

Impacts of speed variations on freeway crashes by severity and vehicle type

Pushpa Choudhary

Research student
Transportation systems engineering,
Department of Civil Engineering,
Indian Institute of Technology (IIT) Bombay
Powai, Mumbai-400 076, India
Tel: +91-22-2576-4348;
E-mail: pushpal1choudhary@gmail.com

Dr Marianna Imprialou

Lecturer (Assistant Professor) in Transport Engineering
School of Civil & Building Engineering
Loughborough University
Loughborough LE11 3TU
United Kingdom
Tel: +44 (0) 1509223422
Email: M.Imprialou@lboro.ac.uk

Dr Nagendra R Velaga (Corresponding Author)

Associate Professor
Transportation Systems Engineering,
Department of Civil Engineering,
Indian Institute of Technology (IIT) Bombay
Powai, Mumbai-400 076, India
Tel: +91-22-2576-7341;
E-mail: n.r.velaga@iitb.ac.in velaga@iitb.ac.in

Dr Alok Choudhary

Reader in Supply Chain Management
School of Business and Economics,
Loughborough University
Loughborough LE11 3TU
United Kingdom
Tel: + 44 (0)1509 228302;
E-mail: A.Choudhary@lboro.ac.uk

44 Abstract

45 Speed variations are identified as potentially important predictors of freeway crash rates;
46 however, their impacts on crashes are not entirely known. Existing findings tend to be
47 inconsistent possibly because of the different definitions for speed variations, different crash
48 type consideration or different modelling and data aggregation approaches. This study explores
49 the relationships of speed variations with crashes on a freeway section in the UK. Crashes split
50 by vehicle type (heavy and light vehicles) and by severity mode (killed/serious injury and slight
51 injury crashes) are aggregated based on the similarities of the conditions just before their
52 occurrence (condition-based approach) and modelled using Multivariate Poisson lognormal
53 regression. The models control for speed variations along with other traffic and weather
54 variables as well as their interactions. Speed variations are expressed as two separate variables
55 namely the standard deviations of speed within the same lane and between-lanes over a five-
56 minute interval. The results, similar for all crash types (by coefficient significance and sign),
57 suggest that crash rates increase as the within lane speed variations raise, especially at higher
58 traffic volumes. Higher speeds coupled with greater volume and high between-lanes speed
59 variation also increase crash likelihood. Overall, the results suggest that specific combinations
60 of traffic characteristics increase the likelihood of crash occurrences rather than their individual
61 effects. Identification of these specific crash prone conditions could improve our understanding
62 of crash risk and would support the development of more efficient safety countermeasures.

63

64 *Keywords:* accidents; speed variation; road safety; crash severity; heavy goods vehicles;
65 multivariate count modelling.

66 **1. Introduction**

67 Speed and speed variations are considered to be among the most important crash contributory
68 factors. Several ITS applications such as Variable Speed Limits (VSL) or cooperative systems
69 are designed to provide speed harmonization anticipating that this will lead to lower crash rates
70 (Farah and Koutsopoulos, 2014; Strömngren and Lind, 2016). However, studies considering
71 speed variations as a contributory factor are relatively low in number and their results are
72 varying (Kockelman and Ma, 2007; Quddus, 2013; Shi et al., 2016). Some of the studies find
73 speed variations to be positively associated with crashes (Quddus, 2013; Tanishita and Wee,
74 2016; Wang et al., 2018) while others find non-significant relationships between speed
75 variations and crash risk (Kockelman and Ma, 2007). Others also report changes in the effects
76 of speed after including speed variance in models (Garber and Gadiraju, 1989).

77 The often-conflicting results of the existing studies may be related to the multiple
78 definitions used to express speed variations, the differences in modelling approaches or data
79 quality and pre-processing methods. All these suggest that further exploration of this
80 contributory factor is needed. Current advances in crash modelling can be proved useful in the
81 examination of the impact of speed variations on crashes. Recently, crash data aggregation has
82 been found to be highly influential on the estimated coefficients of time-varying variables such
83 as speed and traffic flow (Imprialou et al., 2016a; Imprialou et al., 2016b; Xu et al., 2018; Yu
84 et al., 2018). When crashes are aggregated according to the similarities of the traffic conditions
85 just before their occurrence, modelling results appear to be more reliable than in traditional
86 location-based approaches (Imprialou et al., 2016b). Additionally, research has shown that
87 independent variables in crash modelling have unique effects on different crash types and these
88 are more accurately estimated when the correlations between the examined crash types are

89 taken into consideration (i.e. multivariate count models) (Huang et al., 2017; Lord and
90 Mannering, 2010).

91 Although there are many multivariate crash prediction models that examine crashes by
92 severity, there is a very limited number of studies that divides crashes by the involved vehicle
93 types and none of them focuses on heavy goods vehicles. This paper analyses the effects of
94 speed variations along with other traffic and weather variables on different types of crashes
95 and specifically by vehicle types (heavy and light vehicles) and by severity type (killed/serious
96 injuries and slight injuries; Property Damage Only (PDO) crash data were not available and
97 therefore this crash type was excluded from the analysis). Multivariate Poisson lognormal
98 regression models are used to develop the relationships that are applied on a dataset aggregated
99 with the condition-based approach.

100 **2. Literature Review**

101 The impact of traffic characteristics on crash frequency and severity has been widely studied
102 in the literature and has offered useful insight into the development of effective mitigation
103 measures. Speed has received a lot of research attention, but the findings regarding its
104 relationship with crash rates are inconsistent (Aarts and Schagen, 2006). It is clear that higher
105 speed is associated with higher crash severity, but the impact of speed on crash frequency is
106 not clearly defined yet. Some studies suggest a positive relationship between speed and crash
107 frequency (Imprialou et al., 2016a; Imprialou et al., 2016b; Kloeden et al., 2002; Taylor et al.,
108 2000); however, others have shown a negative or an insignificant relationship (Kockelman and
109 Ma, 2007; Quddus, 2013; Stuster, 2004). There is also a common belief that speed does not
110 necessarily lead to more crashes as long as there are no extreme speed differences between
111 vehicles on a roadway section. These differences that are typically referred to as speed
112 variations and have been identified as a potentially significant contributory factor; however,

113 their exact effect on crashes remains inconclusive (Aarts and Schagen, 2006; Kockelman and
114 Ma, 2007; Quddus, 2013; Solomon, 1964). There have been significantly fewer studies
115 focusing on speed variations than on speed and other traffic, geometric or environmental
116 variables (Quddus, 2013). This is may be partially because speed variations are not directly
117 measurable and may be hard to be computed unless the available traffic data are highly spatially
118 and/or temporally aggregated.

119 The effects of speed and its variations were initially studied by Solomon (1964) in a
120 case-control study that suggested that vehicles moving much faster or slower than the modus
121 speed were exposed to higher crash risks introducing the theory “Variance kills”. Some
122 subsequent studies reported that speed variation is so highly influential for triggering crashes
123 that it makes the effect of mean speed negligible, suggesting that “Variance kills, not speed”
124 (Garber and Gadiraju, 1989). This was in line with the findings by Quddus (2013) who found
125 that speed variation is associated positively with the crash rates but, the average speed is not.
126 However, it contradicts the outcomes of other studies that find both speed and speed variance
127 to be significant factors for predicting crash frequency (Levy and Asch, 1989; Tanishita and
128 Wee, 2016). Studies on real-time crash prediction have shown negative associations of average
129 speed with crashes, while a positive relationship between speed variation and crashes (Abdel-
130 Aty et al., 2012; Wang et al., 2016; Wang et al., 2015a; Xu et al., 2016; Yu and Abdel-Aty,
131 2014). Moreover, the effects of speed and speed variations seemed to be related to other traffic
132 variables such as flow (Abdel-aty and Pemmanaboina, 2006; Xu et al., 2016). For instance,
133 Abdel-aty and Pemmanaboina (2006) mentioned that high-speed variation coupled with high
134 occupancy and low variation in volume leads to higher likelihood of a crash, while, Xu et al.
135 (2016) showed that, high-speed variance in high-density traffic flow leads to higher crash risk.

136 The inconsistencies among the results may be related to the differences between
137 analytical methods and also with the definition of speed variations. Speed variation has been

138 represented by multiple different measures such as differences in speed at individual vehicle
139 level (Kloeden et al., 2002; Solomon, 1964), differences at section level traffic characteristics
140 (Quddus, 2013), the difference between the 90th to the 50th percentile of speeds in each lane
141 (Golob et al., 2004), speed differences between and across lanes (Kockelman and Ma, 2007)
142 and others.

