Development of a microscale land use regression model for predicting NO_2 concentrations at a heavy trafficked suburban area in Auckland, NZ

Weissert, LF[1)], Salmond, JA[2)*], Miskell, G[1)], Alavi-Shoshtari, M[1)], Williams, DE[1)]

[1)] School of Chemical Sciences, Faculty of Science, University of Auckland, New Zealand

[2)] School of Environment, Faculty of Science, University of Auckland, Auckland, New Zealand

* Corresponding author. E-mail address: j.salmond@auckland.ac.nz, Tel:+64 9 3737599 ext 88650

Abstract

Land use regression (LUR) analysis has become a key method to explain air pollutant concentrations at unmeasured sites at city or country scales, but little is known about the applicability of LUR at microscales. We present a microscale LUR model developed for a heavy trafficked section of road in Auckland, New Zealand. We also test the within-city transferability of LUR models developed at different spatial scales (local scale and city scale). Nitrogen dioxide (NO_2) was measured during summer at 40 sites and a LUR model was developed based on standard criteria. The results showed that LUR models are able to capture the microscale variability with the model explaining 66% of the variability in NO_2 concentrations. Predictor variables identified at this scale were street width, distance to major road, presence of awnings and number of bus stops, with the latter three also being important determinants at the local scale. This highlights the importance of street and building configurations for individual exposure at the street level. However, within-city transferability was limited with the number of bus stops being the only significant predictor variable at all spatial scales and locations tested, indicating the strong influence of diesel emissions related to bus traffic. These findings show that air quality monitoring is necessary at a high spatial density within cities in capturing small-scale variability in NO_2 concentrations at the street level and assessing individual exposure to traffic related air pollutants.

Keywords: LUR; air pollution; nitrogen dioxide; intra-urban air pollution; transferability; GIS

Introduction

In many cities, personal exposure to air pollution is primarily determined by time spent in the transport micro-environment (Dirks et al., 2012; McNabola et al., 2009; Kaur and Nieuwenhuijsen, 2009). However, time spent in this environment is not limited to commuter activities, and high densities of people are also observed moving through transport corridors as they visit the shops, restaurants and recreational facilities found clustered along busy streets and intersections. In such transport micro-environments the temporal and spatial variability in air pollution concentrations is large, and may be even greater than variability between cities (Hoek et al., 2008; Gurung et al., 2017).

Measuring air pollutant concentrations at local (representative of a neighbourhood or suburb) or microscales (representative of individual roads) in transport corridors is especially challenging as pollutant concentrations are strongly dependent on short-term traffic conditions and the configurations of buildings and streets (Eeftens et al., 2013; Miskell et al., 2015). For example, significant reductions in pollutant concentrations can be observed just a few meters from emission sources (Grange et al., 2014), and roadside concentrations can differ substantially from local background concentrations (Vardoulakis et al., 2011). Further, buildings modify local air flow patterns causing trapping and re-circulatory flows at some locations and increased dispersion of air pollutants at other locations (Salmond and McKendry, 2009; Salmond et al., 2013; Shi et al., 2016). As a result little is known about the relative importance of urban morphology, building design, traffic management and infrastructure (including phasing of traffic lights), and other details such as vegetation or bus stop positions in determining microscale air quality variability. There is therefore a need to improve our understanding about the microscale spatial variability of air pollutants in urban hotspots if we are to develop urban planning and design tools to control and mitigate personal exposure to air pollution in transport corridors, especially at locations where traffic as well as pedestrian activity is high (Borge et al., 2016).

Land use regression (LUR) analysis which emerged as a popular method in epidemiological studies to predict air pollutant concentrations and assess individual exposure levels (Hoek et al., 2008; Jerrett et al., 2005). It has the potential to assist urban planners in identifying the key controls on local air quality in transport corridors. Based on selected land use characteristics (e.g. distance to nearest road, land cover or population density), which are now widely available through geographic information (GIS) systems, LUR models allow estimation of air pollutant concentrations at unmeasured sites based on regression analysis (Hoek et al., 2008; Jerrett et al., 2005). Thus, LUR models are often used to complement regulatory monitoring networks, which are usually sparse due to logistical and financial constraints (Hoek et al., 2008; Vardoulakis et al., 2011). Such models have been primarily developed and applied to urban scale analyses (10^4 - 10^5 m), with some applied to the regional or country scale where they have been effectively used to identify common determinants of air quality for primarily transport related pollutants such as nitrogen dioxide (NO_2). These include factors such as road length, distance to major roads, land cover, traffic volume and density, population density and altitude guide Hoek et al. (2008). Final models typically explain around 60 - 70% of the variability (Beelen et al., 2013) with a range from 51% (Briggs et al., 2000; Gurung et al., 2017; Morgenstern et al., 2007) to 97% (Stedman et al., 1997).

However, there is little evidence to demonstrate their effectiveness (or otherwise) under the highly heterogeneous conditions typical of multi-use transport corridors, and their ability to capture and effectively represent local variability at urban hotspots may be limited (Apte et al., 2017; Ghassoun et al., 2015; Hoek et al., 2008). Further, although LUR models have been used in numerous cities across Europe and North America (Hoek et al., 2008), results from other geographical regions have only recently become available and remain limited (e.g. Australia

(Dirgawati et al., 2015); China (Meng et al., 2015); Nepal (Gurung et al., 2017); New Zealand (Miskell et al., 2015); Iran (Amini et al., 2016)).

