Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 34 (2018) 59–66

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)

Modelling Cycling Flow for the Estimation of Cycling Risk at a Meso Urban Spatial Level

S.Meade [a]*, Dr. K.Stewart [a]

[a]*Edinburgh Napier University, Transport Research Institute, Merchiston Campus,Edinburgh, United Kingdom, EH10 5DT.*

Abstract One of the prevailing challenges in cycling research, or indeed any vulnerable road user research, is the availability of data to ascertain a representative level of 'exposure' or simply how much cycling there is – "when and where". Therefore, it is difficult for researchers and ultimately local authorities to determine if changes in observed accident trends over time are due to increased accident risk, (users or environment becomes more unsafe) or if they are a function of the higher numbers of cyclists using the existing roads and routes resulting in more incidents, i.e. increased exposure. This paper describes the use of recently developed open source transport modelling software and an open source bike routing application to assign realistic cycling flows to the network and validation against observed network link flows. The cyclist flows then provide the 'exposure' variable to examine cyclist safety performance at macro and meso levels using global and local models. The results highlight the need for local level mobility-based exposure metric to describe cyclist safety performance and the superior ability of local models to describe safety performance of cyclists in urban contexts, where more frequently used population based, and global models mask urban spatial patterns and safety performance.

Keywords: Cyclists Flow; Exposure; Macro Model; Meso Model; Geographically Weighted Regression; Safety Performance.

1. Introduction

Cycling as a mode of transport, for any purpose in most countries, is a minority transport choice. Cycling, while beneficial in terms of population health and reducing carbon production, has much higher collision risks, per kilometer

* Corresponding author. *E-mail address:* s.meade@napier.ac.uk

2352-1465
10.1016/j.trpro.2018.11.014

travelled, than for car occupants and despite many countries setting road safety reduction targets, cyclist road safety has lagged behind improvements observed among motorised road users. For example, the UK average risk per billion kilometers travelled, between 2006 and 2015 cyclist killed or serious injury collision risk was almost 20 times higher than for car occupants and over that time cyclist risk has increased by almost 20% while motorised transport risk has experienced an overall improvement (DFT, 2017). Many national authorities seek to increase rates of cycling while at the same time improve road safety, however many authorities lack reliable 'exposure' metrics to calculate collision and injury rates (OECD/ITF, 2013). Detailed traffic data has the greatest potential to improve safety analyses (Lord and Mannering, 2010) however one of the prevailing challenges in cycling research is ascertaining a representative level of 'exposure' or simply "how much cycling happened and where", also traffic exposure is a key determinant of the likelihood of being in a road collision (Loo and Anderson, 2016). Typically trip or population data may be the only information available but the choice, or availability of, 'exposure' variables impact analytical choices when developing accident prediction models (Hauer, 2015).

Transport researchers have noted a correlation between an increase in the number of cyclists (or pedestrians) and a relative reduction of the incident rate of severe/fatal collisions involving cyclists (or pedestrians). The phenomenon ,widely cited in policy, is referred to as the "safety in numbers" (SiN) effect (Jacobsen 2003). At the center of the phenomenon is the observation of non-linearity of risk where, increased 'exposure' results in a less-than-proportional (1:1) increase in the number of collisions (Elvik, 2009). Alternatively, researchers also observed that the risk profiles of cyclists deteriorate if fewer people cycle and more use a car, termed 'risk in scarcity' (Tin Tin et al., 2011). According to the OECD there is little empirical research examining the causal factors that could explain SiN, this may be due to the difficulties surrounding 'exposure'. Furthermore, there is relatively little attention given to the spatial patterns associated with SiN where more pedestrian and cyclist activity leads to lower accident risk. This study aims to add to the understanding of SiN in a spatial context, using a novel approach to obtain 'exposure' estimates and to demonstrate the need to focus on mobility-based exposure measures for cyclists. Further, detailed exposure data may help conclude the debate on the SiN (Bhatia and Wier, 2011, Dozza, 2017).

Researchers have explored national and city level, but there is little research on differences within a city at meso level (Yao and Loo, 2016). Meso level models strike a feasible balance between the level of output information required and cost, where global models may conceal local risk variation and micro level models are often time and cost prohibitive for most cycling budget allocations in highly motorised countries. The modelling approach will be of use to policy makers, engineers and planners who may wish to develop and monitor cycling safety.

