U.S. Department of Transportation

Federal Railroad Administration

A New Model for Highway-Rail Grade Crossing Accident Prediction and Severity

Office of Research, Development, and Technology
Washington, DC 20590

Photo – JJ11 / Shutterstock.com (non-commercial use)

DOT/FRA/ORD-20/40

Final Report
October 2020

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2020	Final Report 10/1/2019 – 7/31/2020

4. TITLE AND SUBTITLE

New Model for Highway-Rail Grade Crossing Accident Prediction and Severity

5. FUNDING NUMBERS

6. AUTHOR(S)

Daniel Brod 0000-0002-1213-0908 -, David Gillen 0000-0001-7047-2669

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Decisiontek, LLC
11821 Parklawn Drive, Ste. 260
Rockville, MD 20852

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research, Development, and Technology
Washington, DC 20590

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DOT/FRA/ORD-20/40

11. SUPPLEMENTARY NOTES

Project Manager: Karen McClure

12a. DISTRIBUTION/AVAILABILITY STATEMENT

This document is available to the public through the FRA website.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The U.S. Department of Transportation Accident Prediction and Severity (APS) model has been used by Federal, State, and local authorities to assess accident risk at highway-rail grade crossings since the late 1980s. The Federal Railroad Administration funded research for the development of a new model that employs current consensus analysis methods and recent data trends. The new model also seeks to address several limitations of the current APS model and to provide a more robust tool for analysts.

This report presents the stages of the new model development, the statistical estimation of the new model, and validations comparing the performance of the new model with the APS. The research shows that the new model described here out-performs the APS by multiple measures.

The new model will support grade crossing management by enabling: more accurate risk ranking of grade crossings, more rational allocation of resources for public safety improvements at grade crossings, and the ability to assess the statistical significance of variances in the measured risk at grade crossings.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Highway-rail grade crossing; risk assessment; accident rates; accident severity; accident types; high-speed rail; highway safety; highway travel; highways; improvements; highway vehicles; railroad crashes; railroad safety; railroad trains	61
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

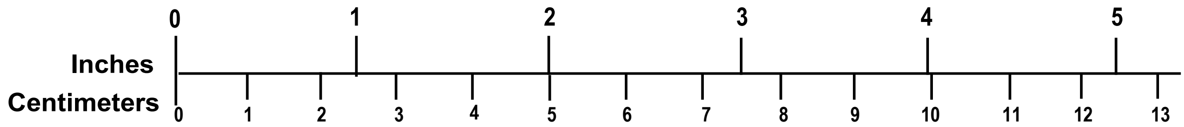
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Acknowledgements

The authors appreciate the insightful review comments by Federal Railroad Administration (FRA) staff Karen McClure and Francesco Bedini. We also acknowledge the ongoing support of Sam Al-Ibrahim of FRA's Office of Train Control and Communication.

Contents

Executive Summary 1

1. Introduction 3
   1.1 Background 3
   1.2 Objectives 7
   1.3 Overall Approach 7
   1.4 Scope 9
   1.5 Organization of the Report 9
2. Preliminary Data Review 11
   2.1 Risk by Warning Device Types 11
3. Data Selection and Analysis 19
   3.1 Data Sources 19
   3.2 Data Selection 19
   3.3 Candidate Variables 21
4. Accident Prediction Model 25
   4.1 The Accident Prediction Model 25
   4.2 ZINB Regression 26
   4.3 Predicting Accidents from the Regression Outputs 29
   4.4 Empirical Bayes Prediction Adjustment 29
5. Accident Severity Model 33
   5.1 Description of the Data 33
   5.2 The Accident Severity Model 33
   5.3 Multinomial Logistic Regression 35
   5.4 Accident Severity Forecast Formulas 37
6. Validation 39
   6.1 Accident Prediction – Cumulative Risk 39
   6.2 Accident Prediction – Risk at Crossings by Accident Count Groups 42
   6.3 Accident Prediction – Accident Risk for Groups of High-Risk Crossings 43
   6.4 Accident Severity – Model Comparisons 44
7. Conclusion 47
8. References 49

Appendix A. Interpreting Regression Outputs 51

Appendix B. Application of the New Model 53

Figures

Figure 2-1. Accidents by Warning Device Type 13
Figure 2-2. Risk per Exposure (Accident Rate) by Warning Device Type 15
Figure 2-3. Relative Risk Levels by Warning Device 16
Figure 3-1. Boxplot of Normalized Crossing Accidents by Grade Crossing Surface Type 22
Figure 3-2. Boxplot of Normalized Crossing Accidents by Grade Crossing Angle 23
Figure 3-3. Boxplot of Normalized Crossing Accidents by Maximum Timetable Speed Deciles 24
Figure 3-4. Boxplot of Normalized Crossing Accidents by Percent Truck Deciles 24
Figure 4-1. ZINB Predicted Accidents by Warning Device Type 29
Figure 4-2. ZINB+EB Predicted Accidents by Warning Device Type 31
Figure 5-1. Severity Predictions for 50 Crossings with the New Model 38
Figure 6-1. Model Comparison (2014–2018, all crossings in sample) 40
Figure 6-2. Model Comparison (2015–2019, all crossings in sample) 40
Figure 6-3. Model Comparison (2014–2018, crossings in sample with non-zero accidents) 41
Figure 6-4. Model Comparison (2015–2019, crossings in sample with non-zero accidents) 41
Figure 6-5. Model Comparison, Accident Counts, and Predictions (2014–2018) 42
Figure 6-6. Model Comparison, Accident Counts, and Predictions (2015–2019) 43
Figure 6-7. Comparison of Predictions for Riskiest Crossings 44
Figure 6-8 Distribution of Predicted Accident Severities with the New Model 45
Figure 6-9. Distribution of Predicted Accident Severities with APS 46

Tables

Table 1-1 Comparison APS vs. New Model .... **Error! Bookmark not defined.**
Table 1-1. Estimates Required for Different Types of Analysis Focus 8
Table 2-1. Warning Device Type Codes and Descriptions 12
Table 3-1. Summary of Data Selection 21
Table 3-2 Candidate Variables for Inclusion in the New Model 21
Table 4-1. ZINB Regression Output 27
Table 5-1. Accident Severity Multinomial Logistic Regression Output 36
Table 6-1. Predicted Severity (Percent of Total) by the New Model and APS 44
Table 6-2. Summary of Severity Category Skewness by Model 46

Executive Summary

This report presents research on a new model as an alternative to the U.S. Department of Transportation grade crossing Accident Prediction and Severity (APS) model, which dates back to 1986. This report follows the steps in developing the new model, presents the modeling results, and validates the new model in comparison to the APS.

In the nomenclature of AASHTO's Highway Safety Manual, the new model is a safety performance function (SPF). SPFs generate metrics (e.g., predicted accidents by severity type) indicating safety (or risk, insofar as more safety means less risk) and have been applied to a range of highway facilities. The SPF approach is applicable to grade crossings, individually and to aggregated collections (i.e., populations)[1], as well.

The new model derives from a policy perspective on grade crossing safety, a review of the data, statistical analysis, and validation. The authors conclude that the new model outperformed the APS, and its adoption would result in more accurate risk ranking of grade crossings, more rational allocation of resources for public safety improvements at grade crossings, and the ability to assess the statistical significance of variances in the measured risk at grade crossings.

Key Conclusions

The preliminary data review indicates a new model could replace the APS based on the key drivers of exposure and grade crossing warning device type (i.e., the data show that risk increases with exposure, and decreases with a more protective warning device type).

There is justification for a single model with warning device type category as a variable rather than separate models for each of the three warning device type categories.

In the U.S. there are 105,377 grade crossings that are public, not closed, not grade separated, and that have non-missing, non-erroneous values for exposure and warning device type. From 2014–2018, there were 8,467 accidents at these grade crossings.

An aggregate analysis of these grade crossings shows that relative to a passive crossing, an average lights crossing had 73 percent less risk per exposure than a passive crossing. An average gated crossing had 63 percent less risk per exposure than a lights crossing.

The findings of the above analysis indicate a functional form with exposure, warning device type, and other grade crossing characteristics.

Model estimation using the zero-inflated negative binomial (ZINB) regression yielded parameters of the expected sign and magnitude, and had strong statistical significance.

The empirical Bayes (EB) method accounted for accident history while correcting for "regression to the mean" bias. Adjusted results with EB produced predictions that more closely track the actual counts than did the APS with its (non-EB) adjustment process for accident history.

[1] The "population" of grade crossings refers to all public grade crossings in the U.S. that are not closed or grade separated. The analysis sample is a large subset (over 100,000) of all grade crossings.

The new model severity component determines the probabilities that an accident will be of one of three severity types: fatal, injury, or property damage only. The severity component of the new model was derived using multinomial logistic regression on the accidents in a 6-year period, 2014–2019. In this period there were 11,131 accidents at public crossings. Of these, there were 9,870 at grade crossings with non-missing, non-erroneous data.

These 9,870 accidents were included in the severity model estimation. The MNL regression shows that the best results were obtained with explanatory variables: rural or urban, maximum time table speed, number of daily trains, and whether a crossing had a lights warning device.

Validations indicate the new model outperformed the APS. One of the validations looks at cumulative risk at crossings, with crossings ordered from greatest to least risk (i.e., accident count). The riskiest crossings in the data sample include 7,822 accidents at 6,409 crossings in 2014-2018. Applying each model (new and APS) to the data, the new model predicted 4,853.3 accidents (62.0 percent of the actual count) whereas the APS predicted 2360.2 accident (30.2 percent of the actual count).

1. Introduction

1.1 Background

1.1.1 About the APS Model

The U.S. Department of Transportation (DOT) Accident Prediction and Severity (APS)[2] model has been used to assess accident risk at highway-rail grade crossings by all levels of government since the late 1980s. The assessments of accident risk at grade crossings are foundational information that guide the management of grade crossings, the identification of high-risk crossings (“hotspots”), and the allocation of resources for improving grade crossing safety.

The APS model was developed in 1986 based on grade crossing and accident data from the preceding 20 years.

Additional modeling efforts intended to support and supplement the APS were conducted more recently by the Volpe National Transportation System Center (Volpe) and FRA. Volpe developed a High-Speed Rail (HSR) Accident Severity Model in 2000[3] to predict accidents and their severity by types of traffic on the highway and railroad. In 2005 the FRA published the final Train Horn Rule (49 CFR 222), which specified “supplementary safety measures” and their impacts on risk reduction. Such measures include: four-quad gates, median barriers, mountable curbs, and new technologies like photo enforcement.

Among its enhancements for assessing grade crossing risk, FRA’s GradeDec.Net online tool gives users access to the HSR Accident Severity Model, and complements the APS model with the supplementary safety measure impacts from the Train Horn Rule.

While these improvements are notable, a new replacement model for the APS is still required to ensure that U.S. DOT, State Departments of Transportation, and local governments efficiently utilize resources for reducing risk at grade crossings.

1.1.2 Grade Crossing Accident Trends and the APS

Grade crossing accidents declined sharply in the 25 years following APS development (from about 3,000 per year to about 2,000 per year). This reduction was due to a number of factors, indicating the relationship between grade crossing characteristics and accidents has likely shifted.

FRA periodically updates the APS normalizing constants[4] so that the national aggregate number of predicted accidents equals the actual number of accidents at the end of the most recent calendar year. While the normalizing constants are applied uniformly within each warning

[2] Farr (1987) describes the APS.

[3] See Mironer, et al. (2000) at https://rosap.ntl.bts.gov/view/dot/8433.

[4] See Farr (1987), 3-7.

device type group, they do not account for the many factors influencing accident risk that have changed in recent years, namely: rail and highway environments, technology, traffic trends, etc.

On the rail side, freight trains are longer, which causes longer block times at crossings. The expansion of intermodal traffic and the growth of intermodal facilities have led to choke points on highways in the vicinity of some major intermodal facilities. Longer waits at crossings contribute to "incentivizing" risky behavior (e.g., driving around lowered gates) by some highway users. In recent years there has been an uptick in grade crossing accidents.

One would also expect changes in highway user behavior to impact safety at crossings. Trends toward larger vehicles (e.g., SUVs and light trucks replacing smaller cars) result in slower queue dispersal at crossings. Changes in traffic mix, increases in number of delivery vehicles, and the rise of ride-sharing – would all contribute to changes in crossing safety and its prediction based on characteristics of grade crossings and traffic volumes by mode.

Moreover, since 1986, new technologies and traffic management measures have been deployed at many crossings, including: constant warning time (CWT) devices, signal pre-emption, and queue cutters

1.1.3 APS Limitations

State and local government agencies have alerted the FRA Office of Research and Development that the APS produces very similar results for a majority of crossings within their jurisdictions, making it difficult to identify the highest-risk highway-rail grade crossings. Limited variance among APS-generated assessments is attributed to the predominance of crossings with no accidents in the preceding 5 years, and similar-site specific characteristics (like traffic counts and warning devices). New consensus methods of analysis (see the Accident Prediction Model section) directly address these issues.

The APS includes three separate models for accident prediction – one for each of the three major grade crossing warning device type categories: passive (signage), flashing lights, and gates. There is no clear rationale for splitting accident prediction into three separate models, as opposed to treating the warning device type as a grade crossing characteristic in a single model for all crossings.

Moreover, the separate models can generate inconsistent outcomes. For example, for some combinations of grade crossing characteristics, the APS calculates higher risk for crossings with the same characteristics except for a more protective warning device. It is easy to see how an analysis of grade crossing risk in a corridor or region could yield results with measures of relative risk between similar crossings with different warning device types that are highly suspect.

Similarly, if seeking to estimate the effect of a warning device upgrade (say, from lights to gates), one could not use the models, segregated by device type category, to estimate risk reduction. The APS resource allocation procedure is to work around this issue by applying a crash modification factor (CMF).[5] A CMF reduces the risk of the unimproved grade crossing by a fixed percentage. The workaround uses the CMF-reduced risk result in place of the APS result for the assessed risk of the improved crossing. The CMF method, while accepted practice, has

---

[5] Farr (1987), p. 11, calls these "effectiveness factors." The term crash modification factor was adopted later.

been critiqued in the safety research literature.[6] Regardless, the model should enable recalculation of the risk at the crossing corresponding to a warning device upgrade without relying on external methods.

Another limitation of the APS model is that it provides no method to determine if risk measures at different crossings differ with statistical significance[7] (e.g., two crossings with predicted annual accidents of, say, 0.21 and 0.23, respectively). If the difference in measured risk at two crossings is not statistically significant, there is no evidentiary basis for treating these crossings differently (e.g., applying an improvement to one of the crossings and not the other). The APS is essentially a scoring model where a statistical model is needed (see the example in Appendix B, Application of the New Model).

1.1.4 Purpose of a New Model

The overarching purpose of a new grade crossing safety model, an alternative to the APS, is to effect evidence-based safety management of grade crossings. The new grade crossing safety model should enable users to:

1. Estimate safety and risk at grade crossings.
2. Estimate safety gains due to prospective improvements to crossings, and support the estimation of benefits from these gains.
3. Screen for high risk crossings and develop strategies and programs for safety improvements.
4. Account for statistical significance of differences in measured risk at crossings.
5. Estimate changes in safety at crossings due to changes in some variable value (e.g., growth of AADT over time).

1.1.5 Policy Perspective of Grade Crossing Safety

Grade crossings are "safety hotspots." Fatalities in grade crossing accidents numbered 260[8] in 2018. While this may seem small in comparison to total U.S. highway fatalities (36,560[9] in 2018), fatalities and accidents at grade crossings are highly significant when considering the amount of highway travel that actually traverses grade crossings.

Transportation agencies at all levels recognize that grade crossings are a significant source of risk and have been singled-out for special programs and safety countermeasures over the years. Accident risk at grade crossings is eliminated by closure or grade separation (closure, however, could possibly re-direct the risk from the closed crossing to other grade crossings). Additional measures like warning device upgrades, supplementary safety measures, and other engineering solutions have been shown to significantly reduce risk at grade crossings.

---

[6] See Hauer (2015), 186-188.

[7] This is similar to asking whether the risk measures of the two crossings are within the "margin of error."

