The Visualisation of *SiN* in Edinburgh

Suzanne Meade, Dr. Kathryn Stewart,
TRI, Edinburgh Napier University

Abstract

The SiN (Safety-in-numbers) effect has become prevalent within cycling research in recent years and describes the phenomenon whereby cycling accident risk may fall when the volume of cyclists increases. Such an effect is clearly attractive to policy makers and campaigners and efforts to increase cycling volumes often now actively assume that a risk reduction and hence potential accident cost saving per cyclist will follow. However, according to UK and Scottish research, the SiN effect can coexist with a decline in cycle safety (Aldred et al., 2017) and evidence for beneficial SiN effect may be highly location dependent. Additionally, cycling risks are significantly higher, per kilometre travelled, than both motorised and pedestrian travel. Therefore, risk factors need to be properly understood in order to mitigate them and the presence or absence of SiN effect should be assessed at a relatively small area level. Specifically a national SiN factor should not be globally applied, and factors derived from other countries should be treated with caution.

To understand and measure SiN cycling flow data at both local and link level is required and this may not be readily available. Strategic cycling models such as Cynemon (TfL, 2017), which is similar to those that have long been used to analyse highway transport issues, are currently not feasible for most authorities due to cost. Cyclist flow is strongly correlated with the number of cyclists' crashes such that more cyclists result in more cyclist crashes (although risk may reduce if SiN effect is present). Cycling flow data enables researchers to unpack risk causation and to unravel cyclist related incidents disaggregated between explanatory factors and exposure.

This paper presents a methodology to estimate cyclist flow patterns by utilising recently developed open source analysis tools (Lovelace et al, 2017) and cycling routing engine applications (www.Cyclestreet.net) which were developed specifically for cyclists. This is illustrated using a case study of Edinburgh City. A combination of traditional (Census and Automatic Traffic Counts) and novel (OpenStreetMap) data was used to produce flow estimates at both link and meso-spatial area levels. The range of novel cycling flow data sources available are discussed, while the case study to be presented utilises census data due to its reliability and lack of third party dependence. The variation of SiN effect at a meso-level within Edinburgh City will then be presented and comparisons made to the global Scotland SiN effect and to comparator regions from the EU.

The purpose of this research is to provide transport planners and policy makers with a quantitative and visual analysis of cycling flow to better understand road safety, risk and the SiN effect at a local area level. If increased risk is not mitigated the magnitude of injury and subsequent public health burden may continue to increase. Poor safety perceptions and outcomes will deter transfer to active modes of transport and future government investment may be threatened if anticipated SiN effects fail to produce desired safety improvements.

1 Introduction

The Cycling Action Plan for Scotland (CAPS) vision aims for 10% of everyday cycling trips by 2020 (TS, 2017), the City of Edinburgh Council (CEC) through its Active Travel Action Plan has a higher aim of 15% because in 2009 CEC signed the Charter of Brussels which also includes the road safety target to reduce the risk of a fatal cyclist accidents by 50% by 2020. At the core of the Safety Plan for Edinburgh 2020 is Vision Zero. The ultimate goal is that all users are safe from the risk of being killed or seriously injured.

Scotland has seen an almost doubling in the distance cycled over the past decade with an overall 41% increase in cyclist traffic, kilometres travelled (TS, 2017). In Scotland, Edinburgh stands out as a city that has increased its modal share and cyclist traffic where in some areas of the city, such as the Meadows, the national target in 2011 has already been reached with 10% of trips to work cycled. Further this growth trend has continued to increase throughout the city.

CAPS provide annual reports on a suite of national indicators to inform the national picture of cycling participation. It also sets out to develop local monitoring, using data from local cycle counts and surveys to develop a coordinated approach to data collection. Local level monitoring of cycling safety is included the City of Edinburgh Council (CEC) Active Travel Action Plan (ATAP) targets to produce a cycling casualty rate index to monitor road safety based on count data commencing 2016. This is part of the Charter of Brussels commitment to reduce the casualty rate for cycling (per km travelled) by 50% from 2010 to 2020 as discussed previously. The most prominent cyclist casualty trend, since the mid 2000s, is the rise in adult cyclist casualties both in terms of hospital admissions and police road accident casualties, hospital admissions have increased by 34% and police incidents by 25% in six years Among adult cyclists, Edinburgh, has more than double the rate of police reported casualties observed in comparison to Scotland's other large cities of Aberdeen, Dundee and Glasgow. Similarly, in terms of hospital admissions there has also been an increased in adults cyclist admissions across Scotland's four largest cities in recent years (GCPH, 2015).

Some of the barriers to cycling include: safety, perceived safety (especially on busy roads); lack of secure cycle parking; hills, weather, cycle theft; lack of information and skills; and finally culture and attitudes. To address these issues ATAP aims to:

- deliver a citywide 'QuietRoutes' network that people perceive as safe and attractive (cater for less confident cyclists);
- reduce traffic speeds; and
- adopt cycle friendly design principles for all streets.

The ATPT also collect and publish monitoring data to evaluate progress against targets and indicators published in the Edinburgh Bike Life (Sustrans, 2017) report. This report is similar to the Bike Account, produced by the Cycling Embassy in Denmark, which provides detailed monitoring information about cycling from several different sources together with new research into one coherent annual report.

Local governments and advocacy groups in Scotland (CTC, 2016; CEC, 2016) promote the increasingly popular transport paradigm 'Safety in Numbers' to encourage active travel through more cycling and walking. The research evidence often cited states that double the cycling or walking volume is associated with only a 32 % increase in the expected accidents (Jacobsen, 2003).

Cycling as a mode of transport, for any purpose, in Scotland is a minority transport choice which accounts for 1.6% of journeys to work (Scottish Census 2011, National Records of Scotland) however it is more prevalent in urban areas, 4% versus 1% in rural areas (Scottish Household Survey, 2015). Consequently, the availability of data to ascertain a representative level of 'exposure' or simply how much cycling there is – "when and where" is very limited and is one of the prevailing challenges in cycling research, or indeed any vulnerable road user research and which this research attempts to address.

