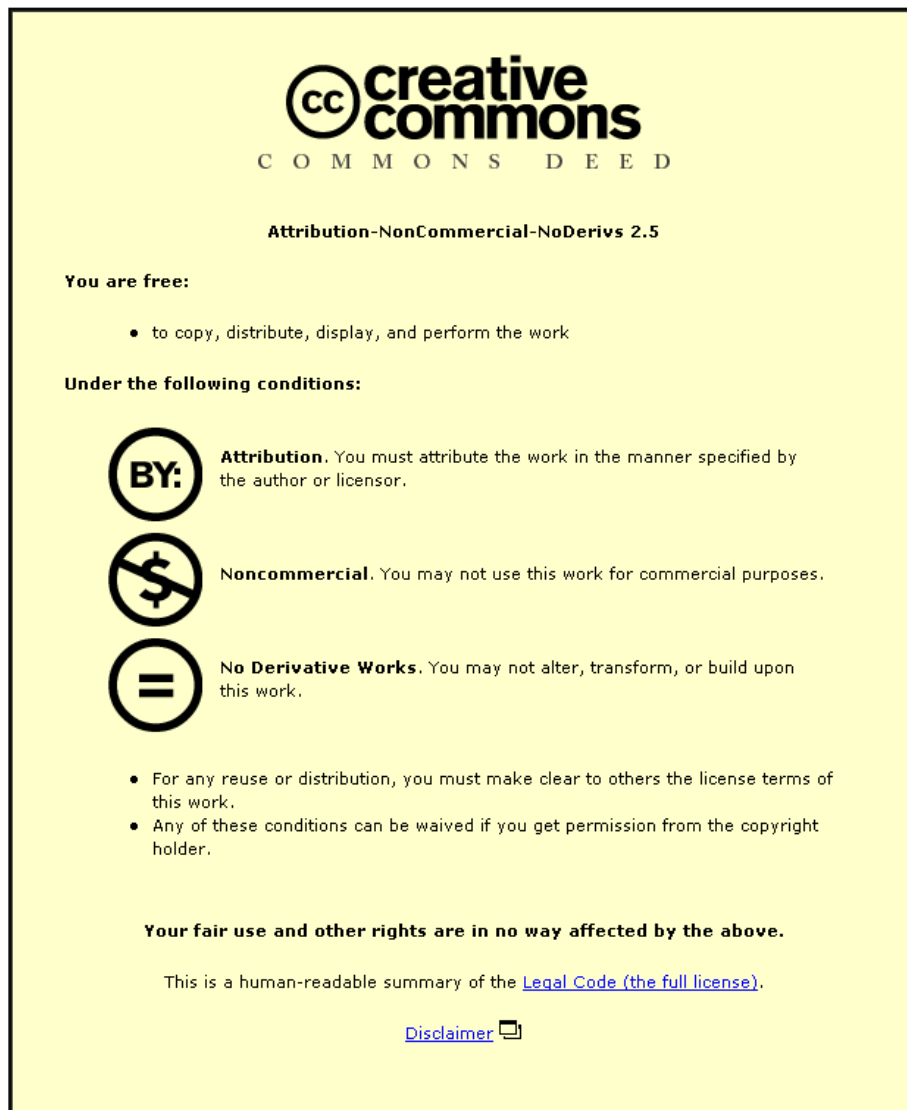


This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



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

A spatiotemporal analysis of the impact of congestion on traffic safety on major roads in the UK

Chao Wang¹, Mohammed Quddus², Stephen Ison³

Transport Studies Group
Department of Civil and Building Engineering
Loughborough University
Loughborough LE11 3TU
United Kingdom

¹ Research Associate, Email: c.wang@lboro.ac.uk; Tel: +44 (0)1509 564682

² Senior Lecturer in Transport Studies, Email: m.a.quddus@lboro.ac.uk; Tel: +44 (0)1509 228545; Fax: +44 (0)1509 223981, **Corresponding Author**

³ Professor of Transport Policy, Email: s.g.ison@lboro.ac.uk; Tel: +44 (0)1509 222605; Fax: +44 (0)1509 223981

Abstract

A spatio-temporal analysis has been conducted aiming to explore the relationship between traffic congestion and road accidents based on the data on the M25 motorway and its surrounding major roads in England during the period 2003-2007. It was hypothesised that increased traffic congestion may be beneficial to road safety as the number of fatal and serious injury (KSI) accidents would be less due to low average speed when congestion is present. If this is confirmed then it poses a potential dilemma for transport policy makers. A series of classical count outcome models (random-effects Negative Binomial models) and spatial models using a full Bayesian hierarchical approach have been developed in this study in order to examine whether congestion has any effect on the frequency of accidents. The results suggest that increased traffic congestion is associated with more KSI accidents and traffic congestion has little impact on slight injury accidents. This may be due to the higher speed variance among vehicles within and between lanes and worse driving behaviour in the presence of congestion. In addition, traffic speeds even within congested situations are likely to be relatively high on major roads compared to other parts of the road network. Some strategies are then proposed to optimise traffic flow which would be beneficial to both congestion and accident reduction.

Keywords: Traffic congestion, Road traffic accidents, spatial autocorrelation, spatio-temporal analysis; Random-effect negative binomial models, Bayesian hierarchical spatial models.

1. Introduction

Addressing road traffic congestion and reducing accidents are important objectives for transport policy makers. An ideal scenario would be to reduce both congestion and accidents simultaneously. However, this may not be possible since it is likely that there is an inverse relationship between traffic congestion and road fatalities (Shefer and Rietveld, 1997). It is possible that there is less fatal and serious injury (KSI) accidents under congested conditions when speed is low. Therefore, it appears that increased traffic congestion may be beneficial to road safety, which poses a potential dilemma for transport policy makers. To address this it is therefore important to fully understand the relationship between traffic congestion and road accidents.

There have been many studies devoted to the relationship between accidents and congestion related factors, such as volume capacity ratio and hourly traffic flow (e.g. Belmont and Forbes, 1953; Gwynn, 1967; Ceder, 1982; Zhou and Sisiopiku, 1997; Lord et al., 2005). It seems, however, that there is a dearth of literature exploring the association between traffic congestion and road accidents. Several studies have investigated the effects of congestion on road accidents by providing both analytical and empirical evidence (Shefer, 1994; Shefer and Rietveld, 1997; Baruya, 1998; Noland and Quddus, 2005; Kononov et al., 2008). For example, Shefer (1994) and Shefer and Rietveld (1997) proposed an inverse relationship between congestion and road fatalities. Based on 63 A and B roads in the UK, Baruya (1998) found that the “degree of congestion” has negative effects on accident frequency (i.e. an inverse association). Noland and Quddus (2005) investigated the relationship between traffic congestion and road accidents using data from London and reported that their results are inconclusive and generally congestion has little effect on accidents according to the data. These studies often have the limitation of using a weak proxy for congestion, such as volume over capacity ratio, “proportion of vehicles slower than half the speed limit”, spatial location (e.g. Inner and Outer London) and employment density. These proxies may not accurately represent levels of traffic congestion, and thus the results from econometric models may be biased. For instance, congestion and traffic density are not equivalent 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 the increase in congestion is not proportional to the increase in density. Noland and Quddus (2005) used an indicator variable for Inner London as a proxy for congestion, and no significant differences were found between Inner London and Outer London and they speculated that speed is generally low in both areas. This suggested that such proxies for congestion may not precisely represent levels of congestion, and therefore a more suitable congestion measurement is required to more accurately represent congestion in an econometric model so as to provide more robust empirical evidence. Noland and Quddus (2005) suggested that instead of an area-wide based study, a road segment based study can be used to better capture the variation of traffic congestion. As such, a more precise congestion measurement is desirable. In terms of methods used in the statistical analysis, Poisson or Negative Binomial models have been used in these studies, so more sophisticated models recently used in safety research such as spatial econometrics could be employed so as to better understand the relationship between traffic congestion and road accidents. Based on urban freeways in California, Colorado and Texas, Kononov et al. (2008) found that total as well as fatal and injury accident rates increase with the increase in traffic congestion. Again, traffic congestion was measured using a proxy in their study: the annual average daily traffic (AADT). In addition, only AADT and number of lanes were considered as risk factors in their study. However, it is essential that other factors affecting road accidents such as road geometry need to be controlled for.

A recent study conducted by the authors examined the impact of traffic congestion on road accidents using a spatial analysis on the M25 motorway in England (Wang et al., 2009). While controlling for other contributing factors such as annual average daily traffic (AADT) and road geometry, it concluded that traffic congestion has no impact on the frequency of accidents on the M25. Several non-spatial models (such as Poisson-lognormal and Poisson-gamma) and spatial models (Poisson-lognormal with conditional autoregressive priors) were employed to investigate the effect of traffic congestion on road accidents both within the peak time and off-peak time period. Congestion was measured using the congestion index (i.e. traffic delay divided by free flow travel time), and this variable was found to be statistically insignificant in all models, meaning that traffic congestion has little or no impact on the frequency of accidents according to the data on the M25. It was speculated that there may be a mixed effect (i.e. both positive and negative) of traffic congestion on road accidents, so congestion ultimately results in little impact on road accidents. The limitations of this previous study, as the authors suggest, are the data used in the analysis. For instance, accident and traffic data were considered for only one major road (i.e., M25) and for only one year (i.e., 2006). There are many other major motorways and A roads connected to the M25. As such, there is a need to consider a road network (rather than a single road) as more spatio-temporal variations are expected on a wider road network. Moreover, data for multiple years should be considered to control for the unobserved effect that changes over time. Therefore, the primary objective of this study is to re-investigate the effect of traffic congestion on road accidents by extending the previous study (Wang et al., 2009) in three ways: (1) extending the study area that includes 13 different motorways and 17 different A roads leading to a total number of 298 road segments, (2) considering traffic and accident data for 5 years (2003-2007) and (3) employing a spatio-temporal Bayesian hierarchical count model that controls for spatial correlation among neighbouring segments and time effects over the years. In this study, the count of annual accidents per road segment is viewed as a function of various factors, and as such several statistical models that are suitable for panel count data are used. It is expected that by extending the previous study both spatially and temporally, more spatio-temporal variations of the level of congestion and accidents will be observed, which would make for a better understanding of the relationship between traffic congestion and road accidents.