2. Study Area

The study area consisted of 111 Scottish Intermediate data zones (IZ) across Edinburgh city. Edinburgh is a compact city with 477,000 inhabitants, 55% of the city's population live within 4 km of the center (CEC, 2014). Edinburgh has experienced a doubling of cycling activity between the years 2001 and 2011 from 2% to 4.8% a trend well ahead of the national average and the mode share varies from 10% to 2.5% within the city.

3. Methodology

The evaluation of cyclist risk includes two parts, firstly the development of a model to provide mobility-based 'exposure' and secondly the specification of global and local models to estimate cyclist risk at a meso spatial scale. Data came from several sources, Department for Transport (DfT) for major and minor roads, City of Edinburgh Council (CEC) automatic counters (AC) at on-road and off-road cycle routes and origin destination (O-D) flow data (ONS, 2014) for the 2011 census, see Table 1 below. Flow models were validated using observed link flows from N=96 counters, N=54 major roads, N=24 minor roads and N=18 on-road and off-road cycle routes. The Department for Transport, STATS19, provided the information on cyclist collisions.

Table 1.Summary of Edinburgh Scottish Intermediate Data Zone - Cyclist trips (ONS, 2014).

Origin-Destination Trips Census 2011	No. trips	(%)	Total	Scottish Intermediate Data Zones *(IZ)*
Inter Zonal (Within Edinburgh)	8808	93		
Inter Zonal (All trips)	9143	96.5		
Intra Zonal	335	3.5	N= 9478	N = 111

3.1. Cyclist Flow Modelling

To estimate population-based cycling exposure (Lovelace et al., 2016) in each IZ formula (1) was used. Where D_{Prod} is the total annual average distance cycled in each IZ, n is the number of people who cycled to work (estimated from Census 2011), f is the frequency of trips (assuming 400 one-way trips per capita each year (Hall et al., 2011)), d is the average trip distance (estimated from TS (2015)) and p is the proportion of bicycle commuter trips (assuming the proportion of commuter trips is one third of all cycling trip purposes (Goodman, 2013; Sustrans, 2017)).

$$D_{Prod} = n \times f \times d \times p \tag{1}$$

Cycling trips to work can be used as a proxy, as in previous research (Lovelace et al, 2016), due to their correlation with all trips (Goodman, 2013). The estimates the actual cyclist routes using O-D information. The study used the functions within the R (CRAN, 2017) package stplanr (Ellison and Lovelace, 2017) developed for sustainable transport planning to estimate mobility-based exposure.

The routing within stplanr uses and external routing engine CycleStreets.net via an application interface program (AIP) developed specifically for cycling based on an Open Street Map (OSM) to replicate the same decisions a knowledgeable cyclist would make to find a route to their destination (Nuttall and Lucas-Smith, 2006). The O-D flows are aggregated in each IZ, then the O-D data is converted into Euclidian flows between O-D pairs, the flow lines are then allocated to the network using CycleStreets.net and finally the overlapping routes aggregated to produce modelled (M) link flows, illustrated in Figure 1 below.

Figure 1. a) IZ with Population Weighted Centroids; b) Euclidean lines between O-D pairs; c) Route allocated flows from stplanr and Cyclestreets.net

Cyclestreets.net has three built-in cycling route options, Fast, Balanced and Quiet to replicate the route choices favoured by fast and experienced utility cyclists to cyclists who may wish to avoid traffic and who are willing to choose less direct routes. All three options were validated against observed (O) cyclist flow volume data, from the N=96 counter locations in Edinburgh. The three models (Fast, balanced and Quiet) M flows were compared to the O link flows using a GEH (Geoffrey Edward Havers) method. The GEH statistic is a modified Chi² statistic used to calculate a value for the difference between O and M flows, it is a widely used criterion (Giuffre et al., 2017) used by UK Highways Agency and Transport for London (TfL) among others (2).

$$GEH_j = \sqrt{\frac{2(O_j - M_j)^2}{O_j + M_j}} \tag{2}$$

Where M is the modelled flow and O_j is the average observed flow. A GEH less than 5.0, for 85% of the model, is acceptable. GEHs between 5.0 and 10.0 may warrant investigation. The data information formats differed, therefore a long-term hourly average flow was used. The GEH has limitations; it does not take account of the variability of the count data and typically uses peak hourly flows to determine 'goodness of fit' (Feldman, 2012). For robustness and to reflect the fact that the GEH is intended for peak hourly motorised traffic flows, the Pearson's correlation coefficient and linear regression were also examined. The best fitting model will provide the vkm variable for the risk models developed.

