Doctoral Dissertation

Submitted in fulfillment of the requirements for

the degree of doctor

at Yokohama National University

Development of Land Use Regression (LUR) model for Nitrogen Dioxide (NO_2) pollution in Lahore, Pakistan

Name: Saimar Pervez

Student ID: 19TG901

Supervisor: Professor Satoshi Nakai

Department: Graduate School of Environment and Information Sciences

Major: Artificial Environment

Acknowledgment

First, I would like to thank Almighty ALLAH (S.W.T), the most merciful, the most gracious, for giving me an opportunity to complete Ph.D. and giving me courage and patience throughout the course of this study.

I would like to express my profound and sincere gratitude to my supervisor Professor Satoshi Nakai for giving me the opportunity to do research and giving valuable guidance throughout my course work and research. His motivation and stimulating suggestions gave me encouragement during the study. It was an honor to research under his guidance. I would also like to thanks to my thesis committee members (Prof. Takashi Kameya, Dr. Hiroki Hondo, Prof. Takeshi Kobayashi, and Dr. Tadahiro Shibutani) for giving me the valuable suggestions and their comments. I gratefully acknowledge Japanese Government (MEXT), Japan for the scholarship and for the research grant to complete the Ph.D. thesis.

Deepest and sincere gratitude goes to my beloved parents, family, brothers and sister for their endless love, prayers, and support. I especially want to express my unbound thanks to my friend and research colleague, Xiao Jun, for her constant encouragement.

Finally, I would like to express my gratitude to all lab fellows for their support during this Ph.D. journey and research work.

Saimar Pervez

Abstract

Pakistan is facing a rapid increase in air pollution. Urban air quality is deteriorating day by day and the people living in the cities facing problems because of air pollution exposure. With the increase in population, industrial development and economic prosperity, air pollution is one of the biggest concerns in Pakistan after the occurrence of recently smog episodes. Especially during the winter, there is a danger from smog and health of the public is at stake. Seasonal variation (Monsoon variation) does affect the air quality and can increase the agglomeration of air pollution, causing adverse situation and problems for the community.

Pakistan being a developing country and it is very difficult to access and monitor air quality throughout the city area due to unavailability of monitoring instruments and lack of funds. Based on some monitoring locations, it is difficult and unreliable to access the exposure from air pollution for the whole city, thus giving a possible solution to cost-effectively option of air pollution modelling, which can provide a reasonable and reliable solution for air pollution exposure assessment and providing the possible solutions to overcome the issue of air pollution mitigation and improvement. Recently, Land use regression (LUR) has gained the importance of an easy, cost-effective, and reliable modeling solution for air pollution exposure assessment. Therefore, in this research work, an attempt was made to develop a LUR model and identification of air pollution hotspots areas with the addition of possibly potential predictor variables affecting the air quality status in the city of Lahore, Pakistan.

The purpose of this study was to develop a Land use regression (LUR) model to provide a better understanding of air exposure and to depict the spatial patterns of air pollutants within the city. Land use regression model was developed using the average seasonal concentration of Nitrogen

Dioxide (NO_2) and considering twenty-two potential predictor variables including road network, land use classification and local specific variable (vehicle maintenance workshop, VMW).

Ambient air pollution can increase the possibility of health risks especially in urban areas of developing countries having the mixture of different air pollution sources. Identification of possible pollution sources within the cities is an important task. Different pollution sources present and yet to be determined in the developing countries, like Pakistan. In this study, Vehicle maintenance workshops (VMW) was considered and identified a possible source of pollution. Distance variable (DV) was used to represent the presence of VMW within the city. All the three developed LUR seasonal (pre-monsoon, monsoon, post-monsoon) models identified the presence of DV, which justified the local pollution source i.e., VMW.

Survey was conducted to gather the information about the workshops present in the study area and their inclusion in the final models improved the efficiency and overall fit of the LUR models. Road network (road length), residential area and distance variable were the influencing factors in the final models, showing the positive association with the pollutant (NO_2) concentration. Spatial characteristics of the predicted concentration showed the city center as the pollution hotspot, surrounding with the road network and residential area, followed by the presence of local vehicle maintenance workshops. Adjusted explained variance of the LUR models was highest for post-monsoon (77%), followed by monsoon (71%), and was lowest for pre-monsoon (70%). This is the first study conducted in Pakistan to explore the applicability of LUR model within a city and hence will offer the application in other cities. The results of this study would also provide help in promoting epidemiological research in future.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation description:

LUR: Land use regression

NO_2: Nitrogen Dioxide

SAP: Secondary air pollutants

GHG: Greenhouse gases

WHO: World Health Organization

EPA: Environmental protection agency

GIS: Geographic Information system

ESCAPE: European Study of Cohorts for Air Pollution Effects

VMW: Vehicle maintenance workshop

SLDS: Specific local data survey

RMSE: Root mean square error

LOOCV: Leave one out cross validation

VIF: Variance inflation factor

CHAPTER 1

1. Introduction

This chapter explains the research background and the importance of this research work. Air quality situation in Pakistan and its related sources have been introduced with the addition of problem statement and objectives of this study.

1.1. Background:

Pollution in developing countries is a very serious issue, especially in big cities. Lahore is the second largest city in Pakistan with the population of around 11.12 million (Pakistan Bureau of Statistics, 2017). Lahore has a semi-arid climate, and the presence of large number of vehicles and industries makes the city's air difficult to breath. This situation gets worse in the winter season, when the smoke and fog mix up in the air and forms a brown haze layer called "smog", causing adverse effects on the people health, and hampering the daily routine of public including the closure of schools.

Reported in the literature, Nitrogen Dioxide (NO_2) is the second most dangerous pollutant, after Particulate matter ($PM_{2.5}$) in Pakistan and its reported levels are much higher than the international standards given by World Health Organization (WHO) (Colbeck, I. et al., 2010). NO_2 is considered as a pre-cursor for the generation of photo-chemical smog and secondary air pollutants (Wang, W. et al. 1995). This pollutant is also corelated well with the traffic densities, giving an indication of increase in the number of vehicles. Number of vehicles increasing in Lahore due to population growth and urbanization.

Pollution in developing countries, predominantly in urban centers is a major issue for public health due to dramatically growing population, automobile numbers, urbanization, and heavy industry. Therefore, urban populations are facing serious health implications such as lung cancer, pneumonia, and acute lower respiratory infections especially for women and children (Jamison, Breman et al. 2006). A large number of studies shows that ambient air pollution is associated with significant adverse effects on public health (Pope III, Dockery et al. 2006). The studies have shown that concentration of NO_2, SO_2 and CO exceeds the National Environmental Quality Standards (NEQS) and one of the studies showed that $PM_{2.5}$ and PM_{10} levels were 6 times higher than the World Health Organization (WHO) guidelines (Niaz et al., 2015; Majid et al., 2012). Pollutants having health concerns in developing countries include particulate matter (PM), nitrogen dioxide (NO_2) and ozone (Brunekreef and Holgate 2002).

Children are more susceptible to diseases because they spend more time on roads playing and in schools which are situated in areas with excessive levels of different air pollutants and because of their physical activities and higher breathing rates (Trasande, Thurston et al. 2005). Different studies show that ambient air pollution can may cause inflammation of the central nervous system which can affect the behavior and performance of school going children (Block, Elder et al. 2012). Ambient air pollution is associated to 3.7 million premature deaths per annum and most of these (88%) are in low and middle income countries (sheet 2014).