Therefore, it is difficult for researchers and 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. According to the ITF/OECD (2013) most authorities lack the factual basis to assess cyclist safety or the impact of 'safety improving' policies.

2 Study Area

This study took place in Edinburgh the capital city of Scotland. The study area consisted of 111 Scottish Intermediate data zones (IZ). Edinburgh is a compact city with 477,000 inhabitants, 55% of the city's population live within 4 km of the centre (CEC, 2013). 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. Within the city mode share varies from 10% to 2.5%.

3 Methodology

The evaluation of cyclist risk detailed in this study includes two parts, firstly a model has been developed which provides mobility-based 'exposure' and secondly global and local models to estimate cyclist risk at meso spatial scale within an urban area have been specified, as shown in Figure 1 below.

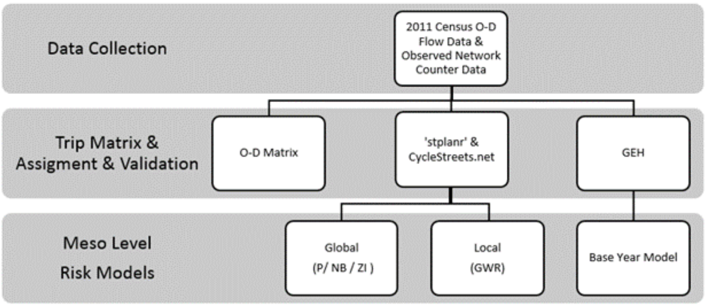


Figure 1. Study procedure.

Data for the study 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 the 2011 census provided the origin destination (O-D) flow data sets (ONS, 2011) for the O-D matrix.

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

There were N=9478 trips to work by bicycle, shown in Table 1, N=9143 trips were within Edinburgh and N=335 (3.5%) of trips remained within their origin IZ. The study used observed data from N=96 counters to validate modeled link flows, N=54 major roads, N=24 minor roads and N=18 on-road and off-road cycle routes, Figure 2 below.

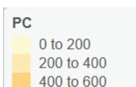
The Department for Transport, STATS19, provided the information on cyclist collisions. The cyclist data varied in metric and completeness, for example; the DfT provided average annual daily flow (AADF) estimates and weekday 12-hour manual counts. The CEC data provided 24hr counts and the O-D flow data considers trips to work on an average day.

3.1 Estimation of cyclist volumes

This section discusses approaches to deriving cycling volumes, first a population-based estimate using modal share data and second a mobility-based estimate from modelled cyclist flows on the road network.

3.1.1 Population-based Estimates

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



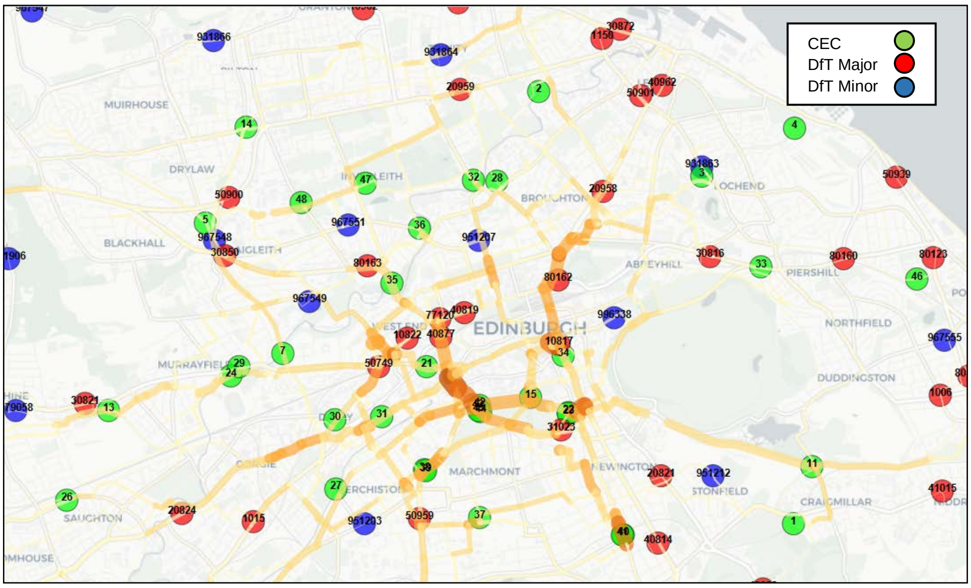


Figure 2. Edinburgh Cyclist counter (N=96) locations and Cyclist flow.

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)).

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

As in previous research (Lovelace et al, 2016) it is assumed that cycling trips to work can be used as a proxy for all cycling trips because they are highly correlated to cycling modal share for all trips (Goodman, 2013).

3.1.2 Mobility-based Estimates

The calculation of the mobility-based exposure estimates the actual cyclist routes using O-D information assigned to the route and cyclist infrastructure using a routing engine algorithm. The study used the functions within the R (2017) package stplanr (Ellison and Lovelace, 2017) developed for sustainable transport planning.

The routing within stplanr uses an 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)

Figure 3. a) IZ with Population Weighted Centroids; b) Euclidean lines between O-D pairs; c) Route allocated flows from stplanr and Cyclestreets.net

The flow volumes were estimated using stplanr::overline function that aggregates overlapping lines (Rowlingson, 2015). Figure 3 illustrates the process, first the O-D flows are aggregated in each IZ, then the O-D data is converted into Euclidian flows between O-D pairs (via matrix estimation using a doubly constrained gravity model), the flow lines are then allocated to the network using CycleStreets.net and finally the overlapping routes aggregated to produce modelled (M) link flows.

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.