[8] https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/Query/AccidentByRegionStateCounty.aspx

[9] https://cdan.nhtsa.gov/tsftables/National%20Statistics.pdf

There is a definitional relationship between risk and exposure. Exposure is a measure of opportunities for accidents to occur. The exposure[10] metric for grade crossing usage is based on coincident arrivals of trains and highway vehicles at a crossing. It is not surprising to find that more heavily trafficked grade crossings, in general, have more protection from warning devices. In this report, Decisiontek analysis examines the relationship between accidents, exposure, and the principal warning device type categories.[11]

The current U.S. DOT APS model has three accident prediction models, one for each warning device type category. For some ranges of input variables, APS calculates higher risk than with a more protective warning device type. (For example, with exposure of 1,000 and maximum timetable speed of 79 mph, the APS predicts more accidents at a gated crossing than at a lights crossing.) This should give pause when considering APS predictions in a region or corridor. If two crossings have similar data with the exception of the warning device type, do we have confidence in the relative measure of their predicted accidents? Moreover, would proposed improvements for the corridor or region be allocated to their most effective use? The new model, based on modern techniques, replaces the three APS models with a single prediction model that incorporates warning device type category as a variable. Its predictions consistently preserve relative magnitudes of risk with different warning devices.

Moreover, the APS resource allocation procedure relies on "effectiveness values"[12] to account for risk reduction with a warning device upgrade (in recent years, these have been renamed "crash modification factors"). The road safety literature indicates that such mixed methods can result in methodological inconsistencies.[13]

The assessment of grade crossing risk and the planning and budgeting for improvements are the sole responsibility of State and local authorities.[14] The public authority assessing grade crossing risk relies on a model like the APS[15] and bases management decisions for improvements, accordingly. The quality of those decisions will rely to a great extent on the quality of the risk assessment.

The new model developed here as an alternative to APS seeks to address the issue of risk assessment quality by:

- Relying upon current data, appropriate data analysis, and statistical methods
- Examining the relationship between exposure, warning device type, and other key grade crossing characteristics

---

[10] Exposure, or exposure to risk, is defined for grade crossings as average annual daily trains times average annual daily highway vehicles at a crossing. This definition is imperfect because accident risk should consider the correlation of vehicle arrivals by mode, accounting for both seasonality and diurnal distributions of traffic.

[11] The APS is defined in Farr (1987). Warning device type categories are: passive, lights, and gates.

[12] Farr (1987) p. 11, Table 3 "Effectiveness Values for Crossing Warning Devices."

[13] See, for example Hauer (2015), Appendix L.

[14] Upon request, the owning railroad grants the public authority easement to build and maintain the road that traverses its track. The railroad bears full responsibility for maintaining warning devices and any equipment within the grade crossing right-of-way.

[15] FRA maintains the APS and provides a web-based version at https://safetydata.fra.dot.gov/webaps/.

- Properly accounting for accident history
- Presenting a fully transparent model that allows for: single crossing estimates, estimates of risk for groups of crossings, and determining whether differences in grade crossing risk warrant similar or different treatment based on statistical measures

1.2 Objectives

The objectives of the research are as follows:

- Develop a new model to serve as an alternative to the current U.S. DOT APS.
- Document the full development process of the model.
- Demonstrate that the model satisfies statistical criteria and is practical for practitioner use.
- Validate the new model by comparing its performance against the APS and actual accident data.

1.3 Overall Approach

1.3.1 About Safety Performance Functions

Since the late 1990s, there has been substantial progress in consensus methods for developing safety prediction models. These new approaches are presented in AASHTO's Highway Safety Manual.[16] In the current mode of thinking, the APS is a type of "safety performance function" (SPF), which yields a metric indicating the safety of a grade crossing. That metric can be either the annual expected number of accidents at a crossing or expected accidents by severity type (e.g., fatal, injury, property damage only – the APS accident severity types).

The SPF is derived in a multi-stage process. The key sources of data for this process are: 1) a set of traits that characterize the facilities under consideration and 2) the 5-year accident history at the grade crossings. The database of traits is the U.S. DOT Grade Crossing Inventory System (GCIS). The database of U.S. DOT Form 57 (a form must be submitted for each highway-rail crossing accident) captures the grade crossing accident history.

The SPF development involves: First, screen the data in the inventory to eliminate irrelevant or erroneous data. Second, discover via analysis the functional forms that best describe the data, and offer hints regarding possible relationships between accidents and traits. Third, derive the safety model from a suitable statistical estimation procedure. Fourth, adjust the number of predicted accidents at each crossing to account for the accident history using empirical Bayes (EB) estimators, which derive from another statistical procedure.

This research covers the development of a new model, namely: the derivation of the SPF, its validation, and the process for estimating safety risk at grade crossings as an alternative to the APS.

---

[16] AASHTO (2010).

1.3.2 Information the SPF Provides

The model, or SPF, provides estimates of four elements for a given set, or population, of grade crossings:

1. $E[\mu_i]$, the expected or predicted number of accidents at crossing i
2. $\sigma[\mu_i]$, the standard deviation of the predicted number of accidents at crossing i
3. $E\{\mu\}$, the mean of all the μs in a population (all crossings or a subset of crossings)
4. $\sigma\{\mu\}$, the standard deviation of all the μs in a population

The following table shows situations for which the above estimates are needed:

Table 1-1. Estimates Required for Different Types of Analysis Focus

Analysis Focus	
Average safety E{μ} for subsets of grade crossings	**Safety (μ, σ) of specific grade crossings**
What is normal for grade crossings with given traits?	Is the crossing "unsafe" or has unusually high risk?
How do the E{μ} vary across subsets of crossings (e.g., by states or region, by device type)?	Can we rank a collection of crossings and divide into high- and low-risk groupings?
What would be the aggregate effect of making an improvement over a population of crossings (e.g., eliminate humped crossings)?	What might be the safety effect and benefit of applying some improvement to a crossing?
Need E{μ} and σ{μ} to answer the questions	Need E[μ] and σ[μ] to answer the questions

Source: based on Hauer (2015).

The estimate of the standard deviation of the safety metric is needed in the case of specific crossings in order to determine whether:

- Predicted accidents are different from zero with statistical significance.[17]
- Safety measures of two crossings are statistically different from one another (i.e., if crossings A and B, say, have predicted accidents of 0.21 and 0.23, respectively, should they be treated differently or with different priority on the basis of the evidentiary data).

To achieve an SPF, data about grade crossing characteristics, or traits, need to be cast as statistical models that explain the accident counts at crossings. In developing a safety model for crossings, there are two clues that the model needs to exploit:

- The first clue is the characteristics (or traits) of the grade crossing. These traits contain information regarding the common features of grade crossings that contribute to accidents.

[17] "Statistical significance" means that a relationship between two or more variables is caused by something besides chance. If the ratio of a crossing's mean predicted accidents to its standard deviation exceeds a threshold value (e.g., 1.65) then the predicted accidents is said to be "statistically significant at the (e.g.) 90% level." This is equivalent to saying that there is a 10 percent probability of a Type I error (falsely rejecting the null hypothesis).

- The second clue available for developing a safety model is the accident history. Accident history captures the unique qualities of each crossing contributing to safety and risk.

As a general approach, the safety model will account for both clues by first predicting accidents based on characteristics, and then adjust the outcome to account for accident history.

The principles outlined in this section guided the development of the new model for grade crossing accident prediction and severity.

1.4 Scope

The analysis of the accident and GCIS data focused, and the development of the new model, focused on methods described in AASHTO's "Highway Safety Manual."[18] The approach the project researchers followed sought to:

- Make best use of their understanding of historical trends, the policy environment, and practice in using the APS.
- Maximize the number of grade crossings included in the regression analysis.

Researchers did not conduct an exhaustive search of alternative approaches, such as: artificial intelligence (AI) methods, like "k nearest neighbors" (KNN); methods for "slicing and dicing" the data into smaller subsets; non-multiplicative (i.e., linear in logs) functional forms, etc. The research team believes that alternative approaches may have merits, but also drawbacks in comparison with the chosen approach.

The focus of the research was on developing the model. The team recognizes that additional work is needed to further operationalize the model and provide guidance for use of the new model by practitioners.[19]

1.5 Organization of the Report

Section 2 is a preliminary data review. The section discusses well-established relationships (e.g., exposure drives risk, upgrading the warning device type at a crossing reduces risk). It concludes with a generic functional form based on the principal drivers of risk (exposure and warning device type) and accommodates additional variables as warranted by data analysis and the estimation process.

Section 3 describes the data selection and data analysis.

Section 4, the Accident Prediction Model, presents the functional form of the new model accident prediction, its estimation using the zero-inflated negative binomial (ZINB) regression method, and the application of the EB method. The section concludes with the new model formulas for accident prediction.

[18] [18] AASHTO (2010).

[19] For example, guidance should provide rules for treating missing data or replacing data from the GCIS with more current or more relevant estimates.

Section 5, the Accident Severity Model, presents the accident severity component of the new model. It describes the multinomial logistic (MNL) regression method used to develop the model.

Section 6, Validation, presents validations of the accident prediction and severity prediction of the new model.

Section 7 is the Conclusion.

2. Preliminary Data Review

In this section, the research team identifies known relationships or well-supported theories relating accident risk at grade crossings to grade crossing traits.

The team explored whether a single model could internalize warning device types and thus avoid having separate models for each class of device. A unified model would ensure that a device upgrade will be accompanied by accurate risk reduction measurements of accidents at grade crossings. This would eliminate the need for employing a "crash modification factor" (CMF)[20] approach to estimate the effect of a device upgrade.

It is intuitively clear, and supported by research[21], that upgrading a warning device type to one that provides a higher level of protection reduces the accident risk at a crossing (given that all other factors remain the same). (That said, it does not follow that a device upgrade is cost-beneficial or even a cost-effective way to improve safety at a crossing.)

There are three warning device type categories: passive, lights, and gates. Within each category, there are several warning device types with somewhat differing risk characteristics than the main category. These will be discussed below.

It is also well understood that risk increases with exposure (although not at a uniform rate for every level of exposure). As one would expect, for a given crossing the greater the exposure and risk, the more likely it is that a local authority will upgrade the warning device. Consequently, nearly all very low-exposure crossings have passive devices and nearly all very high-exposure crossings have gates. The researchers expected to observe a high correlation between device type and exposure at crossings.

This section examines the relationship between accidents, exposure, and device types and concludes with a general functional form for the accident prediction model.

2.1 Risk by Warning Device Types

Table 2-1 shows the warning device codes by super-category (passive, lights, gates) and their meaning in GCIS.

[20] The CMF approach, often based on before-and-after crash studies, provides a factor associated with risk reduction for a particular safety countermeasure. For example, a CMF of 0.12 means that predicted accidents after applying the safety countermeasure will equal predicted accidents before such application times one minus the CMF, i.e., $A_{after}=A_{before}*(1\text{-}CMF)$.

[21] Elvik, R. and Vaa, T.(2004).

Table 2-1. Warning Device Type Codes and Descriptions

Code	Description of Warning Device Type
PASSIVE	
1	No sign or signal
2	Other signs or signals
3	Stop signs
4	Crossbucks
LIGHTS	
5	Non-train-activated special protection
6	Highway traffic signals, wigwags or bells
7	Flashing lights
GATES	
8	Gates
9	4-quadrant gates

Figure 2-1 shows the filtered crossings in the inventory grouped by device type category. The bars indicate the number of crossings with the specified device type having the number of accidents in the period shown on the x-axis. Note that the y-axis uses a log scale.

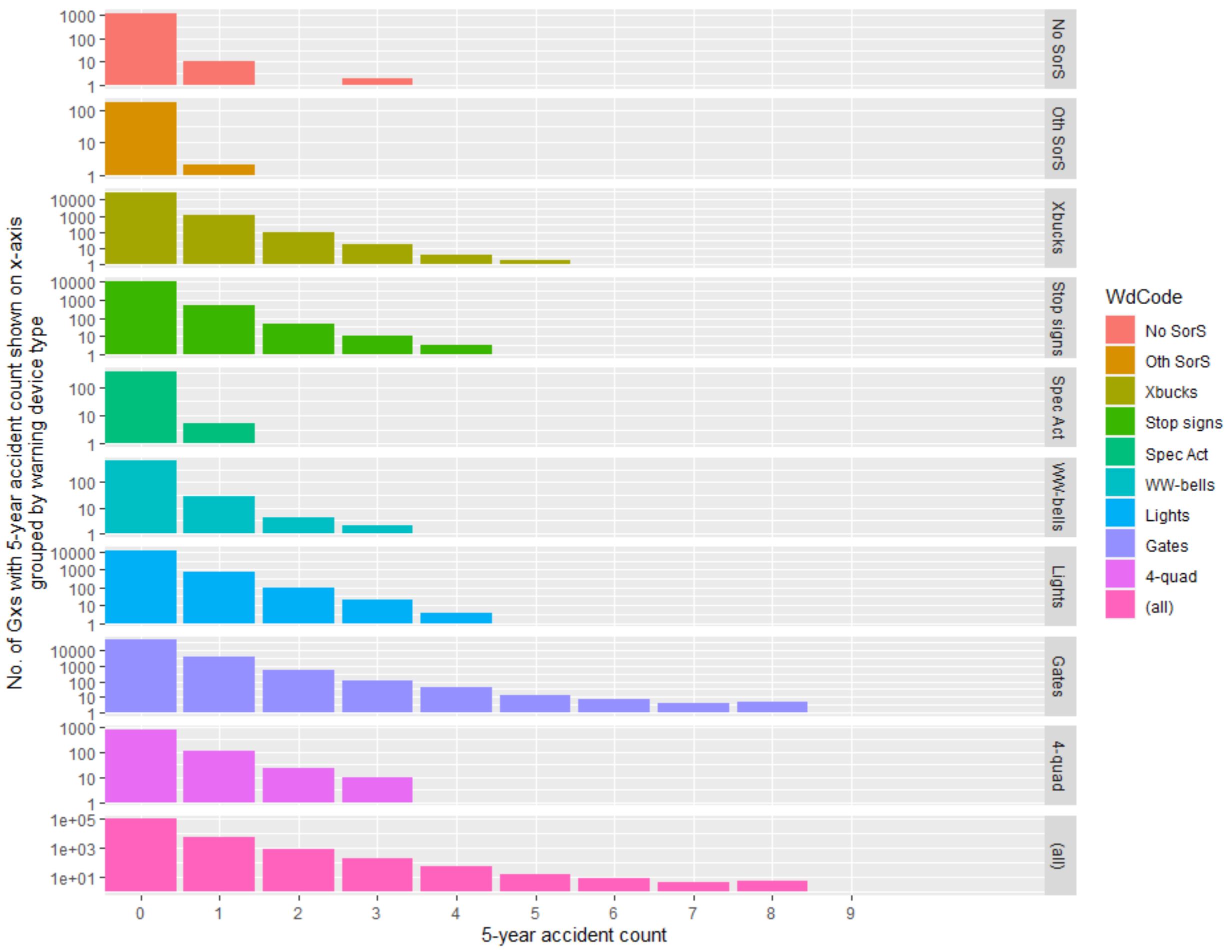


Figure 2-1. Accidents by Warning Device Type

2.1.1 Aggregate Risk Adjusted for Exposure by Warning Device Type

To support an accident prediction model with exposure and warning device type as core variables, the team examined aggregate risk at crossings by warning device type and accident rates (i.e., accident count divided by exposure).

Accident per exposure is the most common way to express accident rates on a facility.[22] Note that the accidents are for 5 years. The exposure data in GCIS[23] are for a typical day. Exposure for the 5-year period is given by:

---

[22] For example, "Highway Statistics 2018, Federal Highway Administration" gives fatality rates in terms of "fatalities per 100 million VMT (vehicle-miles traveled)." VMT is the measure of exposure for general highway use.

[23] As a caveat, note that the GCIS data are reported by State and local agencies with varying data quality. Moreover, some data fields are not maintained as vigorously as others. For example, data for warning device type are, for the most part, current and accurate. Data for the railroad and highway environments at crossings (e.g., AADT, train traffic) tend to be less current and may be out-of-date.

