Environmental Research 155 (2017) 42–48

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Modeling the intraurban variation in nitrogen dioxide in urban areas in Kathmandu Valley, Nepal

Anobha Gurung[a], Jonathan I. Levy[b], Michelle L. Bell[a,*]

[a] School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA
[b] Department of Environmental Health, Boston University School of Public Health, Boston, MA 02118, USA

A R T I C L E I N F O

Keywords:
Land use regression
Traffic
Air pollution
NO_2
Kathmandu Valley
Nepal

A B S T R A C T

Background: With growing urbanization, traffic has become one of the main sources of air pollution in Nepal. Understanding the impact of air pollution on health requires estimation of exposure. Land use regression (LUR) modeling is widely used to investigate intraurban variation in air pollution for Western cities, but LUR models are relatively scarce in developing countries. In this study, we developed LUR models to characterize intraurban variation of nitrogen dioxide (NO_2) in urban areas of Kathmandu Valley, Nepal, one of the fastest urbanizing areas in South Asia.
Methods: Over the study area, 135 monitoring sites were selected using stratified random sampling based on building density and road density along with purposeful sampling. In 2014, four sampling campaigns were performed, one per season, for two weeks each. NO_2 was measured using duplicate Palmes tubes at 135 sites, with additional information on nitric oxide (NO), NO_2, and nitrogen oxide (NOx) concentrations derived from Ogawa badges at 28 sites. Geographical variables (e.g., road network, land use, built area) were used as predictor variables in LUR modeling, considering buffers 25–400 m around each monitoring site.
Results: Annual average NO_2 by site ranged from 5.7 to 120 ppb for the study area, with higher concentrations in the Village Development Committees (VDCs) of Kathmandu and Lalitpur than in Kirtipur, Thimi, and Bhaktapur, and with variability present within each VDC. In the final LUR model, length of major road, built area, and industrial area were positively associated with NO_2 concentration while normalized difference vegetation index (NDVI) was negatively associated with NO_2 concentration (R^2=0.51). Cross-validation of the results confirmed the reliability of the model.
Conclusions: The combination of passive NO_2 sampling and LUR modeling techniques allowed for characterization of nitrogen dioxide patterns in a developing country setting, demonstrating spatial variability and high pollution levels.

1. Introduction

By 2008, over half of the world's population was living in urban areas, with more than 90% of urban population growth by 2030 occurring in developing countries (Obaid, 2007). The majority of the growth in the near future will occur in Asia and Africa, and Asia is predicted to contain more than half of the world's cities with population of 500,000 or more (Seto et al., 2010). High air pollution is one of the many environmental problems present in these urban areas. About 1.4 billion urban residents are currently living in areas exceeding World Health Organization (WHO) air quality guidelines (Zou et al., 2009), and according to WHO, about 3.7 million deaths in 2012 can be attributed to outdoor air pollution worldwide (World Health Organization, 2014). Motor vehicles represent one of the main sources of pollution in urban areas (Health Effects Institute, 2011).

In spite of the increasing urbanization, high traffic volumes, and large health burden, exposure assessments are relatively scarce in developing countries (Han and Naeher, 2006). Studies specifically in developing nations are important as traffic-related air pollution is likely to differ between developing and developed nations due to differences in fuel characteristics, vehicle technology, contributions of heavy polluters (e.g., more motorcycles), driving habits, traffic patterns (e.g., frequent congestion) and roadways (e.g., graded vs. ungraded). In comparison to Western cities, Asian cities also have different built

Abbreviations: HB-HR, high building density and high road density; HB-LR, high building density and low road density; LB-HR, low building density and high road density; LB-LR, low building density and low road density; MPPW, Ministry of Physical Planning and Works; NDVI, normalized difference vegetation index; VDC, village development committee

* Corresponding author.
E-mail address: michelle.bell@yale.edu (M.L. Bell).

http://dx.doi.org/10.1016/j.envres.2017.01.038
Received 24 October 2016; Received in revised form 20 January 2017; Accepted 28 January 2017
Available online 09 February 2017

structure, topography, weather and land use patterns.

For traffic-related air pollutants, the within-city spatial contrast may be as large as the between-city contrast (Hoek et al., 2008), so traditional approaches such as assigning exposures based on values at the nearest monitors (Bell, 2006) are less interpretable. Land use regression (LUR) modeling has been widely used to investigate small scale variation in traffic-related air pollution (Poplawski et al., 2009), but to date, LUR models have been mostly limited to Europe and North America with only a few studies available in developing cities (Allen et al., 2013; Dionisio et al., 2010). In Asia, LUR studies have been in urban centers of China or other locales that may differ from rapidly urbanizing areas of South Asia (Allen et al., 2013; Chen et al., 2010; Kashima et al., 2009; Li et al., 2010).

High air pollution levels in the Kathmandu Valley of Nepal, one of the fastest urbanizing areas of South Asia, were first documented in the 1990s with pollution monitoring by several organizations, including the Kathmandu Valley Vehicular Emission Control Project funded by the United Nations Development Program and the Metropolitan Environment Improvement Project funded by the World Bank (International Center for Integrated Mountain Development, 2006). Vehicles were indicated as the main source of pollution by emission inventories developed in 1993 and 2001 (Asian Development Bank, 2006). Increasing number of vehicles, rapid and unplanned urbanization, population inflow, Valley-centric industrialization, poor road infrastructure, fuel adulteration and lack of regulations or enforcement of regulations contribute to the high air pollution present in the Valley. The Valley's bowl-like topography, low wind speeds limiting air pollution dispersion, and frequent thermal inversions also factor into the high levels of pollution. Despite the high air pollution, insight about health effects are limited by the dearth of air pollution data (Gurung and Bell, 2013), and very limited information is available on the spatial distribution of air pollution in the Valley.

