Modelling the vertical gradient of nitrogen dioxide in an urban area

Marloes Eeftens[1,2*], Danyal Odabasi[1,2], Benjamin Flückiger[1,2], Mark Davey[1,2], Alex Ineichen[1,2], Christian Feigenwinter[2], Ming-Yi Tsai[1,2,3]

Affiliations:

1. Swiss Tropical and Public Health Institute, Basel, Switzerland
2. University of Basel, Basel, Switzerland
3. Dept. of Environmental and Occupational Health Sciences, University of Washington, Seattle, USA

Corresponding author:

Marloes Eeftens

E-mail: marloes.eeftens@swisstph.ch

Department of Epidemiology and Public Health

Swiss Tropical and Public Health Institute

Socinstrasse 57, P.O. Box 4002 Basel, Switzerland

Telephone: +41 61 284 87 25

Fax: +41 61 284 81 05

Keywords:

Air pollution, Vertical gradient, Land Use Regression, NO_2, decay

Sources of financial support:

This work was supported by the Forschungsfonds University of Basel.

Competing interest:

The authors declare they have no competing financial interest.

Abbreviations:

LUR: Land use regression

LOOCV: Leave-one-out cross-validation

NO_2: Nitrogen dioxide

RMSE: Root mean squared error

SVF: Sky view factor

Acknowledgments:

We thank all the study participants who provided a place to perform our measurements.

Abstract

Introduction

Land use regression models environmental predictors to estimate ground-floor air pollution concentration surfaces of a study area. While many cities are expanding vertically, such models typically ignore the vertical dimension.

Methods

We took integrated measurements of NO_2 at up to three different floors on the facades of 25 buildings in the mid-sized European city of Basel, Switzerland. We quantified the decrease in NO_2 concentration with increasing height at each facade over two 14-day periods in different seasons. Using predictors of traffic load, population density and street configuration, we built conventional land use regression (LUR) models which predicted ground floor concentrations. We further evaluated which predictors best explained the vertical decay rate. Ultimately, we combined ground floor and decay models to explain the measured concentrations at all heights.

Results

We found a clear decrease in mean nitrogen dioxide concentrations between measurements at ground level and those at higher floors for both seasons. The median concentration decrease was 8.1% at 10m above street level in winter and 10.4% in summer. The decrease with height was sharper at buildings where high concentrations were measured on the ground and in canyon-like street configurations. While the conventional ground floor model was able to explain ground floor concentrations with a model R^2 of 0.84 (RMSE 4.1 µg/m³), it predicted measured concentrations at all heights with an R^2 of 0.79 (RMSE 4.5 µg/m³), systematically overpredicting concentrations at higher floors. The LUR model considering vertical decay was able to predict ground floor and higher floor concentrations with a model R^2 of 0.84 (RMSE 3.8 µg/m³) and without systematic bias.

Discussion

Height above the ground is a relevant determinant of outdoor residential exposure, even in medium-sized European cities without much high-rise. It is likely that conventional LUR models overestimate exposure for residences at higher floors near major roads. This overestimation can be minimized by considering decay with height.

Introduction

In epidemiological studies done in large populations, exposure to air pollution is typically modelled. Land use regression (LUR) takes actual measurement data obtained from a limited number of monitors distributed throughout the study area and environmental information such as land use, proximity to roads, traffic intensity population density and other predictors, typically obtained from geographical information systems, aiming to predict detailed spatial contrasts (de Hoogh et al., 2016; de Hoogh et al., 2017; Eeftens et al., 2016; Hoek et al., 2008). Typical air pollution LUR models take into account horizontal proximity to a source, but study participants living in the same building but on different floors are attributed the same exposure estimate, which may not be accurate. It is possible that floor of residence is related to health, as suggested by a recent study which found that higher floors have lower mortality rates in Switzerland(Panczak et al., 2013).

Numerous studies have measured vertical gradients for individual streets as was recently reviewed by Sajani et al., (2018). Typically, studies found that in urban areas, levels of many pollutants decrease with increasing height, among them NO_2 (Lufthygieneamt beider Basel, 1993; Meng et al., 2008; Tsai and Chen, 2004; Zauli-Sajani et al., 2018), a marker for fossil fuel combustion by motorized traffic. The majority of these studies have focused on Asian cities which differ in pollutant sources, population density, and street setting (e.g., many high-rise buildings) and are therefore not directly comparable to moderate sized cities in Europe. Moreover, these studies typically measured the vertical gradient at a single or only at several buildings, allowing for the calculation of only one or several decay rates. To date, we only know of three other LUR studies which included height above the ground (floor level) as a model predictor, all are from Asia.(Barratt et al., 2018; Ho et al., 2015; Wu et al., 2014) In two Taiwanese studies, including floor level added substantially to the predictive power of a LUR model for $PM_{2.5}$ (mass of particulate matter with a diameter <2.5 micrometer), and several of its elemental constituents.(Ho et al., 2015; Wu et al., 2014) Another study from Hong Kong calculated vertical decay rates for black carbon and $PM_{2.5}$, but due to a limited number of paired vertical sites, it was not possible to let the decay rate vary depending on local conditions.(Barratt et al., 2018) Instead, decay rates were assumed to be the same across the entire region in the LUR model in the model application for epidemiological exposure estimation.(Barratt et al., 2018) This study aims to characterize vertical gradients in concentrations of NO_2 at many locations and investigate how these gradients may be related to geographical predictors and street configuration. We present and evaluate a three dimensional LUR model for NO_2 in which vertical decay is fully integrated.

