



Macroscopic road safety impacts of public transport: A case study of Melbourne, Australia

Long T. Truong^{a,*}, Graham Currie^b

^a School of Engineering and Mathematical Sciences, La Trobe University, Melbourne, Australia

^b Public Transport Research Group, Monash University, Melbourne, Australia

ARTICLE INFO

Keywords:

Public transport
Transit
Safety
Crashes
Negative binomial
Conditional autoregressive

ABSTRACT

Mode shift from private vehicle to public transport is often considered as a potential means of improving road safety, given public transport's lower fatality rates. However, little research has examined how public transport travel contributes to road safety at a macroscopic level. Further, there is a limited understanding of the individual effects of different public transport modes. This paper explores the effects of commuting by public transport on road safety at a macroscopic level, using Melbourne as a case study. A random effect negative binomial (RENB) and a conditional autoregressive (CAR) model are adopted to explore links between total and severe crash data to commuting mode shares and a range of other zonal explanatory factors. Overall, results show the great potential of public transport as a road safety solution. It is evident that mode shift from private vehicle to public transport (i.e. train, tram, and bus), for commuting would reduce not only total crashes, but also severe crashes. Modelling also demonstrated that CAR models outperform RENB models. In addition, results highlight safety issues related to commuting by motorbike and active transport. Effects of sociodemographic, transport network, and land use factors on crashes at the macroscopic level are also discussed.

1. Introduction

Overall, public transport is a relatively safe travel mode compared to private vehicle, in terms of fatality rates per trips and per passenger distance travelled (Evans, 1994; Beck et al., 2007; Cairney, 2010; Savage, 2013). For example, private vehicle occupants have a fatality rate per person trip rate which is 23 times larger than those for bus passengers (Beck et al., 2007). The fatality rate per passenger miles for car and light truck is 17–66 times higher compared to rail, light rail, and bus (Savage, 2013). As a result, mode shift from private vehicle to public transport is often promoted as a potential means of improving road safety (Litman, 2016).

However, little research has investigated how public transport travel contributes to road safety at a macroscopic level. The little research conducted in this area has mixed results. For example, Moeinaddini et al. (2015) showed that the percentage of journey to work by public transport was negatively associated with the number of road deaths in European cities. In contrast, Tasic and Porter (2016) found that the percentage of journey to work by public transport was positively associated with vehicular crashes, non-motorised crashes, and fatal vehicular crashes in census tract areas in Chicago, US. Likewise, Dong et al. (2016) suggested a higher proportion of commuters

using cars would reduce crashes at the traffic analysis zone level in Florida, US. Wang et al. (2017) found the opposite effects of the proportion of commuting by public transport on motor crashes and pedestrian crashes in Florida, US. While macroscopic road safety effects of public transport have often been explored using journey to work data, it remains relatively unclear if commuting by public transport to or from a zone (i.e. by origin or by destination) has different effects on zonal safety. Furthermore, previous research tends to focus on the road safety effects of overall public transport. As a result, little is known about the individual effects of public transport modes (e.g. train, tram, and bus) at the macroscopic level.

In macroscopic safety studies, road safety measures (e.g. crashes and crash casualties) aggregated at a certain spatial unit (e.g. statistical areas, counties, and traffic analysis zones) are analysed in relation to zonal or area-wide characteristics (Karlaftis and Tarko, 1998; Hadayeghi et al., 2003; Lovegrove and Sayed, 2006; Cai et al., 2017; Lee and Abdel-Aty, 2018). This enables the consideration of road safety from a comprehensive network-wide perspective. A growing body of literature has been looking at total crashes, vehicular crashes, bicycle crashes, and pedestrian crashes at the macroscopic level, considering a wide range of zonal characteristics. Trip generation and mode shares have been shown as crash predictors. For example, Abdel-Aty et al.

* Corresponding author.

E-mail address: L.Truong@latrobe.edu.au (L.T. Truong).

(2011) indicated that trip production and attraction were highly associated with total crashes and peak-hour crashes in traffic analysis zones in Florida. In addition to the effects of public transport mode shares as discussed above, the effects of other modes on crash occurrences have also been investigated. Abdel-Aty et al. (2013) found that commuting by walking was associated with increases in total and severe crashes in Florida. Likewise, Tasic and Porter (2016) suggested commuting by walking would increase motorised and non-motorised crashes in Chicago. Using data from European cities, Moeinaddini et al. (2015) found that commuting by motorbike was associated with an increase in road deaths. Several studies also showed effects of commuting by walking and cycling on increasing pedestrian and bicycle crashes (Amoh-Gyimah et al., 2016; Cai et al., 2016).

Macroscopic safety has also been examined with various socio-demographic characteristics, e.g. population, population density (Noland and Quddus, 2004; Lee et al., 2014b), age groups (Guevara et al., 2004; Aguero-Valverde and Jovanis, 2006; Amoh-Gyimah et al., 2017), income (Noland, 2003; Pirdavani et al., 2012, 2017), and employment (Siddiqui et al., 2012; Coruh et al., 2015). For example, population has been widely employed as an indicator for crash exposure in previous macroscopic safety studies (Karlaftis and Tarko, 1998; Huang et al., 2010; Lee et al., 2014a; Lee and Abdel-Aty, 2018). The proportion of young people has been shown as an important crash predictor (Guevara et al., 2004; Amoh-Gyimah et al., 2017).

There have also been various studies investigating effects of transport network characteristics, e.g. intersection density (Hadayeghi et al., 2003; Lovegrove and Sayed, 2006; Jiang et al., 2016), public transport density (Wang and Kockelman, 2013; Lee et al., 2015a; Tasic and Porter, 2016), road types and speeds (Amoros et al., 2003; Lovegrove and Sayed, 2006; Quddus, 2008; Abdel-Aty et al., 2013), road network structures (Wang et al., 2013), and vehicle kilometre travelled (Karlaftis and Tarko, 1998; Dumbaugh and Rae, 2009; Huang et al., 2010, 2016). For example, crashes tend to increase with higher intersection densities (Hadayeghi et al., 2003) or more signalised intersections (Hadayeghi et al., 2007). Likewise, more public transport stops or higher stop densities tend to be associated with more crashes (Wang and Kockelman, 2013; Tasic and Porter, 2016).

It has been widely accepted that land use in proximity of events influences road crashes (Pulugurtha et al., 2013). For example, crashes tend to increase with a higher proportion of commercial and industrial land areas (Hadayeghi et al., 2007). Pulugurtha et al. (2013) found that mixed use developments were associated with an increase in crashes. There was some evidence about associations between crashes and land use mix, e.g. entropy and balance measures (Wang and Kockelman, 2013; Amoh-Gyimah et al., 2017). The influence of healthcare (e.g. hospital density) and environment characteristics (e.g. rainfall) have been also explored in previous macroscopic safety research (Ng et al., 2002; Coruh et al., 2015; Truong et al., 2016).

In 2018, there were nearly 1150 road deaths in Australia and many more serious injuries, costing Australia around 1.7% of its GDP (BITRE, 2007, 2019). The current national road safety strategy's goal is unlikely to be met by 2020, despite significant investments on road safety. Currently the potential road safety benefit of public transport is largely ignored in current Australian transport policies, road safety action plans, and the Safe System vision statements (ATC, 2011; TIC, 2016).

The aim of this paper is to explore the effects of commuting by public transport on road safety at a macroscopic level, using Melbourne, Australia as a case study. Melbourne's public transport includes trains, buses, and the world's largest tram network. Effects of commuting using public transport (i.e. train, tram, and bus) by origin and by destination are explicitly considered while controlling for sociodemographic, transport network, and land use factors.

