Evaluating Safety Countermeasures at Highway-Railway Grade Crossings: A Review

Mohammadali Zayandehroodi

Barat Mojaradi

Morteza Bagheri

Anastasios Kouvelas

STRC Conference Paper 2025 May 2025

STRC | 25th Swiss Transport Research Conference
Monte Verità / Ascona, May 14-16, 2025

Evaluating Safety Countermeasures at Highway-Railway Grade Crossings: A Review

Mohammadali Zayandehroodi
1-School of Civil Engineering, Iran University of Science and Technology, 2- Institute for Transport Planning and Systems (SVT – Traffic Engineering), ETH University
mzayandehroo@ethz.ch

Barat Mojaradi
School of Civil Engineering, Iran University of Science and Technology

Morteza Bagheri
School of Railway Engineering, Iran University of Science and Technology

Anastasios Kouvelas
Institute for Transport Planning and Systems (SVT – Traffic Engineering), ETH University

May 2025

Abstract

Crash Modification Factors (CMFs) are key metrics used to quantify the effectiveness of safety countermeasures in reducing crashes and prioritizing safety improvements. The primary methods for estimating CMFs include Observational Before-After (BA) studies and the cross-sectional approach, both of which have inherent limitations, such as data requirements. While the BA method accounts for regression-to-the-mean (RTM) bias, the cross-sectional method is often more applicable when BA studies are impractical. Recent advancements in machine learning (ML) have enhanced CMF estimation by providing more accurate statistics. For instance, ML-based clustering can reveal the true impact of countermeasures across different sites, highlighting variations in effectiveness that are masked by traditional methods. In comparison, the BA method tends to underestimate benefits when accounting for changes in traffic. By integrating ML techniques with statistical methods, the augmented approach provides more reliable and precise estimates of countermeasure effectiveness, accounting for heterogeneity across segments. This synthesis of ML and traditional safety analysis methods offers a more comprehensive understanding of safety countermeasures, enabling transportation safety experts to make more informed decisions about which countermeasures are most appropriate for different sites.

Keywords

Crash modification factor 1; Machine Learning 2; Before–After 3

Table of contents

A Introduction 5

B Crash modification factors 5

B.1 Observational before-and-after studies 5

B.2 Cross-sectional method 7

B.3 Propensity Score Method 7

B.4 Machine Learning models 8

C Discussion and conclusions 10

D Reference list 12

A Introduction

The concept of the Crash Modification Factor (CMF) quantifies the safety benefits associated with various countermeasures and depicts the anticipated changes in crashes following their implementation (Davis, 2000). Accurately estimating the CMF requires a robust assessment of the expected number of crashes at a particular site both before and after the implementation of a specific countermeasure (Rudin-Brown et al., 2012). In the last two decades, various approaches have been utilized to predict crashes and gauge the efficacy of countermeasures, with cross-sectional and before-and-after models emerging as the most widely adopted methods (P. Y. Park, 2007). Anticipating further advancements, the upcoming US Highway Safety Manual is expected to provide a comprehensive set of CMFs reflecting the impact of diverse design and operational strategies on highways (AASHTO, 2010).

The literature has put forth a diverse range of statistical approaches for estimating the impact of countermeasures. The next section examines the most commonly employed methods for assessing countermeasures, such as the cross-sectional statistical model and the Before-and-After (BA) method (Al-Marafi & Somasundaraswaran, 2023).

B Crash modification factors

This study aimed to review various methodological approaches for estimating CMFs, offering a comprehensive overview. It examines the key methodologies used in estimating CMFs, highlighting their main advantages and disadvantages. This review serves as an essential precursor to the application of more advanced analytical tools, such as Machine learning (ML) and artificial intelligence, in analyzing road safety outcomes. The structure of the study is as follows: the next section explores the approaches commonly used for implementing observational BA studies.

B.1 Observational before-and-after studies

Observational studies conducted before and after implementing a countermeasure at various sites aim to gauge the change in crash numbers or other risk factors (Gross et al., 2010). These studies address the regression-to-the-mean (RTM) bias, where sites with high crash frequencies in one year tend to decrease in the following year due to the random nature of crashes, regardless of countermeasure or traffic conditions. Several approaches have been employed in these studies, including the, Naïve approach, Empirical Bayes (EB) method, and Full Bayes (FB) method (Yu & Abdel-Aty, 2014; Elvik, 2017; Hauer et al., 2012; Lyon et al., 2020; Shahdah et al., 2014).