Equation 1. Exposure in the Analysis Period (2014–2018)

$$xp = (aadt \cdot dt)_{daily} \cdot 300 \cdot 5$$

where:

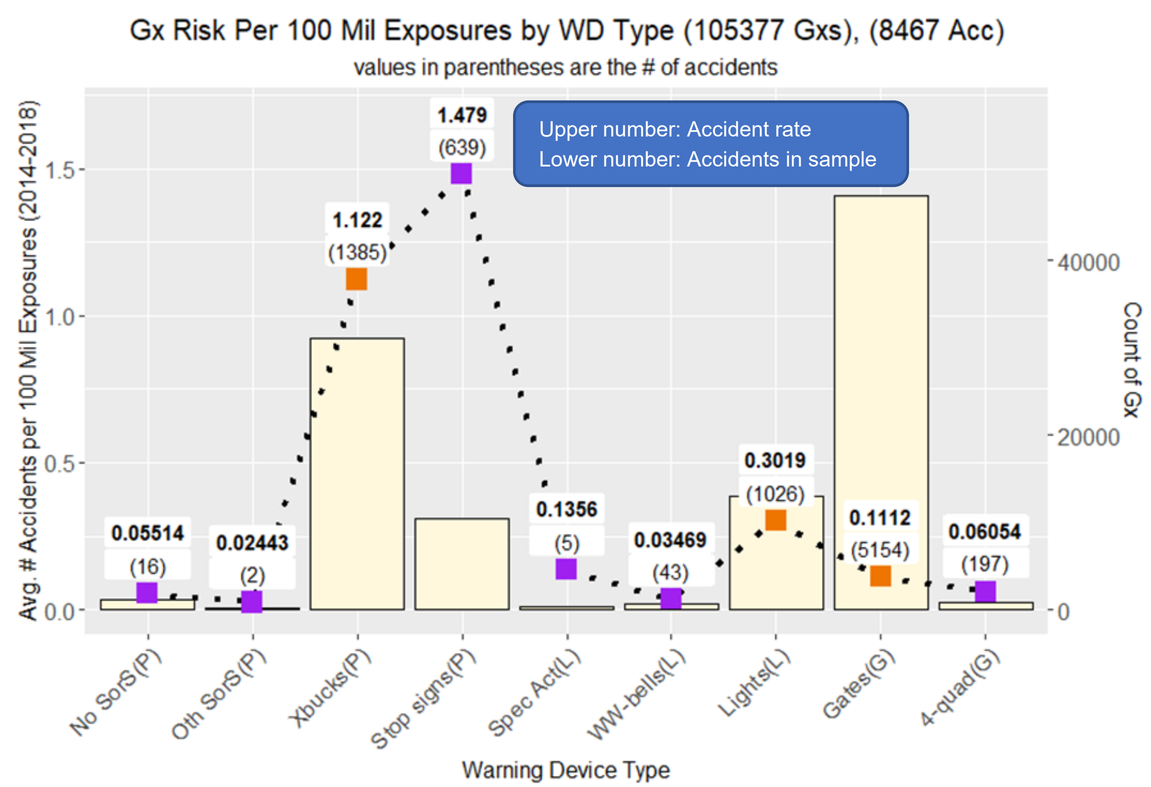
xp	Exposure in 5-year period
aadt	Average annual daily traffic
dt	Daily trains at the crossing
300	Number of annual traffic days
5	Number of years

Figure 2-2 shows the crossing risk divided by exposure for each device type category. The data points (colored purple and orange) show the risk per exposure at each crossing grouped by warning device type. The risk values are shown at the data points in bold and by the left y-axis. The bars are the number of crossings in each group and their values are represented by the right y-axis. Note also the number below the risk value, which is the count of accidents in the period for each grouping of crossings.

Focusing for now on the orange data points, these represent the largest groupings in each of the three super-categories: passive, lights, and gates. A lights crossing has 73 percent less risk per exposure in comparison to the passive crossing. Compared to a lights crossing, the gated crossings have 63 percent less risk per exposure.

The orange points were singled-out because they represent: 1) the main grouping in the super-category, and 2) in each, there is a substantial number of crossings and accidents. The "Stop Signs" category is also sizable and its risk per exposure is not that different than the risk per exposure of the crossbucks grouping (1.122 vs 1.479; in other words, crossbucks are about 75 percent as risky per exposure as stop signs). Moreover, there are over 10,000 crossings in the "Stop Signs" category and initial inspection indicates that it will likely be advantageous to merge the two categories into the "passive" category.

The other warning device type categories within each super-category are somewhat small samples of crossings and accidents with widely different risk characteristics than the main grouping. The crossings with codes for these groupings (1, 2, 5, 6, 9) will be omitted from the analysis. (For accident prediction of these device types, we would use the super-category and then apply a CFM to scale the risk given the best available information).



NOTES TO FIGURE 2-2

The bars in the chart are the number of grade crossings (shown on the right y-axis) for each warning device type (shown on the x-axis). In the x-axis labels, the letters in parentheses indicate the principal warning device type category (P – passive, L – lights, G – gates).

The square markers represent the average number of accidents per exposure at crossings with the warning device type (shown on the left y-axis). Markers are colored orange for the warning device type with the largest number of grade crossings in the warning device type category.

Figure 2-2. Risk per Exposure (Accident Rate) by Warning Device Type

2.1.2 A Generic Functional Form for Accident Prediction

The following generic functional form follows from the above discussion.

Equation 2. Generic Functional Form for Accident Prediction

$$A_{predicted} = e^{[\beta_0+\beta_1 \cdot \log(xp)+ \beta_2 \cdot D2 + \beta_3 \cdot D3]} \cdot f(\boldsymbol{x})$$

where:

xp	Exposure (= daily trains * aadt)
x	Other variables (vector)
D2	1 if crossing warning device is lights, 0 otherwise
D3	1 if crossing warning device is gates, 0 otherwise
Note: If D2 = D3 = 0 then the warning device at the crossing is passive	

From an understanding of the impacts of exposure and warning device types on accident risk, the parameter estimates of coefficients from a statistical estimation process would yield the following:

$$0 > \hat{\beta}_2 > \hat{\beta}_3$$

that is, a crossing with lights warning device has less risk than a crossing with passive device, and a gates crossing has less risk than a lights crossing. (The "hat" diacritical indicates an estimated coefficient of the model.)

The following chart shows the relative risk of an example grade crossing for different warning device types and at different levels of exposure. Note that for very low levels of exposure all crossings have passive warning devices, and at very high levels of exposure grade crossings are gated. Grade crossings with lights fall in the middle range of exposure.

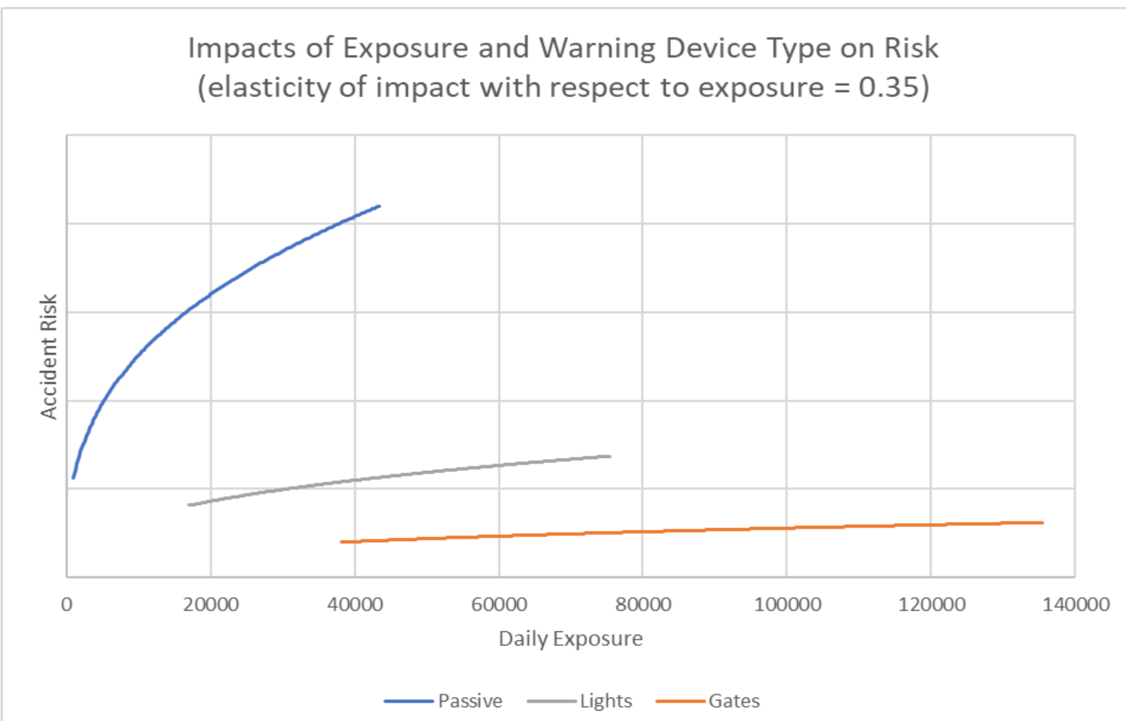


Figure 2-3. Relative Risk Levels by Warning Device[24]

[24] The elasticity of risk with respect to exposure (set to a value of 0.35) is drawn from the current APS and preliminary data analysis. Elasticity is the percent change in one variable (e.g., accident risk) when another variable (e.g., exposure) varies by 1 percent.

The following sections show how this general form, together with additional model variables, will combine in the new accident prediction model.

3. Data Selection and Analysis

The section describes the process of data selection for the development of the new model that will serve as an alternative to APS. The goal was to produce a model that defines an SPF for grade crossings. The first focus was on a model predicting accident occurrence, and later in this document address accident severity prediction given an accident.

Following data analysis and selection of traits for inclusion in the new model, additional filters may be applied to the data to account for missing/erroneous values for the new model traits. An additional consideration that accompanied the data analysis was to retain as many grade crossings in the dataset for model estimation as practical.

The research team sought variables that were likely to support a model. Since researchers proceeded from the assumption that key drivers are represented by exposure and warning device type, they further assumed that f(x) from Equation 2 in the previous section was linear in its variables (which were the explanatory variables the team sought to identify for inclusion in the model).

3.1 Data Sources

The two sources of data for the development of the new model are:

- Grade Crossing Inventory System (GCIS) data. The reference document for the data is "FRA Instructions for Electronic Submission of U.S. DOT Crossing Inventory Data, Grade Crossing Inventory System (GCIS), v2.9.0.0, Released: 7/2/2019." Grade crossing data updates are electronic submissions of Form FRA F 6180.71 by railroads, transit agencies, and States. GCIS uses Open Data (OData), a RESTful (REpresentational state transfer), for data downloads. OData downloads provide a single table that includes all five parts of the inventory – including header information. The data contain one row for each grade crossing in the inventory representing the most current data per the submitting agency's most recent submission.
- The FRA safety data website provides downloading accident data by year. The accident data source is Form 6180.57, which railroads submit to FRA following each grade crossing accident. The Form 6180.57 data download as a single table (in Excel or Access formats) with each accident represented as a single row in the table. For the analysis, researchers looked at accidents in the 5-year period 2014–2018.

We downloaded and inserted the data into SQL server database tables. The tables were merged into a single table with an additional column for total accidents in the period (2014–2018).

3.2 Data Selection

This section describes the process for filtering the data so as to include those crossings that are the focus of the analysis, while eliminating from analysis those crossings that are not of interest (e.g., closed or grade separated). Researchers also filtered out data that had missing or erroneous values for several key analysis variables. Table 3-1 summarizes the data filters along with the number of crossings, accidents, and number of crossings with accidents remaining after applying each filter. The team sought to keep the number of grade crossings in the selection as large as possible so that its practical application in prediction would not require an extensive set of rules to account for missing or erroneous data. For example, if a variable seemed promising for inclusion, yet only, say, 30 percent of grade crossings had data for the variable – researchers opted to exclude it).

3.2.1 Public Crossings Only

GCIS identifies public crossings as those having a value of 3 in the TypeXing field. For private crossings, the roadway is maintained by a private individual or entity. There is no legal obligation for the road maintainers at private crossings to submit data to GCIS. Each year, on average, 14 to 15 percent of accidents occur at private crossings. However, the data of crossing characteristics at private crossings are extremely sparse. Consequently, these have been excluded from the analysis.

3.2.2 At-Grade Crossings Only

Crossings that are grade separated pose no risk of collision between trains and highway vehicles, hence these crossings are excluded. The field PosXing with value set to 1 identifies a crossing at-grade.

3.2.3 Closed Crossings

GCIS identifies closed crossings when the ReasonID (reason for submitting a data update) field is set to value 16. Crossings with ReasonID = 16 have been eliminated from the analysis. Note that it may be the case that a closed crossing was subsequently updated for a different reason, in which case there would be no indicator in GCIS that the crossing was closed.

3.2.4 Missing or Erroneous Values for AADT

Without a value for average annual daily traffic (AADT), risk exposure at the crossing could not be evaluated (defined as AADT times the number of daily trains). Note that AADT, like other variables in GCIS, may be out-of-date.

3.2.5 Missing or Erroneous Values for Number of Daily Trains

As with AADT, crossings that have missing or erroneous data for total number of daily trains have been excluded.

3.2.6 Missing or Erroneous Values for Highway Lanes and Tracks

These two variables are the key descriptors of infrastructure at crossings and may be important predictors of accidents.

Table 3-1. Summary of Data Selection

Filter Criterion (with previous filters)	Number of Crossings Remaining after Filter	Total Number of Accidents 2014-2018 at Remaining Crossings	Of Remaining Crossings, Number with Accidents
None	429,463	10,675	8,814
Public only	266,304	9,147	7,538
At-grade only	220,289	9,110	7,503
Exclude closed	130,107	8,986	7,390
Exclude 0, missing, erroneous AADT	128,378	8,922	7,334
Exclude 0, missing, erroneous highway lanes	127,755	8,895	7,308
Exclude 0, missing, erroneous daily trains	105,383	8,467	6,944
Exclude 0, missing, erroneous total tracks	105,362	8,465	6,942

3.3 Candidate Variables

Variables in the GCIS that were considered candidates for explaining accidents are shown in the table below. Researchers eliminated from the list variables that are already accounted for in the exposure variable (i.e., trains and AADT) and those that are likely highly correlated with these variables. Warning devices were also excluded, as the team included these by default in the new model. The variables are divided into two groups: discrete and continuous.

The analysis assesses whether a variable is a likely candidate for inclusion in the model.

Table 3-2 Candidate Variables for Inclusion in the New Model

Discrete	Continuous
Approach angle	Percent truck
Development type	Passenger train count
Main track?	Hwy speed
Traffic lane type	Max timetable speed
Paved/unpaved	Highway functional class
Crossing surface type	Advanced warning
Urban/rural	

3.3.1 Discrete Explanatory Variables

The discrete variables are essentially category variables that indicate a crossing belongs to a particular category among two or more possibilities. The variables are represented in the data as integer values. However, there is no ordered relationship among the categories represented by the integers.

The method for evaluating the discrete variables for inclusion in the model was to consider crossings with 5-year accident history greater than 0. Researchers then examined a boxplot chart of accidents normalized for exposure and warning device types[25], grouped by the variable by its different levels. If the boxplot indicated significant variance across groupings (i.e., the groupings displayed different medians and other measures indicating variance), then the variable would be considered for inclusion in estimation. If the boxplot displayed no such variance, the team concluded that the variable did not have a strong impact on accident prediction and would be excluded.

As an example, the following chart shows the boxplot for the variable of grade crossing surface type. Researchers aggregated the two categories of "Concrete" and "Concrete and Rubber." This variable displays variance across its categories, so it was flagged for inclusion in the new model.