143 The differences in results could also be related to different crash types. For instance,
144 Kweon and Kockelman, (1996) showed that the effects of speed variation were dependent on
145 crash severity and that specifically slight-injury crashes were associated with high-speed
146 variance. Current crash prediction modelling suggests that separate models for different crash
147 types are not adequate; and therefore, multivariate modelling approaches came into application
148 (e.g. Huang et al., 2017; Imprialou et al., 2016b; Lord and Mannering, 2010; Martensen and
149 Dupont, 2013). Though there are various studies on crash contributory factors by severity
150 levels, there are very few studies focusing on crashes by vehicle type and these are mostly
151 focused on urban environments without making a distinction between heavy and light vehicles
152 (Huang et al., 2017). Whereas, it has been known that due to their unique characteristics
153 (weight, size, stopping distances etc.) heavy vehicles' crash contributory factors should be
154 investigated separately (Wei et al., 2017). Moreover, as per authors' best knowledge, there is
155 no study on investigating the effects of speed variation on heavy vehicle crashes.

156 Other than speed, traffic volume is one of the most studied factors in crash rate
157 predictions (Aarts and Schagen, 2006; Garber and Ehrhart, 2000). Weather conditions could
158 also affect crash risk (e.g. Abdel-aty and Pemmanaboina, 2006; Wang et al., 2015b; Xu et al.,
159 2016). Typically rainy weather is found to be associated with higher crash rates in most of the
160 previous studies (Abdel-aty and Pemmanaboina, 2006; Lee et al., 2003), possibly because, the
161 wetness of pavement reduces friction, making stopping distances longer (Abdel-aty and
162 Pemmanaboina, 2006).

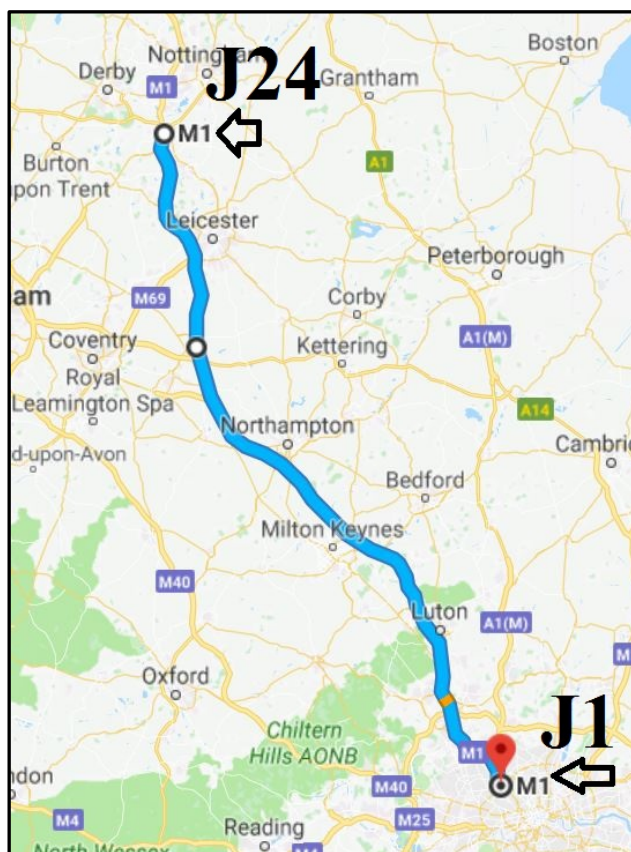
163 This study explores further the relationships of traffic characteristics with crash rates
164 with a special focus on the impact of speed variations (defined as speed differences within and
165 between-lanes). Freeway crashes are split by vehicle types (heavy and light vehicle crashes)
166 and severity (killed or serious and slight injury) and are fitted using multivariate count models.
167 In order to achieve a more accurate representation of the conditions just before crashes, data
168 are aggregated following a condition-based approach (Imprialou et al., 2016b).

169 **3. Data Collection and Preparation**

170 To analyse the impact of speed variations on crashes, traffic and weather data have been
171 employed. The study area was decided to be a section of the South-North motorway M1
172 (Junctions 1-24 (Figure1), located between London and East Midlands Airport) that is one of
173 the most important and busy motorways in England that links London with the North of the
174 country. The length of the study area is 175km per direction and most of its links include three
175 running lanes in each direction. The crash data for three years (from 2013 to 2015) was obtained
176 from the National Road Accident Database of the United Kingdom (STATS 19) (Department
177 for Transport, 2011). Among others, the data included information on severity, involved
178 vehicle types, time, date and location of the crashes. During the study period, there were 1,075
179 fatal and injury crashes in total, of which 11.25% resulted in killed or seriously injured
180 casualties (henceforth: KS crashes) and 88.75% in slight injuries (henceforth: SL crashes). As
181 the study area belongs to the Strategic Road Network of England, that carries almost two-thirds
182 of England's freight, 15.90% of all crashes had at least one commercial vehicle with weight
183 over 3.5 tones involved i.e. heavy vehicles (henceforth these crashes will be referred to as HV-

184 crashes). The rest of the crashes (84.10%) were between mainly passenger vehicles or vans
 185 with weight 3.5 tonnes or less i.e. light vehicles (henceforth: LV-crashes)¹.

186



187

188 **Figure 1: M1 motorway Junctions 1-24, UK (source: Google Maps (2017))**

189

190 Traffic data were obtained from the Motorway Incident Detection and Automatic
 191 Signalling database (MIDAS) of Highways England (Highways England, 2017). The data was

¹ In the present study, a crash is defined as HV-crash if the crash includes at least one heavy goods vehicle. Whereas, LV-crashes are the crashes that involve at least one light vehicle but excluding the crashes which include heavy goods vehicle. Therefore, the crashes which include both heavy goods vehicle and light vehicle are classified as HV-crashes. This definition of crashes by vehicle type has been employed in a number of other studies such as: Chen and Chen, (2011); Lemp et al., (2011) and Zou et al., (2017).

192 collected through 689 inductive loop detectors installed in the study area and provided one-
 193 minute-level traffic data disaggregated by running lane. The traffic variables that were used for
 194 this analysis were traffic volume and mean speed (km/h) by lane. To develop the final dataset
 195 for the analysis, the data were aggregated to the five-minute level and through this aggregation
 196 process the following variables have been developed:

- 197 • *Total volume*: The total volume was estimated by the summation of the number of
 198 vehicles present on a road section between two subsequent loop detectors in each of the
 199 running lanes during a 5 min interval.

$$Total\ volume = \sum_{t=1}^T \left(\sum_{l=1}^L Volume_{t,l} \right) \quad (1)$$

200 where l : lane index (1 to 3) and t : number of minutes (1 to 5).

- 201 • *Average speed*: For each one-minute interval, mean speeds across the lanes were
 202 calculated and then, the average speed for 5 minutes was considered as the average
 203 speed.

$$Average\ Speed = \frac{1}{T} \sum_{t=1}^T \left(\frac{1}{L} \sum_{l=1}^L Speed_{l,t} \right) \quad (2)$$

204 where T : total number of minutes (here $T=5$) and L : the total number of lanes of the road
 205 section.

- 206 • *Between-lanes speed variation*: For each one-minute interval, the standard deviation of
 207 speeds between the lanes was calculated and then, the average of these standard
 208 deviations for 5 minutes was considered as the between-lanes speed variation.

$$Between\ lanes\ speed\ variation = \frac{1}{T} \sum_{t=1}^T \left(\sqrt{\frac{\sum_{l=1}^L (Speed_{l,t} - \overline{Speed}_t)^2}{L}} \right) \quad (3)$$

210 where \overline{Speed}_t : average speed for all lanes for minute t .

- 211 • *Within lane speed variation*: For each lane, the standard deviation of speeds for a 5-
 212 minute interval was calculated and then the average of these standard deviations for all
 213 the three lanes was considered as within lane speed variation.

$$\textit{Within lane speed variation} = \frac{1}{L} \sum_1^L \left(\sqrt{\frac{\sum_1^T (\textit{Speed}_{l,t} - \overline{\textit{Speed}_l})^2}{T}} \right) \quad (4)$$

214 where $\overline{\textit{Speed}_l}$: average speed for 5 minutes within lane l .

- 215 • *Vehicle hours travelled*: Estimated by multiplying the average travel time on each
 216 section (based on average speed) and the corresponding total volume in each 5-minute
 217 interval.

218 Weather conditions were extracted from the open database of MetOffice, the United
 219 Kingdom's national weather service (MetOffice, 2016). The weather data was collected on
 220 hourly basis from eight weather stations which were found adjacent to the study site based on
 221 their geographic locations. Each of the loop detectors in the study area was assigned with one
 222 of these eight stations based on the proximity of the station with the loop detector. For the sake
 223 of simplicity, weather conditions in this analysis were split into two categories indicating
 224 presence or absence of rain. Further, based on the time of the observation of the traffic data, it
 225 was matched with the hourly weather data, to provide the weather information for the same 5-
 226 min interval.