3.2 Validation

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) amongst others (2)

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

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 deemed acceptable whereas GEHs between 5.0 and 10.0 may warrant investigation. The GEH statistic was calculated using the long term average cyclist per hour unit, O_j in equation (2). Due to the different count data formats and to facilitate calculation of, O_j the following assumptions were made, work trips covered a 12hour period between 7am and 7pm and AADF represents 16 hours. Further, the census data was collected in March therefore a 12hour adjusted estimate was also derived to take account of seasonality.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.

3.3 Meso Level Global and Local Risk Models

To estimate the cyclist risk at meso spatial level, safety performance risk models were developed using global and local forms. Global models assess factors effecting risk at a country or regional level and local models assess the effect of factors across small areas to show spatial variation.

The dependent variable was fitted with two exposure variables, trip productions (population-based) and vkm (mobility-based), summary data is shown in Table 2. Three models; Poisson (P), Negative Binomial (NB), Zero Inflated Negative Binomial (ZINB) and generalised linear regression models (GLM) were then developed and finally a Geographically Weighted Regression (GWPR) model was produced. The P, NB and ZINB models provide global estimates of cyclist risk. The GWPR provides local estimates of cyclist risk which vary spatially.

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 lowest Akaike Information Criterion (AIC) indicates goodness of fit between models. AIC values between two models that are less than or equal to 2 don't indicate a substantial difference, an AIC difference of over 10 suggests that the model with the lower AIC is significantly better (Hilbe, 2011).

4 Results and Discussion

This section discusses results from the following:

- Comparison of the three cyclist flow models (Fast, balanced and quiet),
- Observed counts and modelled flows validation,
- Comparison of population-based and mobility-based cyclist volumes
- Cyclist collision risk rates based on mobility-based cyclist volumes, and finally
- Comparison of meso scale risk models using global and local model forms.

4.1 Flow Model Comparisons

The summary statistics of the fast, balanced and quiet flow models are presented in Table 3, below. The trip lengths, measured by distance, show similar tends where the 'fast' model mean trip length is shortest. The 'fast' and 'quiet' flow models are illustrated in Figure 4 and 5 below.

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

CycleStreets.net 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

The 'fast' model vkm totals are smaller than the 'quiet' model total vkm, which reflects the slightly longer and less direct 'quiet' routes. The 'fast' model covers the largest proportion of the available network, flows tend to be higher on busy main roads, which provide directness and have less flows on quieter routes or off-road routes when compared to the 'quiet' model.

Figure 4. Cyclist flow "Fast" (white) option results mapped against quiet roads in green.

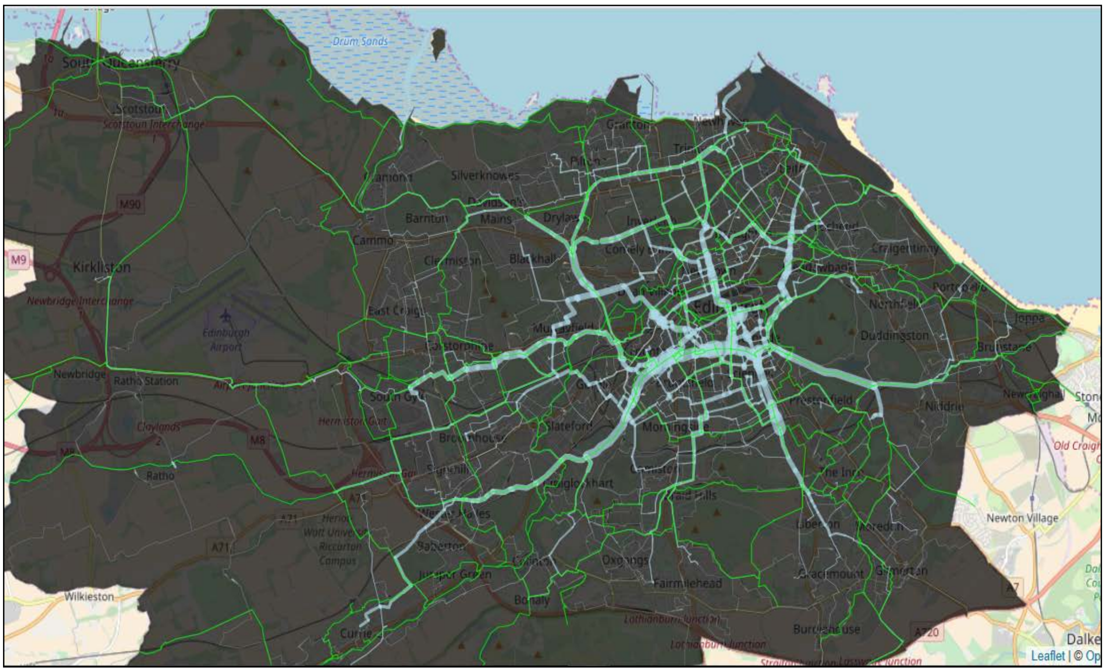


Figure 5. Cyclist flow “Quiet” (blue) option results mapped against quiet roads in green.

4.2 Cyclist Flow Model Validation

Exploratory examination of the three flow models, Figure 6 below, against AADF, 12hr and 12hr adjusted observed data indicates that the ‘balanced’ and ‘quiet’ modelled flows are quite poor predictors of the observer counter data and that the ‘fast’ modelled flows appear be more consistent overall.