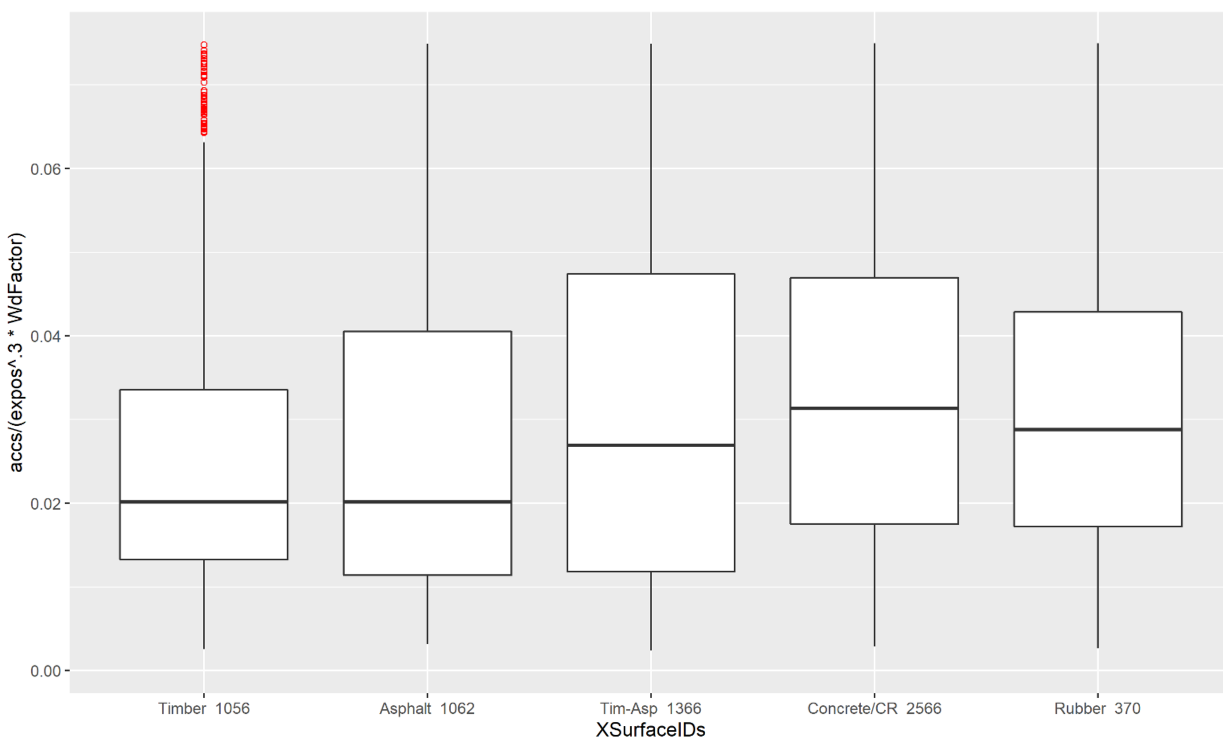
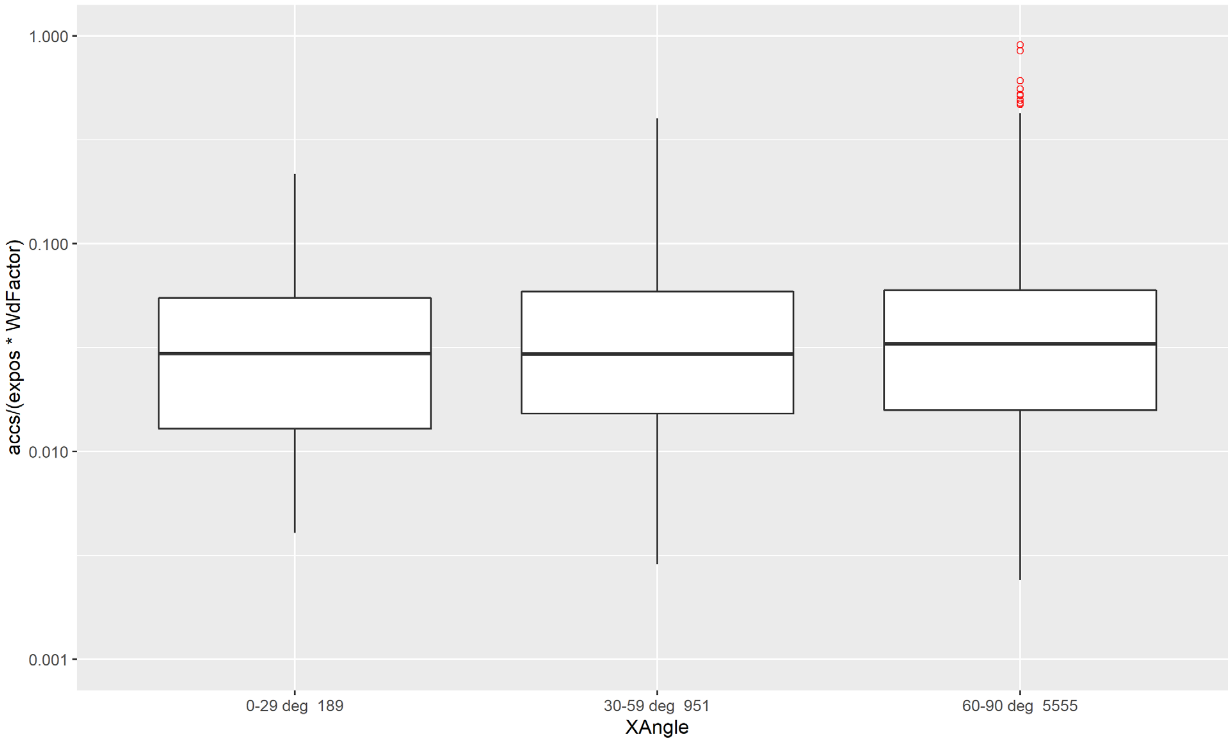


Figure 3-1. Boxplot of Normalized Crossing Accidents by Grade Crossing Surface Type

[25] "Accidents normalized for exposure and warning device types" means accidents in 5-year history divided by the product of exposure and a risk factor for the warning device type. The risk factors used were: passive = 1.0, lights = 0.3 and gates = 0.1. These values are based on the analysis of the previous Section.

The following chart shows the boxplot for the variable of grade crossing angle. There is very little variance across the groupings. Consequently, this variable was excluded from the model.

Figure 3-2. Boxplot of Normalized Crossing Accidents by Grade Crossing Angle



Following the review of the discrete variables, it was found that the following variables warranted inclusion in the model: 1) Crossing surface type, and 2) RuralUrban.

3.3.2 Continuous Explanatory Variables

The grade crossings characteristics that are continuous variables were ordered (i.e., all variable values are comparable, and if values are different, then one is greater than the other). Each can assume a range of values, not necessarily integers. However, data specifications typically restrict the values to integers (e.g., maximum timetable speeds can assume values from 1 to 99).

The method for evaluating the continuous variables for inclusion in the model was to consider crossings with 5-year accident history greater than 0. Researchers then examined a boxplot chart of accidents normalized for exposure and warning device types, grouped by the variable for each of its 10 deciles. If the boxplot indicated a good distribution of the variable, and an observed functional relationship across deciles, then the variable would be considered for inclusion in estimation, otherwise it was not.

The following chart shows the boxplot for the variable of maximum timetable speed.

There was a clear increasing trend for increasing decile. Consequently, this variable was included in the model.

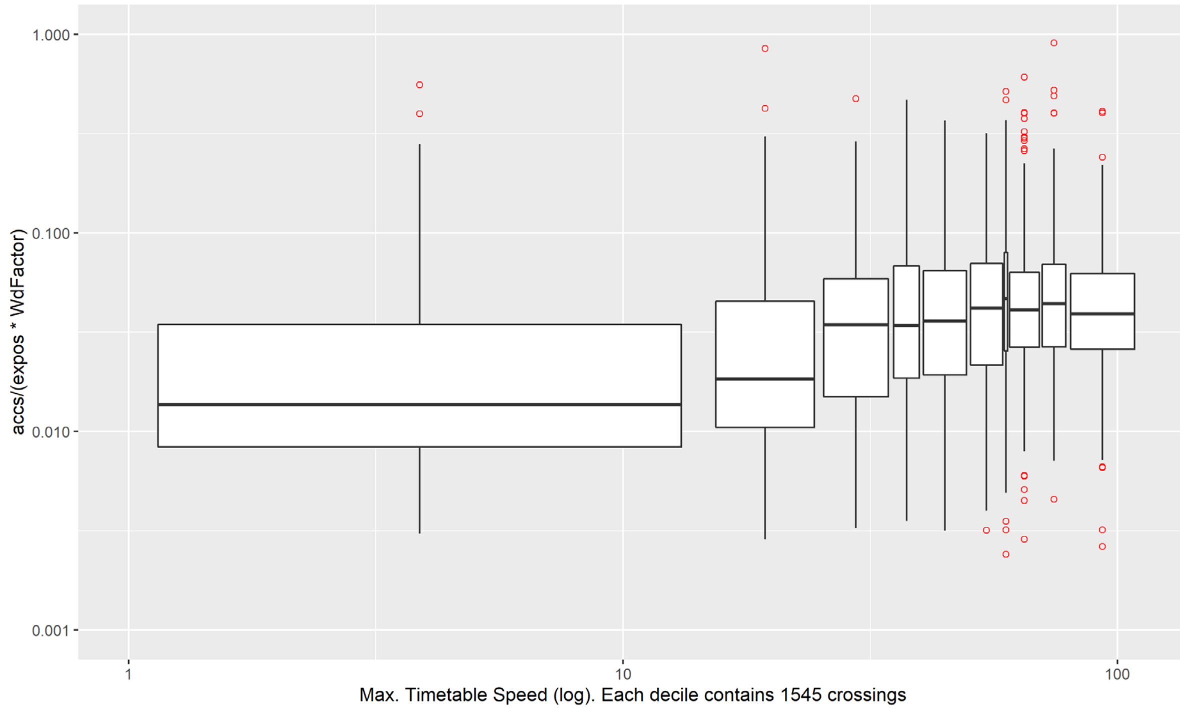


Figure 3-3. Boxplot of Normalized Crossing Accidents by Maximum Timetable Speed Deciles

The following chart shows the boxplot for the variable of percent truck of highway traffic.

There was no clear relationship that changes over deciles of the variable. Consequently, this variable was excluded from the model.

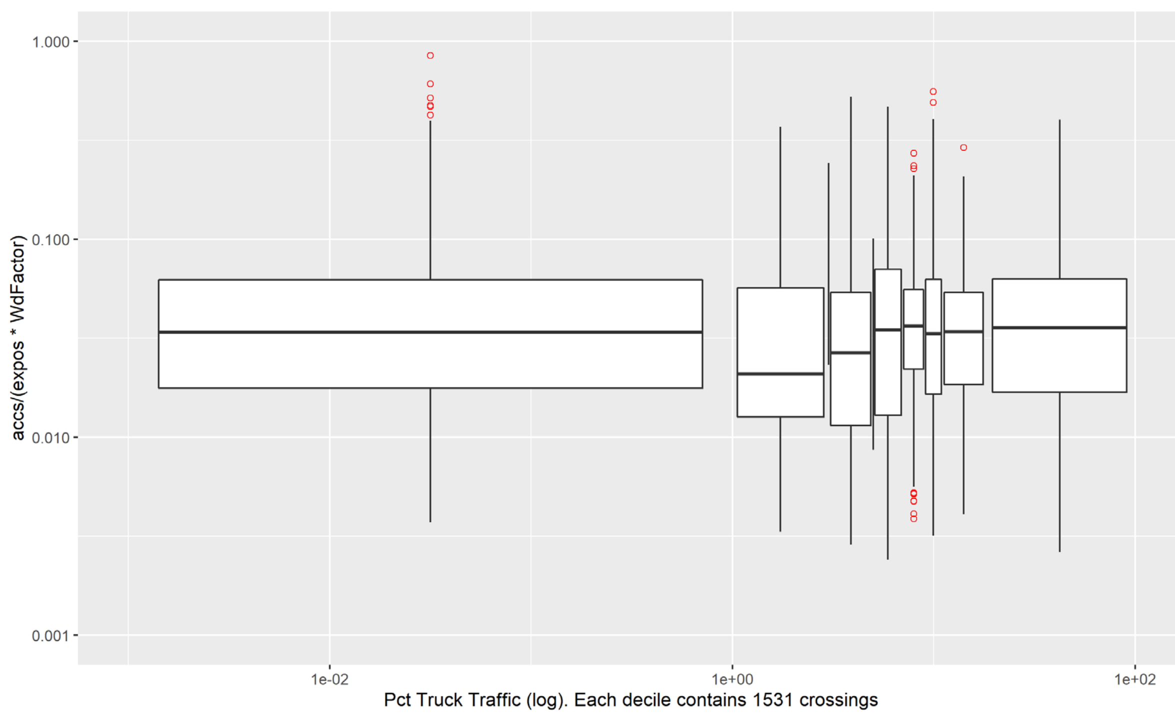


Figure 3-4. Boxplot of Normalized Crossing Accidents by Percent Truck Deciles

4. Accident Prediction Model

This section presents the selected accident prediction model, its regression with the ZINB estimation procedure, and the EB adjustment of the ZINB-predicted values.

ZINB is one type of zero-inflated models. It is used for count variables (e.g., accidents) that exhibit excess zeroes. "Excess zeroes" means that of the many crossings with no accidents in the preceding 5 years, some of those were crossings effectively had no risk of an accident.

The ZINB model assumes that:

- Each crossing has some non-zero probability of being a no-risk crossing.
- Each crossing has an expected number of annual accidents.
- Accident counts for the population of crossings conform to a negative binomial distribution (the standard deviation of accidents for the population is greater than the mean, indicating overdispersion).

ZINB has been adopted in numerous accident studies and is well-suited for the analysis of grade crossing accidents.

The EB method adjusts the estimate of the expected number of accidents so as to account for history, and correct for "regression to the mean"[26] bias. The formula relies on the ZINB regression outputs to estimate a weighting factor. The EB-adjusted estimate is a linear combination of the predicted accidents (from ZINB) and the actual count of accidents. If the accident history indicates no accidents, then the EB adjustment will adjust the expected value of accidents downwards toward zero. For crossings with non-zero accident history, EB will adjust the expected value (usually upward) so that it is closer to the actual count.

R software was used in the model estimation.

4.1 The Accident Prediction Model

Based on the analysis described in the previous sections, the selected accident prediction mode is shown below. The model has two components: 1) a count model and 2) a zero-inflated model.

Equation 3. The ZINB Count Model

$$\mathrm{N}_{CountPredicted} = e^{[\beta_0+\beta_1\cdot \mathrm{lExpo}+\ \beta_2\cdot \mathrm{D}_2+\beta_3\cdot \mathrm{D}_3+\beta_4\cdot \mathrm{RurUrb}+\beta_5\cdot \mathrm{XSurfID2s}+\beta_6\cdot \mathrm{lAadt}+\beta_7\cdot \mathrm{lMaxTtSpd}]}$$

[26] "Regression to the mean" basically means that if a variable is extreme the first time you measure it, it will be closer to the average the next time you measure it. For example, if we randomly selected a crossing that had several accidents in its 5-year history (that is, a very high risk grade crossing), the next random selection would be a crossing whose risk was much closer to the mean for all grade crossings.

Equation 4. The ZINB Zero-Inflated Model

$$P_{InflatedZero} = \frac{z}{1+z}$$
$$z = e^{[\gamma_0 + \gamma_1 \cdot \text{lTotalTrains}]}$$

Equation 5. The ZINB Combined Model

$$N_{Predicted} = N_{CountPredicted} \cdot \left(1 - P_{InflatedZero}\right)$$

where:

$N_{CountPredicted}$	Predicted accidents of count model (data for left-hand side of regression are counts of accidents at crossings in 5-year period 2014–2018)
$P_{InflatedZero}$	The probability that the grade crossing is an “excess zero”
$N_{Predicted}$	Predicted accidents after accounting for excess zeroes
lExpo[1]	Exposure, equal to average annual daily traffic times daily trains
D_2	If warning device type is lights =1, 0 otherwise
D_3	If warning device type is gates =1, 0 otherwise
	(note: if both D_2 and D_3 are zero, then warning device type is passive)
RurUrb	If Rural = 0, if Urban = 1
XSurfID2s	Timber = 1, Asphalt = 2, Asphalt and Timber OR Concrete OR Rubber = 3, Concrete and Rubber = 4
lMaxTtSpd[1]	Maximum timetable speed (integer value between 0 and 99)
lAadt[1]	Average annual daily traffic
lTotalTrains[1]	Total number of daily trains

[1]These variables have been transformed as follows: lx = log(1+αx), where x is the original variable and α is a factor. The factor α was selected so that for the median value of x, ln(1+αx) = ln(x)

4.2 ZINB Regression

The ZINB regression model has two components: the count model and the zero-inflated model. The count model is for predicted accidents before considering the probability of excess zeroes. The zero-inflation model is for estimating the probability of an inflated zero. (An “inflated zero” is a zero accident count that does not derive from a grade crossing’s traits; rather, it is zero because the crossing accident risk is effectively 0.) Note that the explanatory variable for the zero-inflated model is the total number of trains; that is, the fewer trains at a grade crossing the higher the probability of an excess zero.

The predicted (fitted) values of the model are given by f(**x**)*(1-g(**s**)), where f is the count model (operating on the vector of inputs **x** for each observation) and g is the zero-inflation model (operating on the vector of inputs **s** for each observation).

The following table shows the output for the zero-inflated negative binomial regression for the model in the previous section.

Table 4-1. ZINB Regression Output

Pearson residuals

Minimum	1st Quartile	Median	3rd Quartile	Maximum
-0.6559	-0.2742	-0.2072	-0.1504	28.5137

ZINB regression count model coefficients (negative binomial with log link)

Variable	Estimate	Std. Error	Z value	Pr(>|z|) (p-value)	Significance Code
(Intercept)	-8.35922	0.32079	-26.059	< 2e-16	***
lExpo	0.19023	0.02866	6.638	3.18e-11	***
D2	-0.28478	0.04806	-5.926	3.10e-09	***
D3	-0.85770	0.04089	-20.976	< 2e-16	***
RurUrb	0.39346	0.03162	12.444	< 2e-16	***
XSurfaceID2s	0.13182	0.01715	7.686	1.52e-14	***
lMaxTtSpd	0.68760	0.68760	22.702	< 2e-16	***
lAadt	0.10626	0.10626	3.511	0.000446	***
Log(theta)	-0.25934	.08867	-2.925	.003447	**

ZINB regression zero-inflation coefficients (binomial with logit link)

Variable	Estimate	Std. Error	z-value	Pr(>|z|) (p-value)	Significance Code
(Intercept)	1.17084	0.19001	6.162	7.19e-10	***
lTotalTr	-1.01088	0.08452	-11.961	< 2e-16	***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

ZINB regression summary statistics

Log-likelihood	AIC
-2.462e+04	49260.26

Notes on the regression output:

- The values in the column labeled "coefficients" correspond to the βs from the count model in the formulas and γs from the zero-inflation model.
- The column "Std. Error" shows the standard error of the coefficient to the left.