227 3.1 Condition-Based Dataset

228 Data aggregation in crash modelling has been found to influence the results of the analysis
 229 significantly (Imprialou et al., 2016a; Imprialou et al., 2016b). Traditionally, crash count
 230 models are applied onto location-based datasets, where the number of crashes per location unit
 231 (e.g. road link, section or intersection) is modelled against averages of the examined
 232 independent variables (e.g. the annual average of speed, AADT, number of lanes). This

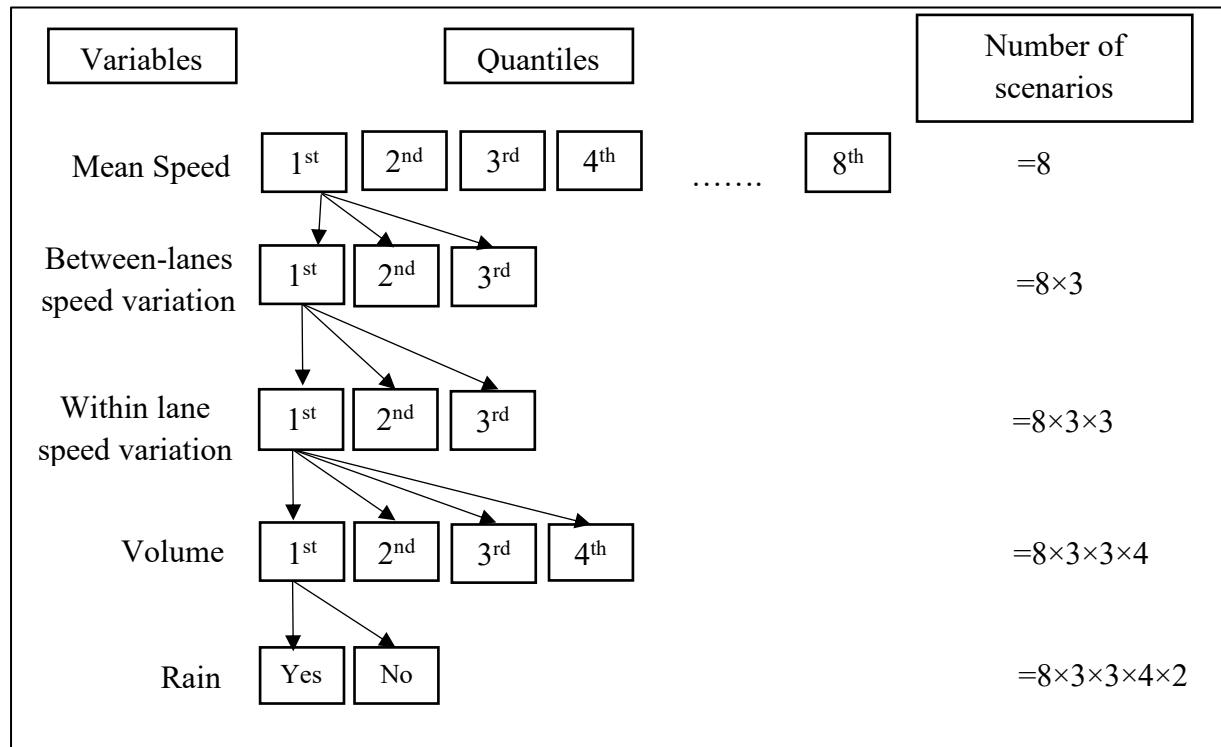
233 approach may be effective for the examination of permanent road characteristics such as road
234 geometry. However, it can be less suitable for understanding the impact of time-varying traffic
235 characteristics on crashes. For instance, in an analysis that employs a location-based dataset,
236 speed variation can be only expressed by the annual variance of speed on the study area which
237 might be not representative of the traffic conditions that are related to crashes.

238 To address this aggregation bias, an alternative condition-based aggregation approach
239 has been proposed, as it indicates the prevailing traffic conditions just prior to the crashes,
240 which can eventually help in identifying the extreme traffic characteristics which might have
241 contributed to crashes. A condition-based model aggregates the crashes based on the similarity
242 of the traffic conditions prior to their occurrence rather than the adjacency of their locations.
243 Therefore, a condition-based dataset includes a number of scenarios that cover all the possible
244 traffic conditions in the study area and each of these scenarios is matched with the respective
245 number of crashes that happened under these conditions (Imprialou et al., 2016b).
246 Consequently, to develop a condition-based dataset, the traffic conditions before each of the
247 examined crashes need to be identified. Pre-crash conditions were defined as the traffic and
248 weather conditions at the closest upstream loop detector to the crash location, five minutes
249 prior to the reported crash time. Some crashes (N=140) were removed from the dataset as they
250 had missing values for the traffic or weather parameters. The final dataset consists of 153 HV
251 crashes (16.37%), 782 LV crashes (83.63%). In terms of severity, 130 crashes (13.9%) were
252 identified as KS crashes, and 805 were identified SL crashes (86.1%).

253 To prepare the pre-crash scenarios, the traffic variables were grouped into equal
254 frequency categories. The reason behind the formation of different scenarios of the traffic and
255 weather data is to represent all possible conditions which could be present in the study area just
256 before the crashes. In order for the scenarios to have equal frequency and to be mutually

257 exclusive, the traffic characteristics were divided into quantiles. The formation of the database
 258 is visualised in Figure 2.

259



260

261 **Figure 2 Flow diagram representing the sequence of scenario creation.**

262 The number of scenarios of the condition-based dataset was empirically defined. The
 263 order of dividing the variables was followed as per the study aim. As the main aim of the study
 264 was to identify the effects of speed and speed variance on the crash frequency, firstly the speed
 265 was split then each speed quantile was divided into different quantiles for between-lanes speed
 266 variations; and similarly, the within lane speed variance was split under each quantile of for
 267 between-lanes speed variations. Further the sequence was followed by splitting the volume and
 268 rain variables respectively. The number of scenarios was determined in order to develop a
 269 dataset with relatively small number of observations so as to avoid generating too many zeros
 270 that might be problematic for the modelling estimations (see (Imprialou, 2015) for a detailed
 271 explanation). During the analysis several other scenario aggregations have been tested but the

272 estimated coefficients did not change from those of the model that will be presented in Section
273 5.

274 As shown in Figure 2, the best scenario combination was achieved by dividing the
275 average speed into eight quantiles (octiles), and further, dividing the between-lanes speed
276 variation for each separate speed quantile into three quantiles (tertiles). Similarly, the within
277 lane speed variation was divided into three quantiles for each quantile of between-lanes speed
278 variation. The volume was divided into four equal frequency groups (quartiles) for each within
279 lane speed variation category. Finally, the grouping was done for weather conditions (rain/no
280 rain). This grouping led to 576 scenarios ($8 \times 3 \times 3 \times 4 \times 2$) which included all possible
281 combinations of variables and each observation represented a distinct traffic and weather
282 scenario. The current study developed and compared the outcomes of two datasets that
283 expressed traffic and weather conditions at two different time intervals prior of each crash in
284 the dataset: (1) 0-5 minute interval, (2) 5-10 minute interval.

285 The traffic characteristics were represented by the median of each quantile. Each crash
286 was then matched with one of the 576 scenarios. The crash frequency for each scenario was
287 presented by vehicle types (HV and LV) and by severity levels (KS and SL). Table 1 shows
288 the descriptive statistics of the study dataset. The exposure on a condition scenario is dependent
289 on the number of vehicles and duration of their movement under these conditions (Imprialou
290 et al., 2016a). Therefore, the total vehicle hours travelled per scenario was selected as the
291 exposure variable for the models.

292

293

294

295 **Table 1 Descriptive statistics for the study dataset**

Variable	Mean	SD	Min	Max
Crash variables				
By transport mode				
HV crashes	0.36	0.98	0.00	8.00
LV crashes	1.36	3.00	0.00	28.00
By Severity levels				
KS crashes	0.23	0.64	0.00	5.00
SL crashes	1.40	3.14	0.00	29.00
Traffic variables				
Speed (km/h)	105.35	11.31	41.07	120.90
Between-lanes speed variation (km/h)	14.12	4.45	3.82	49.81
Within lane speed variation (km/h)	5.56	2.08	2.43	12.83
Volume (vehicles in 5 min interval)	177.33	113.88	27.00	399.00
Speed*Volume (km/h*vehicles)	18556.63	11886.22	2423.25	38519.67
Speed*Between-lanes speed variation (km/h*km/h)	1483.02	428.99	226.50	4014.96
Speed*Within lane speed variation (km/h*km/h)	582.82	219.39	195.94	1407.42
Volume* Between-lanes speed variation (vehicles*km/h)	2473.99	1746.29	152.74	10958.94
Volume* Within lane speed variation (vehicles*km/h)	939.93	700.66	85.79	4030.05
Between-lanes speed variation *Within lane speed variation (km/h*km/h)	78.86	40.07	9.33	279.47
Weather variables				
Rain	0.50	0.50	0.00	1.00

296

297 **4. Methodology**

298 Different crash types sourcing from the same dataset may be potentially correlated. Omission
299 of these correlations from the modelling process, by developing separate count models for each
300 crash type, is likely to lead to erroneous standard errors (Park and Lord, 2007). Multivariate
301 Poisson Lognormal (MVPLN) regression can control for over-dispersion as well as for the
302 correlations between dependent variables; and it has been applied in a number of studies

303 (Huang et al., 2017; Park and Lord, 2007). This study explores the relationships of speed
 304 variations with crash rates by developing two MVPLN models: one that examines the
 305 aforementioned traffic and weather variables by vehicle types (HV and LV crashes) and
 306 another by severity level (KS and SL crashes).

