Chapter 6

Effect of Vehicle Size on the Crash Risk

6.1 Preface

A heterogeneous traffic stream contains vehicles with significantly different static (size, weight) and dynamic (maneuverability) characteristics. Depending on their sizes, vehicles maintain different longitudinal and lateral gaps. In the literature, the effect of vehicle size has been investigated by many researchers while modeling time and space headway. This chapter examines the effect of vehicle size on crash risk in heterogeneous traffic conditions. Bivariate extreme value models were utilized to define conflict and crash risk, as described in Section 3.2. The leader-follower (L-F) based model, as well as global models (not accounting for vehicle type/size), were fitted using the threshold excess approach. Crash risk estimated from both modeling approaches was compared. This study fulfills the objective 2 of this dissertation.

The chapter is organized as follows: Section 6.1 presents the review of relevant literature on crash risk assessment in heterogeneous traffic conditions. Section 6.2 presents the descriptive statistics of conflict data. Section 6.3 presents the bivariate EVT model results. Section 6.4 presents the model estimates and a discussion of the results. Section 6.5 presents the summary of significant findings.

6.2 Crash risk assessment in heterogeneous traffic

This section presents the review of literature on surrogate safety examining the effect of vehicle size in non-lane-based traffic. Weng et al. [150] studied the rear-end crash risk of truck and car traffic in the work zone area, and found that with increase in heavy vehicles, crash risk increases. Also due to better braking capability, cars were found to be at lower risk of being involved in rear end crashes as compared to trucks. Wong and Liao [152] considered three types of vehicles in heterogeneous traffic and identified conflict based on a global TTC threshold of 2 seconds for all vehicle pairs. Although this study considered vehicle type-based conflict counts, it did not incorporate vehicle heterogeneity to define threshold values. In a similar study, Goyani et al. [89] examined the crossing conflict among various vehicle types on unsignalized T-intersections. They used a global PET threshold of 1 s for conflict identification and reported that lighter vehicles are most vulnerable to conflicts and hence crashes. However, lighter vehicles (which are generally small) can safely maintain a smaller gap due to lower speed as well as better braking capability [143]. Therefore, if a global threshold for conflict identification is used, they would appear to be more vulnerable. Kathuria and Vedagiri [217] studied pedestrian and vehicle interactions considering different types of vehicles. Although they found that PET and TTC values were generally lower for interactions between pedestrians and light vehicle types (motorized two and three wheelers), they nevertheless used a single threshold value for identifying conflicting interactions. Charly and Mathew [114] considered negative lateral gap and positive longitudinal gap to identify interacting vehicles in heterogeneous traffic and used 1 sec as threshold value for MTTC (modified time to collision) to identify conflicts. Although they considered lateral gaps for identifying vehicle interactions, they did not explore the effect of different vehicle types while defining conflicts. Zhao and Lee [153] found that distribution of TTC and PET depends on the type of leader-follower (L-F) vehicles. Das et al. [143] found that lead vehicle type and centreline separation are important factors for headway selection in heterogeneous traffic scenarios. The authors suggested that lateral spacing along with longitudinal spacing may be used for defining conflict indicators such as TTC, PET. Das and Maurya [43] studied the effect of centreline separation (CS) on TTC for cars as followers in heterogeneous traffic conditions. They found that TTC decreases with the increase in CS and size of leaders. This indicates that collision risk in heterogeneous traffic varies with CS and size of lead vehicle. They were able to study the effect of lateral separation using CS. However, using a single threshold value of CS may lead to selection bias in identifying vehicle interactions in heterogeneous traffic streams. This is because vehicles with larger width would have higher CS as compared to vehicles with smaller width. Also, they considered only cars as follower vehicle types and did not explore the effect of different vehicle combinations on the TTC. Wang et al. [57] reviewed various SSM and their applicability in defining conflicts in connected and automated vehicle environment. They suggested to incorporate the lateral along with longitudinal parameters while defining conflicts in lane changing and merging interactions. Paul and Ghosh [119] proposed a conflict severity index for safety assessment in heterogeneous traffic conditions. They did not consider the effect of vehicular heterogeneity while selecting threshold for conflict identification, and used a global threshold of 1.5 s. Wang et al. [160] studied the effect of vehicle heterogeneity on surrogate safety measures estimated in car-following scenarios using vehicle trajectory data. They compared various measures (distance and time headway, TTC, safety margin) considering truck and car as vehicle types and found that these measures significantly depend on the type of leader and follower. Furthermore, they suggested that safety margin may be used as a suitable surrogate indicator for car-following traffic with speed less than 90 kmph. They used two different thresholds for trucks and cars and thus accounted for vehicle types in conflict identification. Hu et al. [151] found that with increase in composition of heavy vehicle, probability of crash may decrease. Since they used a global threshold approach for conflict estimation, increase in proportion of heavy vehicles may result in higher TTC hence reducing the overall conflict. They suggested further investigation of effect of vehicle sizes on collision risk.