The increase in number of road traffic and rapid growth of urbanization poses a great health hazard for the surrounding environment and public health. Ambient air pollution is an important environmental issue (WHO, 2016). Each year, two million or more deaths are attributed in the world because of air pollution exposure (Tan, X. et al. 2021). Air pollution is one of the biggest concerns in the modern era because of luxurious lifestyle, which requires more energy and exploration of resources, putting pressure on the generation of toxic air pollutants in the atmosphere. The emission of these air pollutants affects both, climate, and human health (Ashfaq, A. et al. 2012; Pecl, G.T. et al. 2017). Studies have mentioned the effects of air pollution on human health, such as cardiovascular, respiratory, and chronic diseases in adult stage (Karl, T.R. et al. 2009; Kelishadi, R. et al. 2010). One of the Swedish cohort studies reports that exposure to long term air pollution may cause diabetes (Eze, I.C. et al. 2014). Poor air quality is a serious issue in developing countries because of overpopulation, urbanization, and industrialization (Mannucci, P.M. et al. 2017).

1.2. Air pollution in Pakistan:

Air quality in Pakistan is a serious issue that needs to address, and it is continuously deteriorating. A major source of air pollution in Pakistan is emissions from motor vehicles and automobiles. Other sources of ambient air pollution include coal burning, industries and household fuel combustion for domestic use (Ilyas, Nasir et al. 2008). In Pakistan since 2001 –2002 the number of vehicles has grown by 130.3% (Pakistan Economic Survey, 2013-2014). Particulate matter is the most hazardous pollutant. After particulate matter, NO_2 is the second most dangerous pollutant in Pakistan and its reported levels are much higher than the international standards given by WHO (Colbeck, I. et al., 2010).

Pakistan is a South Asian country having a population crossing the figure of 200 million (Rana, I.A. et al. 2018). The rapid increase in the population and unplanned urbanization with the recent development in the industrial units, has worsen the condition of ambient air in the country (UNEP, 2011; Rasheed, A. et al. 2015). Transportation is another source of air pollution emitting 25 times more carbon monoxide (CO) and carbon dioxide (CO_2), and 3.5 times higher sulfur dioxide (SO_2)

as compared to the automobiles in the United States (Barber, N. 2008). Pakistan is the second, among the top ten polluted countries in the world accounting 22,000 premature deaths and 163,432 disability-adjusted life years (DALYS) lost (Ameratunga, S. et al. 2006). A study conducted in the cities of Pakistan (Islamabad, Lahore, Rawalpindi) reported the levels of Nitrogen oxides (NO_x) and Particulate matter (PM_{10}) higher than WHO guidelines (Colbeck, I. et al. 2010). Pakistan Environmental Protection Agency (Pak-EPA) has monitored the level of NO_2 in different cities of Pakistan and estimated that maximum and minimum concentrations were 37.02 ppb and 14.61 ppb in Karachi and Islamabad, whereas another study found that the maximum and minimum concentrations were 37.46 ppb and 2.48 ppb respectively (Colbeck, I. et al. 2010; Shabbir, Y. et al. 2016). Urban air pollution costs the country a loss of about Rs. 65 billion, from total annual loss to the environmental damages which is Rs. 365 billion (Ameratunga, S. et al. 2006). The financial loss occurs from morbidity and mortality linked with cardiovascular and respiratory diseases, lower respiratory illness (LRI) in children, however, if air pollution related problems are taken into consideration on education, malnutrition, and earnings, the financial loss could be higher (Sanchez-Triana, E. et al. 2014).

Road traffic is classified as one of the reasons behind the deterioration of air quality in urban areas (Mage, D. et al. 1996). Increase in the number of vehicles not only causing traffic congestion and greenhouse gas emissions, but imposing significant health impacts, and a source of tropospheric ozone formation (Colvile, R. et al. 2001). Different air pollutants are present in the atmosphere of urban environment like Particulate matter (PM), Nitrogen Dioxide (NO_2), carbon monoxide and dioxide (CO & CO_2) and ozone (O_3), but the pollutant which correlates well with traffic densities, and an important photochemical oxidant is Nitrogen Dioxide (WHO, 2006; Hewitt, C.N. 2001). Nitrogen dioxide (NO_2) has a major role for the generation of secondary air pollutants (SAP) and its concentration is correlated with photochemical smog, acid deposition and ozone variations (Yu, B. et al. 2017). Reported in the literature, after particulate matter, NO_2 is the second most abundant and dangerous air pollutant in Pakistan (Colbeck, I. et al. 2010), therefore it is necessary to have the knowledge of different types of pollution sources and the factors affecting the concentration of pollutant within the city.

One of the studies conducted by Pakistan Environmental Protection Agency (Pak-EPA) with the collaboration of Japan International Cooperation Agency (JICA) in 2000 in the three cities of Pakistan and found that the level of pollutants was higher than the WHO guidelines **(Table 1)** (Pak-EPA/JICA, 2000)

Table 1. Level of pollutants in different cities of Pakistan

Pollutant (Unit)	**City Name**		
	Lahore	Islamabad	Rawalpindi
PM_{10} (ug/m^3)	895	709	520
CO (ppm)	2.82	1.83	1.55
SO_2 (ppb)	44.6	30.7	28.5
NO_2 (ppb)	156.60	74.70	148.50
O_3 (ppb)	8.50	17	10

1.3. Air pollution in Lahore:

Lahore is the largest city in the Punjab Province, and its day-by-day expansion due to migration of people from small cities making the city's air a serious problem. Increase in the annual growth of population, vehicles and industrial units making the air pollution of Lahore is of concern. Other sources of pollution in the city are Brick kilns and vehicle maintenance workshops (VMW), which mostly uses the coal for the process, generating variety of pollutants in the atmosphere. Emissions from vehicle exhaust mostly come to workshops for the maintenance of vehicles e.g., cars, buses and rickshaws are also a significant source of air pollution in the city. Nitrogen dioxide (NO_2) concentration correlates well with the traffic densities (Hewitt, C.N. 2001), increasing number of vehicles may increase the concentration of NO_2. One of the studies conducted in five cities of Pakistan by Pak-EPA with the collaboration of JICA in 2006 and found that the concentration of NO_2 in Lahore was the second highest after Karachi **(Figure 1)** (Pak-EPA/JICA, 2006)

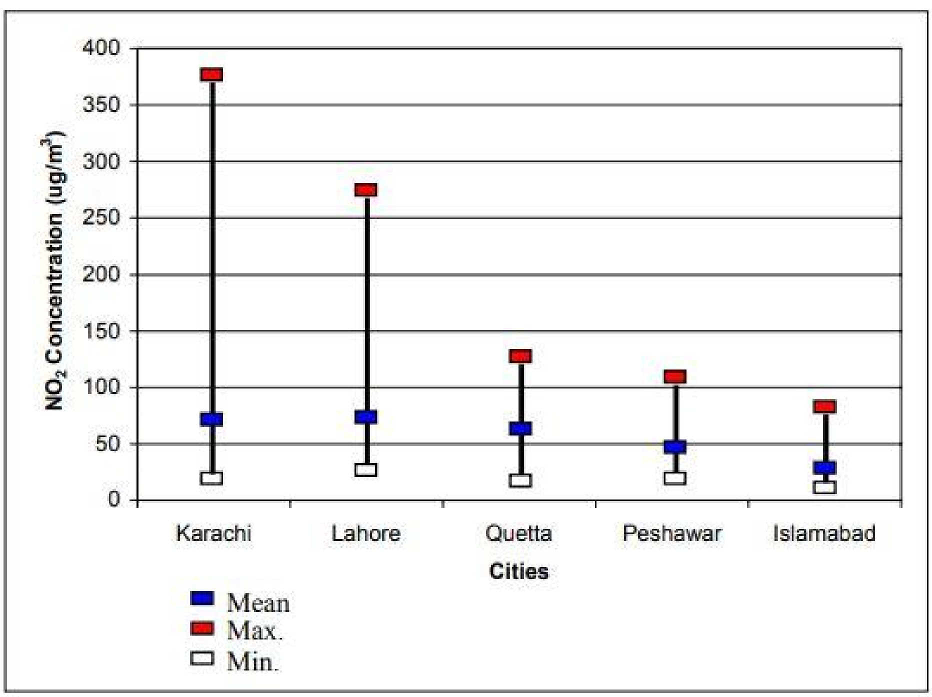


Figure 1. Nitrogen Dioxide pollution in different cities of Pakistan

1.4. Problem statement:

Pakistan is a developing country with an increase in population and economic development. With the rapid increase in industries and vehicles, air pollution in cities is of serious concern. Increase in population and automobiles results in the generation of variety of pollutants. The pollutant which correlates well with traffic densities is, NO_2 and it is also a pre-cursor of photochemical smog.

