

This item was submitted to [Loughborough's Research Repository](#) by the author.  
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

## Impact of traffic congestion on road accidents: a spatial analysis of the M25 motorway in England

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1016/j.aap.2009.04.002>

PUBLISHER

© Elsevier

VERSION

AM (Accepted Manuscript)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Wang, Chao, Mohammed Quddus, and Stephen Ison. 2009. "Impact of Traffic Congestion on Road Accidents: A Spatial Analysis of the M25 Motorway in England". Loughborough University.  
<https://hdl.handle.net/2134/5310>.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

# Impact of Traffic Congestion on Road Safety: A Spatial Analysis of the M25 Motorway in England

Chao Wang, Mohammed A Quddus\*, Stephen G Ison

Transport Studies Group  
Department of Civil and Building Engineering  
Loughborough University  
Loughborough, Leicestershire  
LE11 3TU  
UNITED KINGDOM

\* Corresponding author – email: [m.a.quddus@lboro.ac.uk](mailto:m.a.quddus@lboro.ac.uk)

## ABSTRACT

Traffic congestion and road accidents are two external costs of transport and the reduction of their impacts is often one of the primary objectives for transport policy makers. The relationship between traffic congestion and road accidents however is not apparent and less studied. It is speculated that there may be an inverse relationship between traffic congestion and road accidents, and as such this poses a potential dilemma for transport policy makers. This study aims to explore the impact of traffic congestion on the frequency of road accidents using a spatial analysis approach, while controlling for other relevant factors that may affect road accidents. The M25 London orbital motorway, divided into 70 segments, was chosen to conduct this study and relevant data on road accidents, traffic and road characteristics were collected. A robust technique has been developed to map M25 accidents onto its segments. Since existing studies have often used a proxy to measure the level of congestion, this study has employed a precise congestion measurement. A series of Poisson based non-spatial (such as Poisson-lognormal and Poisson-gamma) and spatial (Poisson-lognormal with conditional autoregressive priors) models have been used to account for the effects of both heterogeneity and spatial correlation.

The results suggest that traffic congestion has little or no impact on the frequency of road accidents on the M25 motorway. All other relevant factors have provided results consistent with existing studies.

**Keywords:** *GIS, traffic congestion, road accidents, spatial modelling and M25 motorway*

## 1 INTRODUCTION

Traffic congestion and road accidents are two important externalities created by road users. Increased travel time caused by traffic congestion imposes social costs to road users, both in terms of economic loss and also the reduced quality of life and mobility. The costs of road traffic accidents to individuals, property, and society in general have also been significant.

Traffic congestion and accidents both impose a burden to society, and as such it is important to reduce their impacts. An ideal solution would be to reduce them simultaneously but this may not be possible, however, since it is speculated that there may be an inverse relationship between traffic congestion and road safety (Shefer and Rietveld, 1997). Shefer and Rietveld (1997) hypothesize that in a less congested road network, the average speed of traffic would be normally high, which is likely to result in more serious injuries or fatalities. On the other hand, in a congested road network, traffic would be slower and may cause less fatalities and serious injuries. This increased traffic congestion may lead to more accidents due to increased traffic volume; however, those accidents may be less severe. This suggests that the total external cost of accidents may be less in a congested situation relative to an uncongested situation. This poses a potential dilemma for transport policy makers since it would appear that traffic congestion can improve road safety; however, traffic congestion reduces mobility which subsequently decreases economic productivity.

It is, therefore, important to understand the association between traffic congestion and road safety so that effective policies can be implemented to control both congestion and road safety. There are few studies in this area and the studies which exist tend to use an analytical approach and a weak proxy for traffic congestion. As such, more robust empirical evidence, and a precise congestion measurement, are required.

The aim of this paper is to explore the effects of traffic congestion on road safety using a spatial analysis approach while controlling for the other contributing factors. The M25 London orbital motorway was used as a case study and disaggregated into 70 road segments. Accident data was obtained from the STATS19 national road accident database, and an appropriate method was developed to assign M25 accidents to the 70 road segments. Traffic characteristic data such as traffic delay, traffic flow and average travel speed for each road segment were obtained from the UK Highways Agency (UKHA<sup>1</sup>). For a precise measurement of the level of congestion on each road segment, a congestion index proposed by Taylor et al. (2000) was employed. The count of the number of traffic accidents for each road segment is viewed as a function of various contributing factors such as the congestion index, traffic flow, traffic speed, road geometry (e.g., radius of curvature and gradient) and road segment length. Several statistical models such as a Poisson-lognormal model, a Poisson-gamma model and a Poisson-lognormal model with conditional autoregressive (CAR) priors are used.

The rest of the paper is structured as follows. First, a summary of current literature on the effect of traffic congestion on road safety and statistical models is presented. This is followed by a description of the data used in this study including the method to assign M25 accidents to the correct road segments. Statistical models used in this study are then discussed followed by a presentation and discussion of the results. Finally, conclusions and recommendations for further research are discussed.

---

<sup>1</sup> <http://www.highways.gov.uk>

## 2 LITERATURE REVIEW

Although the relationship between traffic congestion and road safety is important, there would appear to be a dearth of literature, especially in terms of appropriate quantitative evidence. There is, however, analytical and empirical evidence. For example, Shefer (1994) proposed the hypothesis that there is an inverse relationship between congestion and road fatalities, in which volume over capacity ratio (V/C) was used as a proxy to measure the level of congestion. A further study by Shefer and Rietveld (1997) investigated the link between congestion and safety on highways. They use a similar hypothesis and provide empirical evidence by comparing fatality rates throughout the day finding that during peak hours the fatality rate is obviously lower than other times of the day. Due to data unavailability they examined a proposed model by using a simulated dataset rather than real-world data.

These studies used traffic density as a simple proxy for congestion, which may not represent congestion characteristics properly. Congestion and traffic density are not equivalent concepts and it is unclear how the congestion level evolves with respect to density. It was suggested that a V/C value greater than 0.77 is viewed as congested (Boarnet et al., 1998). It is likely that congestion increases are not proportional to density increases. Additionally, their hypothesis needs to be confirmed with real-world data.