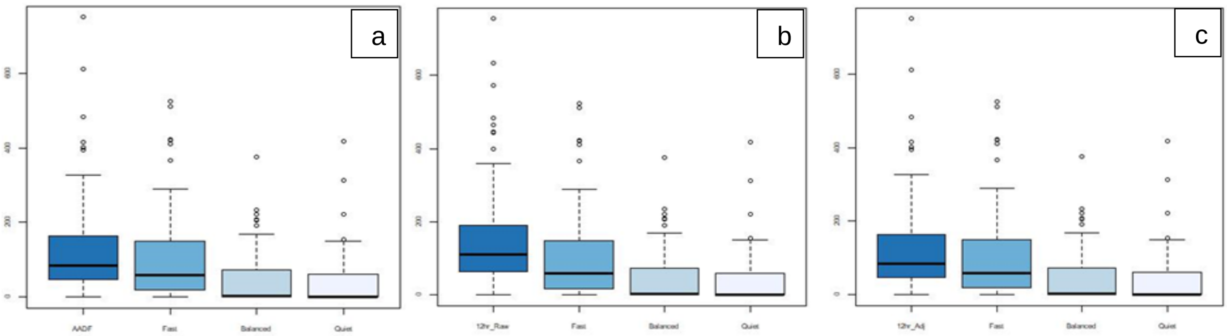


Figure 6. Box Plots 2011 modelled flows versus a) AADF; b) 12hr; c) 12hr adjusted counts

Following from these initial observations, the GEH statistic indicates that the ‘fast’ and ‘balanced’ models have the best fit between the observed and the modelled data. The AADF in combination with the ‘fast’ model produced the highest GEH score, while the combination of the 12hr and 12hr adjusted with the ‘quiet’ model did not meet GEH thresholds, Table 4 below. However, the GEH was not conclusive because several combinations met the thresholds. This may be due to use of the long-term average 0_j instead of a peak hour flow and therefore more robust validation statistic were used.

The Pearson’s and R^2 indicated that the ‘fast’ option, in combination with the 12hr count data, was the best fit for the data with a correlation coefficient of 0.815. The levels of correlation are high, and whilst the use of a long-term hourly average may have hindered the GEH the correlation result is conclusive.

The current understanding of cyclists route preferences suggests that cyclists may prefer routes with less traffic (Lovelace et al., 2016) and better perceived safety, the results here would seem to show a preference for directness. This result is interesting because the more direct routes in this study, where larger flows were observed on the shared road network, typically don't have continuous on-road cycle facilities.

Table 4. Comparison of the CycleStreet.net routing engine options analysis in stplanr.

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

Alternatively, it can be argued that the 'fast' model reflects the existing 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) and thus make less use of the 'quiet' or 'balanced' routes. Another issue to consider is the social imbalance where adults who cycle to work or study in the least deprived decile are 2.7 times higher than the most deprived decile (GCPH, 2015).

While the 'fast' option includes some flows on Edinburgh's 'Quiet Routes' the predominant trend may suggest that these measures may not successfully attract 'fast' cyclists away from more risky routes given that the main reason Scottish people don't cycling to work is because its "*too far to cycle*" (TS, 2017) rather than the perceived lack of quietness of the route. As discussed previously, Table 3 above, 'quiet' routes are slightly longer on average than the 'fast' options. Loo and Anderson (2016) argue that expecting vulnerable road users to avoid travelling on certain routes can be contradictory to promoting their mobility and maintaining equity, cyclists who value directness over safety would likely fall into this category.

A key focus of Edinburgh's cycling investment over the next few years will be the "QuietRoutes" network (Sustrans, 2017) which aims to provide facilities for less confident cyclist and hopefully more unaccompanied 12 year olds so that in time the cycling population may grow and become more age and gender balanced. The development of the flow model here provides data on cyclist flows that can be used to either inform or monitor policies and measures.

4.3 Comparison of the flow model results and population data

A recent study suggests that the total vkm cycled annually is 57.9 mvkm (Sustrans, 2017) which is comparable to the estimate in Table 5, the population estimate is lower 53 mvkm which would over estimate risk and under estimate actual cycling levels.

Table 5. Comparison of the CycleStreet.net routing engine options analysis in stplanr.

CycleStreets.net 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
Sustrans (2017)				57.9			
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.

The 'fast' model provides mobility-based mvkm 'exposure' and cyclist modal share provides population–based 'exposure'. The spatial distributions of the two measures of 'exposure' (population v's distance) differ considerably as highlighted in bold text in Table 6 below. The accident risk for 96.5% of commuters will be attributable to an IZ other than their IZ of origin, where only 3.5% of trips remain in their IZ.

Population based estimates at meso and micro level are likely to misrepresent activity, for example Inverleith and City Centre wards are roughly twice the cyclist volumes compared to the modal share population.

Table 6- Comparison of the CycleStreet.net routing engine options analysis in *stplanr*.

Ward Name	veh_km	% persons aged 16 to 74 who cycle to work	% of mvkm in each ward 'fast'.
Colinton/Fairmilehead Ward	1199.849	4.6	2.5
Portobello/Craigmillar Ward	1937.139	4.6	4.1
Sighthill/Gorgie Ward	3570.222	3.0	**7.5**
Pentland Hills Ward	2322.498	3.4	4.9
Liberton/Gilmerton Ward	1230.482	2.5	2.6
Fountainbridge/Craiglockhart Ward	2353.732	**6.9**	4.9
Meadows/Morningside Ward	5070.379	**9.9**	**10.6**
Inverleith Ward	4484.553	4.5	**9.4**
Forth Ward	2351.24	4.5	4.9
City Centre Ward	5564.933	4.4	**11.7**
Craigentinny/Duddingston Ward	2029.61	4.4	4.3
Drum Brae/Gyle Ward	1264.593	2.9	2.7
Corstorphine/Murrayfield Ward	3001.181	4.5	6.3
Southside/Newington Ward	5102.885	**9.3**	**10.7**
Leith Walk Ward	2258.176	4.6	4.7
Leith Ward	1313.268	4.8	2.8
Almond Ward	2609.887	3.1	5.5

4.4 Cyclist Collision Risk

The collision risk rates were calculated using the 'fast' flow model volumes, discussed earlier, and cyclist collisions, aggregated at IZ level, for both killed and serious injuries (KSI) and all injury collisions, illustrated in Figure 7 below.