Exposure assessment to air pollution at city scale is important due to the increase in the number of vehicles within the cities and around the highways and major roads. Pollutant, like NO_2 being associated with the traffic, the concentration varies within the city and even within distances of 50 m or less (Hewitt, 1991). It is necessary to have such studies to assess the exposure from ambient air pollution to have the knowledge of hotspot area of pollution within a city.

Pakistan being a developing country and monitoring of air pollutants within the city is scarce due to limited availability of technical instrument and funding. Based on some monitoring stations, it is very difficult to access the pollution level for the whole city. Modeling software's has been used for the simulation of air quality in the developed countries, but the problem lies with the use of these models is the cost and complexity involved. Recently, Land use regression (LUR) modeling has also been used in developed countries and it provided a reasonable justification of exposure of pollutants within the cities. LUR model is known as feasible and cost-effective approach to be used for the air pollution modeling, but its applicability is still unclear for a specific city of a developing country, like Pakistan. To fill this research gap, LUR was developed for the city of Lahore, Pakistan.

Seasonal variations emissions occurred both due to shift in the season and meteorological conditions. Change in the seasons tends to affect the quality of air, dispersion and movement of pollutants (Ramachandran et al., 2003). Variation in the temperature and rainy season might affect the concentration of pollution, because when the temperature is low it is difficult for the pollutants to disperse, and they trap near the earth surface and vice versa for the increase in temperature. Decrease in temperature also shifts the usage of burning of fuel to keep the house warm and for heating purpose. Change in these conditions may lead to generate variety of pollutants based on seasonal variations and effect the human health. Therefore, in this study variations in seasonal pollutant concentration were studied to have the better and clear understanding of pollutant level in each season.

To address the need of air pollution exposure assessment, this study will explore the applicability of LUR model to predict spatial variation of NO_2 for Lahore, Pakistan in which emission sources include local industries, household fuel use and automobiles. This would not only provide the evidence of developing LUR models for different air pollutants in Pakistan, but also offer important application of exposure assessment with high representativeness. This study will provide a basic systematic LUR method, which means that this study will serve as a baseline in terms of the identified potential predictor variables that can be explored in other cities with the addition of local pollution sources and hence will promote the wider use in other cities of Pakistan.

1.5. Techniques and model for air pollution:

Different models and techniques are available and has been used to access and model the air pollution concentration **(Figure 2)**, such as Dispersion models (DM), chemistry transport model (CTM), and other techniques like Inverse distance weighting (IDW), Ordinary kriging (OK), Machine learning (ML), but the problem lies with the application of these models and techniques, is the high demand of data requirement, costs and the complexity (Briggs, D. 2007; Ren, X. et al. 2020;Simic, I. et al. 2020). The issue lies with the applicability of interpolation techniques is the assumption that variation in the pollution is dependent on the distance between the sites, which may lead to error in estimating pollution (Mulholland, J.A. et al. 1998). In contrast, Land Use Regression (LUR) model has gained the attention as an easy and effective approach, to provide spatial distributions of air pollutants at intra-urban scale built on a specific number of monitoring sites and predictor variables values, gathered by geographic information system (GIS), (de, Hoogh, K. et al.2014; Zou, B. et al. 2009), providing a reliable solution for air pollution exposure assessment for a developing country like Pakistan (Shi, Y. et al. 2020).

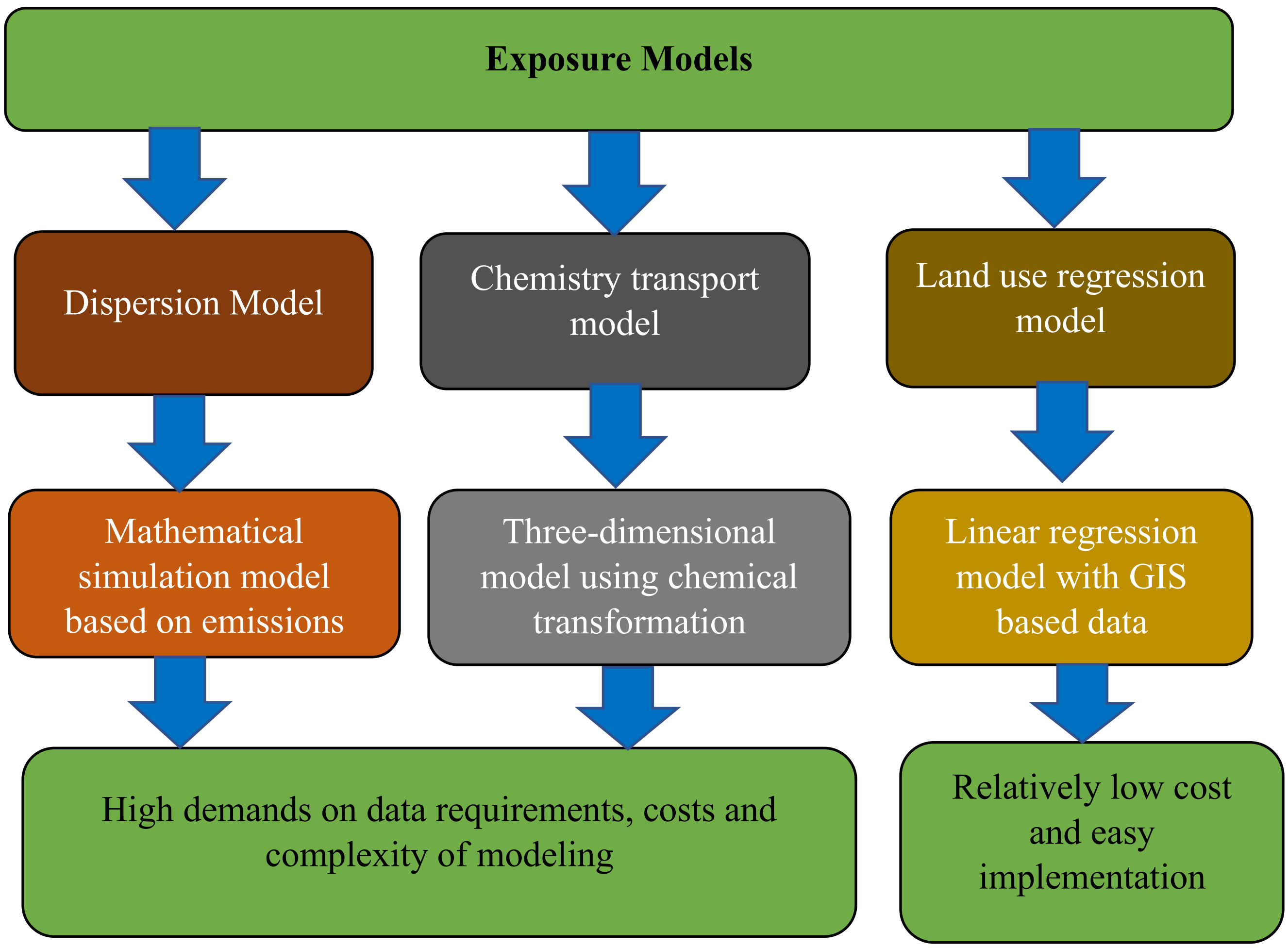


Figure 2. Different exposure models and their differences

1.6. Objectives of the study:

Mostly studies conducted in Pakistan related to air pollution just present the levels of pollutants. There is a lack of knowledge regarding air pollution exposure and the possible factors effecting the concentration within the city.