In this study, we present a LUR model developed for urban microscales and applied to a heavily trafficked suburban street in Auckland, New Zealand. Our study is one of a limited number of studies (such as Miskell et al., 2015) which address local to microscale spatial variability (1-3 km) and use local urban design features as predictor variables (such as presence or absence of shop awnings) rather than standard landuse predictors (such as population and household density) which were homogenous within our study area. In particular, we were interested in examining the transferability of this approach. We also tested the within-city transferability of previously developed LUR models and explored the potential to extend the multi-scale model developed in Auckland's CBD by Miskell et al. (2015) to all spatial scales and sites outside the CBD. This study therefore also offers new insights into the applicability of LUR models developed for a certain area to other locations within the same city at different scales, which has not previously been explored.

Material and Methods

Study area

Auckland is New Zealand's largest and fastest growing city with around 1.5 million inhabitants (Statistics New Zealand, 2013). Vehicle emissions are the largest contributor to air pollution in Auckland with traffic-related NO_x (NO_2, NO) emissions accounting for almost 80% of the total NO_x emissions (Xie et al., 2016). However, pollutants are often dispersed by maritime winds, which occur year-round favoured by Auckland's location on a narrow isthmus (Chappell, 2014; Senaratne and Shooter, 2004). The focus of this study was on a heavy trafficked road (Dominion Road) about 4 km south of the city center (Fig. 1). Dominion Road is a main route for buses and commuters in and out of the city as well as to the main airport

(Auckland Transport, 2017). The area is also well used by pedestrians visiting shops, bars and cafés along the road, making this an interesting area for air pollution measurements due to the high traffic and potential exposure.

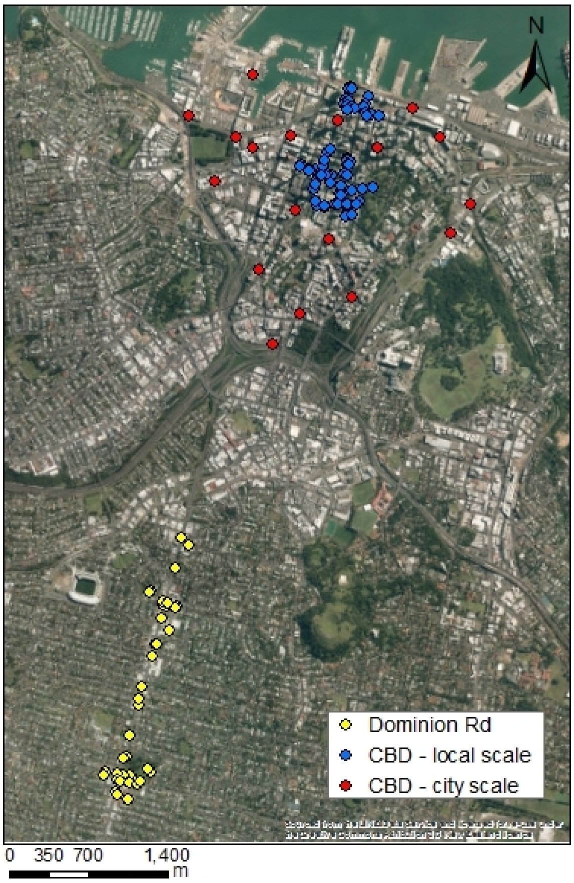

Fig. 1 Study sites.

NO_2 concentration measurements

NO_2 concentrations were measured by Palmes diffusion tubes at 40 sites along a 2 km section of Dominion Road (Fig 1). Sites were chosen to reflect a range of urban design features (such as the presence of building awnings, proximity to bus stops, greenspace, trees and carparks etc.). Sites were also chosen to represent the range of expected spatial variability of air pollution concentrations. The number of sites previous other LUR studies ranged from 14 – 107 (see Hoek et al. 2008, Beelen et al. 2013), with sample sizes of 40 commonly used in the ESCAPE project (which is most commonly referenced as the standard methodology for such studies) (Beelen et al., 2013). Given the size of our sample area, and the number of different environments expected, the choice of a sample size of 40 was deemed sufficient and representative within the context of the resources available.

At each site, we deployed two tubes at a height of approx. 2.5 m for four periods of 14 days between the 18th of November 2016 and the 1st of February 2017. To assess the reliability of the NO_2 measurements during each campaign we used travel and laboratory blanks (AEA Energy and Environment, 2008). Palmes tubes were analysed using a spectrophotometer and NO_2 concentrations calculated following standard methodology (AEA Energy and Environment, 2008). The coefficient of variance (CoV) was used to test the agreement between duplicate readings at each site and results that exceeded a CoV of 0.25 were excluded from the further analysis (Miskell et al., 2015; Mölter et al., 2012). As there was no reference regulatory air quality station near the road section studied here, we were not able to apply a seasonal adjustment to the NO_2 concentrations. Thus, we used seasonally averaged NO_2 concentrations, representative of typical summer conditions in this study, which are likely slightly below the annual average. For comparison, NO_2 measured by routine air quality monitors from 2010 – 2011 by the Auckland Council at another urban road (Khyber Pass, approx. 2 km northeast from Dominion Road) was on average 1 and 3 µg m^{-3} below the annual average in December and January, respectively. Slightly larger differences were observed in Auckland's Central Business District (CBD) where five-year average NO_2 concentrations measured in December and January were around 10 µg m^{-3} below the annual average (Miskell, 2013).

Predictor variables

Predictor variables for the initial stages were chosen based on a previous study undertaken in Auckland (Table 1, (Miskell et al., 2015)) and generated for each site using GIS shape files

(Auckland Council, 2013-2014) and aerial photographs. Each variable was either representative of a pollution source, such as number of lanes, or dispersion modifier, such as presence of building awnings (Table 1). As traffic density was not available at this high spatial resolution we used average weekday traffic congestion during the morning rush hour (06:00 – 09:00), midday (11:00 – 13:00) and evening rush hour (17:00 – 19:00) reported on Google Maps as a proxy. In this area land use, population density, household density and number of buildings show no significant spatially variability thus these predictors are not included in our analysis.