- The “z-value” column is the coefficient divided by the standard error (larger absolute values of z indicate that the coefficient has greater statistical significance).
- “Pr(>|z|)” is the probability of exceeding the absolute value of the z-value (smaller values indicate greater statistical significance).
- The *s in the rightmost column of the table indicate the statistical significance (see “Significance codes”).
- θ[27] is the inverse of the overdispersion (α) parameter of the count model. The estimate of θ is 0.7716 (and the imputed value of α=1.296). α was expected to be greater than 1.
- AIC is the Akaike information criteria for model quality given the dataset.

Key points to note from the regression output:

- The coefficients for lExpo and lAadt have positive signs with expected magnitudes.
- The coefficients for D2 and D3 are negative (i.e., compared to passive devices, lights and gates reduce risk). The coefficient of D3 is about three times that of D2, which conforms to expectations.
- The signs and magnitudes of other coefficients in the count model seem to correspond to expectations.
- The coefficient of lTotalTrains in the zero-inflation model is negative, i.e., the probability of an excess zero decreases with the number of trains, as expected.
- All the coefficients have strong statistical significance.[28]
- The AIC[29] is the least value for all tested models.
- The estimated mean and standard deviations for the population are:
  - Mean: 0.08316
  - Standard deviation: 0.21377

Figure 4-1 is a chart of the ZINB predicted values grouped by device type. The vertical lines on the chart indicate the average log of exposure for each grouping. The horizontal lines on the chart indicate the average predicted 5-year accidents for each grouping. The vertical line indicates the average log of exposure for each grouping.

---

[27] θ is the Greek letter “theta.”

[28] “Strong statistical significance” for an estimated coefficient means there is a very small probability of falsely rejecting the null hypothesis (i.e., the hypothesis that the coefficient is actually 0).

[29] From Wikipedia: The Akaike information criterion (AIC) is an estimator of out-of-sample prediction error and thereby relative quality of statistical models for a given set of data. For a statistical model, let k be the number of estimated parameters in the model. Let L be the maximum value of the likelihood function for the model. Then the AIC value of the model is the following. AIC = 2k - 2*ln(L)

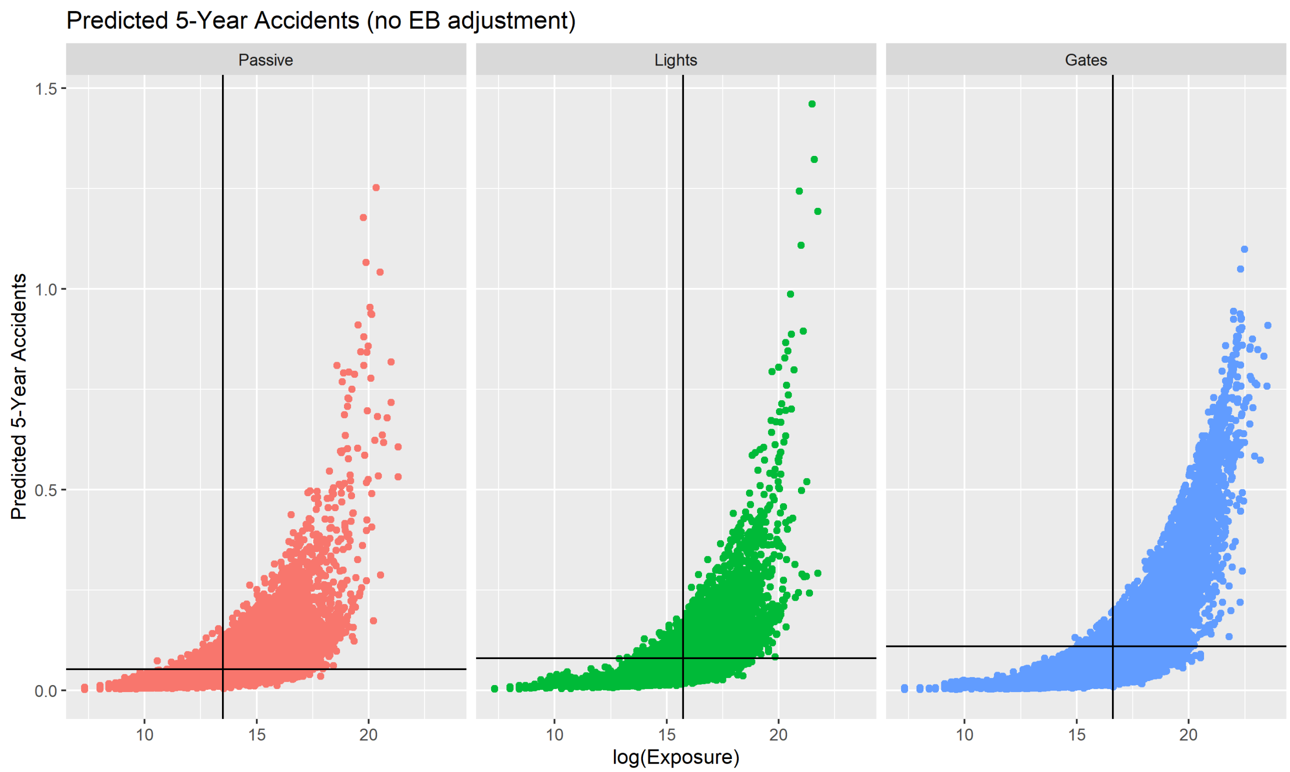


Figure 4-1. ZINB Predicted Accidents by Warning Device Type

4.3 Predicting Accidents from the Regression Outputs

One can apply Equation 2, Equation 3, and Equation 4 above to calculate the predicted accident at a grade crossing (prior to applying the EB adjust described in the following section). The predicted accidents are the fitted values (i.e., $\hat{Y}$) of the model.

The βs in the equations are the ZINB count model coefficient estimates and the γs are the ZINB zero-inflated model coefficients estimates.

4.4 Empirical Bayes Prediction Adjustment

The EB adjustment intends to correct the prediction for "regression to the mean" bias while adjusting the expected value to account for accident history. The process is described in Hauer.[30] For each grade crossing, the expected number of accidents is given by:

Equation 6. Empirical Bayes Adjustment

$$\mathrm{N}_{Expected} = \mathrm{w} \cdot \mathrm{N}_{Predicted} + (1 - \mathrm{w}) \cdot \mathrm{N}_{Observed}$$

where:

[30] E. Hauer, The Art of Regression Modeling in Road Safety, Springer 2015.

$N_{Expected}$	The adjusted number of predicted accidents
$N_{Predicted}$	The number of predicted accidents from the ZINB regression procedure
$N_{Observed}$	The number of observed accidents (i.e., count of accidents at the grade crossing)

and the weighting factor w is given by:

Equation 7. EB Weighting Factor

$$w = \frac{1}{1 + \frac{V[N_{Predicted}]}{N_{Predicted}}}$$

The variance of $N_{Predicted}$ is given by:

Equation 8. Variance of Crossing's Predicted Number of Accidents

$$V[N_{Predicted}] = N_{Predicted} \cdot 1 + \left[N_{CountPredicted} \cdot \left(P_{InflatedZero} + \frac{1}{\theta}\right)\right]$$

where 1/θ, as noted above, is the inverse of the overdispersion parameter α from the ZINB regression (θ is estimated to be 0.7716).

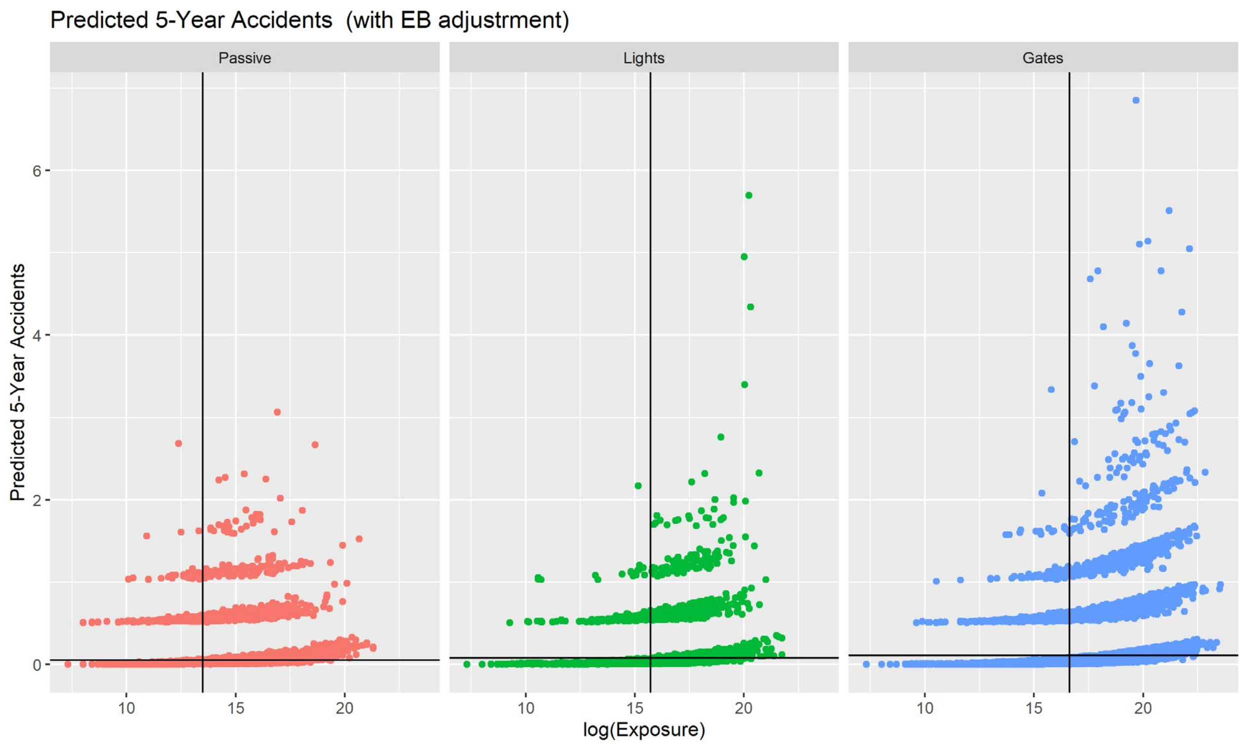
Note that the underlying assumptions of the model indicate that the accident count data for a population of crossings is best described by the NB distribution. The overdispersion parameter describes the overdispersion of data relative to a Poisson distribution (where mean and variance are assumed equal). R software defines the variance of the count variable as $\mu+\mu^2/\theta$.[31] Given this definition of variance, θ should be less than 1 and greater than 0.

Figure 4-2 shows the predicted values grouped by device type, with this chart showing the predicted values including the EB adjustment.

Compared to Figure 4-1, this chart shows the predicted values clustered around the values that represent the accident counts in each grade crossing's 5-year accident history.

[31] Most other software packages (e.g., SAS, Stata, Limdep, SPSS, etc.) define the variance of the count variable as $\mu+ \alpha \cdot \mu^2$. R's θ is equivalent to 1/ α in the other packages. α is the overdispersion parameter of the negative binomial distribution, as defined in these other packages and most of the academic literature.

Figure 4-2. ZINB+EB Predicted Accidents by Warning Device Type



5. Accident Severity Model

Grade crossing management in the U.S. considers three severity categories: fatal, injury and property damage only (PDO). A fatal accident is one with at least one fatality; an injury accident has at least one injury; and a PDO accident has no injuries or fatalities.

The accident severity model seeks to determine the probabilities of prospective accidents at grade crossings belonging to each severity category. The process for predicting accident severity is one of allocating predicted accidents to each severity category. In the APS, there is no process to calibrate accident severity. Over time, accident severity has been fairly stable: fatal accidents are about 10 to 12 percent of the total, injury accidents about 27 percent, and PDO accidents about 61 percent.

The remainder of the section describes the data, the logistic regression process used in the model estimation, and the model results. Some comparisons of the new model with the APS are discussed in the next section.

R software was used in the model estimation.

5.1 Description of the Data

Federal law requires filing a Form 57 accident report for each grade crossing accident. The analysis used the Form 57 report database and GCIS. Researchers examined accidents in the period 2014–2019 (6 years) during which there were 12,983 accidents. They excluded from the model estimation process accidents from the following crossings:

- Private crossings
- Crossings where traits were missing data for key explanatory variables.

There were 11,131 accidents at public crossings. Of these, 9,870 contained all the data for key explanatory variable, and these were included in the model estimation. Of the 9,870 accidents, 1,355 (13.7 percent) were fatal, 2,768 (28.0 percent) were injury accidents, and 5,747 (58.2 percent) were PDO.

These accidents were matched with the grade crossing data from GCIS for each crossing where an accident occurred.

5.2 The Accident Severity Model

For the accident severity model, we seek to estimate the probabilities that given an accident, the accident will be one of three types: fatal, injury or PDO. The explanatory variables for these estimates are grade crossing characteristics. The research sought, therefore, to model three variables:

Equation 9. Probabilities to Estimate – Fatal

$$P(acctype = fatal \mid A)$$

Equation 10. Probabilities to Estimate – Injury

$$P(acctype = injury \mid A)$$

Equation 11. Probabilities to Estimate – PDO

$$P(acctype = PDO \mid A)$$

keeping in mind the following constraint:

Equation 12. Constraint that Severity Probabilities Sum to 1

$$P(acctype = fatal \mid A) + P(acctype = injury \mid A) + P(acctype = PDO \mid A) = 1$$

One of the statistical methods that is well-suited for this type of problem is multinomial logistic regression, which is described in the following section. "Multinomial" refers to the fact that the left-hand side variable of the regression, accident type, can assume more than two values – in this case, three values (fatal, injury, PDO).

The regression analysis requires the selection of a reference level for the accident type variable, and the team selected "fatal." The regression estimates the probability of each of the other two values (injury, PDO) relative to the reference level (fatal). The fact that the three probabilities sum to 1 enables one to readily derive forecast formulas (see the section below on forecast formulas).

Based on analysis of the data, the following is the new accident severity model, with "fatal" selected as the reference level of the left-hand side variable for the regression:

Equation 13. Accident Severity Model – Injury Relative to Fatal

$$ln\left(\frac{\mathrm{P}(acctype = injury \mid \mathrm{A})}{\mathrm{P}(acctype = fatal \mid \mathrm{A})}\right) =$$

$$\beta_{20} + \beta_{21} \cdot \mathrm{lMaxTtSpd} + \beta_{22} \cdot \mathrm{lTrains} + \beta_{23} \cdot \mathrm{RuralUrban} + \beta_{24} \cdot \mathrm{D}_2$$

Equation 14. Accident Severity Model – PDO Relative to Fatal

$$ln\left(\frac{\mathrm{P}(acctype = PDO \mid \mathrm{A})}{\mathrm{P}(acctype = fatal \mid \mathrm{A})}\right) =$$

$$\beta_{30} + \beta_{31} \cdot \mathrm{lMaxTtspd} + \beta_{32} \cdot \mathrm{lTrains} + \beta_{33} \cdot \mathrm{RuralUrban} + \beta_{34} \cdot \mathrm{D}_2$$

where:

P(acctype = fatal | A)	The probability of a fatal accident given an accident A
P(acctype = injury | A)	The probability of an injury accident given an accident A
P(acctype = PDO | A)	The probability of a PDO accident given an accident A
lMaxTtSpd	Natural log of the maximum (rail) timetable speed at the crossing
lTrains	Natural log of the total number of daily trains at the crossing
RuralUrban	1 if crossing is in a rural (non-urban) environment, 0 if in urban
D2	Has value 1 if warning device type is lights, 0 otherwise

5.3 Multinomial Logistic Regression

Multinomial logistic regression is used when the dependent variable in question is nominal (equivalently categorical, meaning that it falls into any one of a set of categories that cannot be ordered in any meaningful way) and for which there are more than two categories. In this study, researchers studied the probability that an accident will be one of three accident types.

The problem is one of statistical classification. There is a dependent variable to be predicted that comes from one of a limited set of accident types. There is, as well, a set of independent variables (also known as features, explanators, traits, etc.), which are used to predict the dependent variable. Multinomial logistic regression is a particular solution to classification problems that uses a linear combination of the observed features and some problem-specific parameters to estimate the probability of each particular value of the dependent variable.