Methods

Study design

We selected 25 buildings, adjacent to a diverse variety of streets in the canton of Basel Stadt, Switzerland, which had a population of 198'000 in 2016.(Statistisches Amt Basel Stadt, 2017) We aimed to represent the full range of traffic densities and street configurations, and over-represent busy and canyon-like streets, where we expected higher concentrations (Figure 1). Measurements took place simultaneously at all sites over a period of two weeks in the winter (25 February to 11 March 2016) and were repeated in summer (13 June to 29 June 2016). Samplers were installed and collected over two consecutive days each round. On each building, we installed three integrated NO_2 passive samplers from the company PASSAM AG(Passam - Laboratory for environmental analysis and air pollution) in a vertical line by attaching the sampler tubes in shelters on the external wall or windows facing the street. Buildings differed in their facade structure, so having consistently the same spacing between ground floor and the higher samplers was not feasible. Instead, the lowest of the three measurements was always at the ground floor, between 1.8 and 2.9 meter above ground and the other two samplers were placed at floors 1 to 6, up to 20 meters above street level, attempting to maximize the vertical spread, but constrained by the residents' willingness to participate. The exact installation height of each sampler was determined using a Bosch PLR 50 Laser Measure. The height of the building to which the samplers were attached, the height of the building on the opposite side of the street and the street width were measured to calculate the canyon aspect ratio: the average building height on both sides of the street (h1 and h2) divided by the street width (w): (h1+h2)/(2*w). During both summer and winter, off-peak (09:00-16:00) traffic density at each site was counted for 30 Minutes. To assess the quality of the measurements, we took a total of 8 field blanks and 12 duplicates during the winter and summer seasons.

Geographical predictor data

A list of 66 spatial predictors were derived in different buffers, following previous land use regression studies in small geographic areas.(Cyrys et al., 2012; Eeftens et al., 2013; Eeftens et al., 2012; Eeftens et al., 2016) The following geographic information was available:

- Information on population density was available from the 100x100m grid the Federal Statistical Office published in 2011. The number of inhabitants was derived in six different buffers of 100, 250, 500, 750, 1000 and 5000m.
- Information on land cover was obtained from the Corine Land Cover 2006 Raster data.(European Environment Agency (EEA)) The original 44 classes were summarized into 7 groups: airport, industry, natural, port, residential, water and urban green spaces, following earlier publications.(Beelen et al., 2009; Beelen et al., 2013; Eeftens et al., 2012) We derived the total area of each land use (in m^2) in six different buffers of 100, 250, 500, 750, 1000 and 5000m.
- A digital national road network (Vector 25) was available from SwissTopo. The SonBase Noise Database, available from the Federal Office for the Environment, provided modelled traffic intensity.(Swiss Bundesamt für Umwelt (BAFU), 2009) The length of each road segment in a buffer was multiplied by the traffic intensity on that segment, after which all segments were summed to obtain the traffic load for that buffer. Correlation between traffic load and locally conducted traffic counts validated this measure (Online Supplement 1). Road length (in m) and traffic load (in vehicles day^{-1}*m) were calculated in eight different buffers of 25, 50, 75, 100, 250, 500, 750, 1000m.

- The Institute of Meteorology, Climatology and Remote Sensing at the University of Basel provided a detailed 1x1m map of the sky view factor (SVF)(Lindberg and Grimmond, 2010), a unit-less measure for the canyon-like structure of urban streets, which was previously shown to improve air pollution models (Eeftens et al., 2013). The higher the sky view factor, the more of the sky is visible, and thus the higher the air exchange rate between the canyon and the air above. Correlations between aspect ratio (measured at each site) and SVF validated this measure (Online Supplement 1).
- Altitude information (in m above sea level) was available from the DHM25 of SwissTopo and was extracted for each point.(Bundesamt für Landestopographie, 2001)

Altitude, natural and urban green land use and sky view factor were assumed to have a negative effect on NO_2 concentration with increasing value and no a priori assumptions were made about the effect of water. All other predictors were assumed to have a positive effect on NO_2 concentration. No a priori assumptions were made on how these predictors affected the vertical decay.

Land use regression modelling

Following an earlier publication(Eeftens et al., 2016), the calculated predictors were screened prior to their evaluation in the LUR model. Predictors were discarded if fewer than five sites had a value different from the most frequently occurring value, if the maximum value was higher than P90+3*(P90-P10) or the minimum value was lower than P10-3*(P90-P10). All these criteria indicate an abnormal distribution and a model including such variables would likely yield unstable coefficients. We followed the same stepwise variable selection procedures as were used by several earlier studies.(Eeftens et al., 2012; Eeftens et al., 2016) Criteria for the inclusion of additional variables were: 1) an improvement in adjusted r-squared by at least 0.01; 2) ensuring the a priori defined direction of effect for all variables; 3) a p value smaller than 0.05; 4) a maximum Cook's D value of 1 for all sites; 5) a maximum Variance Inflation Factor (VIF) of 3 to minimize collinearity; 6) no change in the direction of previously included variables.

Firstly, NO_2 concentration at the ground floor was modelled for both summer and winter separately. We then combined both seasonal ground floor models into a mixed model by including an indicator for season and a random effect for site.

Vardoulakis et al. (2003) suggested that vertical profiles typically decrease exponentially with height. The concentration at any height can thus be determined by the ground concentration, the height and a decay constant *k* following Equation 1. Since we had multiple measurements per building, we fit an exponential model to determine *k* for each measurement site and for the winter and summer seasons. Deriving a second set of models, we then selected the predictors which best determined *k* using the same variable selection method described above, adding a random effect for site. Seasonal models were combined into a model for both seasons, originally retaining all of the selected spatial predictors and the random intercept for site and adding an indicator for season. Any redundant spatial variables were excluded if the increased in AIC was less than 2, indicating that they did not substantially improve the combined model.

Equation 1: $Concentration_{height} = Concentration_{ground} \times e^{k \times height}$.

As a third step, the ground floor model and the model for *k* were combined by plugging the models for ground floor concentration and *k* into Equation 1, thereby extending the LUR to allow for the calculation of concentrations at any height and location in the city.