In this study, we used LUR modeling to characterize intraurban variation of traffic-related air pollution in urban areas of Kathmandu Valley, Nepal. The methods were designed to address issues central to Nepal and similar countries, including limited data availability.

2. Materials and methods

2.1. Study area

The study area covers the urban core area of Kathmandu Valley, Nepal, consisting of the five urban village development committees (VDC). These five urban VDCs form Kathmandu Metropolitan City along with the surrounding municipal towns: Lalitpur, Bhaktapur, Kritipur and Madhyapur-Thimi (Fig. 1). According to Population Census 2011, the urban area in the Valley constitutes 102 km^2 with population of 1,464,164. Annual average population growth in the past ten years was estimated at 4.32% for the Valley (Statistics, 2012). From 2000 to 2010 the number of registered vehicles in the Valley increased 274%, with 93% of the registered vehicles in 2010 being privately owned (Japan International Cooperation Agency, 2011).

2.2. Spatial allocation of sites

2.2.1. Site assignment region

We followed a strategy for site allocation that leveraged available geospatial data to try to maximize contrast among monitors and representativeness of measurements, as done in many previous studies in Europe and North America. A grid of 200 m×200 m was created over the five urban VDCs and each grid cell was assigned to a VDC. When a 200 m×200 m grid cell included the boundary of two or more VDCs, it was split into partial grid cells and a VDC was assigned to each partial grid cell. Similarly, at outer boundaries of the study area, a 200 m×200 m grid cell is partially within and partially outside the study area. Below we refer to complete cells of 40,000 m^2 (i.e., 200 m×200 m) as 'full grid cells' and cells with area less than 40,000 m^2 (i.e., cells split by VDC boundaries or at outer boundaries of study area) as 'partial grid cells'. We refer to all cells (full grid cell and partial grid cells) as 'grid cells'. Input data (e.g., land use, road network) for LUR modeling are unavailable outside the study area or five urban VDCs. Before site allocation, a buffer of 400 m was created at the edge of the study area and going inward from the boundary of the study area to make sure that required input data for LUR modeling are available. We refer to 'site assignment region' as the study area after removing the area in the 400 m buffer at the boundary.

2.2.2. Assigning strata based on built density and road density

Detailed road networks and housing footprint (i.e., location and horizontal dimensions of buildings) data were obtained from the Ministry of Physical Planning and Works (MPPW), Nepal. Next for each grid cell in the site assignment region, building density and road density were calculated. 'Building density' refers to the fraction of built area within a given grid cell where 'built area' refers to the area of the building footprint. Supplement 1 shows the built area for the study area. 'Road density' refers to the density of roads within a given grid cell. Building density was divided into two categories within each VDC: the upper quartile (high building density) and lower three quartiles (low building density). Road density was similarly divided into two categories. By cross-classifying cells, four strata were created: high building density and high road density (HB-HR); high building density and low road density (HB-LR); low building density and high road density (LB-HR); low building density and low road density (LB-LR). Related approaches have been used in earlier studies (e.g., Matte et al., 2013).

2.2.3. Stratified random sampling

Based on the percent area covered within the study area by each of the five urban VDCs, 100 sites were distributed and additional sites added to ensure that each VDC had a minimum of 16 sites leading to a total of 114 monitoring sites. 114 monitoring sites were distributed with almost equal number of sites assigned to each of the four strata of building and road density within a given VDC (Supplement 2). Only grid cells from the site assignment region with area $>20,000$ m^2 were eligible for site selection. Next, area-weighted random sampling was performed without replacement to select grid cells for monitoring sites. Using the area-weighted method, those grid cells with smaller areas (i.e., partial grid cells) had lower probability of being selected in comparison to those with larger areas. When a grid cell was selected, its neighboring grid cells were excluded from subsequent sampling. In Bhaktapur VDC, which has a small number of total grid cells, excluding neighboring grid cells of selected grid cells caused 4 sites (1 HB-LR cell, 3 LB-HR cells) to be added outside of the site assignment region or the study area after removing the area in the 400 m buffer at the boundary.

2.2.4. Purposeful sampling

After placement of 114 monitoring sites using stratified random sampling, an additional 21 sites were distributed in the study area. These 21 sites were allocated at least 300 m from the stratified random sampling sites already selected. Of the purposeful sites, 12 were allocated to major road intersections in the study area. The remaining 9 sites were placed to meet the following ranked criteria below (in the order performed): reduce major geographical gaps, cover regions with high population density, assign monitors in areas of concern, and spatially cover different land use types.

2.2.5. Field site identification

For each of the 135 sampling sites, suitable mounting structures were identified as close to the center of the grid cell as possible. When a grid cell selected for sampling was located in an inaccessible area, the nearest neighboring grid cell with the same stratum of building density and road density was sampled.