Table 1. Predictor variables with defined buffer sizes and expected direction of effect. (Attached at the end of the manuscript)

LUR model development

The LUR model was developed based on stepwise variable selection as outlined by the ESCAPE protocol (Beelen et al., 2013; Brunekreef, 2008). First, each predictor variable was compared to the average NO_2 concentrations measured throughout the study period using univariate regressions. Variables that did not follow the expected slope direction (e.g. an increase in number of traffic lanes is expected to increase air pollutants) were removed from further analysis. The variable with the highest adjusted R^2 was used to start developing the LUR model and predictor variables were then added one at the time and included in the model following standard procedures of the ESCAPE protocol (Brunekreef, 2008). In the final stage, variables with a p-value > 0.1 were removed from the model sequentially.

The final model was tested for multi-collinearity (variance inflation factor > 3), normality, heteroscedasticity, high-leverage points or outliers (Cook's Distance > 1) and spatial autocorrelation (using Moran's I) of the residuals (Brunekreef, 2008). The model was validated using two approaches that are suitable for a small sample size (Dirgawati et al., 2015; Tang et al., 2013). First, we used the 'leave-one-out cross-validation' (LOOCV) method, where the final model was fitted to N – 1 sites and the predicted concentrations were compared to the measured concentration at the left-out site. As the LOOCV tends to overestimate the model performance, we also used the grouped cross-validation, where a random proportion (30%) of the data was used to train a model while the remainder was used for the prediction. This process was repeated 20 times and the average performance was used. The model performance was assessed using the R^2, the root mean square error (RMSE) and the mean squared error (MSE)–R^2. The MSE-R^2 is a more representative metric to assess the goodness of fit around the 1:1 line (Tang et al., 2013) and was calculated as:

$$\mathrm{MSE} - R^2 = 1 - \frac{\mathrm{MSE}}{\left(\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_t)^2\right)}$$

where y_i is the monitored NO_2 concentration at each site and $\bar{y}_t$ the averaged NO_2 concentrations (Tang et al., 2013). Finally, the model results were mapped, with the study area divided into 50 m grid squares and NO_2 concentrations predicted for the centre of each grid square using the LUR model. Inverse distance weighting (IDW) was then used to interpolate the modelled NO_2 concentrations (Ghassoun et al., 2015; Liu et al., 2016). Analysis was done using R (3.2.4) and ArcGIS (v.10.2.2).

Model transferability

Model transferability was tested by transferring the model developed at the microscale to the local scale and city scale dataset collected during a previous study (Miskell et al., 2015). In addition, we applied the multi-scale (city and local scale) model developed by Miskell et al. (2015) to the Dominion Road dataset allowing coefficients to be flexible to explore the potential to extend the multi-scale model to the microscale outside Auckland's CBD. A detailed description of the multi-scale model development is provided by Miskell et al. (2015). In brief, Miskell et al. (2015) developed a strategy to identify those predictor variables that were able to explain spatial variability of NO_2 at different spatial scales so that transferability of the model to either different locations or different scales may be improved. First, two LURs were developed at two different spatial scales (using the local scale and city scale data, Fig. 1) following the standard protocol. Next, these models were used on the other set of data (e.g. local scale model on the city scale data) in order to identify and remove any variables that may be present due to specifics of the data or due to model fitting (e.g. change in slope direction). The local scale model was then used on the city scale data, in order to improve the small-scale explanations, and the city scale model mostly used the local scale predictors. Following the model revision, those variables with p-values > 0.2 were removed, one at a time, to reach a new, revised model, which is referred to as the multi-scale model. This model had adjustable coefficient values for the different spatial scales in order to maximize specific fits and to give comparable performance results to those from their specific LUR models. This illustrated the potential to improve local scale explanations, with a requirement to validate this on a third, independent dataset.

Results and Discussion

Air pollution levels

A summary of NO_2 concentrations measured during each campaign and averaged over the whole study period (summer average) is shown in Table 2. In total, we made 149 measurements at the 40 sites, with some tubes lost or moved during each campaign. Overall, 32 sites had NO_2 concentrations from all campaigns, while the remaining sites were missing NO_2 concentration measurements for one (5 sites) or two campaigns (3 sites). The lower NO_2 concentrations measured during campaign 4 are likely related to the lower traffic due to summer holidays. The duplicates generally agreed well with an average (SD) CoV of 0.052 (0.053) and no sample exceeded the threshold of 0.25. Mean NO_2 concentrations measured along Dominion Road (overall 22 µg m^{-3}) were below concentrations measured in North American and European cities reviewed by Hoek et al. (2008), but similar or above concentrations measured during the ESCAPE study in some European cities (e.g. Oslo, Norway (23 µg m^{-3}), Copenhagen, Denmark (18 µg m^{-3})) (Beelen et al., 2013) and above concentrations observed in streets in Perth, Australia (12 µg m^{-3}) (Dirgawati et al., 2015). The average was below that measured in Auckland's CBD (34 µg m^{-3}), where tall buildings and high bus traffic favour the build-up of pollutants (Miskell et al., 2015; Weissert et al., 2015). Given that studies typically represent annual averages, differences may partly be explained by temporal differences of the measurements as we decided to present a seasonal average representative of summer, when NO_2 concentrations are generally expected to be below the annual average. As expected, higher concentrations were observed at sites around intersections while lower concentrations were observed in park areas or streets away from the main road (Fig. 2).