The paper is organised as follows: first, the data used in this study is briefly described. It is then followed by a discussion of the statistical models. Model estimation results are then presented. This is followed by the discussion of the findings. Finally, conclusions and further research directions are provided.

2. Data description

The M25 and other motorways and A roads⁴ that connect to the M25 were selected. The M25 motorway is an orbital motorway that encircles London, England. The primary reason for selecting the M25 and surround in this study is that the M25 is one of the busiest motorways in Europe, and as such it is expected that there would be large spatial and temporal variation of traffic flow and levels of congestion on the M25 and surround which allows to establish a reliable relationship between accidents and levels of congestion.

Traffic characteristic data were made available from the UK Highways Agency (HA⁵). The HA collects hourly traffic characteristics and road infrastructure data for major

⁴ “A roads” are classified as major roads in England.

⁵ <http://www.highways.gov.uk/>

motorways and A roads at a road segment level in the UK. Traffic characteristics such as average travel time, average travel speed, traffic flow and total vehicle delay for road segments on the M25 and surround during the years 2003 – 2007 have been calculated from the disaggregated traffic data supplied by the HA. The level of segment-based traffic congestion is measured by the total delay (in sec) per kilometre, where the total delay is the traffic delay incurred on all vehicles travelling along a road segment in a year.

Road infrastructure data such as segment length, number of lanes, radius of curvature and gradient for road segments have also been obtained from the HA. According to the HA, a total number of 298 road segments are identified on the M25 and surround, and each road segment starts and ends at a junction. It is worthwhile noting that a particular stretch of a road has two segments – one in each direction. According to the data obtained from the HA, there are a total of about 300,000 hours' traffic delay occurred on the M25 and surround (the total length of roads is around 1,390 kilometres) during 2003-2007.

Accident data for the years 2003-2007 were derived from the STATS19 UK national road accident database. STATS19 database contains information regarding accidents, vehicles and casualties involved in an accident. While accidents were overlaid onto road segments in GIS, mismatches between them are observed due to errors in both accident location (i.e., easting and northing coordinates) and digital road segment data. In order to correctly count the number of accidents occurred on a road segment, the matching technique proposed by the previous study (Wang et al., 2009) was used to match the accident points onto the correct road segment. This mapping method was using a weighting score of the perpendicular distance and the direction of the vehicle relative to a road segment to assign accidents to the correct road segments. Only accidents recorded as occurred on the roadway segments are retained. Similar to the study by Agüero-Valverde and Jovanis (2008), accidents coded as junction accidents (about 30% of total accidents) were excluded from the analysis. This is because major road junctions are complicated in terms of road design (such as flyovers and slip roads) compared to road segments and it is also difficult to obtain a single measure of traffic flow at a junction. Accidents were classified into three categories according to their severity levels in the UK: fatal, serious injury and slight injury accidents. Property damage only (PPO) accidents are not recorded in the STATS19 database.

Summary statistics of the annual segment-level accidents, traffic and road characteristics are presented in Table 1.

Table 1 is about here

Table 1 shows all the variables considered in this study. Although the panel dataset consists of 298 cross-sectional units and 5 temporal units, a total of 1,391 observations were used in the analysis due to missing values. Therefore, the resulting dataset is an unbalanced panel. Due to the low frequency of fatal accidents, they are combined with serious injury accidents. Total traffic delay is normalised by road segment length so as to have a direct measurement of congestion. Average vehicle speed is weighted (by traffic flow) harmonic mean of hourly speed data. Motorway indicator is a dummy variable with 1 representing motorway or A roads with motorway standard such as A1(M); and 0 representing other major A roads. When including the variables in the statistical models, some of the explanatory variables have been transformed into a logarithmic scale to reduce the variance among the variables, including total delay, annual average daily traffic (AADT), road segment length and radius of road curvature. Average speed has been excluded from the model to avoid multicollinearity with the interested variable – total delay per km (correlation coefficient: -0.84). In other words, if both the variables (speed and delay) were included in the model, the model would have the issue of multicollinearity. Since investigating the effect of traffic

congestion is the aim of this paper, average speed has been excluded from the model. Correlation coefficients between other variables (such as AADT and traffic delay) have also been checked and no significant correlations are found.

The spatial distributions of total traffic delay and number of accidents in 2007 are shown in Figure 1. As shown in Figure 1, it is difficult to draw any conclusions regarding the association between accidents and total delay.

Figure 1 is about here

The proportion of accident occurred in each hour over a day is presented in Figure 2. It can be seen that most of the accidents occurred during peak hours when the congestion level is high. For example, around 6.5% of KSI accidents occurred during 07:00–08:00. This seems to suggest that traffic congestion increases the number of accidents.

Figure 2 is about here

Scatter plots showing the relationship between traffic delay and accidents for peak time periods (6:00–20:00 on weekdays and 9:00–20:00 on weekends) and off-peak time periods are presented in Figures 3-4. As can be seen however no clear relationship can be found between accidents and traffic delay.

Figure 3-4 is about here

In this paper the relationship between traffic congestion and the number of road accidents has been examined using two types of statistical model namely: the classical count outcome model and the spatial model using a full Bayesian approach. The two types of models are discussed below.

3. Statistical models

For each category of accidents, two types of statistical models suitable for panel count outcome data have been investigated in this study: (1) classical count outcome models and (2) spatial models using a full Bayesian hierarchical approach. Compared to a classical count model, a spatial model under a full Bayesian framework has the advantage in that they can accommodate spatial correlated effects among neighbouring segments. These two modelling techniques are briefly discussed below.

3.1 Classical count outcome models

As stated, the data used in this study is a panel count dataset. Several classical statistical models suitable for panel count outcome data are considered and tested. This includes: a fixed-effects Negative Binomial (NB) and a random-effects NB model. Both models can be estimated using the maximum likelihood method. The terms “fixed” or “random” refer to the location specific characteristics. The fixed or random-effects NB models have been proposed by Hausman et al. (1984) and are widely used in modelling panel count outcome data (including accident data) in the literature (e.g., Shankar et al., 1998; Chin and Quddus, 2003; Noland and Oh, 2004).

In a fixed- or random-effects model, Y_{it} (the annual number of observed accidents occurred on a road segment i at year t) is assumed NB distributed with parameters θ_i and k_i , where θ_i is the location-specific effect; k_i is the Negative Binomial overdispersion parameter;

and $\lambda_{it} = \exp(\alpha + \beta \mathbf{X}_{it})$ where α is the intercept; \mathbf{X}_{it} is the vector of explanatory variables for a road segment i at year t ; β is the vector of coefficients to be estimated. Y_{it} thus has mean $\theta_i \lambda_{it} / k_i$ and variance $(\theta_i \lambda_{it} / k_i) \times (1 + \theta_i / k_i)$. For a fixed-effects NB model, the location effect θ_i is modelled as a fixed unknown parameter; and for a random-effects NB model, it is treated as a random variable with independently and identically distributed (*i.i.d*) and $(1 + \theta_i / k_i)^{-1}$ is assumed beta distributed with $\text{Beta}(a, b)$.

Therefore the difference between fixed and random-effects models lies in the different specifications of the location specific effect⁶. One can employ a Hausman test (Hausman, 1978) to determine the appropriateness of using a fixed or random-effects model. The Akaike information criterion (AIC) can be used to compare goodness-of-fit and complexity of different models.

3.2 Spatial models using full Bayesian hierarchical approach

The classical count outcome models described above largely ignore the unmeasured spatial correlation among the neighbouring segments. This can be addressed using spatial models that have been used in accident modelling research (Miaou et al., 2003; Aguero-Valverde and Jovanis, 2006; Li et al., 2007; Quddus, 2008; Aguero-Valverde and Jovanis, 2008; El-Basyouny and Sayed, 2009b; Schneider et al., 2009; Persaud et al., 2010; and Lee et al., 2010). According to Aguero-Valverde and Jovanis (2006), compared with the classical count outcome models, spatial models can generally produce consistent results and can better fit the location-specific data. A further study by Aguero-Valverde and Jovanis (2008) employed a similar spatial model on rural two-lane segments in Pennsylvania, finding that spatial models show a significantly better fit than non-spatial models and a spatial model with the simplest neighbouring structure (i.e. the first-order neighbours in their study) is preferable.

In this study a spatial model that is similar to the one used by Aguero-Valverde and Jovanis (2008) has been employed. Compared to their study, in this paper only first-order neighbours were considered, and two additional terms have been added to control for the time effects and space-time interaction. The models are estimated using a full Bayesian hierarchical approach with conditional autoregressive (CAR) priors to account for spatial correlated effects. The form of the model can be expressed as follows:

$$Y_{it} \sim \text{Poisson}(\mu_{it}) \quad (1)$$

$$\log(\mu_{it}) = \alpha + \beta \mathbf{X}_{it} + v_i + u_i + \delta_t + e_{it} \quad (2)$$

where Y_{it} is the annual number of observed accidents occurred on a road segment i at year t ; μ_{it} is the expected accident count on a road segment i at year t ; α is the intercept; \mathbf{X}_{it} is the vector of explanatory variables for a road segment i at year t ; β 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 neighbouring road segment i ; δ_t is the term representing time effects (i.e. year-to-year effects); e_{it} is a random term for extra space-time interaction effects.

