Improving traffic-related air pollution estimates by modelling minor road traffic volumes☆

Miguel Alvarado-Molina [a,*], Ariadna Curto [a,h,i,j], Amanda J. Wheeler [b,c], Rachel Tham [d], Ester Cerin [a,e,f,g], Mark Nieuwenhuijsen [a,h,i,j], Roel Vermeulen [k,l], David Donaire-Gonzalez [l]

[a] Mary Mackillop Institute for Health Research, Australian Catholic University, Melbourne, Australia
[b] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Environment, Melbourne, VIC 3195, Australia
[c] Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS 7000, Australia
[d] Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC 3053, Australia
[e] School of Public Health, The University of Hong Kong, 7 Sassoon Rd., Sandy Bay, Hong Kong
[f] Baker Heart and Diabetes Institute, Melbourne, VIC, Australia
[g] Department of Community Medicine, UiT The Artic University of Norway, Tromsø, Norway
[h] ISGlobal, Barcelona, Spain
[i] Department of Experimental and Health Sciences, Universitat Pompeu (UPF), Barcelona, Spain
[j] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[k] Institute for Risk, Assessment Sciences (IRAS), Division of Environmental Epidemiology (EEPI), Utrecht University, Utrecht, the Netherlands
[l] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

ABSTRACT

Accurately estimating annual average daily traffic (AADT) on minor roads is essential for assessing traffic-related air pollution (TRAP) exposure, particularly in areas where most people live. Our study assessed the direct and indirect external validity of three methods used to estimate AADT on minor roads in Melbourne, Australia.

We estimated the minor road AADT using a fixed-value approach (assuming 600 vehicles/day) and linear and negative binomial (NB) models. The models were generated using road type, road importance index, AADT and distance of the nearest major road, population density, workplace density, and weighted road density. External measurements of traffic counts, as well as black carbon (BC) and ultrafine particles (UFP), were conducted at 201 sites for direct and indirect validation, respectively. Statistical tests included Akaike information criterion (AIC) to compare models' performance, the concordance correlation coefficient (CCC) for direct validation, and Spearman's correlation coefficient for indirect validation. Results show that 88.5% of the roads in Melbourne are minor, yet only 18.9% have AADT. The performance assessment of minor road models indicated comparable performance for both models (AIC of 1,023,686 vs. 1,058,502). In the direct validation with external traffic measurements, there was no difference between the three methods for overall minor roads. However, for minor roads within residential areas, CCC (95% confidence interval [CI]) values were −0.001 (−0.17; 0.18), 0.47 (0.32; 0.60), and 0.29 (0.18; 0.39) for the fixed-value approach, the linear model, and the NB model, respectively. In the indirect validation, we found differences only on UFP where the Spearman's correlation (95% CI) for both models and fixed-value approach were 0.50 (0.37; 0.62) and 0.34 (0.19; 0.48), respectively. In conclusion, our linear model outperformed the fixed-value approach when compared against traffic and TRAP measurements. The methodology followed in this study is relevant to locations with incomplete minor road AADT data.

1. Introduction

Urban air pollution poses a significant public health concern worldwide (Cohen et al., 2017). Recent studies reveal that adverse health effects persist even at low levels of air pollution (Strak et al., 2021). Traffic is a major contributor to urban air pollution (Künzli et al., 2000). Black carbon (BC) and ultrafine particles (UFP) are strong indicators of traffic-related air pollution (TRAP) and can exhibit

substantial variations over small geographic areas, penetrate deep into the respiratory system and even enter the bloodstream (Saha et al., 2019; Van den Hove et al., 2020; van Nunen et al., 2017). Geographical information has been instrumental in advancing our understanding of the health effects of air pollution (Shen et al., 2022; de Hoogh et al., 2018; Eeftens et al., 2012). However, a significant challenge in accurately predicting the difference between low levels of air pollution is the lack of traffic volume data on local roads (hereon called minor roads). Minor roads, along which most urban populations live (e.g. London with 90%), (Morley and Gulliver, 2016; Aldred and Verlinghieri, 2020) contribute significantly to human air pollution exposure, yet their traffic volume is largely unknown (Apronti et al., 2016; Fu et al., 2017). Improving our understanding of minor road traffic volume is crucial to enhancing the accuracy of air pollution exposure assessments to improve public health outcomes.

Five papers were identified from the revised literature regarding traffic volume on minor roads (Table 1), all of which focused on AADT modelling (Morley and Gulliver, 2016; Apronti et al., 2016; Fu et al., 2017; Zhong and Hanson, 2009; Jung et al., 2017). These studies exhibited variations in their geographic settings, encompassing rural, urban, or a combination of both settings. They also differed in the number of measurement points assessed (ranging from 42 to 4462). Additionally, previous studies employed diverse modelling approaches, validation protocols, and performance assessment techniques. Among these five articles focusing on AADT modelling, two did not conduct dedicated analyses for secondary roads (Fu et al., 2017; Jung et al., 2017). Among the remaining three, two conducted only very limited external validations, involving just ten measurements on secondary roads or were conducted in highly rural environments, where 60% of the roads were unpaved (Apronti et al., 2016; Zhong and Hanson, 2009). Notably, none of these studies quantified the impact of modelling AADTs for secondary roads on their correlation with traffic-related air pollutants.

Our case study was developed in Melbourne, Australia using BC, UFP, and traffic count measurements (201 test sites). Melbourne is characterised by an extensive urban sprawl, contrasting with other cities with compact urban designs. We conducted a comprehensive approach to modelling AADT that considers multiple socio-economic and physical attributes that influence AADT variability and BC/UFP levels. The models used open-data sources to be easily reproduced in other geographical locations. The aims were: 1) to compare existing methodologies for modelling minor road AADT primarily using open source and reproducible datasets; 2) to validate externally and directly AADT estimates with real-world traffic count measurements, and indirectly with BC, and UFP; and 3) to quantify the added value of modelling AADT compared to the traditional fixed-value approach. By addressing these aims, our study bridges a crucial gap in air pollution exposure modelling, signifying a novel and valuable advancement in the field globally.

2. Methodology

2.1. Study design

Using the city of Melbourne as an Australian case study, referred to as Melbourne hereon, we generated AADT models for minor roads. We compared the modelled estimates of minor road AADT against real-world measurements of BC and UFP levels and traffic counts collected at 201 sites in Melbourne as detailed in subsection 2.4.2.

2.2. Study area

Melbourne is located in south-eastern Australia, has a temperate oceanic climate (Clarke et al., 2019), is 2705.4 km^2 in size (refer to Appendix B) with a population of 4.60 million in 2021 (4.20 million in 2016) (ABS, 2016a; ABS, 2021). Melbourne's road network infrastructure has a total length of 24,227 km. In 2019, the road network carried

Table 1
Comparison of previous approaches to model minor road AADT.