3.2. Meso Level Global and Local Risk Model Estimation

To estimate the cyclist risk at meso spatial level, safety performance risk models were developed using global and local forms. The dependent variable was fitted with two exposure variables, trip productions (population-based) and vkm (mobility-based), see Table 2 below. Global Poisson (P), Negative Binomial (NB), Zero Inflated (ZI) generalized linear regression models (GLM) were developed and then a local Geographically Weighted Regression (GWPR) model. GWPR refer to a family of regression models where the coefficients may vary spatially (Fortheringham et al., 2002) using coordinates of spatial zone centroids (de Smith et al, 2015). Cyclist risk is assumed to be, non-stationary, rather than stationary as for the global models. Geographically weighted regression (GWR) is an exploratory technique mainly intended to indicate non-stationarity (Bivand, 2017). In general, spatial correlation is one area with a considerable gap between practice and research (Rhee et al., 2016) where GWPR can provide improved model performance (Matkan et al., 2011).

Table 2. Descriptive Statistics of the variables.

Category	Variable	Description	N	Avg	Min	Max	SD
Spatial	IZ	Scottish Intermediate Date Zone	111	-	-	-	-
Collisions	PC	Cyclist Injury (Slight, Serious, Fatal)	240	2	0	25	3
Exposure	Prod	Trip Production in each IZ	9593	86	13	259	56
	vkm	Cyclist Kilometres Travelled per IZ	47688	430	26	1967	392

The P, NB and ZI models provide global estimates whereas the GWPR provided heterogeneous local estimates at each IZ. The models were developed in R (R Development Core Team, 2017) package msme (Hilbe and Robinson, 2015) and GWmodel (Lu et al., 2017). The GWPR models were developed, similarly to the methodology followed by Rhee et al. (2016). Best model fit is the lowest Akaike Information Criterion (AIC), based on the log-likelihood function (Hilbe, 2011), where AIC values differing by less than or equal to 2 are not significant (Nakaya et al., 2005) and AIC differences over 10 suggest that the lower AIC is significant (Hilbe, 2011).

4. Results and Discussion

The fast, balanced and quiet flow models are summarised in Table 3. The vkm totals are smaller for the 'fast' model and higher for the 'quiet' model, which reflects the slightly longer and less direct 'quiet' routes. The estimates are comparable to other research that suggest the total vkm cycled annually is 57.9 mvkm (Sustrans, 2017).

Table 3. Comparison of the CycleStreet.net routing engine options analysis in stplanr.

Route Estimation Method	Segments	Network *(km)*	*vkm*	Annual *mvkm*	Trip Length *(Km) (mean, median, SD)*		
Fast	N=3481	693	47,688	57.2	5.4	4.5	5.7
Balanced	N=3163	675	48,958	58.7	5.6	4.5	6.3
Quiet	N=3207	645	49,348	59.2	5.7	4.6	6.6
D_{Prod}*	-	-	-	53	4.4**	2.1**	-

*Estimated using equation (1) using Census 2011 Table QS701SC (NRS, 2011) data.** TS(2014) Table TD5a, straight line distances.

4.1. Cyclist Flow Model Validation

Exploratory examination of the three flow models, in Figure 2, indicate that the 'balanced' and 'quiet' modelled flows are quite poor predictors of the observer (O) data. The 'fast' modelled flows appear to be more consistent with the AADF and the 12hr adjusted O data. The following assumptions were made, work trips covered a 12hour period between 7am and7pm and AADF represents 16 hours. The census data was collected in March 2011 so a 12hour adjusted estimate was also derived to take account of seasonality. The GEH statistic was calculated using the long term average cyclist per hour unit, O_j in equation (2), and the comparison of validation results are shown in Table 4.