Models are estimated using a software – WinBUGS (Spiegelhalter et al., 2003) and the model specification generally follows the recommendations made in the user manual. A uniform prior distribution is assigned to α ; a highly non-informative normal prior is assumed

⁶ Note the random-effects NB model is equivalent to the random parameter count model used by Anastasopoulos and Mannering (2009) and El-Basyouny and Sayed (2009a) if only the location specific effect is random.

to all β 's with zero mean and 100,000 variance. Other options of priors may also be used, for instance a power prior as an informative prior to utilise previously available data (Lee et al., 2010). The prior distribution for v_i is a normal prior with $N(0, \tau_v^2)$, where τ_v^2 is the precision (1/variance) with a vague gamma prior $Gamma(0.5, 0.0005)$. The spatial correlation term u_i is modelled with a conditional autoregressive (CAR) model proposed by Besag (1974):

$$u_i | u_j, i \neq j \sim N\left(\frac{\sum_j u_j w_{ij}}{w_{i+}}, \frac{\tau_u^2}{w_{i+}}\right) \quad (3)$$

where $w_{ij} = 1$ if segment i and j are adjacent to each other (i.e. shared vertex) and $w_{ij} = 0$ otherwise; $w_{i+} = \sum_j w_{ij}$; and τ_u^2 is a scale parameter assumed as a gamma prior

$Gamma(0.5, 0.0005)$.

For the specifications of the terms δ_t and e_{it} , since the panel data used in this study has a large number of cross-sectional units (i.e., 298 road segments) but relatively short time period (2003-2007), two simple modelling structures for temporal effects (δ_t) have been considered: fixed-effects varying by t ; and random-effects with a first-order random walk (RW (1)) prior. For the fixed time effect model, δ_1 is set to be zero and δ_2 to δ_5 is assigned highly non-informative normal priors with zero mean and 100,000 variance. For the random time effect model, δ_t is assumed to be temporally correlated and assigned a RW (1) prior, which can be modelled using the CAR distribution in WinBUGS (Lagazio et al., 2001; Thomas et al., 2004). Similar to the specification of the prior distribution of u_i , the weight between neighbouring time periods (e.g. 2003 and 2004) is set to be 1 otherwise 0. For the space-time interaction term e_{it} , a vague normal prior $N(0, \tau_v^2)$ where

$\tau_v^2 \sim Gamma(0.5, 0.0005)$ is assumed as suggested by Miaou et al. (2003).

One limitation of using WinBUGS is the limited ability to handle missing values (Kynn, 2006). As the data used in this study is an unbalanced panel data (i.e. missing values for some road segments at a certain year), some road segments are removed from the data to form a balanced panel dataset that can be estimated using WinBUGS. The unbalanced panel data however can be analysed using classical count outcome models.

The Bayesian models are estimated using the Markov Chain Monte Carlo (MCMC) method. The deviance information criterion (DIC), which can be thought as a generalization of the Akaike information criterion (AIC), is used to assess the model goodness-of-fit and complexity (Spiegelhalter et al., 2002). In terms of model fit and complexity, the lower the DIC the better the model.

4. Model estimation results

Since the spatial models capture the spatial correlated effects, it is expected that the spatial models can better fit the data and produce more coherent results. The model estimation results and findings from both the classical count outcome model and spatial model using a full Bayesian approach are presented below.

4.1 Classical count outcome models

A series of fixed and random-effects NB models have been tested for both KSI accidents and slight injury accidents. A Hausman test has been performed and it is found that the random-effects NB model is suitable for the data used in this study. For each type of accidents, models for both balanced and unbalanced panel data are estimated. As such a total number of four classical count outcome models have been estimated. Year dummies have also been included in the models to control for the fixed time effects where year 2003 has been considered as the reference case. The model estimation results are presented in Table 2 and Table 3 (see the “classical count outcome model” column).

Table 2 and 3 are about here

As shown in Tables 2 and 3, model estimation results for balanced and unbalanced panel data are very similar to each other in terms of both the set of statistically significant variables and the magnitudes of their coefficients. *Traffic delay* (sec per km) is found to be statistically significant and positively related to the frequency of both KSI accidents and slight injury accidents. This means that the number of accidents increases with the increase in the level of traffic congestion. The coefficient of $\log(\text{delay in sec per km})$ indicates the elasticity of annual accidents with respect to *traffic delay*, suggesting that a 1% increase in *traffic delay* per km would increase KSI accidents in the region of 0.1% and slight injury accident by 0.05%. The effects of traffic congestion (i.e. delay) on road accidents will be further discussed in the next section.

AADT and road segment length are both statistically significant and positively associated with accidents in all models. This is expected as *AADT* is considered to be the main risk exposure to accidents. The elasticity of *AADT* for the case of KSI accidents is around 0.56 which is in line with the previous study in the UK by Bird and Hashim (2006). The elasticity of *AADT* for slight injury accidents is around 0.15, which appears a bit low compared to other studies (e.g. Aguero-Valverde and Jovanis, 2008). The coefficient of $\log(\text{segment length in m})$ is approximately 1 in all models suggesting that the elasticity of *road segment length* with respect to accidents is about 1. This means a 1% increase in *road segment length* would increase accident frequency by 1%.

As for the road segment characteristics, the minimum *radius* of horizontal curvature was found to be statistically significant (at a 90% confidence level for slight injury accidents) and positively related to accidents. This implies there are more accidents on straighter road segments, which is counter-intuitive at first glance and inconsistent with some studies (see Schneider et al., 2009 for truck accidents); but it is actually in line with previous studies (e.g. Haynes et al., 2007) which found road curvature is protective, especially at highly aggregated spatial units (because system-wide effects can be controlled for). *Gradient* which represents the vertical grade of the road segment was found to be statistically insignificant except in the balanced panel data model for KSI accidents (at a 90% confidence level). *Number of lanes* was statistically significant and positively associated with slight injury accidents suggesting more slight injury accidents would occur on roads with more lanes. *Speed limit* was found statistically insignificant in all models. This may be because there is not enough variation of this variable across the M25 and surround: 268 out of 298 road segments have the speed limit of 112 km/h. *Motorway* was included as a dummy variable to investigate whether accident frequency would be different on motorways. It is found that compared to A roads, motorways tend to have more slight injury accidents but less KSI accidents (at a 90% confidence level in the unbalanced panel data model).

Fixed time effects are significant and negative in KSI models, suggesting that KSI accidents tend to decrease in the years 2004-2007 compared to 2003. The expected number of

slight injury accidents, on the other hand, increases in the years 2004-2005 compared to 2003, and then decreases in the years 2006-2007.

4.2 Spatial models using full Bayesian hierarchical approach

Four spatial models were estimated using the full Bayesian hierarchical approach to take into account both spatial correlation and unobserved heterogeneity. The posterior means and standard deviations (S.D.) of the coefficients for the explanatory variables (β 's), time effects (δ_t), and the standard deviations of other random terms (v_i , u_i , e_{it}) were estimated using the MCMC method. Two chains were simulated with different initial values. The convergence of the two chains was examined by visual inspection of the MCMC trace plots. The initial 30,000 and 180,000 iterations were discarded as burn-ins to achieve convergence for KSI and slight injury accident models respectively. A further 30,000 iterations for each chain were performed and kept to calculate the posterior estimates of interested parameters for both KSI and slight injury accident models. For each category of accident (i.e. KSI accidents and slight injury accidents), a spatial model with fixed time effects and random time effects using a RW (1) prior were estimated. All spatial models are estimated using the balanced panel data. The results are presented in Table 2 and Table 3 (see the "Bayesian spatial model" column).

As indicated in Table 2, model estimation results from the two model specifications (i.e. the models with fixed time effects and RW (1) random time effects) are very similar to each other for the case of KSI accidents. Both model specifications produce similar posterior estimates in terms of coefficients of explanatory variables and standard deviations of random terms (i.e. u , e , and v). The DIC values of the two model specifications are very close to each other meaning that there is no significant difference between these models in terms of statistical fit and model complexity.

In the case of slight injury accidents (Table 3), since the model with RW (1) prior does not fully converge for a long period of simulation, the results from this model are considered unstable. Considering that the RW (1) model does not show any significant difference compared to the fixed time effect model, the results from the fixed time models for both categories of accidents will be used for further interpretation and discussion below.

As can be seen in Tables 2 and 3, the posterior estimates of standard deviation of u ranges from 0.16-0.23, suggesting that the spatial correlation effects (u) are significant for both types of accidents. The spatial correlation is higher for the case of slight injury accidents relative to KSI accidents, which is consistent with previous studies (Aguero-Valverde and Jovanis, 2006). Similarly uncorrelated heterogeneity (v) and space-time interaction effects (e) are also found to be statistically significant, although the space-time integration effects are marginal in the KSI accident models⁷. Much of the heterogeneity effects are captured by the spatial correlation terms.

Compared to the results from the classical count outcome models (i.e. random-effects NB models) for the same balanced panel dataset, most of the explanatory variables that are statistically significant in the classical models are also significant in the spatial models (Tables 2 and 3). The values of the coefficients are also fairly close to the estimates from the classical models. The notable exceptions are the effects of *AADT* and *number of lanes* in KSI

⁷ This may be due to the fact that KSI accidents have a low sample mean. Model estimation results may vary under low- and high-count accident situations. Since KSI accidents suffer from the low sample mean problem, separate tests for low and high KSI counts could not be performed. Models for low- and high-count of slight injury accidents have been tested, showing that for both low and high count situations the level of significance for the slope coefficients generally decreases but the model goodness-of-fit (in terms of AIC values) increases compared to the model for the total accident count. The sample size of the data used in this study is relatively small, so in terms of further studies the use of a larger sample would be useful.