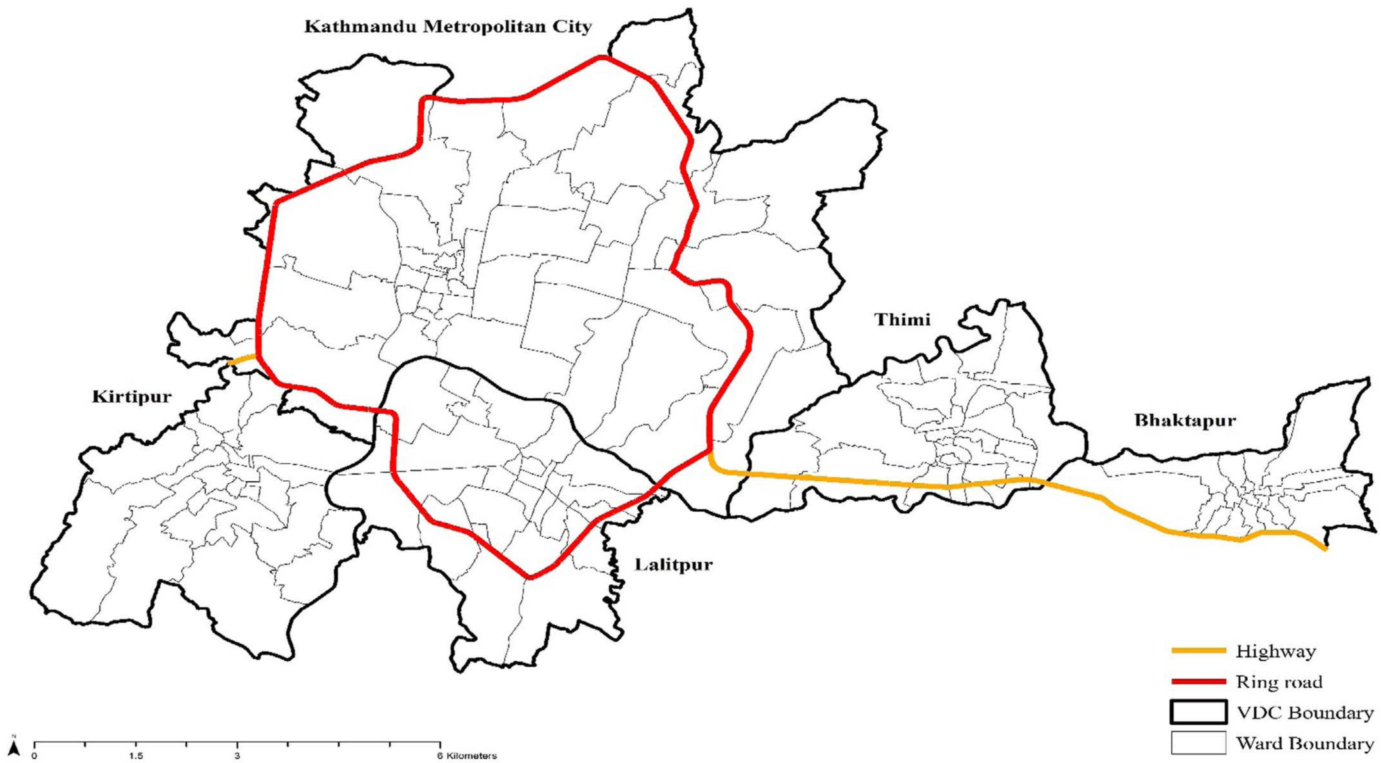

Fig. 1. Five urban VDCs in Kathmandu Valley, Nepal.

2.3. Temporal allocation of sites

Kathmandu Valley experiences subtropical, temperate climate with four distinct seasons a year: pre-monsoon (March – May); monsoon (June – Aug.); post-monsoon (Sep. – Nov.) and winter (Dec. – Feb.). Four sampling campaigns were performed in 2014, one for each season (Jan., May, July, and Nov.) for two weeks each, avoiding unusual events (e.g. major holidays). For a given campaign, samplers were installed within four days, and at the end of the campaign, were collected within four days. All samplers were collected 14 days after their installation in the same order as they were installed.

2.4. Monitoring of air pollutants

NO_2 was selected as the pollutant of interest given its wide use as a surrogate of traffic exposure in LUR modeling (Briggs et al., 1997; Jerrett et al., 2007; Lebret et al., 2000; Ross et al., 2006). Palmes tubes were used for NO_2 sampling at 135 sites with an average height of 2.5 m (breathing level). Each of the tubes was inside a protective cover, at least 20 cm away from any vertical surface and with no overhanging vegetation or buildings. Tubes were also placed at least 100 m away from major sources (e.g., construction work, exhaust flues, vents), 25 m from road intersections, and 2 m away from roadsides.

Duplicate Palmes tubes were used at each site as a measure of precision and backup in case of damage or loss (Aguilera et al., 2008; Briggs et al., 1997; Ross et al., 2006; Wheeler et al., 2008). When duplicate Palmes tubes were available for a given site, NO_2 concentration was calculated as an average of the two NO_2 measurements. Annual mean NO_2 concentration at each site was calculated as the average of the four campaigns. At 28 locations (7 for each of the strata types for building and road density) in Kathmandu VDC, Ogawa badges were used to measure NO, NO_x, and NO_2 at locations collocated with the Palmes tubes. Ogawa badge monitoring sites were randomly selected from the 66 sites allocated to Kathmandu VDC. NO_2 measurements by Ogawa badge were used to validate NO_2 measures made by Palmes tubes. NO_x and NO measurements by Ogawa badges were used to make comparison to NO_2 measures. For each sampling campaign and monitor type, 2 field blanks and 2 lab blanks were collected to correct for any impurities in samplers present before or after measurement (Aguilera et al., 2008; Henderson et al., 2007; Madsen et al., 2007; Su et al., 2008).

2.5. Predictor data

Predictor variables were computed for buffer sizes 25, 50, 75, 100, 150, 200, 250, 300, 350, and 400 m around each monitoring site. For the four Bhaktapur sites added outside of the site assignment region, when portions of the buffers fell outside of the study area, where data for predictor variables were unavailable except for Landsat data, we assumed that the distribution outside the study area (with no data available) was the same as the distribution within the study area (with data available). Table 1 describes the predictor variables including the source of data, units and hypothesized direction of effect in relation to NO_2 levels.

2.6. Statistical analysis

Coefficient of determination (R^2), absolute mean percent difference, and coefficient of variation (CV) were used to assess measurement error by comparing values for duplicate Palmes tubes (135 sites) and to assess coherence among measurement methods by comparing values for duplicate Palmes tubes and Ogawa badges (28 sites in Kathmandu VDC).