Author	Zhong and Hanson	Apronti et al.	Morley and Gulliver	Jung et al.	Fu et al.
Year	2009	2016	2016	2017	2017
Study Area	York County and Beresford CCS in New Brunswick, Canada	Wyoming (23 counties), USA	UK	Incheon City, South Korea	Dublin, Ireland
Sampled points	42 measurement points	476 measurement points (372 points modelling)	4462 measurement points (minor roads)	296 unit sections*	96 measurement points
Study setting	Urban/Rural	Rural	Urban/Rural	Urban	Urban
Road type of interest	Low-class roads/low-volume roads (local roads)	Low volume roads (<400 AADT)	Secondary, tertiary, residential and unclassified roads (OSM road network)	Express ways, urban express ways, major and minor arterials	Low-class roads/low-volume roads (local, secondary, tertiary, and residential roads)
Modelling approach	Travel Demand Model	Linear and logistic regression	Poisson regression	Regression Kriging	Log-linear regression, ordinary least squares regression, and Artificial Neural Network
Dependent variable	Traffic volume	AADT	AADT	VKT and AADT	AADT
Modelling performance	NA	$R^2 = 0.64$ Pearson corr = 0.69 RMSE = 73.4%	UK: 73.6% variance explained	$R^2 = 0.60$	Pearson corr = 0.83–0.93 MAPE = 28.58%
Validation	External validation Traffic count points: *Highway = 12* *Collector = 18* *Local = 12*	External validation *9 counties (104 points)*	External validation *161 sites*	Internal validation 108 ATR *locations and* *292 SPTC locations*	NA
Performance of external validation	Average error: Highway = 9% Collector = 44% Local = 174%	Scatter plot and Pearson correlation = 0.68, 0.61	Major road: Spearman's rho = 0.78; $R^2 = 0.53$; RMSE = 5767 Minor roads: Spearman's rho = 0.84; $R^2 = 0.72$; RMSE = 2774	Average error = 39.05% MAE = 8272.21 MAPE = 39.05	NA

AADT = Annual average daily traffic volume; CCS = Census consolidated subdivision; OSM = OpenStreetMap; VKT = vehicle kilometres travelled; R^2 = Correlation coefficient; RMSE = Root means squared error; MAE = Mean average error; MAPE = Mean absolute percentage error; ATR = Automatic traffic recorders; SPTC = Short periodic traffic counts; * Basic unit for analysis over target road; NA = Not Applicable; corr = correlation.

approximately 15.3 million daily trips, of which 11 million (72%) were private vehicle trips. (VicGov, 2020).

2.3. Measurements and databases

2.3.1. Predictors for minor road AADT regression models

We used existing information on Melbourne's road traffic volume, road network and its characteristics, and sociodemographic characteristics of neighbourhoods (or areas) to model minor road AADT (see Appendix B).

We characterised each minor road according to seven attributes: road type, place of work density, population density, distance to the nearest major road and its AADT, road importance index, and weighted road density.

AADT for Melbourne roads was obtained from Veitch Lister Consulting's 2018 Zenith model (Veitch Lister Consulting, 2020). The Zenith database provided information for all major roads and 18.2% of minor roads. The road network and its characteristics, such as road type or the presence of tunnels and bridges along a road, were downloaded from OpenStreetMap (OSM) for 2018–2020. The 'road type' corresponded to the OSM road classification. For our models, we used seven out of the 14 OSM road types: motorway, trunk, primary, secondary, tertiary, residential and unclassified (Fig. 1). All road links were aggregated to their corresponding main road type to simplify the final model (i.e., secondary links were categorised as secondary road type). Following a road equivalence provided by the UK Department of Transport (Dft), we grouped the motorways, trunks, and primary roads as 'major roads', and secondary, tertiary, residential, and unclassified roads as 'minor roads'. (Morley and Gulliver, 2016).

The AADT was assigned to Melbourne's OSM road network using stepwise sausage buffers of 1.5 m, 5 m, 15 m, and 30 m along the road network. With this procedure, road AADT was assigned to roads within each buffer size (considering the road's geometry and type), starting at 1.5 m. If the road was successfully assigned an AADT, it was removed from the road network. The process was repeated with the remaining roads progressively with all buffer sizes until we reached the 30 m buffer. Then, we converted the OSM data into a routable network using osm2po software (osm2po-core, Pinneberg, Schleswig-Holstein)

(Moeller).

Sociodemographic characteristics, such as population and workplace densities, were obtained from the Australian Bureau of Statistics (ABS) 2016 Census of Population and Housing data. We used the statistical area 1 (SA1) boundaries which generally correspond to areas with 200–800 persons and an average of 400 persons. (ABS, 2016b) The roads were assigned the density of the SA1 they were located in or the average of all SA1s they crossed if they traversed more than one SA1.

In addition, we calculated a road importance index within the network and the surrounding road density weighted by road type, as described below and elsewhere (Morley and Gulliver, 2016; Rose et al., 2009). The road importance index is defined as the total number of times a road is traversed when connecting all road segments within an area of interest, divided by the total number of road segments that form that area. For this purpose, we used the road routing algorithm V.2.2 from PostGIS V.3 (Open Geospatial Consortium, Wayland, MA). (Mikiewicz et al., 2017; pgRouting; OGC).

The surrounding road density was weighted by road type to obtain the weighted road density (WRD), as named by the authors (Rose et al., 2009). This method ranks the roads within a 50 m radius with values from '1' to '3', according to their road type, and multiplies the length of the roads by that value. Finally, the sum is divided by the area. We used the ranking '3' for primary roads, highways, and trunk roads, '2' for secondary and tertiary roads, and '1' for unclassified and residential roads (Rose et al., 2009).

2.3.2. Validation study

We directly validated our AADT estimates by measuring traffic counts and used BC and UFP particle concentration measurements for indirect validation. These real-world measurements were conducted as part of a short-term "mobile monitoring campaign" (MMC) from July 2019 to February 2020 in Melbourne, following the procedures of the EXPOsOMICS project (van Nunen et al., 2017). The measured data, was not part of the dataset used to generate the minor road models.

