Environmental Pollution 216 (2016) 746–754

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Methods to improve traffic flow and noise exposure estimation on minor roads☆

David W. Morley*, John Gulliver

MRC-PHE Centre for Environment & Health, Department of Epidemiology & Biostatistics, Faculty of Medicine, Imperial College London, W2 1PG, London, UK

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form
3 June 2016
Accepted 19 June 2016
Available online 24 June 2016

Keywords:
Noise exposure
CNOSSOS-EU
Road traffic
OpenStreetMap
Geospatial routing
Minor roads

ABSTRACT

Address-level estimates of exposure to road traffic noise for epidemiological studies are dependent on obtaining data on annual average daily traffic (AADT) flows that is both accurate and with good geographical coverage. National agencies often have reliable traffic count data for major roads, but for residential areas served by minor roads, especially at national scale, such information is often not available or incomplete. Here we present a method to predict AADT at the national scale for minor roads, using a routing algorithm within a geographical information system (GIS) to rank roads by importance based on simulated journeys through the road network. From a training set of known minor road AADT, routing importance is used to predict AADT on all UK minor roads in a regression model along with the road class, urban or rural location and AADT on the nearest major road. Validation with both independent traffic counts and noise measurements show that this method gives a considerable improvement in noise prediction capability when compared to models that do not give adequate consideration to minor road variability (Spearman's rho. increases from 0.46 to 0.72). This has significance for epidemiological cohort studies attempting to link noise exposure to adverse health outcomes.

1. Introduction

Exposure to environmental noise can have a number of adverse health effects (WHO, 2011). Noise associated with road traffic is a particular public health concern because it is pervasive and increasing. The link between road traffic noise and health outcome can be investigated via large-scale epidemiological cohort studies; for example, to cardiovascular morbidity as part of the TRAFFIC project (Halonen et al., 2015, 2016) and somatic symptoms in the BioSHaRE project (Zijlema et al., 2015). In order to conduct this type of research, an estimation of noise exposures for each subject must be made. This is often achieved within a geographic information system (GIS) framework by combining spatially referenced data on land cover and road geography with associated traffic flow as inputs to model the emission and propagation of noise from source (Steele, 2001; Murphy and King, 2010).

Following the publication of the European Directive on the Assessment and Management of Environmental Noise (2002/49/EC) (END), a noise model for Europe was proposed to aid standardised noise mapping: Common NOise aSSessment methOdS (CNOSSOS-EU) (Kephalopoulos et al., 2012, 2014). The CNOSSOS-EU methodology allows input data of varying detail. For accurate predictions at a fine spatial scale (for example, city-wide estimations), the highest possible resolution of input data can be feasibly obtained and then processed by the model. Conversely, for epidemiological studies the coverage of modelled predictions is often required at national or international scale. As a result, the high-resolution input data that is available at a localised scale may not exist or be affordable at an international scale. An assessment of the preliminary CNOSSOS-EU road traffic noise model showed that although it could be feasibly parameterised with readily available generalised land cover data, traffic flow information was still needed at the best possible resolution available for noise exposure estimates to be acceptable (Morley et al., 2015).

In terms of noise modelling for the purposes of the END, a comprehensive review of methodological issues with the END has been provided, including a comparison of numerous differences in the calculation methods between countries and how they result in differences in estimated noise levels (Murphy and King, 2010). For example, the UK method (CRTN) does not include attenuation for different frequency bands in the calculation of A-weighted noise, while the French method (NMPB) uses an octave-band approach.

☆ This paper has been recommended for acceptance by Eddy Y. Zeng.

* Corresponding author.
E-mail address: d.morley@imperial.ac.uk (D.W. Morley).

http://dx.doi.org/10.1016/j.envpol.2016.06.042
0269-7491/

Ascari et al. (2015) showed that for Pisa, Italy, the NMPB and Nord2000 methods resulted in estimate differences of up to 3dB, varying by noise frequency (Hz) and noise metric (A- or C-weighted), due to the differences in the method for calculating the contribution to noise levels of light and heavy vehicles. Murphy and King (2010) described how in a previous study (Van Leeuwen and Ouwerkerk, 1997) up to 7 dB difference was observed between five different noise calculation methods. Distributions of exposure estimates will thus depend on the choice of model as well as the study area and quality of input data.

A further issue common to all noise models that has received little attention is they only include traffic on main roads. Neglecting minor roads could lead to underestimation in noise levels especially as they contribute to values higher than 55 dB in many residential areas (Licitra et al., 2012). In the context of the END, noise levels are agglomerated from address-level estimates to calculate population-weighted noise exposures such as G_{den} and G_{night}, and these facilitate the definition of “action zones”. Only accounting for main roads will thus also lead to underestimation in these population-weighted noise metrics (Licitra and Ascari, 2014). Some studies have included minor roads but relied on a fixed value for AADT which may still lead to noise level underestimation, but also overestimation in areas with low traffic flows; for example, in a study (Ögren and Barregard, 2016) using the Nordic method to estimate road traffic noise exposure for the periods 1975–2010 in Gothenburg, Sweden, 250 AADT was used as a default value for minor roads; in our own work in London, UK, we used a higher value of 600 AADT (Gulliver et al., 2015) and reported the potential for variable levels of under- and over-estimation in noise levels, especially in residential areas.

For epidemiological studies, residential areas are important as a greater proportion of cohort members are likely to live there than along major roads. For example, 76% of London’s residential postcode centroids are over 50 m from a major road (Fecht et al., 2016). For this reason, a main aim of future work following our development of a noise exposure model for London was to improve noise estimation on minor roads (Gulliver et al., 2015).

Using estimates of annual average daily traffic (AADT), based on traffic counts on a limited number of roads, application of linear regression modelling can predict AADT traffic flows on road segments without counts (Lowry and Dixon, 2012). Explanatory variables may include number of lanes and road class, nearby population, access to employment or other socio-economic and land use variables (Zhao and Chung, 2001). Improvements to these models can potentially be made with a consideration of the connectivity of the road network. A road network can be conceptualised as an interconnected network graph of nodes (i.e. junctions) and edges (i.e. roads) from which the relative importance of road segments can be defined; it has been established that this relative importance can be directly related to traffic flow (Penn et al., 1998; Paul, 2011, 2012). A relative indication of how well a node is connected to other nodes in a graph translates directly to a road network in reality as these well connected roads would be expected to carry more traffic than roads leading to dead-ends. This quantified importance of each road segment can then be used as a further explanatory variable in AADT regression models (Lowry and Dixon, 2012). Despite the importance of road network connectivity on traffic flow being well established theoretically, there has been few attempts to use this directly for AADT estimation; most applications have centred on finding road hierarchies for automated map generalisation at different scale levels (Gülgen, 2014).