The paper is structured as follows; the next section describes the method and data adopted in the research. Results are then presented. The paper concludes with a discussion of the implications of the findings for policy and research.

2. Data and method

2.1. Data

Crash data in Greater Melbourne, Victoria, Australia between 2014 and 2018 were obtained from Victoria's open data directory (Data.Vic). The number of total crashes and the number of severe crashes (fatal or serious injury) were aggregated at statistical areas level 2 (SA2s). SA2s, which are medium-sized functional areas within the Australian Statistical Geography Standard (ASGS), provide journey to work data by both origin and destination. SA2s in Greater Melbourne has an average area of 32.3 square kilometres, which is larger compared to traffic analysis zones (TAZs) with an average of 3.4 square kilometres. The SA2 is selected for this study because it is the smallest statistical area providing commuting mode shares by both origin and destination. Among a total of 309 SA2s in Greater Melbourne, four SA2s (e.g. airports) have no population in the Australian Bureau of Statistics (ABS) 2016 census and thus excluded from the analysis, considering that population is an important exposure variable.

Independent variables were selected based on the review of literature in the previous section. Sociodemographic variables (i.e. population and the proportion of young people aged 0–14) and land use data (i.e. entropy measure of land use mix and proportion of land use area) were derived from the 2016 ABS census. Journey to work mode shares (i.e. train, tram, bus, walk, bicycle, motorcycle, and car) by origin (i.e. SA2 as the origin) and by destination (SA2 as the destination) were also obtained from the 2016 ABS census. Transport network data (i.e. the number of signalised intersections, number of public transport stops/stations, proportion of roads with above 100 km/h speed limits) were obtained from Data.Vic. Multicollinearity among independent variables were checked through the variance inflation factors (VIF), performed using *car* package (Fox and Weisberg, 2011) in R (R Core Team, 2018).

The entropy measure of land use mix considers the relative proportion of two or more land use categories within an area. In this study, five land use categories (residential, commercial, industrial, parkland, and civic) were employed to calculate the entropy measure; other land use types, e.g. agriculture, were excluded.

$$LUM_{ENT} = - \frac{\sum_{c=1}^{n_c} p_c^L \ln(p_c^L)}{\ln(n_c)} \quad (1)$$

where: LUM_{ENT} is the land use mix entropy measure, p_c^L is the proportion of land use category c in an area, n_c is the number of land use categories in an area. The entropy measure varies from 0 to 1, in which a higher value indicates a higher level of land use mix. A value of 0 indicates a single land use for an area while a value of 1 indicates an ideal balance of land use for an area.

Data were aggregated for each SA2 using ArcGIS 10.6. Table 1 presents a summary of dependent variables and final independent variables (journey to work mode shares and variables that showed strong effects). Density variables (i.e. population density, signalised intersection density, and station/stop density) were considered, but excluded due to high correlations among density variables and between density variables and journey to work by tram based on the VIF analysis. Greater Melbourne's SA2s had an average population of approximately 14,700 persons. Over the 5-year study period (2014–2018), the average numbers of total crashes and severe crashes for each SA2 were 155.5 and 43.1 respectively. Commuting travel either by origin or destination was dominated by car, followed by public transport, walking, cycling, and motorcycling.

Fig. 1 shows the distributions of total crashes and severe crashes by SA2 in Greater Melbourne. A clustering pattern can be seen for both total crashes or severe crashes. For example, zones with lower crashes tend to be clustered outside Melbourne's inner areas. The spatial dependency among crashes in adjacent zones is often tested using Moran's I (Moran, 1950; Huang et al., 2010; Truong and Somenahalli, 2011).

Table 1
Descriptive statistics.

Variable	Mean	Std.	Min	Max
Number of total crashes	155.518	112.694	15.000	981.000
Number of severe crashes	43.082	32.361	3.000	334.000
Population (persons)	14702.810	6266.235	37.000	37344.000
Proportion of people aged 0-14	0.183	0.048	0.032	0.486
Number of signalised intersections	10.393	11.193	0.000	91.000
Number of public transport stops/stations	66.662	35.940	1.000	258.000
Proportion of roads with a speed limit > 100 km/h	0.001	0.006	0.000	0.056
Proportion of industrial area	0.063	0.122	0.000	0.664
Land use mix - entropy measure	0.527	0.167	0.000	0.999
Proportion of commuting by train (origin)	0.127	0.071	0.000	0.355
Proportion of commuting by tram (origin)	0.026	0.061	0.000	0.325
Proportion of commuting by bus (origin)	0.018	0.019	0.000	0.141
Proportion of commuting by cycling (origin)	0.015	0.026	0.000	0.165
Proportion of commuting by walking (origin)	0.034	0.058	0.002	0.413
Proportion of commuting by motorbike (origin)	0.004	0.002	0.000	0.012
Proportion of commuting by car (origin)	0.760	0.151	0.108	0.933
Proportion of commuting by train (destination)	0.046	0.068	0.002	0.549
Proportion of commuting by tram (destination)	0.009	0.021	0.000	0.123
Proportion of commuting by bus (destination)	0.015	0.010	0.000	0.069
Proportion of commuting by cycling (destination)	0.010	0.013	0.000	0.083
Proportion of commuting by walking (destination)	0.033	0.022	0.000	0.138
Proportion of commuting by motorbike (destination)	0.003	0.002	0.000	0.013
Proportion of commuting by car (destination)	0.865	0.108	0.205	0.959

n = 305.

Indeed, the Moran’s I test, conducted using *spdep* package (Bivand et al., 2013) in R, was significant at $p < 0.001$ for both total crashes and severe crashes (Moran I statistic standard deviate = 5.33 and 4.28 respectively).

2.2. Modelling approaches

2.2.1. Random effect negative binomial (RENB) model

The negative binomial (NB) model has been widely adopted to account for over-dispersion in modelling crash frequency (Lord and Mannering, 2010). With the inclusion of random components, it can further account for heterogeneity in crash data, particularly spatial heterogeneity across zones in macroscopic crash analyses (Xu and Huang, 2015; Tasic and Porter, 2016; Truong et al., 2016). Let y_i denote the observed number of crashes in zone i , X_{ik} is the k^{th} variable for zone i , β_k is the coefficient, p is the number of variables. The NB model is derived by assuming:

$$y_i \sim \text{Poisson}(\lambda_i) \tag{2}$$

$$\ln(\lambda_i) = \beta_0 + \beta_1 \ln(\text{pop}_i) + \sum_{k=2}^p \beta_k X_{ik} + \varepsilon_i \tag{3}$$

where λ_i is the Poisson parameter, which is the expected number of fatalities in zone i , pop_i is the population of zone i (exposure variable), $\exp(\varepsilon_i)$ is a gamma-distributed error term with mean one and variance α . With the addition of ε_i , the variance can be larger than the mean as $\text{VAR}(y_i) = \lambda_i + \alpha \lambda_i^2$.

To account for heterogeneity across the zones, the random effect negative binomial (RENB) model can be presented as:

$$\ln(\lambda_i) = (\beta_0 + \omega_i) + \beta_1 \ln(\text{pop}_i) + \sum_{k=2}^p \beta_k X_{ik} + \varepsilon_i \tag{4}$$

where ω_i is a randomly distributed term, e.g. a normally distributed term with mean zero and variance σ^2 (Washington et al., 2011; Greene, 2012). In general, a random effect can be justified if its standard deviation is significantly larger than zero. In this paper, NB and RENB models were estimated using *NLOGIT 5* (Econometric Software, 2012). To assess the impact of independent variables on crash frequency, marginal effects were estimated by averaging individual marginal effects calculated for each observation.