Briefly, the MMC simultaneously measured BC (MicroAethalometer AE51, AethLabs, San Francisco, CA, USA) and UFP (DiSCmini V2.0, Testo, Lenzkirch, Germany) for 30 min, and traffic counts for 15 min at 201 fixed monitoring sites. We followed a stratified random sampling
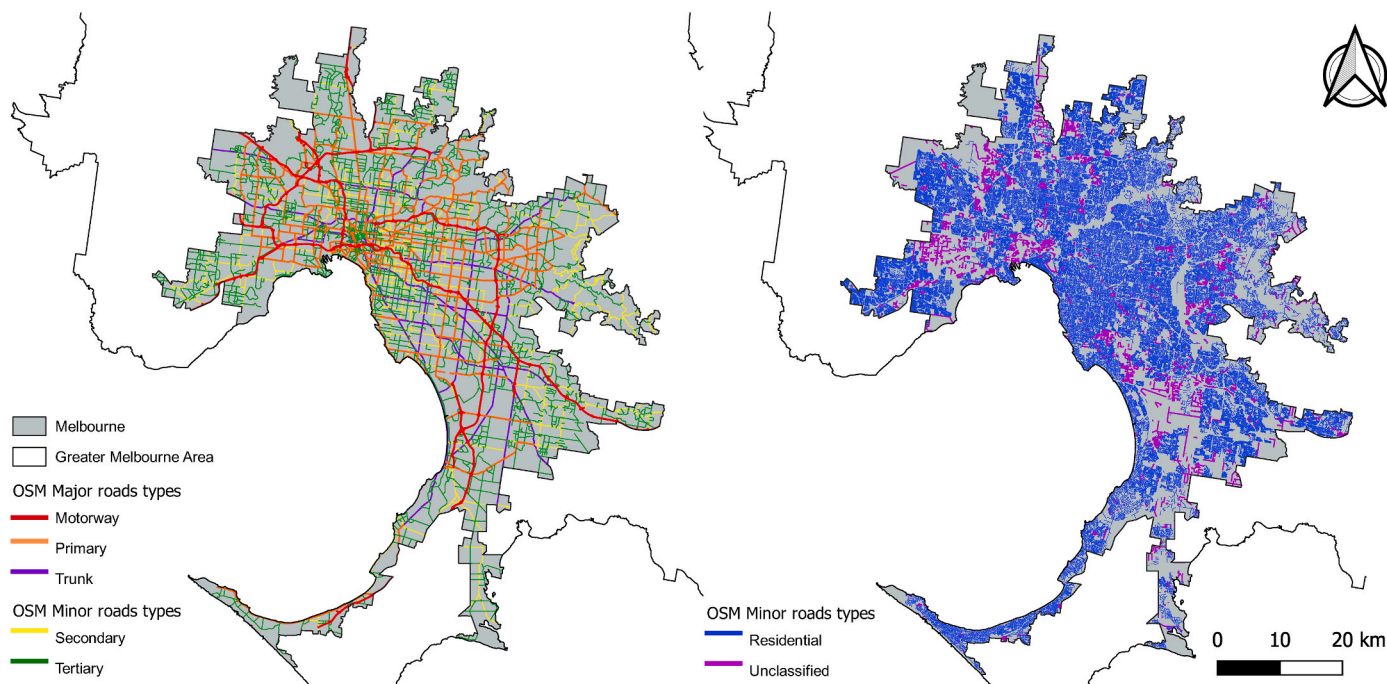

Fig. 1. Melbourne's distribution of roads per OpenStreetMap (OSM) type.

procedure by selecting monitoring sites according to their land use (see Appendix D), traffic volume, built environment, and dwelling and restaurant density. Most of the sites (84%) were >15 m from major roads. Two visits per season over three seasons (i.e., winter (July–August 2019), spring (October–November 2019) and summer (December 2019–February 2020)) were completed for each site producing a comprehensive database of traffic counts, BC and UFP. We took the measurements on working days between 9:00–17:00.

Traffic count data were manually collected with a counter over 15-min intervals at the nearest road (either minor or major). All data were transformed to AADT using the following conversion (Patrick, 2019):

$$AADT = traffic\ count \times 4 \times 12 \times NTF$$

The constants in the equation transform 15-min measurements into hourly data ($\times 4$), and then into 12-h periods ($\times 12$). The night-time factor (NTF) in the equation is assigned a value of 1.27 or 1.47 (mid-point within a range of values) depending on whether the road has low or high levels of traffic, respectively (low AADT $\sim$ 201–750 vehicles/day, and high AADT >2000 vehicles/day for each road lane). (Patrick, 2019).

We collected 1-min BC and 1-s UFP raw data following the procedure described in van Nunen et al. (2017). BC measurements did not require correcting the loading effect of the filter as it was exchanged before reaching an attenuation of 80 (filter overloading at 125). For UFP, we considered all particle concentrations ≤ 500 pt/cm^3 and treated all particle concentration values that differed ten times or more from the previous and next measurements as artifacts. Analyses were performed using datasets with 30-min average values per site.

2.4. Statistical analysis

2.4.1. Generating minor road AADT models

Linear, Poisson and negative binomial (NB) regressions were used to generate minor road AADT models. The latter two regression models were used to accommodate the positively skewed distribution of AADT. In addition, generalized additive models (GAM) were used to identify possible nonlinear relationships between continuous variables and the AADT. We used the likelihood ratio test (LRT) and Variance Inflation Factor (VIF) to find the optimal combination of predictors (within each type of regression) and their transformations (e.g., the natural logarithm of a predictor), while seeking not to overfit the data. For this purpose, we used the package 'mgcv' from R V.3.6.2 (R Foundation for Statistical Computing, Vienna, Austria). (R Core Team, 2019).

The appropriateness of the various regression models (linear, Poisson, NB) was determined by comparing the models' goodness-of-fit measures – namely, the Akaike information criterion (AIC) and Bayesian information criterion (BIC), where lower values indicate a better fitting model (Appendix E). The AADT models were fitted using the same combination of predictors, including road type, natural logarithms of the road importance index, AADT and distance to the nearest major road, population density, workplace density, and WRD. In addition, the selected models were evaluated by comparing their estimated median AADT values for minor roads (secondary, tertiary, unclassified and residential) to the measured median AADT values for the same roads (Appendix F).

2.4.2. Validation study

To assess the validity and added value of AADT modelling, the modelled AADT was compared to a traditional approach where minor roads with unknown AADT are assigned a fixed value of 600 vehicles/day (Morley and Gulliver, 2016). The fixed-value and modelled estimates were compared against measured AADT, BC, and UFP.

Scatter plots of estimated AADT values (from the models or fixed-value approach) vs. measured AADT, BC and UFP values were used to assess and compare the performance of the AADT models and the fixed-value approach. The agreement between estimated and measured AADT values was assessed using the concordance correlation coefficient (CCC), normalized mean bias factor (B'$_{NMBF}$) and normalized mean absolute error factor (E'$_{NMAEF}$) (see descriptions in Table 2). The associations between AADT estimates and BC and UFP measurements were assessed using Spearman's correlation coefficient (rho - 'ρ'). We assessed the model performance using three validation (sub) sets: major and minor roads, minor roads, and residential areas. Residential areas were defined as the sum of the high density residential and low density residential areas according to the 2018 Melbourne Planning Scheme (Department of Transport and Planning, 2018). Table 2 summarises each performance statistic/method, the target value and the datasets used for the validation analysis (mathematical formulas in Appendix G).

3. Results

3.1. Available AADT in Melbourne, Australia

Table 3 shows Melbourne's road network attributes, including the number of road segments, the proportion of segments with traffic count data, and the median and inter-quartile range (IQR) of AADT by road type. In 2018, 88.5% of roads constituting Melbourne's network were classified as minor, and most of them where residential (60.7%). Traffic data was available for only 28.2% of all the road network (i.e., 91,326 roads segments), from which 18.9% were minor roads. Residential roads had a median AADT of 1900 vehicles/day, being higher than the value used in the fixed-value approach when modelling AADT (600 vehicles/day). (Morley and Gulliver, 2016).

3.2. Generating minor road AADT models for Melbourne

The goodness-of-fit statistics for the linear and NB models are presented in Appendix A. The Poisson model was excluded from the analysis because of the overdispersion of the data with a dispersion parameter greater than '1' (2744.38). The NB model had better (i.e.,

Table 2
Performance statistics/methods used to evaluate the minor road AADT models.