To frame cyclist risk within the context of overall road risk it is worth comparing the KSI and all injury risk rates. The average casualty risk in Edinburgh for any severity or mode was 0.47 per mvkm in 2011 and improved slightly to 0.44 per mvkm in 2016 (TS, 2017). The KSI average casualty risk was 0.06 and 0.057 per mvkm over the same period.

By comparison the average cyclist casualty risk was 4.2 per mvkm for all injuries and 0.63 per mvkm for KSI in 2011. That is roughly a seven and ten fold difference respectively.

The figure below shows ranked risk rates according to times over or under the average risk rate for KSI and all injury. Firstly, this illustrates the spatial pattern and second that both the KSI and all injury collision rates can be several times higher than the average in some IZ's.

While it may be argued that drivers travel on average a greater distance per trip, it is worth noting that the average cyclist trip is 4.6 km and driver trip is 10.5 km (TS, 2017) so roughly half, showing that the risk gap is still considerable. Where an individuals main mode of transport is their bike or for those who don't have access to a car or public transport this is a considerable transport risk imbalance.

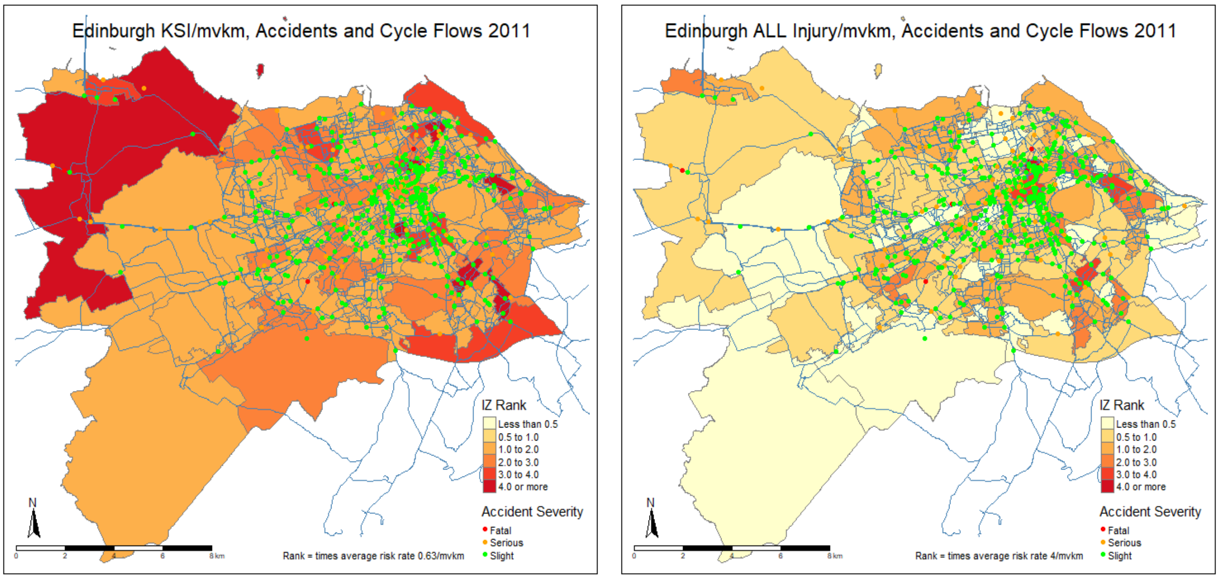


Figure 7. Cyclist risk, collisions (Fatal, Serious and Slight injury) per mvkm.

Given the increase in cycling since 2011 and modest change in the overall number of cyclist collisions suggests that policies such as 20mph, raising awareness and education are having some positive impact on cyclist safety. However, under reporting may conceal part of the absolute change observed.

4.5 Comparison of Meso Level Global and Local Risk Models

As discussed in the introduction, previous research describes a non-linear relationship between cyclist collisions and cycling volume, where more cycling tends to result in less collision risk.

4.5.1 Global models

Global models of various forms were fitted, the NB model form was found to provide the best fit for the data. Models were developed for KSI and all injuries separately and for both the mobility-based exposure and population-based exposure, the results are shown in Table 7 below.

Table 7. Comparison of the CycleStreet.net routing engine options analysis in stplanr.

Model	Intercept β_0	Ln(vkm) β_1	R^2	AIC	Intercept β_0	Ln(Prod) β_1	R^2	AIC
	Mobility-based				Population-based			
Global Models								
NB -ALL	-4.15	0.82		396	-1.75	0.58		435
NB - KSI	-7.57	1.055		143	-4.21	0.69		161
Local Models								
GWPR - ALL	-5.15 to -3.15	0.62 to 0.95	0.39	226	-1.38 to 1.21	0.45 to 0.49	0.07	346
GWPR - KSI	-8.44 to -7.06	0.97 to 1.2	0.3	82	-4.79 to -3.3	0.5 to 0.8	0.07	101

The results illustrated that the choice of exposure variable influences the model results. Global models fitted with mobility-based data (vkm) had coefficient of β_1= 0.82, which show a small “safety in numbers” effect, whereas models fitted with population-based data (Prod) produced results much closer to previous research with coefficient of β_1= 0.58 which suggests the presence of a much higher “safety in numbers” effect. The KSI models indicate that the “safety in numbers” effect is much less pronounced for higher severities.

Furthermore, the models fitted with the vkm from the earlier flow model have significantly better fit than models developed using Prod, the vkm had lower AIC in each case. This implies that the choice of exposure variable in cyclist safety analysis may impact the relative safety rate estimated from the model and hence the “safety in numbers” effect may be over-estimated if population based exposure is used.

4.5.2 Local models

The results of the local model using GWPR are listed in Table 7 above and they are also mapped in Figure 8 below.