From the literature, it has been established that minimum separation as well as different proximity based surrogate measures would be affected by leader and follower vehicle types. Most of the research on conflict measures have ignored vehicle heterogeneity and defined conflicts using a global threshold approach. Also, in staggered car-following scenarios interaction between vehicles is 2-dimensional, and using longitudinal indicators alone is not appropriate in defining conflicts. Although there are few studies which defined SSM based on specific vehicle type, to the best of the authors' knowledge, studies have not considered lateral interactions while defining conflicts. This study aims to estimate crash risk associated with rear-end and side-swipe crashes by considering vehicle types and 2-dimensional interactions.

6.3 Conflict Data

The detail about data collection, vehicle trajectory extraction and descriptive statistics is presented in chapter 4. The interacting vehicles were divided into three major categories namely LV (light motor vehicles), MV (medium motor vehicle) and HV (heavy motor vehicles) based on size as described in Section 4.3. The composition of LV, MV and HV were 71.04%, 23.76% and 5.18% respectively. Speed for different vehicle types were compared using data. Mean speed of LV were lowest and that of MV were highest. Speed variation among different vehicle type is depicted in the Fig.6.1.

6.3.1 Conflict indicators in staggered car-following scenario

Since the definition of crash risk in non-lane-based traffic requires an integration of the longitudinal as well as the lateral indicators, crash risk was defined by combining both of these indicators TTC and Lateral gap Gap_{lat}. TTC was computed along with the corresponding Gap_{lat} as described in Section 5.2. Distributions of TTC and Gap_{lat} among different L-F pairs are presented in Fig.6.2 . From the distribution, it is clear that, minimum TTC and Gap_{lat} depends upon the type of L-F pairs. To avoid temporal dependence among conflict indicators, vehicle trajectories during congestion were not included in this analysis. Nine interaction pairs were identified based on the vehicle categories. The descriptive statistics of TTC and Gap_{lat} for various leader and follower vehicle categories is given in Table. 6.1. The HV-HV and MV-HV cases were not included in the analysis because of smaller sample sizes.

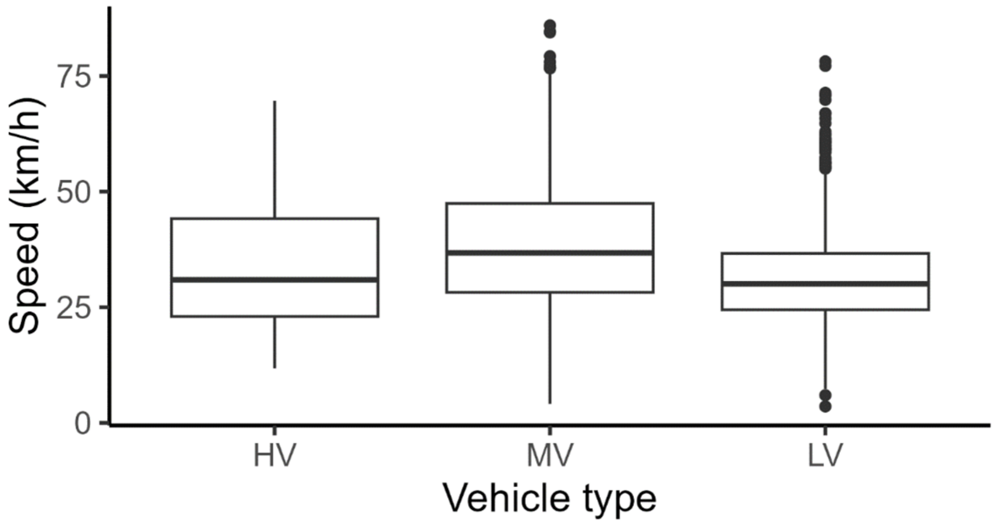


Fig. 6.1 Speed of different vehicle types.

Table 6.1 Descriptive statistics of TTC and Gap_{lat}

Sl. No	L-F pairs	Sample size	TTC (s)			Gap_{lat} (m)		
			Median	Mean	SD	Median	Mean	SD
1	LV-LV	1970	1.2	2.31	3.21	1.1	1.18	1.12
2	LV-MV	853	1.2	2.27	2.91	1.6	1.61	1.26
3	LV-HV	96	2.2	3.18	3.32	1.5	1.56	1.30
4	MV-LV	299	1.9	2.29	3.89	1.0	1.04	1.09
5	MV-MV	214	4.7	4.36	3.01	-0.7	-0.35	1.13
6	MV-HV	13	5.0	5.45	2.67	-0.55	-0.39	1.32
7	HV-LV	139	1.0	1.44	3.46	1.35	1.42	1.39
8	HV-MV	60	3.4	3.26	3.91	0.30	0.18	1.42
9	HV-HV	8	6.2	5.81	2.29	-0.35	-0.36	1.53