Deleted:
Deleted:
Deleted: levels of
Deleted: with
Deleted: all observations
Deleted: may be needed using
Deleted: size

accident models, and *traffic delay*, *radius* and *gradient* in the slight injury accident models. The coefficient of $\log(AADT)$ decreases from 0.57 in the classical model to 0.23 in the spatial model for the case of KSI accidents (Table 2). This may be because the effects of *AADT* has been captured by other variables in spatial models, for example, the effect of the *number of lanes* is statistically insignificant in the classical model but becomes significant in the spatial model (Table 2).

For the case of slight injury accidents, *traffic delay* shows different results between the classical and spatial models. This variable is statistically significant and positively associated with the slight injury accidents in the classical model but becomes statistically insignificant in the spatial model (Table 3). This means *traffic delay* has no impact on the frequency of slight injury accidents according to the results from the spatial model. In addition, *radius* becomes significant at a greater confidence level (95% in the spatial model compared to 90% in the classical model). *Gradient* has now become positive and significant at a 90% confidence level in the spatial model while this variable was negative and statistically insignificant in the classical model. This result (positive and statistically significant) for *gradient* in the spatial model is in line with previous research findings (Shankar et al., 1995; Milton and Mannering, 1998). This suggests that the spatial model can produce more coherent results (i.e. better inference) compared to the classical count outcome models. In addition, the values of the intercepts between classical and Bayesian spatial models are also noticeably different. This may suggest that there is significant heterogeneity in the data and the Bayesian method seems to have captured this. Factors such as weather conditions and driver specific characteristics are not included in the models in this study due to the unavailability of data, which may lead to significant heterogeneity.

The residuals (the differences between observed and predicted values) of these two types of models have been plotted and presented in Figure 5. As illustrated in Figure 5, for KSI accidents, the pattern of residuals are very similar to each other between the two types of models. It can therefore be concluded that the classical and Bayesian models have similar fitting performance for the case of KSI accidents. As for slight injury accidents however, the figure clearly shows that the Bayesian spatial model is much better in terms of model fit compared to the classical Negative Binomial model. As can be seen, while the differences in the residuals are small between the two models for low count regime (observed count less than 20), the differences become larger for higher count regime (observed count greater than 20 and less than 60). The differences in the residual values are especially considerable for high count regime (observed count greater than 60). Considering that KSI accident count is typically in low count regime (less than 10), it appears that the classical count model performs equally well as to the Bayesian model for low count data according to the results in this study.

Figure 5 is about here

Given that the Bayesian spatial models are superior in terms of its underlying theory (e.g. taking account of spatial correlation and prior information), goodness of fit and model inference, it is believed that the Bayesian spatial models more accurately estimated the effects of covariates and better fit the data. Therefore the results from the Bayesian spatial models are preferred, and for the effects of *traffic delay* on the frequency of road accidents, it can be summarised that: *traffic delay* is positively associated with KSI accidents, namely a 1% increase in *traffic delay* (per km) will increase KSI accidents by 0.08%. *Traffic delay* has little or no impact on slight injury accidents.

It should however be noted that under some extreme conditions when traffic is at a very high congestion level (e.g. gridlock) the results above may not hold true. In extremely

congested conditions, vehicles move very slowly, which may reduce the number of KSI accidents. To test this, a random-effects NB model with quadratic terms (for traffic delay) was tested, and the result is presented in Table 4.

Insert Table 4 here

As can be seen in Table 4, the coefficient of *delay* continues to be positive and statistically significant. The coefficient of *delay squared* is however negative (though insignificant), meaning there may be a U-shaped relationship between congestion and the number of KSI accidents. The relationship between the predicted number of KSI accidents and the level of traffic congestion (for the situations on motorways in 2007, while other variables are held at their mean values), is presented in Figure 6.

Insert Figure 6

From Figure 6 it is clear that initially traffic congestion increases the number of KSI accidents. However, when the level of traffic congestion becomes extremely high, i.e. above $1.82\text{E}+06$ (sec per segment per year), the number of KSI accidents decreases. This suggests that improvements in traffic operations (often resulting in reduced traffic congestion) do not always lead to improved road safety. For example, as there may be situations (e.g. at high levels of traffic congestion) when reduced traffic delay may increase accidents. It should however be noted that such a high level of congestion is very rare. In the data used in this study, only 11 out of 1391 observations have traffic congestion above $1.82\text{E}+06$ (sec per segment per year). Therefore it is not surprising that overall traffic congestion increases the number of KSI accidents as implied in Table 2. Considering that such an extreme high level of traffic congestion is very rare, it can be concluded that generally increased traffic congestion increases the number of KSI accidents.

Deleted: that

5. Discussion and policy implication

This paper explores the relationship between traffic congestion and road accidents while controlling for other contributing factors such as AADT, radius and gradient. It has been found that the level of traffic congestion (measured by traffic delay per km length of roadway) is positively associated with the frequency of KSI accidents; while traffic congestion has little impact on slight injury accidents. The effects of other contributing factors have been found to be generally consistent with previous studies.

This result suggests that roadways with a high level of traffic congestion tend to have more KSI accidents. This result is consistent with the study by Kononov et al. (2008) who found that fatal and injury accidents increase with the increase in traffic congestion. . Other previous studies providing empirical evidence either found an inverse relationship between congestion and accidents (Baruya, 1998) or insignificant relationship (Noland and Quddus, 2005). Their studies, however, as discussed often lack an appropriate congestion measurement or suitable statistical methods. In order to evaluate the effects of congestion on road accidents, it is of importance to measure the congestion correctly. Instead of using a proxy for traffic congestion as used in some of the previous studies (Baruya, 1998; Noland and Quddus, 2005; Kononov et al., 2008), this study employed a direct measurement of traffic congestion. It is believed that the congestion measurement used in this study is valid and precise, as it reflects the nature of the congestion – traffic delay. Similar congestion measurements are used and recommended by the UK Department for Transport (DfT), for instance the “driving time lost per mile” (DfT, 2005). Other congestion measurements have

also been tested during this study, for example, “congestion index”⁸ detailed by Taylor et al. (2000) and used by Wang et al. (2009), and generally similar results have found for the unbalanced panel data models. This further confirms the relationship between traffic congestion and road accidents found in this study on the M25 and surround.

Although increased traffic congestion is expected to decrease road fatalities because of lower average speed under traffic congestion as proposed by Shefer and Rietveld (1997), the speed variance among vehicles within and between lanes may not be necessarily low. In contrast, speed variance would be large in congested situations because drivers need to adjust speed frequently (e.g., “stop-go-stop” actions), which significantly increases the complexity of driving. In this scenario drivers may not have enough time to react (i.e. short time between “go-stop” actions), so more conflicts are expected. It has been argued that it is the speed variance that causes safety problem instead of speed itself (Lave, 1985), thus the increased speed variance in congested situations may result in more serious injury accidents. There is also empirical evidence that the go-stop actions in congestion would increase the chances of collisions (Hanbali and Fornal, 1997). In addition, traffic speeds even in congested situations are likely to be relatively high on major roads compared to other parts of the road network. As such, any accidents occurring are likely to be more severe. This may explain the result that traffic congestion has greater effects on KSI accidents than on slight injury accidents on the M25 and surround.

Driving behaviour also plays an important role in increased accidents under congestion conditions. Driving under congestion conditions is different than under free-flow conditions. It is well established that driving behaviour becomes worse in the presence of congestion. For example, it has been found that congestion would result in higher driver stress leading to increased aggressive driving behaviour such as purposeful tailgating and yelling at others (Hennessy and Wiesenthal, 1997). Shinar (1998) showed that congestion is associated with aggressive behaviour such as honking of horns. A recent study by Shinar and Compton (2004) found that there was a strong linear association between congestion and aggressive behaviour and drivers are more likely to behave aggressively during weekday rush hour when the value of time is high, such as cutting across one or more lanes in front of another driver and passing on the shoulder. The results found in this paper would not be surprising since increased congestion would cause more aggressive driving, and aggressive driving is a major road safety concern (Shinar and Compton, 2004) which would subsequently result in more accidents.

The findings from this study added to the debate about the relationship between mobility and safety, showing that mobility and safety can be improved simultaneously. Therefore, there is significant *additional* benefit of reducing traffic congestion, in the sense that it not only reduces the costs relating to increased travel time (i.e. traffic delay) but also benefits involved by preventing accidents.

The findings from this study are useful for transport policy makers in safety improvement. Since traffic congestion imposes safety problems, it is necessary to reduce it for improved road safety. Some measurements could be reinforced to optimise and smooth the traffic flow, which would be beneficial to both congestion and accident reduction. For example, an electronic warning sign can be displayed in real-time at some sites when congestion occurs, making drivers more prepared for the congestion ahead and the associated risk of accidents. Information on congestion has already been displayed on many motorways throughout the UK. Empirical evidences from Hanbali and Fornal (1997) found that adaptive traffic signal systems are very effective in reducing both traffic congestion and road accidents

⁸ Defined as average vehicle delay per free flow travel time.

at intersections, especially the “stop-and-go” driving related collisions. The traffic adaptive traffic signal systems involve the installation of a closed-loop, traffic-responsive signal system at intersections, changing the green-to-cycle (g/c) ratio for the congested roadways to reduce traffic congestion. Another promising measurement is an advanced traffic management system, such as the Active Traffic Management (ATM)⁹ which is currently in operation on the M42 motorway in the UK. The ATM monitors the traffic flows on the roads using sensor loops and automatically calculates the best speed limit for the current traffic. The optimised speed limits are displayed on electronic overhead gantries. A variable speed limit is known to be able to increase homogeneity of driving speeds (i.e. optimise and smooth the traffic flow) and consequently improve road safety (Van Nes et al., 2010). Other information can also be displayed, for example, drivers can be directed to use the hard shoulder during congested periods so as to smooth the traffic flow. The ATM system has originally been introduced for reducing traffic congestion; the preliminary monitoring results showed that the ATM also reduced accidents from 5.1 to 1.8 a month (monthly average based on 6 years’ data) despite initial safety concerns (Highways Agency, 2008). Similar schemes could also be introduced on other major roads in the UK.