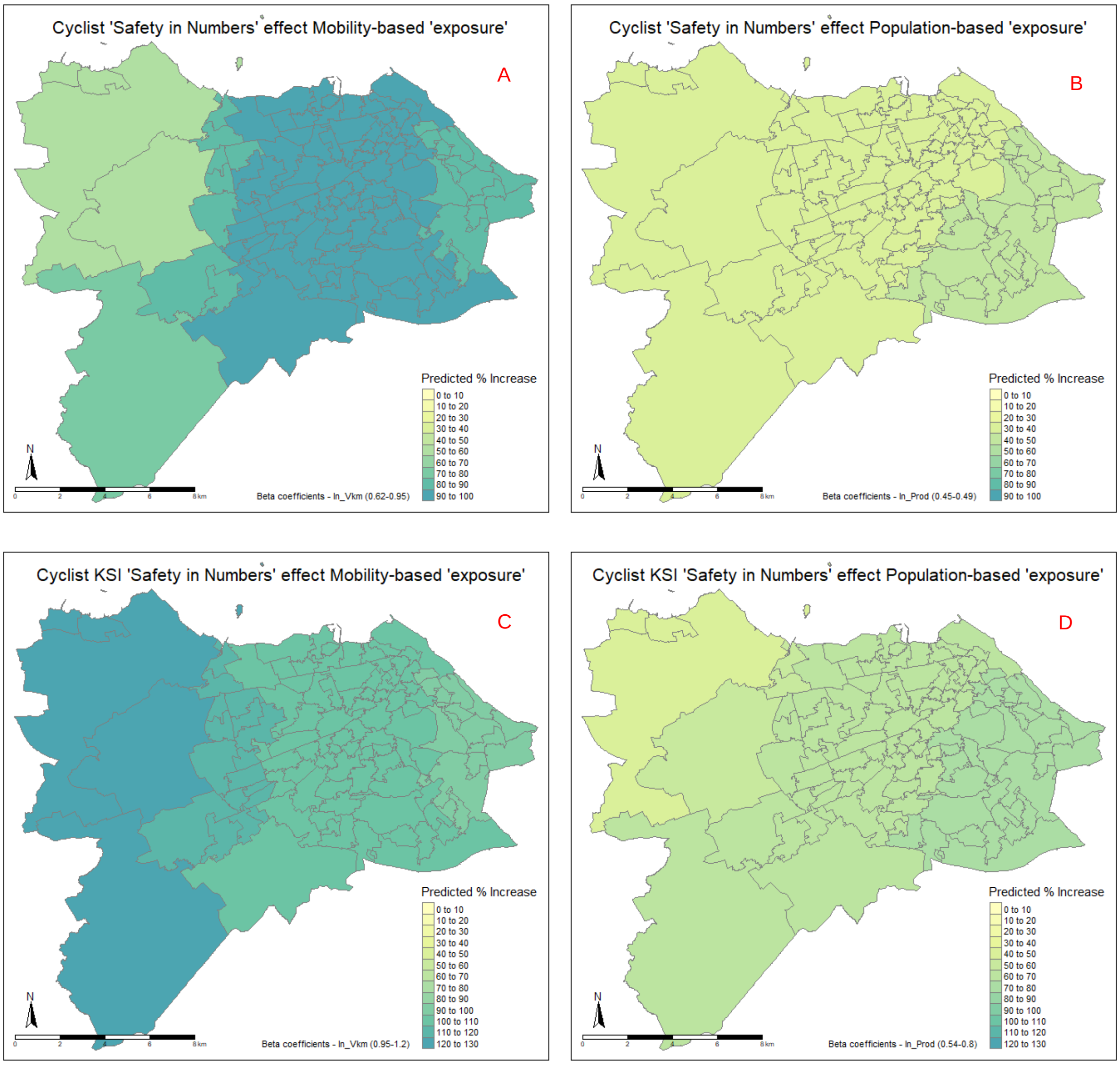


Figure 8. Percentage of expected collisions if cycling doubles with the modelled “Safety in Numbers” effect.

Similar to the global models above, local models were developed for KSI and all injuries separately and for both the mobility-based exposure and population-based exposure. The local model form produces local coefficients for each meso area modelled, in this case the N=111 IZ.

Risk varies spatially, as seen in Figure 7 above, While a global collision rate masks spatial variation, estimating local collision rates at IZ or ward level provides two very different results (see Table 6 below).

The local model coefficients vary between β_1= 0.62 to 0.95 and β_1= 0.97 to 1.2 for all injury and KSI models respectively, which illustrates the spatial variation of the “safety in numbers” effect. Similar to the global models, the local models fitted Prod produced results much closer to previous research with coefficients ranging from β_1= 0.45 to 0.49 and β_1= 0.5 to 0.8 for all injury and KSI respectively which suggests the presence of a substantial "safety in numbers" effect. The KSI models again indicate that the "safety in numbers" effect is much less pronounced for higher severities.

The models fitted with the vkm from the earlier flow model again have significantly better fit than models developed using Prod, the vkm. In addition, the $R^2 =$ 0.07 was very poor for both KSI and all injury fitted with Prod.

Further, comparing the risk rates mapped in Figure 7 with the spatial variation of the model coefficients for vkm, Figure 8 A and D, and Prod, Figure 8 B and C, illustrate that the Prod does not capture the spatial risk variation.

To demonstrate the difference between the mobility-based and population-based model estimates, the coefficients were converted to expected percentage increase given a double increase in cycling volumes in each IZ. Examining the results in Figure 8 A and C, the predicted increases for both the KSI and ALL models range from 60% to over 100% in some locations. Examining the results in Figure 8 B and D, the predicted increases for both the KSI and ALL are considerably lower. The models examined here provide two interesting results:

1. Considering the results in Table7, vkm and Prod represent 'exposure' and they are both linked to increased frequency of cyclist crashes as expected, however the difference between the vkm and Prod was not expected to be so pronounced. In both the global and local models Prod was found to underestimate the effect of cyclist volumes.

2. There is some spatial pattern or spatial dependence present in the data which illustrate that the "safety in numbers" effect varies and can be quantified locally. The GWPR model highlights the non-stationary influence of 'exposure' across IZs. Previous research, using global models, does not capture this.