6.3.2 Dependence structure between TTC and Gap_{lat}

TTC and minimum Gap_{lat} were found to depend on the size of the leader as well as follower (Fig. 6.3a and Fig. 6.3b). Gap_{lat} was minimum for MV-MV interactions and maximum for LV-HV interactions. For the same leader category, TTC decreases with decrease in size of the follower vehicle as it needs a shorter time to stop in case of a near-crash event. Based on two-way ANOVA (p-value < 3e-12 for TTC, and p-value<2e-16 for Gap_{lat}) mean TTC and Gap_{lat} were significantly different and found to depend on the vehicle category of L-F pairs.

In the present study, L-F based models were proposed for safety assessment. Two modelling approaches, L-F model based on individual leader and follower vehicle category and global model with all vehicles were used for comparison of conflict and crash risk. In this study, conflicts were defined using TTC and Gap_{lat} together with the help of the BGP model as described in Section 5.3. The probability of crash (R), can be calculated using estimated model parameters using Eqn.5.4.

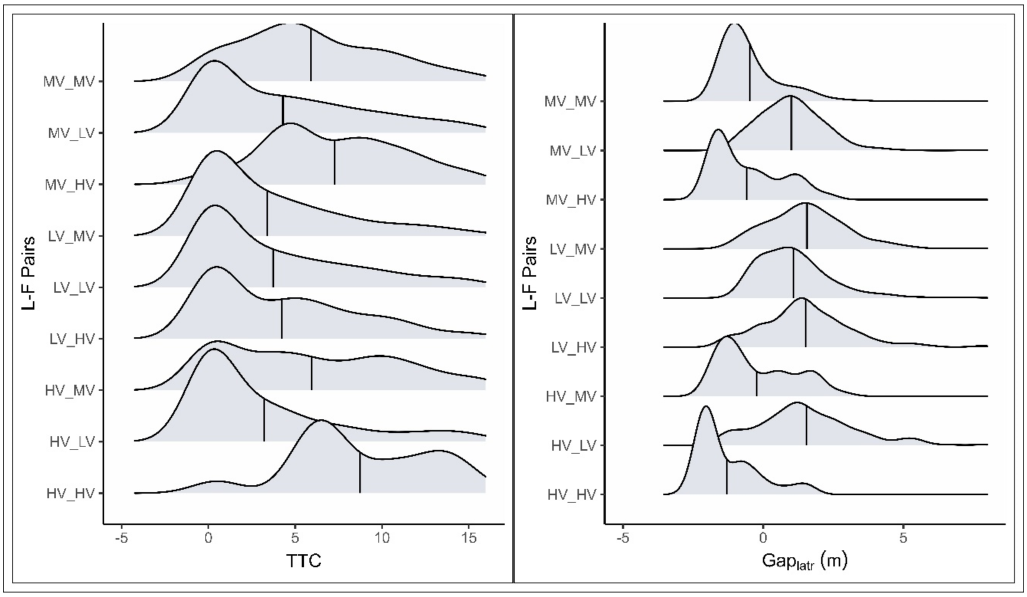


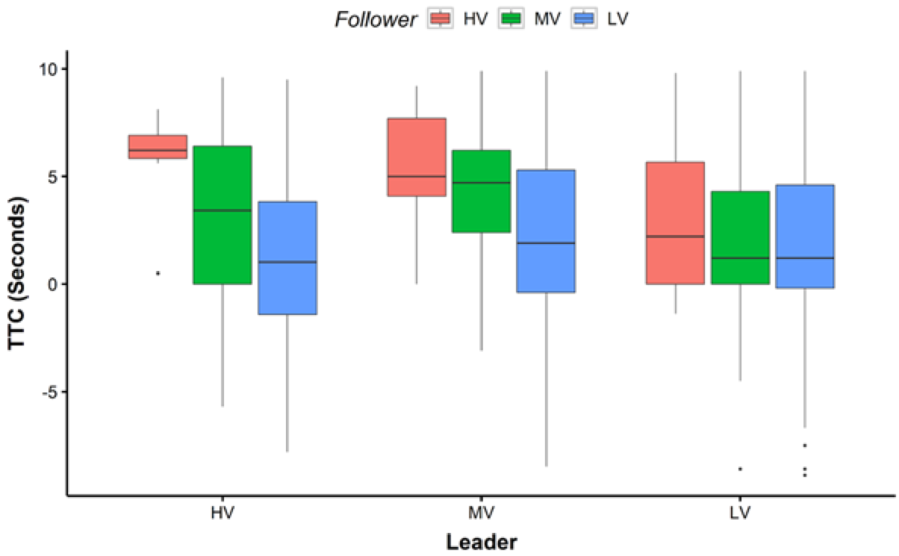
Fig. 6.2 Speed of different vehicle types.