307 In an MVPLN, the number of crashes by type (vehicle type or severity) for a dataset
 308 with n observations (i.e. condition scenarios) follows a Poisson distribution:

$$y_{ik} \sim \text{Poisson}(\lambda_{ik}), \quad i = 1, 2, 3, \dots, n; k = 1, 2, \dots, K \quad (5)$$

309 where i : index of observation, k : index of crash type, y_{ik} : observed number of crashes for k
 310 crash type for i^{th} observation and λ_{ik} expected mean for k type crashes for i^{th} observation.

311 Following is the link function for λ_{ik} :

$$\ln(\lambda_{ik}) = \beta_{k0} + \sum_{m=1}^m \beta_{km} X_{ikm} + \ln(e_i) + \varepsilon_{ik} \quad (6)$$

312 where β_{k0} intercept of k crash type; β_{km} : coefficient of m^{th} explanatory variable for k crash
 313 type, X_{ikm} : value of m^{th} explanatory variable for i^{th} observation for k crash type. ε_{ik} :
 314 unobserved heterogeneity for i^{th} observation for k crash type.

315 ε_i is assumed to follow multivariate normal (MVN) distribution and controls for the
 316 correlations within the unobserved heterogeneity:

$$\varepsilon_{ik} \sim \text{MVN}(0, \Sigma), \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \dots & \sigma_{kk} \end{pmatrix} \quad (7)$$

317 where Σ is the variance–covariance matrix of the unobserved heterogeneity.

318 The model's parameters estimation was done using Markov chain Monte Carlo
 319 (MCMC) in a Bayesian framework because the direct computation of the marginal distribution
 320 of accident counts is not possible to be obtained directly (for more information see: Ma, 2006;

321 Park and Lord, 2007; Imprialou et al., 2016b; Wang et al., 2015a). The prior distribution for β
 322 is multivariate normally distributed:

$$\beta \sim MVN(\beta_0, R_{\beta_0}) \quad (8)$$

323 The conjugate prior distribution of the inverse of the variance-covariance matrix for the
 324 heterogeneity and the spatial correlation follows a Wishart distribution (Huang et al., 2017;
 325 Park and Lord, 2007):

$$\sum^{-1} \sim Wishart(R, d) \quad (9)$$

326 where β_0 , R_{β_0} and R are known non-informative hyper parameters and d is equal to the degrees
 327 of freedom (number of the examined crash types, in this case $d = k = 2$).

328 5. Results and Discussion

329 The models were fitted using WinBUGS software which incorporates full Bayes model
 330 estimation approach using the Markov Chains Monte Carlo (MCMC) method (Spiegelhalter et
 331 al., 2003). Each model was developed with 200,000 iterations of two Markov chains and the
 332 initial 50,000 iterations were discarded from the final model estimates. The actual functional
 333 forms of the relationships between traffic variables and crashes are not known and potential
 334 interactions between traffic variables cannot be ruled out. Therefore, the present study
 335 examined the effects of speed variations using several interaction-term combinations in
 336 addition to the individual traffic variables.

337

338 All the traffic variables along with all their multiplicative interaction combinations² and rain
 339 were taken as explanatory variables in both multivariate models in various combinations. The
 340 final models that are presented here, were chosen based on the lowest DIC (Deviance
 341 Information Criterion) value.

342 The best-fitting models for vehicle type and severity type are presented in terms of posterior
 343 means, standard deviations (SD), MC Error and the 95% credible intervals of the estimated
 344 coefficients in Tables 2 and 3, respectively. The correlations between the crash types in each
 345 model were also calculated and it was found that both the models showed very high correlations
 346 (0.981 and 0.980 for the crash types by vehicle type and by severity levels, respectively). This
 347 suggests that the different crash types are related to each other and should be modelled using
 348 multivariate models. For both the models, the best fitted variable combination included all
 349 traffic and weather variables plus the following interactions: a) volume and speed, b) volume
 350 and within lane speed variation and c) Speed and between-lanes speed variation.

351

352

353

354 **Table 2 Multivariate model results for crash rates by vehicle type (HV and LV crashes)**

HV crashes						
Variables	Mean	SD	MC Error	2.50%	Median	97.50 %

² Possible multiplicative interaction combinations: (i) Volume* Speed, (ii) Volume* Between-lanes speed variation, (iii) Volume* Within lane speed variation (iv) Speed* Between-lanes speed variation (v) Speed* Within lane speed variation (vi) Between-lanes speed variation* Within lane speed variation

Speed	-0.1292	0.016	0.001	-0.161	-0.129	-0.098
Volume	-0.03544	0.008	0.000	-0.051	-0.035	-0.021
Within lane speed variation	-0.4776	0.091	0.002	-0.654	-0.477	-0.300
Between-lanes speed variation	-0.2538	0.087	0.003	-0.420	-0.256	-0.079
Rain	6.537	0.673	0.016	5.357	6.485	7.993
Volume*Speed	0.000183	0.000	0.000	0.000	0.000	0.000
Volume*Within lane speed variation	0.002204	0.000	0.000	0.001	0.002	0.003
Speed*Between-lanes speed variation	0.004118	0.001	0.000	0.002	0.004	0.006
Intercept	-1.11	1.391	0.043	-3.788	-1.147	1.648
Ln(exposure)	1	Vehicle hours travelled				
LV crashes						
Variables	Mean	SD	MC Error	2.50%	Median	97.50 %
Speed	-0.1226	0.015	0.001	-0.152	-0.122	-0.094
Volume	-0.04516	0.007	0.000	-0.061	-0.045	-0.032
Within lane speed variation	-0.4173	0.069	0.002	-0.552	-0.418	-0.282
Between-lanes speed variation	-0.241	0.079	0.003	-0.404	-0.242	-0.085
Rain	7.994	0.595	0.019	6.937	7.958	9.292
Volume*Speed	0.000269	0.000	0.000	0.000	0.000	0.000
Volume*Within lane speed variation	0.002449	0.002	0.000	0.000	0.002	0.003
Speed*Between-lanes speed variation	0.003549	0.004	0.001	0.000	0.002	0.005
Intercept	-1.293	1.323	0.046	-3.902	-1.267	1.319
Ln(exposure)	1	Vehicle hours travelled				
Model performance parameters						
\bar{D}	1353.49					
p_D	198.213					
DIC	1551.7					

355 Note: Boldface indicates statistically significant coefficients at the 95% credible interval.

356 **Table 3 Multivariate model results for crash rates by severity levels (KS and SL crashes)**

KS crashes						
Variables	Mean	SD	MC Error	2.50%	Median	97.50%
Speed	-0.1332	0.02 0	0.001	-0.167	-0.133	-0.096
Volume	-0.04594	0.00 9	0.000	-0.061	-0.046	-0.028
Within lane speed variation	-0.3616	0.09 7	0.002	-0.521	-0.361	-0.174
Between-lanes speed variation	-0.3801	0.13 6	0.005	-0.619	-0.368	-0.137
Rain	6.546	0.71 6	0.017	5.456	6.491	8.088
Volume*Speed	0.000257	0.00 0	0.000	0.000	0.000	0.000
Volume*Within lane speed variation	0.002648	0.00 1	0.000	0.002	0.003	0.004
Speed*Between-lanes speed variation	0.005087	0.00 1	0.000	0.002	0.005	0.008
Intercept	-1.008	1.62 4	0.05152	-3.594	-1.059	2.283
Ln(exposure)	1	Vehicle hours travelled				
SL crashes						
Variables	Mean	SD	MC Error	2.50%	Median	97.50%
Speed	-0.1506	0.01 7	0.001	-0.178	-0.152	-0.116
Volume	-0.03954	0.00 7	0.000	-0.051	-0.039	-0.026
Within lane speed variation	-0.4756	0.07 3	0.002	-0.593	-0.477	-0.331
Between-lanes speed variation	-0.4706	0.10 0	0.004	-0.632	-0.472	-0.284
Rain	8.334	0.69 2	0.023	7.273	8.280	9.800

Volume*Speed	0.000214	0.00 0	0.000	0.000	0.000	0.000
Volume*Within lane speed variation	0.002501	0.00 0	0.000	0.002	0.003	0.003
Speed*Between-lanes speed variation	0.006268	0.00 1	0.000	0.004	0.006	0.008
Intercept	0.7937	1.50 2	0.053	-1.785	0.886	3.433
Ln(exposure)	1	Vehicle hours travelled				
Model performance parameters						
\bar{D}	1334.65					
p_D	198.067					
DIC	1532.71					

