018) Accident/incident data. [https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on\\_the\\_fly\\_download.aspx](https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx). Accessed 10 Oct 2018
5. Austin RD, Carson JL (2002) An alternative accident prediction model for highway–rail interfaces. *Accid Anal Prev* 34(1):31–42
6. Hu S, Li C, Lee CK (2010) Investigation of key factors for accident severity at railroad grade crossings by using a logit model. *Saf Sci* 48(2):186–194
7. Yan X, Richards S, Su X (2010) Using hierarchical tree-based regression model to predict train–vehicle crashes at passive highway–rail grade crossings. *Accid Anal Prev* 42(1):64–74
8. Chadwick S, Zho N, Saat MR (2014) Highway–rail grade crossing safety challenges for shared operations of high-speed passenger and heavy freight rail in the U.S. *Saf Sci* 68:128–137
9. Rezvani AZ, Peach M, Thomas A et al (2015) Benefit-cost methodology for highway–railway grade crossing safety protocols as applied to transportation infrastructure project prioritization processes. *Transp Res Procedia* 8:89–102
10. Lu P, Tolliver D (2016) Accident prediction model for public highway–rail grade crossings. *Accid Anal Prev* 90:73–81
11. Elzohairy Y, Benekohal R (2000) Evaluation of expected accident frequency formulas for rail–highway crossings. A technical report prepared for Illinois Department of Transportation. <http://www.idot.illinois.gov/Assets/uploads/files/Transportation-System/Research/Illinois-Transportation-Research-Center/2000.09.01%20-%20Evaluation%20of%20Expected%20Accident%20Frequency%20Formulas%20for%20Rail-Highway%20Crossings%20-%20VC-HR1%20FY98.pdf>. Accessed 24 Sept 2018
12. Dulebenets MA, Moses R, Sobanjo J et al (2020) Development of the optimization model for improving safety at rail crossings in Florida. A technical report prepared for Florida Department of Transportation. <https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/research/reports/fdot-bdv30-977-26-rpt.pdf>. Accessed 1 Mar 2020
13. Oh J, Washington SP, Doohee N (2006) Accident prediction model for railway–highway interfaces. *Accid Anal Prev* 38(2):346–356
14. Fu B, Giossi RL, Persson R et al (2020) Active suspension in railway vehicles: a literature survey. *Railw Eng Sci* 28(1):1–33
15. Zeng J, Wei L, Wu P (2016) Safety evaluation for railway vehicles using an improved indirect measurement method of wheel–rail forces. *J Mod Transp* 24(2):114–123
16. Al-Douri YK, Tretten P, Karim R (2016) Improvement of railway performance: a study of Swedish railway infrastructure. *J Mod Transp* 24(1):22–37
17. Zhao Y, Chen S, Tan X et al (2013) New technologies for high-risk tunnel construction in Guiyang–Guangzhou high-speed railway. *J Mod Transp* 21(4):258–265
18. Celebi BT, Kaymakci OT (2016) Verifying the accuracy of interlocking tables for railway signaling systems using abstract state machines. *J Mod Transp* 24(4):277–283
19. Faghri A, Demetsky M (1986) Evaluation of methods for predicting rail–highway crossing hazards. A technical report prepared for Virginia Department of Transportation. [http://www.virginiadot.org/vtrc/main/online\\_reports/pdf/86-r32.pdf](http://www.virginiadot.org/vtrc/main/online_reports/pdf/86-r32.pdf). Accessed 24 Sept 2018
20. Bowman BL (1994) Assessment of the State of Alabama rail–highway safety program. A technical report prepared for Alabama Department of Transportation. <http://www.eng.auburn.edu/files/centers/hrc/Multimodal%20Transportation.pdf>. Accessed 01 Oct 2018
21. Qureshi M, Virkler M, Bernhardt K et al (2003) highway rail highway–rail grade crossing project selection. A technical report prepared for Missouri Department of Transportation. <https://library.modot.mo.gov/RDT/reports/Ri01010/RDT03017.pdf>. Accessed 24 Sept 2018
22. Dulebenets MA (2012) Highway–rail grade crossing identification and prioritizing model development. Master thesis, University of Memphis
23. U.S. DOT (2007) Rail–highway grade crossing handbook, second edition 2007. <https://www.fra.dot.gov/Elib/Document/1464>. Accessed 24 Sept 2018
24. Weissmann AJ, Weissmann J, Kunissety JL et al (2013) Integrated prioritization method for active and passive highway–rail crossings. A technical report prepared for Texas Department of Transportation. <https://static.tti.tamu.edu/tti.tamu.edu/documents/0-6642-1.pdf>. Accessed 01 Oct 2018