Model diagnostics

We derived model R^2 and RMSE, as well as cross-validation R^2 and RMSE for all models. We further validated all models by iteratively leaving each of the 25 buildings out entirely and predicting it using the remaining 24 buildings. For models M1 and M2, where we had a single observation per building, this was done using leave-one-out cross-validation (LOOCV), iteratively leaving out each one of 25 sites and predicting its concentration based on the coefficients obtained from a model fit on the remaining 24 sites (Table 1). For the mixed models (M3-M9), this cross-validation was done by simultaneously leaving out all observations from that building from both seasons and using only the fixed effects obtained from the remaining sites to predict up to six concentrations for that building (two seasons * up to three different heights).

For the four models which were built using stepwise selection (M1, M2, M4 and M5), we assessed the chance that similarly high model R^2 values could be obtained solely by chance, as suggested by Basagaña et al., 2012.(Basagaña et al., 2012) Using the dependent variables from the study (NO_2 or k), we simulated the model building process 10,000 times using unique sets of randomly generated "predictor" variables, from a normal distribution with mean = 0 and standard deviation of 1. For each model, we then simulated the chance of finding a model based on random variables with an equally high or higher R^2 than the model we selected.

Used software

All geographical predictor variables were derived in ArcGIS10. All statistical analyses were done in R version 3.3.2, using packages lme4 for the mixed models and ggplot for plotting.(R Core Team)

Results

Measurements

We found that NO_2 concentrations were substantially higher in the winter season with a ground-floor median of 35.1 µg/m³ (InterQuartile Range (IQR): 32.5-43.9) than in the summer season (median 23.2 µg/m³, IQR: 20.2-34.8) (Figure 2). All field blanks indicated concentrations under or at the detection limit (≤0.4 µg/m³) indicating negligible contamination. Primary and duplicate samples were highly correlated (R^2: 0.94) and showed a low mean absolute difference of 1.6 µg/m³, indicating high reproducibility (Online Supplement 2). NO_2 concentrations generally decreased with increasing height above the ground in both winter and summer. The decay constant *k* was calculated at a median of -0.0084 m^{-1} in winter (IQR: -0.011 m^{-1} to -0.00068 m^{-1}) and at a median of -0.011 m^{-1} (IQR: -0.017 m^{-1} to -0.0015 m^{-1}) in summer. This translates to a median concentration reduction of 8.1% at 10m above street level (IQR: 0.7% to 10.2%) in winter and 10.4% (IQR: 1.5% to 15.2%) in summer (see Online Supplement 3).The decrease with height was sharper at buildings where higher NO_2 concentrations were measured at ground level: for an IQR increase in ground level NO_2, *k* decreased by 0.011 m^{-1} in winter and by 0.0078 m^{-1} in summer. The spatial contrast in NO_2 was highest at the ground floor, and decreased with increasing height (Figure 2). The average temperature in winter was 4.0 °C with a relative humidity of 77%, in summer this was 17.8 °C and 75% relative humidity.

LUR models

Out of the 67 predictors derived for the study, 27 were dropped in the variable screening process, leaving 40 eligible for evaluation. The screening process mostly flagged variables with a limited number of <5 points with a different value than the most common; e.g. few sites were within 100, 250 or 500m of an airport, port or industrial area and thus these variables were mostly 0. The on-site measurements of the aspect ratio and the SVF derived from GIS were highly correlated, as were the manual traffic counts and the GIS derived traffic load in 25m (R^2 = 0.60 and R^2 = 0.81 respectively, Online Supplement 1).

The ground floor models identified for the winter (M1) and summer (M2) included the same two predictor variables (traffic load in a 25m buffer and residential land use in a 5000m buffer) with very similar coefficients, and yielded similar R^2's of 0.71 and 0.79, respectively (Table 1). The combined mixed model for both seasons (M3, with random effects for site) showed a large increase of 11.9 $\mu g/m^3$ for season and again yielded similar coefficients to the season-specific models and a slightly higher R^2 of 0.84. No interactions between season and GIS variables were significant. Ground floor models performed reasonably well in leave-one-out (M1 and M2) cross validation, yielding a 10-11% lower predictive power and slightly higher similar RMSE's of 4.9 $\mu g/m^3$ and 4.6 $\mu g/m^3$. Leave-one-group-out (M3) cross validation on the mixed model (M3), yielded a similar RMSE of 4.5 $\mu g/m^3$ and only slightly lower predictive power, due to the relatively large but easy to predict difference in season.

The decay constant did not differ notably between the seasons: a median decay of -0.0084m^{-1} was calculated for winter and -0.011m^{-1} for summer (Figure 3). Typically, *k* was lower for sites surrounded by high traffic load and in streets with a low SVF. The winter model for the decay constant *k* (M4) explained a substantial part of the variability in *k* (R^2 of 0.69), and included variables related to traffic load (25m) and road length (100m) in the close vicinity, residential land use (250m), as well as sky view factor (Table 1). The summer model (M5) did not perform as well (R^2 of 0.37), but included the same traffic load (25m) predictor and sky view factor, which were selected for both models with very similar coefficients, indicating that the busier and the more canyon-like the street, the lower *k* was. When the season-specific models were combined (M6), season was evaluated as an additional predictor but was not included because it was not significant. Similarly, residential land use and road length were no longer significant in the model for both seasons, and were dropped. The combined-season model (M6) for *k* had a lower R^2, but performed similarly well in leave-one-group-out cross-validation.

For the models M7, M8 and M9, R^2 was very similar to the ground level models (R^2's of 0.67, 0.78, 0.84, respectively). Leave-one-group-out cross-validation predictions were able to estimate the variability in measured concentrations well (Table 1).

For the four models which were built using stepwise selection (M1, M2, M4 and M5), we repeatedly simulated the chance of obtaining a model with at least the same R^2 by evaluating 40 randomly generated predictor variables. For the ground floor models M1 and M2, only 2.3% and 0.1% of simulated models had R^2 values equal to or higher than 0.71 and 0.79 respectively, none of which included as few as two predictors. This indicated that the chance of obtaining similarly good ground floor models by chance was very small. The winter model for *k* (M4) had an R^2 of 0.69, and a similarly low chance of 2.6% (1.2% with 4 variables or fewer) of obtaining this result by chance. However, a model similar to the summer model for *k* (M5), which included 2 predictors and yielded an R^2 of 0.37, was obtained by chance in 26% of simulations (12% with 2 predictors or fewer). This indicated that it was more likely that the M5 model was due to chance.