A recent study on the effects of traffic congestion on road safety was conducted by Noland and Quddus (2005), who investigate congestion and safety in London using a spatial analysis approach. London was disaggregated into 15 366 area-wide spatial units and a series of Negative Binomial models were used for analysing peak time and off-peak time accidents, while controlling for other contributing factors. Congestion levels were measured using several proxy variables including an indicator variable for Inner and Outer London (spatially), proximate employment and employment density. Their results are indeterminate and the proxy variables for congestion are generally statistically insignificant in their models, suggesting that there is little effect of traffic congestion on road safety. This may be due to, as the authors suggest, the weakness of the proxies used for congestion as congestion can be highly localized and time-of-day specific, as such a more precise congestion measurement should be used to better understand the effects of congestion on safety. In addition, this study is based on urban conditions, and the results may be different on high speed roads (e.g. motorway).

Besides these studies aiming to investigate the effect of traffic congestion on road safety, other studies provide evidence from various other aspects. For example, by investigating single and multi-vehicle highway crash rates, Ivan et al. (2000) found that the morning peak period is the safest time. A study on crash rates on French motorways (Martin, 2002) reveals that crash rates are highest in light traffic compared to heavy traffic and night-time and light-traffic hour crashes are much worse in terms of serious crashes, hence light traffic is a safety problem both in terms of crash rate and severity. Shinar and Compton (2004) investigated accidents through drivers' behaviour and they found that a linear relationship exists between congestion and the frequency of aggressive behaviours which may affect road safety.

In terms of statistical models used in spatial analysis, Poisson based models are often used to establish a relationship between traffic accident frequency and factors that contribute to accident occurrence as they are suitable for count data. The Poisson based models include the Poisson-lognormal model (e.g. Lord and Miranda-Moreno, 2007; Aguerro-Valverde and Jovanis, 2008); and the Poisson-gamma model (also known as Negative Binomial model) which is widely used in accident analysis (e.g. Shankar et al, 1995; Milton and Mannering, 1998; Abdel-Aty and Radwan, 2000; Lord, 2000; Amoros et al., 2003; Noland and Quddus,

2004; Kim et al., 2006; Haynes et al., 2007). Poisson-lognormal and Poisson-gamma models have the advantage in that they can accommodate overdispersion in accident data.

Recent studies have developed methods using spatial econometrics to address the issues of unmeasured spatial correlation among neighbouring spatial units. Such studies were primarily based on a Bayesian framework in which Conditional Autoregressive (CAR) models are often employed to take into account spatial dependence among neighbouring spatial units. This method was initially used in ecological analysis and disease mapping (e.g. Ghosh et al., 1999; MacNab and Dean, 2001), and then recently in road accident analysis (Miaou et al., 2003; Aguero-Valverde and Jovanis, 2006; Li et al., 2007; Aguero-Valverde and Jovanis 2008). According to Aguero-Valverde and Jovanis (2006 and 2008), compared with non-spatial Poisson models, spatial models can generally produce consistent results and in some cases can better fit the data.

This study aims to address the gap mentioned above by using the M25 motorway in England as a case study, in which relevant data is obtained, a precise congestion index is employed and a series of statistical models are tested.

### 3 DATA DESCRIPTION

In this study the M25 motorway in England was chosen to conduct this research. The M25 motorway is an orbital motorway which encircles London, and is one of the busiest motorways in Europe. Hourly traffic characteristic data for the year 2006, such as average travel time, average travel speed, traffic flow and total vehicle delay for road segments of the M25 were obtained from the UK Highways Agency. According to the Highways Agency, 72 road segments with both directions are identified on the M25 and a road segment starts or ends at a junction. Road segment infrastructure data, such as direction (clockwise and anticlockwise), number of lanes, segment length, radius of curvature and gradient were also obtained from the same source.

Accident data were derived from the STATS19 UK national road accident database. In order to avoid a lot of motorway segments with zero or low accident counts, especially for the case of fatal and serious accidents, STATS19 data for 2004 to 2006 were aggregated. This can also ease the variability of accident frequency from year to year and this is also a common practice used in the previous studies (e.g. Milton and Mannering, 1998; Abdel-Aty and Radwan, 2000; Graham and Glaister, 2003; Haynes et al., 2007). Only accidents recorded as occurring on the M25 motorway are retained. Accidents coded as junction accidents (about 15% of total accidents) were also excluded from the analysis. This is because motorway junctions are complicated in terms of road design (such as fly-overs and slip roads) compared to road segments and it is also difficult to obtain a measure of traffic flow at motorway junctions. In the STATS19 database, accident data are provided with reference to a location measured as points (the easting and northing coordinates in a local British National Grid Coordinate system). While this accident data is overlaid onto spatial motorway segment (centre-line) data, mismatches between them are observed. This is due to the error in both accident data and motorway segment data. An appropriate method is needed for assigning accidents to the correct road segments. This is to ensure that the counts of the number of accidents for each segment are correct. Otherwise, the modelling results would be misleading. This method is briefly discussed.

### 3.1 Mapping accidents to the correct motorway segments

As discussed, due to the error in accident location and the fact that accident data and spatial motorway network data are obtained from different sources, it is highly likely that there are mismatches. This is shown in Figure 1.

**Insert figure 1 here**

In Figure 1, the dots show the locations where accidents occurred. Two solid lines (AB and CD) represent the centre-line of two carriageways of the motorway and the dotted line denotes the central barrier of the carriageways. If M refers to the location of an accident then a robust method is required to assign this accident correctly either on link AB or link CD. One can employ two variables from the available information: (1) the perpendicular distance from the accident point to the segment and (2) the angular difference (assuming  $\Delta\theta$ ) between the direction of the vehicle just before the accident and the link direction. The perpendicular distance and the link direction can be obtained from the coordinates of the start and end nodes of a segment and the direction of the vehicle just before the accident can be obtained from STATS19 data. A segment is more likely to be the correct segment if the distance is short and the angular difference is small. Therefore, a weighting score ( $WS$ ) is developed based on these two factors:

$$WS_i = \frac{1}{d_i} + \cos(\Delta\theta_i) \quad d_i \neq 0$$

where  $d_i$  is the perpendicular distance (in metres) from an accident point to a road segment,  $i$ , and  $\Delta\theta_i$  is the angular difference between the direction of an accident and the direction of a link  $i$  ( $0 - 180^\circ$ ). The minimum value of  $d_i$  is set to be 1 metre and the  $WS_i$  for a segment ranges from -1 to +2. If the  $WS$  for a segment is high then it is considered as the correct segment.