B.1.1 Naïve approach

The fundamental assumption of the naïve approach is that crash frequencies before countermeasure implementation will align with expectations (Yu & Abdel-Aty, 2014). In this method, expected crashes are calculated by determining the ratio of road crashes to the number of years before countermeasure, and this ratio is then applied to estimate expected crashes after countermeasure, considering only the post-countermeasure period. However, according to (Gross & Jovanis, 2007), the naïve approach tends to overestimate the countermeasure effect due to the RTM problem. Additionally, the naïve approach inaccurately predicted a total reduction in crashes after a hypothetical countermeasure without any actual effect. This inaccuracy stems from the failure to account for RTM bias in this approach (Fawcett et al., 2017).

B.1.2 Empirical Bayes approach

The EB approach aims to enhance road safety estimation. This method improves estimation precision, addressing the primary limitation of the Comparison Group (CG) and Naïve approaches by incorporating the RTM effect (Y. Park & Saccomanno, 2005; Shen & Gan, 2003). It achieves this by calculating a weighted average of observed and predicted crashes (Hauer et al., 2012).

In general, utilizing an EB before-and-after approach offers two significant advantages for estimating countermeasure effects. Firstly, the EB model helps to mitigate or eliminate much of the RTM bias inherent in the naïve approach, thereby producing more dependable estimates of CMFs. Secondly, the EB approach takes into consideration changes in traffic volume during the before-and-after study period. These variations in traffic volume serve as a surrogate variable, representing both observed and unobserved factors at study sites, such as unrecorded weather conditions. The EB before-and-after approach has been widely embraced as a primary safety evaluation tool by numerous US and Canadian agencies, including the Interactive Highway Safety Design Model and the Highway Safety Manual (D.W. Harwood, F.M. Council, E. Hauer, W.E. Hughes, 2000).

Despite their relative success in addressing RTM bias, EB before-after models still possess certain limitations, which are outlined below:

1) EB before-after models necessitate extensive data, such as yearly-based exposure data over a specified study period, which escalates the cost and duration of analyses. However, researchers often lack the resources to gather the requisite inputs, leading to incomplete results. For instance, databases for sites in Canada and the US lack yearly-based exposures like traffic volumes, posing challenges for EB approach utilization.

2) EB before-after models rely on the assumption that changes in traffic volume between before and after periods account for the effects of all unobserved factors. However, other factors beyond traffic volumes, such as reporting biases, may independently or in combination affect crash frequency. Although the EB before-after analysis itself does not mandate this assumption, most researchers predominantly focus on traffic volume as the primary input factor.

3) EB before-after models typically examine only one countermeasure at a time, neglecting the simultaneous estimation of multiple countermeasure effects.

4) EB before-after models yield average effects rather than tailored effects for specific sites, such as sites. This poses challenges for decision-makers addressing site-specific issues, as the average effectiveness of a countermeasure might not suffice. Evaluating individual

countermeasures for local application via before-after models requires substantial time and resources, rendering it (Al-Marafi & Somasundaraswaran, 2023).

B.1.3 Full Bayes approach

The FB approach, similar to the EB method, utilizes non-treated reference sites to draw inferences and account for potential influences unrelated to the countermeasure. (El-Basyouny & Sayed, 2006) highlighted a key distinction between the FB and EB approaches: in the FB approach, predicted crash frequencies without countermeasures are derived from the Crash Prediction Model (CPM) estimated using data from both the pre-countermeasure period of treated sites and reference sites, while the EB approach utilizes data solely from reference sites to estimate the CPM. Recent recommendations by researchers suggest employing the FB approach for evaluating the impact of safety countermeasures (El-Basyouny & Sayed, 2006; Sayed et al., 2016). This approach offers several advantages over others, including the ability to account for all data uncertainties, requiring less data, allowing more flexibility in selecting crash frequency distributions, providing more detailed causal inferences, and considering spatial correlation effects among sites in the model formulation. Sacchi and Sayed (2015) compared the results of naïve, EB, and FB approaches in estimating countermeasure effectiveness using two methods for selecting hypothetical countermeasure sites: random selection to mitigate selection bias and non-random selection based on abnormal crash frequency (black spots). For randomly selected sites, all approaches yielded reasonable results. However, for non-randomly selected sites, the FB approach demonstrated superior performance compared to the naïve and EB approaches. It's important to note that the complexity of the FB approach may render the EB approach more appealing to researchers for practical implementation (Shahdah et al., 2014).