LUR predictions

Figure 4 shows the correlation between the measured NO_2 concentrations and the predictions as obtained from the models M3 (Figure 3A) and M9 (Figure 3B). Model predictions by the "conventional" ground floor LUR model (M3) differ by building but are otherwise the same, regardless of the height. In contrast, the predictions by model M9 show a clear effect of height within each building, and that the decay with height is site-dependent. Predictions from the model M9, which considers the height, agree better with the measured data, showing a correlation of 0.84 with the measured data (versus 0.79 for M3) and a lower RMSE of 3.78 µg/m³ (versus 4.25 µg/m³ for M3). Applying the winter and summer ground floor models (M1 and M2) to predict all 72 and 69 sites measured in the respective seasons, regardless of their height, yielded R^2's of 0.62 (RMSE 4.13 µg/m³) for winter and 0.71 (RMSE 4.09 µg/m³) for summer. Both of the seasonal models considering height (M6 and M7) also explained more variability with R^2's of 0.67 (RMSE 3.86 µg/m³) for winter and 0.78 (3.63 µg/m³) for summer.

Figure 5 shows the prediction errors by both models M3 and M9 as a function of floor of residence. At the ground floor level, both models perform similarly well, but the ground floor model M3 is less able to accurately predict concentrations at higher floors, showing a significant overestimation for floors for the 3rd, 4th and higher floors. The model M9 which includes height as a predictor is able to derive unbiased concentration estimates for both the ground floor and all higher floors.

Discussion

We found a clear decrease in NO_2 concentration with increasing height in both winter and summer, which was more pronounced in streets with high ground-floor concentrations. We showed that conventional "ground floor" LUR models can be extended to include height, by introducing a decay constant, *k*, which—like the ground floor concentration—depends on locally important environmental characteristics, such as the street configuration (e.g., SVF) and the local traffic load. We showed that a ground floor LUR model for NO_2 developed in the conventional way over-predicted concentrations at higher floors, whereas our LUR model considering concentration decay with height was able to predict ground floor and higher floor concentrations with similar accuracy.

Determinants of vertical decay

The study confirms earlier suggestions from a review by Hoek et al (2008) that LUR models give biased estimates at higher floors for buildings near major roads, where concentration decrease sharply with height, and that vertical gradients at urban background locations are limited.(Hoek et al., 2008) In addition, this study suggests that the lower SVF (the more canyon-like the street), the sharper the decay with height. This is likely due to a lower air exchange rate between the air inside the canyon and the air above the building canopy, leading to substantial pollutant build-up at lower levels.(Vardoulakis et al., 2003) The median decay constants of $-0.0084m^{-1}$ (winter) and $-0.011m^{-1}$ (summer) were similar to the $-0.012m^{-1}$ applied for NO_2 in Barrat et al. (2018), and those reported by Zauli-Sajani et al. (2018), estimated to be $-0.022m^{-1}$ in winter and $-0.007m^{-1}$ in summer[1]. While all three previous LUR studies which considered height hypothesized that the decay rate may depend on street configuration, they were unable to derive any patterns.(Barratt et al., 2018; Ho et al., 2015; Wu et al., 2014) In 1993, a study by the Lufthygieneamt beider Basel (Air hygiene office of Basel) also reported a decay with height in a canyon-like street.(Lufthygieneamt beider Basel, 1993) An inverted vertical gradient was observed for the heating season in that study, because of the chimneys emitting NO_2 at rooftop level.(Lufthygieneamt beider Basel, 1993) We indeed found that k was slightly higher (closer to 0 m^{-1}) in winter across the whole city, but this was not statistically significant. Although this may suggests that vertical decay is indeed slightly reduced in the wintertime because of additional NO_2 emissions at rooftop level, we did not see a flattening or inversion of the vertical gradient in the winter time as reported in the study from 1993. This could be due to more homes switching to district heating and/or cleaner alternative fuels since 1993 making ground floor emissions from traffic relatively important determinants of vertical decay.

[1] These numbers were not directly reported by Zauli-Sajani et al. (2018), but derived using Equation 1 from the heights (ground ~1.5m, 15m, 26m, 44m, 65m), the winter NO_2 concentrations (70.5 µg/m³, 63.6 µg/m³, 62.9 µg/m³, 38.7 µg/m³, 18.1 µg/m³) and the NO_2 summer concentrations (41.2 µg/m³, 34.4 µg/m³, 32.1 µg/m³, 30.0 µg/m³, 25.3 µg/m³).

Implications for exposure science

Vertical gradients and street configuration are typically thought to be relevant mostly to high-density high-rise cities.(Barratt et al., 2018; Ho et al., 2015; Wu et al., 2014) Our measurements show that the concentration drop in NO_2 is also relevant for medium-sized European cities, and that considering decay with height can also improve LUR models commonly applied to assess residential air pollution exposure in epidemiological studies. Similar vertical gradients may exist in other cities, but this depends on whether (as in Basel) most local air pollutants are emitted by traffic at street level, rather than e.g. by industry or home heating installations, which typically emit at rooftop level. In addition, vertical gradients may depend on weather patterns and local topography. This is one of the first studies looking at vertical gradients; therefore, we advise caution in translating these findings into local policies, or assuming impacts on indoor and personal air pollution exposure.

