Road Safety Assessment in Heterogeneous and Non-Lane-Based Traffic Using Surrogate Safety Measures and Statistical Modeling

Thesis submitted in partial fulfilment
for the Award of

DOCTOR OF PHILOSOPHY (PHD)

in

CIVIL ENGINEERING

by

ASHUTOSH KUMAR

DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY (BHU)
VARANASI - 221005, INDIA

ROLL NUMBER
18061005

YEAR OF SUBMISSION
2024

Chapter 8

Conclusions

8.1 Preface

Road traffic crashes have emerged as a major health issue that has been a concern all over the globe. To address this issue, researchers have successfully developed methods and tools to quantify and improve road safety. There are crash-based methods that require observing the total number of crashes occurring within a given time period. However, reliable crash data is not available, especially in developing countries. To this end, researchers have developed non-crash-based safety assessment methods to quantify and subsequently improve road safety. Non crash-based methods are also desirable since these do not require waiting for several years for crashes to happen. Among these methods are surrogate safety measures (SSMs) that have been widely accepted for proactive safety assessment. SSM could be a promising tool for quantifying and improving road safety in developing countries.

Traffic in most developing countries is heterogeneous in nature. The two key aspects that contrast heterogeneous traffic with homogeneous traffic are vehicular heterogeneity and non-lane-based movements that lead to two-dimensional interactions. Researchers have used SSM in these conditions but without accounting for vehicular heterogeneity and two-dimensional vehicular interaction. Most of these studies utilize conflict indicators that are defined in homogeneous traffic conditions where vehicular interactions are one-dimensional. They have also assumed that conflicting vehicles have their trajectory overlap at least at a point. However, in non-lane-based traffic, it has been shown that vehicles move freely without maintaining any lane discipline. Consequently, the assumptions of overlapping trajectories would not hold true for such traffic conditions. Therefore, using one-dimensional conflict indicators is inadequate in defining conflict in non-lane-based traffic conditions. In addition, previous studies usually ignored vehicular heterogeneity in crash risk assessment. Therefore, for more accurate risk assessment it is essential to incorporate non-lane-based two-dimensional vehicular interactions and vehicular heterogeneity in defining traffic conflict. This study aims to address these gaps. To this end, video data were collected at four uncontrolled intersections on four-lane divided highways in India. Conflict indicators were estimated from vehicle trajectories obtained from those video recordings. The selected sites were high-crash locations with heterogeneous and non-lane-based traffic conditions. Therefore, it was possible to capture interaction among different vehicle pairs, making them suitable for this study.

This chapter summarizes the findings of the dissertation and provides future research directions. The main findings and their applications are discussed in the next section. Further, the contributions made by this dissertation is presented. Finally, the limitations and extension of the current work are provided.

8.2 Research Findings and Conclusions

8.2.1 Defining conflict in non-lane-based traffic

The conflict indicators defined for one-dimensional interactions may not be suitable for non-lane-based traffic where vehicular interactions are two-dimensional (longitudinal and lateral). This study proposes a methodology to define conflicts considering two-dimensional vehicle interactions. A bivariate extreme value approach was used to define conflict in two-dimensional vehicular interaction using Time-to-collision (TTC) and lateral gap. The results show that incorporating lateral and longitudinal conflict indicators into the bivariate extreme value models can significantly improve conflict-based risk assessment. In this study, the site-based and joint-site models were used to compare crash risk. The confidence intervals for crashes at four sites were estimated to be [9.2, 26], [18.6, 40.4], [9.2, 26], and [28.6, 54.4], respectively. The confidence interval for joint-site was found to be [81.4, 122.7]. For the joint-site model, the crash estimate was close to the observed crashes, and the majority of estimated crash confidence interval overlapped with the Poisson confidence interval of observed crashes. This outcome supports the applicability of the bivariate GPD model for defining conflict and crash risk in non-lane-based traffic conditions. Further, the site-based estimates overestimate the crashes. The possible reason for this deviation could be the short duration of the observation period. The uncertainty in crash estimates gets amplified by the large difference in conflict and crash observation period. Further, a separate site-based model tends to overestimate the total number of crashes each year. The estimated range is between 244 and 532 crashes at each site. The 95% confidence interval for the joint-site model is much smaller than the separate site-based model, suggesting a more precise estimate. This result suggests that by increasing the sample size through complete pooling the prediction capability of the fitted model improves significantly. The accuracy of the joint-site model is better than site-based models, and it is expected to improve further when data from a sufficiently large sample size is available.

8.2.2 Examining the effect of vehicle size on crash risk

Previous literature suggests that minimum space and time headway varies with leader-follower types. Most common conflict indicators (like TTC, PET, and MTTC) are derived using these microscopic traffic parameters. The computation of TTC depends on the space headway between the interacting vehicles. Since, space headway is a function of vehicle type, the effect of vehicle type must be considered in estimating conflict based on TTC. Previous studies have used a global threshold for these indicators without accounting for the vehicle sizes. Inappropriately selecting the threshold would likely lead to a biased estimation of conflict and hence safety.