Univariate regression modeling was performed for each predictor variable in separate models for each buffer size (25–400 m) against the annual average NO_2 concentration. The univariate model with the highest adjusted R^2 was selected as the initial model. To this initial model, the remaining variables (those variables with ≤0.80 correlation with the initial predictor variable) were added separately and the effect on adjusted R^2 noted. The predictor variable with the highest increase in adjusted R^2 was maintained in the model if the increase in adjusted R^2 was $> 1\%$, the regression coefficient for that variable conformed to

Table 1
Potential predictor variables for land use regression modeling for five urban VDCs in Kathmandu Valley, Nepal.

Data source	Predictor variable	Unit	Hypothesized direction of effects on NO_2[b]
Road network			
MPPW[a]	Road length of all roads in a buffer	m	+
MPPW[a]	Road length of all roads in a buffer by road type	m	+
MPPW[a]	Distance to nearest major road	m	−
MPPW[a]	Distance to nearest road	m	−
Land use			
MPPW[a]	Slum area	m^2	+
MPPW[a]	Urban green (parks, cultivation, plantation)	m^2	−
MPPW[a]	Industrial area	m^2	+
MPPW[a]	Mixed area (commercial, residential)	m^2	+
MPPW[a]	Heritage area	m^2	+
MPPW[a]	Institutional area	m^2	+
MPPW[a]	Open area	m^2	−
Footprint of housing			
MPPW[a]	Built Area	m^2	+
Brick kilns			
CartoDB	Number of brick kilns	number	+
CartoDB	Distance to nearest brick kilns	m	−
Vegetation and soil reflectance			
Landsat	Brightness, greenness	n/a	+
Landsat	NDVI	n/a	−
Population			
Census 2011	Population density	person/m^2	+
Census 2011	Housing density	household/m^2	+
Elevation			
Field work	Elevation	m	−

[a] MPPW Ministry of Physical Planning and Works, Nepal/
[b] Here + indicates an anticipated association where a higher value of the predictor is associated with higher level of NO_2, while − indicates an anticipated association where higher value of the predictor is associated with lower level of NO_2.

the hypothesized direction (Table 1), the directions of regression coefficients for predictors already in the model did not change, and the variance inflation factor (VIF) of the variable did not exceed 3. Further addition of predictor variables in supervised stepwise process was performed until there were no remaining predictors meeting the above criteria. Finally, evaluation was performed on the significance of the variables in the model based on their associated regression coefficients. Any variables with p-value > 0.10 were sequentially removed from the model starting with the least significant (Aguilera et al., 2008; Beelen et al., 2013). This final model was selected as the LUR model for predicting annual ambient NO_2 concentration.

Sensitivity analysis was performed to see how interpolation of predictor data outside of the study area impacted selection of the final model. Cook's D was used to examine influential observations (Beelen et al., 2013; Jerrett et al.al., 2007). High Cook's D value (> 1) was taken to indicate an influential observation. In that case, the final model was rerun excluding the site with high Cook's D value. Then the individual site was excluded in subsequent analysis if the site had an influential observation based on Cook's D and the model changed substantially (model R^2 increase > 0.05) (Beelen et al., 2013). Monitor measurements and model residuals were plotted to evaluate whether residuals displayed heteroscedasticity. Residual spatial autocorrelation was evaluated to assess the independence assumption (spatial autocorrelation error violates the assumption) using Moran's I (Aguilera et al., 2008; Beelen et al., 2013; Jerrett et al.al., 2007). Two types of cross-validation were performed to evaluate the exposure model. The first method involved removing one of the sites, recreating the model and predicting the concentration at the omitted site. This procedure was performed for each of the sampling sites and root mean squared error (RMSE) was calculated (Brauer et al., 2003). The second method involved testing the model in three random selections of 85% of the sample to predict concentration at the remaining 15% (Aguilera et al., 2008).

3. Results

Kathmandu and Lalitpur had higher average building density and road density (Supplement 3 and 4, respectively) compared to the three other urban VDCs (Bhaktapur, Kirtipur, and Thimi) in the Valley. Supplement 2 summarizes the distribution of monitoring sites based on strata (four strata based on building density and road density), sampling method (stratified vs. purposeful), and five urban VDCs. Kathmandu VDC, being the only metropolitan city and covering the largest area (49% of the study area), has the largest number of total sites allocated (66 sites). Fig. 2 shows the exact location of monitoring sites within the five urban VDCs.

The proportion of sites with completed sampling by campaign was high for both Palmes tubes (90.4–95.6%) and Ogawa badges (78.6–100%). Agreement between collocated duplicate Palmes tubes was high (R^2=0.95). The absolute mean percent difference and CV for duplicate Palmes tubes across campaigns and sites were 12.0% and 13.8%. Agreement between collocated Palmes tube and Ogawa badges at 28 sites was high with R^2 value of 0.76 and absolute mean percent difference of 21.1%.

Annual average NO_2 by site (based on Palmes tubes) ranged from 5.7–120 ppb for the study area with maximum NO_2 observed in Kathmandu as shown in Table 2.

NO_2 concentrations as measured by Palmes tubes were lowest on average during monsoon and highest during pre-monsoon (Table 3). Similar NO_2 seasonal trends were observed based on Ogawa badges, with highest average NO_2 of 23.7 ppb during pre-monsoon and lowest average NO_2 of 9.9 ppb during monsoon (Supplement 5). Highest average NO and NO_x of 31.4 ppb and 57.2 ppb, respectively, were observed in winter. Lowest average NO and NO_x occurred during the monsoon season.

Table 4 shows the final LUR model for predicting intraurban variation in NO_2 for the urban area of Kathmandu Valley. Only 133 sites were included in the model; one site was excluded as measurements were not available for all four campaigns and one site was excluded due to high Cook's D value (> 1). Length of major road within 50 m, built area within 400 m, and industrial area within 75 m were observed to have positive associations with NO_2 concentration. NDVI within 50 m was negatively associated with NO_2 concentration.