The aim of this study is to improve the coverage of roads with traffic flows on a national scale for application in estimating road traffic noise exposures in epidemiological studies. A nationwide set of AADT counts is used to formulate a relationship between the hierarchical importance of a road and expected AADT. Road importance is defined from simulated journeys using data and a routing algorithm often found in in-car satellite navigation devices. Both the estimated traffic flows and subsequent noise estimations are validated by comparing them to observed traffic counts and noise measurements obtained during the course of this study. We highlight the importance of minor roads in noise exposure modelling by revealing possible exposure misclassification if only traffic flows on major roads are considered in detail. This work has wider significance in terms of European-wide exposure studies and road traffic noise prediction in the context of the adoption of a standard modelling approach.

2. Methods

2.1. Road and traffic data

We developed three traffic flow data sets for minor roads, as shown in Table 1. Each of these data sets were based on traffic count data freely downloadable from the UK Department for Transport (DfT).

Here we use the most recent DfT data from 2013 where AADT is defined on major and minor roads and split into vehicle categories. The major road (motorways and A-roads) data set consists of a near complete coverage of point counts (n = 18,012) for all junction to junction links in the UK, while for minor roads (B-roads and below) only a small sample is available (n = 4462). AADT counts for specific roads are supplied at point locations; these have to be assigned to a spatial representation of the road network, which is done here with the freely available road geography from OpenStreetMap (OSM). The DfT defined major roads of motorway, principal, and trunk equate to one of the OSM classes of motorway, primary, and trunk. Similarly, the DfT minor classes of B, C, and unclassified correspond to one of the OSM classes of secondary, tertiary, residential, or unclassified. An unclassified road does not imply an unknown type, rather this is a DfT term for a road of importance lower than ‘B’ classification. Differences in nomenclature is not an issue in this study as DfT count points are allocated to an OSM road segment by spatial nearest neighbour assignment. For the rest of this study we refer to the road classification as defined by OSM. Due to the near completeness of the DfT major road counts, we used the same values of AADT for major roads for the three models in Table 1. For other European countries, major road data is often available on a national scale; for example, Eeftens et al. (2012) obtained data from 20 countries to develop air pollution models. Data on traffic flows on minor roads at national scale is, however, rarely available.

2.2. Assignment of AADT on minor roads

2.2.1. Method 1: fixed

Method 1 (Fixed) assumes a constant traffic flow as used in our earlier implementation of CNOSSOS-EU (Morley et al., 2015); all minor roads in the UK are assumed to have an AADT of 500. Methods 2 and 3 aim to improve traffic flow estimation on minor roads and subsequently the accuracy of noise exposure estimates.

2.2.2. Method 2: classified

There are clear contrasts in AADT values by urban or rural locations and by OSM minor road class (Fig. 1). According to Wilcoxon signed rank tests, there are significant differences between urban and rural AADT for secondary, tertiary and unclassified roads ($p = 0.00$) while there is no significant difference for residential roads ($p = 0.38$). This can be expected as for residential areas the urban-rural distinction is harder to define precisely. In addition, Kruskal–Wallis tests indicate that AADT on the four road classes is

Table 1
Summary of the three approaches used to classify AADT on minor roads for noise exposure assessment.

	Model	Minor road treatment
1	Fixed	All minor roads assigned AADT of 500.
2	Classified	Minor roads assigned median AADT according to DfT counts by sub-class (residential, unclassified, tertiary or secondary).
3	Routed	Minor roads assigned predicted AADT based on a regression model including road importance.

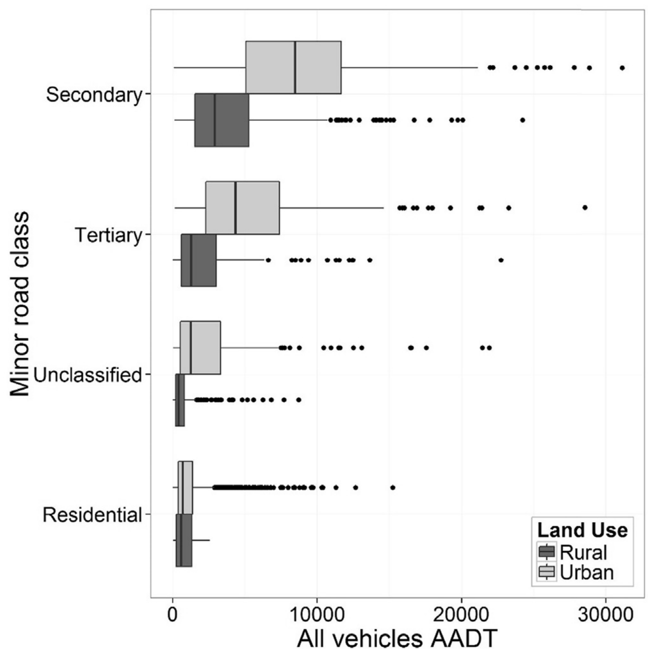

Fig. 1. Department for Transport (DfT) AADT 2013 counts on urban and rural minor roads by OSM road class.

significantly different when considered by either urban or rural situation ($p = 0.00$). As a result, method 2 (Classified) assigns variable traffic flows on minor roads using national median AADT values by minor road sub-class, dependent on the road segment being in an urban or rural area (Table 2).

2.2.3. *Method 3: routed*

As the basis for estimating the variability of traffic flow on minor roads it is possible to simulate probable journeys through the road network. Here, we use the open source database PostgreSQL with the GIS extensions of PostGIS and pgRouting. OSM road data can be routing-enabled within pgRouting and a feature of this is the availability of a cost attribute for each road segment. This cost is calculated based on road class, maximum speed limit, segment length and whether the road is one-way; the logic being that faster movement on main thoroughfares offers the most attractive routes to drivers and leads to higher traffic flows. Starting with every point where a minor road intersects the major road network (for which AADT has already been defined), the shortest driving route to every other minor road accessible without crossing another major road is calculated using the Dijkstra shortest path algorithm (Dijkstra, 1959). This is one of the most common routing algorithms used in GIS (Worboys and Duckam, 2004). Routes can be calculated using the full road network, so journeys are not restricted to the minor road network if partial travel on major roads is deemed more efficient by the routing algorithm. This is practical as we are attempting to model driver behaviour. In reality drivers would not necessarily limit themselves to minor roads only, but rather use the most effective combination of major and minor routes. A count of each time a minor road segment is traversed through these simulated journeys is kept and thus gives an indication of importance for traffic movement (Thomson and Richardson, 1995). To account for the fact it is possible to travel to more destinations from some initial starting points, these scores are standardised by dividing by the total number of accessible destinations accordingly.