357 Note: Boldface indicates statistically significant coefficients at the 95% credible interval.

358

359 As the main aim of the study is to examine the relationships of speed variations with
360 crashes, the discussion focuses on these effects. Both the variations have negative coefficients
361 but, as both are also present in interaction terms, direct interpretation of the individual
362 coefficients is not possible. To facilitate the interpretation of the interaction of volume and
363 within lane speed variation, the crash rates are plotted against the entire range of within lane
364 speed variations and volume in Figures 3A, 3B, 3C and 3D for HV, LV, KS and SL crashes
365 respectively. The effects of other variables are kept constant (at their mean) while estimating
366 crash rates. For example, the equation used for developing the graph for the HV crash model
367 (Figure 3A) is:

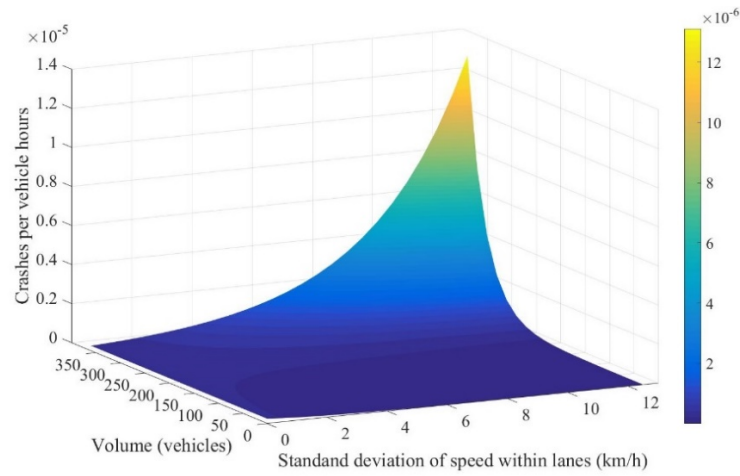
368

$$\begin{aligned}
\frac{HV \text{ Crashes}}{Veh \text{ hours}} = & \exp(-0.1292 \cdot \overline{Speed} - 0.03544 \cdot \overline{Volume} \\
& - 0.4776 \cdot \text{Within lane speed variation} \\
& - 0.2538 \cdot \overline{Between lanes speed variation} + 6.537 \cdot \text{rain} \\
& + 0.000183 \cdot (\overline{Volume} \cdot \overline{Speed}) \\
& + 0.002204 \cdot \text{Volume} \cdot \text{Within lane speed variation} \\
& + 0.004118 \cdot (\overline{Speed} \cdot \overline{Between lanes speed variation}) - 1.11)
\end{aligned} \tag{10}$$

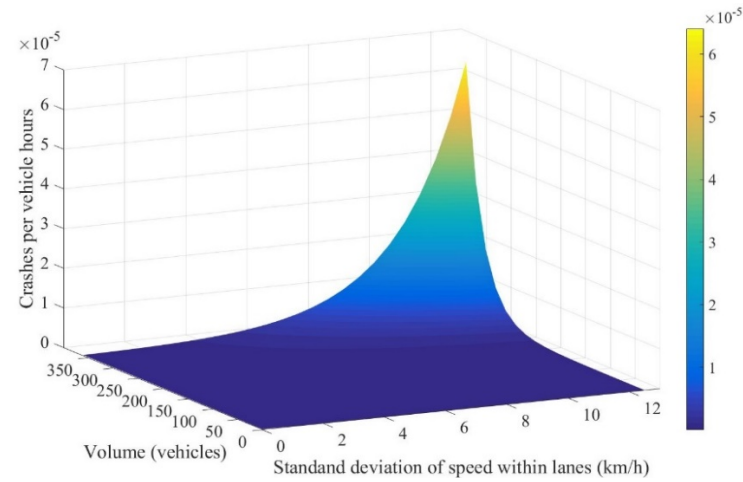
369

370

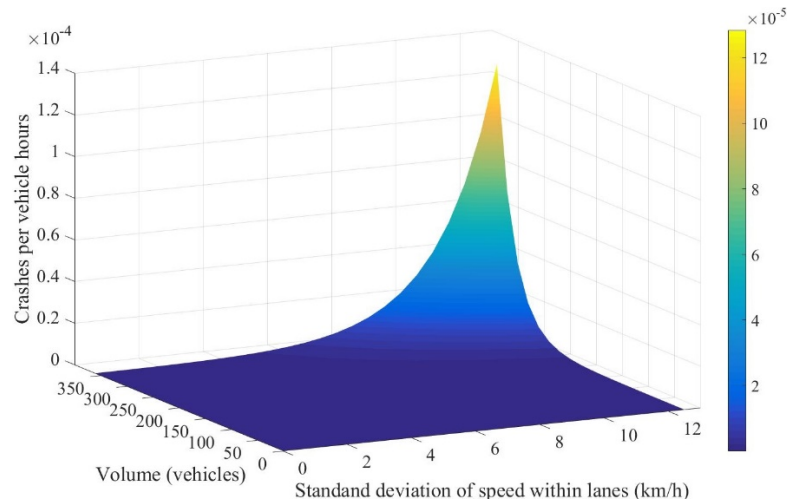
371



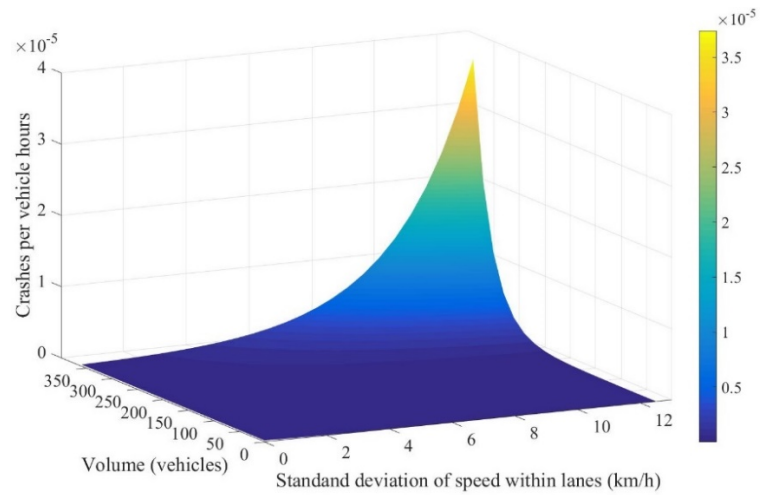
(3A) HV crashes



(3B) LV crashes



(3C) KS crashes



(3D) SL crashes

Figure 3 3d Contour graphs of crashes per vehicle hours as a function of within lane speed variation and volume

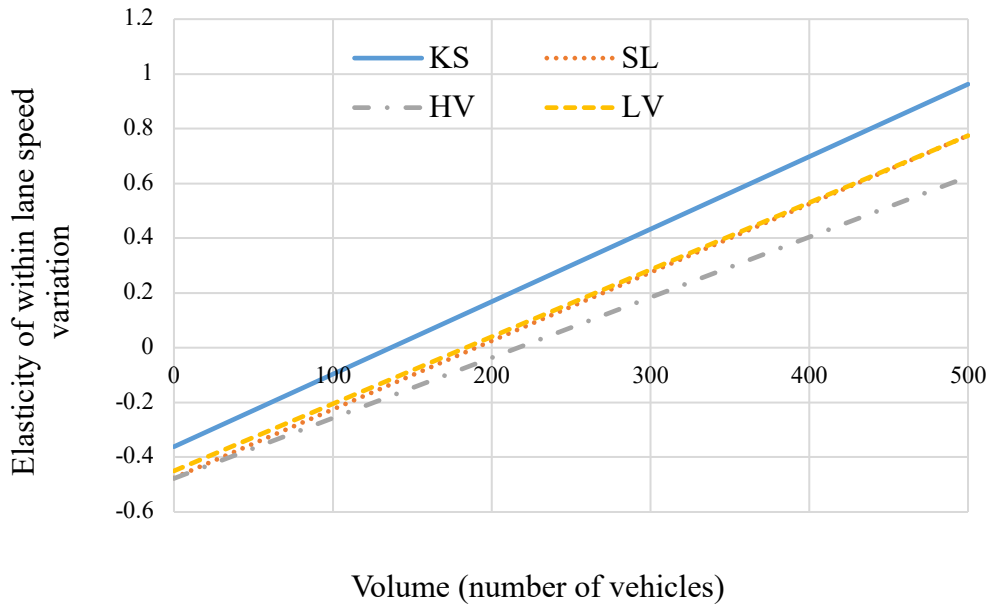
372 The curves show that the relationship of crash rates with the *within lane speed variation*
373 varies according to the volume conditions on the road. More specifically, all crash types seem
374 to be triggered by within lane speed variation at higher volumes. This is more clearly
375 demonstrated by Figure 4A that shows the elasticity of within lane speed variation as a function
376 of traffic volume. The threshold values of volume where the elasticities become positive are
377 216, 183, 132, and 187 for HV, LV, KS, and SL crashes respectively. This means that when
378 traffic volume is higher than these values, increases in within lanes speed variation are likely
379 to lead to more crashes. The present study results are in line with the previous study Garber
380 and Erhart (2000) who showed that high variation in speed results into higher crash rates in the
381 presence of high flow per lane, whereas the lower volume may not affect the crash rate
382 significantly. Further, the present results show that the KS crashes have higher elasticities
383 compared to the SL ones (Figure 4A). One of the possible reasons behind this could be that the
384 route analysed in the study is a freeway, as the literature shows that the crashes on the roads
385 characterised with high speed limits are more prone to severe crashes (Zhu and Srinivasan,
386 2011). In high speed conditions, increase in the within lane speed variance can further worsen
387 the situations in terms of severity.