B.2 Cross-sectional method

Although observational before-after (BA) studies are typically considered the preferred approach for estimating CMFs, practical constraints are associated with their implementation. These constraints include the need for accurate knowledge of the countermeasure date to delineate the before and after periods and the requirement for several years of post-countermeasure data collection. Additionally, assessing safety effects becomes complex when multiple countermeasures are applied at a single site (Hauer, 1983; Al-Marafi & Somasundaraswaran, 2023). In such cases, the cross-sectional method offers an alternative for estimating CMFs due to its simpler data acquisition process compared to observational BA studies.

The cross-sectional method, also known as Safety Performance Function (SPF), establishes a relationship between crash occurrences and the geometric characteristics and traffic volumes of a roadway (Rahman et al., 2020).

B.3 Propensity Score Method

Past researchers have frequently treated countermeasures as exogenous variables in their modeling approaches. However, some scholars, including Kim and Washington (2006), have sought to tackle countermeasure selection bias by introducing the concept of endogeneity. They aim to address endogenous relationships to gain a better understanding of the true effects of various countermeasures on crashes. Drawing from crash analysis results using highway intersection data from Georgia, they argued that the inconsistent results often seen in countermeasure evaluations stem from a lack of control for potential endogeneity issues between crash rates and countermeasures (Kim & Washington, 2006).

In the realm of sites, Austin and Carson (2002) acknowledged that the presence of warning devices, like flashing lights or gates, could be endogenous to crash rates. This implies that crash rates are frequently used as a basis for installing warning devices such as flashing lights or gates at sites. To address countermeasure selection bias effectively, it is essential to understand the criteria used to select specific sites, such as sites, for various types of countermeasures. However, traditional before-and-after and cross-sectional models fail to directly investigate why particular sites are chosen for countermeasures. Consequently, their estimations are influenced by the non-randomness inherent in the countermeasure selection process. One significant issue in this regard is known as Regression-To-the-Mean (RTM) bias, stemming from the tendency to apply countermeasures at sites with a history of high crash rates. Countermeasure selection bias can still persist if there is a systematic bias in how sites are chosen for improvements. For example, if all selected sites exhibit train speeds above a certain threshold or have a minimum number of tracks, these criteria could introduce a systematic selection bias (Austin & Carson, 2002).

B.4 Machine Learning models

Recent advances in ML offer new opportunities to improve the reliability of safety evaluations at Highway Railway Grade Crossings (HRGCs). ML models excel at identifying complex patterns and interactions in large datasets, which can complement and improve traditional statistical approaches. Zayandehroodi’s studies introduce innovative ML-based frameworks for safety analysis, aiming to overcome the limitations of statistical methods. The first study (Zayandehroodi et al., 2024) proposed a ML framework for evaluating countermeasures by integrating data-driven techniques into each step of the analysis. It used XGBoost (extreme gradient boosting) algorithm to first identify the most influential factors associated with crash severity and frequency. Next, it applied a Deep Latent Class Analysis (DLCA) clustering method to group similar sites based on those important features. By clustering sites into more homogeneous categories, the study could account for heterogeneity across HRGCs. Finally, within each cluster, the study computed CMFs and their standard errors for various countermeasures (such as gates, lights, bells, and crossbucks) by analyzing the crash data in that cluster. The second study (Zayandehroodi et al., 2025) builds upon this framework by introducing a hybrid Negative Binomial–LSTM model to improve crash prediction and explicitly model uncertainty. This approach combines the statistical limitation of a Negative Binomial (NB) regression with the sequence-learning capability of a Long Short-Term Memory (LSTM) neural network, implemented within the EB context. In essence, the hybrid NB-LSTM model uses historical crash time-series data to capture temporal dependencies (e.g. trends or seasonality in crash occurrence) that a static NB model might miss, while still outputting predictions in a form compatible with the EB method. By better predicting the expected crashes in HRGCs (including quantifying the prediction uncertainty), the EB-adjusted CMFs can be estimated with greater accuracy and stability.

Table 1 furnishes a compilation of methods employed for estimating CMF, accompanied by their respective advantages and disadvantages.