This derived relative road importance indicator is unitless in terms of actual traffic flow. In order to relate this to real AADT values, it is subsequently used in the regression model with a training set of traffic counts on minor roads. Here we use the available sample of the DfT minor road dataset ($n = 4462$) as described in section 2.1. In addition to the routing importance of each minor road segment, several other predictor variables are considered: Due to the clear differences in AADT by OSM minor road type and by urban and rural locations (Fig. 1), these were included in the regression model as factors. Traffic count locations were classified as urban or rural using the settlement attribute of Ordnance Survey Meridian 2™. To account for known variability in traffic flow on main roads feeding the minor networks, the AADT on the nearest major road is included. In an extended model, further terms influencing traffic volumes such as local population density, vicinity to commercial areas or other points of interest could be included. As this study is a proof of concept of routing importance, this is the simplest representation of traffic volumes in the vicinity is used.

As traffic counts are observed to follow a Poisson distribution (Zhang et al., 2011), a generalised linear model (GLM) of Poisson family can be fitted to the data to predict AADT (Equation (1)).

$$\begin{aligned} \text{AADT} = {} & \log(\text{route importance}) + (\text{OSM Road Type}) \\ & + \log(\text{AADT on nearest major road}) \\ & + (\text{Urban or Rural}) \end{aligned} \tag{1}$$

Outliers were removed as those AADT values greater than 1.5 of the inter-quartile range for each road class split by urban or rural. As the routing importance and the traffic flow on the nearest major road were extremely skewed, these were log transformed prior to modelling.

Table 2
Department for Transport (DfT) 2013 median AADT (and interquartile range in brackets) for OSM road classes split by rural or urban location as used in the classified method of minor road assignment.

OSM road class	Urban	Rural
Secondary	8488 (6606)	2907 (3725)
Tertiary	4333 (5101)	1274 (2390)
Unclassified	1251 (2784)	409 (575)
Residential	710 (993)	588 (1059)

3. Model evaluation

3.1. *Noise modelling*

Predictions of noise exposure are made from an implementation of the CNOSSOS-EU road traffic noise model (Kephalopoulos et al.,

2012) using the three different traffic data sets described above in individual model runs, while all other inputs remain constant. The CNOSSOS-EU noise model is currently in the final phases of testing and validation (Kephalopoulos et al., 2014) and a series of software libraries exist to implement the algorithms. Here, our custom implementation follows the guidelines and empirical equations published in (Kephalopoulos et al., 2012). Due to data requirements, we have made several simplifications to the model (Morley et al., 2015).

Our version of the CNOSSOS-EU model takes account of the composition of traffic flow according to vehicle classes. For the fixed mode1, 10% of AADT is classified as heavy vehicles in accordance to other modelling methods where vehicle class composition data are not available (Department of Transport Welsh Office, 1988). In this study we varied the proportioning of AADT into specific vehicle categories for the classified and routed models depending on the corresponding average proportions of light and heavy goods vehicles, cars and motorbikes for each road class with increasing total AADT, as defined by the DfT minor road dataset. For Norwich, where we measured noise levels to evaluate noise predictions on minor roads in this study, average traffic flow is dominated by light vehicles (cars) making up over 95% of AADT on average for all classes of minor road. See Supporting Information for the details of the method to variably proportion AADT and further details of the noise model used in this study.

3.2. Model evaluation

A two-fold approach to evaluation was followed. Firstly, to demonstrate the effectiveness of the estimated road importance as an explanatory variable, a set of 81 15-min traffic counts were made in the town of Leamington Spa (UK). These counts focused on localised residential networks of minor roads and attempted to emphasise the variability within interconnected roads. Secondly, to verify the modelled AADT on a broader scale, 161 sites covering a full range of traffic conditions were chosen in the larger city of Norwich (UK). At these sites, 30-min traffic counts were taken along with co-located noise measurements. Noise measurements (L_{Aeq}) were made using a Cirrus Research Optimus sound level meter (model CR:171B). This was fitted with a wind-shield, located on a tripod at a height of 1.5 m and situated on the pavement as far back from the roadside as possible without entering private property. The locations of all sample sites were recorded with a handheld GPS. Each site was visited three times and all traffic counts were converted to AADT using temporal expansion factors according to those defined in National Roads Authority (2012). Using the average daily and monthly profiles provided by the DfT, expansion factors allow short (30 min) interval traffic counts to be converted to AADT using a reference to an hourly flow profile for an average day (here we use counts from 2013 provided by the DfT). Following this, these AADT values can be further standardised to adjust for weekly (weekends compared to weekdays) and seasonal variability (summer and winter). Full details of the field data collection are outlined in the Supporting Information.

How the inaccuracies in noise exposure estimation with regard to minor road AADT treatment translate to epidemiological cohort studies is demonstrated by estimating exposure for every building in the Norwich study area. The three AADT road datasets, summarised in Table 1, were used to make noise exposure estimates at receptor points taken to be 1 m from the road-facing façade; in total 106,280 noise exposure receptor points were identified. All statistical analysis was performed in R 3.2.2 (R Core Team and Core Team, 2015). Noise modelling was carried out in PostgreSQL/PostGIS.

4. Results and discussion

4.1. Local road importance

Simulation of journeys over the minor road network gives a standardised index of importance for each road segment in terms of expected traffic load. Fig. 2 shows the relative estimated road importance for Leamington Spa; better connected roads and trunk routes are clearly defined with higher indices (thicker lines), while dead-ends and poorly linked roads have lower significance. The relation of traffic counts made at the locations in Leamington Spa and predicted road importance is shown in Fig. 3. In terms of hierarchical ranking, observed counts of vehicles correlate well to road importance (Spearman's rho. = 0.78, R^2 = 0.68). The association between counted traffic and estimated road importance allows the development of a nationwide model of minor road AADT estimates.

4.2. Nationwide AADT prediction

Using the road importance index along with OSM defined road class, AADT on nearest major road and urban or rural situation, a Poisson GLM was fitted to the observed AADT counts on a sample of minor roads obtained from the DfT. All model terms were significant at the 99% level and of the total variance explained of 73.6%, 40.3% is explained by the routing importance alone. Adding the OSM road type increases this to 67.9% while the reaming 5.7% is explained by urban location and nearby major traffic flow.