Table 2. Descriptive statistics of the NO_2 concentrations measured during summer 2016/2017 along Dominion Road (units are in µg m^{-3}).

Sampling campaign	*n* (measurements)	Mean	Standard Deviation (SD)	Median	Min	Max	Range
1	35	23	7	22	12	34	22
2	39	21	6	22	12	35	23
3	36	27	7	26	15	42	27
4	39	16	6	17	7	28	21

Overall	40		22	6	22	12	34	22

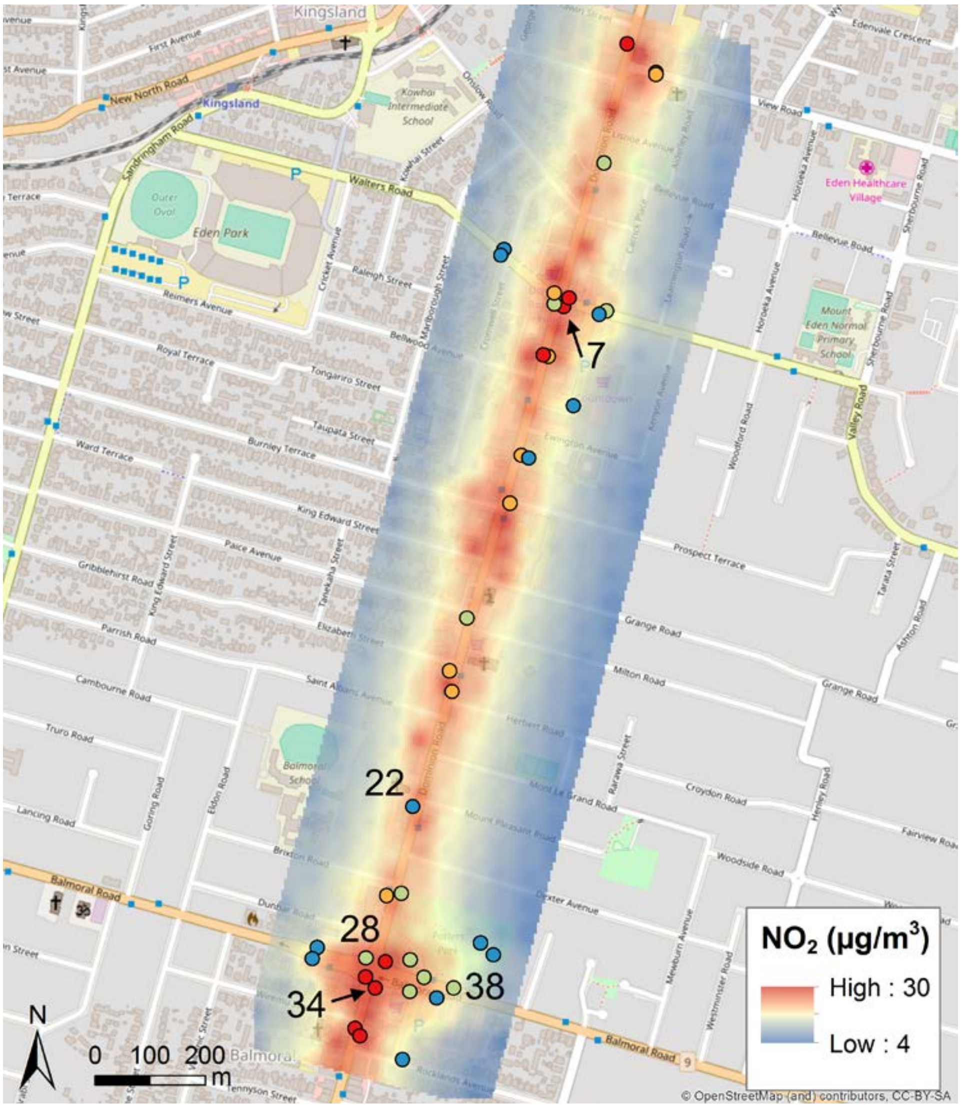

Fig. 2 NO_2 concentrations measured at the 40 sites (represented by the dots, with numbered dots showing sites with largest discrepancies between modelled and measured NO_2 concentrations) and modelled along Dominion Road.

Predictor variables

The predictor variables used in the final model were distance to major road, number of bus stops within a 100 m buffer, presence of awnings and street width (Table 3). Of these, presence of awnings (dispersion determinants) had the highest proportional contribution (β * (90th percentile – 10th percentile)) to the modelled NO_2 concentrations (79.08%), followed by number of bus stops within a 100 m buffer (16.52%) (Table 3). Distance to major road or street width only had a minor influence on modelled NO_2 concentrations. When comparing the predictor variables with city or regional scale LUR models (e.g. Beelen et al., 2013; Hoek et al., 2008), where predictor variables are usually related to traffic and land cover, it becomes evident that local scale street and building configurations become also important at the microscale. It is interesting to note that similar variables were also observed to be important in a local scale model in Auckland's CBD (e.g. number of bus stops, presence of awnings) (Miskell et al., 2015). Likewise, Tang et al. (2013) showed that including building and street configurations can be used to account for pollution dispersion and accumulation patterns in urban areas and improve the performance of LUR models. Following the ESCAPE protocol we removed two predictor variables from the model development (distance to tree, number of carparks) as they did not follow the expected pattern or slope of effect. In this study 'distance to tree' was not a significant variable in the model. This may be because trees can both act to increase and decrease air pollution concentrations depending on the dominant process. The presence of trees may improve air quality through enhanced deposition processes but trees may also decrease the dispersion of pollutants resulting in a local increase of NO_2 concentrations (Janhäll, 2015; Salmond et al., 2013). We also removed 'number of carparks' from the model because the effect of this parameter was variable depending on buffer size.