5 Conclusions

The study may provide transport planners and policy makers with a quantitative and visual analysis of cycling flow to gain insight into cyclist road safety, risk rates and the "safety in numbers" effect at a local area level using Edinburgh as a case study.

5.1 Cycling Flow Modelling

Bespoke micro-simulation type network models have typically been required to provide a mobility-based measure of 'exposure'. This study utilised census O-D data and open source software stplanr and CycleStreet.net to develop a strategic cycling flow model. Available data from several sources were combined and compared using a long-term hourly average flow to facilitate validation between observed and modelled flows. This combined approach offers policy makers and planners empirical information, simply "how much cycling happened and where". This information may be useful to monitor and compare national and local cycling performance indicators.

CEC use parallel approaches to cycle infrastructure in the city, in the first instance 'Quiet Routes' network to cater for less confident cyclists and secondly a move towards a Cycle Friendly City. The study found that the 'fast' cyclists routing algorithm produced link-flow estimates that best fit to observed count data; this may indicate uptake bias where faster cyclists are the majority, typically male and more affluent, and demonstrates the importance of validating cyclists behavioural routing to local conditions. While the 'fast' option includes some cycle traffic on Edinburgh's 'Quiet Routes' the predominant trend suggests that measures such as ''Quiet Routes" may not attract 'fast' cyclists, furthermore the main reason cited for not cycling to work is "too far to cycle" (TS, 2017) suggesting that longer quiet routes may not be attractive.

The flow model developed enables monitoring of flow balance between 'fast' and 'quiet' routes which should change in time as the cyclist population grows to include a wider more balanced cycling population with more women and child cyclists. This methodology can be developed to generate a cycling safety index mapping at a link or meso area level, see Figure 7 above. In this way, the network can be monitored and ranked for safety or Level of Service and benefit cost ratios of proposed schemes.

5.2 Accident Models:

This study demonstrates that a "safety in numbers" effect is present but that it is strongly location dependent and has a much less pronounced effect than expected from EU studies.

The GWPR spatially disaggregated models provided local estimates of cyclist risk and mobility-based 'exposure' using mvkm provided a better model fit with significantly lower AIC than population-based data. The GWPR models show that the "safety in numbers" effect varies spatially and that some locations show little or no effect. Therefore, global "safety in numbers" factors should not be assumed or applied at a local level, and furthermore research from other countries should be treated with caution.

Models that use population-based 'exposure', where data availability may have restricted analytical choices, should be cognisant of spatial variation and the exposure variable specified when drawing inference about "safety in numbers". The results presented here suggest that the "safety in numbers" effect may be overestimated if a population based exposure measure is used which is consistent with the absolute increase in casualties recorded in hospital admissions and police records.

Given the current prevalence of "safety in numbers" in cycling policy and advocacy, overestimating the effect may be counterproductive particularly where absolute risk remains high or where cycling 'exposure', levels are low. Visualising the model results across local area zones provides a more accessible platform to communicate information to non-technical practitioners and decision makers.

5.3 Future work

A fundamental benefit of meso analysis is the ability to merge socioeconomic information with spatial variation. This Evaluating urban areas at a meso level using local models spatially disaggregate the effects of independent variables, whereas global models report an average. Mesoscopic models also strike a balance between the level of output information and cost, where global models do not provide enough detail at a local level and microscopic models are time and cost prohibitive such as the Cynemon (TfL, 2017).

Finally, this research used all collision severities, modelling only killed and serious injury (KSI) collisions would hinder the reliability of the models due to small sample size and prevalence of zero KSIs in small geographic areas, however previous cyclist safety performance research typically considers only fatal and serious collisions. Categorising collision frequency by severity by developing a casualty-based cost-weighting to different severities (Yao and Loo, 2012) may 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: can we make valid or practical inferences from available evidence? Accident Analysis and Prevention, 43, 1, 235-240.

Lu, B., Harris, P., Charlton. M., Brunsdon, C., Nakaya, T., Gollini, I., 2017. GWmodel Geographically Weighted Models. R package version 2.0-5[Online], CRAN, https://cran.r-project.org/web/packages/GWmodel/GWmodel.pdf. [Accessed 12 Jan 2018].

Bivand, R.,2017. Geographically Weighted Regression. [Online], CRAN, https://cran.r-project.org/web/packages/spgwr/vignettes/GWR.pdf. [Accessed: 2 Jan 2018].

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

City of Edinburgh, 2016. Active Travel Action Plan 2016. [ONLINE]

CTC, 2016, Source; http://www.cyclinguk.org/campaign/safety-in-numbers.

Department for Transport, 2017. Reported Road Casualties Great Britain: 2016 Annual Report, Moving Britain Ahead. [Online], Department for Transport, Scottish Government, Welsh Assembly Government, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/668504/reported-road-casualties-great-britain-2016-complete-report.pdf. [Accessed: 21 Feb 2018].

Dozza, M.,2017. Crash risk: How cycling flow can help explain crash data. Accident Analysis and Prevention, 105, 21-29. http://www.sciencedirect.com/science/article/pii/S0001457516301464.

Elvik, R., 2009. The non-linearity of risk and the promotion of environmentally sustainable transport. Accident Analysis and Prevention, 41, 849–55.

Elvik, R., 2017. Exploring factors influencing the strength of the safety-in-numbers effect. Accident Analysis and Prevention,100, 75-84. http://www.sciencedirect.com/science/article/pii/S000145751630450X.

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. [Online] https://aetransport.org/en-gb/past-etc-papers/conference-papers 2012?search=GEH&abstractId=3836&state=b.

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. http://www.mdpi.com/2071-1050/9/11/2084.

Glasgow Centre for Population Health, 2015. Trends in pedestrian and cyclist road casualties in Scotland. [Online] http://www.gcph.co.uk/assets/0000/5206/Pedestrian_and_cyclist_casualties_analysis_FINAL.pdf. [Accessed 17 Mar 2018].