The multinomial logistic regression output for the model described in the previous section is shown in Table 5-1.

Table 5-1. Accident Severity Multinomial Logistic Regression Output (with “fatal” as the selected reference level)

Regression Output – Accident Severity				
Part A – For a given accident, probability of an injury accident (relative to a fatal accident)				
Variable	**Estimate**	**Std. Error**	**z-value**	**Pr(|z|)>0**
Intercept	5.248627	0.355109	14.78032	0
lMaxTtSpd	-0.92544	0.097943	-9.44876	0
lTrains	-0.28326	0.042458	-6.6716	2.53e-11
RuralUrban	-0.27408	0.072886	-3.76042	0.00017
D2	0.489354	0.141041	3.469598	0.000521
Part B – For a given accident, probability of a PDO accident (relative to a fatal accident)				
Variable	**Estimate**	**Std. Error**	**z-value**	**Pr(|z|)>0**
Intercept	6.957135	0.339015	20.52161	0
lMaxTtSpd	-1.23128	0.092907	-13.2528	0
lTrains	-0.22114	0.039411	-5.61125	2.01e-08
RuralUrban	-0.24085	0.067191	-3.58462	0.000338
D2	0.330487	0.135769	2.434192	0.014925
Part C – Summary Statistics				
Residual Deviance:	17986.35			
AIC:	18006.35			

The coefficient estimates are highly significant. The exception is the coefficient for D2 in Part B, whose significance level is 97.6 percent (i.e., there is a 2.4 percent probability of a type I error[32]). The value for the AIC is the least among all of the variable combinations that were tested.

Interpreting the coefficient estimates can be tricky. The estimates (except for the intercept and D2) are negative, indicating that the relative risk of injury or PDO to fatal decreases – which means that the risk of a fatal accident (the most severe) increases.

[32] A type I error occurs when rejecting a true null hypothesis.

5.4 Accident Severity Forecast Formulas

Equations 15–17 show the forecast formulas for the accident severity model.

Equation 15. Accident Severity Forecast Formulas - Fatal

$$\Pr(Y_i = fatal \mid A) = \frac{1}{1 + \sum_{k=2}^{3} e^{\beta_k \cdot X_i}}$$

Equation 16. Accident Severity Forecast Formulas - Injury

$$\Pr(Y_i = injury \mid A) = \frac{e^{\beta_2 \cdot X_i}}{1 + \sum_{k=2}^{3} e^{\beta_k \cdot X_i}}$$

Equation 17. Accident Severity Forecast Formulas - PDO

$$\Pr(Y_i = PDO \mid A) = \frac{e^{\beta_3 \cdot X_i}}{1 + \sum_{k=2}^{3} e^{\beta_k \cdot X_i}}$$

Notes to formulas:

The subscript k indicates accident type: fatal = 1, injury = 2, PDO = 3.

The subscript i indicates a grade crossing.

Y_i is the variable indicating accident type (fatal, injury or PDO).

$\boldsymbol{\beta}$s are the vectors of coefficient estimates.

$\boldsymbol{\beta}_{2j}$ is the coefficient estimate vector for probability of injury accident relative to fatal, and Part A of Table 5-1 contains the coefficient estimate values.

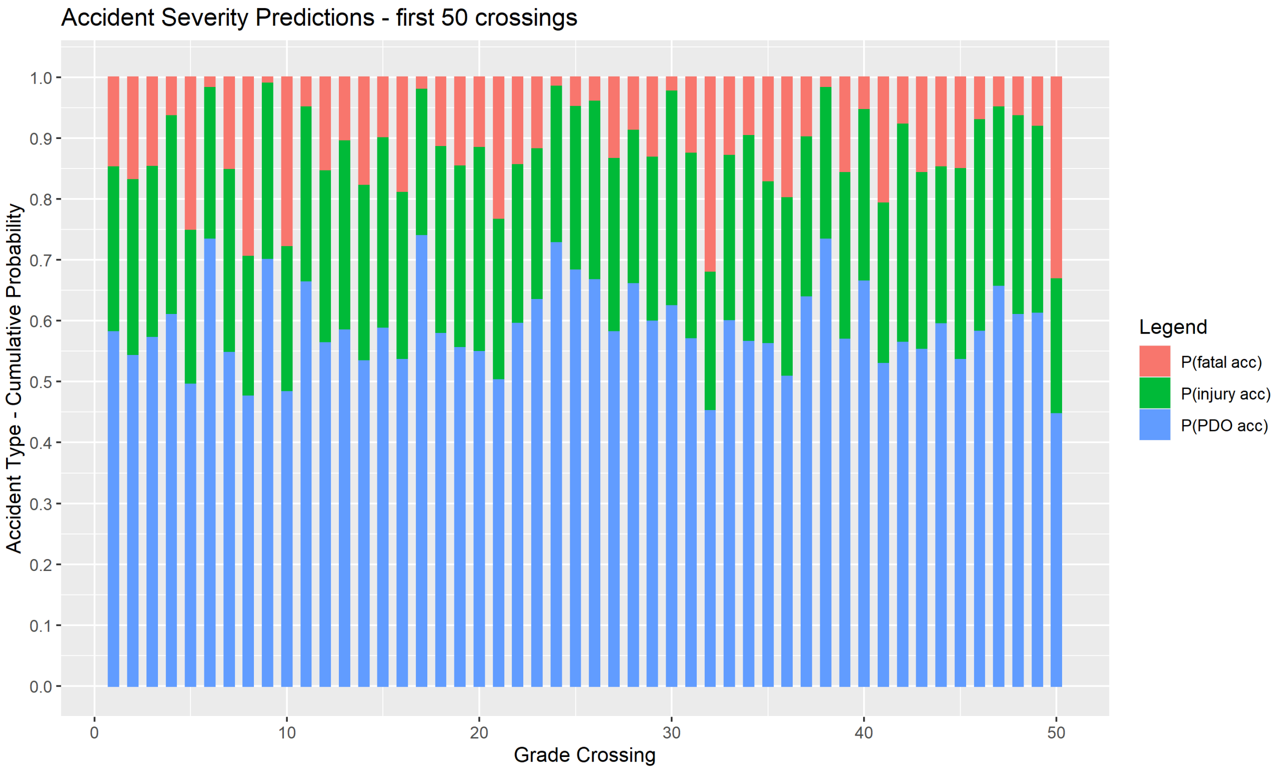
$\boldsymbol{\beta}_{3j}$ is the coefficient estimate vector for probability of PDO accident relative to fatal, and Part B of Table 5-1 contains the coefficient estimate values.

The subscript j (0 to 4) indicates the explanatory variable to which the $\boldsymbol{\beta}$ element corresponds. For example, 0 is for the intercept, 1 is for the variable lMaxTtSpd, and so on.

$\mathbf{X}_i$ is the vector of crossing traits that explain accident severity.

The following chart shows forecast severity for 50 accidents with the new model.

Figure 5-1. Severity Predictions for 50 Crossings with the New Model



6. Validation

The section presents validations for the new model (estimated with the ZINB and EB methods). Note here that the term “prediction” means the expected value of accidents at the crossing. In general, accidents are rare and the (annualized) expected value of accidents at a crossing will be a real value between 0 and 1. A non-zero accident count will be larger in most cases than the expected value of accidents at a crossing, which reflects the fact that the observed count in a previous year is not expected to repeat frequently in subsequent years.

The first validation compares cumulative predicted accidents by the new model and the APS with the actual risk as measured by accident counts.

The second validation shows the predicted accidents for the new model and the APS for crossings grouped by accident count.

The third comparison examines the model results (the new model and APS) for different groupings of high-risk crossings and shows the results in a chart. In this case, researchers counted accidents at the 50 highest-risk crossings (and then at the subsequent groupings of highest-risk crossings). The better of the two models will predict accidents at the groupings of crossings that is closer to the actual accident counts.

For the severity model, this report shows comparisons of the model performance with that of the APS.

6.1 Accident Prediction – Cumulative Risk

For this validation we order the grade crossings from high risk to low risk (according to total accidents in 5-year history). The y-axis on the charts below shows the actual cumulative risk and the predicted risk with each model. The better model is the one that tracks closer to the actual cumulative risk.

The four charts below represent two cases and two periods. The first case displays cumulative accident count and predictions for all crossings in the estimation sample (which includes 94,029 crossings). The second case focuses on the crossings with non-zero accidents. The first period is the estimation period 2014–2018. The second period is the following year, which covers 5-year accidents from 2015–2019.

The vertical line indicates the boundary between those crossings with non-zero accidents in the period (to the left of the line) and those with zero accidents in the period (to the right of the line).

Figures 6-1 and 6-2 show the counts and predictions, ordered from high to low risk, for the complete set of crossings in the estimation sample. Figure 6-1 is for the period 2014–2018. Figure 6-2 is for the period 2015–2019.

Figures 6-3 and 6-4 show the same chart data as Figures 6-1 and 6-2, but limit the data displayed to those crossings with non-zero accident history.

The charts demonstrate that the new model was the better predictor of accident risk than the APS.

Figure 6-1. Model Comparison (2014–2018, all crossings in sample)

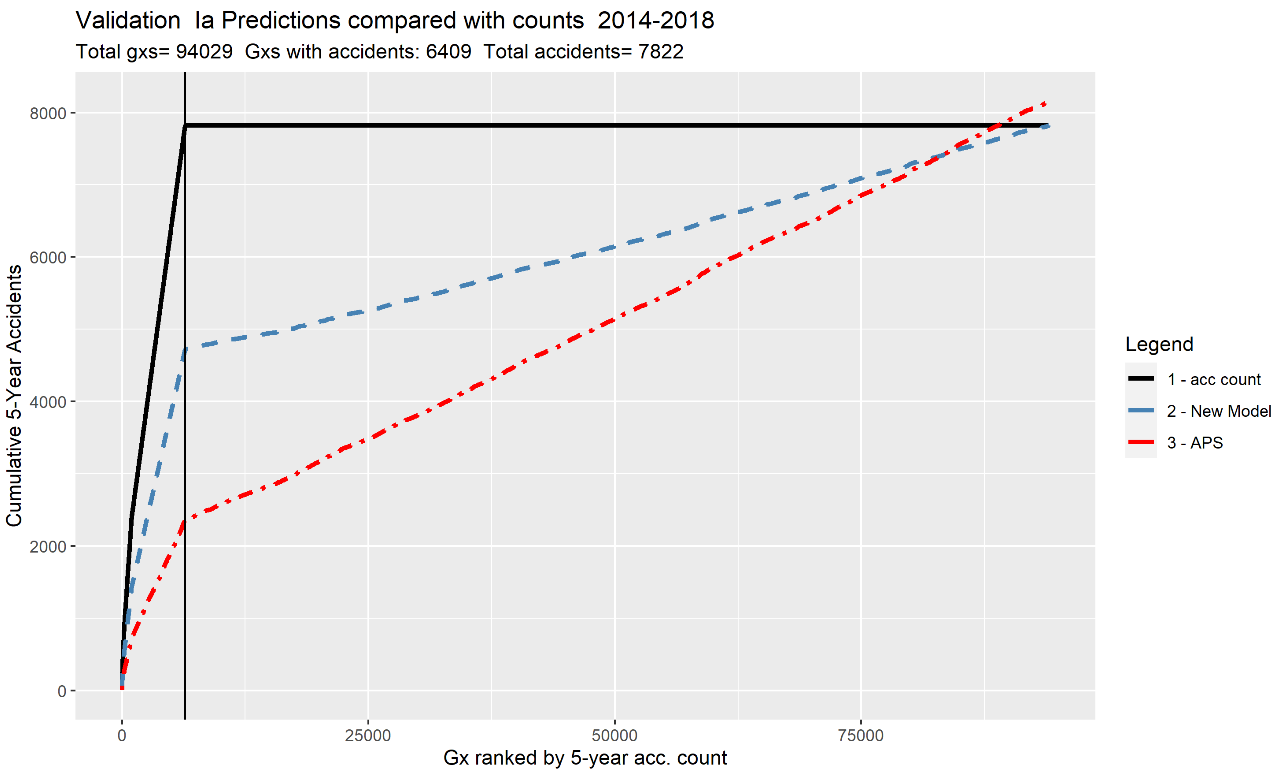


Figure 6-2. Model Comparison (2015–2019, all crossings in sample)

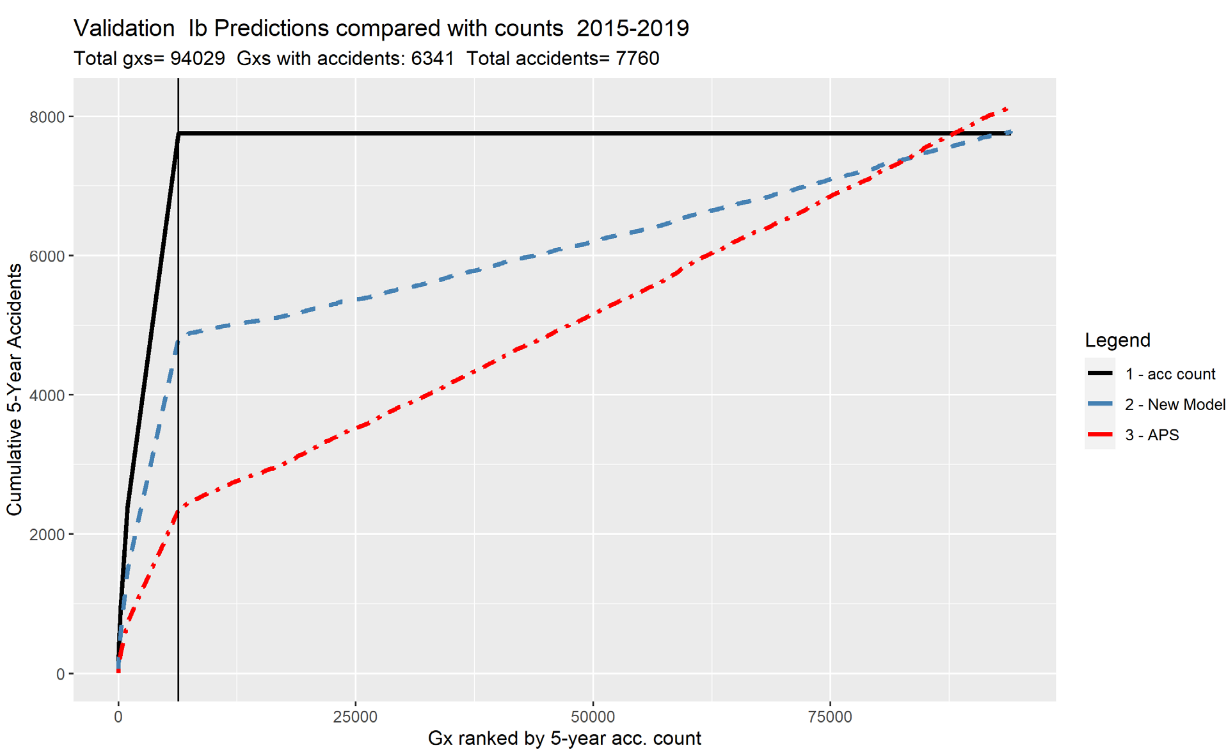


Figure 6-3. Model Comparison (2014–2018, crossings in sample with non-zero accidents)