This study examines the effect of vehicle sizes on crash risk associated with rear-end and side-swipe crashes. The Bivariate EVT model (proposed in Chapter 5) was utilized to define conflict in two-dimensional vehicular interaction using Time-to-collision (TTC) and lateral gap. Interacting vehicle pairs were divided into nine groups based on the size of the leader and the follower (L-F). The L-F-based models as well as the global model were fitted using bivariate Generalized Pareto Distribution (GPD). The threshold for conflict indicators was estimated for each L-F model separately.

Minimum TTC as well as lateral gap were found to depend on the size of the L-F pair. Therefore, using global thresholds for identifying conflicts is not appropriate in traffic streams that consists of multiple vehicle types. TTC was lowest for lighter vehicle interactions. Since lighter vehicles have shorter stopping distances due to their small size and lower speed, they tend to maintain shorter gaps which results in smaller TTC. Also, the minimum lateral gap was found to be the lowest for MV-MV interactions, as those vehicles tend to follow the inline car-following scenario. Although the proportion of interaction among lighter vehicles was highest, the estimated probability of conflicts as well as crashes involving LV-LV pairs was low. This may be attributed to lower speed and smaller braking distance of lighter vehicles. Further, MVs move at a relatively higher speed and maintain smaller lateral gaps than others, the probability of crash as well as conflict is high among them. Since lighter vehicles are able to maintain relatively smaller gaps (TTC and lateral gap) because of their small size and shorter stopping distance. Therefore, conflict analysis based on a global threshold would be biased since such normal interactions would be identified as critical even though they are safe.

In addition, the L-F-based model incorporates the influence of vehicle types on crash risk. It highlights that the probability of conflicts and crashes depends on the interaction between different vehicle pairs and helps identify high-risk vehicle groups. For instance, although the probability of LV-LV being involved in a conflict and crash is the lowest, these pairs account for the highest number of conflicts due to their significantly higher traffic share (53.9%). On the other hand, MV-MV pairs are identified as the most critical, as they exhibit the highest probability of conflict in a given interaction. This study highlights the importance of considering vehicle types and two-dimensional interactions when evaluating road safety.

8.2.3 Incorporating heterogeneity using Bayesian hierarchical modeling framework

Vehicular heterogeneity is an important aspect of heterogeneous traffic that prevails in low and middle-income countries. Further, researchers have utilized data from multiple sites to get a representative sample. Crash risk varied across sites. Previous studies have not incorporated these heterogeneities into the model framework. Instead, a pooled modeling approach was adopted in most studies. In pooled models, the population is treated as homogeneous and a universal model is assumed for all subgroups in the population. This study proposes a hierarchical model framework for incorporating vehicular and site-based heterogeneity in risk assessment of rear-end crashes.

The finding from this study suggests that crash risk is different among each vehicle group. Therefore, incorporating the heterogeneity in crash risk estimation will lead to more accurate crash risk assessment and hence mitigation strategies. The modified time to collision (MTTC) depends on the L-F pairs. Particularly, MTTC was found to be relatively small for interactions involving two and three-wheelers and large for cars and LCVs. This implied that vehicle type must be considered in modeling crash risk, especially in heterogeneous traffic scenarios. Further, both pooled and hierarchical models were fitted to compare the crash risk among subgroups. If there is heterogeneity (LF-based model) in the population, the pooled model leads to a biased estimate. In contrast, for the homogeneous case (site-based model), the hierarchical and pooled model leads to comparable crash risks. Incorporating vehicular heterogeneity into the Bayesian hierarchical model can help identify critical vehicle groups. Findings from this study revealed that interactions involving cars or LCVs with other vehicles are riskier than others. A possible reason is that these vehicles maintain a higher speed than others, resulting in higher speed differences. Safety measures like enforced speed limits and speed calming devices (rumble strips) may be deployed at these intersections to reduce the speed difference hence conflict and crash risk due to cars and LCVs.

8.3 Thesis Contributions

The novelty of this study is in defining conflict and crash risk considering two-dimensional vehicular interactions in mixed traffic environments. This research has proposed a novel methodology for defining conflict in traffic streams that do not follow lane discipline. Further, the effect of vehicle sizes on crash risk was also examined. Unlike previous studies, the crash risk was found to depend on L-F pairs. The findings of this dissertation can be utilized for calibrating vehicle warning systems in these scenarios. Further, this research has outlined a Bayesian hierarchical framework to incorporate vehicular and site-based heterogeneity into the model. The proposed methodology can be utilized for more accurate risk assessment in heterogeneous traffic environments.

8.4 Applications and Practical Implications of the Research

This section briefly highlights the practical implications derived from the findings of this thesis.