2.2.2. Spatial model with conditional autoregressive priors (CAR)

To account for possible spatial dependency among crashes in adjacent zones, the spatial model with conditional autoregressive (CAR) priors can be employed (Aguero-Valverde and Jovanis, 2006; Quddus, 2008; El-Basyouny and Sayed, 2009; Huang et al., 2010; Truong et al., 2016). In this paper, the model proposed by Leroux et al. (2000) was utilised to account for spatial dependency:

$$y_i \sim \text{Poisson}(\lambda_i) \tag{5}$$

$$\ln(\lambda_i) = \beta_0 + \beta_1 \ln(\text{pop}_i) + \sum_{k=2}^p \beta_k X_{ik} + \phi_i \tag{6}$$

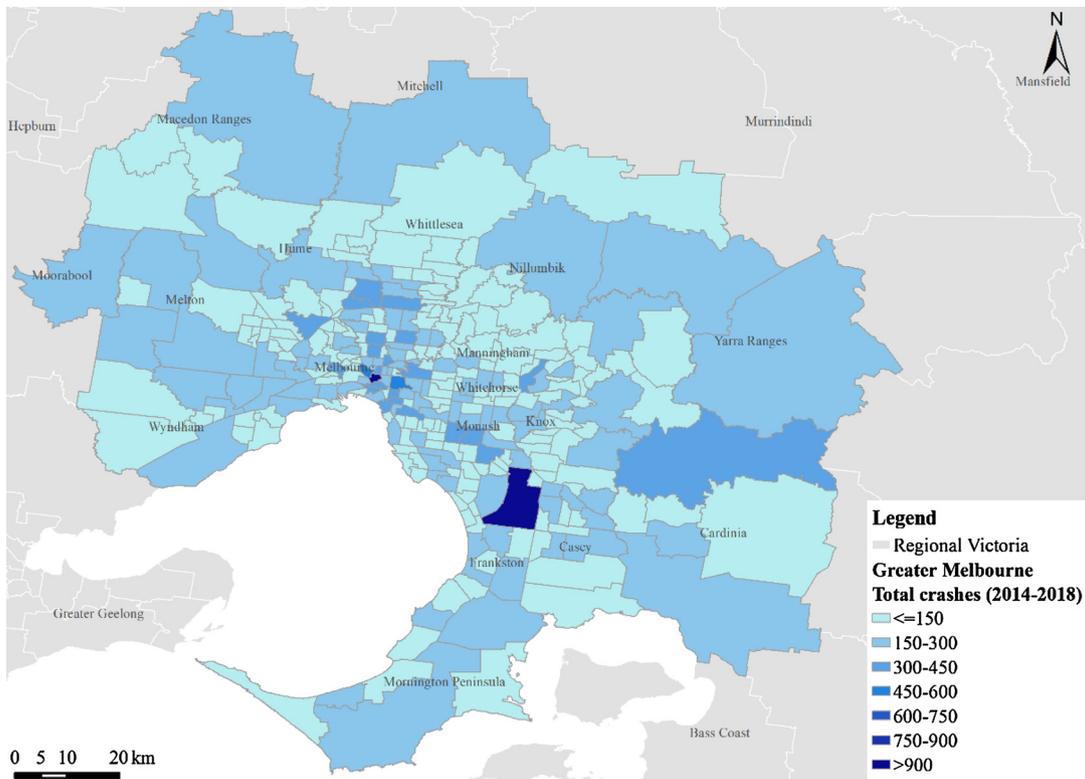
$$\phi_i | \phi_{-i}, W, \tau^2, \rho \sim N \left(\frac{\rho \sum_{j=1}^n w_{ij} \phi_j}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho} \right) \tag{7}$$

where: ϕ_i are random effects that account for residual spatial dependency in the data after the effects of covariates have been removed, ϕ_{-i} is the vector of random effects except for ϕ_i , $W = \{w_{ij}\}$ is the $n \times n$ adjacent matrix ($w_{ij} = 1$ if zones i and j are adjacent; 0 otherwise), τ^2 is the parameter controlling the variance of random effects, which is assumed as a prior inverse-gamma (1, 0.01), and ρ is the spatial parameter, $\rho \sim U(0,1)$. While $\rho = 1$ corresponds to the intrinsic CAR prior (Besag et al., 1991) for strong spatial correlation, $\rho = 0$ corresponds to independence (random effects with constant mean and variance).

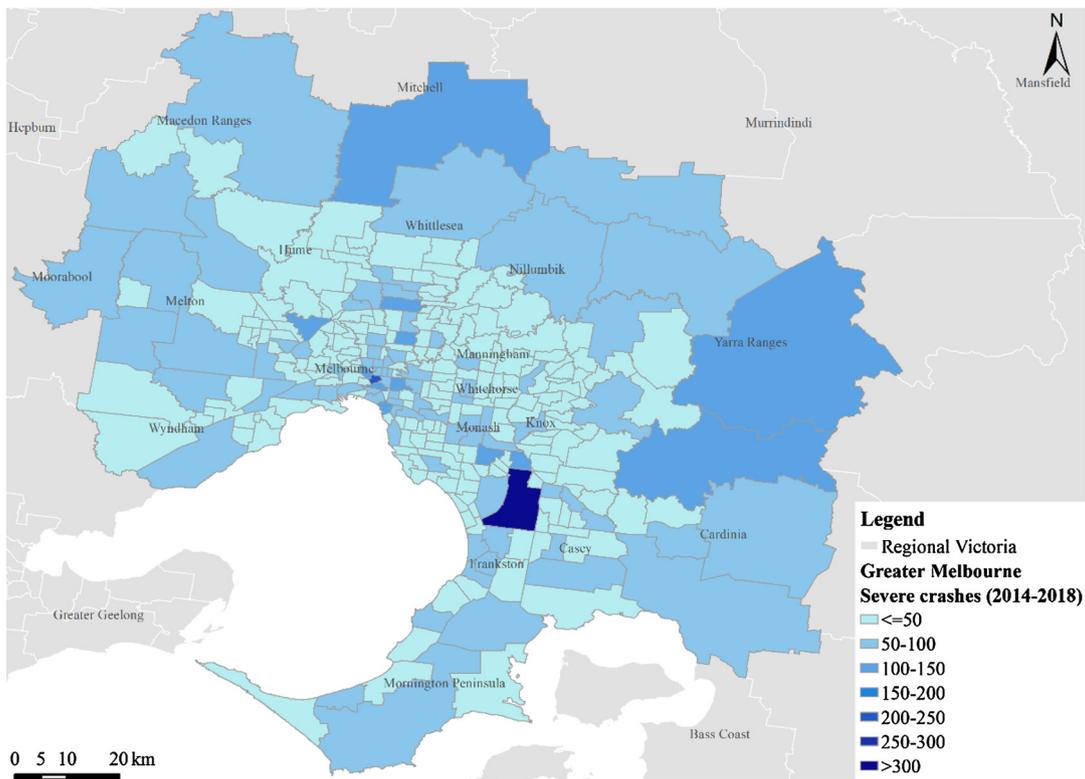
Parameters of the CAR model were estimated using Markov Chain Monte Carlo (MCMC) simulation. A burn-in period was set as 100,000 iterations. Parameter estimates were based on 500,000 samples with a thinning level of 20. Convergence of the CAR model was investigated by visual diagnostic and Geweke convergence diagnostic. Model fit was assessed using Deviance Information Criteria (DIC) (Spiegelhalter et al., 2002). Modelling was performed using the *CARBAYes* package (Lee, 2013) in R.