Strengths and limitations

In this study, we only measured NO_2, which is a good marker for traffic-related air pollutants. It is likely that other air pollutants which decay more slowly with increasing horizontal distance, such as particulate matter smaller than 2.5 or 10 µm ($PM_{2.5}$ or PM_{10}), may also exhibit a less pronounced vertical gradient (Karner et al., 2010). In contrast, pollutants which decrease more rapidly than NO_2 in the horizontal plane, such as CO, NO and ultrafine particles, may do so in the vertical dimension as well due to their chemical transformation and coagulation properties.(Karner et al., 2010)

As we used two-week time-integrated measurements of NO_2 in two different seasons, we were unable to test how the ground floor concentrations and vertical gradients may depend on individual meteorological factors such as precipitation, relative humidity, sunshine hours, temperature, wind speed and wind direction. The two seasons captured substantial differences in many of these weather parameters (see Online Supplement 4). The ground floor concentrations clearly differed substantially between the winter and summer seasons, and hence season is an important predictor for ground floor concentration. However, the median and IQR of *k* only differed slightly by season, and season was not ultimately a significant determinant. This indicates that in our study, *k* was likely not substantially influenced by weather.

Another limitation of the study is its limited sample size (25 buildings) in the horizontal plane, which was due to the limited scope and budget available for the study. Several earlier papers have raised concern about overfitting in LUR models based on small datasets.(Basagaña et al., 2012; Wang et al., 2012) We took care to restrict the number of variables using a screening and a priori set standards for the acceptance of variables into the model (see methods), reducing the chance that unrelated variables would be included merely by chance. In addition, we quantified the chance of obtaining models with similar explanatory power using 40 randomly generated predictors, showing that for the ground level models, this chance was very low. For the models for k, there was a higher chance, but considering that both seasonal models for *k* independently selected SVF and a small (25m) buffer traffic load estimate supports that these predictors are likely actually predictive of *k*.

Furthermore, it was challenging to recruit multiple apartments in a single building to take part in the study, and especially to find buildings on major roads with relatively high ground floor concentrations. Due to the limited number of three different heights per building, we only considered an exponential decay shape for the vertical gradient and did not consider shapes with a higher number of parameters.

Conclusion

Height above the ground is a relevant determinant of outdoor residential exposure to traffic related air pollutant NO_2. The decrease with height is clearer in places with a high ground floor concentration. This study is one of the first studies to develop a land use regression model which incorporates height above the ground as a predictor variable. This model outperformed a conventional ground-floor LUR model fitted using only measurements at street level in predicting concentrations at higher floors. The conventional LUR model overestimated exposure at higher floors, particularly for residences where relatively high NO_2 concentrations were measured at ground level near major roads. This overestimation may be minimized by considering decay with height.

References

Barratt B, Lee M, Wong P, Tang R, Tsui TH, Cheng W, et al. A Dynamic Three-Dimensional Air Pollution Exposure Model for Hong Kong. HEI Research Report number 194 2018.

Basagaña X, Rivera M, Aguilera I, Agis D, Bouso L, Elosua R, et al. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmospheric Environment 2012; 54: 634-642.

Beelen R, Hoek G, Pebesma E, Vienneau D, de Hoogh K, Briggs DJ. Mapping of background air pollution at a fine spatial scale across the European Union. Science of The Total Environment 2009; 407: 1852-1867.

Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. 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 2013; 72: 10-23.

Bundesamt für Landestopographie. DHM25: Das digitale Höhenmodell der Schweiz. 2001.

Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelen R, et al. Variation of NO_2 and NO_X concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmospheric Environment 2012; 62: 374-390.

de Hoogh K, Gulliver J, van Donkelaar A, Martin RV, Marshall JD, Bechle MJ, et al. Development of West-European $PM_{2.5}$ and NO_2 land use regression models incorporating satellite-derived and chemical transport modelling data. Environmental Research 2016; 151: 1-10.

de Hoogh K, Héritier H, Stafoggia M, Künzli N, Kloog I. Modelling daily $PM_{2.5}$ concentrations at high spatio-temporal resolution across Switzerland. Environmental Pollution 2017.

Eeftens M, Beekhuizen J, Beelen R, Wang M, Vermeulen R, Brunekreef B, et al. Quantifying urban street configuration for improvements in air pollution models. Atmospheric Environment 2013; 72: 1-9.

Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PM_{coarse} in 20 European study areas; results of the ESCAPE project. Environmental Science & Technology 2012; 46: 11195-205.

Eeftens M, Meier R, Schindler C, Aguilera I, Phuleria H, Ineichen A, et al. Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. Environmental Health 2016; 15: 53.

European Environment Agency (EEA). Implementation and achievements of CLC2006. Available from: https://www.eea.europa.eu/data-and-maps/data/clc-2006-raster-4 [Last accessed 21 August 2018].

Ho C-C, Chan C-C, Cho C-W, Lin H-I, Lee J-H, Wu C-F. Land use regression modeling with vertical distribution measurements for fine particulate matter and elements in an urban area. Atmospheric Environment 2015; 104: 256-263.

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

Karner AA, Eisinger DS, Niemeier DA. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. Environmental Science & Technology 2010; 44: 5334-5344.

Lindberg F, Grimmond C. Continuous sky view factor maps from high resolution urban digital elevation models. Climate Research 2010; 42: 177-183.

Lufthygieneamt beider Basel. Höhenabhängigkeit der NO_2-Belastung (in German). Study report 1993.

Meng ZY, Ding GA, Xu XB, Xu XD, Yu HQ, Wang SF. Vertical distributions of SO_2 and NO_2 in the lower atmosphere in Beijing urban areas, China. Science of The Total Environment 2008; 390: 456-465.

Panczak R, Galobardes B, Spoerri A, Zwahlen M, Egger M. High life in the sky? Mortality by floor of residence in Switzerland. European Journal of Epidemiology 2013; 28: 453-462.

Passam - Laboratory for environmental analysis and air pollution. Company website. Available from: http://www.passam.ch/wp/ [Last accessed 21 August 2018].

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.r-project.org/ [Last accessed 21 August 2018].

Statistisches Amt Basel Stadt. Statistisches Jahrbuch des Kantons Basel-Stadt, Chapter 1: Bevölkerung (in German). 2017.

Swiss Bundesamt für Umwelt (BAFU). SonBase–The GIS Noise Database of Switzerland. Technical bases [English version] 2009.