LUR model results and limitations

The final model explained 66% of the variability in NO_2 concentrations with a RMSE of 3.317 µg m^{-3} (Table 3). On average, the modelled NO_2 concentrations were the same as the measured NO_2 concentrations with an almost equal number of over- and underestimated sites (18 and 17 sites, respectively). The largest difference between measured and modelled NO_2 concentrations were observed at site 22 and 28 where modelled NO_2 concentrations were 7 µg m^{-3} above measured NO_2 concentrations. Observations at both sites were unexpectedly low given their location adjacent to Dominion road and close proximity bus stops. A further measurement campaign may be required to account for the discrepancy at these sites. In contrast, modelled NO_2 concentrations at sites 34, 38 and 7 were 6 µg m^{-3} lower than those measured (Fig. 2). Again, the observed measurements at these sites were unexpected. Site 7 and 34 had higher measured NO_2 concentrations than sites in their surroundings with similar land use characteristics. Site 38 is located relatively far away from Dominion road, but is still located along a busy road, but this road is not accounted for in the model.

The R^2 (MSE-R^2) and RMSE of the LOOCV validation were 0.60 (0.61) and 3.839 µg m^{-3}, respectively (Table 4). A slightly lower MSE-R^2 (0.60) was achieved from the LGOCV method (Table 4). Nevertheless, both R^2 are similar to the model R^2 (Table 3) indicating that the model performed well under internal validation. The adjusted R^2 is within the range of those achieved by LUR in European cities (55% - 92%) (Beelen et al., 2013). The RMSE, on the other hand, was lower than the RMSE of most cities in the ESCAPE study (Beelen et al., 2013), indicating a better overall accuracy of the LUR model due to data being less spread around the best-fit line.

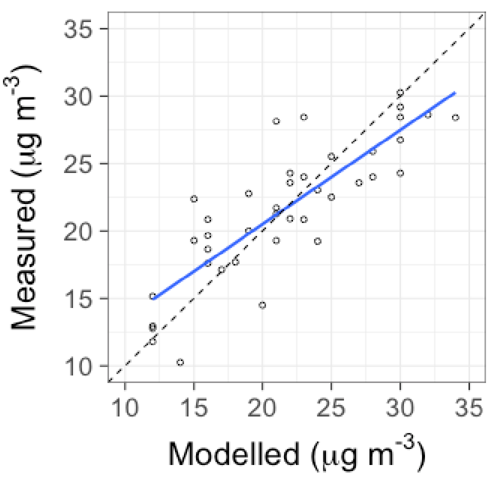

Fig. 3 Modelled vs. measured NO_2 concentrations. The blue solid line represents the best-fit and the dashed line shows the 1:1 line.

The diagnostic tests conformed to the requirements for regression analyses, with VIF's below 3 (Table 3), and no high-leverage points or outliers (max. Cook's D = 0.27) observed. The Moran's *I* showed no spatial autocorrelation between the residuals (p = 0.915). NO_2 concentrations, as well as the residuals, were normally distributed (Shapiro Wilk, $p > 0.05$). The mapped NO_2 concentrations indicate high NO_2 concentrations underneath building awnings, which may explain the higher concentrations visible adjacent to the road rather than on the road. The presence of awnings combined with the density of bus stops also likely explains the spatial variability in NO_2 concentrations along Dominion Rd (Fig. 2).

Table 3. Final LUR model.

LUR model	β	Std. Error	*p*-value	VIF	Proportional contribution[1)] (%)
Intercept	17.210	2.131	< 0.01		
Distance to major road	-0.055	0.013	< 0.01	1.57	0.78
Nr. of bus stops within a 100 m buffer	1.400	0.687	0.056	1.44	16.52
Awnings	5.436	1.815	< 0.01	1.15	79.08
Street width	0.248	0.075	< 0.01	1.04	3.61
Adj. R^2	0.66				
R^2	0.70				
RMSE (μg m^{-3})	3.317				
MSE-R^2	0.71				

[1)] β * (90^{th} percentile – 10^{th} percentile)

Table 4. Cross-validation results.

Validation method[1)]	RMSE	R^2	MSE	MSE-R^2	Iterations
LOOCV	3.839	0.60	14.738	0.61	-
LGOCV	3.886	0.65	15.101	0.60	20

[1)] LOOCV = Leave-one-out cross-validation; LGOCV = grouped (leave-30%-out) cross-validation.

A limitation of the LUR model presented here, and microscale models in general, is the availability of traffic data at sufficient spatial and temporal scale. Google Maps traffic information only gives information about the typical traffic flow and is categorised into only four categories. Given the high spatial variability of air pollutants at the microscale future studies should also test the accuracy of the modelled NO_2 concentrations mapped in Fig. 2 and how these agree with exposure measurements.

Within-city transferability of LUR models

The model developed in this study had relatively good performance when scaled up and applied to data previously reported at local and city scales in Auckland's CBD (Miskell et al., 2015), with an adjusted R^2 of 0.57 and 0.76 and a Spearman rank correlation of 0.68 and 0.89, respectively (Table 5). Interestingly, the model performed better at the city scale, which was also the case when insignificant (p-value > 0.1) predictor variables (all except number of bus stops within 100 m) were removed (adj. R^2 = 0.76, RMSE = 2.89). A slightly lower adjusted R^2 and Spearman rank correlation (0.68) was obtained when the microscale model was applied to the local scale dataset in the CBD, but apart from street width all predictor variables were significant (Table 5). If street width were removed following the ESCAPE protocol, the adjusted R^2 was 0.54 and the RMSE was 3.891. This suggests that unlike standard LUR models which often perform poorly when applied to different areas of the city or scaled down, the microscale variant developed here has reasonably good transferability in terms of both space and scale.