Sensitivity analysis was performed by running the model using only 123 sites (including only sites where the entire buffer area for all buffers of 25–400 m were within the study area) and separately using 132 sites (including sites where the entire buffer areas were within the study area and those sites with buffer areas $< 25\%$ falling outside the study area for all buffers of 25–400 m). Both sensitivity analyses produced similar covariates and adjusted R^2 in comparison to the final model (Supplement 6).

For the final model, an analysis comparing monitor measurements to residuals showed absence of heteroscedasticity. Moran's I was not statistically significant ($p > 0.05$), showing no evidence of spatial autocorrelation and indicating that the included predictor variables account sufficiently for the spatial variation within each area (Eeftens et al., 2012). Two cross-validation analyses were performed to examine the validity of the regression models. The first method (removing one of the sites, recreating the model, and predicting the concentration at the omitted site) estimated RMSE of 6.3 ppb. The second method (model using 85% of the sample to predict concentration of at the remaining 15% sites) estimated RMSE of 5.0 ppb. Application of the final model

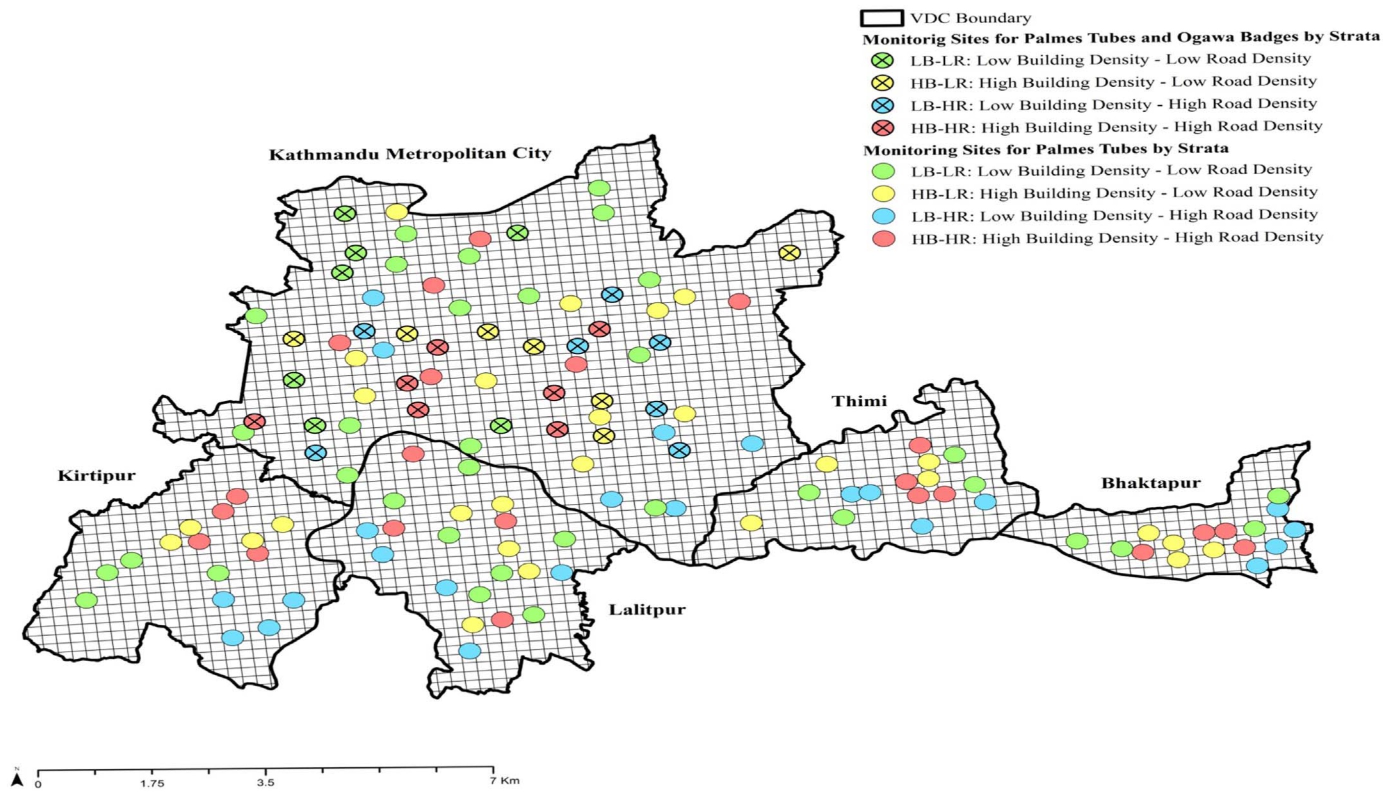

Fig. 2. Location of monitoring sites for five urban VDCs in Kathmandu Valley, Nepal. Note: Based on upper and lower three quartiles of road density and building density by VDC, grid cells were cross-classified into four strata. HB-HR: High Building Density – High Road Density; HB-LR: High Building Density – Low Road Density; LB-HR: Low Building Density – High Road Density; LB-LR: Low Building Density – Low Road Density. Each dot in the map is centered at the exact location of a monitoring site. For each sampling site, a suitable mounting structure was identified as close to the center of the grid cell as possible. When selected grid cells were inaccessible, the nearest neighboring grid cell with the same stratum was selected.

Table 2
Summary of annual average NO_2 concentrations (ppb) based on Palmes tubes by VDCs and study area for five urban VDCs in Kathmandu Valley, Nepal.