Tsai MY, Chen KS. Measurements and three-dimensional modeling of air pollutant dispersion in an Urban Street Canyon. Atmospheric Environment 2004; 38: 5911-5924.

Vardoulakis S, Fisher BE, Pericleous K, Gonzalez-Flesca N. Modelling air quality in street canyons: a review. Atmospheric Environment 2003; 37: 155-182.

Wang M, Beelen R, Eeftens M, Meliefste K, Hoek G, Brunekreef B. Systematic evaluation of land use regression models for NO_2. Environmental Science & Technology 2012; 46: 4481-9.

Wu C-F, Lin H-I, Ho C-C, Yang T-H, Chen C-C, Chan C-C. Modeling horizontal and vertical variation in intraurban exposure to $PM_{2.5}$ concentrations and compositions. Environmental Research 2014; 133: 96-102.

Zauli-Sajani S, Marchesi S, Trentini A, Bacco D, Zigola C, Rovelli S, et al. Vertical variation of $PM_{2.5}$ mass and chemical composition, particle size distribution, NO_2, and BTEX at a high rise building. Environmental Pollution 2018; 235: 339-349.

Table 1

Table 1: Model results for ground-floor models, the decay constant *k* and the final LUR including height, for winter and summer seasons separately and combined.

Dependent variable	Model	Season	N	Structure [a]	R^2 RMSE	R^2 LO(G)OCV RMSE [c]
Ground floor NO_2	M1	Winter	25	$NO_{2,\ winter}$ = 37.5+ 6.32*traffic load (25m)+ 3.47*residential land use (5000m)	0.71 4.2	0.59 4.9
	M2	Summer	25	$NO_{2,\ summer}$ = 25.5+ 7.17*traffic load (25m)+ 3.36*residential land use (5000m)	0.79 3.7	0.68 4.6
	M3	Both seasons [b]	50	$NO_{2,\ both\ seasons}$ = 25.8+ 11.9*winter + 7.09*traffic load (25m)+ 3.34*residential land use (5000m)	0.84 4.1	0.79 4.5
k (the decay constant)	M4	Winter	25	k_{winter} = -0.00956+ -0.00988 * traffic load (25m) + -0.00512 * residential land use (250m) + 0.00448 * SVF [d] + -0.00345 * road length (100m)	0.69 0.0055	0.61 0.0066
	M5	Summer	25	k_{summer} = -0.0125 + 0.00811 * SVF [d] + -0.00632 * traffic load (25m)	0.37 0.011	0.23 0.012
	M6	Both seasons [b]	50	$k_{both\ seasons}$ = -0.0112 + -0.00750 * traffic load (25m) + 0.00632 * SVF [d]	0.43 0.0067	0.33 0.010
Models including height	M7	Winter [b]	72	NO_2 = ground floor $concentration_{winter}$ * $exp(k_{winter}$*height)	0.67 3.9	0.55 4.5
	M8	Summer [b]	69	NO_2 = ground floor $concentration_{summer}$ * $exp(k_{summer}$*height)	0.78 3.6	0.68 4.4
	M9	Both seasons [b]	141	NO_2 = ground floor $concentration_{both\ seasons}$ * $exp(k_{both\ seasons}$*height)	0.84 3.8	0.79 4.0

[a] Effects are shown for standardized variables, so the relative effect sizes of the predictors can be compared.

[b] Mixed effects model with random intercept for site, reported model diagnostics are shown for fixed effects only.

[c] LO(G)OCV = Leave One (Group) Out Cross Validation: all observations from one building (all heights, all seasons) were left out simultaneously, which meant a single one for M1, M2, M4 and M5, two for M3 and M6, up to three for M7 and M8, and up to six for M9. Concentrations were estimated using only observations from other buildings.

[d] Sky view factor, see the method section; geographical predictor data.

Figure 1: Map of measurement sites: black dots represent the 25 sampling sites in the canton of Basel Stadt, Switzerland.