The multi-scale model presented by Miskell et al. (2015) also showed poor results when scaled down to microscales and applied to the data collected in this study. It could only explain 35% of the variability in NO_2 concentrations along Dominion Road, and distance to traffic light was not a significant predictor (*p*-value > 0.1) (Table 5). Thus, although the multi-scale model performed well for areas within the CBD it was not able to capture the variability of NO_2 concentrations outside the CBD where building and road configurations can be different (less and wider spaced traffic lights, lower buildings, etc.). What is interesting to note is that the only variable that was relevant in the multiscale model at all scales within and outside the CBD is the number of bus stops within 100 m. In Auckland, buses are almost exclusively run by diesel, which is the main source of NO_2 in Auckland. At smaller spatial scales (local and microscale) dispersion variables, such as presence of awnings, also becomes relevant.

Table 5. External validation of the Dominion Road LUR model and performance of multi-scale (local/city scale) model applied at different scales (Miskell et al., 2015). (Attached at the end of the manuscript)

Implications

The findings from this study have important implications for urban development indicating the importance of considering street and building configuration to minimize individual exposure to traffic related air pollutants. This might involve limiting the use of awnings near busy roads or developing pedestrian areas and walkways away from multi-lane roads. The strong influence of bus stops on NO_2 concentrations also supports the need to introduce electric and hybrid buses, which will be trialled in Auckland in 2017 (Auckland Council, 2017). A recent study in

Singapore also identified bus stops as hotspots of individual exposure and suggested to set bus shelters further away from the major road (Velasco and Tan, 2016). The results further indicate that models developed at city scales may not be able to capture the small scale variability in NO_2 concentrations along the road and that there is a need to consider dispersion features such as presence of awnings, supporting previous findings by Tang et al. (2013). The advantage of microscale models as presented in this study is the potential of estimating individual or population exposure at urban hotspots. These results may be used to assess differences in exposure depending on which side of the street pedestrians are walking on or to identify route choices with minimal exposure to traffic related air pollutants. Such detail is generally not available from LUR models developed at city or regional scales, which are commonly used to estimate individual exposure based on the residential address (Jerrett et al., 2007; Urman et al., 2014).

While the model developed at the microscale performed relatively well when transferred to different spatial scales in the city, it showed that the importance of predictor variables changes depending on location and spatial scale, limiting within-city transferability. Thus, as has been suggested by previous studies (Allen et al., 2011; Marcon et al., 2015; Vienneau et al., 2010), LUR models provide better results when developed locally and caution is required when transferring LUR models, even within cities, unless street and building configurations are similar. This also has some important implications for air quality monitoring, suggesting that future research should focus on monitoring NO_2 concentrations at a high spatial resolution within urban environments in order to obtain representative small-scale variability of pollutant concentrations at urban hotspots.

Conclusions

In this study, we assessed the performance of LUR modelling at the microscale representative of a busy street in Auckland, New Zealand. We showed that NO_2 concentrations can be modelled at this scale with a good performance of the model (adj. R^2 = 0.66, RMSE = 3.317μg m^{-3}). Unlike LUR models developed at the city or regional scale, this study has shown that building and street configuration, such as presence of awnings, is important predictors for NO_2 concentrations at the street level. The microscale model performed well when transferred to the city and local scale within Auckland's CBD, although the only significant predictor variables at all spatial scales was the number of bus stops within 100 m. The study indicated implications related to urban development, exposure assessments at urban hotspots and air quality monitoring, highlighting the importance of high density measurements or micro and local scale models to capture the small scale variability in NO_2 concentrations.

Acknowledgments

The authors thank Auckland Council for providing the air quality data obtained from the routine monitoring network. We also thank Auckland Transport for permission to use lamp-posts for tube installation along Dominion Road and Alana Chester for helping with the laboratory analysis and preparation of figures.

References

AEA Energy and Environment, 2008. Diffusion tubes for ambient NO_2 monitoring: Practical guidance for laboratories and users (Report to DEFRA and the Devolved Administrations), Didcot, UK.

Allen, R.W., Amram, O., Wheeler, A.J., Brauer, M., 2011. The transferability of NO and NO_2 land use regression models between cities and pollutants. Atmospheric Environment 45, 369 - 378.

Amini, H., Taghavi-Shahri, S.M., Henderson, S.B., Hosseini, V., Hassankhany, H., Naderi, M., Ahadi, S., Schindler, C., Kunzli, N., Yunesian, M., 2016. Annual and seasonal spatial models for nitrogen oxides in Tehran, Iran. Sci Rep 6, 32970.

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. Environ Sci Technol.

Auckland Council, 2017. Auckland Transport to trial electric buses.

Auckland Transport, 2017. Dominion Road upgrade.

Beelen, R., Hoek, G., Vienneau, D., ... , Brunekreef, B., de Hoogh, K., 2013. Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. Atmospheric Environment 72, 10-23.

Borge, R., Narros, A., Artinano, B., Yagüe, C., Gomez-Moreno, F.J., de la Paz, D., Roman-Cascon, C., Diaz, E., Maqueda, G., Sastre, M., Quaassdorff, C., Dimitroulopoulou, C., Vardoulakis, S., 2016. Assessment of microscale spatio-temporal variation of air pollution at an urban hotspot in Madrid (Spain) through an extensive field campaign. Atmospheric Environment 140 432 - 445.

