Modelling Spatial Variability of Intra-Urban Concentration of NO_2: Situation Assessment in Indian Cities

Introduction

Air quality is a cause of concern in India, particularly in cities where swelling urban population, unplanned urban and industrial development and increased volume of motorized traffic have resulted in severe air pollution level often exceed the National Ambient Air Quality Standards (NAAQS), which in turn affecting the surrounding environment and human health. Cities in India have grown haphazardly with little consideration of the functioning of urban systems as a whole. The country's urban areas often lack adequae regional transport networks, for example. Large swaths of informal settlements have emerged in vacant inner-city districts and suburban peripheries, compromising environmental conditions, public health, and personal safety. In recent decade, India is struggling to improve its urban air quality. To tackle this deteriorating situation the country requires a method to measure properly and model pollution levels.

Different studies evident spatial and temporal distribution of air pollution. Various studies have documented significant variation of outdoor air pollution at small scale within urban areas for important pollution such as NO_2 (Fisher, 2000; Kingham, Briggs, Elliott, Fischer, & Erik Lebret, 2000; Lebret et al., 2000; Monn, 2001; Jerrett et al., 2005; Zhu, Hinds, Kim, Shen, & Sioutas, 2002). In some settings, the distribution greatly vary within a city than the between cities (Jerrett, et al., 2005; Miller et al., 2007). Like every large cities in the world, Indian cities have varied land use pattern. Due to that, the spatial distribution of air pollution in cities influenced by many factors, including distance from source, reactivity of pollutants with other chemicals in the air, wind pattern, temperature, elevation and a strong influence of urban design (building height and density). Further, regional air pollution mitigation has focused mainly on emission reduction through economic (odd-even license plates access in Delhi, ban of sale of above 2000cc diesel vehicles in Delhi), or technological (catalytic convertors for cars, clean coal for power plants) means. However, strategies for mitigating the intra-urban impact of air pollution have not been identified or investigated mainly due to unknown and unexplored situation of spatial distribution of air pollution. The role of land use features in regulate ambient air pollution concentration plays an important role. It rises an interesting question: to what extent can urban land use be managed or modified to decrease the impact of air pollution on human health? However, the relationship among air pollution, local land use pattern and associated health impacts has not been quantified or investigated systematically before.

To address such challenges, Land Use Regression (LUR) model, which successfully combines measurements of air pollution and stochastic modelling using land use variables obtained through geographic information systems (GIS) to construct predictive model of ambient air pollutants concentration. Further, the spatial interpolation technique gives a smooth surface with a continuous map of pollution concentration. Modelling of spatial variation in air quality and estimating human exposure of urban India provides a better opportunity of understanding the health risk and to inform the necessary policies and research to address air pollution level. The value of this modelling in India is not just because air quality is poor in cities like Delhi and Mumbai, but also because these cities are in high density, high rise cities featuring a predominantly vertical population growth. The combination of densification, complex topography and vertical stratified growth in the sense of high level regional pollution suggesting a more complex variation in pollution concentration than seen in the cities of developed world. Findings from these cities may not be applicable to the Indian content, or even to other cities of developing countries. Therefore, this research seek to predict the concentration of NO_2 by using land use variables at unsampled locations of Delhi, Mumbai and Navi Mumbai on the basis of sampled monitoring data and also tried to see the distributional variation of NO_2 concentration within cities.

Materials and Methods

The monthly average of nitrogen dioxide (NO_2) concentration (μg/m3) were obtained directly from CPCB official website, for thirteen monitoring stations in Delhi, six monitoring stations in Mumbai and Navi Mumbai between the year 2015 and 2016. All of these stations are located in proximity to industrial area, high traffic intensity area, residential and other area (CPCB, 2016). To estimate independent variables for LUR model, the different spatial data have been used in this study are consisted in five broad categories: Physical geography– altitude (meter); Land use- industrial area and commercial area; Road– road type and length of different roads; Traffic– traffic intensity, density of traffic signals, density of bus stops; Demographic– population density. The altitude data used in this study estimated through Google Earth. For land use data of industries and commercial area, OpenStreetMap data of India have been used. Density of bus stops and traffic signals data also obtained through map vectorization from Google Earth. The roads of the cities classified under three broad categories ; i) Primary roads- express ways and national highways (NH); ii) secondary roads- state highways (SH) and major district roads (MDR); iii) Tertiary roads- other district roads (ODR) (Kadiyali & Lal, 2005). Traffic intensity data consists of number of vehicles per kilometre on a major road of Delhi and Mumbai which was obtained from Delhi Traffic Police website and Transport Department Statistics of Mumbai respectively. Ward wise population data of Delhi and Mumbai in 2001and 2011 are taken from Census of India for the purpose of projection of 2015 population for both the cities.