4.3. Validation of AADT estimates

A comparison of modelled (minor roads) and count estimated AADT for Norwich survey points are shown in Fig. 4. Also shown are the AADT values obtained from the DfT major road dataset as assigned to the OSM road geography and compared to the counts. Counts are shown as the median with error bars representing the high and low of the three repeat visits. Both modelled AADT on minor roads and assigned major road AADT show good agreement to the traffic count and fall along the identity line. Major road rankings are well captured with Spearman's rho of 0.78, R^2 of 0.53, and RMSE of 5767. For both major and minor roads, the routed method moderately under-predicts the observed AADT; the gradient of the best fit lines shown in Fig. 4 are comparable at 0.57 and 0.45 for major and minor classes respectively. As both major and minor predictions rely on the DfT AADT data, the fact that the relative under-prediction is similar for both the minor roads and assigned major roads (see section 2.1), indicates that the routed modelling approach is capturing the true variability in the DfT data.

There are three clear outliers in the major road assignment which can be explained by incompleteness in the DfT dataset, with the nearest count point used for the road actually being beyond several major junctions on the same road, thereby not accounting for traffic leaving the road segment at these points. Removal of these outliers would improve the fit of the relationship (to R^2 of 0.80, RMSE of 3311), however Spearman's rho. remains at 0.78 indicating that the DfT dataset represents suitably the overall hierarchy of major roads by AADT. The routing model performs well in capturing the grading of minor roads in Norwich with a Spearman's rho of 0.84, R^2 of 0.72 and RMSE of 2774.

4.4. Validation of noise estimates

Results of the CNOSSOS-EU noise model for the 161 monitoring sites in Norwich compared to median measured hourly L_{Aeq} for the three road network representations are shown in Fig. 5, with

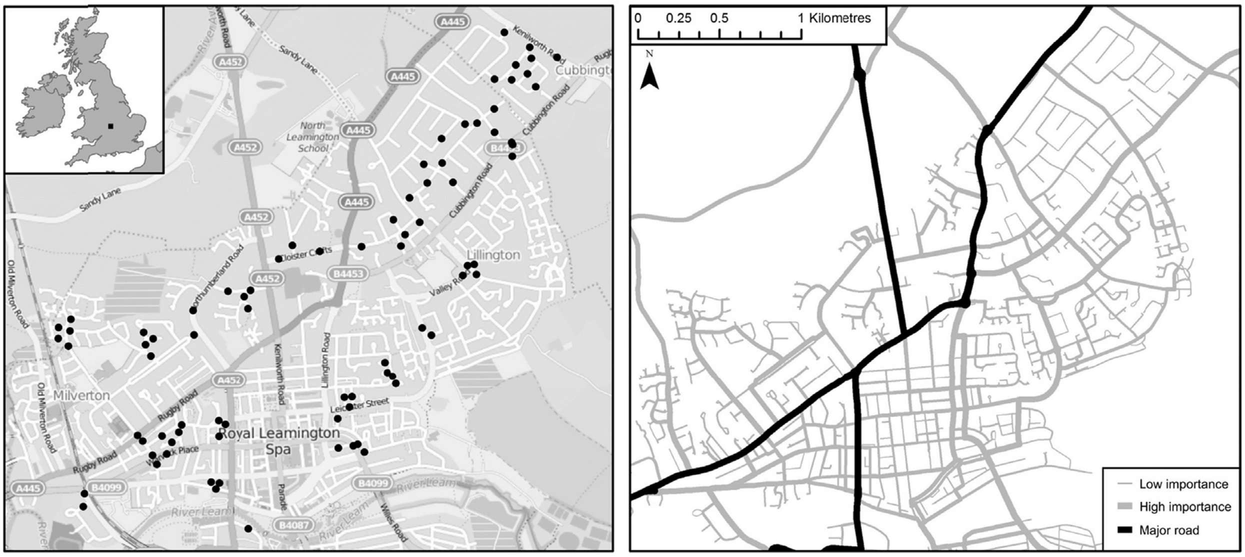

Fig. 2. Study area of Leamington Spa showing traffic count locations (black points; left) and corresponding estimated routing importance on minor roads (right) with thicker grey lines indicating more important roads. Major roads are shown for reference in black and are not symbolised by importance.

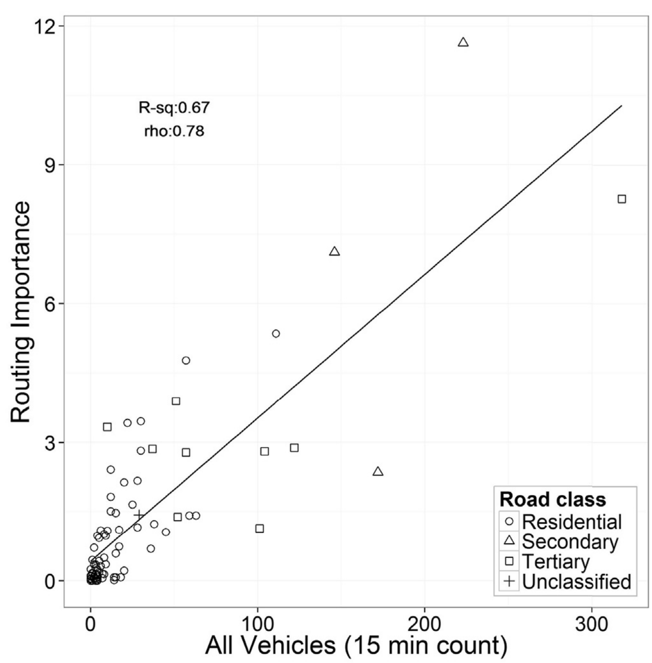

Fig. 3. Relationship of routing estimated minor road importance (by OSM road class) and 15-min traffic counts in Leamington Spa with linear best fit line.