6.4 Analysis and Results

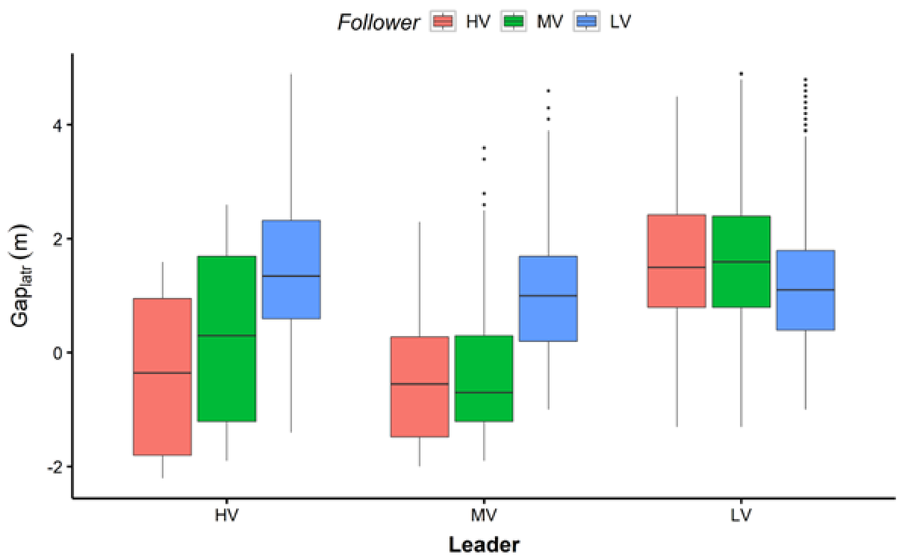
This section presents the modelling of threshold excess for conflict indicators (TTC and Gap_{lat}) using the EVT approach.

6.4.1 EVT model results

For fitting BGP models, suitable thresholds should be estimated for the two variables (marginal distributions). In this study, threshold selection was done using (1) threshold stability plot, and (2) spectral measure plot. More details about these plots are presented in Section 3.2. An example of the spectral measure plot and threshold stability plot for negated TTC and negated Gap_{lat} is depicted in Fig.6.4 and Fig.6.5. Spectral measure plot was used for estimating the threshold for the joint distribution. Further, point estimate for modelling was selected by taking the intersection of both estimates from the respective threshold stability plot and the combined spectral measure plot [185, 190]. The estimated thresholds for L-F based models as well as global model are provided in Table. 6.2.



(a)



(b)

Fig. 6.3 Comparison of (a) TTC and (b) Gap_{lat} among L-F pairs

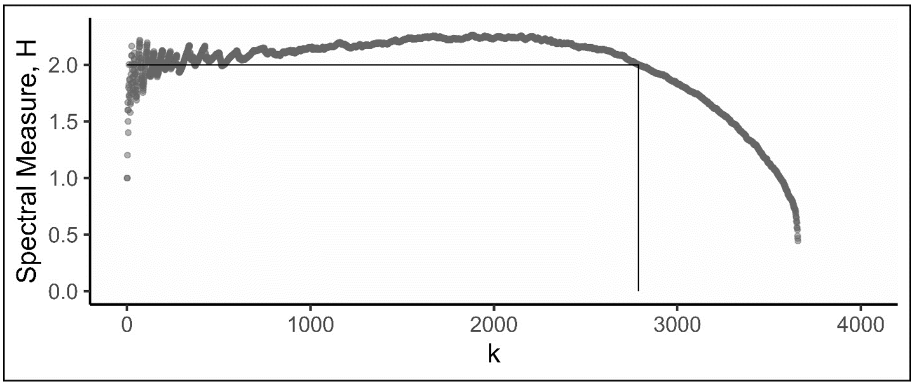
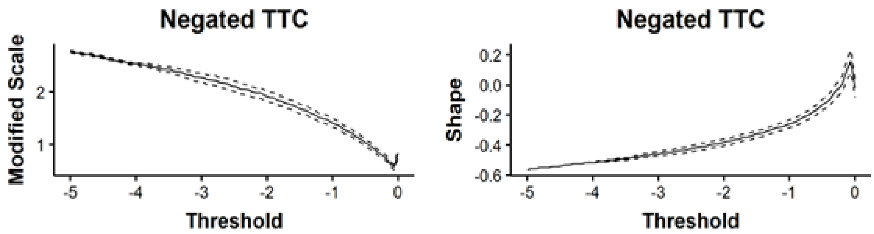
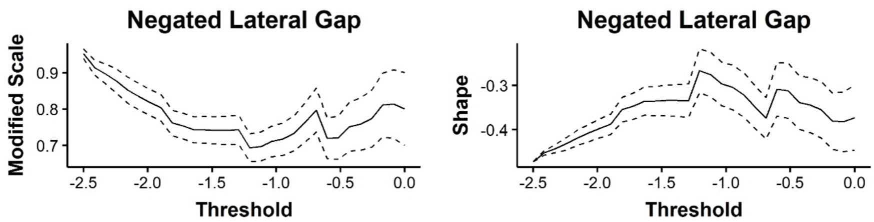


Fig. 6.4 Spectral measure plot for negated TTC vs. negated $\mathrm{Gap}_{\mathrm{lat}}$.