Under the Classical Linear Model, upheld the assumptions of Ordinary Least Square (OLS) regression, land use regression was developed for accurate exposure assessment at those locations where monitoring stations are not situated in the our study area. The Land Use Regression model requires data of air pollution concentration which will be placed as dependent variable and land use variables such as industrial area, commercial area, are under vegetation, length of road, traffic intensity and other topographical variables like wind speed and temperature is used as independent variables. The land use pattern at monitoring sites and area within a buffer constructed around with the help GIS software. For each monitoring sampled station locations, we constructed a series of concentric circle, known as buffers, with radii of 100, 300, 500, 750, 1000 and 2000 meters, using ArcGIS (after reviewing previous literature). All input variables such as land use data, road data, and population data was computed by estimating the area, length and quantity of the land use characteristics within certain buffer. Due to unavailability of traffic intensity data of all streets for the cities modelled in this study, the traffic intensity was estimated within certain buffer by multiplying length of roads in certain buffer and number of vehicles per km. of the city on a day. Population in each buffer was estimated by computing the area of buffer, which was further multiplied with population density of the census ward in a city. A total of 58 variables in eight categories were created for LUR model development.

Ordinary Kriging approach was used to visualize the surface of the LUR model using Geostatistical wizard in ArcGIS 10.2. It is a weighted combination of measurements at surrounding of measured sample locations. Kriging assigns weights at each concentration by exploiting the spatial correlation among the observed measurements. The robustness of the surface map obtained, was established by estimating the congruency between predicted values and observed values through scatter plot. The assumption for normality is checked using Q-Q plot.

Important Results

NO_2 sample collected from different monitoring stations of Delhi and Mumbai had an arithmetic mean of 62.61 ($\mu g/m^3$) with values ranging from 19 to 185.71 ($\mu g/m^3$) and 45.15 ($\mu g/m^3$) with values ranging from 6.14 to 129.18 ($\mu g/m^3$) respectively. After elimination of other variables that are correlated (Pearson's $r \geq 0.06$) with most highly ranked variables, 10 variables remained from the total of 58 variables in Delhi and 7 out of 58 variables remained in Mumbai and Navi Mumbai. The multivariate analysis evidenced variables with statistical significance. In the context of Delhi, the final model explained 46% of NO_2 variance. Traffic intensity in primary and secondary roads within the buffer of 1000 meters and 500 meters radius respectively, and length of primary roads at 1000 meters radius buffer exhibits the strongest interaction with level of NO_2 ($p < 0.05$) than the interaction with density of traffic signals within the buffer of 2000 meter. For instance, in case of traffic signals, as the number of traffic signal increases in 2000 meters radius buffer area concentration of NO_2 rises. However, traffic intensity in primary road at 1000 meters radius buffer shows negative relation with the coefficient which may be resulting from too wide distance from the monitoring stations **(See Table 1)**.

In Mumbai, the model explained 55% of the variability in measured concentration of NO_2. Industrial and commercial area within 750 meters radius buffer shows the strongest relation to NO_2 ($p < 0.001$) concentration in the city. However, commercial area shows negative relation with NO_2 concentration. The traffic intensity of primary roads within 1000 meters radius buffer and population within 300 meters radius buffer shows comparatively less significant relation with NO_2 concentration **(See Table 2)**.