The aim of the present study was to develop LUR model based on seasonal variation [Pre-monsoon (April, May, and June), Monsoon (July, August), and Post-monsoon (September, October, and November)] of NO_2 concentration for Lahore city. This research work tried to develop and apply LUR for the first time in the specific city of Pakistan, with the limited availability of pollutant concentration data and the difficulties to get the data of potential predictor variables and of the local pollution source i.e., Vehicle maintenance workshop (VMW).

To have the knowledge of different pollution variables effect on the concentration of pollutant for each season, and to depict its spatial distribution, LUR was applied in Pakistan. It is necessary to compare the three period models for better understanding of local emission sources and the effect of different potential predictor variables for each season because change in season affect the dispersion of pollutants and quality of air within the city. Therefore, in this study variation in seasonal pollutant concentration were studied to have the better understanding of pollutant level in each season.

In this study, a possible source of pollution was identified for the city of Lahore, which was vehicle maintenance workshops (VMW). Due to the lack of exposure data in Pakistan, this study tried to provide exposure assessment for the Lahore city, and which will provide the long-term epidemiological studies in air pollution in future.

1.7. Justification of the research work:

Air pollution is a serious threat and its exposure to humans causes significant health problems. Different exposure models are available in the literature and have been used in different countries, but the problem lies with their applicability is the complexity and cost involved (Ren, X. et al. 2020). Land use regression model has recently gained a reputation as one of the easy and cost-effective method for the air pollution exposure assessment technique (Zou, B. et al. 2009). LUR model has been used in developed countries, but its applicability still yet to be explored in developing countries, especially in Pakistan with the limited availability of air quality monitoring network and lack of air pollution data. Different sources of air pollution are present within the cities and their identification is an important task. In this study, a local source of pollution, vehicle maintenance workshop was identified and found to be a significant source in all the three seasons. Final developed LUR models showed a positive correlation with the road network, residential area and the VMW.

CHAPTER 2

2. Literature Review

2.1. Overview of air pollutants:

Different air pollutants are present in the atmosphere of urban environment like Particulate matter (PM), Nitrogen Dioxide (NO_2), carbon monoxide and dioxide (CO & CO_2) and ozone (O_3) which may cause the health effects and even deaths greater than respiratory and cardiovascular diseases (Sahsuvaroglu et al., 2006; Garrett & Casimiro, 2011; World Health Organization, 2000). These pollutants have different type of sources and generation of the pollutants from these sources resulting in the formation of various problems in the atmosphere, the main one which is smog, a brown haze layer in the atmosphere during the winter season reducing the visibility, hammering every aspect of life, and carrying serious human health effects (Lodhi et. al., 2009). **(Table 2)** showing the most prominent air pollutants in the order of their occurrence in the atmosphere of Pakistan.

2.2. Sources and Health impacts of air pollutants:

The prominent sources of air pollution include fossil fuel burning in household, industrial emissions, vehicular exhaust, emissions from brick kilns, burning of solid waste and dust generation from construction activities. Many studies have classified industrial pollution, indoor air pollution, suspended particulates and increase in traffic density as the important sources for the degradation of ambient air quality in the country (World Bank, 2006; Pakistan Millennium Development Goals Report 2005). The emissions of greenhouse gases (GHG) mostly from power sector and the emissions of Carbon dioxide (CO_2) likely to reach 250 Mt by 2020 which may grow to 650 Mt (Ahmad et al., 2011).

Table 2. Prominent pollutants in the atmosphere of Pakistan

Pollutant	**Sources**	**Reference Study**
Particulate Matter (PM)	Combustion of fossils fuel/ biomass, vehicular exhaust, power plants and industrial emissions, vehicles with two-stroke engines	Measurement and analysis of fine particulate matter ($PM_{2.5}$) in urban areas of Pakistan was conducted and the values were found to be exceeding the Pakistan National Environmental Quality Standards (Rasheed et al., 2015)
Nitrogen Dioxide (NO_2)	Vehicle exhaust, fuel burning,	Measurement of NO_2 was done at selected public schools of Lahore and the results concluded that values were significantly higher than the standard values, therefore school children were at risk of developing health complications (Mehmood et al., 2015)
Carbon Monoxide (CO)	Traffic, Power Plants, Industrial emissions	Investigation of Carbon Monoxide at heavy traffic intersections of Karachi (Pakistan) using GIS to evaluate potential risk areas for respiratory and heart diseases was conducted to know the levels of CO and of Carboxyhemoglobin (COHb) and their associated health effects (Durdana et al., 2011)

Sulfur Dioxide (SO_2)	Diesel usage in vehicular and for industrial purposes	Spatio-Temporal analysis of atmospheric Sulfur Dioxide column densities over Pakistan by using SCIAMACHY data was done to present the atmospheric SO_2 column densities indicating that an overall increase of about 70% in SO_2 column densities over Pakistan during the time of 2004-2012 (Khattak et al., 2014)
Ozone (O_3)	Vehicle emissions (Nitrogen oxides) and volatile organic compounds chemically react under sunlight	Long term temporal trends and spatial distribution of total ozone over Pakistan was measured and revealed that linear relationship exists between year-to-year total column ozone and solar activity (L. Rafiq et al., 2017)
Dust	Construction activities, vehicular traffic, marble cutting	Air pollution by roadside dust and automobile exhaust at busy road crossing of Lahore was carried out for the assessment of dust pollution levels prevalent in relation to ever increasing roadside automobile exhaust and reported that levels of roadside dust are quite high, and this is a health hazard to the public (Jafary and Faridi, 2006)

Construction industry has a major role in development of the country and brick kilns is among them because it provides bricks, which is the most important building material for the construction of buildings in Pakistan. The fuel mostly used in the brick kilns is coal with the addition of wood that not only generates variety of pollutants (SO_2, NO_X, CO, CO_2, PM) but also a cause of deforestation (Saeed, 2017). Exhaust from the brick kilns is degrading the environment by adding harmful gases from the emissions and posing great health effects to plants, animals, soil and the people living in the vicinity of the brick kilns (Nafees et al., 2012; Kamal et al., 2014)

2.3. Health Effects:

Air pollution not only effects the human health and quality of life but also it causes degradation of environment (Aziz & Bajwa, 2004; World Bank, 2006). Food and water effected by air pollution can enter to human body through ingestion (Thron, 1996). Children living nearby the densely traffic areas have increased respiratory tract complications (Ciccone et al., 1998). Air pollutants relates to increase in deaths and hospital admissions (Kampa & Castanas, 2008), health impacts can be chronic or acute and several organs of body can be affected but mainly it influenced respiratory and cardiovascular system (Kunzli and Tager, 2005; Cohen et al., 2005)

2.4. Health related air pollution studies in Pakistan:

Presence of variety of air pollution sources and their respective pollutants poses a great health effect. Different studies have been conducted to assess the health impacts due to these pollutants. We have summarized the studies done in Pakistan in reference to air pollutants measurements and their respective impacts.