In this study, if  $WS_1$  was significantly greater than  $WS_2$  (i.e.  $WS_1 - WS_2 > 0.3^2$ ) then the accident was assigned to link 1. In the case where there was no significant difference between  $WS_1$  and  $WS_2$  (i.e.  $|WS_1 - WS_2| \leq 0.3$ ) then the accidents are assigned randomly such as all assigned to the segments with the clockwise direction. There were about 2% such accidents in our data. In order to investigate the impact of randomly assigned accidents, a sensitivity analysis was conducted and no significant difference in the modelling results was found.

### 3.2 Congestion Index

The primary objective of this study is to investigate the effects of traffic congestion on road safety. Previous studies (Shefer and Rietveld, 1997; Noland and Quddus, 2005) often use a weak proxy for congestion and so a more precise congestion measurement is required. In this study, we employed the following equation proposed by Taylor et al. (2000) to estimate segment-level traffic congestion. This is known as the congestion index (CI):

$$CI = (T - T_0) / T_0 \quad (T_0 \neq 0)$$

---

<sup>2</sup> 0.3 is equal to 10% of the range of  $WS$ . Other threshold values such as 0.15 and 0.45 (i.e. 5% and 15% of the range of  $WS$  respectively) for the difference between  $WS_1$  and  $WS_2$  have also been tested. The results are not significantly different.

where  $T$  is the actual travel time and  $T_0$  is the free flow travel time on a particular road segment. CI is dimensionless and independent of road segment length or road geometry, so it can be compared between different road segments. CI is non-negative and the higher the CI value the higher the level of congestion. Vehicle delay and average travel time data are available from the UKHA. Free flow travel time is calculated by average travel time minus average vehicle delay (weighted by traffic flow) for the year 2006.

Summary statistics of the M25 motorway accidents at the segment-level, traffic characteristic and road infrastructure data are presented in Table 1:

**Insert table 1 here**

As can be seen in Table 1, there are 70 observations for each variable. Two road segments were excluded from the total 72 road segments because data for some key variables for these two segments was missing. Average vehicle speed is weighted (by traffic flow) harmonic mean of hourly speed data. Direction is a dummy variable with 0 representing the anti-clockwise and 1 representing the clockwise direction.

The spatial distribution of CI and segment-level fatal and serious injury accidents is shown in Figure 2. Note that the road segments in different directions (i.e. clockwise and anticlockwise) are close to each other and therefore, they seem to merge together in Figure 2. This figure gives a visual comparison between traffic congestion and accidents. It is noticeable that the southern segments of the M25 show a low level of traffic congestion (Figure 2a) but a high level of fatal and serious injury accidents (Figure 2b) suggesting that there may be an inverse relationship between traffic congestion and fatal and serious injury accidents.

**Insert figure 2 here**

A further analysis of data has been conducted to see whether there is a relationship between interested variables. This is presented in Figures 3 and 4.

**Insert figures 3 and 4 here**

Generally, a positive relationship can be noticed between traffic accidents and AADT (Figure 3). However, no clear association can be seen between traffic accidents and the congestion index (CI) as indicated in Figure 4.

## 4 STATISTICAL MODELS

Several statistical models suitable for count data have been considered and investigated in this study. Poisson based models, such as a Poisson-lognormal, a Poisson-gamma and a Poisson-lognormal with conditional autoregressive (CAR) priors have been recommended and widely used in the literature (e.g. Shankar et al, 1995; Milton and Mannering, 1998, Abdel-Aty and Radwan, 2000; Lord, 2000; Amoros et al., 2003; Miaou et al., 2003; Kim et al., 2006; Lord and Miranda-Moreno, 2007; Aguero-Valverde and Jovanis 2008; Quddus, 2008). The base form of Poisson based model can be expressed as follows:

$$Y_i \sim \text{Poisson}(\mu_i)$$

$$\log(\mu_i) = \alpha + \beta \mathbf{X}_i + v_i + u_i$$

where,

$Y_i$  is the observed number of accidents occurred on road segment  $i$ ;

$\mu_i$  is the expected Poisson accident rate at road segment  $i$ ;

$\alpha$  is the intercept;

$\mathbf{X}_i$  is the vector of explanatory variables for road segment  $i$ ;

$\boldsymbol{\beta}$  is the vector of coefficients to be estimated;

$v_i$  is a random term which captures the heterogeneity effects for road segment  $i$ ;

$u_i$  is a random term which captures the spatially correlated effects for road segment  $i$ .

All models are estimated under a full hierarchical Bayesian framework by using a software package WinBUGS (Spiegelhalte et al., 2003). Models are differentiated by different specifications of the random terms (i.e.  $v_i$  and  $u_i$ ). The specification of each model is as follows:

**Poisson-lognormal model:** It is of interest to test the model with heterogeneity effects only, so the spatially correlated effects term  $u_i$  is excluded in this model. Model specification follows the recommendations used in the WinBUGS user manual (Spiegelhalte et al., 2003). A uniform prior distribution is assigned to  $\alpha$ ; a highly non-informative normal prior is assigned to all  $\beta$ 's with zero mean and 100,000 variance. The prior distribution for uncorrelated heterogeneity term  $v_i$  is a normal prior with  $N(0, \tau_v^2)$ , where  $\tau_v^2$  is the precision (1/variance) with a vague gamma prior  $Gamma(0.5, 0.0005)$ . Note under this parameterization a gamma distribution  $\theta \sim Gamma(a, b)$  is defined with mean  $E(\theta) = a/b$  and variance  $Var(\theta) = a/b^2$ .

**Poisson-gamma model:** Similar to Poisson-lognormal model, the spatial correlation term ( $u_i$ ) is excluded in this model. Also, the same prior distributions are assigned to  $\alpha$  and  $\beta$ 's. The term,  $\exp(v_i)$ , is assigned to a gamma prior i.e.,  $\exp(v_i) \sim Gamma(\phi, \phi)$ , where  $\phi$  is assigned to a non-vague hyper prior with  $Gamma(0.1, 1.0)$  as suggested by Lord and Miranda-Moreno (2007).

**Poisson-lognormal CAR model:** This model accommodates both heterogeneity and spatial correlation effects (i.e.  $v_i$  and  $u_i$ ). The same priors are assigned to  $\alpha$ ,  $\beta$ 's and  $v_i$  as in the Poisson-lognormal model. The spatial correlation term  $u_i$  is modelled with a conditional autoregressive (CAR) model proposed by Besag (1974):

$$u_i | u_j, i \neq j \sim N\left(\frac{\sum_j u_j w_{ij}}{w_{i+}}, \frac{\tau_u^2}{w_{i+}}\right),$$

where  $w_{ij}$  denotes the weight between road segment  $i$  and  $j$ ;  $w_{i+} = \sum_j w_{ij}$ ; and  $\tau_u^2$  is a scale parameter assumed as a gamma prior  $Gamma(0.5, 0.0005)$ .