388 Traffic conditions with high volume and high speed variation within the same lane
389 represent conditions with lower levels of service and therefore unstable flow. These conditions
390 can create higher crash risk because of the limited spacing between vehicles and therefore lower
391 time to react to sudden changes in nearby vehicle speeds (Li et al., 2018; Xu et al., 2016). So,
392 as expected, under these conditions, more coordinated traffic would be safer. On the other hand,
393 the results for low-volume conditions (i.e. lower crash rates during higher within lane speed
394 variation) are less straightforward to explain. Typically, lower volume conditions are mainly
395 associated with free flow conditions with low demand, however, it can also be observed at slow
396 moving conditions due to congestion during the peak periods. A possible explanation for the

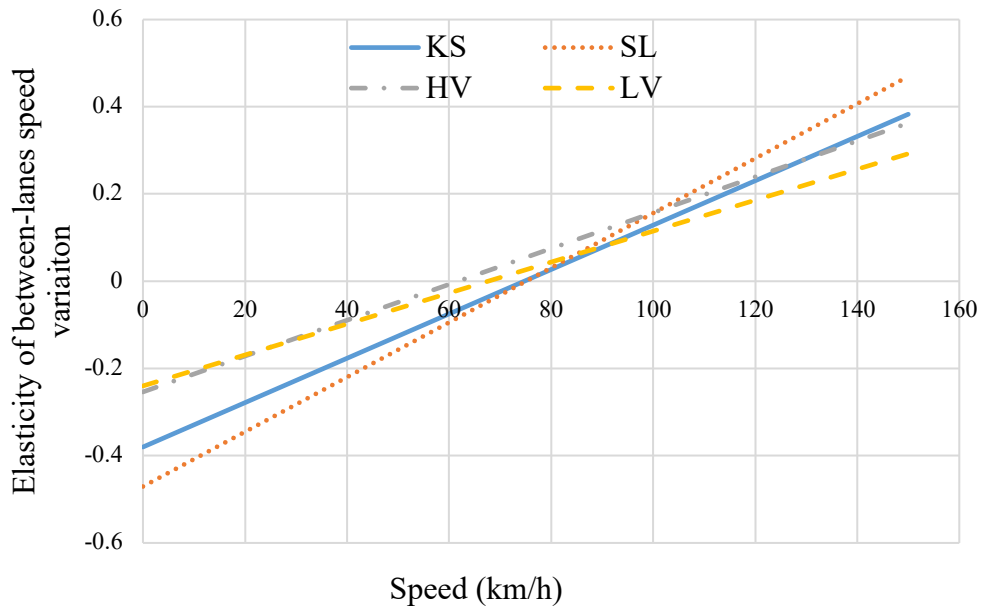
397 first lower volume condition may be that in these conditions, drivers have more freedom to
398 select their comfortable speeds and maintain sufficient spacing from other vehicles. Therefore,
399 even if the differences in speeds within the same lane are high, these do not lead to frequent
400 crash-prone interactions. In the second scenario, slow moving conditions due to congestion
401 during the peak periods, restricts the freedom of the drivers to vary the speed, therefore, it
402 cannot be accounted for the high-speed variance conditions. The results regarding flow and
403 within lane speed variation are consistent with some previous studies which found that crashes
404 happen more in the presence of high-speed variation during congested flow conditions (Golob
405 et al., 2004). The within lane elasticity curve shown in Figure 4A exhibits that an increase in
406 the within lane speed variation and volume will lead to a more sharp increase for KS crashes
407 than the SL crashes. As the higher within lane speed variation relates with the situations of
408 more extreme speeds on the roadway (too slow and too fast), this could be the possible reason
409 for the sharp increase in the crash rate for higher speed variations.

410 Interaction effects of between-lanes speed variation and speed on the HV, LV, KS and
411 SL crash rates are shown in Figure 5A, 5B, 5C and 5D respectively. The shape of the curves
412 (for all crash types) show that the effect of between-lanes speed variation on the crash risk
413 changes in the presence of different average speeds. Figure 4B shows the elasticities of
414 between-lanes speed variations with respect to average speeds. It indicates that the crash risk
415 increases when both the average speed and the between-lanes speed variations are increasing.
416 Specifically, when average speeds are higher than 61, 67, 75 and 75 km/h for HV, LV, KS and
417 SL crashes respectively crash risk is constantly positively associated with increased between-
418 lanes speed variations. In fact, traffic conditions with speeds lower than these thresholds are
419 particularly rare in the study area as in more than 97% of the time the speed is higher than
420 70km/h. Comparing the elasticities, it is observed that between-lanes speed variation cause

421 higher crash risk for HV crashes than the LV crashes; and surprisingly, the SL elasticities are
 422 higher than the KS crashes under these circumstances.



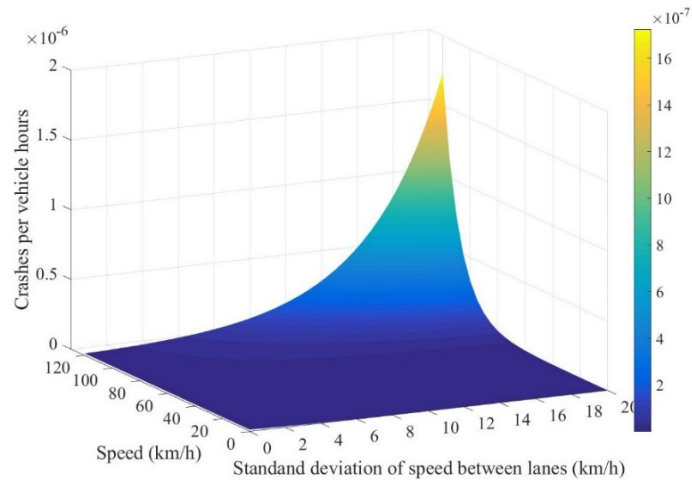
4(A) Elasticity of within lane speed variation across different types of crash rates for a range of volume values



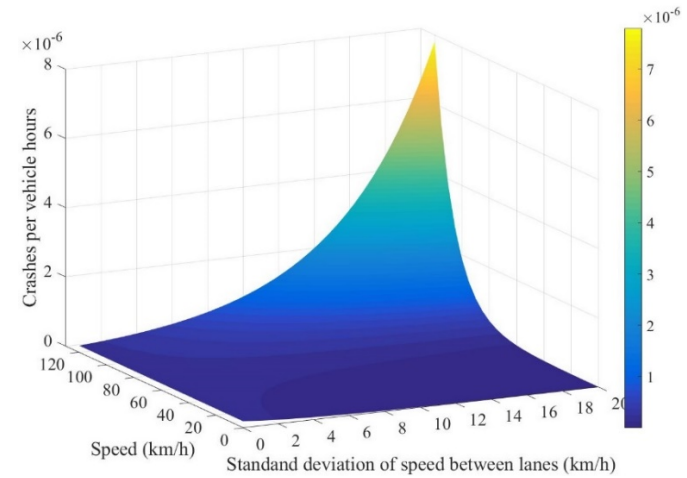
4(B) Elasticity of between-lanes speed variation across different types of crash rates for a range of speed values

423 **Figure 4 Elasticity plots of within lane (4(A)) and between-lanes (4(B)) speed variations**

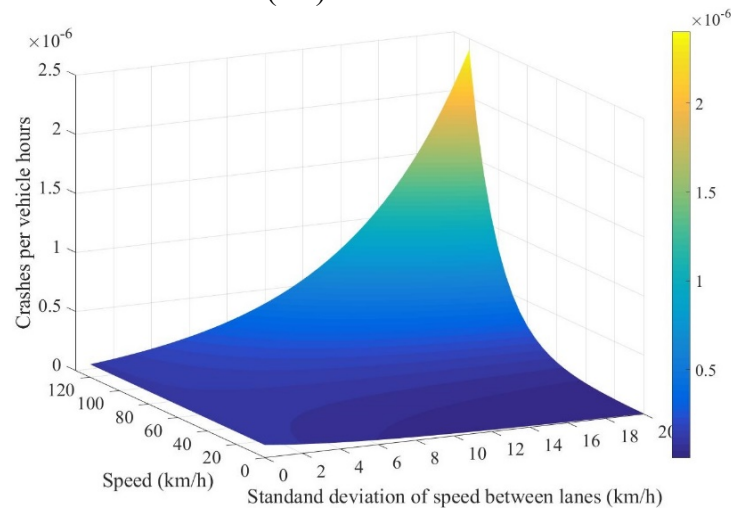
424 The positive relationship between crash rates and the between-lanes speed variation practically
425 at all speed conditions in the study area possibly indicates crashes related to lane changing or
426 overtaking manoeuvres (Ma et al., 2017; Potts et al., 2007; Wang et al., 2017). Overtaking
427 manoeuvres tend to be more frequent under high speed conditions and if manoeuvres are
428 combined with higher speed differences between the lanes, may trigger more side impacts.
429 Higher between-lanes speed variation may be caused by the presence of heavy goods vehicles
430 on the road, which tends to be slower than the rest of the traffic, especially at free-flow
431 conditions. The LV are more likely to change lanes than HV to increase speed because it's
432 easier to manoeuvre for LV, this increases the instances of encounter of LV with the HV.
433 Subsequently the crashes involving multivehicle (HV and LV) increase because of the higher
434 between-lanes speed variation. But as the present the study terms multivehicle crashes as HV
435 crashes if at least one HV is involved in the crash, this can explain the fact that the elasticity
436 for HV crashes is higher than for LV.