(a)



(b)

Fig. 6.5 Threshold stability plot for (a) negated TTC and (b) negated $\mathrm{Gap}_{\mathrm{lat}}$

Due to small sample size of MV-HV and HV-HV interactions, they were not considered in the analysis. The extreme values of the joint distribution (TTC and $\mathrm{Gap}_{\mathrm{lat}}$) were obtained by using those threshold values to the vehicle specific data and the combined data respectively. Three symmetric distributions namely Husler-Reiss, Negative Logistic and Logistic as described in Section 3.2.2, were considered for modelling the extreme values of joint distribution (TTC and Gap_{lat}). These distributions come under the umbrella of generalized pareto distribution. Logistic model was selected for further analysis since Akaike information criterion (AIC) value was lowest among all other models as presented in Table. 6.3. This is corroborated by recent studies where Logistic distribution was found to be most suitable for modelling bivariate conflict indicators [185, 190, 205].

With the estimated BGP models, the probability of crashes each year at four sites were computed for the global model as well as the L-F specific models using Eqn.5.4. Also, the 95% confidence intervals for the estimated crashes were computed using the simulation as suggested by Zheng et al. [182]. Model estimates and predicted crashes are presented in Table. 6.4. The dependence parameter (α) varies as 0.75   0.99. This may imply a weak dependence between the variables (TTC and Gap_{lat}) at extreme values. These variables also have a weak dependence (based on correlation coefficient) throughout the range of values. The dependence structure (weak dependence) being consistent with the correlation analysis makes the fitted bivariate model appropriate.

Table 6.2 Threshold for marginal and joint distribution

Indicators	Leader-Follower (L-F) Based Model							Combined Model
	LV-LV	LV-MV	LV-HV	MV-LV	MV-MV	HV-LV	HV-MV	Global
-TTC (Interval Estimate)	(-3, -2.2)	(-3, -2.3)	(-3, -2.5)	(-4, -3.3)	(-4, -3.5)	(-3.8, -3.1)	(-3.5, -3)	(-3, -2.5)
-TTC (Point Estimate)	-2.2	-2.3	-2.7	-3.3	-4.6	-3.1	-3.8	-2.6
-Gap_{lat} (Interval Estimate)	(-1.5, -0.5)	(-1.5, -1)	(-1.5, -1.1)	(-1.5, -0.7)	(-1.4, -1)	(-1.8, -1.2)	(-1.6, -0.8)	(-1.6, -1.25)
-Gap_{lat} (Point Estimate)	-1.5	-1.8	-1.8	-1.7	-1.3	-1.9	-1.8	-1.6
Selected Threshold (-TTC, -Gap_{lat})	(-2.2, -0.5)	(-2.3, -1)	(-2.5, -1.1)	(-3.3, -0.7)	(-3.5, -1)	(-3.1, -1.2)	(-3, -0.8)	(-2.5, -1.25)

Table 6.3 Model selection based on AIC

Distribution	Global Model	LV-LV	LV-MV	LV-HV	MV-LV	MV-MV	HV-LV	HV-MV
Husler-Reiss	21335.56	9252.55	4209.84	528.49	1701.91	1260.15	989.79	431.49
Negative Logistic	21330.14	9217.39	4210.05	504.92	1714.55	1243.53	962.64	425.73
Logistic	2156.65	9249.13	4401.32	496.86	1701.85	1233.57	953.99	420.46