Briggs, D.J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., Smallbone, K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. The Science of the Total Environment 253, 151 - 167.

Brunekreef, B., 2008. ESCAPE Study Manual, in: Institute of Risk Assessment Sciences, U.U. (Ed.), The Netherlands.

Chappell, P.R., 2014. The Climate and Weather of Auckland, in: NIWA (Ed.), 2 ed.

Dirks, K.N., Sharma, P., Salmond, J.A., Costello, S.B., 2012. Personal concentration to air pollution for various modes of transport in Auckland, New Zealand. Open Atmospheric Science Journal 6, 84-92.

Dirgawati, M., Barnes, R., Wheeler, A.J., Arnold, A.-L., McCaul, K.A., Stuart, A.L., Blake, D., Hinwood, A., Yeap, B.B., Heyworth, J.S., 2015. Development of Land Use Regression models for predicting exposure to NO_2 and NO_x in Metropolitan Perth, Western Australia. Environmental Modelling & Software 74, 258-267.

Eeftens, M., Beekhuizen, J., Beelen, R., Wang, M., Vermeulen, R., Brunekreef, B., Huss, A., Hoek, G., 2013. Quantifying urban street configuration for improvements in air pollution models. Atmospheric Environment 72, 1 - 9.

Ghassoun, Y., Ruths, M., Lowner, M.O., Weber, S., 2015. Intra-urban variation of ultrafine particles as evaluated by process related land use and pollutant driven regression modelling. Sci Total Environ 536, 150-160.

Grange, S.K., Dirks, K.N., Costello, S.B., Salmond, J.A., 2014. Cycleways and footpaths: What separation is needed for equivalent air pollution dose between travel modes? Transportation Research Part D: Transport and Environment 32, 111-119.

Gurung, A., Levy, J.I., Bell, M.L., 2017. Modeling the intraurban variation in nitrogen dioxide in urban areas in Kathmandu Valley, Nepal. Environ Res 155, 42-48.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution Atmospheric Environment 42, 7561 - 7578.

Janhäll, S., 2015. Review on urban vegetation and particle air pollution - deposition and dispersion. Atmospheric Environment, 130 - 137.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 15, 185-204.

Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N., Finkelstein, M.M., 2007. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. J Toxicol Environ Health A 70, 200-212.

Kaur, S., Nieuwenhuijsen, M.J., 2009. Determinants of personal concentration to PM2.5, ultrafine particle counts, and CO in a transport microenvironment. Environ. Science Technology 43, 4737-4743.

Liu, C., Henderson, B.H., Wang, D., Yang, X., Peng, Z.R., 2016. A land use regression application into assessing spatial variation of intra-urban fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) concentrations in City of Shanghai, China. Sci Total Environ 565, 607-615.

Marcon, A., de Hoogh, K., Gulliver, J., Beelen, R., Hansell, A.L., 2015. Development and transferability of a nitrogen dioxide land use regression model within the Veneto region of Italy. Atmospheric Environment 122, 696 - 704.

McNabola, A., Broderick, B.M., Gill, L.W., 2009. A principal components analysis of the factors effecting personal concentration to air pollution in urban commuters in Dublin, Ireland. J. Environ. Sci. Health 44, 12, 1219-1226.

Meng, X., Chen, L., Cai, J., Zou, B., Wu, C.F., Fu, Q., Zhang, Y., Liu, Y., Kan, H., 2015. A land use regression model for estimating the NO_2 concentration in Shanghai, China. Environ Res 137, 308-315.

Miskell, G., 2013. Identifying urban design details affecting the spatial variability of micro-scale ambient nitrogen dioxide for the Auckland CBD area, School of Environment. University of Auckland, Auckland, NZ.

Miskell, G., Salmond, J., Longley, I., Dirks, K.N., 2015. A Novel Approach in Quantifying the Effect of Urban Design Features on Local-Scale Air Pollution in Central Urban

Areas. Environ Sci Technol 49, 9004-9011.

Mölter, A., Sindley, S., de Vocht, F., Agius, R., Kerry, G., Johnson, K., Ashmore, M., Terry, A., Dimitroulopoulou, S., Simpson, A., 2012. Performance of a microenvironmental model for estimating personal NO_2 exposure in children. Atmospheric Environment 51, 225 - 233.

Morgenstern, V., Zutavern, A., Cyrys, J., Brockow, I., Gehring, U., Koletzko, S., Bauer, C.P., Reinhardt, D., Wichmann, H.E., Heinrich, J., 2007. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occup Environ Med 64, 8-16.

Salmond, J.A., McKendry, I.G., 2009. Chapter 2 Influences of Meteorology on Air Pollution Concentrations and Processes in Urban Areas, Air Quality in Urban Environments. The Royal Society of Chemistry, pp. 23-41.

Salmond, J.A., Williams, D.E., Laing, G., Kingham, S., Dirks, K., Longley, I., Henshaw, G.S., 2013. The influence of vegetation on the horizontal and vertical distribution of pollutants in a street canyon. Sci Total Environ 443, 287-298.

Senaratne, I., Shooter, D., 2004. Elemental composition in source identification of brown haze in Auckland, New Zealand. Atmospheric Environment 38, 3049-3059.

Shi, Y., Lau, K.K., Ng, E., 2016. Developing Street-Level PM2.5 and PM10 Land Use Regression Models in High-Density Hong Kong with Urban Morphological Factors. Environ Sci Technol 50, 8178-8187.

Statistics New Zealand, 2013. 2013 census usually resident population counts.