The implementation of these measurements (e.g. ATM) is traditionally justified in terms of congestion reduction. The results from this study re-confirm the benefit of these measures, as in addition to congestion reduction, these measurements are also beneficial to road safety, adding the additional benefits of implementing these measurements. Therefore, the findings from this study are helpful for transport policy makers in devising safety programmes and allocating highway funds.

As discussed above, it was argued that the positive association between traffic congestion and the number of KSI accidents on major roads is partially due to the increased speed variance and worse driving behaviours in congested situations. Therefore some measures could be introduced to reduce speed variance and improve driving behaviour in congested situations. As for reducing speed variance, “average speed check cameras” can be installed on a stretch of a roadway that has been identified as an accident hotspot to enforce a suitable “average speed” on the roadway. This ensures that all vehicles travel at a similar and consistent speed on the stretch of the roadway, which reduces speed variance and improves road safety. Transport Scotland (2009) recently reported that the enforcement of a consistent speed of 40mph (64km/h) on the M80 had successfully enhanced road safety during M80 road works. In addition to the enforcement of average speed, inappropriate speed should be avoided. The inappropriate speed includes excessive speed and very low speed. Slow drivers are as much hazardous as fast ones since they increase speed variance (Lave, 1985), which is especially a concern for major roads with high posted speed limits. Therefore *minimum* speed limits could be introduced to enforce the minimum speed that drivers are required to drive on major roads. Minimum speed limits are operated in many countries outside the UK. Similar policies for minimum speed limits could be introduced to the UK to ban inappropriate low speed on major roads for safety.

Bad driving behaviour in congested situations may be more difficult to tackle. As suggested by Shinar and Compton (2004), some measures could be implemented to reduce aggressive driving behaviours in congested situations, such as providing more information to urge drivers to take alternative routes during rush hour, encouraging car sharing and flextime, enhancing law enforcement and promoting driver education. A campaign highlighting the dangers of aggressive driving in congested situations may be useful for improving road safety on busy major roads.

⁹ See Highways Agency website (<http://www.highways.gov.uk/knowledge/1361.aspx>) for more details.

6. Conclusions and further research

This paper has examined the relationship between traffic congestion and frequency of road accidents on the M25 and surround in England. Several classical count models and spatial models using a full Bayesian approach have been developed to investigate the effects of traffic congestion on road accidents. While the results from the classical and spatial models are generally similar to each other, there are some inconsistencies between them for some variables, for example, the effects of traffic congestion on slight injury accidents. The results from the spatial models are argued to be preferable as they accommodate spatial correlations and better fit the data.

From the model estimation results it has been found that traffic congestion is positively associated with the frequency of KSI accidents: a 1% increase in traffic delay per km would increase KSI accidents by about 0.1%. Traffic congestion has little impact on slight injury accidents. The reasons for this may be due to increased speed variance among vehicles within and between lanes and bad driving behaviour under congested situations. Other contributing factors have also been controlled for and have provided consistent results with previous studies. For example, a 1% increase in traffic flow has found to be associated with a 0.2% increase in KSI accidents. Based on this finding, some measures could be introduced to improve road safety.

There are some limitations in this study. Firstly, data have been aggregated to an annual level in order to obtain sufficient count of accidents. An analysis using aggregated data also has the benefit of taking into account “system-wide effects” (Noland and Oh, 2004). Such aggregation may however be subject to ecological fallacy (Davis, 2002) and thus the results may be biased. In addition, the number of accidents on the road segments in this study is rather low, especially for the KSI accidents and slight injury accidents may suffer from the problem of under-reporting.

This paper is focusing on the M25 and its surrounding major roads (motorways or A roads). Therefore, further research is required to investigate the effect of traffic congestion on road accidents on other major roads in the UK and other countries. Road infrastructure and traffic characteristics in other parts of the UK or in other countries may be different to the M25 and surround, for instance roads in rural or urban settings; and the effect of traffic congestion on different road users (e.g., motorised and non-motorised transport) may also be different. Also there is not enough variation in the speed limit on major roads (most road segments employed in this study have speed limit of 112 km/h), which means that the speed limit cannot be used, as a proxy for any other speed related effects. Further studies on both major and minor roads where the speed limit varies greatly would be beneficial. Further studies for other major roads can test the model developed here and offer further empirical evidence, which may eventually provide a conclusive statement on the relationship between traffic congestion and road accidents. This would widen the potential use of the findings from this paper. In addition to major roads, the impact of traffic congestion on other types of roads or junctions also needs to be examined.

Deleted: of
Deleted: are
Deleted: in this study
Deleted: could not act
Deleted: n appropriate
Deleted: , so f
Deleted: to a greater extent
Deleted: may

References

- 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 with spatial models. *Transportation Research Record: Journal of the Transportation Research Board*, **2061**, pp. 55-63.

Anastasopoulos, P.C. and Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention*, **41**(1), pp. 153-159.

Baruya, A., 1998. *MASTER: Speed-accident relationships on different kinds of European roads*. Deliverable D7. UK: Transport Research Laboratory.

Belmont, D.M. and Forbes, T.W., 1953. Effect of average speed and volume on motor-vehicle accidents on two-lane tangents, *Proceedings of the Highway Research Board*. pp383-395.

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**, pp. 93-99.

Cameron, A.C. and Trivedi, P.K., 1998. Regression analysis of count data. Econometric Society Monographs. Cambridge: Cambridge University Press.

Ceder, A., 1982. Relationships between road accidents and hourly traffic flow—II : Probabilistic approach. *Accident Analysis & Prevention*, **14**(1), pp. 35-44.

Chin, H.C. and Quddus, M.A., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accident Analysis & Prevention*, **35**(2), pp. 253-259.

Davis, G.A., 2002. Is the claim that ‘variance kills’ an ecological fallacy? *Accident Analysis & Prevention*, **34**(3), pp. 343-346.

Department for Transport, 2005. *Perceptions of congestion: report on qualitative research findings (Appendix E: Congestion measure sheets)*.

El-Basyouny, K. and Sayed, T., 2009a. Accident prediction models with random corridor parameters. *Accident Analysis & Prevention*, **41**(5), pp. 1118-1123.

El-Basyouny, K. and Sayed, T., 2009b. Urban Arterial Accident Prediction Models with Spatial Effects. *Transportation Research Record: Journal of the Transportation Research Board*, **2102**, pp. 27-33.

Gwynn, D.W., 1967. Relationship of accident rates and accident involvements with hourly volumes. *Traffic Quarterly*, **21**(3), pp. 407-418.

- Hanbali, R. and Fornal, C., 1997. Methodology for evaluating effectiveness of traffic-responsive systems on intersection congestion and traffic safety. *Transportation Research Record: Journal of the Transportation Research Board*, **1603**(-1), pp. 137-149.
- Hausman, J.A., 1978. Specification tests in econometrics. *Econometrica*, **46**(6), pp. 1251-1271.
- Hausman, J.A., Hall, B.H. and Griliches, Z., 1984. Econometric models for count data with an application to the Patents-R & D relationship. *Econometrica*, **52**(4), pp. 909-938.
- 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.
- Hennessy, D.A. and Wiesenthal, D.L., 1997. The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics*, **40**(3), pp. 348.
- Highways Agency, 2008. *ATM Monitoring and Evaluation 4-Lane Variable Mandatory Speed Limits 12 Month Report (Primary and Secondary Indicators)*. Available: <http://www.ha-research.gov.uk/projects/projectdocuments.php?method=download&ID=262>.
- Kononov, J., Bailey, B. and Allery, B., 2008. Relationships between safety and both congestion and number of lanes on urban freeways. *Transportation Research Record: Journal of the Transportation Research Board*, **2083**, pp. 26-39.
- Kynn, M., 2006. *Missing data in winBUGS*. Centre for Applied Statistics, Lancaster University.
- Lagazio, C., Dreassi, E. and Biggeri, A., 2001. A hierarchical Bayesian model for space-time variation of disease risk. *Statistical Modelling: An International Journal*, **1**(1), pp. 17-29.
- Lave, C.A., 1985. Speeding, coordination, and the 55 MPH limit. *The American Economic Review*, **75**(5), pp. 1159-1164.
- Lee, S., Choi, J. and Kim, S.W., 2010. Bayesian approach with the power prior for road safety analysis. *Transportmetrica*, **6**(1), pp. 39.
- 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., Manar, A. and Vizioli, A., 2005. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accident Analysis & Prevention*, **37**(1), pp. 185-199.
- 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.

- Noland, R.B. and Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. *Accident Analysis & Prevention*, **36**(4), pp. 525-532.
- 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.
- Persaud, B., Lan, B., Lyon, C. and Bhim, R., 2010. Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. *Accident Analysis & Prevention*, **42**(1), pp. 38-43.
- 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.
- Schneider, W., Zimmerman, K., Van Boxel, D. and Vavilikolanu, S., 2009. Bayesian Analysis of the Effect of Horizontal Curvature on Truck Crashes Using Training and Validation Data Sets. *Transportation Research Record: Journal of the Transportation Research Board*, **2096**, pp. 41-46.
- Shankar, V., Albin, R., Milton, J. and Mannering, F., 1998. Evaluating median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effects Negative Binomial model. *Transportation Research Record: Journal of the Transportation Research Board*, **1635**(-1), pp. 44-48.
- 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., 1998. Aggressive driving: the contribution of the drivers and the situation. *Transportation Research Part F: Traffic Psychology and Behaviour*, **1**(2), pp. 137-160.
- 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.