There are several methods to define the weights ( $w_{ij}$ ) between road segments depending on the consideration of different neighbour structures. The weighting scheme could use contiguity based weights, for example,  $w_{ij} = 1$  if spatial unit  $i$  and  $j$  are adjacent (i.e. shared border and/or vertex) and  $w_{ij} = 0$  otherwise. Alternatively, distance based weights can be used, for example, the shorter the distance between  $i$  and  $j$ , the larger the weight ( $w_{ij}$ ). In this study contiguity based weights are employed, and as suggested by Aguero-Valverde and Jovanis (2008), two different neighbouring structures are considered: first order neighbours and second-order neighbours. First-order neighbours are defined as road segment  $j$  is directly

connected to segment  $i$  and  $w_{ij} = 1$ ; second-order neighbours are defined as road segment  $j$  is connected to first-order neighbours of segment  $i$  and  $w_{ij} = 1/2$ .  $w_{ij} = 0$  if segment  $i$  and  $j$  are not neighbours to each other (first or second order).

All models discussed in this section can be estimated using the Markov Chain Monte Carlo (MCMC) method under the full hierarchical Bayesian framework. The deviance information criterion (DIC), which can be thought as a generalization of the Akaike information criterion (AIC), can be used to compare goodness-of-fit and complexity of different models estimated under a Bayesian framework (Spiegelhalter et al., 2002). In terms of goodness-of-fit, the lower the DIC the better the model.

## 5 MODEL ESTIMATION RESULTS AND DISCUSSIONS

The main objective is to develop a series of models to investigate the relationship between traffic congestion and the frequency of different road accidents using the data from the 70 segments of the M25 motorway. Since only few fatal accidents occurred on each segment, these accidents were combined with serious injury accidents. This resulted in two categories of accidents: (1) fatal and serious injury accidents and (2) slight injury accidents. As such, two models were estimated for each of the four specifications such as (1) Poisson-lognormal, (2) Poisson-gamma, (3) Poisson-lognormal with CAR priors (1<sup>st</sup> order neighbour) and (4) Poisson-lognormal with CAR priors (2<sup>nd</sup> order neighbour). Therefore, a total of eight models were estimated.

To reduce the large variation among the explanatory variables, annual average daily traffic (AADT) and radius of road curvature were transformed into a logarithmic scale. Other forms of explanatory variables have been also tested, for example, radius of curvature and gradient are transformed into indicator (dummy) variables and it is found that they give similar results. Average vehicle speed was excluded as it was found to be highly correlated with the congestion index (correlation coefficient: -0.71). The correlation coefficients between other independent variables were also examined and the maximum value was found to be 0.59 suggesting that multicollinearity is not a problem for the rest of the explanatory variables.

The posterior means and standard deviations of the coefficients for the explanatory variables, the standard deviations of heterogeneity<sup>3</sup> and spatial correlation were estimated using the MCMC method. Two chains were simulated with different initial values and the initial 20 000 iterations (for the case of fatal and serious injury accident model) and 75 000 iterations (for the case of slight injury accident model) were discarded as burn-ins to achieve the convergence of the two chains. Then a further 75 000 iterations for each chain were performed and kept to calculate the summary statistics of interested parameters such as posterior means and standard deviations. The results for eight models are presented in Tables 2 and 3.

### Insert tables 2 and 3 here

As can be seen in Table 2, all specifications of Poisson models produce similar results in terms of the set of statistically significant variables and the values of their coefficients. In all specifications for the case of fatal and serious injury accidents, the statistically significant variables are *log of AADT*, *length of the segment* and *maximum vertical grade (%)*. These variables are also found to be statistically significant across all specifications in explaining

---

<sup>3</sup> For Poisson-gamma models the standard deviations of  $\exp(v)$  are estimated.

the variation in the frequency of slight injury accidents on the segments of the M25 motorway (see Table 3). However, the variable *number of lanes* becomes significant in all specifications except one (see Table 3). The posterior mean of the standard deviation of uncorrelated heterogeneity ( $v$ ) is found to be statistically significant across all models suggesting that the effect of heterogeneity does exist in our accident data. Similarly, the posterior mean of the standard deviation of spatial correlation ( $u$ ) for both types of road accidents is also statistically significant suggesting that road accidents are spatially correlated among neighbouring road segments. The effect of spatial correlation for the case of fatal and serious injury accidents is generally found to be slightly higher than for the case of slight injury accidents.

The DIC values for different specifications are found to be similar for both categories of accidents suggesting that there is no significant difference among different specifications tested in this study in terms of model goodness-of-fit. Generally, the Poisson-gamma model fitted the data slightly better, especially for the case of slight injury accidents. However, the better statistical fit does not necessarily mean the model could better reflect the theory of accidents and the actual effects of relevant factors (Lord et al., 2005). The effects of various explanatory variables in the models are discussed below.

## 5.1 Congestion index

The congestion index was calculated for each of the M25 motorway segments to precisely represent the level of congestion. In the case for fatal and serious injury accidents, this variable showed the expected negative sign suggesting that the increased level of congestion is associated with the decreased level of fatal and serious injury accidents. However, this variable was found to be statistically insignificant in all forms of Poisson models for both categories of accidents. This means that the level of traffic congestion has no impact on the frequency of road accidents according to the data on the M25. Therefore, spatial differences of congestion among road segments of the M25 cannot explain the variation in road accidents.

Therefore, our results do not confirm the hypothesis proposed by Shefer and Rietveld (1997) that traffic congestion may have an effect on road safety. This result is however in line with the findings of Noland and Quddus (2005) who investigated the association between congestion and safety in London based on area-wide data, and did not find a firm evidence that supported the hypothesis.