Stedman, J.R., Vincent, K.J., Campbell, G.W., Goodwin, J.W.L., Downing, C.E.H., 1997. New high resolution maps of estimated background ambient NO_x and NO_2 concentrations in the UK. Atmospheric Environment 31, 3591 - 3602.

Tang, R., Blangiardo, M., Gulliver, J., 2013. Using building heights and street configuration to enhance intraurban PM_{10}, NO(x), and NO_2 land use regression models. Environ Sci Technol 47, 11643-11650.

Urman, R., McConnell, R., Islam, T., Avol, E.L., Lurmann, F.W., Vora, H., Linn, W.S., Rappaport, E.B., Gilliland, F.D., Gauderman, W.J., 2014. Associations of children's lung function with ambient air pollution: joint effects of regional and near-roadway pollutants. Thorax 69, 540-547.

Vardoulakis, S., Solazzo, E., Lumbreras, J., 2011. Intra-urban and street scale variability of BTEX, NO_2 and O_3 in Birmingham, UK: Implications for exposure assessment. Atmospheric Environment 45, 5069-5078.

Velasco, E., Tan, S.H., 2016. Particles exposure while sitting at bus stops of hot and humid Singapore. Atmospheric Environment 142, 251-263.

Vienneau, D., de Hoogh, K., Beelen, R., Fischer, P., Hoek, G., Briggs, D., 2010. Comparison of land-use regression models between Great Britain and the Netherlands. Atmospheric Environment 44, 688 - 696.

Weissert, L.F., Salmond, J., Friedel, A., LaFave, M., Schwendenmann, L., 2015. Temporal and spatial patterns of carbon dioxide mixing ratios in a subtropical urban environment during spring. Weather and Climate 35, 13 - 32.

Xie, M., Zhu, K., Wang, T., Chen, P., Han, Y., Li, S., Zhuang, B., Shu, L., 2016. Temporal characterization and regional contribution to O and NO at an urban and a suburban site in Nanjing, China. Sci Total Environ 551-552, 533-545.

Table 1. Predictor variables used for the model development with defined buffer sizes, impact and expected direction of effect.

Variable	Variable description	Impact	Unit	Direction of effect	Buffer size radius (m)
Kerb length	Length of kerb within buffer size	Modifier	m	+	10, 25, 50, 100
Distance to traffic light	Distance to nearest traffic light	Source	m	-	NA
Distance to intersection	Distance to nearest intersection	Source	m	-	NA
Distance to major road	Distance to major road (including arterial roads going in and out of the city, major urban roads and motorways)	Source	m	-	NA
Height-width ratio	Height of surrounding buildings divided by street width	Modifier	ratio	+	NA
Car parks	Sum of large car parks within a buffer size	Source	count	+	25, 50, 100
Nr. of lanes	Sum of traffic lanes along nearest road	Source	count	+	NA
Distance to bus stop	Distance to nearest bus stop	Source	m	-	NA
Nr. of bus stops	Sum of bus stops within buffer size	Source	count	+	10, 25, 50, 100
Nr. of bus lanes	Sum of bus lanes within buffer size	Source	count	+	10, 25, 50, 100
Tree density	Tree density within buffer size	Source	Trees/m^2	-	10, 25, 50, 100
Distance to tree	Distance to nearest tree	Modifier	m	+	NA
Vegetation	Presence of tree within buffer size	Modifier	Y = 1, N = 0	-	10, 25, 50, 100
Side of street	Left or right with a north/east orientation	Modifier	L = 1, R = 0	NA	NA
Street width	Distance from one side of the street to the other	Modifier	m	+	NA
Awnings	Presence of awnings within 10 m	Modifier	Y = 1, N = 0	+	NA
Building footprint	Area of buildings within a buffer size	Modifier	m	+	25, 50, 100
Morning rush hour traffic	Average traffic flow estimated from GoogleMaps during weekday morning rush hour	Source	1: Slowest, 4: Fastest	+	NA
Midday traffic	Average weekday traffic flow at midday estimated from GoogleMaps	Source	1: Slowest, 4: Fastest	+	NA
Evening rush hour traffic	Average traffic flow estimated from GoogleMaps during weekday evening rush hour	Source	1: Slowest, 4: Fastest	+	NA

Table 5. External validation of the micro scale Dominion Road LUR model and performance of multi-scale (local/city scale) model applied at different scales (Miskell et al., 2015). [1)]

Model	Dataset	Nr. Locations	Spatial scale	Variables[1)]	Adj. R^2	RMSE ($\mu g\ m^{-3}$)	Spearman rank correlation
Micro scale (Dominion Road)	Dominion Rd	40	Micro	Distance to major rd*** Nr. of bus stops within 100 m Awnings** Street width**	0.66	3.317	0.84
	CBD	62	Local	Distance to major rd* Nr. of bus stops within 100 m*** Awnings*** Street width	0.57	4.852	0.68
	CBD	21	City	Distance to major rd Nr. of bus stops within 100 m*** Awnings Street width**	0.76	2.758	0.89
Multi-scale (Miskell et al., 2015)	Dominion Rd	40	Micro	Nr. of lanes Nr. of bus stops within 100 m** Distance toward traffic light	0.35	4.647	0.59
	CBD	62	Local	Nr. of lanes*** Nr. of bus stops within 100 m*** Distance toward traffic light**	0.54	5.058	0.67
	CBD	21	City	Nr. of lanes*** Nr. of bus stops within 100 m*** Distance toward traffic light*	0.79	2.661	0.91

* *p*-value < 0.1, ** *p*-value < 0.05, *** *p*-value < 0.01