Method	Target Value	Description	Validation Dataset
Coefficient of determination			
(R^2)	1	Determines the proportion of variance of the dependent variable explained by the independent variable.	Traffic counts, BC, UFP
Spearman's correlation coefficient (rho)			
(ρ)	1	Nonparametric measure of the strength and direction of the correlation between variables.	Traffic counts, BC, UFP
Concordance correlation coefficient with confidence intervals			
[CCC (CI)]	± 1	Absolute agreement statistic which evaluates the consistency and bias of the measurements with confidence intervals.	Traffic counts
Normalized mean bias factor			
(B'$_{NMBF}$)	0	Indicates over/underestimation of estimated values.	Traffic counts
Normalized mean absolute error factor			
(E'$_{NMAEF}$)	0	Represents the ratio of the mean absolute gross error and the mean observed value for the case of overestimation or the mean modelled value for the case of underestimation.	Traffic counts
Visual assessment			
Scatter Plot	Compact cloud along 45° diagonal	Shows the agreement between estimated and observed values.	Traffic counts

AADT = Annual average daily traffic; BC = Black carbon; UFP = Ultrafine particles; Performance statistics mathematical formulas in Appendix G.

Table 3
Descriptive statistics for major and minor roads in Zenith (2018) database and OSM road type classifications.

OSM Road type	Number of road segments (%)*	Road segments with AADT (%)**	AADT Median (IQR) (Vehicles/day)
All	323,722 (100)	91,326 (28.2)	8700 (15,200)
Major roads	37,089 (11.5)	37,089 (100.0)	18,900 (13,600)
Trunk	11,256 (3.5)	11,256 (100.0)	22,725 (12,200)
Primary	20,299 (6.3)	20,299 (100.0)	16,850 (11,050)
Minor roads	286,633 (88.5)	54,237 (18.9)	4300 (7300)
Secondary	20,177 (6.2)	11,673 (57.9)	12,000 (11,100)
Tertiary	50,749 (15.7)	23,813 (46.9)	4800 (6400)
Unclassified	19,077 (5.9)	1282 (6.7)	1900 (3300)
Residential	196,630 (60.7)	17,469 (8.9)	1900 (3100)

* Percentage of the total number of road segments; ** Percentage of the total of road segments with AADT; Zenith = Annual average daily traffic (AADT) dataset; OSM = OpenStreetMap; IQR = Interquartile range.

lower) AIC and BIC values than the linear model. However, the linear model had more deviance explained and had a median for residential roads closer to those previously reported (i.e., 710 vehicles/day) (Morley and Gulliver, 2016). Due to the similar performance of the linear and NB models, we used both for the validation analysis.

3.3. Validation study

In the validation study, we compared the AADT estimates from the linear and NB models and the fixed-value approach against independent measures of AADT, BC and UFP levels. Table 4 and Fig. 2 show the agreement between the estimated and measured AADT by three (sub) sets of the road network: all roads, minor roads, and minor roads within residential areas. Across all AADT estimation methods, the CCC decreased as the variability of the measured values decreased, while the B'$_{NMBF}$ and the E'$_{NMAEF}$, in general, increased. The CCC decrease was greater for the fixed-value approach than for the linear or the NB models (Table 4). The increase in the B'$_{NMBF}$ was greater for the NB model than for the linear model or for the fixed-value approach (Table 4).

The difference between the B'$_{NMBF}$ and the E'$_{NMAEF}$ (0.16–0.37) in the linear and NB models (all [sub]sets), indicates the over-compensation of values above and below the mean due to over-estimates compensating underestimates (Gustafson and Yu, 2012). Similar B'$_{NMBF}$ and E'$_{NMAEF}$ values were observed for the linear model and NB models when using all roads, with the minor road (sub)set models having higher absolute differences than the linear model. However, the largest differences were observed when modelling with fixed values, showing absolute differences between 0.23 and 1.01.

Table 5, Figs. 3 and 4 show the associations of estimated AADT with

Table 4
Agreement between estimated and measured AADT values by method of estimation.

Statistic	Linear	NB	Fixed-value
	All roads, n = 194 (51% of the AADT values were estimated)		
CCC (95% CI)	0.97 (0.97–0.98)	0.97 (0.96–0.98)	0.97 (0.96–0.98)
B$_{NMBF}$	0.06	0.11	0.01
E$_{NMAEF}$	0.23	0.27	0.24
	Minor roads, n = 137 (72% of the AADT values were estimated)		
CCC (95% CI)	0.90 (0.87–0.93)	0.88 (0.84–0.91)	0.85 (0.79–0.89)
B$_{NMBF}$	0.13	0.37	−0.17
E$_{NMAEF}$	0.50	0.68	0.62
	Residential area, n = 70 (91% of the AADT values were estimated)		
CCC (95% CI)	0.47 (0.32–0.60)	0.29 (0.18–0.39)	−0.001 (−0.17–0.18)
B$_{NMBF}$	1.08	2.47	0.06
E$_{NMAEF}$	1.42	2.68	1.07

AADT = Annual average daily traffic; CCC (CI) = Concordance correlation coefficient with confidence intervals; B'$_{NMBF}$ = Normalized mean bias factor; E'$_{NMAEF}$ = Normalized mean absolute error factor; n = number of sites; NB = Negative binomial.

measured BC and UFP concentrations for the three AADT estimation methods used in this study. The BC minor roads linear model yielded the strongest correlation (linear ρ = 0.54, NB ρ = 0.52, and fixed-value ρ = 0.46), and the UFP linear and NB models yielded the same correlation (linear ρ = 0.50, NB ρ = 0.50, and fixed-value ρ = 0.34). The relationships of AADT with BC and UFP for minor roads within residential areas were not statistically significant. However, the linear model had the highest ρ estimates (BC ρ = 0.18 and UFP ρ = 0.22).

4. Discussion

We have shown that minor road AADT modelling increases the estimation accuracy of AADT compared to the commonly used fixed-value approach. Our AADT estimates showed better agreement with measured AADT. We also had stronger correlations with BC and UFP when using the linear and NB models than the fixed-value approach. While the linear and NB models showed comparable performance for the overall road network, the linear model showed lower B'$_{NMBF}$ and E'$_{NMAEF}$ on all road (sub)sets. All (sub)set models showed lower B'$_{NMBF}$ than their respective E'$_{NMAEF}$ indicating over-compensation of the estimates. Lower compensation observed in the linear and NB models between over and under-predicted values indicates smoother AADT changes in the road network, which is an expected traffic flow behaviour in urban areas (Morley and Gulliver, 2016). On the contrary, fixed values produce abrupt changes of AADT in the network, which is less desirable when modelling TRAP as it is not representative of real-world conditions.

Our models overestimated AADT due to minor road traffic data being underrepresented and mainly having AADT on minor roads with values higher than the previously reported median AADT. We found that only 18.9% of minor roads within Melbourne's road network had AADT. Nevertheless, other geo-referenced databases necessary for AADT modelling, e.g., administrative areas, urban/rural areas and workplace statistics, had the required detail and depth of information.