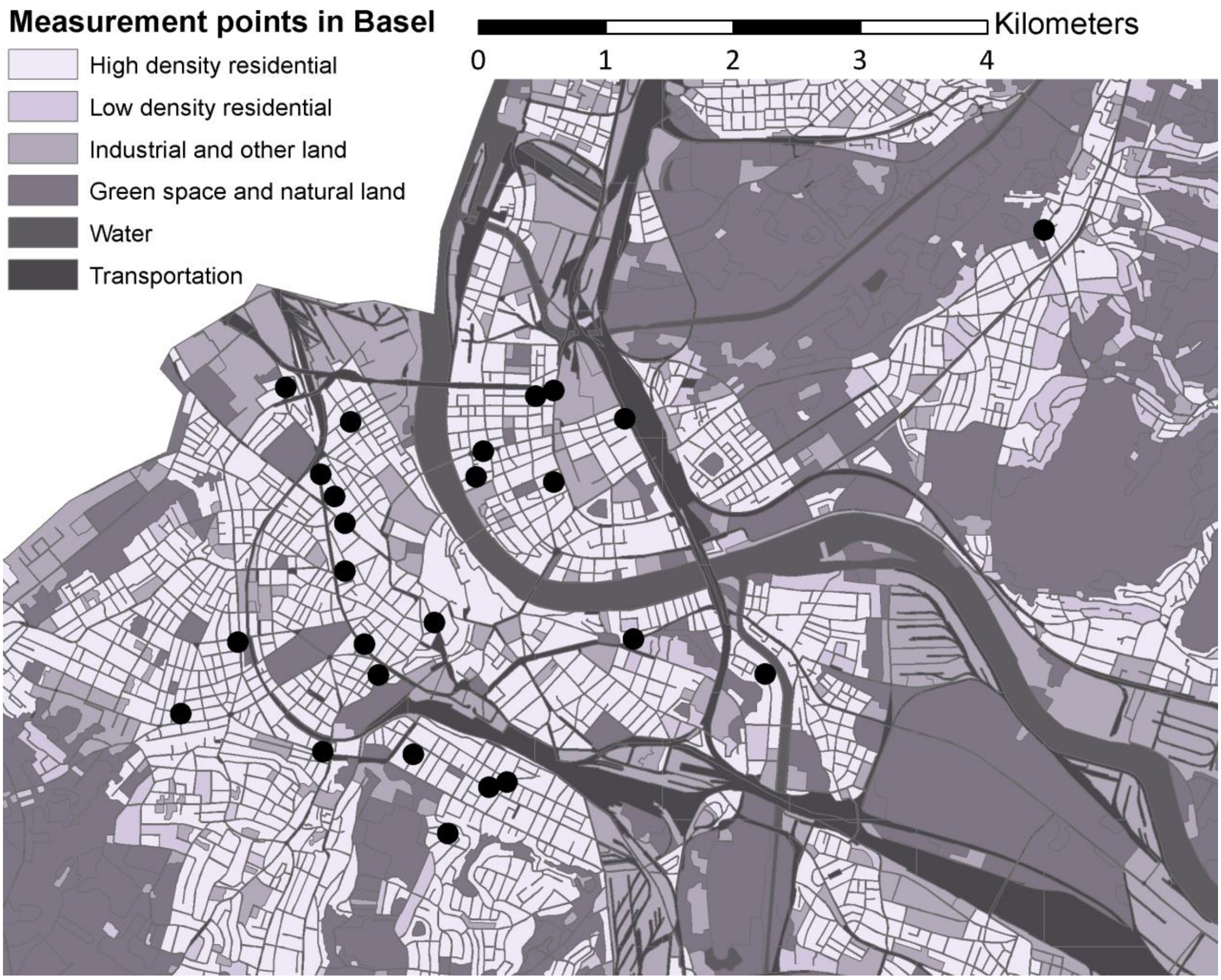

Figure 2: NO_2 concentration by season: each line connects the concentrations measured at the different floors of the same building. The ribbons represent the general decreasing trend with increasing height above the ground, estimated by locally weighted regression (LOESS).

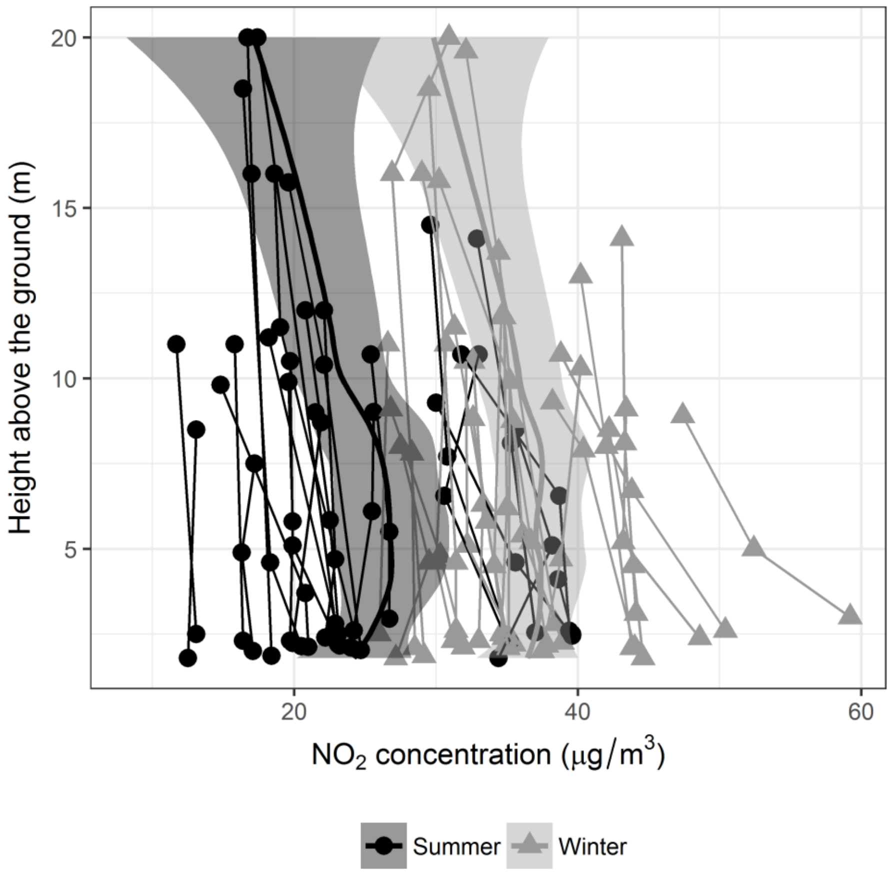

Figure 3: The decay constant *k* as calculated for all sites in the summer and winter period. Each point represents *k* as calculated for one of 25 sites measured in each season. The larger the point, the higher the traffic intensity around it: sites surrounded by more traffic generally have a lower *k* (i.e. a stronger decay). The darker the point, the more canyon-like the street, and thus the lower the SVF: sites in canyon-like streets generally have a lower *k* (i.e. a stronger decay). The diamond represents the mean, boxes the 25th, 50th and 75th percentiles, and whiskers extend to the smallest observation ≥ the 25th percentile - 1.5 * IQR (Interquartile Range) and the largest observation ≤ the 75th percentile + 1.5 * IQR. The Grey lines connect the same sites in different seasons.