Table 6.4 Estimation results of BGP model

Vehicle Pairs	$-$TTC(u_x)	$-$Gap$_{\text{lat}}$(u_y)	σ_x (SE)	ξ_x (SE)	σ_y (SE)	ξ_y (SE)	α (SE)	Joint exceedances (Proportion)	Log-likelihood value (LL)	Estimated crashes with 95% C.I.
LV-LV	-2.2	-0.5	2.266 (0.064)	-0.166 (0.017)	0.853 (0.03508)	-0.562 (0.0240)	0.99 (2e-06)	93 (0.0472)	-4619	0.30 (0.28, 0.31)
LV-MV	-2.3	-1	2.415 (0.108)	-0.191 (0.019)	1.451 (2e-06)	-0.630 (2e-06)	0.83 (0.024)	68 (0.080)	-2195	0.27 (0.25-0.29)
LV-HV	-2.5	-1.1	3.881 (1e-06)	-0.995 (1e-06)	2.572 (1e-06)	-1.071 (1e-06)	0.75 (0.051)	4 (0.042)	-243	0.29 (0.27-0.30)
MV-LV	-3.3	-0.7	4.848 (0.424)	-0.3569 (0.056)	1.406 (0.159)	-0.818 (0.102)	0.99 (2e-06)	41 (0.137)	-846	0.40 (0.35-0.44)
MV-MV	-3.5	-1	3.797 (0.415)	-0.553 (0.071)	2.950 (2e-06)	-1.017 (2e-06)	0.75 (0.037)	46 (0.214)	-612	0.57 (0.55-0.57)
HV-LV	-3.1	-1.2	5.448 (0.007)	-0.499 (2e-06)	2.037 (0.253)	-0.773 (0.1015)	0.75 (0.0466)	33 (0.235)	-472	0.42 (0.38-0.45)
HV-MV	-3.2	-0.8	4.850 (0.856)	-0.500 (0.1168)	2.822 (2e-06)	-1.045 (2e-06)	0.75 (0.0537)	8 (0.13)	-205	0.50 (0.45-0.53)
Global	-2.5	-1.25	2.757 (0.059)	-0.225 (0.0083)	1.263 (0.0339)	-0.3396 (0.00192)	0.99 (2e-06)	843 (0.230)	-10673	0.33 (0.32-0.34)

The estimated conditional probabilities Pr(conflict|interaction) and Pr(crashes|conflict) for vehicle L-F pairs are obtained from the BGP model and depicted in Fig.6.6. The conditional probabilities of conflicts and crashes are related, and as the risk of conflict increases, crash risk also increases across vehicle pairs. This validates the fitted extreme value BGP models even further as suggested by Tarko [218]. The suitability of BGP model along with TTC and Gap_{lat} as conflict measures in modelling crash risk in heterogeneous traffic conditions is supported by the above findings. Probability of conflict and crash are both smaller when lighter vehicles are leaders as compared to situations when they are followers. Probability of conflict and crash are both higher among medium sized vehicles as compared to others. Even though, proportion of light vehicle interactions is highest (53.9%), the probability of conflict as well as crashes are lowest among them. There does not seem to be a relationship between the proportion of interactions observed in the data, and the proportion of conflicts in those interactions. Crash based validation of the models could not be performed due to unavailability of detailed crash data. Therefore, model validation was performed based on rationalism of estimates [188]. Comparisons were made between global and L-F based models based on estimated conflict probability (Fig.6.7). The global threshold model under predicts the risk for medium and heavy vehicles and over predicts the risk in case of lighter vehicles as compared to the L-F based models.

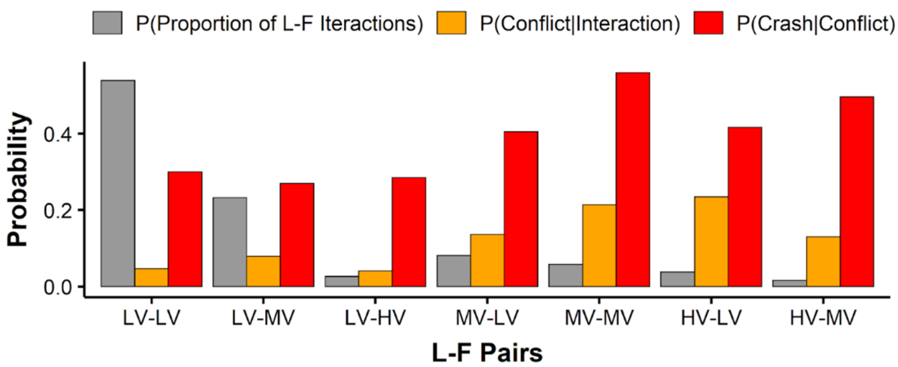


Fig. 6.6 Proportion of conflicts in interactions and proportion of crashes in conflicts.