Thomas, A., Best, N., Lunn, D., Arnold, R. and Spiegelhalter, D., 2004. GeoBUGS user manual version 1.2.

Transport Scotland, 19/10/2009-last update, average speed cameras to enhance road safety during M80 works | transport scotland. Available: <http://www.transportscotland.gov.uk/news/average-speed-cameras-road-safety-M80-works> [19/12/2009, 2009].

Van Nes, N., Brandenburg, S. and Twisk, D., 2010. Improving homogeneity by dynamic speed limit systems. *Accident Analysis & Prevention*, **42**(3), pp. 944-952.

Wang, C., Quddus, M.A. and Ison, S.G., 2009. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. *Accident Analysis & Prevention*, **41**(4), pp. 798-808.

Zhou, M. and Sisiopiku, V., 1997. Relationship between volume-to-capacity ratios and accident rates. *Transportation Research Record: Journal of the Transportation Research Board*, **1581**(-1), pp. 47-52.

Table 1 Summary statistics of the variables.

Variable	Mean	Standard Deviation	Min	Max
<i>Accident</i>				
Fatal and serious injury accidents	1.071	1.418	0	10
Slight injury accidents	7.774	8.999	0	93
<i>Traffic characteristics</i>				
Total delay (sec per km)	197,279.7	242,742.9	622.865	1,900,374
Annual average daily traffic (AADT) (vehicles)	45,675.48	20,667.55	5.918	98,394.83
Average vehicle speed (km/h)	84.918	14.757	33.663	118.134
<i>Road segment characteristics</i>				
Segment length (km)	4.980	3.611	0.32	22.08
Minimum radius (m)	674.331	364.787	4.94	2,000
Maximum gradient (%)	3.175	1.337	0.6	8
Number of lanes	2.904	0.715	1	6
Speed limit (km/h)	109.628	7.997	64	112
Motorway indicator	0.689	0.463	0	1

Table 2 Model estimation results for KSI accidents

Classical count outcome model					Bayesian spatial model			
Model for balanced panel data			Model for unbalanced panel data		Model with fixed time effects		Model with RW(1) time effects	
Variables	Coefficient	z value	Coefficient	z value	Mean	S.D.	Mean	S.D.
log(delay in sec per km)	0.115**	2.93	0.120**	3.17	0.081*	0.044	0.088**	0.044
log(AADT)	0.568**	5.01	0.560**	5.14	0.232**	0.079	0.244**	0.095
log(segment length in m)	0.941**	13.99	0.945**	14.39	0.974**	0.073	0.991**	0.061
log(minimum radius)	0.216**	3.07	0.203**	2.97	0.252**	0.069	0.255**	0.069
Maximum gradient (%)	0.061*	1.67	0.052	1.47	0.077*	0.041	0.079*	0.042
Number of lanes	0.018	0.24	0.036	0.5	0.207**	0.079	0.206**	0.079
Speed limit	0.009	1.22	0.009	1.32	0.01	0.01	0.013	0.009
Motorway	-0.149	-1.38	-0.173*	-1.7	-0.089	0.176	-0.111	0.175
Year 2003	-	-	-	-	0	-	0.189**	0.05
Year 2004	-0.195**	-2.34	-0.210**	-2.57	-0.187**	0.081	0.038	0.046
Year 2005	-0.221**	-2.64	-0.219**	-2.66	-0.211**	0.08	-0.011	0.045
Year 2006	-0.390**	-4.46	-0.383**	-4.46	-0.375**	0.085	-0.134**	0.05
Year 2007	-0.280**	-3.25	-0.285**	-3.36	-0.284**	0.083	-0.082	0.051
Intercept	-15.152**	-9.29	-15.168**	-9.64	15.071**	1.504	-15.995**	1.536
Parameter <i>a</i>	75.662**	2.476	77.716**	2.533	0.164** 0.079 0.067** 0.058 0.344** 0.104 - - 3225.6		0.170**	0.075
Parameter <i>b</i>	6.444**	4.157	6.799**	4.142			0.062**	0.053
Log likelihood	-1624.98		-1701.9				0.342**	0.096
S.D. (<i>u</i>)							0.177**	0.098
S.D. (<i>e</i>)								
S.D. (<i>v</i>)								
S.D. (<i>t</i>)								
AIC/DIC	3279.954		3433.801				3224.31	
<i>N</i>	1330		1391		1330		1330	

* Statistically significantly different from zero ($p < 0.1$ or 90% credible sets show the same sign),

** Statistically significantly different from zero ($p < 0.05$ or 95% credible sets show the same sign)

Table 3 Model estimation results for slight injury accidents

Variables	Classical count outcome model				Bayesian spatial model			
	Model for balanced panel data		Model for unbalanced panel data		Model with fixed time effects		Model with RW(1) time effects†	
	Coefficient	z value	Coefficient	z value	Mean	S.D.	Mean	S.D.
log(delay in sec per km)	0.048**	2.02	0.050**	2.15	0.044	0.032	0.02	0.034
log(AADT)	0.144**	3.44	0.149**	3.57	0.129**	0.036	0.148**	0.052
log(segment length in m)	0.852**	13.84	0.870**	14.23	0.925**	0.067	0.726**	0.207
log(minimum radius)	0.094*	1.77	0.096*	1.84	0.138**	0.07	-0.019	0.213
Maximum gradient (%)	-0.002	-0.05	-0.014	-0.4	0.068*	0.04	0.115*	0.099
Number of lanes	0.424**	6.37	0.419**	6.41	0.455**	0.071	0.421*	0.176
Speed limit	-0.003	-0.6	-0.004	-0.76	-	-	-	-
Motorway	0.260**	2.75	0.230**	2.51	0.278**	0.143	0.348**	0.222
Year 2003	-	-	-	-	0	-	-0.031	0.026
Year 2004	0.118**	3.32	0.121**	3.4	0.115**	0.037	0.068**	0.023
Year 2005	0.081**	2.25	0.086**	2.39	0.079**	0.038	0.042	0.022
Year 2006	0.024	0.64	0.029	0.78	0.028	0.039	-0.006	0.023
Year 2007	-0.066*	-1.73	-0.056	-1.5	-0.048	0.04	-0.073**	0.026
Intercept	-7.765**	-8.3	-7.889**	-8.53	10.61**	0.857	-13.473**	6.137
Parameter <i>a</i>	13.964**	6.029	13.288**	6.26	0.230** 0.061 0.179** 0.018 0.499** 0.048 - - - -		0.276** 0.169 0.182** 0.018 0.716** 0.357 0.128** 0.069	
Parameter <i>b</i>	3.620**	8.602	3.549**	8.832				
Log likelihood	-3281.42		-3437.06					
S.D. (<i>u</i>)								
S.D. (<i>e</i>)								
S.D. (<i>v</i>)								
S.D. (<i>t</i>)								
AIC/DIC	6592.842		6904.122		6133.92		6154.35	
<i>N</i>	1330		1391		1330		1330	

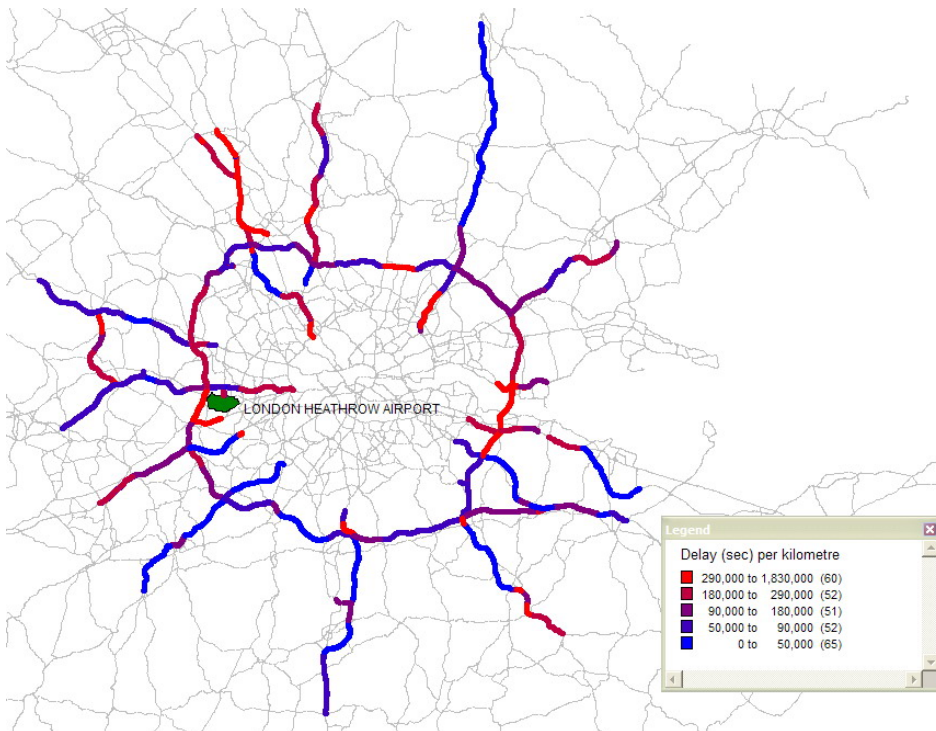
* Statistically significantly different from zero ($p < 0.1$ or 90% credible sets show the same sign),** Statistically significantly different from zero ($p < 0.05$ or 95% credible sets show the same sign)