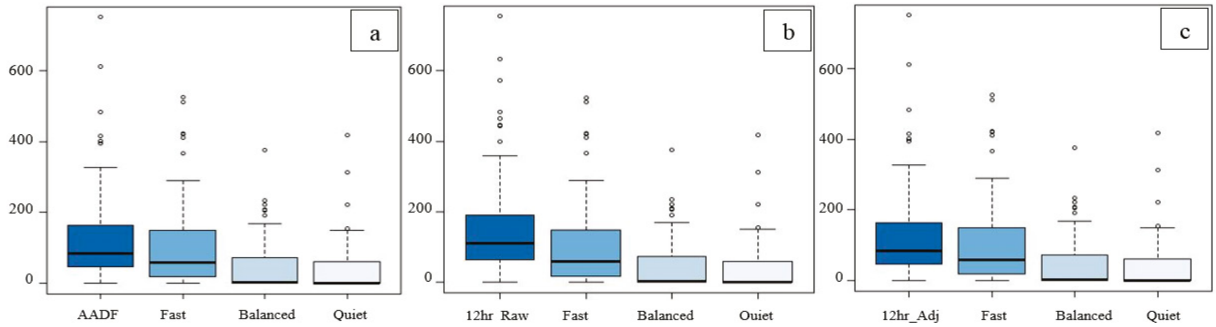

Figure 2. Box Plots 2011 modelled flows versus a) AADF; b) 12hr; c) 12hr adjusted counts.

The GEH statistic indicates that the 'fast' and 'balanced' models have the best fit between the O and the M data. The AADF in combination with the 'fast' model has the highest GEH score. The 'quiet' model comparison to the 12hr and 12hr adjusted do not meet the GEH thresholds. The GEH was not conclusive, this may be due to use of the long term average O_j instead of a peak hour flow. The Pearson's and R^2 however indicated that the 'fast' option, in combination with the 12hr count data, provided the best validation results with a correlation coefficient of 0.815.

While the 'fast' model showed best correlation with O data, the current understanding of cyclist route preferences suggests that cyclists prefer routes with less traffic (Lovelace et al., 2016) such as the 'quiet' or 'balanced' models. Alternatively, it can be argued that the 'fast' model reflects the gender and age bias in Edinburgh (towards young/middle aged affluent men) where fewer women and children or retired people cycle to work (Sustrans, 2017).

Table 4. CycleStreet.net routing engine options validation using GEH/Pearson's/R^2

Validation Statistic	'Fast'	'Balanced'	'Quiet'
GEH(AADF)	97.9%	91.7%	91.7%
GEH(12hr)	90.6%	84.4%	81.3%
GEH(12hr) Adjusted	91.7%	87.5%	85.4%
Pearson's Correlation coefficient (AADF)	0.745	0.616	0.577
Pearson's Correlation coefficient (12hr)	0.815	0.5	0.437
Pearson's Correlation coefficient (12hr) Adjusted	0.694	0.699	0.685
R^2(AADF)	0.551	0.373	0.326
R^2 (12hr)	0.661	0.242	0.183
R^2(12hr) Adjusted	0.476	0.484	0.464

Cyclists, in 2011, appear to favor more direct routes in this study, this suggests that measures such as 'quiet streets or quiet routes' may not successfully attract cyclists in Edinburgh, furthermore the main reason Scottish people cite for not cycling to work is "*too far to cycle*" (TS, 2017) rather than the perceived quietness of the route. However, given gender imbalance this may not always be the case, but it does also suggest that measures or policies aimed to improve cycling safety should focus on links or areas with higher volumes rather than simply aiming to offset routes elsewhere that are assumed to be less dangerous. Loo and Anderson (2016) make the salient point that asking vulnerable road users to avoid travelling on certain routes can be contradictory to promoting mobility and equity.

4.2. Meso Level Risk Models

An initial spatial inspection of the 'fast 'model flows and the cyclist modal share at IZ level indicated that the levels of exposure vary spatially and differ considerably, as illustrated in Figure 3. While a global collision rate will be substantially the same, estimating local collision rates at IZ or ward level, using equation (2) above, provides two very different results. Therefore, collision rates that use population data hold true if the population under review travelled only within the subject area. If we consider trip data, presented in this study and summarised in Table 1, of the total 9478 trips, only 3.5% (N=335 trips) occur within its IZ zone. This illustrates the importance of firstly using vkm as an exposure measure and secondly the need to account for spatial variation.

Considering previous research that examined non-linear relationship between cyclist collisions and cycling volume, the models examined here show mixed results. The global models fitted with vkm had coefficients ranging from β_1= 0.89 to 0.81 suggesting a weak SiN effect. However, the global models fitted with Prod produced results much closer to previous research with coefficients ranging from β_1 = 0.49 to 0.58. In all cases the vkm provided a better prediction than Prod (trip production), Table 5 below. The NB and ZINB global models had the lower AIC (AIC=396) which indicates a significantly better fit than the global models fitted using Prod. These results suggest that the choice of exposure variable may impact the relative safety rate estimated and may overestimate the SiN effect.