There may be a number of reasons for the insignificance of traffic congestion in our models. First, one may think that the effects of congestion might be captured by other factors such as *speed variance* and *traffic flow*. Literature suggests that *speed variance* is an important factor in explaining the occurrence of traffic accidents (Aljanahi et al., 1999; Ossiander and Cummings, 2002). This may also be true in our case in the sense that the congestion index (CI) used in this study does not explicitly represent how traffic speed on a specific segment varies at different times. Therefore, our intention was to include *speed variance* in the model. Speed variation in the literature, however, is measured by *Acceleration noise* (AN) which is also regarded as a congestion measurement (Taylor et al., 2000). Moreover, AN requires a considerable amount of data and due to the fact that speed variation is affected not just by traffic conditions (e.g. congestion) but also by driver behaviour this variable is not considered in this study. The effect of traffic flow is discussed in the following section.

One could also explain this result by the laws of accident causation proposed by Elvik (2006) who offered some general regularities in order to explain the relationship between the risk factors and accidents. For example, *the law of rare events* implies that “rare events” such as environmental hazards would have more effect on accident rates than “regular events”; *the*

*law of complexity* states that the more complex traffic situation road users encounter, the higher probability accidents would happen; and *the law of cognitive capacity* implies that accidents are more likely to happen as cognitive capacity approaches its limits. As such it can be speculated that traffic congestion may have a mixed effect on road safety. On the one hand, traffic congestion is increasingly common (i.e. “regular events”) in a modern transport network, especially on a busy motorway such as the M25, so people are increasingly aware of it and in addition, the delay caused by congestion could give drivers more time to make decisions; on the other hand, traffic congestion increases the “complexity” of the traffic situation and drivers may need to make frequent changes (e.g. adjusting speeds) in a congested traffic situation. Therefore, due to these possible mixed effects, overall congestion may ultimately result in little or no impact on road safety.

## 5.2 AADT and road segment length

*AADT* and *road segment length* are the two most important factors explaining road accident frequency in the models. *AADT* and *road segment length* are both statistically significant and positively associated with accidents in all models. This is to be expected, as *AADT* and *segment length* are considered to be the main exposure to accident risks.

The coefficient of  $\log(AADT)$  indicates the elasticity of accidents with respect to *AADT* suggesting that 1% increase in *AADT* would increase fatal and serious injury accidents by 1.21% to 1.86%. The coefficient of  $\log(AADT)$  for the case of slight injury accidents is found to be close to that for the case of fatal and serious injury accidents ranging from 1.03 to 1.53. The elasticity of *AADT* appears a little high in this study compared with some of the previous studies which reported that the elasticity ranges from 0.6 – 0.7 (e.g. Abdel-Aty and Radwan, 2000; Bird and Hashim, 2006; Agüero-Valverde and Jovanis, 2008). This may be because these studies were undertaken under different road conditions. For example, the study conducted by Abdel-Aty and Radwan (2000) was based on a “principal arterial” passing through the centre of Orlando in Florida. In addition, *AADT* was normalised by the number of lanes (i.e. *AADT* per lane) in this study. Bird and Hashim (2006) used a sample of rural two-lane single carriageways. A similarly study undertaken by Agüero-Valverde and Jovanis (2008) was also based on rural two-lane roads. Mitra et al. (2007) further showed that the effects of *AADT* on different roads are different while looking at junction accidents, and it was found that generally the coefficient of *AADF* for “major” roads is higher than “minor” roads. Our study was based on a major motorway which is one of the busiest in Europe and therefore, the higher value of the coefficient of *AADT* can be seen as reasonable.

The estimates of coefficients for *segment length* are generally similar among different model specifications, which is around 0.13 for fatal and serious injury accidents and 0.15 for slight injury accidents. The corresponding mean elasticity<sup>4</sup> of *segment length* is 0.68 for fatal and serious injury accidents and 0.79 for slight injury accidents. These values are found to be lower than some existing studies (Bird and Hashim, 2006) and higher than other studies (Abdel-Aty and Radwan, 2000). It can be speculated that the relationship between accident frequency and road segment length is non-linear. According to *the universal law of learning* (Elvik, 2006), the accident rate tends to decline as the number of kilometres travelled increases suggesting that such a non-linear relationship exists. Therefore, the elasticity of accidents with respect to segment length can be different in different scenarios.

## 5.3 Other contributing factors

---

<sup>4</sup> Mean elasticity is defined as  $E = \frac{\partial \mu}{\partial x} \cdot \frac{\bar{x}}{\mu} = \beta \mu \cdot \frac{\bar{x}}{\mu} = \beta \bar{x}$ .

Other contributing factors considered in the models include the *radius of road curvature*, *gradient (%)*, *number of lanes* and *direction* (clockwise and anti-clockwise) of the motorway. The *radius of road curvature* that reflects the degree of horizontal curvature of a road segment is included as literature suggests that this may have an impact on accidents (Milton and Mannering, 1998). However, this variable was found to be statistically insignificant in all models. This may be due to the fact that there is a mixed effect of road curvature (Milton and Mannering, 1998; Haynes et al., 2007), especially on a highly aggregated road segments. This may also be due to the fact that there is not enough variation in horizontal curvature among the M25 road segments. An indicator variable representing the high road curvature (i.e., the radius of curvature is less than 500m) was also tested but found to be statistically insignificant.

*Gradient (%)* which represents the vertical grade of the segment was also included in the models and found to be statistically significant and positively associated with accidents in all models. This is consistent with the study by Milton and Mannering (1998) who used an indicator variable to represent vertical grade. This variable was found to be more significant (95% confidence level) in the Poisson-lognormal CAR models for the case of fatal and serious injury accidents.

The variable representing the *number of lanes* is found to be statistically insignificant in all fatal and serious injury accident models but becomes significant and positively related to the frequency of slight injury accidents in all specifications except one which is the Poisson-lognormal with CAR priors (1<sup>st</sup> order neighbour). *Direction* is included in the models as a dummy variable (clockwise and anticlockwise) to investigate whether there is an association between the frequency of accidents on the M25 and its directions. Li et al. (2007) suggested that the roadway directions need to be differentiated to better evaluate roadway risks. In this study, *direction* was found to be statistically insignificant in all models suggesting that this variable does not have any effect on the accidents.

Some other important factors such as the weather conditions and road conditions (e.g. pavement roughness) were not included in this study due to the data unavailability. We do not have access to the information about the scheduled road works (SRW) on the M25 during the study period. This may affect both accidents and traffic characteristics and as such the findings of this study.