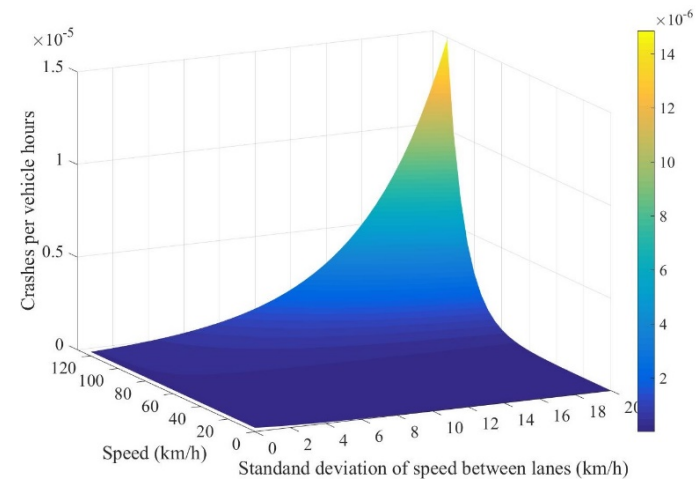
(5A) HV crashes



(5B) LV crashes



(5C) KS crashes



(5D) SL crashes

Figure 5 3d Contour graphs of crashes per vehicle hours as a function of between-lanes speed variation and speed

437 The model results show that the impact of speed on crashes is associated with volume
438 and between-lanes speed variations, which complicates its interpretation. The interactions
439 suggest that under high volume and low between-lanes speed variation, higher speeds are
440 associated with lower crash rates. As the between-lanes speed variation increases though,
441 higher speeds will lead to more crashes. These results extend some previous findings (Garber
442 and Ehrhart, 2000; Kloeden et al., 2002; Tanishita and Wee, 2016) which observed that higher
443 crash rates are observed if higher speeds are coupled with high variation in speed.

444 Aligning with the previous studies, it was shown that the presence of the rain increases
445 crash risk (Abdel-aty and Pemmanaboina, 2006; Naik et al., 2016). The coefficient of the rain
446 variable shows that the effect of rain is higher for the LV crashes when compared to the HV
447 crashes. A possible reason behind higher crash risk for the LV during the rain could be related
448 to the better training of heavy goods vehicle drivers in driving under rainy conditions.
449 Surprisingly the results suggest that rain has higher effects on SL crashes than the KS crashes,
450 which is different from the previous findings but could be explained by the lower speed during
451 driving in rainy weather that might result in less serious crashes.

452 Both the datasets (0-5 minute prior of crashes and 5-10 minute prior of crashes) resulted
453 in similar models in terms of main effects of traffic variables and therefore, for brevity only
454 the first model was presented in this section. The main difference was observed in the weather
455 variable "Rain". More specifically, presence of rain was found to be negatively associated with
456 the probability of a crash occurring in the following 5-10 minutes. This difference in results
457 might imply more careful driving behaviour during rainy period but it can also be attributed to
458 inaccuracies in weather data as some of the weather stations were situated quite far away from
459 some crash locations.

460

461 **6. Conclusions**

462 This study focused on modelling the effects of speed variations on freeway crash rates by
463 vehicle type (HV and LV) and for different levels of severity (KS and SL). Crash data were
464 aggregated following a condition-based data aggregation approach in order to achieve better
465 representation of time-varying variables. The crash frequencies of a three-year period on a
466 segment of M1 motorway were modelled using Multivariate Poisson lognormal regression.
467 The traffic variables along with their interactions and weather variables were investigated for
468 their possible influence on crash risk. All the examined variables were found to have a
469 statistically significant impact on crash rates and the signs of the estimated coefficients were
470 identical for all the four examined crash types. Following are the main contributory findings of
471 the study:

- 472 a) The study results showed that the crash rate increases with increase in the within lane
473 speed variances at higher volume conditions.
- 474 b) The crash rate also increases with increase in the between-lane speed variances at high
475 speed conditions.
- 476 c) The within lane speed variance is identified as a higher risk for LV crashes than the HV
477 crashes and the chances of KS crashes are higher than the SL crashes.
- 478 d) Whereas, the between-lane speed variance is related with higher crash risk for HV
479 crashes than the LV crashes.

480 Overall, the results suggest that the speed and its variations are not solely responsible
481 for the higher crash rates, but the combination of specific traffic conditions play an important
482 role in crash occurrences. Additionally, the results show that the speed variation should be
483 considered in two different dimensions (between-lanes and within lane) to better interpret the
484 crash triggering situations and to develop better and more precise safety measures.

485 These results could be helpful in understanding crash risk at different traffic conditions
486 and to that end in the development of more efficient countermeasures for traffic management
487 agencies and the road freight industry. The outcomes of this study could also contribute to the
488 design of in-vehicle crash warning systems applicable to both commercial and private vehicles.

489 As this analysis focused on a busy freeway section that does not include extreme
490 geometry, in order to generalise the outcomes of the models it could be beneficial to consider
491 a larger and more diverse road network and to incorporate geometric data in the models.
492 Additionally, the present study did not analyse the PDO crashes, therefore, further research
493 should include PDO crashes, so that the results can be generalised for crashes of all severity
494 types. The current study also does not examine differences in single and multi-vehicle crashes
495 separately owing to the limited number of single vehicle crashes in the study area. Therefore,
496 a future study is required to obtain more insights into the impacts of speed variations on
497 different collision types.

498 **Acknowledgement**

499 We would like to thank Highways England and Mott Mac Donald for providing us
500 access to Motorway Incident Detection and Automatic Signalling database (MIDAS) for
501 conducting this research project.

502 This research is partially supported by the European Union Europe Aid-funded Project
503 “EU-India Research & Innovation Partnership for Efficient and Sustainable Freight
504 Transportation (REINVEST),” Contract Number: R/141842. The contents of this publication
505 are the sole responsibility of the authors of this paper and can in no way be taken to reflect the
506 views of the European Union.

507 **References**

- 508 Aarts, L. and Schagen, I. Van (2006) 'Driving speed and the risk of road crashes : A review',
 509 *Accid. Anal. Prev.*, 38, 215–224.
- 510 Abdel-Aty, M. A., Hassan, H. M. and Ahmed, M. (2012) 'Real-Time Analysis of Visibility
 511 Related Crashes: Can Loop Detector and AVI Data Predict Them Equally?', in *Transp. Res.
 512 Board 91st Annu. Meet.*
- 513 Abdel-aty, M. A. and Pemmanaboina, R. (2006) 'Calibrating a Real-Time Traffic Crash-
 514 Prediction Model Using Archived Weather and ITS Traffic Data', *IEEE Trans. Intell.
 515 Transp. Syst.*, 7(2), 167–174.
- 516 Chen, F. and Chen, S., 2011. 'Reliability-based assessment of vehicle safety in adverse driving
 517 conditions.' *Transp. Res. Part C*, 19(1), 156-168
- 518 Department for Transport (2011) *STATS19 road accident injury statistics*.
- 519 Farah, H. and Koutsopoulos, H. N. (2014) 'Do cooperative systems make drivers? car-
 520 following behavior safer?', *Transp. Res. part C Emerg. Technol.*, 41, 61–72.
- 521 Garber, N. J. and Ehrhart, A. A. (2000) 'Effect of Speed , Flow , and Geometric Characteristics
 522 on Crash Frequency for Two-Lane Highways', *Transp. Res. Rec. J. Transp. Res. Board*,
 523 1717, 76–83.
- 524 Garber, N. J. and Gadiraju, R. (1989) 'Factors Affecting Speed Variance and Its Influence on
 525 Accidents', *Transp. Res. Rec. J. Transp. Res. Board*, 1213, 64–71.
- 526 Golob, T. F., Recker, W. W. and Alvarez, V. M. (2004) 'Freeway safety as a function of traffic
 527 flow', *Accid. Anal. Prev.*, 36, 933–946.
- 528 Google maps (2017), <https://www.google.co.in/maps> Accessed Jul. 31, 2017
- 529 Highways England (2017) *MIDAS Data* <https://www.midas-data.org.uk/> Accessed Jul. 3,
 530 2017.