For defining the most common conflict indicators such as PET, TTC, MTTC, DRAC and TA, it is assumed that the trajectories of conflicting vehicles overlap at one or more points. These indicators are more suitable for defining conflict in lane-based traffic where trajectories overlap more often. However, traffic in most developing countries (e.g., China and India) is characterized by heterogeneous vehicular mix and lane-free traffic. In such traffic streams vehicle interaction is 2-dimensional characterized by longitudinal and lateral interactions. Therefore, to define conflict in such traffic stream, it is important to consider lateral as well as longitudinal movements. The methodology proposed for defining traffic conflict in non-lane-based traffic scenarios considering longitudinal and lateral vehicle interactions represent a significant advancement toward non-crash-based safety assessment in non-lane-based traffic. The findings from this study show that incorporating lateral and longitudinal indicators together into the bivariate models can significantly improve the conflict-based risk assessment in these traffic conditions.

Vehicular heterogeneity is another important aspect of traffic stream in developing countries. The effect of vehicle type on crash risk is investigated in this thesis. The finding reveals vehicle type dependent crash risk. Minimum TTC as well as lateral gap were found to depend on size of LF pair. Therefore, using global thresholds for conflict segregation is not appropriate in traffic streams consisting of multiple vehicle types. This study suggests to consider different thresholds for different vehicle type combinations rather than using global thresholds. Modern driving-assistance technologies, such as advanced driver assistance systems (ADAS) are complained for false warning in heterogeneous traffic conditions. The existing systems provide alerts for potential risks or imminent collisions based on a fixed headway irrespective of vehicle type. However, the findings of this thesis reveal that threshold for critical interactions depends upon LF type. Therefore, utilizing LF based threshold for such warning systems could enhance road safety by providing accurate warnings to drivers. Further, this research highlights the need to extend traditional collision avoidance models—which primarily account for longitudinal dynamics—to include lateral traffic dynamics for evaluating safety in non-lane-based traffic conditions.

Further, this study proposed a Bayesian hierarchical EVT model to account for vehicular heterogeneity in crash risk assessment. A completely pooled model is shown to produce biased estimates of crash risk in heterogeneous populations, whereas separate models for individual LF combinations suffer from limited sample sizes. The findings reveal that while the pooled model is inappropriate for heterogeneous populations, the hierarchical model provides a more robust estimate. For homogeneous cases, however, both hierarchical and pooled models yield comparable results. The methodology proposed in this study can also be extended to incorporate roadway geometry and driver behavior-based heterogeneity, enabling more comprehensive crash risk assessments.

8.5 Limitations and Direction for Future Research

This study presented a novel methodology for non-crash-based safety assessment in the non-lane-based traffic stream. There are several limitations, which could be explored in future research and are presented below:

The proposed bivariate model that defines conflict and crash risk could improve non-crash-based safety assessment in two-dimensional vehicular interactions. However, literature suggests that proximity indicators may be used along with evasive action-based conflict indicators to define conflict frequency and severity. This study utilizes a bivariate EVT model to combine two proximity indicators. In future studies, proximity and evasive action-based conflict indicators may be integrated into a multivariate EVT model for better assessment of conflict frequency and severity.

The definition of conflict in the present study assumes two vehicle interactions only. However, traffic interactions in non-lane-based traffic often involve multiple vehicles. Therefore, defining conflict in multivehicle interactions could be another aspect of future research.

The hierarchical model framework is utilized to incorporate vehicular and site-based heterogeneity. The proposed model framework could be further extended to incorporate roadway related (facility type such as signalized, unsignalized intersection, and mid-block section or road type such as divided and undivided highway, rural and urban road) and driver related (gender, age and driving experience) heterogeneity in crash risk assessment.

Tracking vehicles and capturing conflict indicators is challenging in developing countries because of the presence of multiple vehicle types (both motorized and non-motorized) and non-lane-based driving behavior in the traffic stream. In the present study, vehicle identification and tracking were performed semi-manually. This approach was time-consuming and error-prone due to miss-clicks. Automated trajectory extraction tools developed for heterogeneous traffic may alleviate these shortcomings. Automated conflict detection systems can be developed using artificial intelligence and machine learning technologies. These systems could provide real-time application of the safety assessment methodology presented in this dissertation.

Validation of predicted crashes using SSM based EVT model requires detailed (frequency, severity) crash data. The data available for this study only included fatal crashes. Therefore, relative crash risk among different vehicle groups and sites was investigated using the proposed models. Future research may use detailed crash data when available to validate the proposed model and predict future crashes.

This study utilized Extreme Value Theory (EVT) for crash risk assessment. Other lesser explored statistical models such as survival analysis models to capture time-dependent risk, copula model for multivariate modeling, and application of machine learning models for predicting conflict and crash risk could be implemented in future research.