Land Use Regression Modelling in an Industrialised Setting: A South African Study

S. Muttoo, MMedSci, PhD [1], R. Naidoo, MBChB, PhD [1], K. Meliefste [2], L. Ramsay, PhD [1], R. Beelen, PhD [2]
[1] Univ of KwaZulu-Natal, Durban, South Africa; [2] Institute for Risk Assessment Sciences, Utrecht, the Netherlands

Introduction

Durban, specifically the South Durban (SD) region, situated on the east coast of South Africa, has historically been plagued with air pollution concerns, attributed to the juxtaposition of residential homes among an array of polluting industries including approximately including two major oil refineries, a sugar manufacturing plant, a paper manufacturing plant, sewage treatments works, a cluster of chemical industries, major petrochemical and chemical storage facilities, metal smelting industry, breweries, and paint and textile industries among others[1]. It also includes the Port of Durban, known to be one of the busiest in Africa, and is a focal point of many of the city's major transport routes[1]. Often referred to as the South Durban Industrial Basin due to its basin like topography that is conducive to air pollution accumulation, and coastal proximity, which accompanied by meteorological influences further enhance air quality impacts.

This toxic composition of polluting sources is a known risk factor for adverse public health outcomes, as highlighted in previous studies[2-4] and of further concern is the lack of current exposure data for this region, which is an important consideration when assessing the burden of impact on public health, particularly in vulnerable groups such as pregnant women.

In preparation for the MACE birth cohort study that was conducted in Durban, South Africa, this study presents the exposure assessment component using Land Use Regression modelling to predict exposure of the selected cohort to NOx levels.

Figure 1: South Durban Industrial Basin

Aim and Objectives

To characterize exposure of pregnant females residing in the South Durban region, to ambient NOx, using land use regression modelling. The specific objectives were:

- To determine ambient NOx levels in the South Durban.
- To identify geographic predictor variables likely to influence variation in individual exposure.
- To develop a LUR model for the South Durban.
- To predict exposure of pregnant females residing in south Durban using parameter estimates of the developed LUR model.

Methods & Materials

Study Location

The study area, located within the SD region of the Ethekwini Metropolitan District, of Kwa-Zulu Natal, South Africa has an approximate demographic of 595 601[5].

NOx Monitoring

Ambient NOx was measured over two two-week sampling periods at 32 sites in the SD using Ogawa badges. Quality control during sampling was achieved by duplicate sampling & co-locating samplers with existing Air Quality Monitoring Stations (AQMS). To achieve an annual adjusted average for NOx, the ESCAPE[6] method of adjustment was used.

Geographic Predictor Variables

Various categories of geographic predictor variables at varying buffer distances, were created, in accordance with those published in other studies. Variable categories included road type/length (length of major & minor roads); land use (area of open space, industrial land use & the harbour); physical geography (elevation); population & housing density.

LUR Model Development

Standard linear regression techniques were used to develop the prediction model. Using the ESCAPE[6] approach variables were either included / excluded from the model based increase in the adjusted R^2 > 1% and "a priori" direction of effect of existing variables (e.g. as NOx is associated with traffic emissions, regression models are expected to indicate a positive slope for traffic related variables) and those that were added in a stepwise manner. The leave-one-out cross validation (LOOCV) technique was used to validate the model.

Participant Exposure Assessment

The LUR model was then used to predict exposure at the residential addresses of study participants. For all the addresses, data was collected on all geographic predictors that were used in model development. In some cases values were truncated to avoid over-prediction.

Results

NOx Measurements

A high correlation (R=0.99) was observed between duplicate samplers for both sampling periods. For passive samplers that were co-located with an existing AQMS, differences of between 8 – 10 μg/m³ were observed. Using the ESCAPE[6] protocol the NOx levels were adjusted to an annual average

Table 1: Descriptive Summary of Annual Adjusted Average of NOx Results

N = 32	NOx(μg/m³)
Mean	33.1
Median	34.0
Standard Deviation	9.4
Minimum	13.7
Maximum	52.4

Geographic Predictor Variables

Table 2: A summary of significant geographic variables.