Table 1. Summary of Methodologies to Estimate CMF

Method	**Advantages**	**Disadvantages**
Naïve approach	Ease of application	Does not consider RTM bias; may overestimate the countermeasure effect; lacks control over the effects of external causal factors.
Empirical Bayes approach	Addressing RTM bias; does not require a large number of reference sites.	Challenging to gather adequate data. Yields more precise estimates compared to naive comparison method.
Full Bayes approach	Addressing RTM bias; capability to incorporate all uncertainties in the data; does not necessitate a large number of reference sites; able to accommodate temporal and spatial variations.	Challenges in application; collecting suitable data is demanding. Serves as a complex alternative to the EB approach.
Cross-sectional method	Addressing RTM bias; considers variations in countermeasure effectiveness over time.	Does not consider the effects of unaccounted elements; necessitates a sufficient sample size, particularly when large explanatory variables are integrated into the developed model.

Propensity Score Method	- Addresses systematic selection bias - Includes multiple countermeasure selection criteria	-Assesses the impact of individual countermeasures separately - Offers only the average effect of countermeasures
Machine Learning models	-ML models can provide more accurate predictions by learning complex relationships in the data that traditional methods may miss. -ML models can process and analyze large amounts of data efficiently, allowing for better insights and predictions. -ML models can incorporate various types of data (e.g., traffic, environmental conditions, countermeasure types) and learn from them to improve the prediction of crash reductions.	- The "black-box" nature of these models makes it challenging to understand how specific inputs lead to certain predictions.

C Discussion and conclusions

Various methods for estimating CMFs each offer distinct advantages and disadvantages when applied in transportation safety analysis. The methods range from simple approaches like the Naïve approach to more sophisticated techniques such as ML models. Understanding the trade-offs between these methods is essential for selecting the most appropriate approach based on available data, research goals, and the need for precision.

The Naïve approach is widely recognized for its simplicity and ease of application, making it a good starting point for many studies. However, it has significant drawbacks, including its failure to account for RTM bias and its tendency to overestimate the effect of countermeasures.

Additionally, it lacks the ability to control for external factors that may influence the results, reducing its reliability in complex settings.

In contrast, the EB approach addresses RTM bias and provides more precise estimates compared to the Naïve method. One of its primary advantages is that it does not require a large number of reference sites, making it more feasible in certain scenarios. However, collecting sufficient data for EB analysis can be challenging, as it still requires accurate and extensive information to generate reliable results.

The Full Bayes approach expands on EB by incorporating uncertainties in the data, thus offering a more comprehensive understanding of the effects of countermeasures. This method can accommodate variations over time and space, which is particularly valuable for dynamic environments. However, it is more complex to apply and requires demanding data collection processes, making it less accessible for some studies. Despite these challenges, it provides a more flexible and robust alternative to the EB method.

The Cross-sectional method also accounts for RTM bias and includes the ability to assess how countermeasure effectiveness may vary over time. However, it struggles with unaccounted external factors and often necessitates a large sample size, especially when integrating many explanatory variables into the model. This requirement can make the method less practical in smaller or less data-rich studies.

The Propensity Score Method is effective at addressing systematic selection bias and allows for the inclusion of multiple countermeasure criteria. However, it has limitations, including its focus on the average effect of countermeasures rather than assessing individual countermeasures. This reduces its ability to capture the full range of potential variations across different contexts.

Finally, ML models offer several key advantages, including the ability to handle large datasets, uncover complex relationships in the data, and incorporate diverse types of input (e.g., traffic, environmental factors, and countermeasure types). These capabilities make ML models particularly powerful in predicting crash reductions and offering more accurate estimates than traditional methods. However, their primary disadvantage lies in the "black-box" nature of many ML algorithms, such as deep learning. This lack of interpretability can make it difficult to understand the rationale behind specific predictions, limiting the practical application of these models in safety-critical scenarios where transparency is essential.

D Reference list

AASHTO. (2010). An introduction to the Highway Safety Manual. In *Washington D.C.: Federal Highway Administration*. http://www.highwaysafetymanual.org/

Al-Marafi, M. N., & Somasundaraswaran, K. (2023). A review of the state-of-the-art methods in estimating crash modification factor (CMF). *Transportation Research Interdisciplinary Perspectives*, *20*(May). https://doi.org/10.1016/j.trip.2023.100839

Austin, R. D., & Carson, J. L. (2002). An alternative accident prediction model for highway-rail interfaces. *Accident Analysis and Prevention*, *34*(1), 31–42. https://doi.org/10.1016/S0001-4575(00)00100-7

D.W. Harwood, F.M. Council, E. Hauer, W.E. Hughes, and A. V. (2000). Prediction of the Expected Safety Performance of Rural Two-Lane Highways. In *US Department of Transportation* (Issue December).