The interpolated annual spatial concentration of NO_2 using the Kriging approach have created a smooth surface of intra-urban distribution for Delhi. The average of LUR predicted NO_2 was 53.63 $\mu g/m^3$ (SD= 10.85 $\mu g/m^3$). The interpolated NO_2 concentrations map gests that city centre or the main city are the major pollution source in Delhi. While the peripheral areas may play a role in decreasing air pollutant concentrations. Modelled ambient concentrations of NO_2 within Delhi city range from 42 $\mu g/m^3$ to 100 $\mu g/m^3$, well below the National Ambient Air Quality standards i.e. 40 $\mu g/m^3$ **(See Figure 1)**.

Similarly, kriging was applied to construct the NO_2 concentration surface in Mumbai. In comparison to Delhi the LUR-predictive mean concentration is lesser with lower standard deviation. The low standard deviation value reflects the more homogeneity in terms of NO_2 spread across the city. In terms of spatial distribution of pollution concentration, it is more in south-east and north part of Mumbai. Whereas, Navi Mumbai have lesser concentration compare to Mumbai. The surface map reflects the heterogeneity of NO_2 concentration across different areas. For instance, concentration can be observed to be more in south-east and north part of Mumbai whereas, Navi Mumbai have lesser concentration compare to Mumbai. Modelled ambient concentrations of NO_2 within Mumbai city range from 29 to 61 μg/m3, which is also below the National Ambient Air Quality standards **(See Figure 1)**.

The validity of LUR model can only be established through coherency or relatedness of predicted and observed values. The scatter plot describes the relatedness or the correlation between the predicted values and observed values of NO_2. The correlation between the values came out to be 0.49 **(See Figure 2)**.

Additionally, the kriging interpolation is based on assumption of normality of observed values. In order to check whether the assumption has been valid in the analysis, Q-Q plot is used. Majority of the points found to be closer or lying on the straight line passing through the origin which establishes the normality assumption. However, extreme values seem to be placed quite far in the plot which is believed to be not affecting the results **(See Figure 3)**.

Conclusion

We can summarise the attempt with the estimated value for air pollution concentration, the predicted surface map for Delhi, Mumbai and Navi-Mumbai, using the concentration data from CPCB and predictor variables from census and Google Earth. The study used multiple secondary data for study area to construct and evaluate the land use regression model. In general, it can be said that the modelling procedure has brought good estimates in terms of accuracy and suitableness with the theory of environmental change. The methodology has its limitations with respect to its applicability which lies within a small area, the availability of data is another challenge. Amidst these limitations and challenges, to our knowledge, this research work is first application of land use regression and spatial interpolation (Kriging) to predict the surface map for two most important cities in India using the secondary data available in public domain. However, the modelling should be modified to inculcate other predictors and improvise on the estimates obtained for air pollution concentration.

References

CPCB, (2016). Air Quality Monitoring, Emission Inventory and Source Apportionment Study for Indian Cities: National Summary Report. New Delhi: Government of India.

Fisher, K. (2000). Uncertain aims and tacit negotiation: Birth control practices in Britain, 1925-50. *Population and Development Review*. https://doi.org/10.1111/j.1728-4457.2000.00295.x

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., … Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*. https://doi.org/10.1038/sj.jea.7500388

Kadiyali , L., & Lal, N. (2005). *Principles and practices of highway engineering.* New Delhi: Khanna Publishers.

Kingham, S., Briggs, D., Elliott, P., Fischer, P., & Erik Lebret. (2000). Spatial variations in the concentrations of traffic-related pollutants in indoor and outdoor air in Huddersfield, England. *Atmospheric Environment*. https://doi.org/10.1016/S1352-2310(99)00321-0

Lebret, E., Briggs, D., Van Reeuwijk, H., Fischer, P., Smallbone, K., Harssema, H., … Elliott, P. (2000). Small area variations in ambient NO2concentrations in four European areas. *Atmospheric Environment*. https://doi.org/10.1016/S1352-2310(99)00292-7

Miller, K. A., Siscovick, D. S., Sheppard, L., Shepherd, K., Sullivan, J. H., Anderson, G. L., & Kaufman, J. D. (2007). Long-term exposure to air pollution and incidence of cardiovascular events in women. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa054409

Monn, C. (2001). Exposure assessment of air pollutants: A review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. *Atmospheric Environment*. https://doi.org/10.1016/S1352-2310(00)00330-7