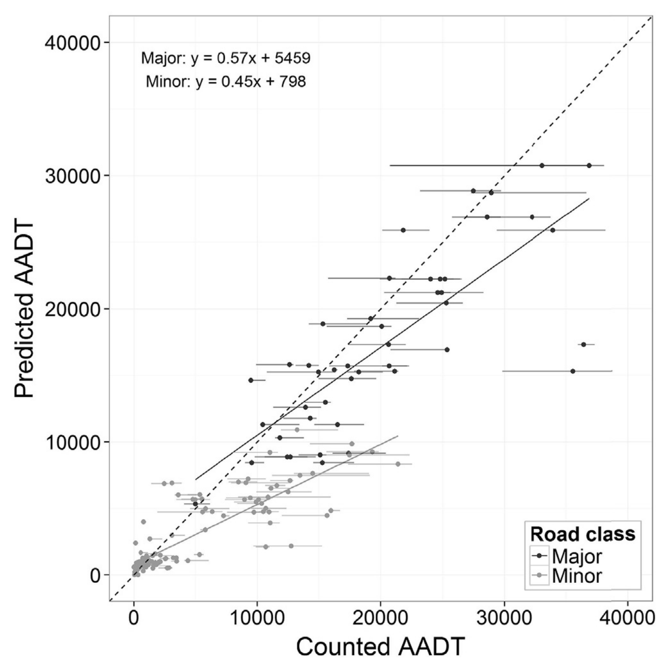

Fig. 4. Predicted AADT at the 161 count sites in Norwich using the minor roads routing model compared to observed AADT estimated from 30-min traffic counts. The median observed AADT is shown of three visits together with the minimum and maximum error bars. Sites are categorised as either being situated on a major or minor road. The dashed line represent a one-one relationship between measured and modelled values. Solid lines represent the best fit for the relationship between counted and predicted AADT split by road class, the equations of these lines are also quoted.

correlation statistics in Table 3. When looking at major roads only, model performance is almost identical. This is expected as all three models use the same major roads data, and any minor variability is due to different AADT assignment on nearby minor roads which also contribute to the modelled noise. However, it is clear that attributing all minor roads with a constant AADT value is inadequate, as the fixed model demonstrates a lack of variability in predicted L_{Aeq} and fails to predict well for noisier sites (~>62 dBA) (rho = 0.46, RMSE = 6.28). The classified model improves AADT predictions greatly by more accurately dealing with the more trafficked secondary and tertiary roads (rho = 0.63, RMSE = 5.01), and considering routing importance gives even greater refinement in predictions (rho = 0.72, RMSE = 4.80). This RMSE is valid in the context of the inherent error of our parameterisation of the CNOSSOS-EU methodology (Morley et al., 2015) to predict noise levels. This can be quantified using the variability of the repeated field measurements taken in Norwich. On minor roads, the average difference between the minimum and maximum of the three repeats at each site is 3.5 dBA (σ = 2.3 dBA). There is a difference of 1.79 (RMSE) between the error for major roads (3.01) and minor

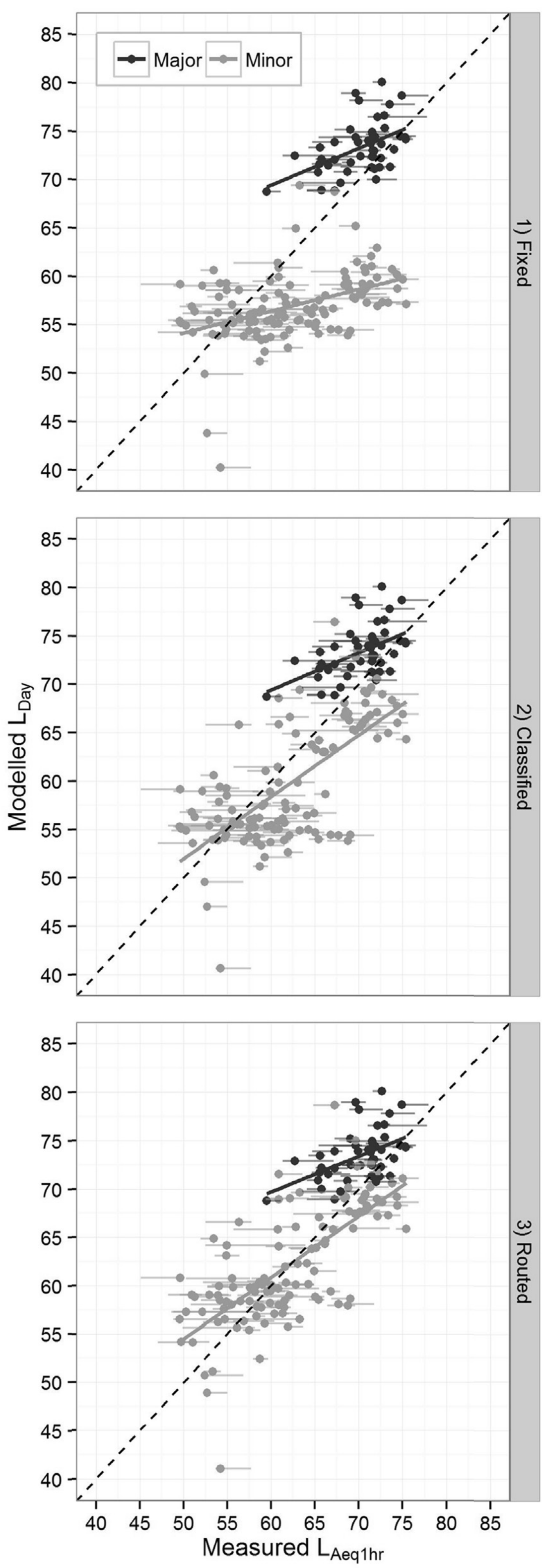

Fig. 5. Norwich measured hourly L_{Aeq} (median, minimum and maximum of three repeat visits) and modelled L_{day} (07:00 to 19:00). Sites are split by minor and major road locations for the three different approaches of minor road AADT assignment. The dashed lines represent a one-one relationship between measured and modelled values.

roads (4.80) (Table 3). In addition to the possible variable error from our routed method, an unknown proportion of the noise error is also due to deficiencies in the traffic flow data (e.g. errors in traffic flows on roads surrounding the measurement sites, variable speeds along road links for which we only able to use fixed values). Although in designing the field measurements we tried to avoid locations with other sources of noise there could be situations where they contribute to model error by confounding the comparison of measured (an integral of all present noise sources) and modelled (noise due to traffic, only) noise values (e.g. aircraft, birdsong, construction, human voices, railways).