2.2.3. Measures of model prediction performance

Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used to measure prediction performance. Models with lower RMSE and MAPE indicates a better prediction performance.



a) Distribution of total crashes



b) Distribution of severe crashes

Fig. 1. Distribution of crashes in Greater Melbourne.

Table 2
Results of negative binomial (NB) models - TOTAL CRASHES.

Variable	Commuting origin		Commuting destination	
	Estimate	Std. Error	Estimate	Std. Error
Log of population	0.278	***	0.162	***
Proportion of people aged 0-14	-2.438	***	-1.939	***
Number of signalised intersections	0.016	***	0.020	***
Number of public transport stops/stations	0.005	***	0.005	***
Proportion of roads with a speed limit > 100 km/h	10.697	***	11.527	***
Proportion of industrial area	0.997	***	0.684	**
Land use mix - entropy measure	-0.274	*	-0.108	
Proportion of commuting by train (origin)	-1.559	***		
Proportion of commuting by tram (origin)	-2.135	*		
Proportion of commuting by bus (origin)	-3.756	*		
Proportion of commuting by cycling (origin)	3.452			
Proportion of commuting by walking (origin)	2.829	***		
Proportion of commuting by motorbike (origin)	34.396	*		
Proportion of commuting by train (destination)			-1.193	0.793
Proportion of commuting by tram (destination)			1.951	3.253
Proportion of commuting by bus (destination)			-8.360	**
Proportion of commuting by cycling (destination)			2.107	3.899
Proportion of commuting by walking (destination)			2.193	1.359
Proportion of commuting by motorbike (destination)			58.944	***
Intercept	2.368	***	3.087	***
Dispersion parameter	0.135	***	0.134	***
Log likelihood	-1621.643		-1620.901	
Log likelihood (intercept only)	-1784.972		-1784.972	
AIC	3273.3		3271.8	
MAPE	36.176%		36.373%	
RMSE	85.121		113.763	

n = 305; * p < 0.05; ** p < 0.01; *** p < 0.001; all models were significant at p < 0.001.

$$RMSE = \sqrt{\frac{1}{n_o} \sum_{j=1}^{n_o} (y_j^O - y_j^P)^2} \tag{8}$$

$$MAPE = \frac{1}{n_o} \sum_{j=1}^{n_o} \left| \frac{y_j^O - y_j^P}{y_j^O} \right| \tag{9}$$

where: y_j^O is the observed number of crashes, y_j^P is the predicted number of crashes, and n_o is the number of observations.

3. Results

3.1. Random effect negative binomial (RENB) models

In this section, results of the baseline NB models are presented, followed by results of the RENB models. Commuting mode shares by origin and by destination were considered in separate models due to multicollinearity. Specifically, including both commuting mode shares by origin and destination in the same model led to excessive VIFs (e.g. above 11). As a common rule of thumb, a VIF value of greater than 10 indicates severe multicollinearity (O'Brien, 2007). Table 2 summarises results of two NB models for total crashes, considering commuting mode shares by origin and by destination respectively. Both models and associated dispersion parameters were significant at p < 0.001. It is noted that the proportion of commuting by car was excluded due to redundancy, which can be used as a reference point when evaluating the effects of commuting more shares. VIFs were well below 5 for all variables, except the proportion of commuting by tram. For the commuting origin model, the VIF for the commuting by tram variable was 5.16. However, there was no evidence of multicollinearity as the signs and significance levels of coefficients were consistent while the model fit was better when compared to the model without this variable (Chi-square = 6.98, df = 1, p < 0.01). Likewise, no evidence of multicollinearity was found with regards to the commuting destination model, although the VIF for the commuting by tram variable was 5.4.

Table 3 presents results of NB models for severe crashes with

commuting mode shares by origin and by destination. Like total crash models, the proportion of commuting by tram was included since there was no evidence of multicollinearity. Both models and dispersion parameters were significant at p < 0.001.

Table 4 summarises results of two RENB models for total crashes with commuting mode shares by origin and by destination respectively. Both models were significant at p < 0.001 with significant random effects and dispersion parameters. Likelihood ratio tests showed that RENB total crash models were not significantly better than NB total crash models (Chi-square = 3.394, df = 1, p = 0.07 and Chi-square = 1.246, df = 1, p = 0.26 for the commuting origin and destination models respectively). However, RENB models had substantially lower MAPE and RMSE when compared to NB models. Furthermore, RENB models had more significant variables. Overall, RENB models performed better than NB models.

All variables of the commuting origin model were significant. As indicated by positive coefficients, total crashes were likely to increase with higher population, more signalised intersections or public transport stops/stations, and higher proportions of industrial areas or roads with above 100 km/h speed limits. In contrast, fewer total crashes were associated with a higher proportion of people aged 0–14 or a higher level of land use mix. Furthermore, total crashes in a zone tended to decrease with higher proportions of commuting from this zone by all public transport modes (train, tram, and bus). However, total crashes in a zone tended to increase with higher proportion of commuting from this zone by active transport (walking and cycling) or motorbike.

Land use mix and the proportions of commuting by train, tram, walking, and cycling were not significant in the commuting destination model. Otherwise, results were consistent with the commuting origin model. Fewer total crashes in a zone were associated with a higher proportion of commuting to this zone by bus. However, a greater number of total crashes in a zone were associated with a higher proportion of commuting to this zone by motorbike.

Table 5 presents results of RENB models for severe crashes, which considered SA2 as the commuting origin and destination respectively. Both models were significant at p < 0.001 with significant random

Table 3
Results of negative binomial (NB) models- SEVERE CRASHES.

Variable	Commuting origin			Commuting destination		
	Estimate		Std. Error	Estimate		Std. Error
Log of population	0.343	***	0.052	0.175	***	0.036
Proportion of people aged 0-14	- 3.349	***	0.617	- 2.520	***	0.578
Number of signalised intersections	0.015	***	0.004	0.021	***	0.004
Number of public transport stops/stations	0.003	***	0.001	0.004	***	0.001
Proportion of roads with a speed limit > 100 km/h	14.101	***	3.759	15.327	***	3.383
Proportion of industrial area	1.077	***	0.281	0.590		0.318
Land use mix - entropy measure	- 0.494	**	0.151	- 0.169		0.173
Proportion of commuting by train (origin)	- 2.573	***	0.432			
Proportion of commuting by tram (origin)	- 3.954	***	0.956			
Proportion of commuting by bus (origin)	- 5.495	**	2.113			
Proportion of commuting by cycling (origin)	3.329		2.503			
Proportion of commuting by walking (origin)	3.602	***	0.720			
Proportion of commuting by motorbike (origin)	38.020	*	17.656			
Proportion of commuting by train (destination)				- 1.889	*	0.958
Proportion of commuting by tram (destination)				1.392		3.701
Proportion of commuting by bus (destination)				- 13.046	***	2.948
Proportion of commuting by cycling (destination)				- 3.782		3.997
Proportion of commuting by walking (destination)				4.668	***	1.352
Proportion of commuting by motorbike (destination)				77.579	***	12.806
Intercept	1.033		0.532	1.932	***	0.396
Dispersion parameter	0.165	***	0.018	0.162	***	0.016
Log likelihood	- 1276.521			- 1274.856		
Log likelihood (intercept only)	- 1400.975			- 1400.975		
AIC	2583.0			2579.7		
MAPE	45.532%			45.524%		
RMSE	24.394			26.168		

n = 305; * p < 0.05; ** p < 0.01; *** p < 0.001; all models were significant at p < 0.001.

effects and dispersion parameters. All variables were significant in the commuting origin model and their effects on severe crashes were in alignment with those on total crashes. Likelihood ratio tests showed that RENB severe crash models were significantly better than NB severe crash models (Chi-square = 9.636, df = 1, p < 0.01 and Chi-

square = 7.156, df = 1, p < 0.01 for the models for commuting origin and destination respectively). Compared to NB models, RENB models also had much better MAPE and RMSE, and more significant variables.