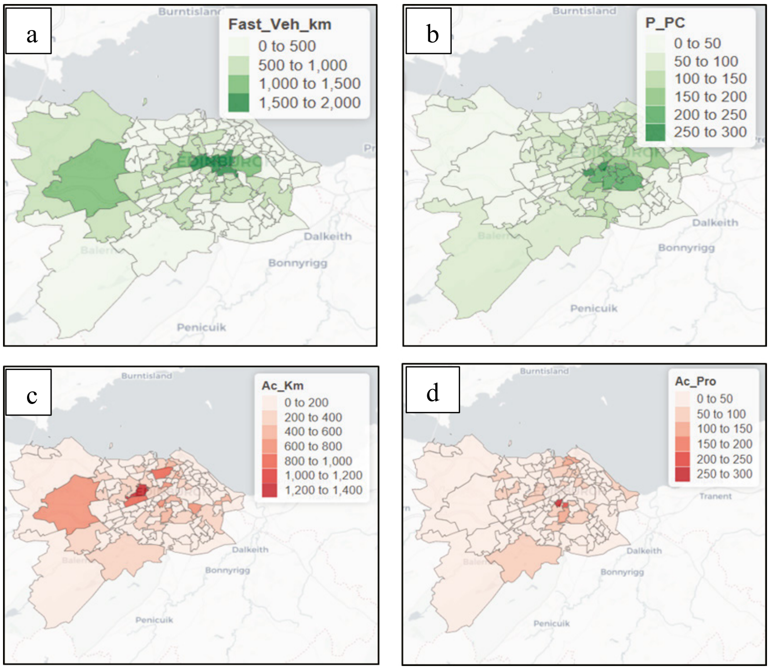

Figure 3. (a) Spatial distribution of vkm, (Fast_Veh_Km); (b) Spatial distribution of Trip Production (P_PC); (c) Spatial distribution of aggregate vkm; (d) Spatial distribution of Trip Production.

A recent longitudinal examination of the effect noted that while results indicated a SiN non-linearity in the data the researchers also noted that the number of collisions increased during the same period (Aldred et al., 2017). The coefficients produced using the vkm would seem more intuitive given the continued increase in cyclist collisions.

Table 5. Meso Level Risk Model Results

Model	Intercept β_0	Ln(vkm) β_1	Quasi-R²	AIC	Intercept β_0	Ln(Prod) β_1	Quasi-R²	AIC
P	-4.62	0.89		440	-1.37	0.49		554
NB	-4.15	0.82		396	-1.75	0.58		435
ZINB	-4.09	0.81		399	-1.60	0.55		439
GWPR	(5.15, -4.83, 3.15)	(0.62, 0.93, 0.95)	0.39	226	(-1.38, -1.32, -1.21)	(0.45, 0.48, 0.49)	0.07	346

Mapping the collision rates, Figure 4(a), illustrates that there is some spatial pattern, the global models don't capture this, where local risk variation or autocorrelation may be present. The Moran I statistic for the dependent variable is positive (I = 0.235) which suggests spatial effects and some clustering with significant p-values (p=0.05) in the central

IZ, Figure 4(b) below. The GWPR estimates local regression equations for each IZ, the coefficients in Table 5 describe the minimum, median and maximum, whereas the P, NB and the ZI provided single coefficient estimate. The GWPR vkm coefficients vary between β_1 = 0.62 to 0.95 and the Prod coefficients vary between β_1 = 0.45 to 0.49 however in the Prod model R^2=0.07 indicating a poor fit in addition to an AIC significantly higher than the vkm fitted model.

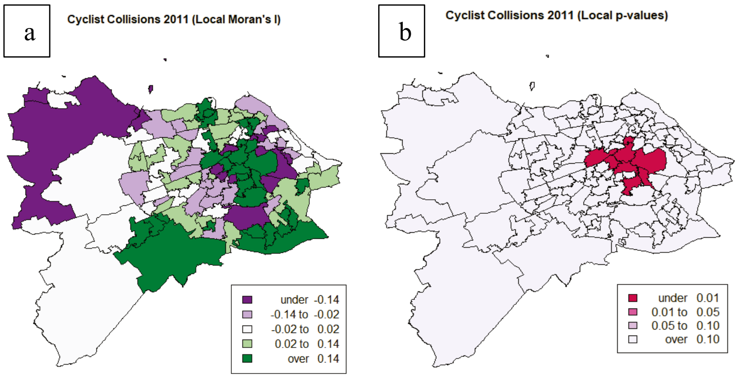

Figure 4. Cyclist collisions, a) Moran's I statistic; b) local P-values.