4.5. *Sources of prediction error*

Addition of the routing importance leads to an improved discrimination of sites in the mid-range of measured noise, although in some cases the noise model is still making large prediction errors. A possible reason for this is inconsistencies in how OSM editors distinguish road classes; for the UK there are no strict OSM guidelines and the choice can be subjective. Alternatively, a road may have changed function and not been updated yet or simply just classified in error within OSM. This is illustrated by a site with a recorded noise level in $L_{Aeq,1hr}$ of 54.2 dBA but an under-prediction at 40.9 dBA. According to OSM, this site is located on a pedestrian only street and is assigned a traffic flow of zero by the model. In reality, this street is open to residential traffic (observed AADT of 343 in the field) which accounts for the measured noise level. An additional problem is shown by a site with an under-prediction of 10 dBA; this is located on an unclassified road and has an observed AADT of 5346 while the model predicts AADT of 1548. Being greater than the median AADT of 1200 for an unclassified urban road, this would be a feasible estimate and describes the road section to be busier than the average. In reality, however, this road segment could also accurately be considered as tertiary based on its characteristics (wide, non-residential with clear road markings) and, accordingly, the median AADT for an urban tertiary road is 4219, which is better agreement to the observed traffic flow. The contrary is also true for several sites where the noise model over-predicts. Here roads classed as tertiary by OSM could be subjectively classed as residential or unclassified and therefore receive a lower model assigned AADT estimate.

The OSM route-enabled network also does not account for ‘service roads’ (e.g. private roads). Traffic routing on such roads therefore could not be considered but may be significant around Norwich Airport, industrial estates on the outskirts of the city or large private complexes such the University of East Anglia campus. The model also does not directly account for the higher traffic volumes around such points of interest. These cases are of less significance here as the focus of this work is to make exposure estimates at residential locations only. Another possible cause of prediction error is the inability of the model to account for driver behaviour. Some monitoring sites were located in areas popularly used as ‘cut-throughs’ or ‘rat-runs’ between major roads or busy areas by drivers. As a result, modelled traffic flow on these segments were lower than the observed. Anecdotal evidence from discussions with residents complaining about high traffic volumes in their area combined with the observed higher than expected manual traffic counts support presence of this characteristic.

4.6. *Population noise-level exposure*

Distributions of noise exposure estimates for each model at every building façade in Norwich are shown in Fig. 6 in 1 dBA classes, percentage classified into 5 dBA bands in Table 4, and as a density plot in Fig. S3. Maps of residential noise exposures are also shown in Supporting Information Fig. S2. For the Fixed and

Table 3
RMSE, Pearson's product moment correlation (PPMC), Spearman's rho. and R^2 of Median measured hourly L_{Aeq} in Norwich and modelled L_{day} (07:00 to 19:00) (as shown in Fig. 5) for the three minor road AADT models. Statistics are presented for all count sites and separated by major and minor road class.

	All				Major				Minor			
	RMSE	PPMC	rho	R^2	RMSE	PPMC	rho	R^2	RMSE	PPMC	rho	R^2
Fixed	5.59	0.62	0.67	0.38	3.01	0.49	0.47	0.24	6.28	0.42	0.46	0.18
Classified	4.55	0.77	0.75	0.59	2.99	0.50	0.48	0.25	5.01	0.69	0.63	0.48
Routed	4.36	0.79	0.78	0.62	3.01	0.48	0.48	0.23	4.80	0.72	0.72	0.52

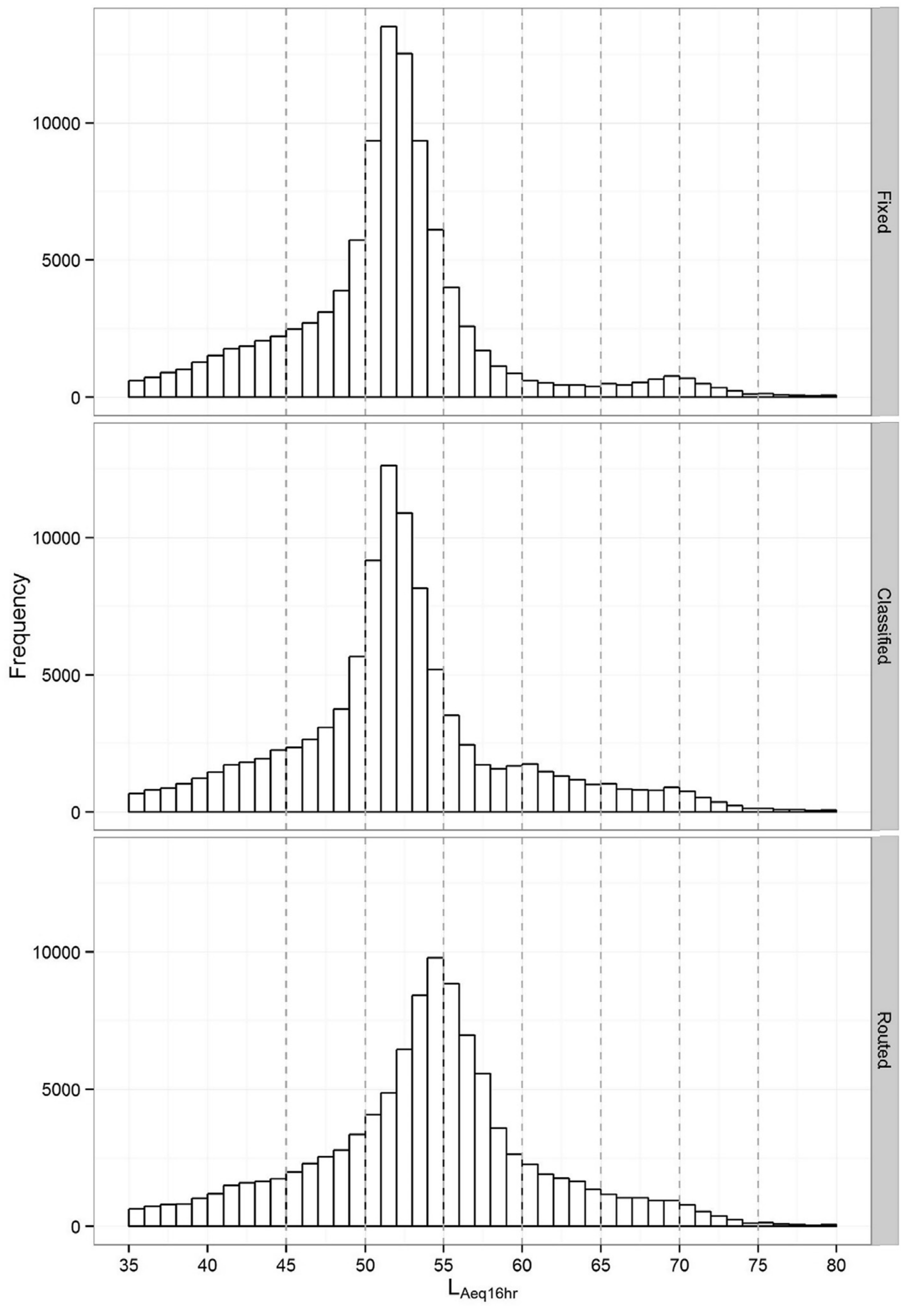

Fig. 6. Comparison of modelled $L_{Aeq,16hr}$ for Norwich building facades using the three different traffic flow approximations (fixed, classified and routed) binned into 1 dBA classes. Dashed vertical lines represent 5 dBA intervals.