2.4.1. Impacts on Human health:

A questionnaire-based survey was conducted with the monitoring of ambient air quality and noise level in different areas of Lahore, Pakistan and concluded that residents living nearby dense traffic areas, workers, laborers, shopkeepers, and traffic wardens have prevalence of diseases because of bad air quality and noise levels. In addition, the study found that wardens affected by respiratory diseases were 21%, by skin problems were 8%, 10% affected by cardiovascular and 13% by hearing problem (Hamid et al., 2019)

Dental teaching institutions in 12 cities of Pakistan was surveyed in a study to ask about the human health issues with respect to mercury amalgam waste and the response was to ban on mercury amalgam to protect public health, children, and the environment (Khawaja et al., 2016).

Acute Respiratory Infection (ARI) and lung diseases also linked with the outdoor air pollution and the number of patients increased from 1998-2000 in Punjab **(Table 3)**, which is the largest province in Pakistan, increasing population with an increase in industry and traffic volume is the main reason for the ARI to be at third number after diarrhea and dysentery in children (PECC/SDPI, 2001)

Table 3. Acute respiratory issue in Punjab, Pakistan

Year	**Children less than five years**	**All age groups**
1998	837,693	2,826,000
1999	1,024,004	3,304,000
2000	1,536,514	4,989,000

Another health survey-based study was conducted to access the health effects due to air pollution for the people living nearby brick kilns, heavy traffic and away from these sites. It was concluded that 33% around brick kilns and 31% around heavy traffic area were affected from asthma as compared to people (13%) who were living away from the sites. Emissions from industrial units poses great health effects which is evident from the study done nearby an old cement plant for dwellers of Babri Banda town (Kohat, NWFP) and concluded that people living nearby the areas were highly affected lung, chest and respiratory problems caused by the emissions of dust and gases (GTZ-NWFP, EPA, 2000).

Effect of Particulate matter (PM) was conducted in children to check the variation of blood pressure and found that children studying in school that were situated in high traffic related area was having higher arterial blood pressure as compared to children studying in less polluted area (Sughis et al., 2012)

People who spend most of the time around the area of pollution are mostly affected, such a study was conducted to access the effect of air pollution on traffic wardens, and it was investigated that different problems were associated because of the time spent on the roads and showing the signs of lung problems, chronic ear-nose-throat (ENT) and hyper-tension (Pak EPA, 2005).

Association between black carbon and cardiovascular disease was investigated in largest metropolitan city of Pakistan, Karachi and concluded that there was a link between black carbon exposure and the hospital admissions and further it elaborated the increase is especially in females as compared to males indicating the age range from 41-60 years (Malashock et al., 2018)

Similar study was conducted to access the level of $PM_{2.5}$ and the relation with hospitalization and recorded that there was a significant increase in rates of admissions in hospitals which were related to the elevated level of particulate matter (Khwaja, H., et al.,2012)

2.4.2. Effects on Plants:

Pakistan is basically an agricultural country with mostly population based on the production of different crops, vegetables, and fruits but due to an abrupt increase in population and availability of facilities in cities forced the people to move towards the urban areas, reflecting a great deal with the environmental degradation and creating air pollution problems. Bad air quality effects not only human health but it also poses great threat to the plants. Plant growth reduced when they are exposed to air having pollution level excessive from the recommended guidelines. Study have shown that there was a reduction in growth and leaf senescence, photosynthetic rate and photosynthetic efficiency was affected under unfiltered air as compared to filtered air (Wahid et al., 1995a). When the concentration of ozone increased from the safe limits, it has negative effects in the leaves of potatoes, onions and cotton and the aphid attack increased, a study done in northwest of Pakistan showed (Ahmad et al., 2013). In another study, it was found that there was a reduction in yield from 37% to 42% of rice and 34.8% to 46.7% in wheat because of the poor air quality around the surrounding (Wahid et al., 1995b).

2.5. Land use regression (LUR) overview:

LUR is a statistical method of air pollution modeling. It is commonly used to estimate variations in air pollutant concentrations for population exposure assessment. The technique links spatially heterogeneous air quality measurements with geospatial predictors **(Figure 3)**. LUR models provide a comparatively robust method for spatial prediction, while having a lower sampling effort compared to geo-statistical models, and a lower data requirement than dispersion models (Hoek, G. et al. 2008).

LUR model incorporates air pollutants data, geographic predictors such as land use, population density and road traffic network data around the monitoring points, and after multiple linear regression, it provides spatial annual or seasonal pollution level, at un-monitored locations **(Figure 4)** (Jerrett, M. et al. 2007). The model provides a simple and cost-effective method for air pollution exposure assessment at regional and intra-urban level by substituting expensive dispersion models (Hoek, G. et al. 2008; Ghassoun, Y. et al. 2015).

Recently, LUR model is mostly applied in developed countries in North America and Europe, and the exposure assessment of different air pollutants have been successfully conducted (Novotny, E.V. et al. 2011; Beelen, R. et al. 2013; Cattani, G. et al. 2017; Ross, Z. et al. 2007). Still, developed countries remain the focus of conducting air pollution studies related to exposure assessment and its health effects (Tonne, C. et al. 2017). Therefore, it is necessary to have studies related to exposure from air pollution, which can be helpful for epidemiological evidence in developing country conditions (Ma, J. et al. 2017), having air pollution-related health problems.

Challenges and constraints in the development of LUR models in these situations, includes inadequate availability of GIS data, emissions from air pollution sources not well interrelated, and deficiency of routine monitoring concentration data. Pollutant's concentration data gained from the national monitoring locations are used to characterize the air pollution of the entire city which would lead to evaluation error in public air pollution exposure (Huang, L. et al. 2017). It would be meaningful to perform LUR studies to explore the performance of model, thus an efficient and economical air pollutant concentration model can be obtained.