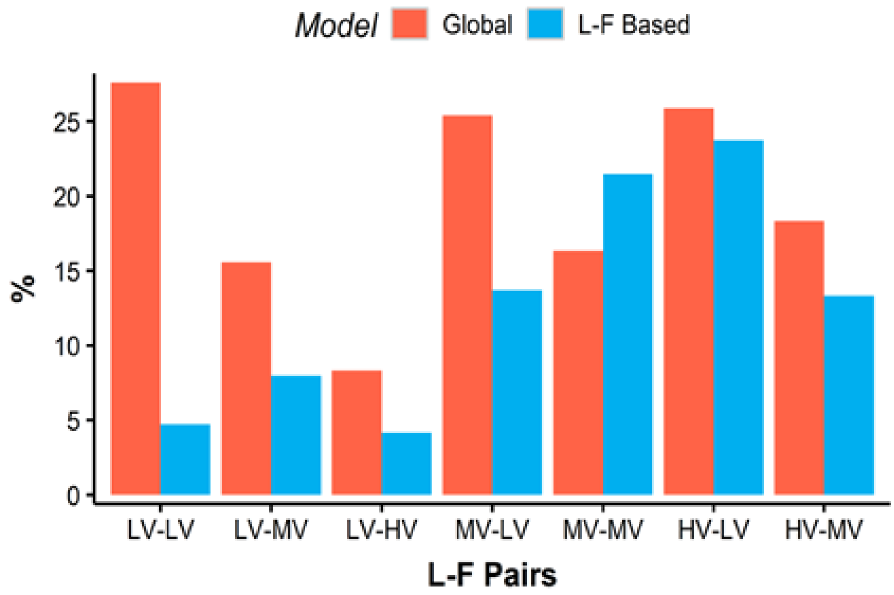


Fig. 6.7 Comparison of proportion of conflicts in total interactions based on the two modelling approaches (global and L-F).

6.4.2 Sensitivity analysis

Sensitivity of crash probability to threshold selection was analysed for checking the accuracy of selected thresholds. If the selected thresholds are large enough, the exceedances will follow GPD and will result in almost the same crash probability [218]. Threshold values given in Table. 6.2 were used for comparison of crash probabilities. Estimated probability of crash with selected thresholds (Threshold-1) were compared with the crash probability using thresholds from spectral measure plot (Threshold-2), for each model as depicted in Fig.6.8. The variation in results obtained from different thresholds were small which justifies the selection of thresholds in this study.

6.5 Discussion

The present study demonstrated the use of bivariate extreme value models for safety assessment in two-dimensional vehicle interactions with vehicular heterogeneity. Due

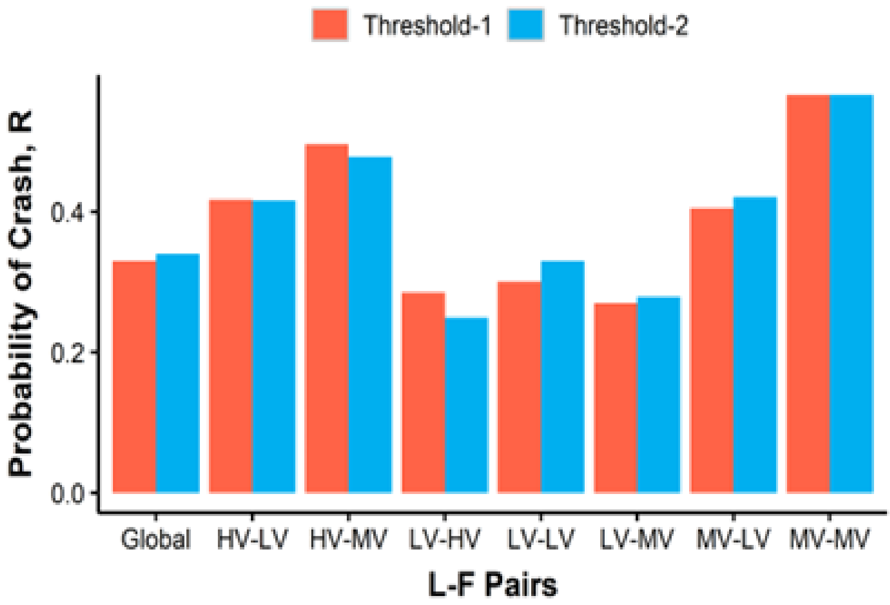


Fig. 6.8 Sensitivity analysis of crash probability with different thresholds.

to significant heterogeneity (two-way ANOVA (p-value < 3e-12 for TTC, and p-value < 2e-16 for Gap_{lat})) among vehicle pairs, L-F based models were used. The findings from this study are discussed here.

6.5.1 Minimum TTC and Gap_{lat} depends upon the L-F vehicle type

TTC was found to be significantly different among different L-F pairs. This finding is similar to the finding of Zhao and Lee [153]. TTC was least among lighter vehicle interactions (Fig.6.3). Since lighter vehicles have shorter stopping distance due to small size and lower speed, they tend to maintain shorter gaps which results in smaller TTC. Also, minimum Gap_{lat} was found to be the lowest for MV-MV interactions, as those vehicles tend to follow the inline car-following scenario. Kashyap et al., (2020) also found similar results with cars in heterogeneous traffic conditions. The minimum Gap_{lat} was found to be higher for interactions involving LV as leader since LV tend to give way while followed by heavy vehicles (Fig.6.3).