VDCs	Mean[a] ± SD (min, max)	Interquartile range (IQR)
Kathmandu	26.0 ± 13.6 (10.6, 120)	9.2
Lalitpur	25.0 ± 9.0 (12.1, 43.1)	12.8
Bhaktapur	16.0 ± 5.3 (7.3, 24.4)	7.1
Thimi	18.7 ± 4.2 (13.3, 25.2)	8.7
Kirtipur	11.0 ± 3.4 (5.7, 29.4)	3.9
Study Area (All VDCs)	22.1 ± 11.7 (5.7, 120)	11.8

[a] Mean refers to the annual average NO_2 based on all sites with NO_2 measures available during the four campaigns for a given VDC or the total study area. Some sites were missing data due to tubes being lost, voided or identified as an outlier.

Table 3
Summary of NO_2 measurements (ppb) based on Palmes tubes by season for five urban VDCs in Kathmandu Valley, Nepal.

Season	No of sites	Mean[a] ± SD (min, max)	Interquartile range (IQR)
Winter	133	27.9 ± 8.7 (1.6, 54.0)	9.6
Pre-Monsoon	128	29.9 ± 31.4 (5.6, 349)	14.2
Monsoon	124	15.2 ± 7.5 (3.4, 38.8)	9.1
Post-Monsoon	130	18.6 ± 7.7 (3.7, 49.5)	9.5

[a] Mean refers to the average NO_2 by season based on all sites with NO_2 measures available during the four campaigns. Some sites did not have data for all four campaigns due to tubes being lost, voided or identified as an outlier.

Table 4
Final land use regression model for five urban VDCs in Kathmandu Valley, Nepal.

Variable (n=133, Adjusted R^2 =0.51)[a]	Coefficient	Standard error	t	Probability > t
Intercept	27.1	1.89	14.3	< 0.000100
Length of major road within 50 m	0.100	0.0200	5.06	< 0.000100
NDVI within 50 m	−24.5	3.64	−6.74	< 0.000100
Built area within 400 m	0.0000310	0.0000160	1.94	0.0600
Industrial area within 75 m	0.000581	0.000174	3.34	0.00100

[a] Here 133 sites were included with interpolation of predictor variables for buffer areas falling outside of the study area for all buffers 25–400 m irrespective of the size of the buffer area falling outside the study area.

to each grid cell within the urban area of Kathmandu Valley yielded a spatial surface with significant variability between and within VDCs (Fig. 3).

4. Discussion and conclusion

We successfully developed and applied a cost-effective air quality monitoring campaign and used those measurements to develop an LUR model in urban areas of Kathmandu Valley. We observed high levels of NO_2 (annual average 22.1 ppb) with substantial within-city variation (5.7–120 ppb). By comparison, annual average NO_2 was 14.9 ppb in Southern California (range 8.1–28.7 ppb) (Ross et al., 2006), 12.4 ppb in Windsor, Ontario (range 6.9–20.2 ppb) (Wheeler et al., 2008), and 10.7 ppb in Ulaanbaatar, Mongolia (Allen et al., 2011). We found higher NO_2 levels in the two urban VDCs with higher building and road

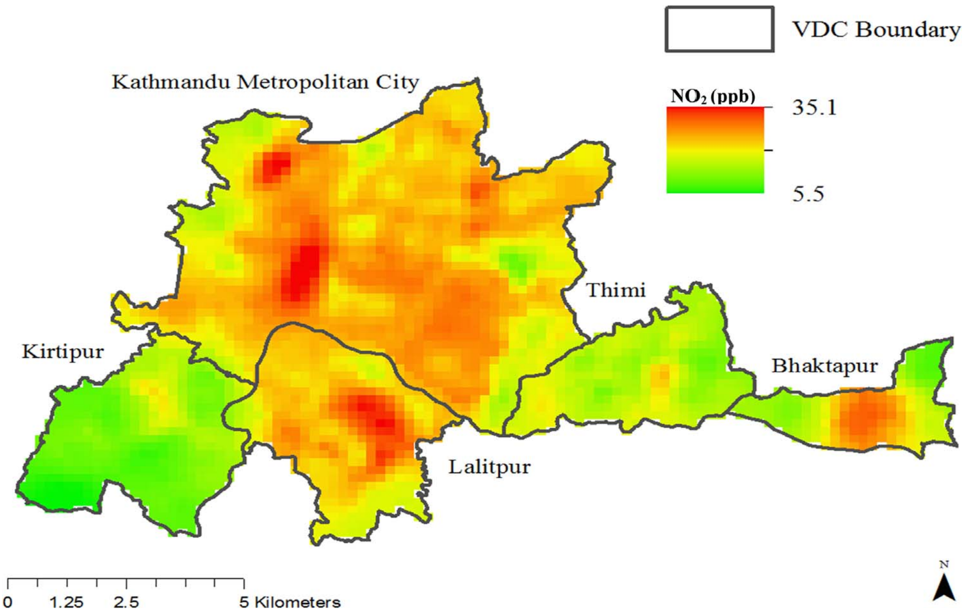

Fig. 3. Final NO_2 predicted surface for five urban VDCs in Kathmandu Valley, Nepal.

density (Kathmandu and Lalitpur) in comparison to the other three urban VDCs. The two VDCs with higher NO_2 also have higher population density (Statistics, 2012), which both contributes to NO_2 emissions and increases population exposures. Variability is also observed within each VDC, in agreement with the literature (Hoek et al., 2008).

Season had a large influence on NO_2 levels in the Kathmandu Valley, with lower concentrations during monsoon when higher precipitation and higher surface temperature may have promoted atmospheric conditions conducive to dispersing pollutants (Aryal et al., 2008; Giri et al., 2008). Similarly, lower NO and NO_x levels were observed during monsoon. Highest NO_2 was observed in pre-monsoon season, which may be associated with relatively less humidity. High NO and NO_x levels were observed during winter. During winter season, strong stable cold air is experienced in the Valley (Regmi et al., 2003) such that there is lower surface temperature and lower wind speed, restricting dispersion (Aryal et al., 2008).