The GWPR Prod models produce coefficients close to previous research (Jacobsen, 2003; Schepers et al., 2011; Aldred et al., 2017), but the vkm models have lower AIC's and their coefficients suggest that the majority of the IZ coefficients tend towards linearity with a median β_1 = 0.93. This result is in agreement with previous findings that suggest vkm should be used (Matkan et al., 2011; Dozza, 2016; Rhee et al., 2016) which holds for both global and local models.

The GWPR model highlights the non-stationary influence of 'exposure' across IZs. The results suggest that global models may be limited in their ability to explain SiN where spatial autocorrelation is present. As suggested by Dozza, (2017) the number of accidents by the cyclist flow for different geographic regions differentiates and quantifies the SiN effect which this paper demonstrates at a meso spatial level.

5. Conclusions

Bespoke micro-simulation type network models are typically required to provide a mobility-based measure of 'exposure'. This study modelled census O-D data using open source software stplanr and CycleStreet.net validated using a long-term hourly average. This combined approach offers policy makers and planners empirical information, simply "how much cycling happened and where", to monitor cycling safety using normalised 'exposure'.

The global and local models fitted with population-based 'exposure' produce coefficients similar to previous research. However, the mobility-based 'exposure' provided a better model fit with significantly lower AIC results and the coefficients indicated that a much less pronounced apparent SiN effect.

The GWPR improved the model fit compared to the P, NB and ZINB. GWPR provides local parameter estimates that illustrate spatial variations, which are assumed to be stationary in global models often without question (Fortheringham et al., 2002). This study illustrated that the "safety in numbers" effect has spatial variation where parameter estimates ranged from β_1 = 0.62 to 0.95 across N=111 IZ and that meso level analysis may help to explain 'where' the SiN effects manifests. Given the current prevalence of SiN in cycling policy and advocacy, a less than expected or desired effect may be counterproductive, where absolute relative risk remains high in most places or where cycling 'exposure' levels are low. Therefore, global models and models that use population-based 'exposure' should be cognisant of spatial heterogeneity when drawing inference about SiN.

Extending the univariate models to multivariate models may provide further insight. Meso level analysis is also useful because it can merge socioeconomic information with spatial variation and strike a balance between the level of output information and cost.

A limitation of the GWPR maybe its ability to fully captured overdispersion. However, recent research (Silva and Rodrigues, 2014; Rhee et al, 2017) proposes Geographically Weighted Negative Binomial Regression which should be investigated further. Finally, categorising collision frequency by severity using casualty-based cost-weighting to different severities (Yao and Loo, 2012) may also prove beneficial.

References

Aldred, R, Goel, R, Woodcock, J. and Goodman, A., 2017.Contextualising Safety in Numbers: a longitudinal investigation in cycling safety in Britain, 1991-2001 and 2001-2011. Injury and Prevention, 0, 1-6.

Bhatia, R., Wier, M., 2011. Safety in numbers re-examined, Acc.Anal. & Prev., 43, 1, 235-240.

Bivand, R.,2017.Geographically Weighted Regression. [Online]

City of Edinburgh, 2014. Census 2011, City Trends. [Online]

Department for Transport, 2017. Reported Road Casualties Great Britain: 2016 Annual Report, Moving Britain Ahead.

De Smith, M., Goodchild, M. and Longley, P., 2015. Geospatial Analysis, 5th Edition, Winchester, United Kingdom, The Winchester Press.

Dozza, M.,2017. Crash risk: How cycling flow can help explain crash data. Acc.Anal. & Prev., 105, 21-29.

Ellison, R. and Lovelace, R. 2017. stplanr: Sustainable Transport Planning. R package version 0.2.2. [Online]

Elvik, R., 2009. The non-linearity of risk and the promotion of environmentally sustainable transport. Acc.Anal. & Prev., 41, 849–55.

Elvik, R., 2017. Exploring factors influencing the strength of the safety-in-numbers effect. Acc.Anal. & Prev., 100, 75-84.

Feldman, O., 2012. The GEH measures and quality of the highway assignment models. European Transport Conference, Oct 8-10 2012, Scotland, Glasgow, Association For European Transport.