Compared to previous studies, including Morley and Gulliver's (2016) study, which had the largest traffic count sample, the present study had access to significantly more AADT values on secondary roads (54,237 vs. 4462 points in Morley and Gulliver (2016)) (Morley and Gulliver, 2016). The predictors (i.e., variables entered in the models of AADT) and their effects on AADT were very similar to other studies on this topic, suggesting that traffic determinants are independent of a city's culture or local driving behaviour (Morley and Gulliver, 2016; Apronti et al., 2016; Fu et al., 2017). However, in contrast to the study by Apronti et al. (2016) in Wyoming, we found no differences by road surface, possibly because approximately 98.8% of the urban roads in Victoria are paved (Apronti et al., 2016; Greaves, 2021). Unlike Zhong and Hanson's (2009) study, our AADT estimates did not consider destinations of interest because appropriate data were not available and this would have possibly limited the importance of this variable in the models (Zhong and Hanson, 2009). Unlike Morley and Gulliver (2016), we could not apply the Poisson model to estimate minor road AADT because AADT data were over dispersed. To our knowledge, Morley and Gulliver's (2016) study and our study are the only ones that examined a representative set of minor roads.

Our AADT models showed better external validity than the models presented in previous studies. We obtained comparable AADT estimate errors and achieved stronger correlations with air pollutants (Morley and Gulliver, 2016; Apronti et al., 2016; Fu et al., 2017; Zhong and Hanson, 2009; Jung et al., 2017). This is the first study to quantify the added value and improvement in the strength of associations of estimated AADT in residential areas with BC and UFP concentrations compared to the fixed-value approach.

4.1. Strengths and limitations

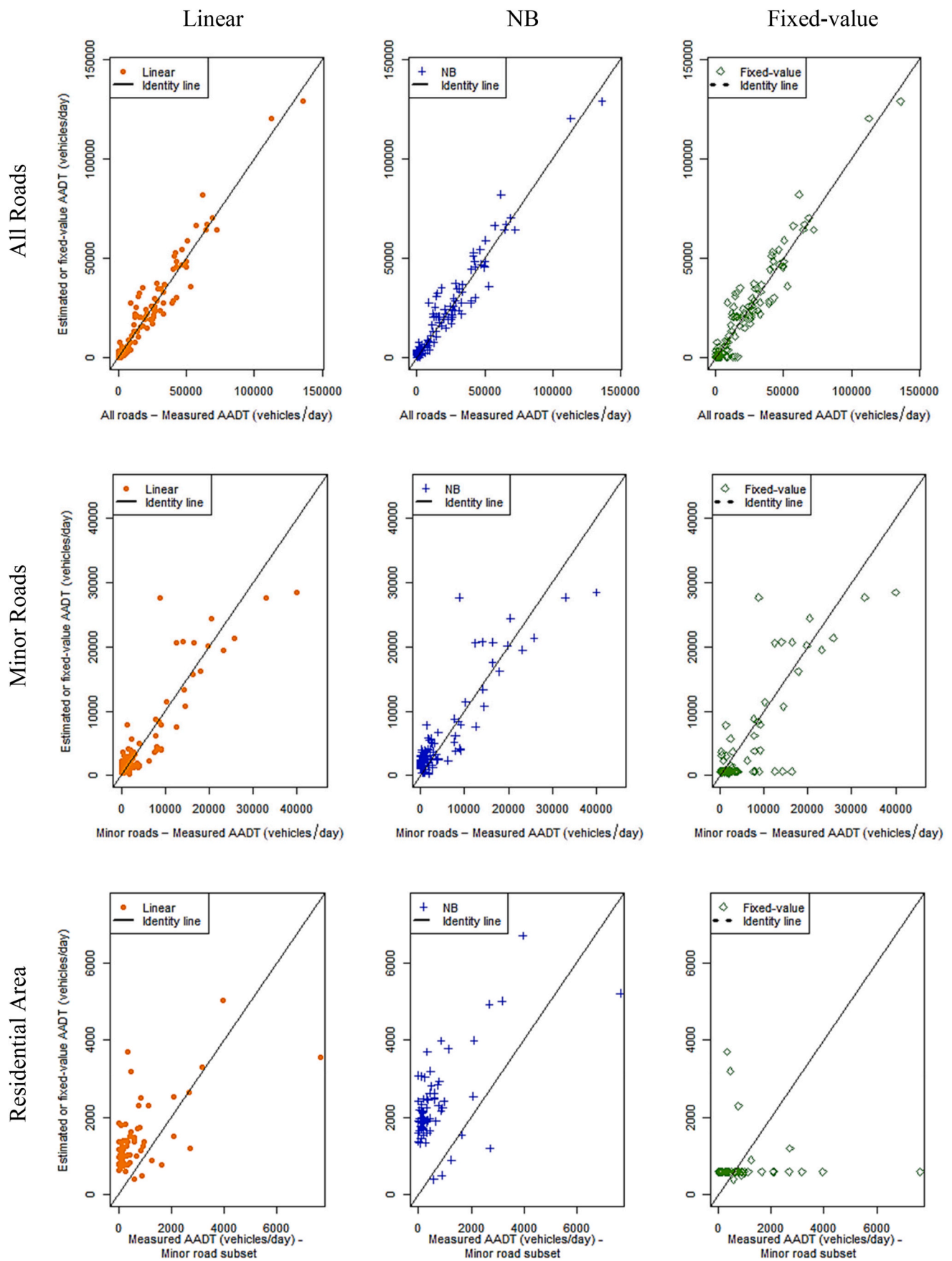This is the first time this AADT modelling approach has been

Fig. 2. Scatter plot of estimated vs. measured annual average daily traffic (AADT); NB = Negative binomial.

Table 5
Associations of estimated AADT with BC and UFP concentrations by method of AADT estimation (values represent Spearman's correlation coefficients and their 95% CIs).

Parameter	Linear	NB	Fixed-value
	All roads, n = 194 (50% of the AADT values were estimated)		
BC	0.64§ (0.55; 0.61)	0.62§ (0.53; 0.70)	0.61§ (0.51; 0.69)
UFP	0.59§ (0.49; 0.68)	0.58§ (0.48; 0.67)	0.53§ (0.42; 0.62)
	Minor roads, n = 143 (74% of the AADT values were estimated)		
BC	0.54§ (0.41; 0.64)	0.52§ (0.39; 0.62)	0.46§ (0.32; 0.58)
UFP	0.50§ (0.37; 0.62)	0.50§ (0.37; 0.62)	0.34§ (0.19; 0.48)
	Residential area, n = 70 (91% of the AADT values were estimated)		
BC	0.18 (−0.06; 0.40)	0.12 (−0.11; 0.35)	0.17 (−0.07; 0.39)
UFP	0.22$^+$ (−0.02; 0.43)	0.20 (−0.04; 0.42)	0.07 (−0.16; 0.34)

AADT = Annual average daily traffic; BC = Black carbon; UFP = Ultrafine particles; n = number of sites; CIs = Confidence intervals; Statistical significance codes (p): <0.001 '§' <0.001 '**' <0.01 '&' <0.05 '+'.

undertaken in a sprawling metropolis in Australia (see Appendix A). The strengths of our study lie in the comprehensive evaluation of two different statistical regression methods compared with the traditional fixed-value approach, including the assessment of associations between estimated and measured AADT, and measured air pollution. Furthermore, the models were evaluated for both systematic and random errors, enabling correction of the estimates for systematic error prior to implementation.