† Model not fully converged

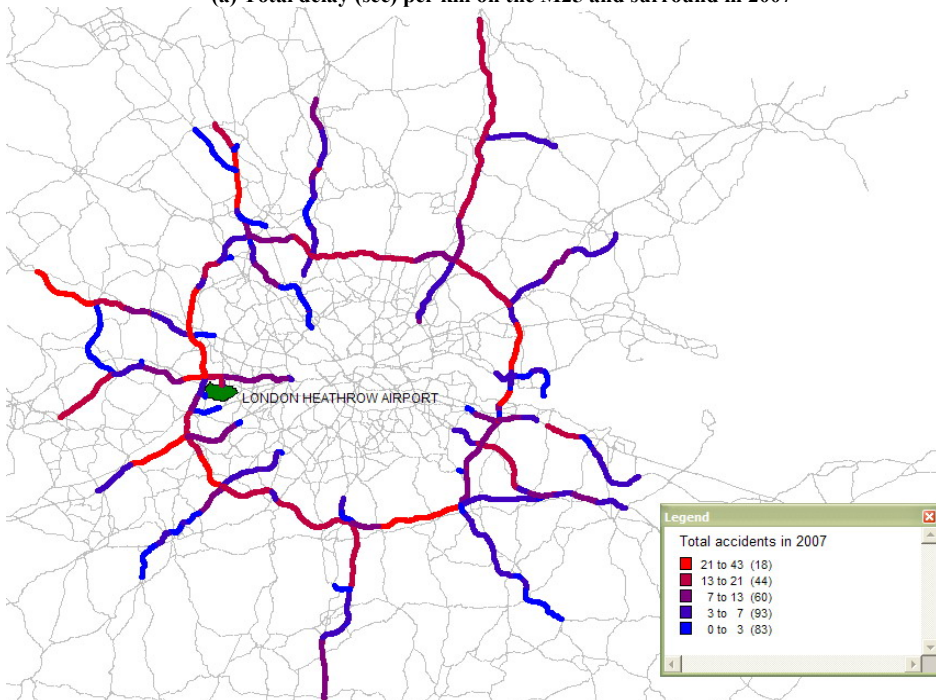
Table 4 Random-effects NB model with quadratic terms for KSI accidents

Variables	Coefficient	Std. Err.	z value
Delay (hour per km)	0.003**	0.001	2.09
Delay ² (hour per km) ²	-0.000003	0.000004	-0.73
log(AADT)	0.594**	0.107	5.53
log(segment length in m)	0.943**	0.065	14.43
log(minimum radius)	0.206**	0.068	3.01
Maximum gradient (%)	0.057	0.035	1.63
Number of lanes	0.018	0.071	0.25
Speed limit (km/h)	0.010	0.007	1.48
Motorway	-0.160	0.102	-1.57
Year 2004	-0.213**	0.082	-2.6
Year 2005	-0.219**	0.082	-2.65
Year 2006	-0.391**	0.086	-4.53
Year 2007	-0.288**	0.085	-3.39
Intercept	-14.390**	1.462	-9.84
Parameter <i>a</i>	77.940**	30.423	2.56
Parameter <i>b</i>	6.994**	1.711	4.09
Log likelihood	-1701.853		
AIC	3435.706		
<i>N</i>	1391		

** p<0.05



(a) Total delay (sec) per km on the M25 and surround in 2007



(b) Total number of accidents on the M25 and surround in 2007

Figure 1 Spatial distribution of total delay and number of accidents on the different road segments on the M25 and surround.