Regarding the commuting destination model, the proportions of commuting by tram and cycling were not significant. Effects of

Table 4
Results of random effect negative binomial (RENB) models - TOTAL CRASHES.

Variable	Commuting origin			Commuting destination		
	Estimate		Std. Error	Estimate		Std. Error
Log of population	0.283	***	0.030	0.173	***	0.028
Proportion of people aged 0-14	- 2.440	***	0.394	- 1.956	***	0.430
Number of signalised intersections	0.015	***	0.002	0.020	***	0.002
Number of public transport stops/stations	0.005	***	0.001	0.005	***	0.001
Proportion of roads with a speed limit > 100 km/h	10.262	***	2.313	10.885	***	2.607
Proportion of industrial area	1.026	***	0.155	0.723	***	0.213
Land use mix - entropy measure	- 0.255	*	0.102	- 0.122		0.136
Proportion of commuting by train (origin)	- 1.364	***	0.255			
Proportion of commuting by tram (origin)	- 1.986	**	0.610			
Proportion of commuting by bus (origin)	- 3.315	***	0.947			
Proportion of commuting by cycling (origin)	3.816	***	1.153			
Proportion of commuting by walking (origin)	2.765	***	0.467			
Proportion of commuting by motorbike (origin)	30.724	**	10.095			
Proportion of commuting by train (destination)				- 1.048		0.614
Proportion of commuting by tram (destination)				1.871		2.490
Proportion of commuting by bus (destination)				- 7.707	***	2.270
Proportion of commuting by cycling (destination)				2.663		3.021
Proportion of commuting by walking (destination)				2.029		1.220
Proportion of commuting by motorbike (destination)				55.188	***	9.073
Intercept	2.236	***	0.285	2.967	***	0.296
Standard deviation of parameter distribution	0.272	***	0.016	0.193	***	0.019
Dispersion parameter	15.970	***	1.467	10.252	***	0.882
Log likelihood	- 1619.946			- 1620.278		
Log likelihood (intercept only)	- 1784.972			- 1784.972		
AIC	3271.9			3272.6		
MAPE	16.753%			26.106%		
RMSE	38.245			79.086		

n = 305; * p < 0.05; ** p < 0.01; *** p < 0.001; all models were significant at p < 0.001.

Table 5
Results of random effect negative binomial (RENB) models- SEVERE CRASHES.

Variable	Commuting origin		Commuting destination	
	Estimate	Std. Error	Estimate	Std. Error
Log of population	0.352	***	0.207	***
Proportion of people aged 0-14	-3.269	***	-2.631	***
Number of signalised intersections	0.014	***	0.020	***
Number of public transport stops/stations	0.003	***	0.004	***
Proportion of roads with a speed limit > 100 km/h	14.006	***	14.690	***
Proportion of industrial area	1.086	***	0.621	***
Land use mix - entropy measure	-0.408	***	-0.160	*
Proportion of commuting by train (origin)	-2.220	***		
Proportion of commuting by tram (origin)	-3.857	***		
Proportion of commuting by bus (origin)	-4.687	***		
Proportion of commuting by cycling (origin)	3.919	***		
Proportion of commuting by walking (origin)	3.450	***		
Proportion of commuting by motorbike (origin)	31.609	***		
Proportion of commuting by train (destination)			-1.495	***
Proportion of commuting by tram (destination)			1.436	
Proportion of commuting by bus (destination)			-12.109	***
Proportion of commuting by cycling (destination)			-2.670	
Proportion of commuting by walking (destination)			3.274	***
Proportion of commuting by motorbike (destination)			68.247	***
Intercept	0.756	**	1.599	***
Standard deviation of parameter distribution	0.393	***	0.391	***
Dispersion parameter	107.023	***	118.964	***
Log likelihood	-1271.703		-1271.278	
Log likelihood (intercept only)	-1400.975		-1400.975	
AIC	2575.4		2574.6	
MAPE	8.987%		9.059%	
RMSE	3.441		3.350	

n = 305; * p < 0.05; ** p < 0.01; *** p < 0.001; all models were significant at p < 0.001.

sociodemographic and transport network variables, commuting by bus and by motorbike were consistent compared to the commuting destination model for total crashes. In addition, results clearly showed that severe crashes in a zone tended to decrease with a higher proportion of commuting to this zone by train, but increase with a higher proportion of commuting to this zone by walking.

Overall, there was some evidence about the effect of land use mix measured by entropy in RENB models, which suggested that a higher level of land use mix was associated with fewer crashes. Land use mix measured by entropy has been found to be negatively associated with pedestrian crashes (Wang and Kockelman, 2013), but positively associated with bicycle crashes (Barajas, 2018). A higher level of land use mix would lead to more active travel, but shorter travel distances for all modes. Reduced crash exposure from the later would overcome potential negative safety effects from the former, given low active mode shares in Melbourne. This would be the reason for safety benefits from a higher level of land use mix.

The Moran's I test was conducted for the RENB models' residuals, which were significant at p < 0.01. This suggested that the CAR models, accounting for spatial dependency, would perform better.

3.2. Conditional autoregressive (CAR) models

Table 6 presents results of CAR models for total crashes, considering commuting mode share by origin and by destination respectively. Compared to RENB models, CAR models were significantly better with much lower MAPE and RMSE, suggesting that spatial dependence in crash data was better accounted for by CAR model. In both CAR models, 95% Bayesian credible intervals of parameters for population, people aged 0–14, signalised intersections, public transport stops/stations, roads with above 100 km/h speed limits, and industrial land use area had the same sign, indicating credible effects of these variables on total crashes. These results were in alignment with those from RENB models. With regards to the commuting origin model, credible effects of commuting by train, tram, bus, walking, and motorbike were evident. With

regards to the commuting destination model, commuting by bus and motorbike also had credible effects on total crashes. Overall, credible effects of these variables were consistent compared to RENB models for total crashes.

Table 7 shows results of CAR models for severe crashes. Both CAR models for commuting origin and commuting destination had lower MAPE and RMSE when compared to RENB models. In the commuting origin model, credible effects were found for most variables, except for land use mix and commuting by cycling. In the commuting destination model, credible effects were not found for industrial land use area, land use mix, commuting by tram, cycling, and walking. Directions of these effects were similar to those of the CAR model for total crashes.

4. Discussion and conclusions

The following discussions were based on marginal effects estimated mainly from CAR models given their better performance. The changes in crashes are per 5 years since the dependent variables were based on crashes during the 5-year period (2014–2018).