In terms of generalizability, we acknowledge that the specific characteristics of the study area (i.e., extensive urban sprawl) and data availability could limit the applicability of our findings to different settings. Nevertheless, the overarching principles and methodologies we outline, have the potential for adaptation and implementation in regions also lacking data for minor roads. In addition, the models primarily used open-data sources to ensure reproducibility in other geographical locations.

The traffic data used to generate the models had a low percentage of residential roads with available AADT (8.9%) and, as a result, they overestimated AADT relative to what was expected. These two factors limited the representativeness of these roads for residential areas and reduced the accuracy of the models. Future studies should include alternative modelling data in their design to address this issue, such as satellite imagery or stratified random surveys of traffic counters. In
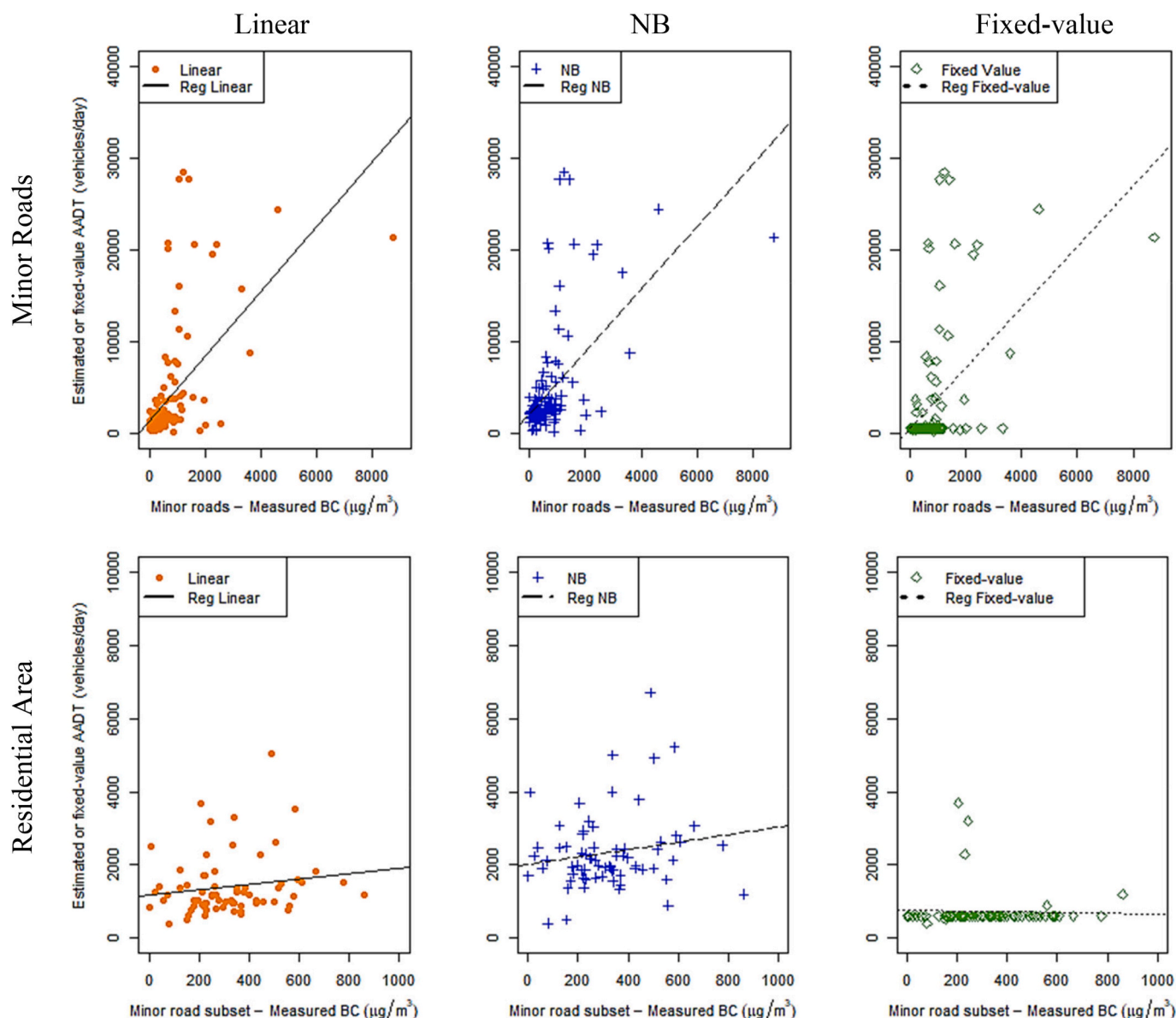
Fig. 3. Scatter plots of estimated annual average daily traffic (AADT) vs. measured black carbon (BC) concentrations; NB = Negative binomial; Reg = regression.
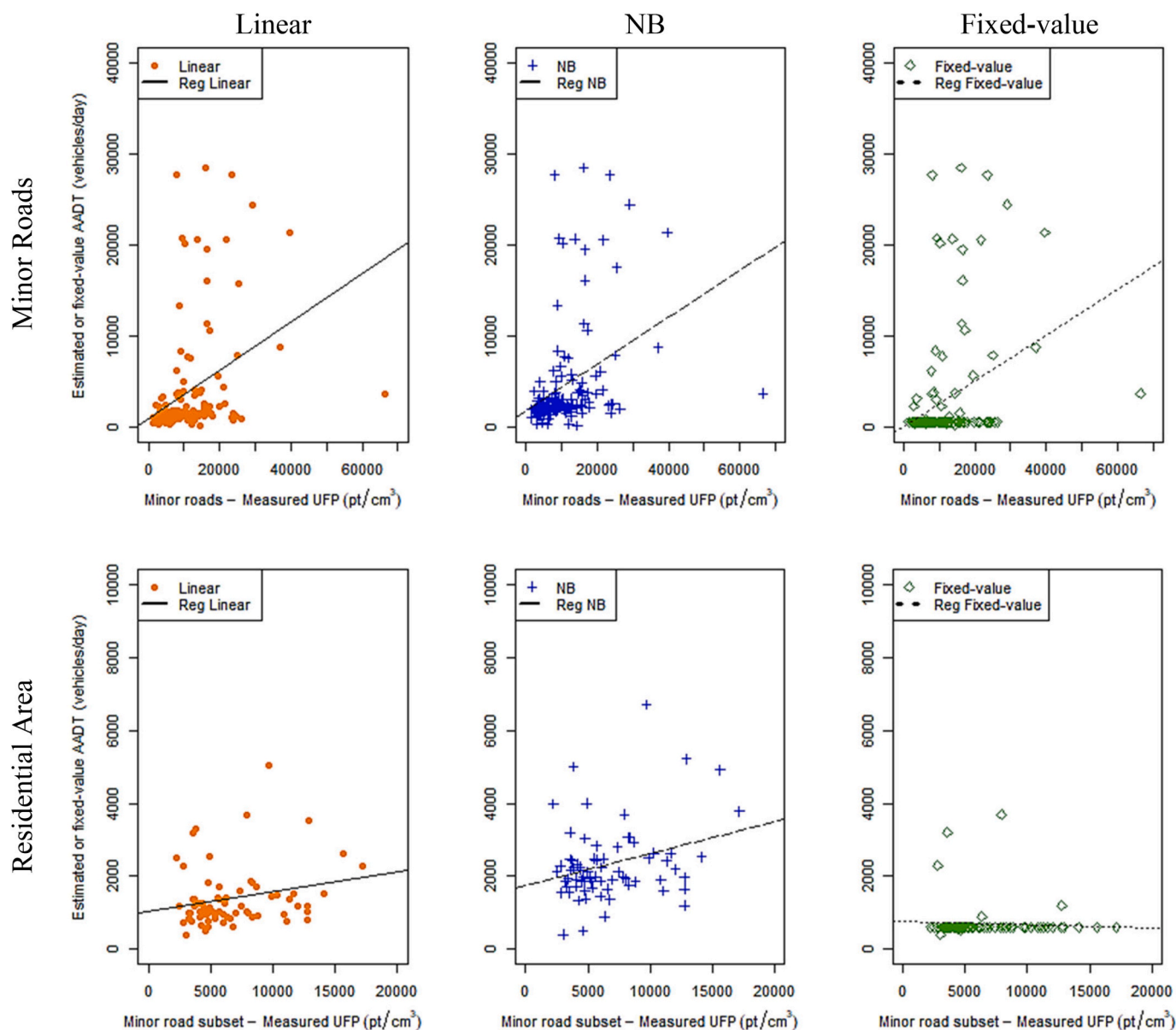
Fig. 4. Scatter plot of estimated annual average daily traffic (AADT) vs. measured ultrafine particle (UFP) concentrations; NB = Negative binomial; Reg = regression.