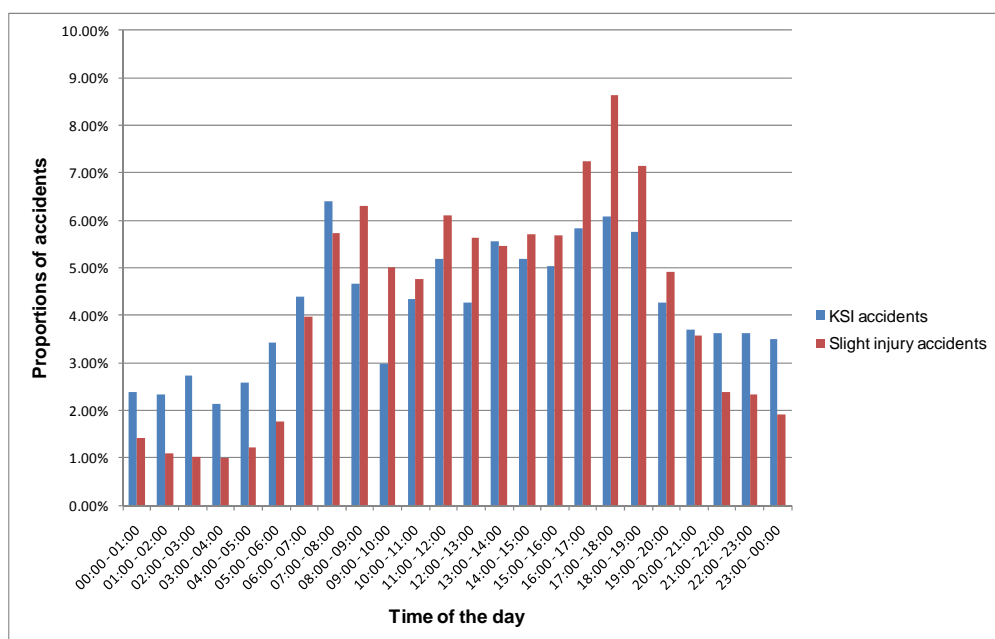


Figure 2 Proportions of accidents occurred over a day

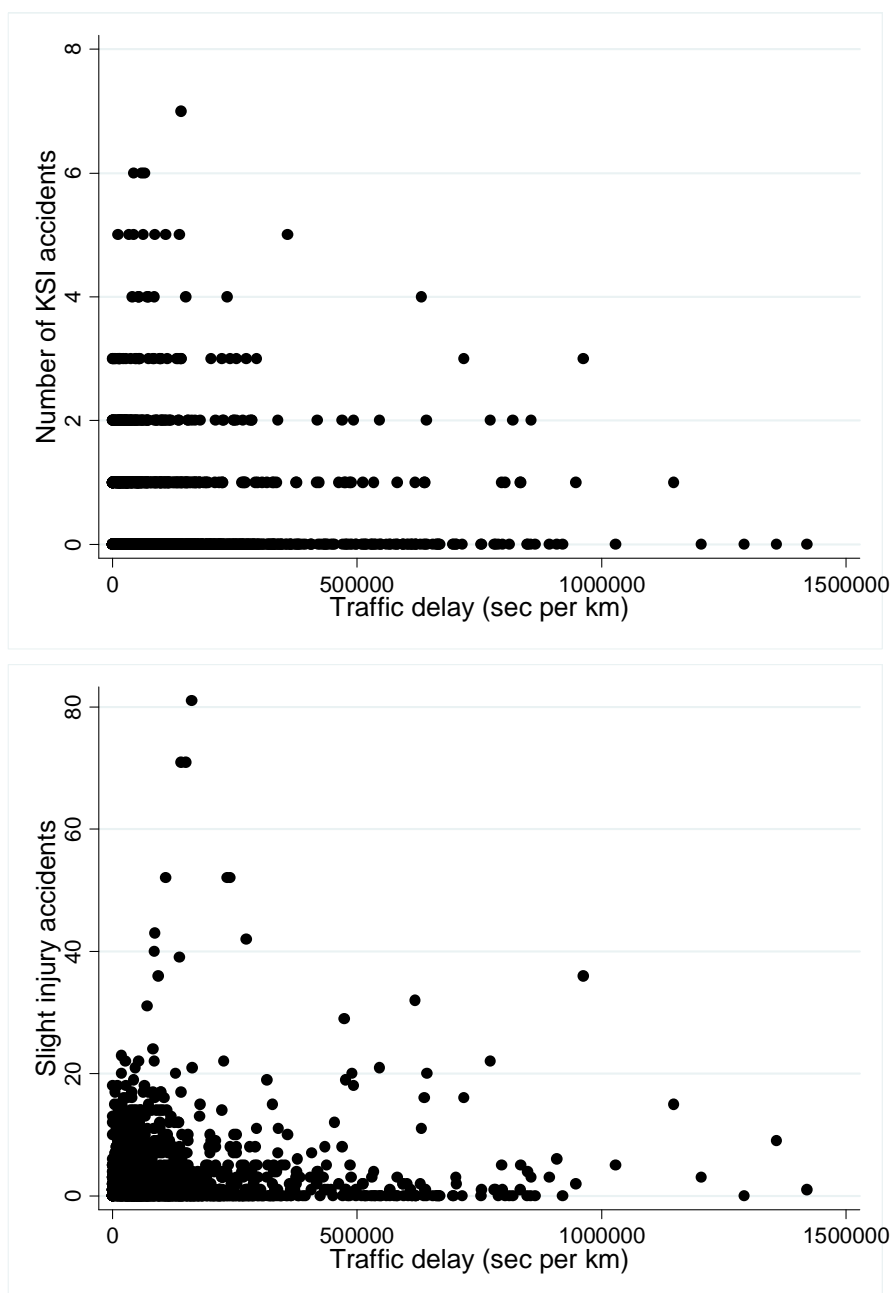


Figure 3 Road accidents and traffic delay during peak time

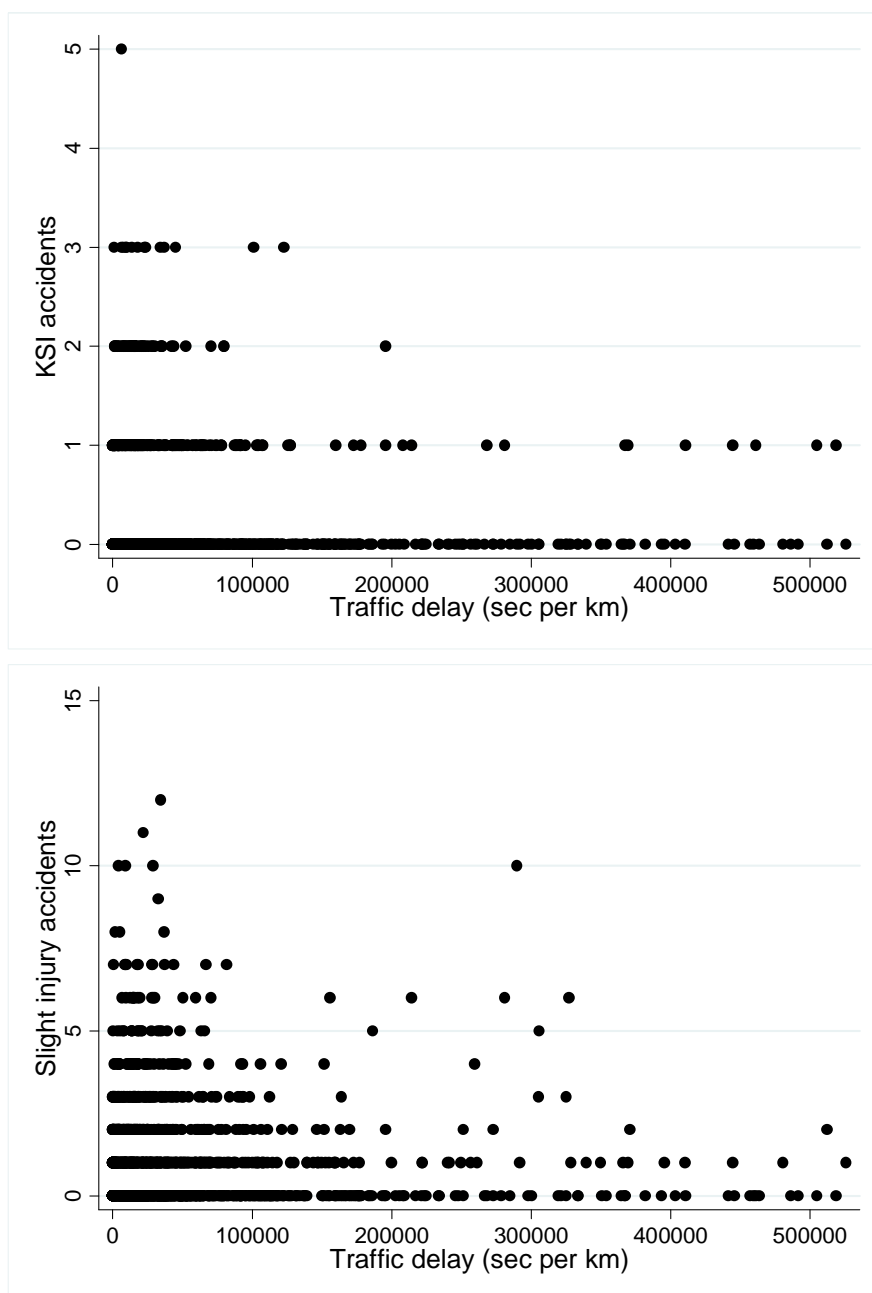


Figure 4 Road accidents and traffic delay during [off-peak](#) time

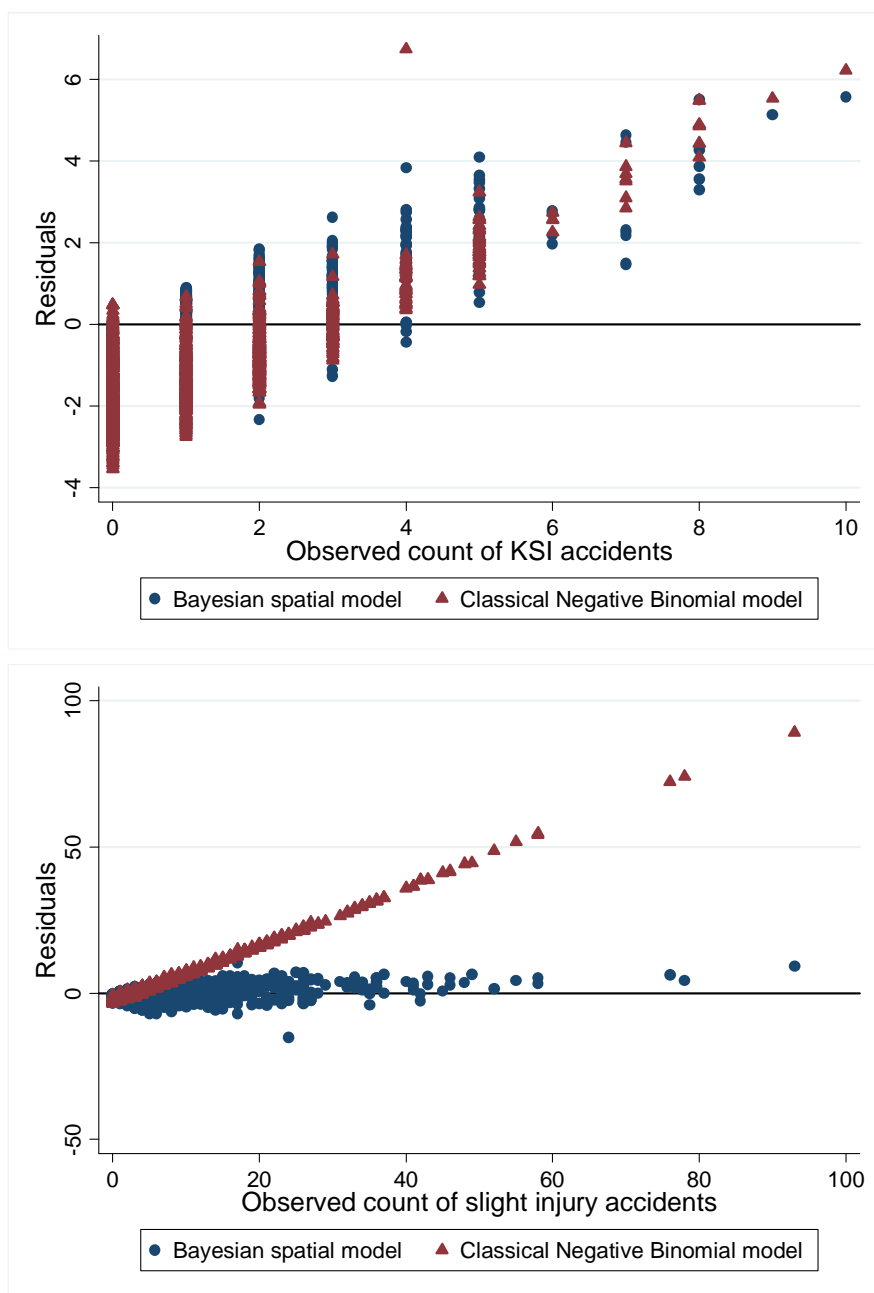


Figure 5 Comparison of residuals: Bayesian spatial model vs. classical NB model

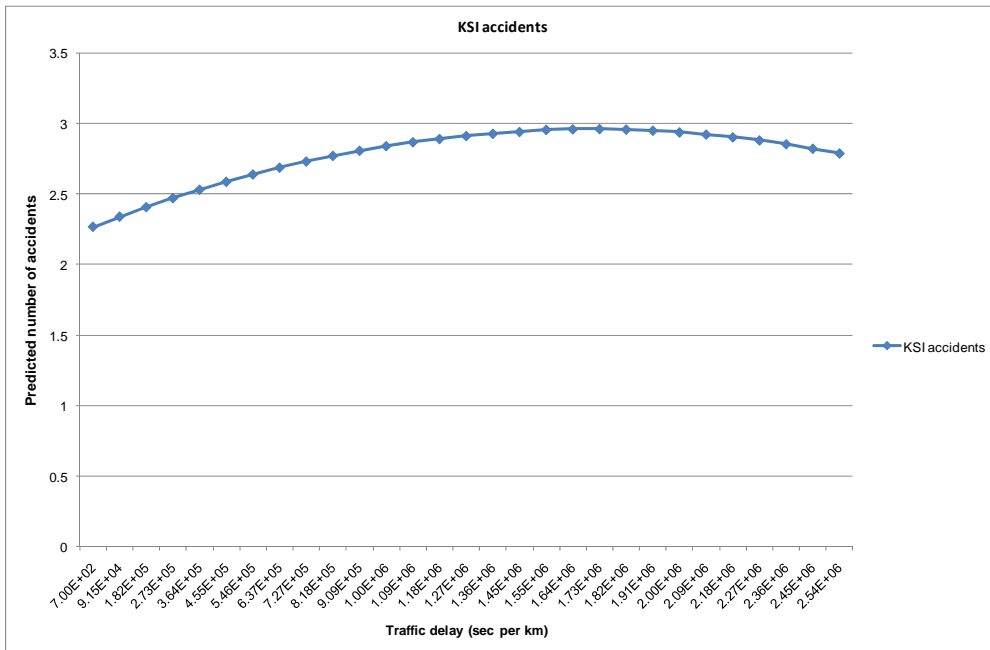


Figure 6 Predicted numbers of KSI accidents with respect changes in traffic delay