In earlier studies, LUR modeling explained 60–70% of NO_2 variability based on predictors including traffic, population density, altitude, and land use (Hoek et al., 2008). In this study, the final model explained 51% of NO_2 variability using four predictors, and the model performed well in cross-validation analyses. Variations in R^2 among studies can arise due to variability in the measured concentrations, quality of the predictor data utilized, and modeling approach (Briggs, 2007). For developing cities like Kathmandu, data availability is a limitation. For example, we did not have a continuous reference site to adjust measured NO_2 values for temporal trends, so we instead focused on a simple average of the four seasonal measurements in our models. In addition, model variability reveals the complexity of cities and how cities differ in characteristics (e.g., topography, emission sources) (Briggs, 2007).

The relationships between NO_2 and covariates were similar to those observed in LUR models for other regions. As in our study, NO_2 levels were associated with length of major roads in Amsterdam (Briggs et al., 1997), Montreal (Gilbert et al., 2005), Toronto (Jerrett et al.al., 2007), and Windsor, Ontario (Wheeler et al., 2008). Like in many LUR studies, no traffic count data were available for our study area (Brauer et al., 2003; Henderson et al., 2007; Madsen et al., 2007). Similarly, our finding regarding the significance of built area is in agreement with studies in Amsterdam, Prague, and Huddersfield, England (Briggs et al., 1997) and in New York City (Clougherty et al., 2013). Associations with industrial land use were also seen in Toronto (Jerrett et al.al., 2007) and in Hamilton, Ontario (Sahsuvaroglu et al., 2006). Our finding regarding the significance of NDVI reinforces use of remote sensing data with global coverage and its usage in places like Barcelona, Spain (Dadvand et al., 2015) and Tianjin, China (Chen et al., 2012).

In general, our study demonstrates the application of LUR modeling approaches developed for cities in Europe and North America to urban areas in developing countries. Knowledge of intraurban variation of pollutants is important to avoid potential misclassification (Ito et al.al., 2005; Pinto et al., 2004) and errors associated with estimation of exposure and subsequently health effects (Cyrys et al., 2012). Studies using more simplistic methods of exposure, such as distance to roadways as a proxy of exposure, are useful but have potential for exposure misclassification and do not include the importance of other variables that may influence exposure. On the other hand, methods such as dispersion models are very data intensive, and comprehensive and high-resolution emissions inventories and meteorological fields are not readily available in settings like Kathmandu. For places like Kathmandu Valley, Nepal, LUR modeling is an attractive approach, as it does not require as many data inputs as dispersion models but can include various factors important for understanding exposure (Ross et al., 2006). While LUR models contain appreciable uncertainties, including the possibility of bias in predictions given the wide range of measured concentrations, they likely capture the general spatial contrast across the area and provide a foundation for more refined models in the future.

The output from the LUR model can be utilized to investigate vulnerable populations and development of policies related to land use, road design, air pollution control, and human health protection. Our study reinforces the value of cost-effective monitoring campaigns coupled with statistical analysis within developing cities.

Acknowledgement

This work was supported by Environment Protection Agency (Grant number: FP-91765001-0) Science to Achieve Results Fellowship, Yale Institute of Biospheric Studies, and Hixon Center for Urban Ecology.

The authors declare they have no actual or potential competing financial interests.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.envres.2017.01.038.

References

Aguilera, I., Sunyer, J., Fernandez-Patier, R., Hoek, G., Aguirre-Alfaro, A., Meliefste, K., et al., 2008. Estimation of outdoor NO_x, NO_2, and BTEX exposure in a cohort of pregnant women using land use regression modeling. Environ. Sci. Technol. 42, 815–821.

Allen, R.W., Gombojav, E., Barkhasragchaa, B., Byambaa, T., Lkhasuren, O., Amram, O., et al., 2013. An assessment of air pollution and its attributable mortality in Ulaanbaatar, Mongolia. Air Qual. Atmos. Health 6, 137–150.

Aryal, R.K., Lee, B.-K., Karki, R., Gurung, A., Kandasamy, J., Pathak, B.K., et al., 2008. Seasonal PM_{10} dynamics in Kathmandu Valley. Atmos. Environ. 42, 8623–8633.

Asian Development Bank CAI-A, 2006. Country Synthesis Report on Urban Air Quality Management. Asian Development Bank, Clean Air Initiative-Asia, Philippines.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., et al., 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. Atmos. Environ. 72, 10–23.

Bell, M.L., 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia region of the United States. Environ. Int. 32, 586–593.

Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., 2003. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. Epidemiology 14, 228–239.

Briggs, D.J., 2007. The use of GIS to evaluate traffic-related pollution. Occup. Environ. Med. 64, 1–2.

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., et al., 1997. Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 11, 699–718.

Chen, L., Wang, Y., Li, P., Ji, Y., Kong, S., Li, Z., et al., 2012. A land use regression model incorporating data on industrial point source pollution. J. Environ. Sci. 24, 1251–1258.

Chen, L., Bai, Z., Kong, S., Han, B., You, Y., Ding, X., et al., 2010. A land use regression for predicting NO_2 and PM_{10} concentrations in different seasons in Tianjin region, China. J. Environ. Sci. 22, 1364–1373.

Clougherty, J.E., Kheirbek, I., Eisl, H.M., Ross, Z., Pezeshki, G., Gorczynski, J.E., et al., 2013. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: the New York City Community Air Survey (NYCCAS). J. Expo. Sci. Environ. Epidemiol. 23, 232–240.

Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., et al., 2012. Variation of NO_2 and NO_x concentrations between and within 36 European study areas: results from the escape study. Atmos. Environ. 62, 374–390.

Dadvand, P., Rivas, I., Basagana, X., Alvarez-Pedrerol, M., Su, J., De Castro Pascual, M., et al., 2015. The association between greenness and traffic-related air pollution at schools. Sci. Total Environ. 523, 59–63.