Results suggest that mode shift from private vehicle travel to public transport travel would result in substantial reductions in the number of total crashes as well as severe crashes. Specifically, a percentage point increase in the proportion of commuting from a zone by train would reduce the number of total crashes by 2.2 and severe crashes by 0.86 in this zone, holding all other variables (including proportions of commuting by tram, bus, walking, cycling, and motorbike) constant. The one-percentage-point increase in the proportion of commuting by train therefore implies a one-percentage-point reduction in the proportion of commuting by car, as other modes (i.e. ferry) are negligible. Similarly, increases in the proportions of commuting by tram, bus, cycling, walking, and motorbike in the following discussions imply mode shift from car travel.

In addition, a percentage point increase in the proportion of commuting from a zone by tram would reduce total crashes by 3.1 and severe crashes by 1.5. Commuting from a zone by bus has a larger effect

Table 6
Results of conditional autoregressive (CAR) models – TOTAL CRASHES.

Variable	Commuting origin			Commuting destination		
	Mean	Credible interval		Mean	Credible interval	
		2.5%	97.5%		2.5%	97.5%
Log of population	0.305	0.205	0.407	0.184	0.094	0.271
Proportion of people aged 0-14	-2.930	-4.477	-1.399	-2.140	-3.651	-0.734
Number of signalised intersections	0.017	0.011	0.023	0.020	0.013	0.027
Number of public transport stops/stations	0.005	0.003	0.007	0.006	0.004	0.008
Proportion of roads with a speed limit > 100 km/h	10.661	1.941	19.487	10.914	2.188	19.488
Proportion of industrial area	0.964	0.491	1.409	0.687	0.151	1.193
Land use mix - entropy measure	-0.282	-0.621	0.069	-0.158	-0.516	0.187
Proportion of commuting by train (origin)	-1.409	-2.401	-0.420			
Proportion of commuting by tram (origin)	-1.990	-3.931	-0.113			
Proportion of commuting by bus (origin)	-3.669	-6.916	-0.417			
Proportion of commuting by cycling (origin)	1.448	-2.316	5.420			
Proportion of commuting by walking (origin)	4.208	2.718	5.847			
Proportion of commuting by motorbike (origin)	71.850	38.640	109.056			
Proportion of commuting by train (destination)				-0.477	-1.867	1.051
Proportion of commuting by tram (destination)				2.329	-3.329	7.582
Proportion of commuting by bus (destination)				-6.343	-12.344	-0.056
Proportion of commuting by cycling (destination)				1.162	-6.386	8.509
Proportion of commuting by walking (destination)				1.232	-2.368	4.639
Proportion of commuting by motorbike (destination)				69.455	39.345	102.218
Intercept	1.908	0.956	2.837	2.791	1.946	3.637
τ^2	0.340	0.200	0.589	0.237	0.155	0.388
ρ	0.295	0.091	0.660	0.132	0.016	0.340
DIC	2656.6			2664.7		
MAPE	2.128%			2.187%		
RMSE	2.537			2.413		

n = 305; italics are used for parameters that were not significantly different to zero, based on the 95% credible intervals.

as a percentage point increase in the proportion of commuting from a zone by bus would reduce total crashes by 5.7 and severe crashes by 1.8. It is noted that incidence rate ratio based on the exponent of a variable coefficient can also be used for interpretation. For example, if

the proportion of commuting from a zone by bus were increased by one percentage point, the rate for severe crashes would be expected to decrease by a factor of 0.958, i.e. $exp(-4.259 \times 0.01)$. In other words, a percentage point increase in the proportion of commuting from a zone

Table 7
Results of conditional autoregressive (CAR) models – SEVERE CRASHES.

Variable	Commuting origin			Commuting destination		
	Mean	Credible interval		Mean	Credible interval	
		2.5%	97.5%		2.5%	97.5%
Log of population	0.312	0.196	0.426	0.200	0.099	0.303
Proportion of people aged 0-14	-3.757	-5.402	-2.047	-2.851	-4.675	-1.117
Number of signalised intersections	0.016	0.009	0.023	0.020	0.012	0.028
Number of public transport stops/stations	0.004	0.002	0.006	0.005	0.003	0.007
Proportion of roads with a speed limit > 100 km/h	11.191	0.624	21.554	17.700	7.228	28.556
Proportion of industrial area	1.096	0.567	1.611	0.429	-0.181	1.029
Land use mix - entropy measure	-0.305	-0.727	0.085	-0.168	-0.591	0.251
Proportion of commuting by train (origin)	-2.002	-3.234	-0.658			
Proportion of commuting by tram (origin)	-3.479	-5.792	-1.039			
Proportion of commuting by bus (origin)	-4.259	-7.998	-0.379			
Proportion of commuting by cycling (origin)	2.313	-1.666	6.343			
Proportion of commuting by walking (origin)	3.593	1.694	5.472			
Proportion of commuting by motorbike (origin)	67.310	28.577	107.500			
Proportion of commuting by train (destination)				-1.739	-3.478	-0.142
Proportion of commuting by tram (destination)				2.408	-4.300	8.939
Proportion of commuting by bus (destination)				-11.240	-18.946	-4.122
Proportion of commuting by cycling (destination)				-9.053	-17.920	0.453
Proportion of commuting by walking (destination)				2.922	-1.189	7.064
Proportion of commuting by motorbike (destination)				150.696	109.343	196.325
Intercept	0.923	-0.155	2.029	1.495	0.483	2.506
τ^2	0.460	0.278	0.708	0.382	0.237	0.599
ρ	0.442	0.167	0.818	0.234	0.055	0.520
DIC	2227.9			2238.1		
MAPE	7.152%			6.530%		
RMSE	2.328			2.110		

n = 305; italics are used for parameters that were not significantly different to zero, based on the 95% credible intervals.

by bus would reduce severe crashes by 4.2%. In terms of commuting by destination, a percentage point increase in the proportion of commuting to a zone by train would reduce severe crashes by 0.75. A percentage point increase in the proportion of commuting to a zone by bus would reduce total crashes by 9.9 and severe crashes by 4.8.

Interestingly, commuting by bus had the largest safety benefits, in terms of reducing both total crashes and severe crashes. Most public transport trips include active transport (e.g. for getting to and from stops/stations) that has a relatively high casualty rate compared to public transport (Beck et al., 2007). It is widely accepted that bus and tram passengers are willing to cover a shorter distance using active transport, particularly walking, compared to train passengers. The risk associated with active transport as part of public transport trips would therefore be lower for bus and tram passengers than for train passengers. Since about two thirds of Melbourne's 250 km tram network share the road with general traffic, the majority of trams operate as streetcars (Currie and Shalaby, 2007). The road safety issues of streetcars, e.g. heavy vehicles operating in a complex environment with multiple other road users (Naznin et al., 2018), would be attributed to lower benefits of commuting by tram compared to by bus. In addition to active transport, private transport (e.g. park and ride, and kiss and ride) can also be part of many train commuting trips. The greater risks associated with access to and from train stations would be a main reason for the lowest benefits of commuting by train.

Results also indicate safety issues around active transport. A percentage point increase in the proportion of commuting from a zone by walking was associated with an increase of 6.5 in total crashes and an increase of 1.5 in severe crashes. This is consistent with previous research, which found commuting by walking was associated with increases in vehicular and non-motorised crashes in Chicago (Tasic and Porter, 2016) and in total and severe crashes in Florida (Abdel-Aty et al., 2013). While commuting by cycling was not significant in CAR models, it was significant in RENB models. Based on RENB models for commuting by origin, a percentage point increase in the proportion of commuting from a zone by cycling would lead to 5 total crashes and 1.4 severe crashes additionally. The negative safety impact of commuting by cycling was also found in Melbourne by Amoh-Gyimah et al. (2017). Nevertheless, results of this study contradict findings from European cities (Moeinaddini et al., 2015), where commuting by walking or cycling was associated with reductions in road deaths. This difference would be explained by different traffic environments in US and Australian cities (i.e. car dependent cities with a less friendly environment for pedestrians and cyclists) compared to European cities.