classified models, the mode of estimated noise levels is 52 dBA (Fig. 6). The routed model has a mode of 54 dBA and an increase in the number of higher noise level estimates. The metric used here is $L_{Aeq,16hr}$ (average dBA between 07:00 and 23:00) which is commonly used in epidemiological studies (van Kempen and Babisch, 2012; Halonen et al., 2016; Tonne et al., 2016). With no

Table 4
Percentage of Norwich area residential locations (n = 106,280) in 5 dBA noise ($L_{Aeq,16hr}$) classes for the three models of minor road classification.

dBA	<45	45–<50	50–<55	55–<60	60–<65	65–<70	70–<75	≥75
Minor fixed	17.6	16.8	47.8	9.6	2.3	2.8	1.8	1.3
Minor classified	16.5	16.4	43.3	10.3	6.3	4.0	1.9	1.3
Minor routed	13.6	12.2	31.6	26.0	8.4	4.9	1.9	1.4

special consideration of minor roads, the noisier areas are clustered along main roads while residential areas experience relatively constant low noise levels. In the Classified model, buildings along secondary roads are now classified with higher noise exposure estimates and an increased number of predictions in the range 55–70 dBA. In the Routed model, there is a general shift to higher noise estimates. The greatest density of noise estimates for the Routed model is increased by 2.5 dBA, accordingly there is an increase of 16% of residential locations in the range 55–60 dBA, and also an increase of 8% of residential locations in the range 60–70 dBA. These increases corresponds to 17,430 and 8715 residential locations, respectively, in Norwich. In all three minor road treatments, estimates greater than 70 dBA are consistent as the major road traffic flows are the same in each case. The real effect of a more considered approach to minor roads is highlighted by this example of a population in Norwich. Although the validation with the 161 observed noise levels shows a modest increase in model fit between the Classified and Routed methods, when applied to a cohort study, the change in the distribution of exposure values (Fig. 6) becomes significant.

4.7. Strengths and limitations

In this study we have used a limited set of traffic flow counts and demonstrated two methods (Classified and Routed) to estimate AADT for every road segment in the UK. Particularly on a national scale, the Routed approach is a substantial improvement on noise exposure classifications where no consideration is given to variability in the minor road network due to the lack of a suitable dataset. Using a small set of known AADT counts and index of connectivity in the road network, a regression model (Equation (1)) can allow this to be expanded to make nationwide estimates. A major strength of this study is the large number of independent and repeated traffic counts with co-located noise measurements used to validate the models. Traffic flow datasets created in this way fit with the rationale of the CNOSSOS-EU method in that standardised data can be used to make comparable noise exposure assessments with national and international scope. The methods developed here for estimating traffic flows on minor roads also have potential importance to studies developing models of air pollution exposures such as land use regression (LUR) (Eeftens et al., 2012). Variables on minor road length as a contributory predictor of the variability in pollutant concentrations might be better served by a variable describing both road length and traffic flows (i.e. estimated using the Classified or Routed approaches outline here) on minor roads. Additional major strengths of this study are the national scale of the data used to develop methods and that we used independent data (i.e. hold-out-validation) to evaluate both the performance of the classified and routed methods for predicting AADT and noise estimates.

Our study does however have a number of limitations The main limiting factor of this method is the small size of the national DfT dataset which accounts for less than 0.17% of all minor roads in the UK. With access to a much larger training set, predictions in unsampled areas would be greatly improved. In addition, as OSM is in a continual community-driven state of development, further improvements to the road network classification including better representations of roads types and one-way sections will allow better predictions to be made. In this approach, we considered traffic flow on the nearest major road as an explanatory variable, but the model relied mostly on road connectivity. The main disadvantage of this being that the simulated journeys through the network here are not necessarily rational in that the goal of the trips were not considered. Including a weighting to journeys that could be expected based on their start and end location would give a more real estimation of network importance. For example, routes between areas of high population density or between commercial zones will be more common than routes starting and ending at points with no features of interest. Future work will focus on including these other predictors in the regression model. Our evaluation of the methods found a difference of 1.79 in RMSE between the noise predictions on major (3.01) and minor roads (4.80). In addition to information provided in section 4.4 on validation of the noise estimates, we also recognise that the performance of the models may be limited by the use of relatively short-term traffic counts and noise measurements, as highlighted by the variability between the repeated traffic counts shown in Fig. 4. Nonetheless, we have developed methods to produce a national scale dataset of traffic flows on minor roads which has potential for use in other countries in both noise modelling and air pollution models (i.e. improved estimation of the contribution of air pollution emissions from background roads). In order to facilitate wider acceptance of the methodology we aim to undertake further data collection in other locations, undertaking a greater number of repeat measurements to improve the robustness of the comparisons, to calibrate the models and expand to information on model validation.

In summary, using the results of this study we would expect an average noise exposure error, from our implementation of CNOSSOS-EU, of ~3.0 dB close to major roads and ~4.8 dB close to minor roads (Fig. 5). The method presented here on improving traffic flow estimation on minor roads has for people living on minor roads both reduced the error from ~6.3 dB to ~4.8 dB and improved the Spearman's correlation from 0.46 to 0.72. The increase in correlation especially is a substantial improvement in an epidemiological context where relative ranking of exposures is important. Using the predicted traffic flows on minor roads also leads to significant changes in absolute values of noise with a tendency for noise levels on minor roads to be higher (Fig. 6). We developed a traffic estimation method for minor roads that others can adopt.

Acknowledgements

This work was funded by the European FP7 EXPOsOMICS project; Grant number: 308610. All OpenStreetMap data are licensed under the Open Database License 1.0 and © OpenStreetMap contributors (see www.openstreetmap.org/copyright for full details). OSM data for the UK were downloaded from Geofabrik (www.geofabrik.de). Author contributions: DM and JG designed the research and conducted fieldwork, DM performed the analysis.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envpol.2016.06.042.

References

Ascari, E., Licitra, G., Teti, L., Cerchiai, M., 2015. Low frequency noise impact from road traffic according to different noise prediction methods. Sci. Total Environ. 505, 658–669.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/.