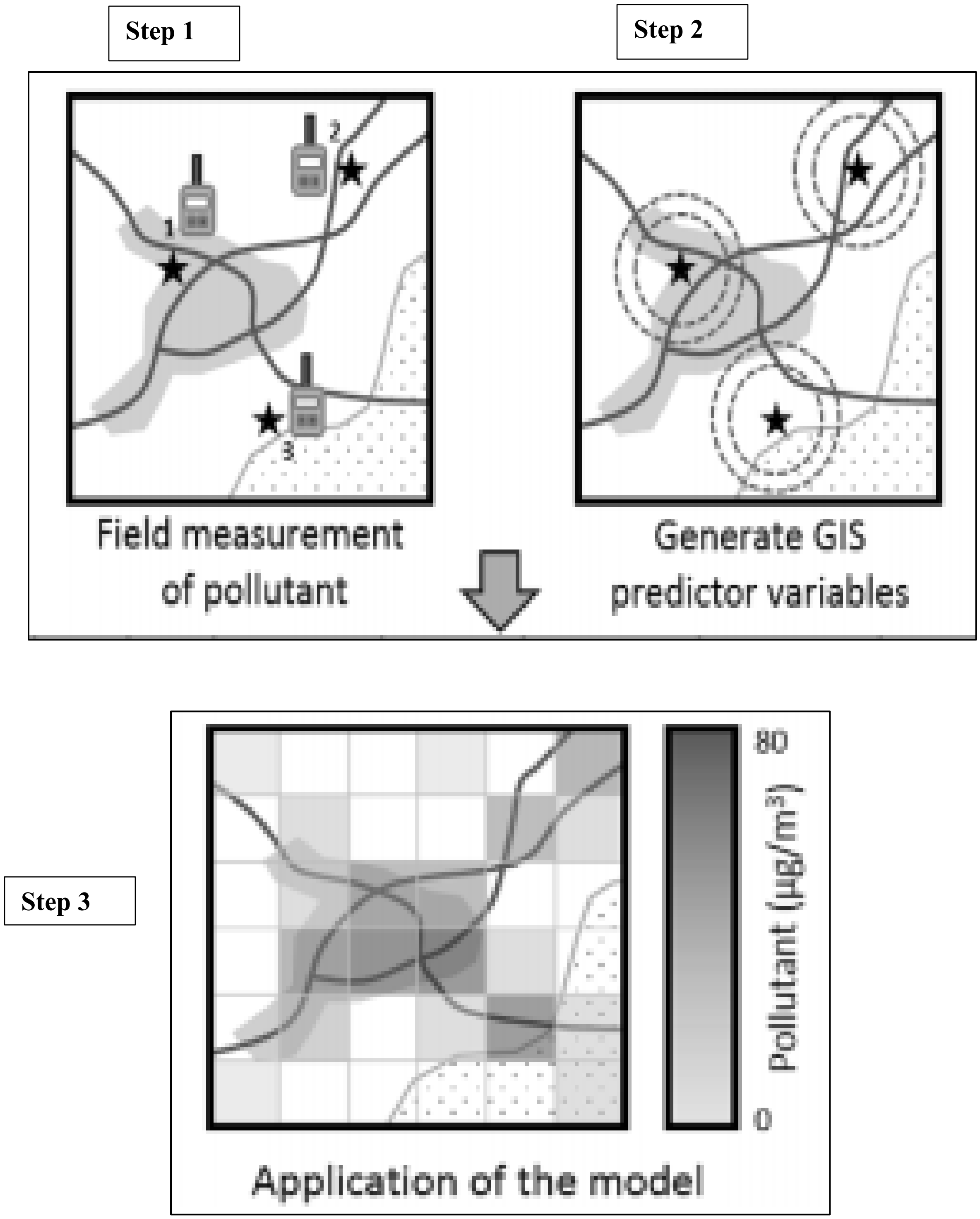


Figure 3. General overview and application of LUR model

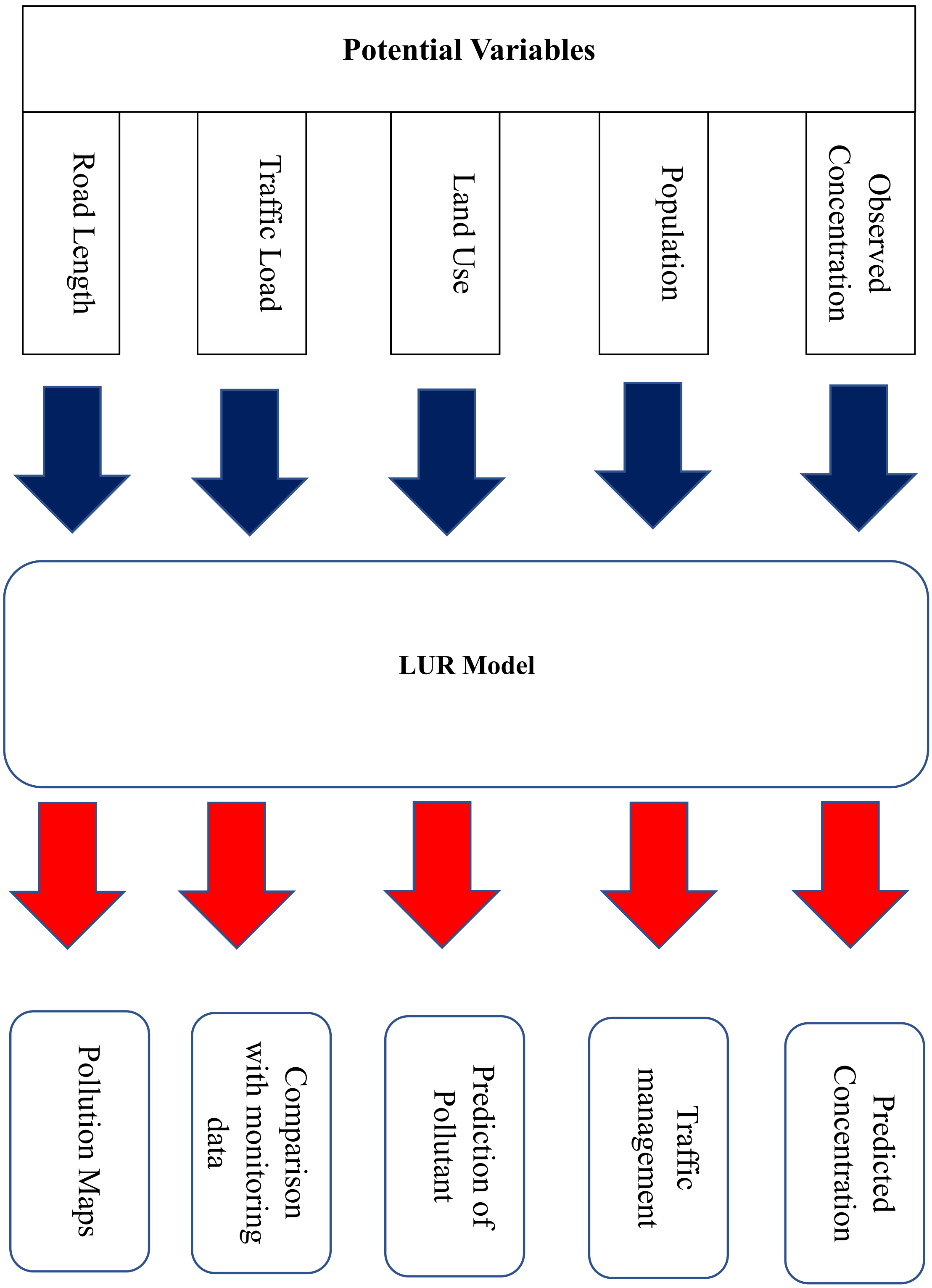


Figure 4. Description of potential variables and LUR model use

2.6. LUR Studies in the world:

Different studies have been conducted related to the development of LUR and its application for the prediction of pollutants and exposure assessment. This section describes the studies that have been done in different countries of the world.

2.6.1. Title of the study: Estimate annual and seasonal PM_1, $PM_{2.5}$ and PM_{10} concentrations using land use regression model (M. Miri, et al. 2019)

In this study, LUR models were developed for seasonal and annual concentrations of Particulate matter pollutants based on 26 monitoring locations in the city of Sabzevar, Iran. Different predictor variables were used including land use, urban morphology, geographic location **(Table 4)**

Table 4. Predictor variables used to develop LUR models

Potential predictor variables to develop LUR models.

Variable	Description	Sub-categorie	Buffer (m)	Direction of effect
Street	Total length of road types (m)	1 = Side streets	100, 200, 300, 500, 1000	+
		2 = Main streets		+
		3 = Major roads		+
		4 = All roads		+
Land Use	Total area of 10 land use (LU) types (m2)	1 = Transportation	100, 200, 300, 400, 500	+
		2 = Commercial/official		+
		3 = Urban facilities and equipment		NA
		4 = Remedial/health care		+
		5 = Education		–
		6 = Green space		–
		7 = Cultural / religious		+
		8 = Industrial		+
		9 = Residential		–
		10 = Others		NA
Urban morphology	Maximum and average height of buildings (m)	1 = Maximum height	100, 200, 300, 400, 500	+
		2 = Average height		+
Distance variables	Distance (DIST)various features (m)	1 = All Traffic Surrogate and Land Use variables		–
		2 = Squares		–
		3 = Bust terminal		–
Population	Density of population (persons per m^2)	All population	500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000	+
Geographic location	Physical location	1 = elevation (m)		NA
		2 = slope (degree)		NA