Fotheringham, A.S., Brunsdon, C., and Charlton, M.E., 2002. Geographically Weighted Regression the analysis of spatially varying relationships. Chichester, Wiley.

Giuffrè, T., Trubia, S., & Canale, A., and Persaud, B., 2017. Using Microsimulation to Evaluate Safety and Operational Implications of Newer Roundabout Layouts for European Road Networks. Sustainability, 9, 2084.

Goodman, A., 2013. Walking, Cycling and Driving to Work in the English and Welsh 2011 Census:Trends, Socio-Economic Patterning and Relevance to Travel Behaviour in General. PLoS ONE, 8, 8, e71790.

Hall, C. M., Hultman, J., Gössling, S., Mcleod, D. and Gillespie, S., 2011. Tourism mobility, locality and sustainable rural development. In Sustainable Tourism in Rural Europe: Approaches to Development. London, Routledge, 28–42.

Hauer, E., 2015. The art of regression modeling in road safety.USA, Springer.

Hilbe, J.M, 2011. Negative Binomial Regression, 2nd Edition, Cambridge, Cambridge University Press.

Hilbe, J.M and Robinson, A., 2015. Methods of Statistical Model Estimation. R Package Version 0.5.1. [Online]

Jacobsen, PL., 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Injury and Prevention, 9, 3, 205–9.

Loo, B. and Anderson, T.K., 2016. Spatial Analysis methods of Road Traffic Collisions. United Kingdom, Cornwall, CRC Press.

Lord, D. and Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy Pract., 44, 5, 291–305.

Lovelace, R., Goodman, A., Aldred, R., Berkoff, N., Abbas, A. and Woodcock, J., 2016. The Propensity to Cycle Tool: An open source online system for sustainable transport planning. Journal Of Transport And Land Use, 10, 1.

Lu, B., Harris, P., Charlton. M., Brunsdon, C., Nakaya, T., Gollini, I., 2017. GWmodel Geographically Weighted Models. R version 2.0-5. [Online]

Matkan, A., Afshin, SM., Mirbagheri, B. and Shahri, M., 2011. Explorative spatial analyis of traffic accidents using GWPR model for urban safety planning. In 3rd International Conference on Road Safety and Simulation, September 14-16 2011, Indianapolis, USA, TRB.

Nakaya, T. , Fotheringham, AS., Brunsdon, C. and Charlton, M., 2005. Geographically weighted Poisson regression for disease association mapping. Stat Med, 24,17, 2695–717.

National Records of Scotland, 2011. Census aggregate data. UK Data Service (Edition: June 2016). (Census Table QS701SC).

Nuttall and Martin Lucas-Smith. 2006. CycleStreets. [Online] https://www.cyclestreets.net/.

OECD/International Transport Forum, 2013. Cycling, Health and Safety, OECD Publishing/ITF.

Office for National Statistics, 2014. Census: Special Workplace Statistics (United Kingdom) [WU03BSC_IZ2011_Scotland]

Rhee, K., Kim, J., Lee, Y. and Ulfarsson, G., 2016. Spatial regression analysis of traffic crashes in Seoul. Acc.Anal. & Prev., 91,190-199.

Schepers, J.P., Kroeze, P.A., Sweers, W. and Wüst J.C., 2011. Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections Acc.Anal. & Prev., 43, 853-86.

Silva, A. and Rodrigues. T., 2014. Geographically Weighted Negitive Binomial Regresion Incorporation overdispersion. Statistics and Computing, 24, 769-783.

Sustrans, 2017. Bike Life Edinburgh 2017. [Online]

Tin Tin S., Woodward, A., Thornley, S. and Ameratunga, S. 2011. Regional variations in pedal cyclist injuries in New Zealand: safety in numbers or risk in scarcity?. Aust. N. Z. J. Public Health, 35, 357-363.

Transport Scotland, 2014. Transport and Travel in Scotland 2013, Annex A:Table TD5a: [Distance]

Transport Scotland, 2015. Transport and Travel in Scotland 2014, Annex A:Table TD5a: [Distance] Distance (road network)

Transport Scotland, 2017, "Transport and Travel in Scotland 2016 -26 September 2017 Table 26: [Cycling]

Yao.S, and Loo.S, 2016. Safety in numbers for cyclist beyond national-level and city-level data: a study of the non-linearity of risk within the city of Hong Hong. Injury Prevention, 22, 6, 379-378.