6.5.2 Comparison of conflicts using L-F based threshold and global threshold model

The global threshold would underestimate conflicts for some vehicle pairs and overestimate it for others since vehicle specific information is lost due to aggregation. As seen in Fig.6.7, the percentage of conflicts estimated using the global threshold model is high for LV-LV interactions. Goyani et al. [89] used global threshold and reported that interactions among lighter vehicles (m2w and m3w) were more critical. Even though the lighter vehicles have ability to maintain smaller gaps (due to smaller sizes and weights, and lower speeds), based on a global threshold approach these fall in the categories of conflicting vehicles. Using a single threshold would designate normal interactions between those pairs as conflicts. Fig.6.9 depicts the single global threshold-based conflict segregation.

In the present study, the threshold for TTC was found to be lower in lighter vehicles as compared to medium and heavy vehicles. This finding is similar to Das and Maurya [43]. Also, TTC threshold among heavy vehicles interactions was found maximum, as they maintain higher longitudinal gap due to large sizes [149, 154]. If a global threshold is used to identify conflicts, the intensity of interaction in case of HV-HV and LV-LV vehicle pairs would be treated in the same way which may not be appropriate [25]. Therefore, using a global threshold is not appropriate in a traffic stream with heterogeneous vehicle sizes. The L-F based approach would enable identifying critical vehicle groups so that countermeasures specific to those groups may be deployed.

6.5.3 Effect of vehicle size on conflicts and crashes

Based on conditional probability of crashes (Fig.6.6), it can be argued that probability of being in a crash is a function of L-F pairs. Although the proportion of interaction among lighter vehicles was highest (Table. 6.4), the estimated probability of conflicts as well as crashes involving LV-LV pairs was low. This may be attributed to lower speed and smaller

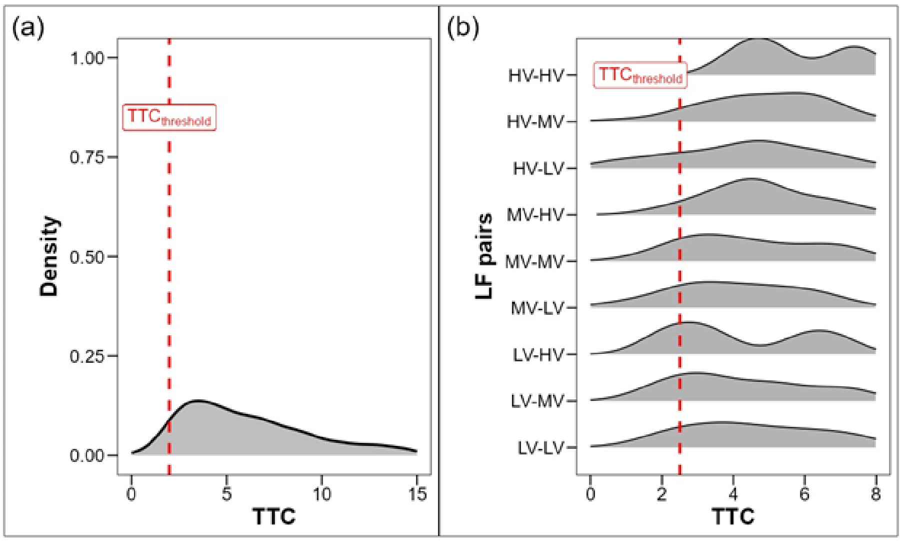


Fig. 6.9 Single threshold based conflict segregation.

braking distance of lighter vehicles. Also, the estimated crash probability for LV as a leader is lower than the same being a follower vehicle. This reveals the risky driving behaviour of lighter vehicle drivers in heterogeneous traffic. Since MVs move at a relatively higher speed, and maintain smaller Gap_{lat} than others, probability of crash as well as conflict is high among them.

For a global model (not incorporating vehicle heterogeneity), the threshold of conflict indicator is affected by vehicle composition [219]. Therefore, the conflict probability will also depend on vehicle composition. Also, results from the L-F based model confirms that high proportions of interactions do not necessarily imply high probability of conflicts (Fig.6.7). For example, even though the proportion of LV-LV pairs is highest in the observed data, the probability of conflict is lowest among them. Therefore, global threshold which depends on vehicular composition, may not be appropriate for conflict prediction. As traffic composition tends to vary from site to site the importance of L-F based models becomes even more important for safety assessment.

6.6 Chapter Summary

This study examines the effect of vehicle types on crash risk in two-dimensional interactions which prevails in heterogeneous traffic scenarios. Interacting vehicle pairs were divided into 9 groups based on the size of leader and follower (L-F) vehicles. The L-F based model as well as global models (not accounting for vehicle size) were fitted using bivariate Generalized Pareto Distribution (GPD). Threshold selection is an important step in modeling extremes. Thresholds were selected using both threshold stability plots and spectral measure plots. The following were the major findings of the study:

1. TTC varied significantly among different L-F pairs. TTC was the minimum for lighter vehicle interactions.

2. 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.

3. The estimated crash probability for LV as a leader is lower than the same being a follower vehicle.

This study demonstrates the importance of incorporating vehicle type as well as 2-dimensional interactions in safety assessment.