Dionisio, K.L., Arku, R.E., Hughes, A.F., Vallarino, J., Carmichael, H., Spengler, J.D., 2010. Air pollution in Accra neighborhoods: spatial, socioeconomic, and temporal patterns. Environ. Sci. Technol. 44, 2270–2276.

Eeftens, M., Beelen, R., Hoogh, K.E., Bellander, T., Cesaroni, G., Cirach, M., et al., 2012. Development of landuse 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. Environ. Sci. Technol. 46, 11195–11205.

Gilbert, N.L., Goldberg, M.S., Beckerman, B., Brook, J.R., Jerrett, M., 2005. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada with a landuse regression model. J. Air Waste Manag. Assoc. 55, 1059–1063.

Giri, D., Krishna Murthy, V., Adhikary, P.R., 2008. The influence of meteorological conditions on PM_{10} concentrations in Kathmandu Valley. Int. J. Environ. Res. 2, 49–60.

Gurung, A., Bell, M.L., 2013. The state of scientific evidence on air pollution and human health in Nepal. Environ. Res. 124, 54–64.

Han, X., Naeher, L.P., 2006. A review of traffic-related air pollution exposure assessment studies in the developing world. Environ. Int. 32, 106–120.

Health Effects Institute, 2011. Concentration of Air Toxics in Motor Vehicle – Dominated Environments. Health Effects Institute, Boston, MA.

Henderson, S., Beckerman, B., Jerrett, M., Brauer, M., 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environ. Sci. Technol. 41, 2422–2428.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., 2008. A review of landuse regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42, 7561–7578.

International Centre for Integrated Mountain Development, 2006. Kathmandu Valley: Environment Outlook, Kathmandu, Nepal

Ito, K., De Leon, S., Thurston, G., Nadas, A., Lippmann, M., 2005. Monitor-to-monitor temporal correlation of air pollution in the contiguous us. J. Expo. Sci. Environ. Epidemiol. 15, 171–184.

Japan International Cooperation Agency, 2011. Data Collection Survey on Traffic Improvement in Kathmandu Valley: Inception Report. Japan International Cooperation Agency, Kathmandu.

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

Kashima, S., Yorifuji, T., Tsuda, T., Doi, H., 2009. Application of land use regression to regulatory air quality data in Japan. Sci. Total Environ. 407, 3055–3062.

Lebret, E., Briggs, D., Reeuwijik, H.V., Fischer, P., Smallbone, K., Hendrik, H., et al., 2000. Small area variations in ambient NO_2 concentrations in four European areas. Atmos. Environ. 34, 177–185.

Li, C., Du, S.Y., Bai, Z.P., Shao-fei, K., Yan, Y., Bin, H., Dao-wen, H., Li, Z.Y., 2010. Application of land use regression for estimating concentrations of major outdoor air pollutants in Jinan, China. J. Zhejiang Univ. Sci. A 11, 857–867.

Madsen, C., Carlsen, K.C.L., Hoek, G., Oftedal, B., Nafstad, P., Meliefste, K., 2007. Modeling the intra-urban variability of outdoor traffic pollution in Oslo, Norway—a ga2len project. Atmos. Environ. 41, 7500–7511.

Matte, T.D., Ross, Z., Kheirbek, I., Eisl, H., Johnson, S., Gorczynski, J.E., et al., 2013. Monitoring intraurban spatial patterns of multiple combustion air pollutants in New York City: design and implementation. J. Expo. Sci. Environ. Epidemiol. 23, 223–231.

Obaid, T., 2007. State of World Population 2007. United Nation Population Fund

Pinto, J.P., Lefohn, A.S., Shadwick, D.S., 2004. Spatial variability of $PM_{2.5}$ in urban areas in the United States. J. Air Waste Manag. Assoc. 54, 440–449.

Poplawski, K., Gould, T., Setton, E., Allen, R., Su, J., Larson, T., et al., 2009. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. J. Expo. Sci. Environ. Epidemiol. 19, 107–117.

Regmi, R.P., Kitada, T., Kurata, G., 2003. Numerical simulation of late wintertime local flows in Kathmandu Valley: implication for air pollution transport. J. Appl. Meteorol., 389–403.

Ross, Z., English, P.B., Scalf, R., Gunier, R., Smorodinsky, S., Wall, S., et al., 2006. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses. J. Expo. Sci. Environ. Epidemiol. 16, 106–114.

Sahsuvaroglu, T., Arain, A., Kanaroglou, P., Finkelstein, N., Newbold, B., Jerrett, M., et al., 2006. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. J. Air Waste Manag. Assoc. 56, 1059–1069.

Seto, K.C., Sánchez-Rodríguez, R., Fragkias, M., 2010. The new geography of contemporary urbanization and the environment. Annu. Rev. Environ. Resour. 35, 167–194.

Statistics CBo, 2012. Preliminary Results of National Population and Housing Census. Census Bureau of Statistics, Kathmandu, Nepal.

Su, J.G., Brauer, M., Buzzelli, M., 2008. Estimating urban morphometry at the neighborhood scale for improvement in modeling long-term average air pollution concentrations. Atmos. Environ. 42, 7884–7893.

Wheeler, A.J., Smith-Doiron, M., Xu, X., Gilbert, N.L., Brook, J.R., 2008. Intra-urban variability of air pollution in Windsor, Ontario–measurement and modeling for human exposure assessment. Environ. Res. 106, 7–16.

World Health Organization, 2014. 7 Million Premature Deaths Annually Linked to Air Pollution. World Health Organization.

Zou, B., Wilson, J.G., Zhan, F.B., Zeng, Y., 2009. Air pollution exposure assessment methods utilized in epidemiological studies. J. Environ. Monit. 11, 475–490.