## 6 CONCLUSIONS AND FURTHER RESEARCH

This paper has investigated the effects of traffic congestion on road accidents while controlling for other contributing factors such as traffic flow, segment length, number of lanes, curvature and gradient. A robust technique has been developed to assign M25 accidents onto its segments and a congestion index has been used to represent the level of traffic congestion on each segment. A series of relevant statistical models has been employed to develop the relationship between the frequency of accidents and the congestion index. Results from the various model specifications have been generally found to be consistent with each other. It has been found that traffic congestion has no impact on the frequency of accidents (either for fatal and serious injury accidents or for slight injury accidents). This may be due to the impact of other factors such as traffic flow and speed variation on traffic congestion or their impacts on road accidents. Other factors included in the models have provided expected results.

In terms of the specifications of the models, the Poisson-gamma model appeared to fit the data slightly better, especially for the case of slight injury accidents. Although the spatial model (i.e. the Poisson-lognormal with CAR priors) has implied that the spatial correlation

among neighbouring segments of the M25 motorway is significant, the results are very similar to the non-spatial models.

One limitation of this study is that only road segments from the M25 motorway have been included in the analysis. There are many other motorways and A roads (major roads in England) connected to the M25 and it is reasonable to believe that a study including all roads connected to the M25 motorway may provide a better understanding on the impact of traffic congestion on road safety, as there will be more spatio-temporal variations in the level of traffic congestion and the frequency of accidents. In addition, the effects of congestion at junctions on accidents should be explored.

Given the data availability, data for multiple years can be collected and a spatio-temporal analysis can be employed to ensure that the time effects are controlled for. Clearly, further research is needed to fully understand the effects of traffic congestion on road safety.

## ACKNOWLEDGEMENTS

The authors would like to thank the UK Highways Agency for providing traffic characteristics data for the M25 motorway. The content of the paper however does not necessarily express the views of the Highways Agency and the authors take full responsibility for the content of the paper and any errors or omissions.

## REFERENCES

- Abdel-Aty, M.A. and Radwan, A.E., 2000. Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, **32**(5), pp. 633-642.
- Aguero-Valverde, J. and Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis & Prevention*, **38**(3), pp. 618-625.
- Aguero-Valverde, J. and Jovanis, P.P., 2008. Analysis of road crash frequency using spatial models, *Proceedings of the Paper Presented at the 87th Annual Meeting of the Transportation Research Board*.
- Aljanahi, A.A.M., Rhodes, A.H. And Metcalfe, A.V., 1999. Speed, speed limits and road traffic accidents under free flow conditions. *Accident Analysis & Prevention*, **31**(1-2), pp. 161-168.
- Amoros, E., Martin, J.L. and Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. *Accident Analysis & Prevention*, **35**(4), pp. 537-547.
- Besag, J., 1974. Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, **36**(2), pp. 192-236.
- Bird, R. and Hashim, I., 2006. Exploring the Relationship between Safety and the Consistency of Geometry and Speed on British Roads, *Proceedings of the Paper Presented at the 85th Annual Meeting of the Transportation Research Board*.
- Boarnet, M., Kim, E. and Parkany, E., 1998. Measuring Traffic Congestion. *Transportation Research Record*, **1634**(-1), pp. 93-99.

- Elvik, R., 2006. Laws of accident causation. *Accident Analysis & Prevention*, **38**(4), pp. 742-747.
- Ghosh, M., Natarajan, K., Waller, L.A. and Kim, D., 1999. Hierarchical Bayes GLMs for the analysis of spatial data: An application to disease mapping. *Journal of Statistical Planning and Inference*, **75**(2), pp. 305-318.
- Graham, D.J. and Glaister, S., 2003. Spatial Variation in Road Pedestrian Casualties: The Role of Urban Scale, Density and Land-use Mix. *Urban Studies*, **40**(8), pp. 1591-1607.
- Haynes, R., Jones, A., Kennedy, V., Harvey, I. and Jewell, T., 2007. District variations in road curvature in England and Wales and their association with road-traffic crashes. *Environment and planning A*, **39**(5), pp. 1222-1237.
- Ivan, J.N., Wang, C. and Bernardo, N.R., 2000. Explaining two-lane highway crash rates using land use and hourly exposure. *Accident Analysis & Prevention*, **32**(6), pp. 787-795.
- Kim, K., Brunner, I. and Yamashita, E., 2006. Influence of land use, population, employment, and economic activity on accident's. *Safety Data, Analysis, and Evaluation*, (1953), pp. 56-64.
- Li, L., Zhu, L. and Sui, D.Z., 2007. A GIS-based Bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes. *Journal of Transport Geography*, **15**(4), pp. 274-285.
- Lord, D., 2000. *The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models*, Department of Civil Engineering, University of Toronto, Toronto.
- Lord, D. and Miranda-Moreno, L.F., 2007. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: a Bayesian perspective, *Proceedings of the Paper Presented at the 86th Annual Meeting of the Transportation Research Board*.
- Lord, D., Washington, S.P. and Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis & Prevention*, **37**(1), pp. 35-46.
- Macnab, Y.C. and Dean, C.B., 2001. Autoregressive Spatial Smoothing and Temporal Spline Smoothing for Mapping Rates. *Biometrics*, **57**(3), pp. 949-956.
- Martin, J., 2002. Relationship between crash rate and hourly traffic flow on interurban motorways. *Accident Analysis & Prevention*, **34**(5), pp. 619-629.
- Miaou, S., Song, J.J. and Mallick, B., 2003. Roadway traffic crash mapping: a space-time modeling approach. *Journal of Transportation and Statistics*, **6**(1), pp. 33-57.
- Milton, J. and Mannering, F., 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation*, **25**(4), pp. 395-413.