Davis, G. A. (2000). Accident reduction factors and causal inference in traffic safety studies: A review. *Accident Analysis and Prevention*, *32*(1), 95–109. https://doi.org/10.1016/S0001-4575(99)00050-0

El-Basyouny, K., & Sayed, T. (2006). Comparison of two negative binomial regression techniques in developing accident prediction models. *Transportation Research Record*, *1950*, 9–16. https://doi.org/10.3141/1950-02

Elvik, R. (2017). Road safety effects of roundabouts: A meta-analysis. *Accident Analysis and Prevention*, *99*, 364–371. https://doi.org/10.1016/j.aap.2016.12.018

Fawcett, L., Thorpe, N., Matthews, J., & Kremer, K. (2017). A novel Bayesian hierarchical model for road safety hotspot prediction. *Accident Analysis and Prevention*, *99*, 262–271. https://doi.org/10.1016/j.aap.2016.11.021

Gross, F., & Jovanis, P. P. (2007). Estimation of the safety effectiveness of lane and shoulder width: Case-control approach. *Journal of Transportation Engineering*, *133*(6), 362–369. https://doi.org/10.1061/(ASCE)0733-947X(2007)133:6(362)

Hauer, E. (1983). An application of the likelihood/bayes approach to the estimation of safety countermeasure effectiveness. *Accident Analysis and Prevention*, *15*(4), 287–298. https://doi.org/10.1016/0001-4575(83)90053-2

Hauer, E., Bonneson, J., Council, F., Srinivasan, R., & Zegeer, C. (2012). Crash modification factors. *Transportation Research Record*, *2279*, 67–74. https://doi.org/10.3141/2279-08

Kim, D. G., & Washington, S. (2006). The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis and Prevention*, *38*(6), 1094–1100. https://doi.org/10.1016/j.aap.2006.04.017

Lyon, C., Persaud, B., Merritt, D., & Cheung, J. (2020). Empirical bayes before-after study to develop crash modification factors and functions for high friction surface treatments on curves and ramps. *Transportation Research Record*, *2674*(12), 505–514.

https://doi.org/10.1177/0361198120957327

Park, P. Y. (2007). *Estimating Effectiveness of Countermeasures Based on Multiple Sources : Application to Highway-Railway Grade Crossings By*. University of Waterloo.

Park, Y., & Saccomanno, F. F. (2005). Evaluating Factors Affecting Safety at Highway–Railway Grade Crossings. *Transportation Research Record*, *1*, 1–9.

Rahman, M. A., Sun, X., & Das, S. (2020). Reconfiguring Urban Undivided Four-Lane Highways to Five-Lane: A Nonideal but Very Effective Solution for Crash Reduction. *Journal of Transportation Engineering, Part A: Systems*, *146*(10), 04020116. https://doi.org/10.1061/jtepbs.0000422

Rudin-Brown, C. M., Lenné, M. G., Edquist, J., & Navarro, J. (2012). Effectiveness of traffic light vs. boom barrier controls at road-rail level crossings: A simulator study. *Accident Analysis and Prevention*, *45*, 187–194. https://doi.org/10.1016/j.aap.2011.06.019

Sayed, T., Sacchi, E., & DeLeur, P. (2016). Evaluating safety benefits of the Insurance Corporation of British Columbia road improvement program using a full Bayes approach. *Transportation Research Record*, *2582*(2582), 26–33. https://doi.org/10.3141/2582-04

Shahdah, U., Saccomanno, F., & Persaud, B. (2014). Integrated traffic conflict model for estimating crash modification factors. *Accident Analysis and Prevention*, *71*, 228–235. https://doi.org/10.1016/j.aap.2014.05.019

Shen, J., & Gan, A. (2003). *Development of Crash Reduction Factors Methods , Problems , and Research Needs*. *03*, 50–56.

Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science*, *63*, 50–56. https://doi.org/10.1016/j.ssci.2013.10.012

Zayandehroodi, M., Mojaradi, B., & Bagheri, M. (2024). Evaluating the effectiveness of safety countermeasures at highway–railway grade crossing based on a machine learning framework. *Traffic Injury Prevention*, 1–8. https://doi.org/10.1080/15389588.2024.2387713

Zayandehroodi, M., Mojaradi, B., & Bagheri, M. (2025). Improving reliability of safety countermeasure evaluation at highway-rail grade crossings through aleatoric uncertainty modeling with machine learning techniques. *Reliability Engineering and System Safety*, *261*(November 2024), 111082. https://doi.org/10.1016/j.ress.2025.111082