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. https://doi.org/10.1371/journal.pone.0071790.

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], CRAN, https://cran.r-project.org/web/packages/msme/msme.pdf. [Accessed 20 Feb 2018].

Jacobsen, PL., Ragland, DR. and Komanoff C., 2015. Safety in numbers for walkers and bicyclists: exploring the mechanisms. Injury and Prevention, 21,4, 217–20. http://injuryprevention.bmj.com/content/21/4/217.long.

Jacobsen, PL., 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Injury and Prevention, 9, 3, 205–9. http://injuryprevention.bmj.com/content/9/3/205.

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. https://ceprofs.civil.tamu.edu/dlord/Papers/Lord-Mannering_Review.pdf.

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, . https://www.jtlu.org/index.php/jtlu/article/view/862.

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

Lovelace, R. and Ellison, R., 2017. stplanr: Sustainable Transport Planning. R package version 0.2.2. [Online], CRAN, https://CRAN.R-project.org/package=stplanr. [Accessed: 02 Jan 2018].

Matkan, A., Afshin, SM., Mirbagheri, B. and Shahri, M., 2011. Explorative spatial analysis 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, Transport Research Board. [Online] http://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS/1/Matkan,A.pdf.

National Records of Scotland, 2011. Census aggregate data. UK Data Service (Edition: June 2016). 2011 Usual method of travel to work, All residents aged 16-74 in employment in the week before Census; main mode of travel by distance (Census Table QS701SC). [Online] http://dx.doi.org/10.5257/census/aggregate-2011-1. [Accessed: 12 Oct 2017].

Nakaya, T. , Fotheringham, AS., Brunsdon, C. and Charlton, M., 2005. Geographically weighted Poisson regression for disease association mapping. Stat Med, 24,17, 2695–717. https://www.ncbi.nlm.nih.gov/pubmed/16118814.

Nuttall and Martin Lucas-Smith. 2006. CycleStreets. [ONLINE] https://www.cyclestreets.net/. [Accessed 20 January 2018].

OECD/International Transport Forum, 2013. Cycling, Health and Safety, OECD Publishing/ITF. [ONLINE] http://dx.doi.org/10.1787/9789282105955-en.

Office for National Statistics, 2014. Census: Special Workplace Statistics (United Kingdom) [WU03BSC_IZ2011_Scotland, "Location of usual residence and place of work by method of transport - All usual residents aged 16 and over in employment the week before the census - 2011 Intermediate Zones]. UK Data Service Census Support. [ONLINE] https://wicid.ukdataservice.ac.uk. [Accessed: 12 Oct 2017].

R Core Team, 2017. R: A language and environment for statistical computing. [Online] http://www.R-project.org/

Rhee, K., Kim, J., Lee, Y. and Ulfarsson, G., 2016. Spatial regression analysis of traffic crashes in Seoul. Accident Analysis and Prevention, 91,190-199. http://www.sciencedirect.com/science/article/pii/S0001457516300562.

Rowlingson, B., 2015. Overlaying lines and aggregating their values for overlapping segments. [ONLINE] http://gis.stackexchange.com. [Accessed 30 October 2017].

Transport Scotland, 2014. Transport and Travel in Scotland 2013, Annex A:Table TD5a: [Distance] Distance summary statistics by mode of transport 2013 - Glasgow. [Online] https://www.transport.gov.scot/media/31945/j333840.pdf. [Accessed: 12 Mar 2018].

Transport Scotland, 2015. Transport and Travel in Scotland 2014, Annex A:Table TD5a: [Distance] Distance (road network) summary statistics by mode of transport 2014, Glasgow. [Online] https://www.transport.gov.scot/publication/transport-and-travel-in-scotland-2014/j389989-64. [Accessed: 12 Mar 2018]

Transport Scotland, 2017, "Transport and Travel in Scotland 2016 -26 September 2017 Table 26: [Cycling] Reasons why do not cycle to work, 2009-2014", Glasgow, Accessed 08/02/2018, Source: https://www.transport.gov.scot/publication/26-september-2017-transport-and-travel-in-scotland-2016/table-26-cycling-reasons-why-do-not-cycle-to-work-2009-20141/.

Transport Scotland, 2017, "Transport and Travel in Scotland 2016 -26 September 2017: Table TD5a: [Distance] Distance (straight line) summary statistics by mode of transport, 2016. [Online] Glasgow, https://www.transport.gov.scot/publication/26-september-2017-transport-and-travel-in-scotland-2016/table-td5a-distance-distance-straight-line-summary-statistics-by-mode-of-transport-2016/. [Accessed: 17 Mar 2018]

Transport Scotland, 2017. Scottish Transport Statistics No 36 2017

Schepers, J.P., Kroeze, P.A., Sweers, W. and Wüst J.C., 2011. Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections. Accident Analysis and Prevention, 43, 853-86.

Schepers.P, Hagenzieker.M ,Methorst.R, and van Wee.B, 2013, "A conceptual framework for road safety and mobility applied to cycling safety". Accident Analysis and Prevention, 62, 331-340.

Silva, A.R. and Rodrigues. T.C.V., 2014. Geographically Weighted Negative Binomial Regression Incorporation overdispersion. Statistics and Computing, 24, 769-783.

Sustrans, 2017. Bike Life Edinburgh 2017. [ONLINE] https://www.sustrans.org.uk/sites/default/files/file_content_type/bike-life-edinburgh-2017-report.pdf. [Accessed: 8 Jan 2018].

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, 2017. Cycling Action Plan for Scotland 2017-2020, Cycling as a form of transport. [ONLINE]

Transport for London, 2017. Cycling Network Model for London. [ONLINE]

Wegman, F.C.M., Zhang, F., and Dijkstra, A., 2012, "How to make more cycling good for road safety?" Accident Analysis and Prevention 44 (1), 19–29.

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. http://doi.org/10.1136/injuryprev-2016-041964.