- Mitra, S., Washington, S. and Schalkwyk, I.V., 2007. Important omitted spatial variables in safety models: understanding contributing crash causes at intersections, *Proceedings of the Paper Presented at the 86th Annual Meeting of the Transportation Research Board*.
- Noland, R.B. and Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. *Accident Analysis & Prevention*, **36**(6), pp. 973-984.
- Noland, R.B. and Quddus, M.A., 2005. Congestion and safety: A spatial analysis of London. *Transportation Research Part A: Policy and Practice*, **39**(7-9), pp. 737-754.
- Ossiander, E.M. And Cummings, P., 2002/1. Freeway speed limits and traffic fatalities in Washington State. *Accident Analysis & Prevention*, **34**(1), pp. 13-18.
- Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accident Analysis & Prevention*, **40**(4), pp. 1486-1497.
- Shankar, V., Mannering, F. and Barfield, W., 1995. Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention*, **27**(3), pp. 371-389.
- Shefer, D. and Rietveld, P., 1997. Congestion and Safety on Highways: Towards an Analytical Model. *Urban Studies*, **34**(4), pp. 679-692.
- Shefer, D., 1994. Congestion, air pollution, and road fatalities in urban areas. *Accident Analysis & Prevention*, **26**(4), pp. 501-509.
- Shinar, D. and Compton, R., 2004. Aggressive driving: an observational study of driver, vehicle, and situational variables. *Accident Analysis & Prevention*, **36**(3), pp. 429-437.
- Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D., 2003. WinBUGS User Manual Version 1.4.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Linde, A.V.D., 2002. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **64**(4), pp. 583-639.
- Taylor, M.A.P., Woolley, J.E. and Zito, R., 2000. Integration of the global positioning system and geographical information systems for traffic congestion studies. *Transportation Research Part C: Emerging Technologies*, **8**(1-6), pp. 257-285.

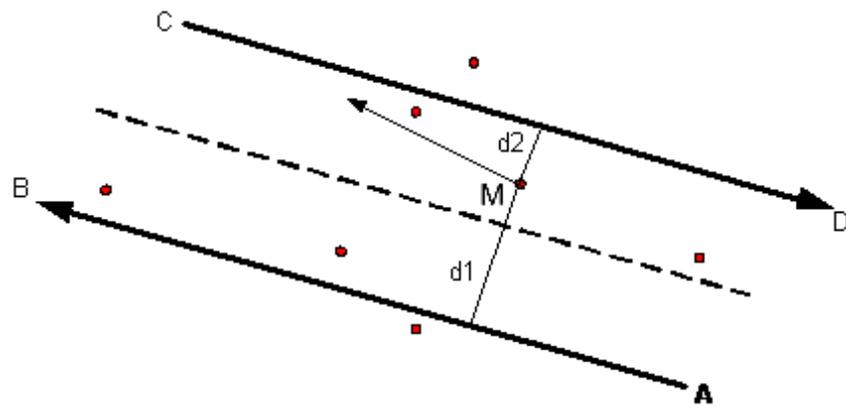
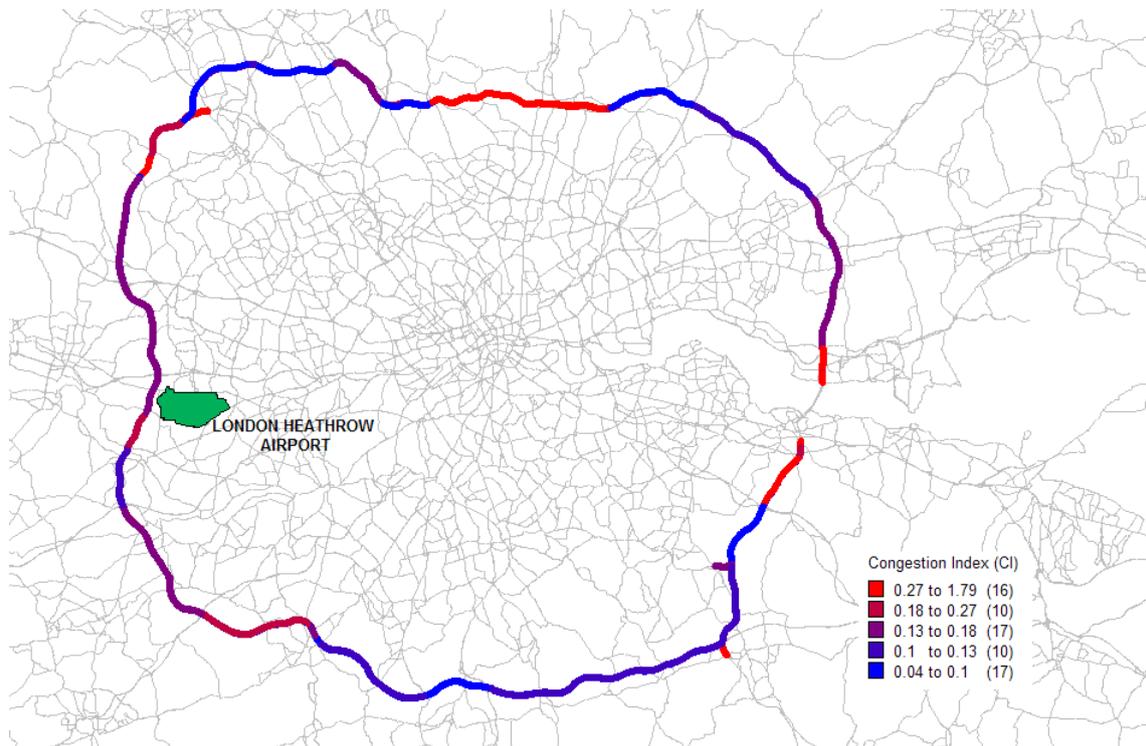
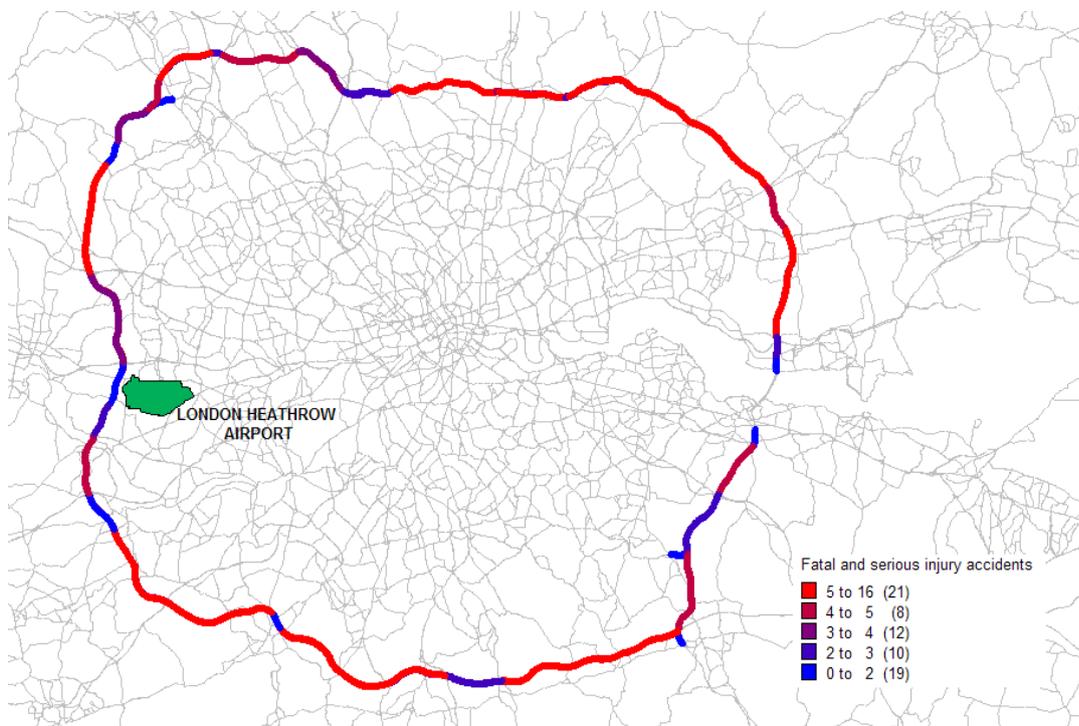