Zhu, Y., Hinds, W. C., Kim, S., Shen, S., & Sioutas, C. (2002). Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment*. https://doi.org/10.1016/S1352-2310(02)00354-0

Tables and figures

Table 3.3: Results of LUR model with chosen variables based on criteria, Delhi

NO2	β	Std. Error	t	P>t	95% CI	
					LCL	UCL
Population (300m)	0.00533	0.00343	1.55	0.123	-0.0015	0.0121
Industrial Area (750m)	-0.00006	0.00004	-1.41	0.16	-0.0001	0.0000
Commercial Area (750m)	0.00034	0.00022	1.57	0.119	-0.0001	0.0008
Traffic Signals (1000m-2000m)	0.86892	0.49565	1.75	0.082	-0.1111	1.8489
Traffic signals (500m)	0.82338	3.80933	0.22	0.829	-6.7065	8.3533
Traffic Intensity Primary Road (750m-1000 m)	-0.03686	0.01722	-2.14	0.034	-0.0709	-0.0028
Traffic Intensity Secondary Road (750 m-1000 m)	-0.03436	0.02813	-1.22	0.224	-0.0900	0.0213
Traffic Intensity Secondary Road (500m)	0.00693	0.00185	3.33	0.026	-0.0053	0.0192
Primary Road (1000m)	0.00873	0.00375	2.33	0.022	0.0013	0.0161
Secondary Road (750 m)	0.01240	0.00834	1.49	0.139	-0.0041	0.0289
Constant	42.34590	11.95442	3.54	0.001	18.7100	65.9820

R^2= 0.4689, N=156

Table 3.4: Results of LUR model with chosen variables based on criteria, Mumbai-Navi Mumbai

NO2	β	Std. Error	t	P>t	95% CI	
					LCL	UCL
Population (300m)	0.00073	0.00037	1.99	0.05	0.00000	0.00146
Industrial Area (750m)	0.00014	0.00004	3.64	0.01	0.00006	0.00021
Commercial Area (750m)	-0.00036	0.00008	-4.52	0.01	-0.00052	-0.00020
Primary Road Traffic Intensity(1000m)	0.00080	0.00092	0.87	0.389	-0.00103	0.00263
Secondary Road Traffic Intensity(750m-1000m)	-0.02075	0.00374	-0.67	0.532	-0.02820	-0.01331
Secondary Road Traffic Intensity (500 m)	0.00904	0.00198	4.56	0.01	0.00510	0.01298
Secondary Road (500m-750m)	0.04212	0.00975	4.32	0.01	0.02274	0.06151
Constant	45.02315	13.15620	3.42	0.001	18.84643	71.19980