The developed models performed well within the range of R_2 of 0.68 to 0.75. the minimum mean concentration was seen in spring, and the maximum concentration was observed in autumn. Road length and distance to roads were found in all the models showing the positive association with the concentration of pollutants. Industrial land use also showed the positive association with the concentration of pollutant. The regression maps of PM1 showed that the maximum concentration was observed in the south part of the city around beltway. The maps of $PM_{2.5}$ showed that maximum concentration was observed in the east and south part of the city near the roads and transportation land use. The maps of PM_{10} showed that the maximum concentration was observed in the west of the city and around major roads. The spatial distribution was reasonable, considering traffic as a main source of PM emissions. This study showed that pollutant concentrations were higher than World Health Organization (WHO) guidelines, and cause problem for public. People living near the roads and industrial areas are exposed to pollutants (M. Miri, et al. 2019)

2.6.2. Title of the study: Development of land use regression models for $PM_{2.5}$, SO_2, NO_2 and O_3 in Nanjing, China (L. Huang et al., 2017)

This study explored the applicability of LUR models based on the data obtained from the national monitoring sites in the city of Nanjing, China. Models for four different important pollutants were developed and they performed well related to the explanation of pollutants concentrations in the city showing that it is feasible to have the better simulation of the air quality data for areas having different characteristics. The location of study area for LUR model is shown **(Figure 5)**

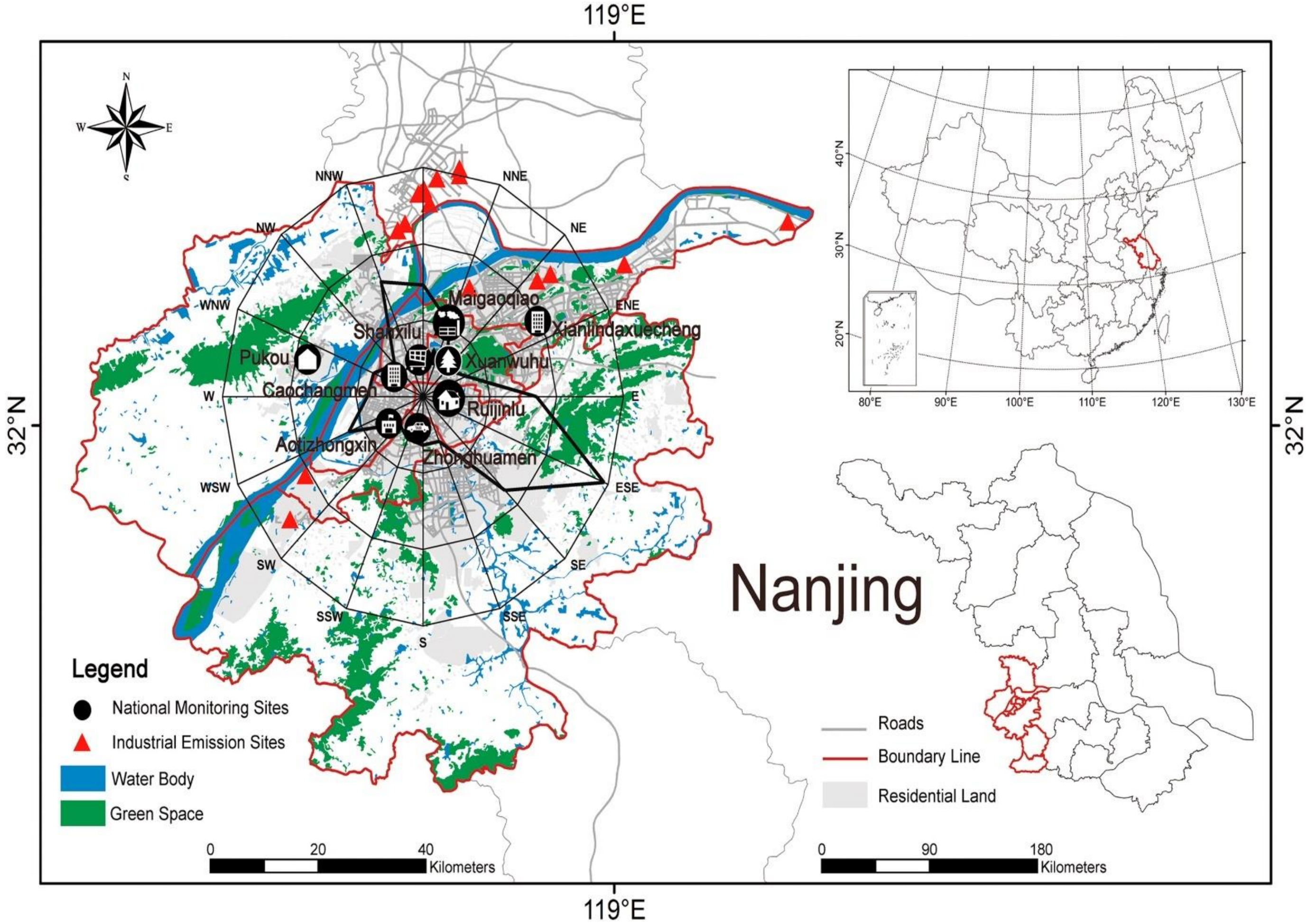


Figure 5. Location of study area

LUR model was developed based on the nine monitoring locations, using the average annual concentration of pollutants including different predictor variables. Developed LUR models explained moderate to good variance for the pollutants. Adjusted explained variance was highest for NO_2 (87%), followed by SO_2 (83%), and was lower for $PM_{2.5}$ (72%) and O_3 (65%). Spatial distribution maps of SO_2 and NO_2 showed that the high concentration was observed in the center of the city due to the usage of fossil fuel in the surrounding urban districts because of the high population density. High road traffic density was also another factor for showing the high concentration of pollutants. For the $PM_{2.5}$ pollutant, road length showed the positive association and considered as an important factor. O_3 concentration was heavily affected by regional influence with the potential predictor variable of longitude.

The developed LUR models in this study were of good quality, indicating that it is possible to develop model based on national monitoring sites. This study showed that with comparison to other models, LUR model can provide better resolution of spatial distribution maps with simple procedure (L. Huang et al., 2017)

2.6.3. Title of the study: Land use regression modelling estimating nitrogen oxides exposure in industrial south Durban, South Africa (S. Muttoo et al., 2018)

This study tried to explore the exposure characteristics in the South Durban area by developing a LUR model and to understand the association between health outcomes and exposure due to air pollution. The area of this study was South Durban region of the eThekwini Municipality of KwaZulu-Natal, South Africa, having a population of 595,601. Potential predictor variables that were identified including major road length within the buffer radius of 300m, minor road length within buffer radius of 1000m showing the positive association with the pollutant, whereas area of open space within the radius of 1000m showed the negative association with the concentration of pollutant.