Figure 1 Accident location and motorway centre-line data



(a) CI on the different segments of the M25



(b) Fatal and serious injury accidents on the different segments of the M25

Figure 2 Spatial distribution of CI and fatal and serious injury accidents on the different segments of the M25

**Fatal and serious injury accidents**

**Slight injury accidents**

**Figure 3 Traffic accidents and AADT**

**Fatal and serious injury accidents**

**Slight injury accidents**

**Figure 4 Traffic accidents and Congestion Index (CI)**

**Table 1 Summary statistics of the variables**

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>Accident</i>					
Fatal and serious injury accidents	70	3.8429	3.5207	0	16
Slight injury accidents	70	38.6429	30.3592	0	128
<i>Traffic characteristics</i>					
Congestion index (CI) (‰)	70	0.2242	0.2436	0.0407	1.7874
Annual average daily traffic (AADT) (vehicles)	70	64699.39	19543.48	11343.1	96000.99
Average vehicle speed (km/h)	70	83.2113	13.1741	43.9865	106.4163
<i>Road segment infrastructure</i>					
Road segment length (km)	70	5.2603	3.42	0.7584	15.3992
Minimum radius (m)	70	749.6679	296.8777	24.96	1815.26
Maximum gradient (%)	70	3.0971	1.1528	0.7	6.1
Number of lanes	70	3.3338	0.8311	2	6
Direction (dummy variable)	70	0.5143	0.5034	0	1

**Table 2 Models for fatal and serious injury accidents on the M25**

Variables	Poisson-lognormal		Poisson-Gamma		Poisson-lognormal CAR (1st order neighbour)		Poisson-lognormal CAR (2nd order neighbour)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Congestion index (%)	-0.5877	0.6958	-0.4225	0.7935	-0.6934	0.7256	-0.7836	0.7485
log(AADT)	<b>1.212**</b>	0.4942	<b>1.856**</b>	0.5465	<b>1.435**</b>	0.4316	<b>1.227**</b>	0.5195
Segment length (km)	<b>0.1351**</b>	0.0260	<b>0.1475**</b>	0.0337	<b>0.1362**</b>	0.0241	<b>0.1362**</b>	0.0247
log(minimum radius)	0.234	0.2121	0.2916	0.2289	0.2525	0.1955	0.2966	0.3380
Maximum gradient (%)	<b>0.1868*</b>	0.0993	<b>0.2103*</b>	0.1130	<b>0.2026**</b>	0.0912	<b>0.2099**</b>	0.1290
Number of lanes	-0.0961	0.1601	-0.2023	0.1869	-0.0416	0.1544	0.0023	0.1892
Direction	0.0543	0.1585	0.0855	0.1957	-0.1469	0.3115	-0.1596	0.3178
Intercept	<b>-14.77**</b>	5.0330	<b>-22.08**</b>	5.28	<b>-17.49**</b>	4.511	<b>-15.64**</b>	6.009
S.D. ( $u$ )					<b>0.1244**</b>	0.07992	<b>0.2376**</b>	0.1429
S.D. ( $v$ )	<b>0.3229**</b>	0.1249	<b>0.5285**</b>	0.0790	<b>0.1759**</b>	0.1427	<b>0.149**</b>	0.1681
DIC	283.561		281.021		282.358		284.038	

\* Statistically significantly different from zero (90% credible sets show the same sign)

\*\* Statistically significantly different from zero (95% credible sets show the same sign)

**Table 3 Models for slight injury accidents on M25**

Variables	Poisson-lognormal		Poisson-Gamma		Poisson-lognormal CAR (1st order neighbour)		Poisson-lognormal CAR (2nd order neighbour)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Congestion index (‰)	0.1513	0.4382	-0.1181	0.4199	0.3731	0.5493	0.2966	0.4477
log(AADT)	<b>1.492**</b>	0.2919	<b>1.026**</b>	0.2346	<b>1.525**</b>	0.6250	<b>1.376**</b>	0.3756
Segment length (km)	<b>0.1505**</b>	0.0231	<b>0.1582**</b>	0.0269	<b>0.1587**</b>	0.0225	<b>0.1584**</b>	0.0206
log(minimum radius)	0.0244	0.1754	0.0526	0.1399	0.0220	0.1858	0.0647	0.1473
Maximum gradient (%)	<b>0.2287**</b>	0.0871	<b>0.1896**</b>	0.0804	<b>0.2285**</b>	0.0780	<b>0.2291**</b>	0.0663
Number of lanes	<b>0.1912*</b>	0.1074	<b>0.262**</b>	0.1159	0.1893	0.1461	<b>0.2237*</b>	0.1204
Direction	0.0046	0.1328	0.0049	0.1421	-0.0788	0.2459	-0.0664	0.2250
Intercept	<b>-15.53**</b>	3.0960	<b>-10.55**</b>	2.3570	<b>-15.93**</b>	6.0140	<b>-14.66**</b>	3.4880
S.D. ( $u$ )					<b>0.0964**</b>	0.0882	<b>0.1341**</b>	0.1303
S.D. ( $v$ )	<b>0.4777**</b>	0.0600	<b>0.5291**</b>	0.0571	<b>0.4294**</b>	0.0957	<b>0.4184**</b>	0.0790
DIC	490.053		482.197		490.286		489.130	

\* Statistically significantly different from zero (90% credible sets show the same sign)

\*\* Statistically significantly different from zero (95% credible sets show the same sign)