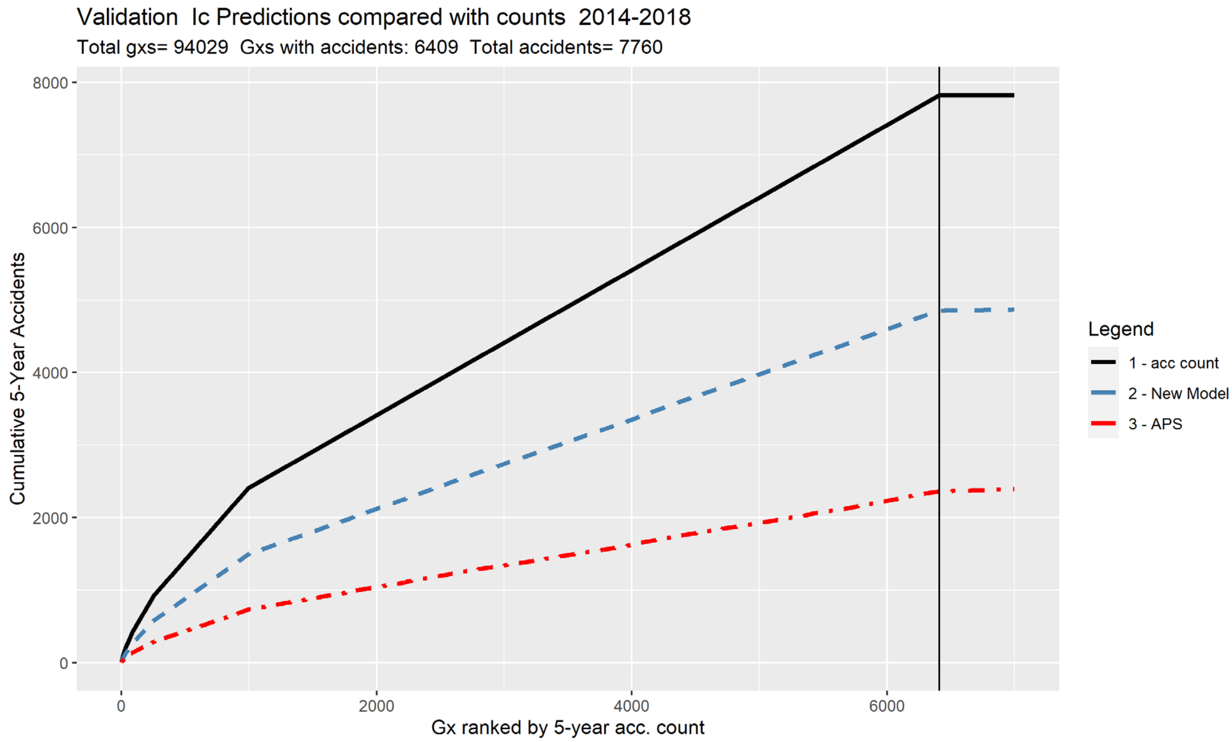
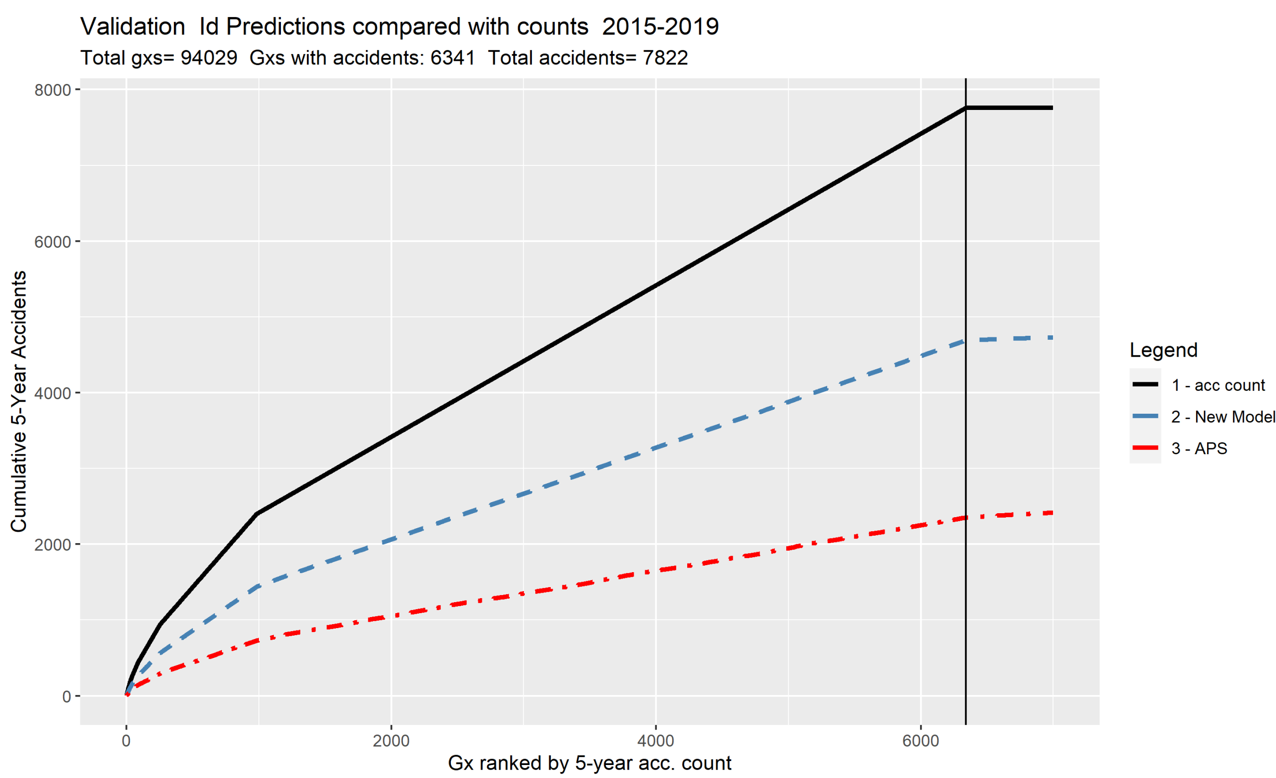


Figure 6-4. Model Comparison (2015–2019, crossings in sample with non-zero accidents)



On the riskiest crossings, the new model (ZINB+EB) predicted cumulative accident risk much better than APS.

6.2 Accident Prediction – Risk at Crossings by Accident Count Groups

In the second validation, researchers grouped the crossings by the number of accidents in the 5-year history. The chart shows the number of accidents in the grouping on the x-axis

The orange square markers show mean predicted accidents with the APS given traits at the crossings with the specified accident history (shown on the x-axis). The square blue markers show mean predicted accidents with the new model. The lines below and above the markers indicate the 10^{th} and 90^{th} percentiles, respectively. The lines also indicate the bounds of the 80 percent confidence interval of the prediction for crossings in the period.

For example, in Figure 6-5 below (displaying the period 2014–2018) at crossings having three accidents the new model predicted between 1.6 and 2.0 accidents. The APS predicted 0.5 to 1.4. The new model better predicted the crashes at crossings for each level of accident risk than the APS.

Figure 6-6 shows the results for the period 2015–2019.

Figure 6-5. Model Comparison, Accident Counts, and Predictions (2014–2018)

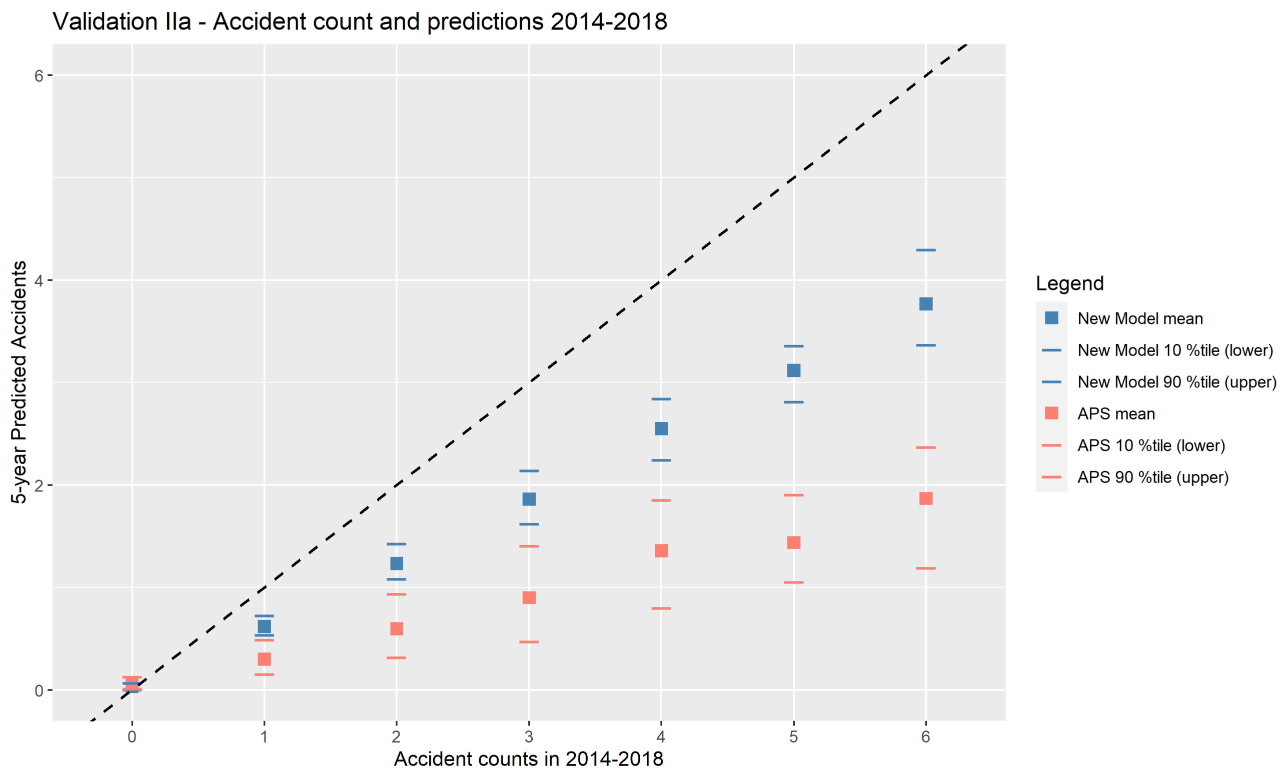
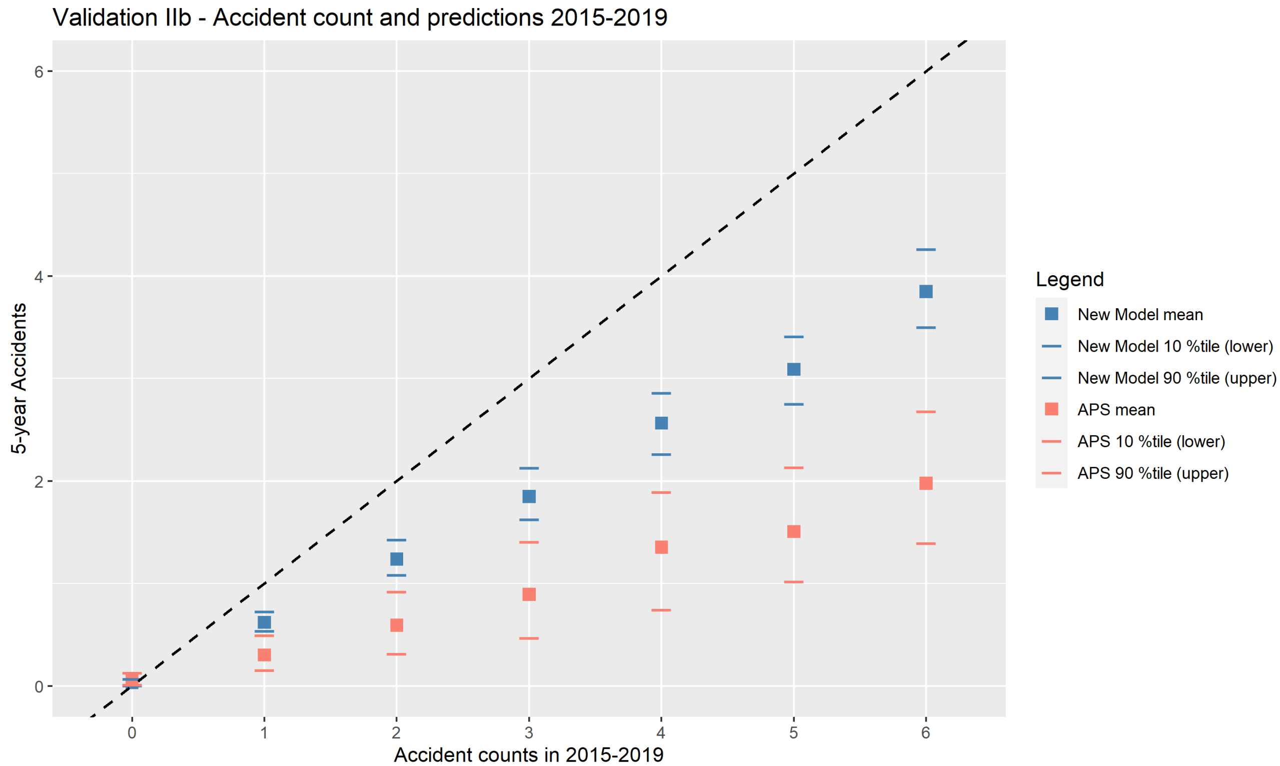


Figure 6-6. Model Comparison, Accident Counts, and Predictions (2015–2019)



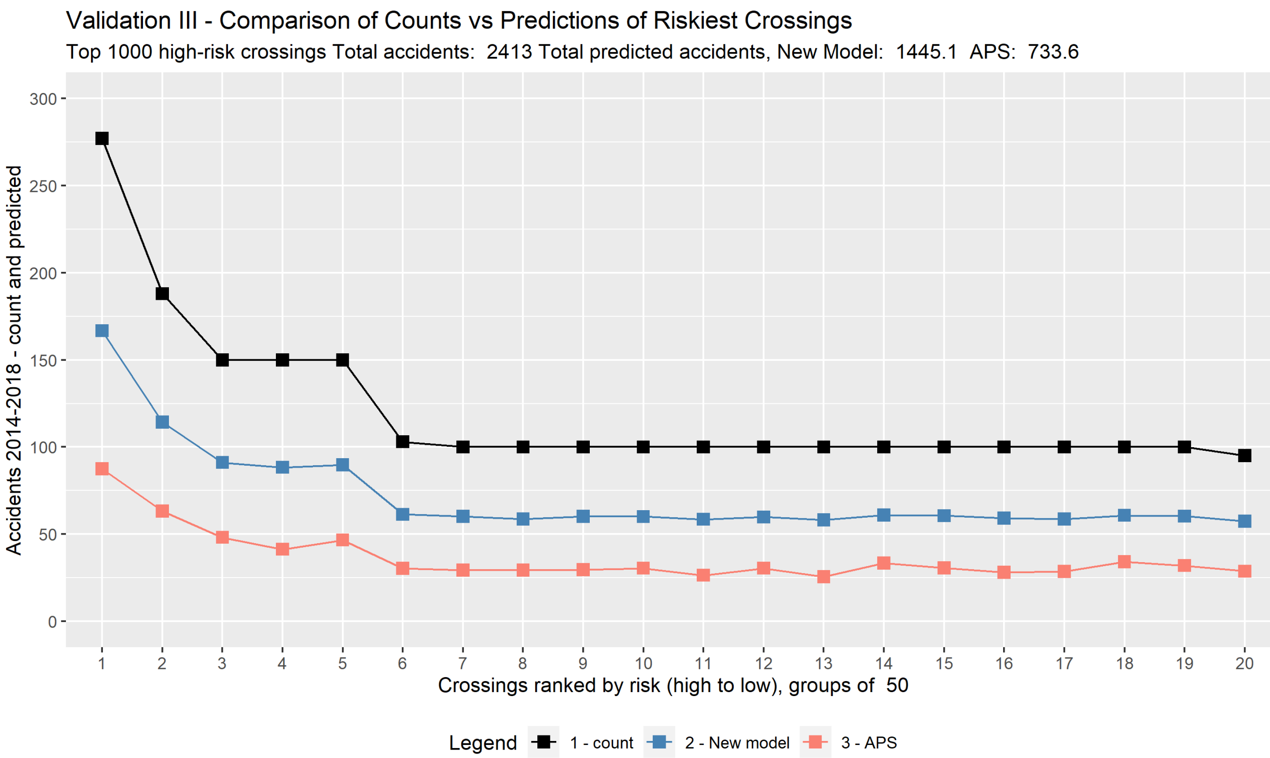
6.3 Accident Prediction – Accident Risk for Groups of High-Risk Crossings

The third validation examines the model results (APS and new model) for groupings of high-risk crossings and shows the results in a chart. The better of the two models will predict accidents at each grouping of crossings that is closer to the actual accident counts.

Crossings in the estimation sample were ordered by decreasing risk, and then divided into groups of 50. In the figure below, the x-axis shows groupings 1 to 20 (20 groups of 50 equals total of 1,000). The y-axis shows the actual and predicted crossings by model (new model and APS) for each grouping.

For each grouping, the new model performed better than the APS. For the top 1,000 high-risk crossings in 2014–2018 the accident count was 2,578 accidents. The APS predicted 791.3 accidents while the new model predicted 1,518.0 accidents at these 1,000 high-risk crossings.

Figure 6-7. Comparison of Predictions for Riskiest Crossings



6.4 Accident Severity – Model Comparisons

The table below shows the predicted accident severity for all accidents and by each accident type in the severity estimation sample.