25. Iowa DOT (2006) Use of a benefit-cost ratio to prioritize projects for funding. [https://www.iowadot.gov/iowarail/assistance/130/130SelectionProcess\\_final.pdf](https://www.iowadot.gov/iowarail/assistance/130/130SelectionProcess_final.pdf). Accessed 24 Sept 2018
26. Hans Z, Albrecht C, Johnson P et al (2015) Development of railroad highway grade highway–rail grade crossing consolidation rating formula. A technical report prepared for Iowa Department of Transportation. [https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1061&context=intrans\\_techtransfer](https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1061&context=intrans_techtransfer). Accessed 24 Sept 2018
27. Ryan C, Mielke A (2017) Development of revised grade highway–rail grade crossing hazard index model. A technical report prepared for Nevada Department of Transportation. <https://www.nevadadot.com/home/showdocument?id=9133>. Accessed 24 Sept 2018
28. Sperry B, Naik B, Warner J (2017) Evaluation of grade crossing hazard ranking models. In: Ohio transportation engineering conference. [http://www.dot.state.oh.us/engineering/OTEC/2017Presentations/75/Sperry\\_75.pdf](http://www.dot.state.oh.us/engineering/OTEC/2017Presentations/75/Sperry_75.pdf). Accessed 30 Sept 2018
29. Hao W, Daniel J (2014) Motor vehicle driver injury severity study under various traffic control at highway–rail grade crossings in the United States. *J Saf Res* 51:41–48
30. Hao W, Kanga C, Yang X et al (2016) Driver injury severity study for truck involved accidents at highway–rail grade crossings in the United States. *Transp Res Part F Traffic Psychol Behav* 43:379–386
31. Hao W, Moghimi B, Yang X et al (2017) Effects of foggy conditions on driver injury levels in US highway–rail grade crossing accidents. *Case Stud Transp Policy* 5(4):627–633
32. Khan I, Lee E, Khan M (2018) Developing a highway rail grade crossing accident probability prediction model: a North Dakota case study. *Safety* 4(2):22
33. Saccomanno F, Fu L, Miranda-Moreno L (2004) Risk-based model for identifying highway–rail grade crossing blackspots. *Transp Res Rec* 1862:127–135
34. Saccomanno F, Park P, Fu L (2007) Estimating countermeasure effects for reducing collisions at highway–railway grade crossings. *Accid Anal Prev* 39(2):406–416
35. Evans AW (2011) Fatal accidents at railway level crossings in Great Britain 1946–2009. *Accid Anal Prev* 43(5):1837–1845
36. Dent M, Marinov M (2019) Introducing automated obstacle detection to British level crossings. In: Fraszczyk A, Marinov M (eds) *Sustainable Rail transport*. Springer, Cham
37. Borsos A, Gabor M, Koren C (2016) Safety ranking of railway crossings in Hungary. *Transp Res Procedia* 14:2111–2120
38. Laapotti S (2016) Comparison of fatal motor vehicle accidents at passive and active railway level crossings in Finland. *IATSS Res* 40(1):1–6
39. Liang C, Ghazel M, Cazier O et al (2017) Risk analysis on level crossings using a causal Bayesian network based approach. *Transp Res Procedia* 25:2167–2181
40. Liang C, Ghazel M, Cazier O et al (2017) A new insight on the risky behavior of motorists at railway level crossings: an observational field study. *Accid Anal Prev* 108:181–188
41. Larue GS, Wullems C (2015) Human factors evaluation of a novel Australian approach for activating railway level crossings. *Procedia Manuf* 3:3293–3300
42. Beanland V, Salmon PM, Filtress AJ et al (2017) To stop or not to stop: contrasting compliant and non-compliant driver behaviour at rural rail level crossings. *Accid Anal Prev* 108:209–219
43. Larue GS, Naweed A, Rodwell D (2018) The road user, the pedestrian, and me: investigating the interactions, errors and escalating risks of users of fully protected level crossings. *Saf Sci* 110(Part B):80–88
44. U.S. DOT (2016) Federal-aid policy guide. <https://www.fhwa.dot.gov/legisregs/directives/cfr23toc.htm>. Accessed 24 Sept 2018
45. Pasha J, Dulebenets MA, Abioye OF et al (2020) A comprehensive assessment of the existing accident and hazard prediction models for the highway–rail grade crossings in the State of Florida. *Sustainability* 12(10):4291
46. FRA (2010) Accident prediction and resource allocation procedure normalizing constants 2010. <https://www.fra.dot.gov/Elilib/Document/1488>. Accessed 02 Oct 2018