In addition, a percentage point increase in the proportion of commuting from a zone by motorbike was associated with increases of 111.7 in total crashes and 29 in severe crashes. In terms of commuting by destination, a percentage point increase in the proportion of commuting to a zone by motorbike tends to increase total crashes by 108 and severe crashes by 64.9. It is noted that the share of commuting by motorbike was quite low, with an average of 0.3-0.4% in Melbourne. The effect of commuting by motorbike on increasing road deaths was also evident in European cities (Moeinaddini et al., 2015). While the overrepresentation of motorcycle deaths has been reported in Melbourne, this study further highlights the issue from a macroscopic safety perspective.

Results also shows significant effects of sociodemographic, transport network, and land use variables, which are in alignment with previous research. Considering both commuting by origin and destination models, a percentage point increase in the proportion of people aged 0-14 would reduce total crash by 3.3-4.6 and severe crashes by 1.2-1.6 in a zone. Guevara et al. (2004) also suggested relationship between fewer crashes and a higher proportion of young people. Zones with a higher proportion of people aged 0-14 would generate less commuting traffic, leading to lower crash exposure. An additional signalised intersection would lead to increases of 2.6-3.1 in total crashes and 0.7-0.9 in severe crashes, while an additional public transport stops/stations

would lead to increases of 0.8-0.9 in total crashes and 0.2 in severe crashes. More signalised intersections indicate a more complex traffic environment while more public transport stops/stations additionally indicate a higher level of pedestrian exposure. Positive relationships between crashes and intersections and public transport stops/stations have been reported in previous research (Hadayeghi et al., 2007; Tasic and Porter, 2016).

Furthermore, a percentage point increase in the proportion of roads with above 100 km/h speed limits would increase the number of total crashes by 16.6-17 and the number of severe crashes by 4.8-7.6. This finding is in alignment with previous research, which suggested a significant negative safety effect of roads with high speed limits (Abdel-Aty et al., 2013). Regarding land use factors, a percentage point increase in the proportion of industrial area was associated with increases of 1.1-1.5 in total crashes and 0.47 in severe crashes. Earlier studies also indicated that zones with a higher proportion of industrial areas tend to have more crashes (Hadayeghi et al., 2007; Pulugurtha et al., 2013), which would be attributed by higher activities of heavy vehicles.

Overall, this study suggests much potential for public transport as a road safety solution through a macroscopic analysis of road safety in Melbourne, Australia. It is evident that mode shift from private vehicle to public transport, including train, tram, and bus, for commuting would reduce not only total crashes, but also severe crashes. A practical implication of the findings is that public transport should be actively supported as a road safety solution, in addition to its widely accepted role as a congestion relief solution. Thus, it should be integrated into road safety strategies and action plans.

A limitation of this study is that the macroscopic analysis was conducted for only one spatial unit, i.e. SA2, considering that SA2 is the smallest statistical area providing commuting mode shares by both origin and destination. Nevertheless, while many findings of this study corroborate previous research, new findings regarding macroscopic safety effects of public transport modes are also presented. The macroscopic safety analysis based on SA2s can be incorporated with strategic transport planning by predicting crashes for SA2s using aggregated data from TAZs, given that SA2s are larger than TAZs. Another limitation is that safety effects of public transport was considered through commuting mode shares, which was based on the census day only. This would be addressed in future work by utilising other data sources, such as household travel surveys. As this study considered the aggregated number of public transport stops/stations, the effects of stops/stations by individual public transport mode should be explored in future work. Future research should also explore possible spatial variations of impacts of public transport modes on macroscopic safety. In addition, multivariate modelling approaches should be considered in future work for jointly modelling the frequency of crashes by severity (Park and Lord, 2007; Ma et al., 2008; Aguerro-Valverde, 2013; Lee et al., 2015b) or the proportion of crash types (Lee et al., 2018).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research is partly supported by La Trobe University's Building Healthy Communities RFA Grant.

References

- Abdel-Aty, M., Lee, J., Siddiqui, C., Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. *Transp. Res. Part A Policy Pract.* 49, 62-75.
- Abdel-Aty, M., Siddiqui, C., Huang, H., Wang, X., 2011. Integrating trip and roadway