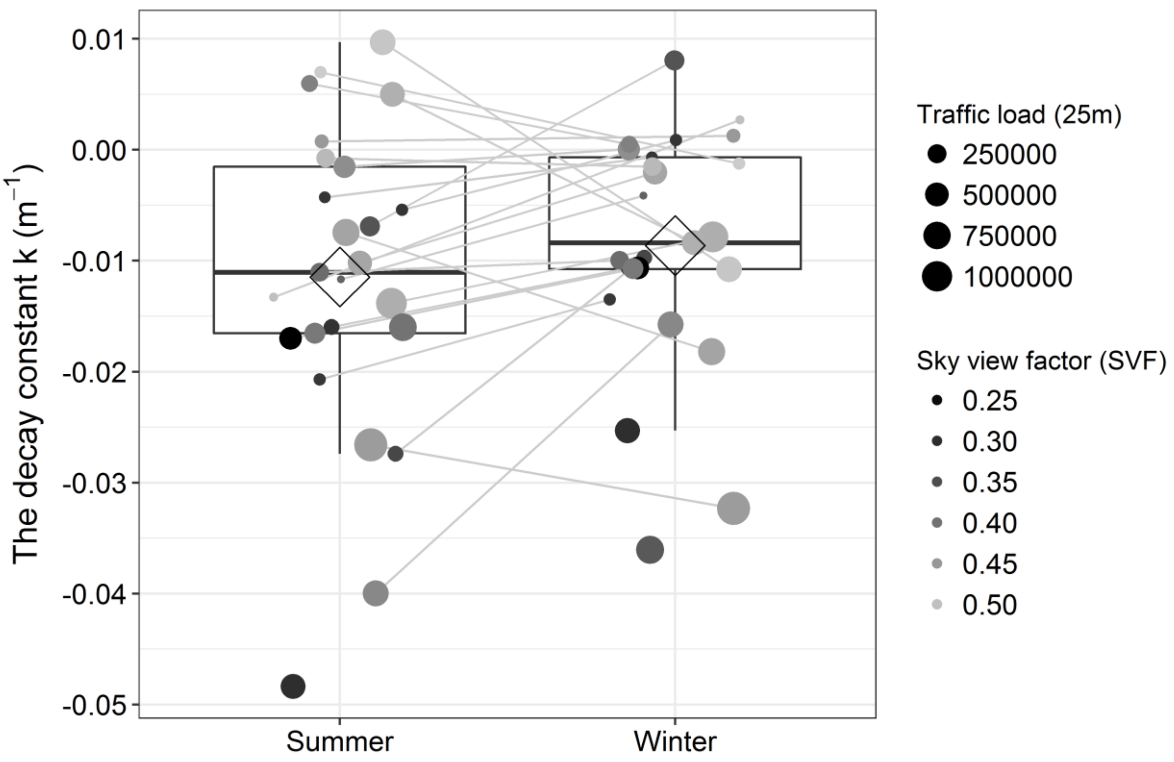

Figure 4: Model predictions for all 141 measurements by the ground floor model M3 (left) and by the model including height M9 (right). Connected points depict observations done in the same season at the same building. The ground floor model estimates the same concentration for the entire building facade, regardless of height, whereas the model including height shows a more realistic decay and predicts the observed decrease in NO_2 with height more accurately.

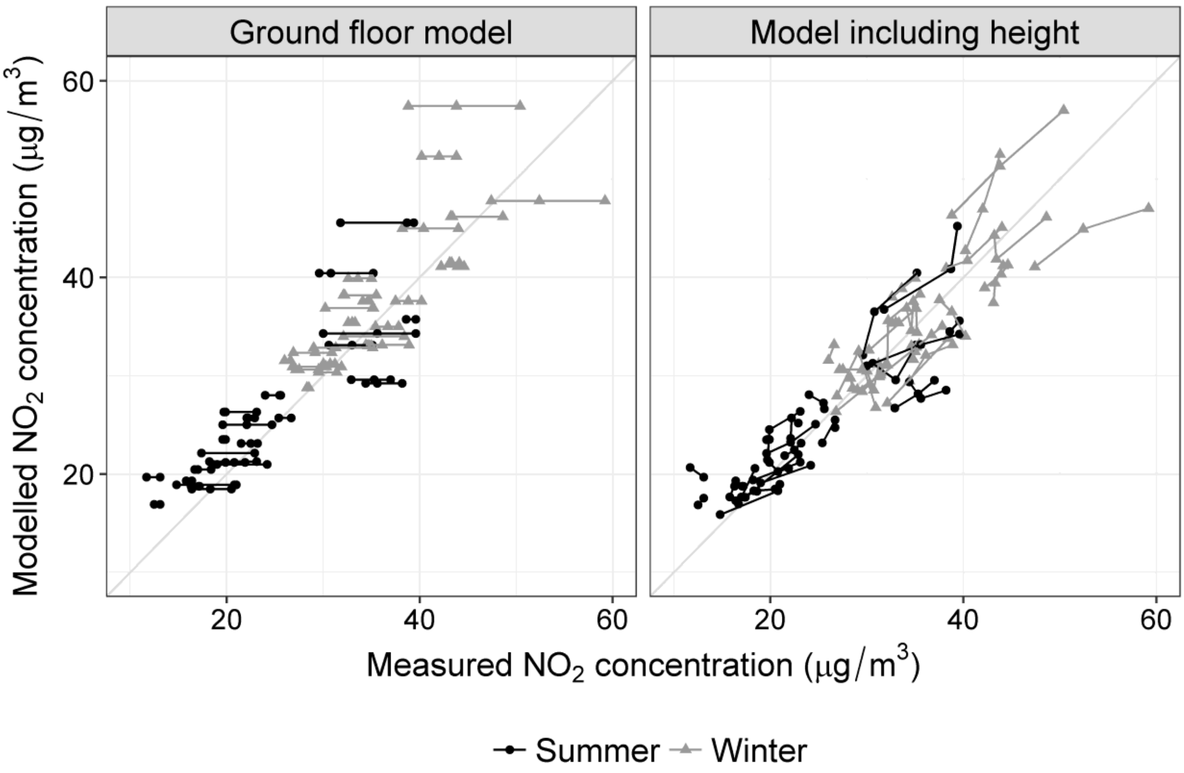

Figure 5: Model prediction error by floor for the model which did not include height and the model which considers decay with height. Note that the bias of the light grey model is increasingly positive for higher floors, indicating overestimation of the NO_2 concentration. The box shows the 25th, 50th and 75th percentiles, whiskers extend to the smallest observation ≥ the 25th percentile - 1.5 * IQR (Interquartile Range) and the largest observation ≤ the 75th percentile + 1.5 * IQR.