- 531 Highways England (2017) *MIDAS Data*. <https://www.midas-data.org.uk/> Accessed Jul. 3,
532 2017.
- 533 Huang, H., Zhou, H., Wang, J., Chang, F. and Ma, M. (2017) ‘Analytic Methods in Accident
534 Research A multivariate spatial model of crash frequency by transportation modes for urban
535 intersections’, *Anal. Methods Accid. Res.*, 14, 10–21.
- 536 Imprialou, M.-I. (2015) *Developing Accident-Speed Relationships Using a New Modelling
537 Approach*. PhD thesis, Loughborough University, United Kingdom.
- 538 Imprialou, M.-I. M., Quddus, M. and Pitfield, D. E. (2016a) ‘Predicting the safety impact of a
539 speed limit increase using condition-based multivariate Poisson lognormal regression’,
540 *Transp. Plan. Technol.*, 39(1), 3–23.
- 541 Imprialou, M.-I. M., Quddus, M., Pitfield, D. E. and Lord, D. (2016b) ‘Re-visiting crash-speed
542 relationships: A new perspective in crash modelling’, *Accid. Anal. Prev.*, 86, 173–185.
- 543 Kloeden, C. N., McLean, J. and Glonek, G. F. V. (2002) *Reanalysis of travelling speed and the
544 risk of crash involvement in Adelaide South Australia*. Road safety research report
545 (Australian Transport Safety Bureau), CR207, Australian Transport Safety Bureau.
- 546 Kockelman, K. M. and Ma, J. (2007) ‘Freeway Speeds and Speed Variations Preceding Crashes
547 , Within and Across Lanes’, *J. Transp. Res. Forum*, 46(1), 43–61.
- 548 Kweon, Y. and Kockelman, K. M. (1996) ‘Safety Effects of Speed Limit Changes Use of Panel
549 Models , Including Speed , Use , and Design Variables’, *Transp. Res. Rec. J. Transp. Res.
550 Board*.
- 551 Lee, C., Hellinga, B. and Saccomanno, F. (2003) ‘Real-time crash prediction model for
552 application to crash prevention in freeway traffic’, *Transp. Res. Rec. J. Transp. Res. Board*,
553 (1840), 67–77.
- 554 Lemp, J.D., Kockelman, K.M. and Unnikrishnan, A., 2011. ‘Analysis of large truck crash
555 severity using heteroskedastic ordered probit models’ *Accid. Anal. Prev.* 43(1), 370-380.

- 556 Levy, D. T. and Asch, P. (1989) 'Speeding, coordination, and the 55-mph limit: Comment',
557 *Am. Econ. Rev.*, 79(4), 913–915.
- 558 Li, Y., Zheng, Y., Wang, J., Kodaka, K. and Li, K. (2018) 'Crash probability estimation via
559 quantifying driver hazard perception', *Accid. Anal. Prev.*, 116(May 2017), 116–125.
- 560 Lord, D. and Mannering, F. (2010) 'The statistical analysis of crash-frequency data : A review
561 and assessment of methodological alternatives', *Transp. Res. Part A*, 44(5), 291–305.
- 562 Ma, J. (2006) 'Bayesian multivariate Poisson-Lognormal regression for crash prediction on
563 rural two-lane highways'. PhD thesis, University of Texas, Austin
- 564 Ma, Z., Zhang, H., Steven, I., Chien, J., Wang, J. and Dong, C. (2017) 'Predicting expressway
565 crash frequency using a random effect negative binomial model: A case study in China',
566 *Accid. Anal. Prev.*, 98, 214–222.
- 567 Martensen, H. and Dupont, E. (2013) 'Comparing single vehicle and multivehicle fatal road
568 crashes : A joint analysis of road conditions , time variables and driver characteristics',
569 *Accid. Anal. Prev.*, 60, 466–471.
- 570 MetOffice (2016) *UK climate - Historic station data*. Accessed Jul. 2, 2016
- 571 Naik, B., Tung, L., Zhao, S. and Khattak, J. (2016) 'Weather impacts on single-vehicle truck
572 crash injury severity', *J. Safety Res.*, 58, 57–65.
- 573 Park, E. and Lord, D. (2007) 'Multivariate Poisson-lognormal models for jointly modeling
574 crash frequency by severity', *Transp. Res. Rec. J. Transp. Res. Board*, (2019), 1–6.
- 575 Potts, I., Harwood, D. and Richard, K. (2007) 'Relationship of lane width to safety on urban
576 and suburban arterials', *Transp. Res. Rec. J. Transp. Res. Board*, (2023), 63–82.
- 577 Quddus, M. (2013) 'Exploring the Relationship Between Average Speed , Speed Variation ,
578 and Accident Rates Using Spatial Statistical Models and GIS', *J. Transp. Saf. Secur.*, 5(1),
579 27–45.
- 580 Shi, Q., Abdel-Aty, M. and Yu, R. (2016) 'Multi-level Bayesian safety analysis with

- 581 unprocessed Automatic Vehicle Identification data for an urban expressway', *Accid. Anal.*
582 *Prev.*, 88, 68–76.
- 583 Solomon, D. (1964) 'Accidents on main rural highways related to speed, driver, and vehicle'.
- 584 Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003) 'WinBUGS user manual'.
- 585 Strömberg, P. and Lind, G. (2016) 'Harmonization with variable speed limits on motorways',
586 *Transp. Res. Procedia*, 15, 664–675.
- 587 Stuster, J. (2004) *Aggressive Driving Enforcement: Evaluations of Two Demonstration*
588 *Programs*.
- 589 Tanishita, M. and Wee, B. Van (2016) 'Impact of vehicle speeds and changes in mean speeds
590 on per vehicle-kilometer traffic accident rates in Japan', *IATSSR*, 0–5.
- 591 Taylor, M. C., Lynam, D. A. and Baruya, A. (2000) 'The effect of drivers' speed on the
592 frequency of road accidents', *Transp. Res. Lab.*, 56.
- 593 Wang, L., Abdel-Aty, M. and Lee, J. (2017) 'Safety analytics for integrating crash frequency
594 and real-time risk modeling for expressways', *Accid. Anal. Prev.*, 104, 58–64.
- 595 Wang, L., Abdel-Aty, M., Shi, Q. and Park, J. (2015a) 'Real-time crash prediction for
596 expressway weaving segments', *Transp. Res. Part C*, 61, 1–10.
- 597 Wang, L., Shi, Q. and Abdel-Aty, M. (2015b) 'Predicting crashes on expressway ramps with
598 real-time traffic and weather data', *Transp. Res. Rec. J. Transp. Res. Board*, (2514), 32–38.
- 599 Wang, X., Yu, R. and Zhong, C. (2016) 'A field investigation of red-light-running in Shanghai,
600 China', *Transp. Res. Part F Traffic Psychol. Behav.*, 37, 144–153.
- 601 Wang, X., Zhou, Q., Quddus, M., Fan, T. and Fang, S. (2018) 'Speed, speed variation and crash
602 relationships for urban arterials', *Accid. Anal. Prev.*, 113(November 2016), 236–243.
- 603 Wei, Z., Xiaokun, W. and Dapeng, Z. (2017) 'Truck crash severity in New York city: An
604 investigation of the spatial and the time of day effects', *Accid. Anal. Prev.*, 99, 249–261.
- 605 Xu, C., Liu, P. and Wang, W. (2016) 'Evaluation of the predictability of real-time crash risk

- 606 models', *Accid. Anal. Prev.*, 94, 207–215.
- 607 Xu, P., Huang, H. and Dong, N. (2018) 'The modifiable areal unit problem in traffic safety:
608 Basic issue, potential solutions and future research', *J. Traffic Transp. Eng. (English Ed.)*,
609 5(1), 73–82.
- 610 Yu, R. and Abdel-Aty, M. (2014) 'Using hierarchical Bayesian binary probit models to analyze
611 crash injury severity on high speed facilities with real-time traffic data', *Accid. Anal. Prev.*,
612 62, 161–167.
- 613 Yu, R., Quddus, M., Wang, X. and Yang, K. (2018) 'Impact of data aggregation approaches
614 on the relationships between operating speed and traffic safety', *Accid. Anal. Prev.*, 120,
615 304–310.
- 616 Zou, W., Wang, X. and Zhang, D., 2017. 'Truck crash severity in New York city: an
617 investigation of the spatial and the time of day effects' *Accid. Anal. Prev.* 99, 249-261.
- 618 Zhu, X. and Srinivasan, S., 2011. 'A comprehensive analysis of factors influencing the injury
619 severity of large-truck crashes' *Accid. Anal. Prev.*, 43(1), 49-57.