Geographic Predictor Variables	Mean	Std Dev	Min	Med	Max
Length of Major Road within 300m	95,9	367,3	0	0	1656,6
Length of Minor Road within 1000m	28182,9	5364,6	15917,5	27795,7	35739,0
Area of Open Space within 1000m²	237179,1	223212,9	0	170237,6	767935,6
*Variables presented are those that were found to be significant out of a total of 33 variables *Variables referred to are in context of their respective buffer distances					

Results (continued)

Model Building & Validation

Variables were individually regressed against NOx concentrations in the univariate analysis. In the stepwise linear regression the start model yielded an adjusted R^2 of 0.41 and this sequentially increased to 0.73 as non-significant variables were removed from the model in a stepwise manner. The final model presented is a combination of 3 variables. The LOOCV yielded an R^2 of 0.59

Table 3: Summary of Regression Model Predicting NOx

Variable	β	SE	t	P
Intercept	18,81385	6,72687	2,8	0,009
Length of Minor Road within 1000m	0,0006708	0,0002085	3,22	0,003
Length of Major Road within 300m	0,0082245	0,0025539	3,22	0,003
Area of Open Space within 1000m	-9,08E-06	5,13E-06	-1,77	0,088
R^2 = 0.73; RMSE = 4.97				

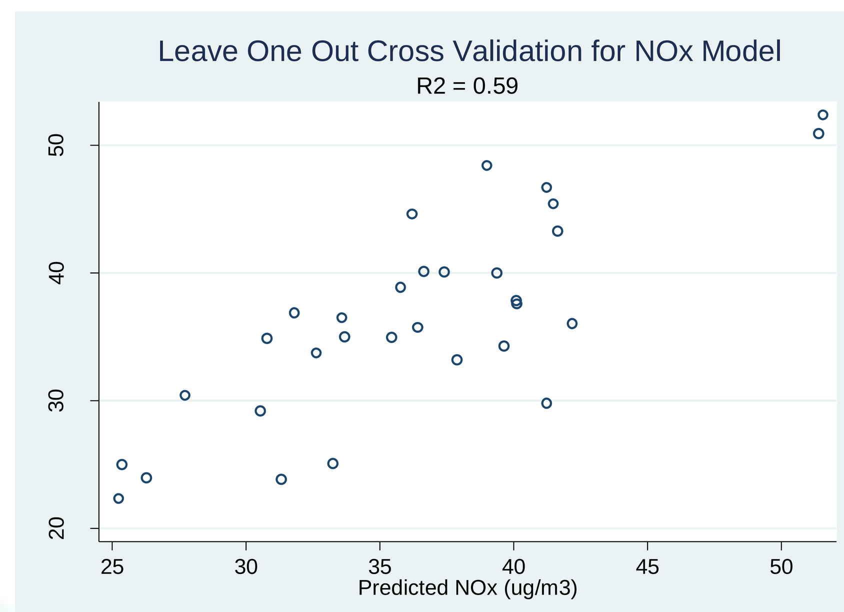

Figure 2: Leave one out cross validation for NOx

In the final model it was noted that all predictor variables had acceptable levels of VIF (<5). No outliers were identified in the Cook's distance test.

The Pearson's correlation coefficients where determined for significant predictor variables identified in the model & as per predefined "a priori" the road length variables were positively associated with NOx while the open space variable was negatively associated with NOx

Participant Exposure Assessment

The final model was applied to participant addresses. Truncation was only applied to the variable major road within 300m, though minimally. However, the correlation between the un-truncated and truncated model was observed to be high at 0.99.

Participant exposure estimates ranged from 19,95 – 53,24 μg/m³ with a mean of 38,97 μg/m³ in comparison to measured NOx levels with a range of 22,91 – 50,94 μg/m³ and a mean value of 36,35 μg/m³.