Table 6-1. Predicted Severity (Percent of Total) by the New Model and APS

		For Accidents of Severity Type			
		All Accidents	Fatal	Injury	PDO
New Model Predictions	Fatal	13.7	18.7	13.1	12.9
	Injury	28.4	27.3	28.4	28.0
	PDO	58.2	54.0	58.4	59.1
APS Predictions	Fatal	6.9	7.2	6.9	6.9
	Injury	27.3	26.9	27.4	27.4
	PDO	65.7	65.9	65.6	65.7

With the new model, the aggregate percentage of accidents of each accident type exactly equaled the percentages in the sample (as expected). The APS predictions in the aggregate diverged somewhat from the sample data; for example, APS predicted the percent of fatal accidents to be half of the actual percentage.

We gauge the predictive performance of the severity model by estimating the predicted percentage of a severity category. We do this while only considering those accidents in that category. That value should well exceed the percentage of a severity category when considering all accidents.

Table 6-1 shows that predicted fatal accidents with the new model increased from a mean of 13.7 percent for all accidents to 18.7 (a 36 percent increase). When considering only accidents that were actually fatal. The comparable change with APS was 6.9 to 7.2 (a 4 percent increase). Overall, the new model performed better, with more significant movements in the correct direction when restricting to accidents of a particular type.

Figures 6-8 and 6-9 below show boxplot charts of predicted accident severities for the new model and APS.

Figure 6-8 Distribution of Predicted Accident Severities with the New Model

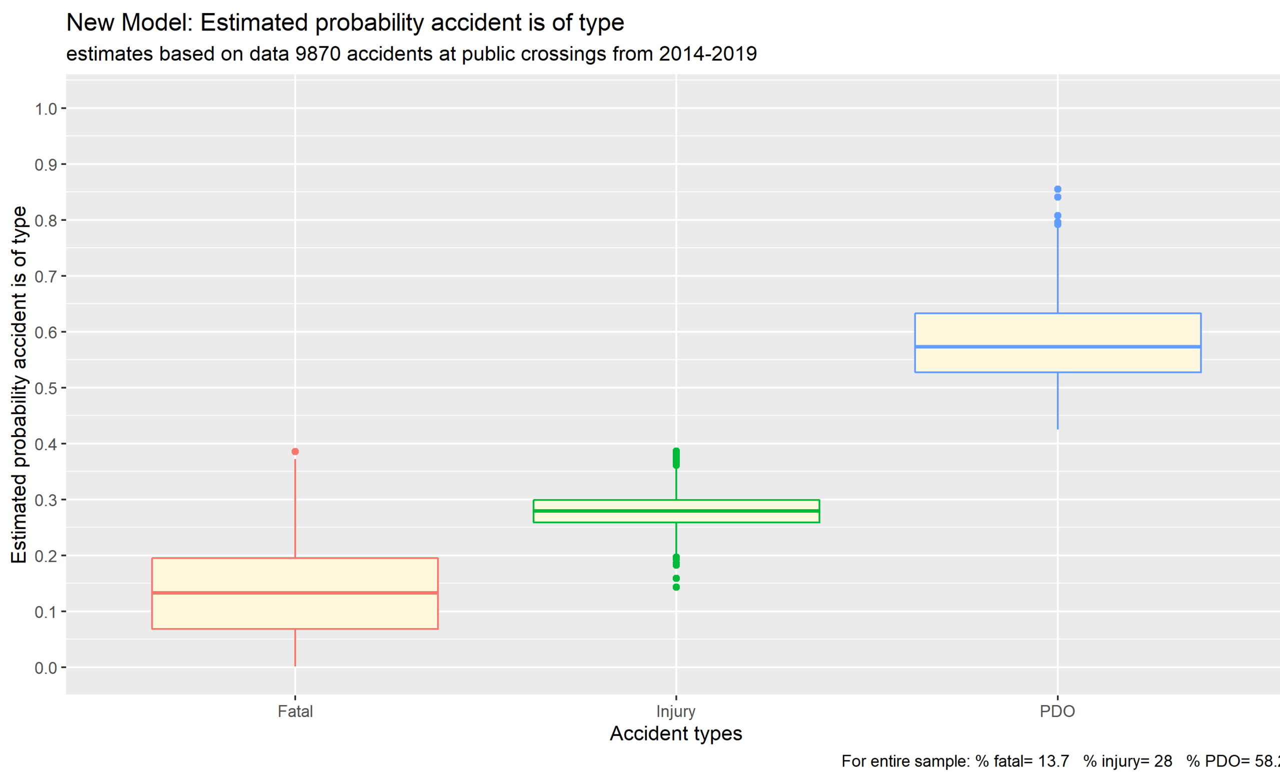
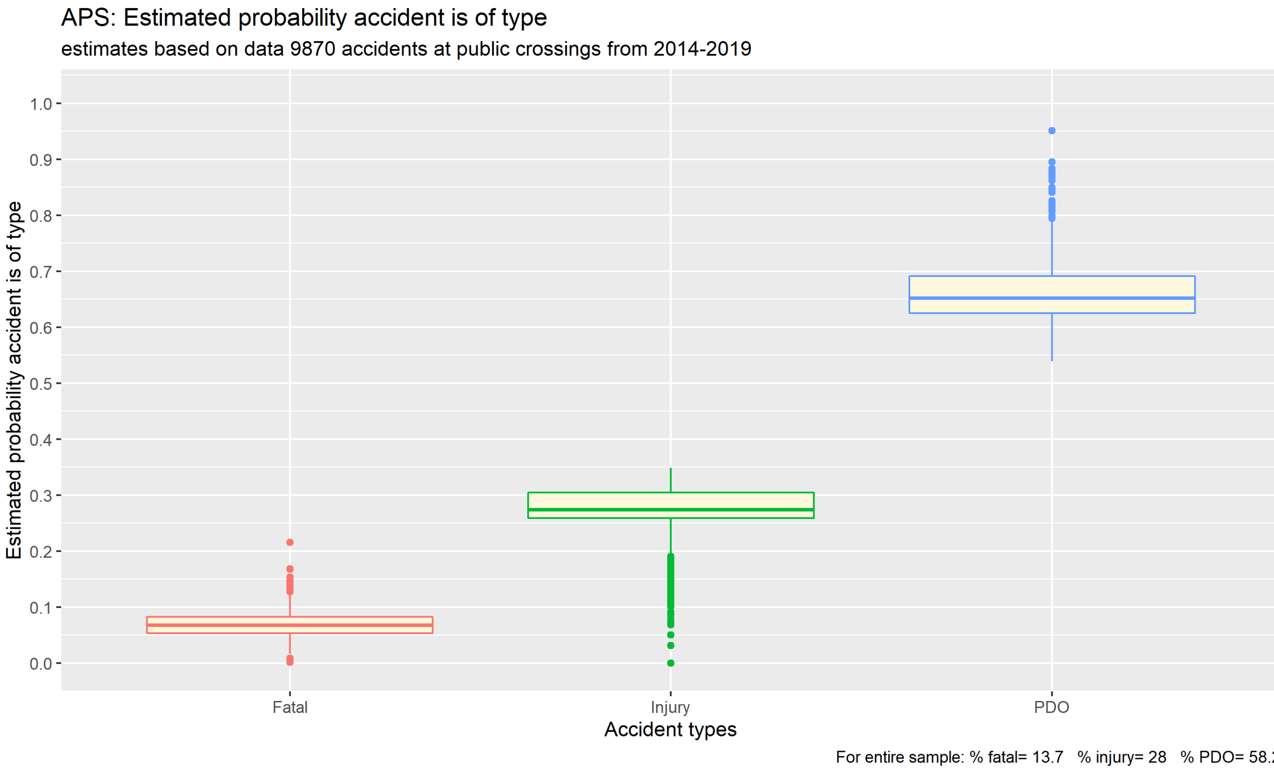


Figure 6-9. Distribution of Predicted Accident Severities with APS



The two charts indicate:

- The mean value from APS for fatal was about half that of the new model, while the means for injury and PDO accidents were similar.
- The new model had a higher variance for the fatal and PDO categories, with smaller variance for the injury category. (Standard deviations were 0.08182, 0.02986, and 0.07511 for fatal, injury, and PDO, respectively.)
- APS had a small variance for fatal, somewhat larger for injury, and a bit larger still for PDO. (Standard deviations were 0.0208, 0.03704, and 0.0455 for fatal, injury, and PDO, respectively.)
- APS had more and more unbalanced outliers. The injury category skewed downward, and the PDO category skewed up. The table below shows a summary of the skewness values:

Table 6-2. Summary of Severity Category Skewness by Model

	New Model	APS
Fatal	0.3132	0.5734
Injury	0.2782	-0.9523
PDO	0.4544	0.6588

7. Conclusion

The preliminary data review indicates that a new model could replace the APS based on the key drivers of exposure and grade crossing warning device type. In other words, the data show that risk increases with exposure and more protective warning device type reduces risk.

Other findings include:

- There is justification for a single model with category of warning device type as a variable rather than separate models for each of the three warning device type categories.
- Grade crossings that are public, not closed, not grade separated, and that have non-missing, non-erroneous values for exposure and warning device type, number 105,377 nationally. In the period 2014–2018 there were 8,467 accidents at these grade crossings.
- An aggregate analysis of these grade crossings showed that relative to a passive crossing, a lights crossing had 73 percent less risk per exposure. A gated crossing had 63 percent less risk per exposure than a lights crossing.
- The findings of the above analysis indicate a functional form with exposure, warning device type, and other grade crossing characteristics.
- The analysis indicates additional variables that are likely to explain accident occurrence: grade crossing is in rural or urban area, maximum timetable speed, and grade crossing surface types.
- Model estimation using ZINB regression yielded parameters of the expected sign and magnitude, and had strong statistical significance.
- Including the number of daily trains and the AADT at the crossing, which are components of the exposure metric, improved the regression results as indicated by the AIC.
- The EB method accounts for accident history while correcting for "regression to the mean" bias. Adjusted results with EB produced predictions that more closely track the actual counts than did the APS adjustment process for accident history.
- The new model severity component determined the probabilities that an accident would be of one of three severity types: fatal, injury or PDO.
- The severity component of the new model was derived using multinomial logistic regression on the accidents in the 6-year period 2014–2019.
- In the period there were 11,131 accidents at public crossings. Of these, the crossings where these accidents occurred had non-missing, non-erroneous data for 9,870 grade crossings. The accidents at these crossings were included in the severity model estimation.
- The multinomial logistic regression showed that the best results were obtained with explanatory variables: rural or urban, maximum time table speed, number of daily trains, and whether a crossing has a lights warning device.

- Validations showed that the new model performed better than the APS by multiple measures.

8. References

1. Farr, E.H. (1987). Summary of the DOT Rail-Highway Crossing Resource Allocation Procedure – Revisited [DOT/FAR/OS-87/05].
2. Hauer, E. (2015). The Art of Regression Modeling in Road Safety. Springer
3. Hauer, E. (2001). Overdispersion in modeling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis and Prevention 33,* 799–808.
4. Hauer, E. (2004). Statistical Safety Modeling. *Transportation Research Record 1897.* Washington, DC: National Academies Press, 81–87.
5. American Association of State Highway and Transportation Officials. (2010). Highway Safety Manual, 1st edition.
6. Brod, D., Weisbrod, G., Moses, S.B., Gillen, D., & Martland, C.D. (2013) Comprehensive Costs of Highway-Rail Grade Crossing Crashes. NCHRP Report #755. Transportation Research Board.
7. Federal Railroad Administration. (2019). GradeDec.Net Reference Manual.
8. Mironer, M., Coltman, M. & McCown, R. (2000). Assessment of Risks for High-Speed Rail Grade Crossings on the Empire Corridor [DOT-Volpe-FRA-00-03]. Washington, DC: U.S. Department of Transportation.
9. Elvik, R. & Vaa, T. (2014). Handbook of Road Safety Measures. Oxford, UK: Elsevier.

Appendix A. Interpreting Regression Outputs

A regression analysis is a set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables. A dataset contains a number of observations for each variable.

The independent variable is often called the left-hand side (LHS) variable because it is written to the left of the equals sign. The dependent variables (also called explanatories) are the right-hand side (RHS) variables.

In regression analysis, the analyst develops a model linking the LHS with RHS variables and "runs" a regression. A statistical program examines the dataset and finds the values of model coefficients that meet optimization criteria.[33]

The regression output table contains general statistics along with coefficient estimates and statistics.

The following describes the columns in the regression output table that relate to the coefficient estimates:

Column Name	**Column Description**
Variable	Each row contains the name of a model variable. If the model has a constant, the row will usually say "constant" or "intercept," depending upon the software used.
Estimate	The estimate of the variable model coefficient (in this report, coefficients are subscripted and shown in model formulas as lowercase Greek letters β (beta) and γ (gamma)
Std. Error	The standard deviation of the coefficient estimate
z-value	This is the estimate divided by the standard error.
Pr(>|z|) (p-value)	In statistical significance testing, the p-value is the largest probability of obtaining test results at least as extreme as the results actually observed, under the assumption that the null hypothesis is correct (i.e., assuming the coefficient is actually 0). This is equivalent to the probability of falsely rejecting the null hypothesis (also called a Type I error).

[33] The two broad classes of regression techniques are least squares (LS) and maximum likelihood estimation (MLE). With LS, the regression minimizes the sum of squared residuals ("residuals" are the differences between the LHS values and the "fitted" (calculated values of the model). With MLE, the regression seeks the point of maximum of a likelihood function that is constructed from all the data observations. The data sets under consideration will usually determine which technique is most appropriate.

Column Name	Column Description
Significance codes	Corresponding to the p-value is the level of significance. For example, if the p-value is between .001 and .01, it will show ‘**’. With ‘**’ it can be said that “the probability of a Type I error is less than 1 percent”, or, “the coefficient is significantly different from 0 with 99 percent confidence.”

The general statistics include descriptive statistics of the regression and its residuals. This study examines the AIC, which enables model quality comparison and whose value is least for the better model specification with the given set of data.

Appendix B. Application of the New Model

The APS enables risk ranking of grade crossings (in a corridor or region). However, it cannot inform when two grade crossings with similar risk scores (e.g., predicted annual accidents) should be treated the same or differently. The new model provides descriptive statistics of the population of grade crossings, and these can be used to determine if scores are close enough to warrant same or different treatment.[34]

For example, suppose we have two grade crossings A and B, and the new model estimates they have predicted annual accidents of 0.21 and 0.26, respectively. From the analysis of data in developing the model, we know that:

1. Mean value of 5-year accidents for the population of grade crossings is E{μ} = 0.08319
2. The variance of 5-year accidents for the population of grade crossings is V{k} = 0.1220627.
3. The standard deviation of 5-year accidents for the population of grade crossings is:

$$\hat{\sigma} = \sqrt{\mathrm{V}\{k\} - \mathrm{E}\{\mu\}} = \sqrt{0.1220627 - 0.08319} = 0.1972562$$

Since the standard deviation is for 5-year accidents, divide by 5 for the standard deviation of predicted annual accidents:

$$\hat{\sigma}_{annualized} = \frac{0.1972562}{5} = 0.03945124$$

Crossing A has predicted annual accidents of 0.21, then adding the standard deviation to the value 0.21 + 0.03945124 = 0.24945124. Crossing B has predicted annual accidents of 0.26, which is greater than the previous value and outside a band of one standard deviation from the mean value of predicted annual accidents of A. We would conclude that the predicted annual accidents of the two crossings differ significantly and, therefore, the two warrant different treatment based on the new model.

[34] Following Hauer (2015) Chapter 2, “A Safety Performance Function for Real Populations.”

Abbreviations and Acronyms

AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
AIC	Akaike Information Criterion (a measure of the relative quality of a model for a given set of data)
APS	Accident Prediction and Severity
CMF	Crash Modification Factor (a safety countermeasure's ability to reduce crashes and crash severity)
CFR	Code of Federal Regulations
CWT	Constant Warning Time (device at grade crossings with active warning devices that ensures the time between initial warning and a train's arrival at the roadway is constant, regardless of the speed of the train)
DOT	Department of Transportation
EB	Empirical Bayes (procedure for statistical inference in which prior distributions are derived from data)
FRA	Federal Railroad Administration
GCIS	Grade Crossing Inventory System
GX	Grade crossing (used in this document's figures)
HSR	High-Speed Rail
MLE	Maximum Likelihood Estimation (a class of model estimation procedures)
MNL	Multinomial Logistic (a regression analysis method)
NB	Negative Binomial (a probability distribution)
PDO	Property Damage Only (a severity type of train-highway vehicle accident at a grade crossing)
SPF	Safety Performance Function (a function for evaluating the safety of a transportation facility, or population of facilities, from a set of facility traits and accident history)
TRB	Transportation Research Board
Volpe	Volpe National Transportation Systems Center
ZINB	Zero-Inflated Negative Binomial (a regression analysis method)