Department for Transport (DfT) Traffic Counts; http://www.dft.gov.uk/traffic-counts/download.php (accessed January 2016).

Department of Transport Welsh Office, 1988. Calculation of Road Traffic Noise. HMSO, London. U.K.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271.

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dedele, A., Eriksen, K., Galassi, C., Grazuleviciene, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Kraemer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Moelter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskoevi, E., Weinmayr, G., Brunekreef, B., Hoek, G., 2012. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2-Results of the ESCAPE project. Atmos. Environ. 62, 303–317.

Fecht, D., Hansell, A., Morley, D., Dajnak, D., Vienneau, D., Beevers, S., Toledano, M.B., Kelly, F.J., Anderson, R.H., Gulliver, J., 2016. Spatial and temporal associations of road traffic noise and air pollution in London: implications for epidemiological studies. Environ. Int. 88, 235–242.

Gülgen, F., 2014. Road hierarchy with integration of attributes using fuzzy-AHP. Geocarto Int. 29 (6), 688–708.

Gulliver, J., Morley, D., Vienneau, D., Fabbri, F., Bell, M., Goodman, P., Beevers, S., Dajnak, D., Kelly, F.J., Fecht, D., 2015. Development of an open-source road traffic noise model for exposure assessment. Environ. Model. Softw. 74, 183–193.

Halonen, J.I., Hansell, A.L., Gulliver, J., Morley, D., Blangiardo, M., Fecht, D., Toledano, M.B., Beevers, S.D., Anderson, H.R., Kelly, F.J., Tonne, C., 2015. Road traffic noise is associated with increased cardiovascular morbidity and mortality and all-cause mortality in London. Eur. Heart J. 36, 2653–2661.

Halonen, J.I., Blangiardo, M., Toledano, M.B., Fecht, D., Gulliver, J., Ghosh, R., Anderson, H.R., Beevers, S.D., Dajnak, D., Kelly, F.J., Wilkinson, P., Tonne, C., 2016. Is long-term exposure to traffic pollution associated with mortality? A small-area study in London. Environ. Pollut. 208 (A), 25–32.

Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F., 2012. Common Noise Assessment Methods in Europe (CNOSSOS-EU) EUR 25379 EN. Publications Office of the European Union, Luxembourg.

Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F., van Maercke, D., Shilton, S., Jones, N., 2014. Advances in the development of common noise assessment methods in Europe: the CNOSSOS-EU framework for strategic environmental noise mapping. Sci. Total Environ. 482–483, 400–410.

Licitra, G., Ascari, E., 2014. Gden: an indicator for European noise maps comparison and to support action plans. Sci. Total Environ. 482–483, 411–419.

Licitra, G., Ascari, E., Brambilla, G., 2012. Comparative analysis of methods to estimate urban noise exposure of inhabitants. Acta Ascut United Ac 98, 659–666.

Lowry, M., Dixon, M., 2012. GIS Tools to Estimate Average Annual Daily Traffic. National Institute for Advanced Transportation Technology: University of Idaho. Final Report KLK725 N12-03.

Morley, D.W., de Hoogh, K., Fecht, D., Fabbri, F., Bell, M., Goodman, P.S., Elliott, P., Hodgson, S., Hansell, A., Gulliver, J., 2015. International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. Environ. Pollut. 206, 332–341.

Murphy, E., King, E.A., 2010. Strategic environmental noise mapping: methodological issues concerning the implementation of the EU Environmental Noise Directive and their policy implications. Environ. Int. 36, 290–298.

National Roads Authority, 2012. Project Appraisal Guidelines. Unit 16.2 Expansion Factors for Short Period Traffic Counts. National Roads Authority, Dublin.

Ögren, M., Barregard, L., 2016. Road traffic noise exposure in Gothenburg 1975-2010. PLoS One 11 (5), e0155328.

OpenStreetMap (OSM); http://www.openstreetmap.org/.

Ordnance Survey Meridian 2; https://www.ordnancesurvey.co.uk/opendatadownload/products.html.

Paul, A., 2011. Axial analysis: a syntactic approach to movement network modeling. Institute of town planners. India J. 8 (1), 29–40.

Paul, A., 2012. Unit-segment analysis: a space syntax approach to capturing vehicular travel behavior emulating configurational properties of roadway structures. Eur. J. Transp. Infrastructure Res. 12 (3), 275–290.

Penn, A., Hillier, W.R.G., Banister, D., Xu, J., 1998. Configurational modeling of urban movement networks. Environ. Plan. B Plan. Des. 25, 59–84.

pgRouting; http://pgrouting.org/.

PostGIS; http://postgis.net/.

PostgreSQL; http://www.postgresql.org/.

Steele, C., 2001. A critical review of some traffic noise prediction models. Appl. Acoust. 62, 271–287.

Thomson, R.C., Richardson, D.E., 1995. A graph theory approach to road network generalization. In: In Proc. 17th Int. Cartographic Conference, pp. 1871–1880.

Tonne, C., Halonen, J.I., Beevers, S.D., Dajnak, D., Gulliver, J., Kelly, F.J., Wilkinson, P., Anderson, H.R., 2016. Long-term traffic air and noise pollution in relation to mortality and hospital readmission among myocardial infarction survivors. Int. J. Hyg. Environ. Health 219, 72–78.

van Kempen, E., Babisch, W., 2012. The quantitative relationship between road traffic noise and hypertension. J. Hypertens. 30, 1075–1086.

Van Leeuwen, H., Ouwerkerk, M.A., 1997. Comparison of Some Prediction Models for Railway Noise Used in Europe. A The Hague DGMR. Report L.94.0387.

WHO, 2011. Burden of Disease from Environmental Noise: Quantification of Healthy Life Years Lost in Europe. WHO, Copenhagen.

Worboys, M.F., Duckam, M., 2004. GIS: a Computing Perspective. CRC Press, Boca Raton.

Zhang, Z., Mao, G., Anderson, B.D.O., 2011. On the information propagation process in mobile vehicular Ad Hoc networks. IEEE Trans. Veh. Technol. 60 (5), 708–712.

Zhao, F., Chung, S., 2001. Contributing factors of annual average daily Traffic in A Florida county - exploration with geographic information system and regression models. Pavement Manag. Monit. Accel. Test. 1769, 113–122.

Zijlema, W., Morley, D.W., Stolk, R.P., Rosmalen, J.G., 2015. Noise and somatic symptoms: a role for personality traits? Int. J. Environ. Health 218 (6), 543–549.