addition, although we followed a validated methodology, the traffic count data collected for the external validation analyses were still subject to a degree of random measurement error due to inter- and intra-day variations in traffic volume or unobserved factors in the built and social environment. Consequently, the association between measured and estimated AADT may have been underestimated.

This study showed that our minor road AADT model could better identify differences in traffic and BC and UFP concentrations on residential roads compared to the traditional fixed-value approach. After correcting for systematic errors, the AADT estimates could be integrated with existing traffic count data to study TRAP exposure. Thus, the proposed methodology, being a modelled estimation, is likely to be mainly affected by the Berkson type error rather than the classical error. The Berkson type error is due to using the group mean exposure (e.g. road types, land use) instead of individual values (Nieuwenhuijsen, 2015). However, the Berkson-type error causes little or no bias in the measured effects, thus allowing us to quantify the actual effects of the AADT. In addition, by disaggregating residential traffic levels relative to the fixed-value approach, the methodology also provided increased precision and statistical power. Both elements are particularly relevant for

studying the health effects resulting from traffic emissions in residential areas.

5. Conclusions

Minor road AADT, especially on residential roads, needs to be more accurate and representative when used to model road networks. Accuracy and representativeness are important to improve for the assessment of potential population health effects. The models used have proven valid in estimating minor roads' AADT for Melbourne and quantified the added value of modelling AADT compared to the traditional fixed-value approach. Furthermore, they increased the estimation accuracy and statistical power to study the effects of the AADT in residential areas, which are particularly relevant in the absence of health exposure thresholds for BC and UFP air pollution. In addition, we found that the current traffic data for Melbourne covered only 28.2% of the road network and a very low percentage of minor roads (18.9%) and residential roads (8.9%). Future studies should consider this limitation, as residential AADT needs to be more representative to improve the accuracy of the estimates. Finally, our linear model outperformed the

fixed-value approach when compared against traffic and TRAP measurements. The methodology followed in our study is relevant to locations with incomplete minor road AADT data.

Funding

CRediT author statement

Miguel Alvarado Molina: Conceptualization; Methodology; Software; Formal analysis; Investigation; Data curation; Writing - Original draft preparation; Visualization; **Ariadna Curto:** Resources; Methodology; Investigation; Writing - Review & Editing; **Amanda J. Wheeler:** Conceptualization; Methodology; Writing - Review & Editing; **Rachel Tham:** Conceptualization; Writing - Review & Editing; **Ester Cerin:** Funding acquisition; Writing - Review & Editing; **Mark Nieuwenhuijsen:** Writing - Review & Editing; **Roel Vermeulen:** Writing - Review & Editing; **David Donaire Gonzalez:** Conceptualization; Methodology; Formal analysis; Investigation; Writing - Review & Editing; Project administration.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used "Grammarly" and "GPT-4-OpenAI" in order to check grammar and improve readability of paragraphs. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2023.122657.

References

ABS, 2016a. *Census QuickStats: Melbourne (UCL)*. Melbourne Code UCL201001 (UCL). https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/UCL201001?opendocument.

2016b ABS. 2016b. Main Features - Statistical Area Level 1 (SA1). https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Statistical%20Area%20Level%201%20(SA1)~10013 (accessed 2022-10-17).

ABS, 2021. Melbourne, Census All Persons QuickStats. Australian Bureau of Statistics. https://abs.gov.au/census/find-census-data/quickstats/2021/UCL201001.

Aldred, R., Verlinghieri, E., 2020. LTNs for All? Mapping the Extent of London's New Low Traffic Neighbourhoods. Possible. https://policycommons.net/artifacts/1439909/ltnsforall3/2067510/.

Apronti, D., Ksaibati, K., Gerow, K., Hepner, J.J., 2016. Estimating traffic volume on Wyoming low volume roads using linear and logistic regression methods. J. Traffic Transp. Eng. Engl. Ed. 3 (6), 493–506. https://doi.org/10.1016/j.jtte.2016.02.004.

Clarke, J., Thatcher, M., Hernaman, V., Heady, C., Round, V., Rafter, T., Trenham, C., Wilson, L., 2019. Victorian Climate Projections 2019 Technical Report. CSIRO, Victoria. Technical 978-1-76077-736–4. https://www.climatechangeinaustralia.gov.au/en/projects/victorian-climate-projections-2019/vcp19-publications/.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. Lancet Lond. Engl. 389 (10082), 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6.

de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U.A., Katsouyanni, K., Klompmaker, J., Martin, R.V., Samoli, E., Schwartz, P.E., Stafoggia, M., Bellander, T., Strak, M., Wolf, K., Vienneau, D., Brunekreef, B., Hoek, G., 2018. Spatial PM2.5, NO2, O3 and BC models for western europe – evaluation of spatiotemporal stability. Environ. Int. 120, 81–92. https://doi.org/10.1016/j.envint.2018.07.036.

Department of Transport and Planning. *Melbourne Planning Scheme - Ordinance*; Planning Scheme, 2018. Department of Transport and Planning, Melbourne. https://planning-schemes.app.planning.vic.gov.au/Melbourne/ordinance.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dèdèlè, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S. von, Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of land use regression models for PM2.5, PM2.5 absorbance, PM10 and PMcoarse in 20 European study areas; results of the ESCAPE project. Environ. Sci. Technol. 46 (20), 11195–11205. https://doi.org/10.1021/es301948k.