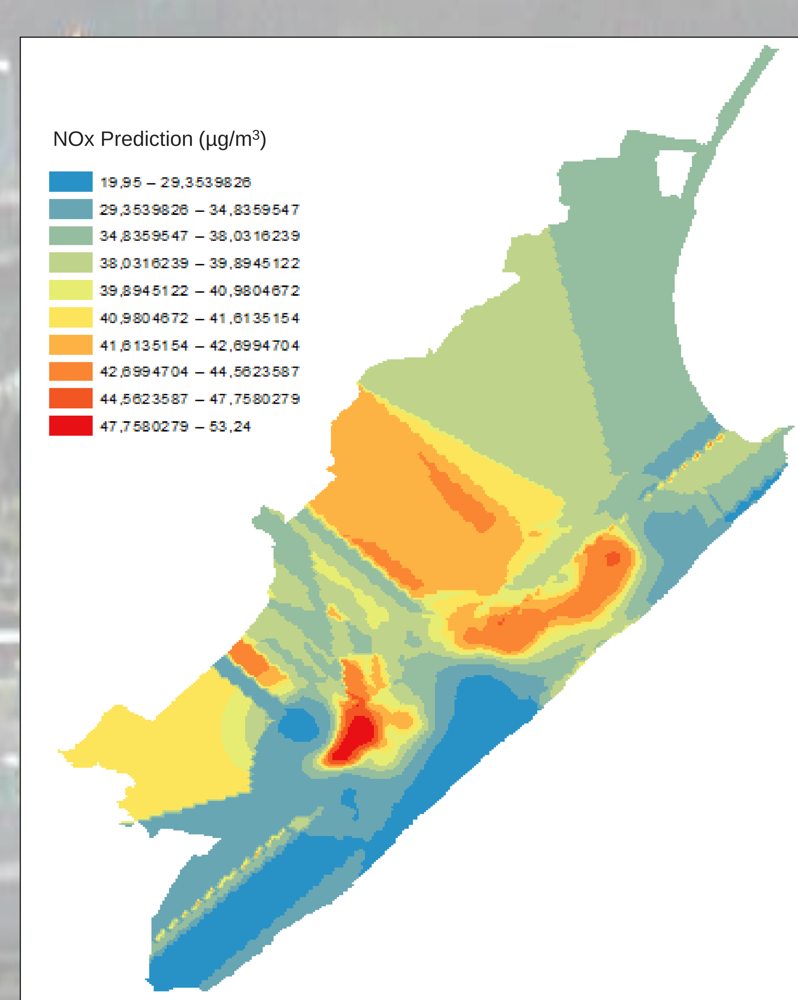

Figure 3: Concentration Map of NOx Exposure Estimates

Discussion

The LUR model presented for the South Durban region, includes three variables accounting for 73% of the variation in the observed NOx levels. The significant variables identified in this study are comparable to previous studies, in which length of a major road is the common factor.

As with previous studies, NOx levels were mostly influenced by proxy traffic markers such as length of road variables. Industrial land use showed no significant association with NOx levels, which could be attributed to the lack of distinction made between industries which emit NOx and those that do not in the modelling. Modelled and predicted mean NOx levels were highly correlated, thus indicating the strength of the model to accurately predict exposure at un-monitored locations.

Though there were some shortcomings in model development relating to the number of samplers and method of installation, the overall model performed well in comparison to previously published work with a high predictive capability.

The LUR modelling technique has proven to be one that works well in low-resource settings and its purpose design nature optimises its predictive capability.

This LUR modelling technique developed for this urban industrialised setting in South Africa, to predict exposure at participant addresses was fairly successful and the results of this preliminary assessment shall be related to biological determinants of adverse birth outcomes in subsequent research which forms part of a birth cohort study.

References

1. Guastella L, Knudsen S. South Durban Basin Multi-Point Plan Case Study Report. Durban, South Africa: Department of Environmental Affairs and Tourism; 2007.
2. Naidoo R, Gqaleni N, Batterman S, Robins T. South Durban Health Study. South Africa: University of Kwa-Zulu Natal; 2007.
3. Kistnasamy EJ, Robins TG, Naidoo R, et al. The relationship between asthma and ambient air pollutants among primary school students in Durban, South Africa. International Journal of Environment and Health 2008;2:365-85.
4. Naidoo RN, Robins TG, Batterman S, et al. Ambient pollution and respiratory outcomes among schoolchildren in Durban, South Africa. South African Journal of Child Health 2013;7:127-34.
5. South Africa - Ethekwini Population Census. In: Statistics South Africa, 2016 AfriGIS (Pty) Ltd, Google; 2011.
6. ESCAPE Exposure Assessment Manual. In; 2010.

Declaration: All authors of this poster have read the definition of Financial Conflict of Interest and certify there are no conflicts of interest to declare.

Funding Sources: Medical Research Council, South Africa; National Research Foundation, South Africa