R^2= 0.5586, N=48

Figure 1: Concentration of NO2 (µg/m3) in Delhi, Mumbai-Navi Mumbai, 2015

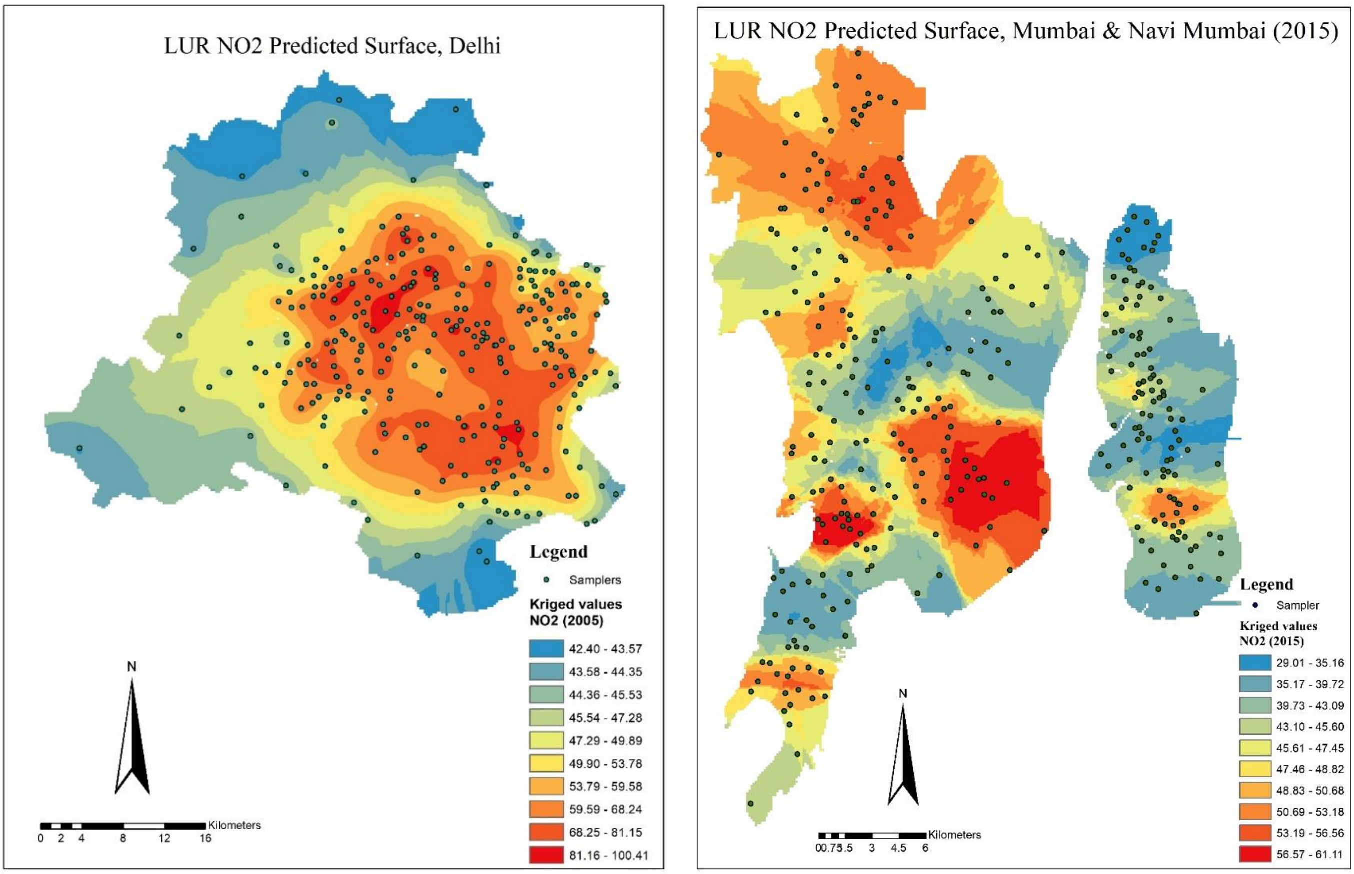

Figure 2: Cross validation of LUR model for Delhi and Mumbai-Navi Mumbai

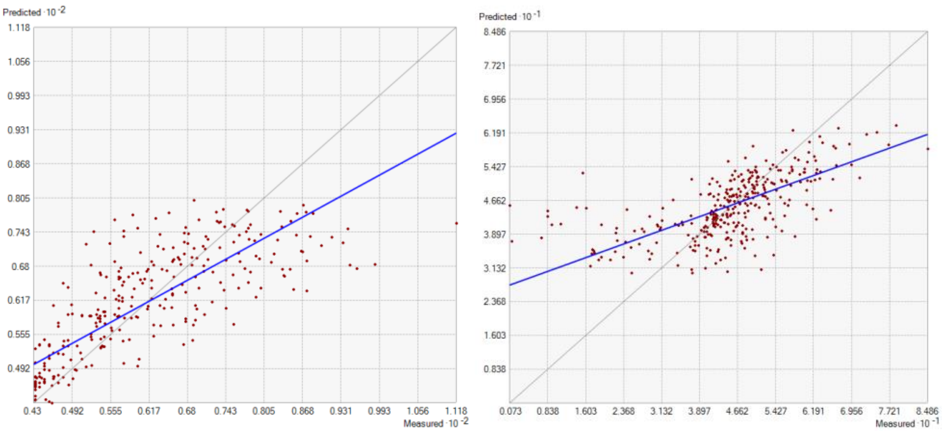

Figure 3: Q-Q plot for normality check before kriging interpolation, Delhi and Mumbai-Navi Mumbai