Fu, M., Kelly, J.A., Clinch, J.P., 2017. Estimating annual average daily traffic and Transport emissions for a national road network: a bottom-up methodology for both nationally-aggregated and spatially-disaggregated results. J. Transport Geogr. 58, 186–195. https://doi.org/10.1016/j.jtrangeo.2016.12.002.

Greaves, A., 2021. *Maintaining Local Roads*; 978-1-925678-90–1. Victorian Auditor-General's Office, Melbourne, Australia, p. 84. https://www.audit.vic.gov.au/sites/default/files/2021-03/20210317-Local-Roads-report.pdf?, 2021.

Gustafson Jr., W.I., Yu, S., 2012. Generalized approach for using unbiased symmetric metrics with negative values: normalized mean bias factor and normalized mean absolute error factor. Atmos. Sci. Lett. 13 (4), 262–267. https://doi.org/10.1002/asl.393.

Jung, S., Kim, J., Kim, J., Hong, D., Park, D., 2017. An estimation of vehicle kilometer traveled and on-road emissions using the traffic volume and travel speed on road links in incheon city. J. Environ. Sci. 54, 90–100. https://doi.org/10.1016/j.jes.2015.12.040.

Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., Puybonnieux-Texier, V., Quénel, P., Schneider, J., Seethaler, R., Vergnaud, J.C., Sommer, H., 2000. Public-health impact of outdoor and traffic-related air pollution: a European assessment. Lancet Lond. Engl. 356 (9232), 795–801. https://doi.org/10.1016/S0140-6736(00)02653-2.

Mikiewicz, D., Mackiewicz, M., Nycz, J., 2017. Mastering PostGIS Create, Deliver, and Consume Spatial Data Using PostGIS.

Moeller, C. osm2po - openstreetmap converter and routing engine for java. http://osm2po.de/ (accessed 2022-December-16)..

Morley, D.W., Gulliver, J., 2016. Methods to improve traffic flow and noise exposure estimation on minor roads. Environ. Pollut. 216, 746–754. https://doi.org/10.1016/j.envpol.2016.06.042.

Nieuwenhuijsen, M.J., 2015. Exposure Assessment in Environmental Epidemiology. Oxford University Press.

OGC. The Home of Location Technology Innovation and Collaboration | OGC. https://www.ogc.org/ (accessed 2021-07-15)..

Patrick, S., 2019. Guide to Pavement Technology Part 4K: Selection and Design of Sprayed Seals. Austroads.

pgRouting. pgRouting Project - Open Source Routing Library. pgRouting Project. http://pgrouting.org/ (accessed 2020-0January-10).

R Core Team, 2019. *R*: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rose, N., Cowie, C., Gillett, R., Marks, G.B., 2009. Weighted road density: a simple way of assigning traffic-related air pollution exposure. Atmos. Environ. 43 (32), 5009–5014. https://doi.org/10.1016/j.atmosenv.2009.06.049.

Saha, P.K., Zimmerman, N., Malings, C., Hauryliuk, A., Li, Z., Snell, L., Subramanian, R., Lipsky, E., Apte, J.S., Robinson, A.L., Presto, A.A., 2019. Quantifying high-resolution spatial variations and local source impacts of urban ultrafine particle concentrations. Sci. Total Environ. 655, 473–481. https://doi.org/10.1016/j.scitotenv.2018.11.197.

Shen, Y., de Hoogh, K., Schmitz, O., Clinton, N., Tuxen-Bettman, K., Brandt, J., Christensen, J.H., Frohn, L.M., Geels, C., Karssenberg, D., Vermeulen, R., Hoek, G., 2022. Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression. Environ. Int. 168, 107485 https://doi.org/10.1016/j.envint.2022.107485.

Strak, M., Weinmayr, G., Rodopoulou, S., Chen, J., de Hoogh, K., Andersen, Z.J., Atkinson, R., Bauwelinck, M., Bekkevold, T., Bellander, T., Boutron-Ruault, M.-C., Brandt, J., Cesaroni, G., Concin, H., Fecht, D., Forastiere, F., Gulliver, J., Hertel, O., Hoffmann, B., Hvidtfeldt, U.A., Janssen, N.A.H., Jöckel, K.-H., Jørgensen, J.T., Ketzel, M., Klompmaker, J.O., Lager, A., Leander, K., Liu, S., Ljungman, P., Magnusson, P.K.E., Mehta, A.J., Nagel, G., Oftedal, B., Pershagen, G., Peters, A., Raaschou-Nielsen, O., Renzi, M., Rizzuto, D., van der Schouw, Y.T., Schramm, S., Severi, G., Sigsgaard, T., Sørensen, M., Stafoggia, M., Tjønneland, A., Verschuren, W. M.M., Vienneau, D., Wolf, K., Katsouyanni, K., Brunekreef, B., Hoek, G., Samoli, E., 2021. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis. BMJ 374, n1904. https://doi.org/10.1136/bmj.n1904.

Van den Hove, A., Verwaeren, J., Van den Bossche, J., Theunis, J., De Baets, B., 2020. Development of a land use regression model for black carbon using mobile monitoring data and its application to pollution-avoiding routing. Environ. Res. 183, 108619 https://doi.org/10.1016/j.envres.2019.108619.

van Nunen, E., Vermeulen, R., Tsai, M.-Y., Probst-Hensch, N., Ineichen, A., Davey, M., Imboden, M., Ducret-Stich, R., Naccarati, A., Raffaele, D., Ranzi, A., Ivaldi, C., Galassi, C., Nieuwenhuijsen, M., Curto, A., Donaire-Gonzalez, D., Cirach, M., Chatzi, L., Kampouri, M., Vlaanderen, J., Meliefste, K., Buijtenhuijs, D., Brunekreef, B., Morley, D., Vineis, P., Gulliver, J., Hoek, G., 2017. Land use regression models for ultrafine particles in six European areas. Environ. Sci. Technol. 51 (6), 3336–3345. https://doi.org/10.1021/acs.est.6b05920.

Veitch Lister Consulting, 2020. *Traffic Flow Zenith 2018*; Traffic Flow Zenith 2018; 4699355722. Centre of air pollution, energy, and health research. https://cloudstor.aarnet.edu.au/plus/f/4699355722.

VicGov. 2020. Victorian *Integrated Survey Of Travel & Activity (VISTA)*. Tableau Software. https://public.tableau.com/views/VISTA-Trips-timeseriesAccess/Trips-methodoftravel?%3Aembed=y&%3AshowVizHome=no&%3Adisplay_count=y&%3Adisplay_static_image=y&%3AbootstrapWhenNotified=true&%3Alanguage=en&:embed=y&:showVizHome=n&:apiID=host0#navType=0&navSrc=Parse (accessed 2020-0August-21)..

Zhong, M., Hanson, B.L., 2009. GIS-based travel demand modeling for estimating traffic on low-class roads. Transport. Plann. Technol. 32 (5), 423–439. https://doi.org/10.1080/03081060903257053.