- characteristics to manage safety in traffic analysis zones. *Transportation Research Record: Journal of the Transportation Research Board*. 2213, 20–28.
- Agüero-Valverde, J., 2013. Multivariate spatial models of excess crash frequency at area level: case of costa rica. *Accid. Anal. Prev.* 59, 365–373.
- Agüero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in pennsylvania. *Accid. Anal. Prev.* 38 (3), 618–625.
- Amoh-Gyimah, R., Saberi, M., Sarvi, M., 2016. Macroscopic modeling of pedestrian and bicycle crashes: a cross-comparison of estimation methods. *Accid. Anal. Prev.* 93, 147–159.
- Amoh-Gyimah, R., Saberi, M., Sarvi, M., 2017. The effect of variations in spatial units on unobserved heterogeneity in macroscopic crash models. *Anal. Methods Accid. Res.* 13, 28–51.
- Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some french counties. *Accid. Anal. Prev.* 35 (4), 537–547.
- ATC, 2011. National road safety strategy 2011–2020. Australian Transport Council.
- Barajas, J.M., 2018. Not all crashes are created equal: associations between the built environment and disparities in bicycle collisions. *J. Transp. Land Use* 11 (1), 865–882.
- Beck, L.F., Dellinger, A.M., O'Neil, M.E., 2007. Motor vehicle crash injury rates by mode of travel, united states: using exposure-based methods to quantify differences. *Am. J. Epidemiol.* 166 (2), 212–218.
- Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* 43 (1), 1–20.
- BITRE, 2007. Estimating Urban Traffic and Congestion Cost Trends for Australian Cities. Working Paper 74.
- BITRE, 2019. Australian Road Deaths Database. Bureau of Infrastructure, Transport and Regional Economics.
- Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. *Applied Spatial Data Analysis With R*, second ed. Springer, NY, USA.
- Cai, Q., Abdel-Aty, M., Lee, J., Eluru, N., 2017. Comparative analysis of zonal systems for macro-level crash modeling. *J. Safety Res.* 61, 157–166.
- Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: incorporating spatial spillover effects in dual state count models. *Accid. Anal. Prev.* 93, 14–22.
- Cairney, P., 2010. The Road Safety Consequences of Changing Travel Modes. Austroads, Sydney, NSW, Australia.
- Coruh, E., Bilgic, A., Tortum, A., 2015. Accident analysis with aggregated data: the random parameters negative binomial panel count data model. *Anal. Methods Accid. Res.* 7, 37–49.
- Currie, G., Shalaby, A., 2007. Success and challenges in modernizing streetcar systems: experiences in melbourne, australia, and toronto, canada. *Transp. Res. Rec.* 2006 (1), 31–39.
- Dong, N., Huang, H., Lee, J., Gao, M., Abdel-Aty, M., 2016. Macroscopic hotspots identification: a bayesian spatio-temporal interaction approach. *Accid. Anal. Prev.* 92, 256–264.
- Dumbaugh, E., Rae, R., 2009. Safe urban form: revisiting the relationship between community design and traffic safety. *J. Am. Plan. Assoc.* 75 (3), 309–329.
- Econometric Software, 2012. *Nlogit. Version 5 Reference Guide*.
- El-Basyouny, K., Sayed, T., 2009. Urban arterial accident prediction models with spatial effects. *Transportation Research Record: Journal of the Transportation Research Board* 2102 (1), 27–33.
- Evans, A.W., 1994. Evaluating public transport and road safety measures. *Accid. Anal. Prev.* 26 (4), 411–428.
- Fox, J., Weisberg, S., 2011. *An {r} Companion to Applied Regression*. Sage, Thousand Oaks, CA.
- Greene, W.H., 2012. *Econometric Analysis*, 7th ed. Pearson Prentice Hall, Upper Saddle River, NJ.
- Guevara, F.L., Washington, S., Oh, J., 2004. Forecasting crashes at the planning level: simultaneous negative binomial crash model applied in tucson, arizona. *Transportation Research Record: Journal of the Transportation Research Board*. 1897, 191–199.
- Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record: Journal of the Transportation Research Board*. 1840, 87–95.
- Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety prediction models: proactive tool for safety evaluation in urban transportation planning applications. *Transportation Research Record: Journal of the Transportation Research Board*. 2019, 225–236.
- Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-level crash risk analysis in florida. *Transportation Research Record: Journal of the Transportation Research Board*. 2148, 27–37.
- Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., Abdel-Aty, M., 2016. Macro and micro models for zonal crash prediction with application in hot zones identification. *J. Transp. Geogr.* 54, 248–256.
- Jiang, X., Abdel-Aty, M., Hu, J., Lee, J., 2016. Investigating macro-level hotzone identification and variable importance using big data: a random forest models approach. *Neurocomputing*. 181, 53–63.
- Karlaftis, M.G., Tarko, A.P., 1998. Heterogeneity considerations in accident modeling. *Accid. Anal. Prev.* 30 (4), 425–433.
- Lee, D., 2013. Carbayes: an r package for bayesian spatial modeling with conditional autoregressive priors. *J. Stat. Softw.* 55 (13), 24.
- Lee, J., Abdel-Aty, M., 2018. Macro-level analysis of bicycle safety: focusing on the characteristics of both crash location and residence. *Int. J. Sustain. Transp.* 12 (8), 553–560.
- Lee, J., Abdel-Aty, M., Choi, K., 2014a. Analysis of residence characteristics of at-fault drivers in traffic crashes. *Saf. Sci.* 68, 6–13.
- Lee, J., Abdel-Aty, M., Choi, K., Huang, H., 2015a. Multi-level hot zone identification for pedestrian safety. *Accid. Anal. Prev.* 76, 64–73.
- Lee, J., Abdel-Aty, M., Jiang, X., 2014b. Development of zone system for macro-level traffic safety analysis. *J. Transp. Geogr.* 38, 13–21.
- Lee, J., Abdel-Aty, M., Jiang, X., 2015b. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accid. Anal. Prev.* 78, 146–154.
- Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M., Cai, Q., 2018. Analysis of crash proportion by vehicle type at traffic analysis zone level: a mixed fractional split multinomial logit modeling approach with spatial effects. *Accid. Anal. Prev.* 111, 12–22.
- Leroux, B.G., Lei, X., Breslow, N., 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence. *Statistical Models in Epidemiology, the Environment, and Clinical Trials*.
- Litman, T., 2016. *The Hidden Traffic Safety Solution: Public Transportation*. American Public Transportation Association (APTA).
- Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transp. Res. Part A Policy Pract.* 44 (5), 291–305.
- Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighbourhood traffic safety. *Can. J. Civ. Eng.* 33 (5), 609–621.
- Ma, J., Kockelman, K.M., Damien, P., 2008. A multivariate poisson-lognormal regression model for prediction of crash counts by severity, using bayesian methods. *Accid. Anal. Prev.* 40 (3), 964–975.
- Moieinaddini, M., Asadi-Shekari, Z., Sultan, Z., Zaly Shah, M., 2015. Analyzing the relationships between the number of deaths in road accidents and the work travel mode choice at the city level. *Saf. Sci.* 72, 249–254.
- Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. *Biometrika*. 37 (1/2), 17–23.
- Naznin, F., Currie, G., Logan, D., 2018. Exploring road design factors influencing tram road safety – melbourne tram driver focus groups. *Accid. Anal. Prev.* 110, 52–61.
- Ng, K., Hung, W., Wong, W., 2002. An algorithm for assessing the risk of traffic accident. *J. Safety Res.* 33 (3), 387–410.
- Noland, R.B., 2003. Traffic fatalities and injuries: the effect of changes in infrastructure and other trends. *Accid. Anal. Prev.* 35 (4), 599–611.
- Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in england. *Accid. Anal. Prev.* 36 (6), 973–984.
- O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* 41 (5), 673–690.
- Park, E.S., Lord, D., 2007. Multivariate poisson-lognormal models for jointly modeling crash frequency by severity. *Transp. Res. Rec.* 2019 (1), 1–6.
- Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., 2012. Application of different exposure measures in development of planning-level zonal crash prediction models. *Transportation Research Record: Journal of the Transportation Research Board*. 2280, 145–153.
- Pirdavani, A., Daniels, S., van Vlietden, K., Brijs, K., Kochan, B., 2017. Socioeconomic and sociodemographic inequalities and their association with road traffic injuries. *J. Transp. Health* 4, 152–161.
- Pulugurtha, S.S., Duddu, V.R., Kotagiri, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. *Accid. Anal. Prev.* 50, 678–687.
- Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of london crash data. *Accid. Anal. Prev.* 40 (4), 1486–1497.
- R Core Team, 2018. *R: A Language and Environment for Statistical Computing*. R foundation for statistical computing, Vienna, Austria.
- Savage, I., 2013. Comparing the fatality risks in united states transportation across modes and over time. *Res. Transp. Econ.* 43 (1), 9–22.
- Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accid. Anal. Prev.* 45, 382–391.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Series B Stat. Methodol.* 64 (4), 583–639.
- Tasic, J., Porter, R.H., 2016. Modeling spatial relationships between multimodal transportation infrastructure and traffic safety outcomes in urban environments. *Saf. Sci.* 82, 325–337.
- TIC, 2016. National Road Safety Action Plan 2015–2017. Transport and Infrastructure Council.
- Truong, L.T., Kieu, L.-M., Vu, T.A., 2016. Spatiotemporal and random parameter panel data models of traffic crash fatalities in vietnam. *Accid. Anal. Prev.* 94, 153–161.
- Truong, L.T., Somenahalli, S.V.C., 2011. Using gis to identify pedestrian-vehicle crash hot spots and unsafe bus stops. *J. Public Trans.* 14 (1), 99–114.
- Wang, J., Huang, H., Zeng, Q., 2017. The effect of zonal factors in estimating crash risks by transportation modes: motor vehicle, bicycle and pedestrian. *Accid. Anal. Prev.* 98, 223–231.
- Wang, X., Wu, X., Abdel-Aty, M., Tremont, P.J., 2013. Investigation of road network features and safety performance. *Accid. Anal. Prev.* 56, 22–31.
- Wang, Y., Kockelman, K.M., 2013. A poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accid. Anal. Prev.* 60, 71–84.
- Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL.
- Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. *Accid. Anal. Prev.* 75, 16–25.