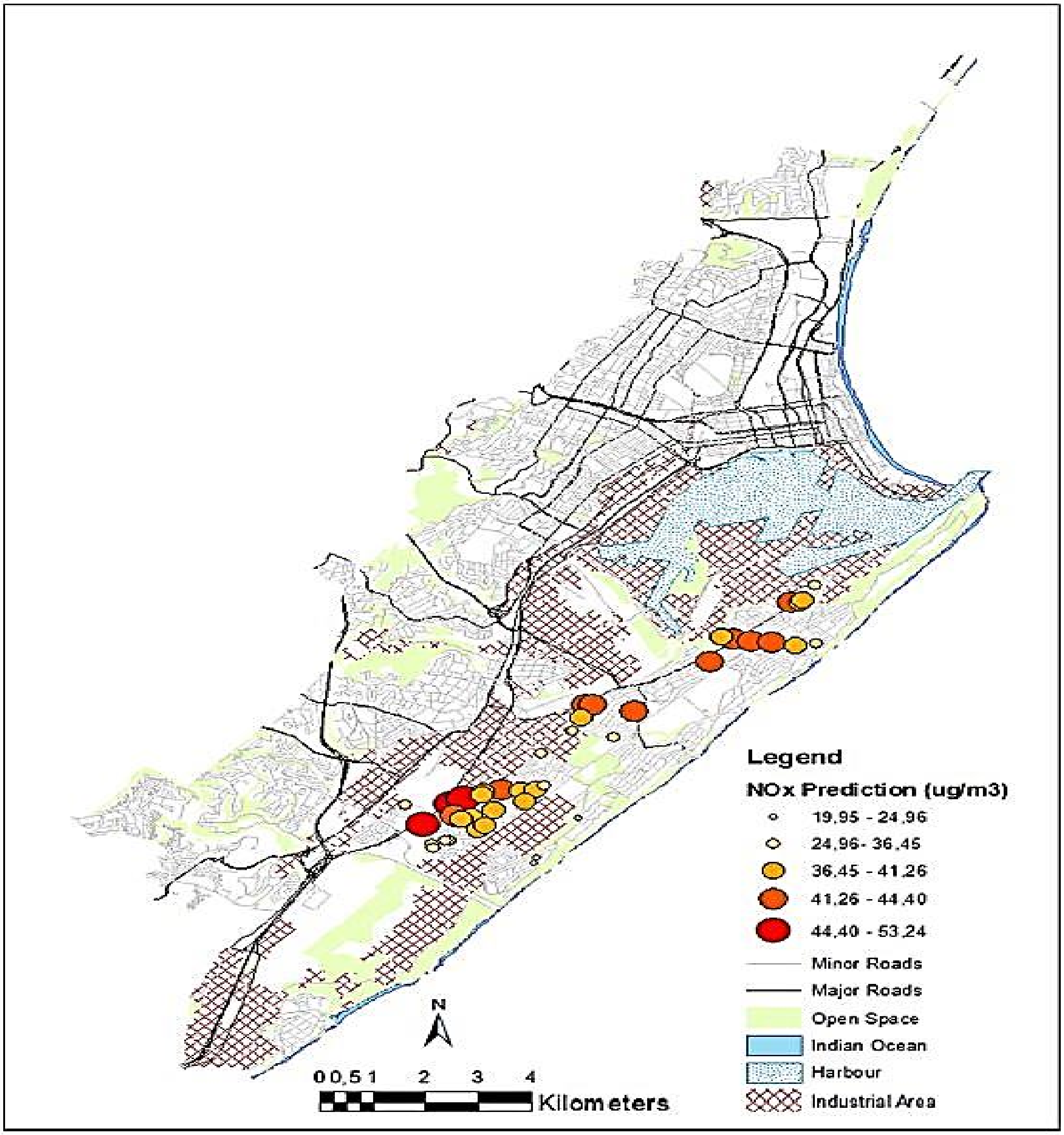


Figure 6. NOx estimation at participant address points in South Durban

The exposure at participant addresses is shown in the (**Figure 6)** showing the higher concentration near the road traffic (S. Muttoo et al., 2018)

2.6.4. Title of the study: Development of land use regression model and health risk assessment for NO_2 in different functional areas: A case study of Xi'an, China (Z. Liu, et al., 2019)

In this study, Xi'an was the study area having a large population (6.4 million people). For the development of LUR model, forty potential predictive variables including land use, width and type of road, pollution sources inclusion etc. with the annual average concentration of Nitrogen Dioxide (NO_2) monitored at 10 locations were used. Models were developed for three seasons: Annual,

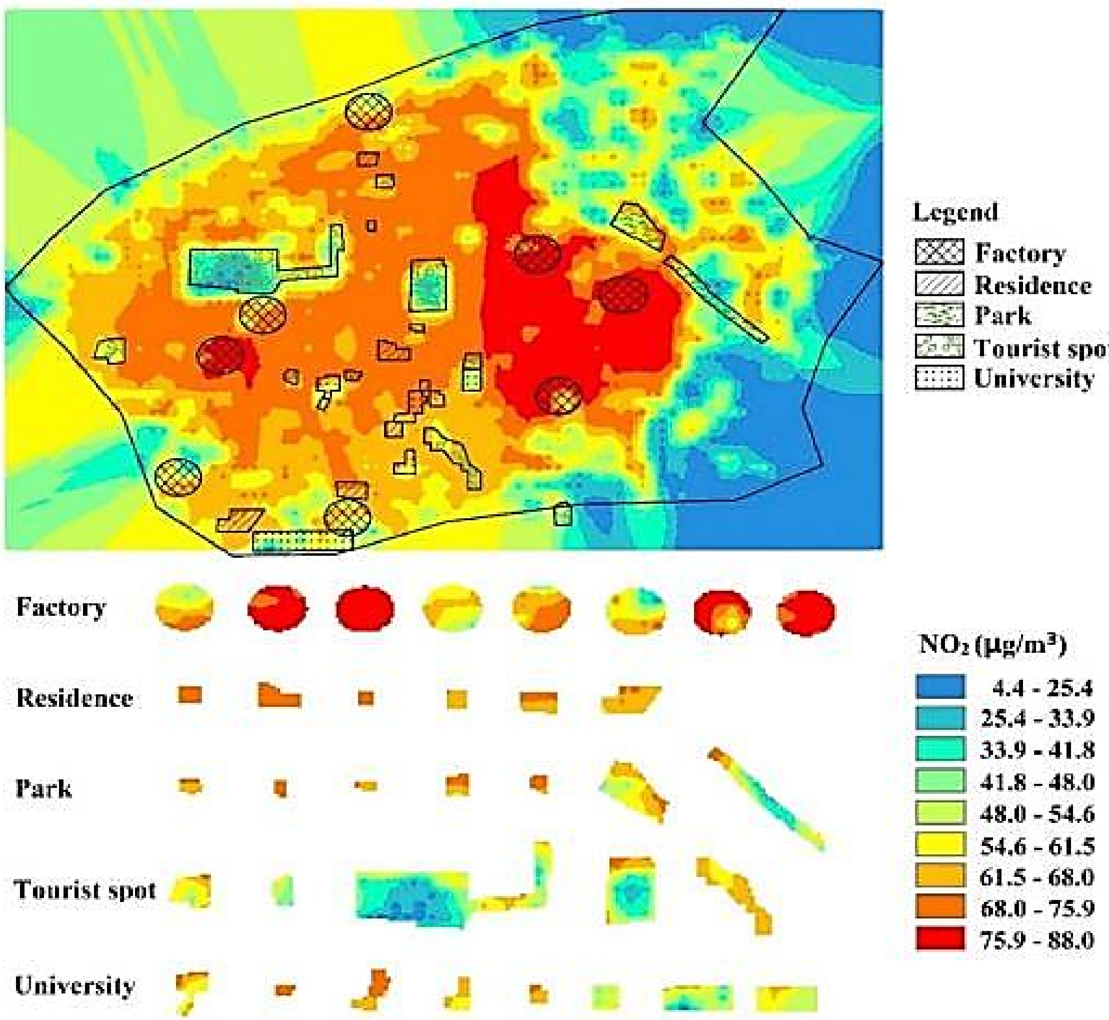


Figure 7. Spatial distribution of predicted NO_2 concentration in Xi'an

non-heating period (June-October) and the heating period (November-April). The spatial distribution of predicted NO_2 showed that higher concentration was observed around the dense factory area, and lower in the tourist spots **(Figure 7)**. Emissions from industries considered as one of the major sources of NO_2 pollution. Residential area found to be the second important source after industrial emissions (Z. Liu, et al., 2019)

2.6.5. Title of the study: Land use regression models for ultrafine particles, fine particles, and black carbon in Southern California (R.R. Jones et al., 2020)

The aim of this study was to provide spatial resolution of ultrafine particles (UFP) for exposure

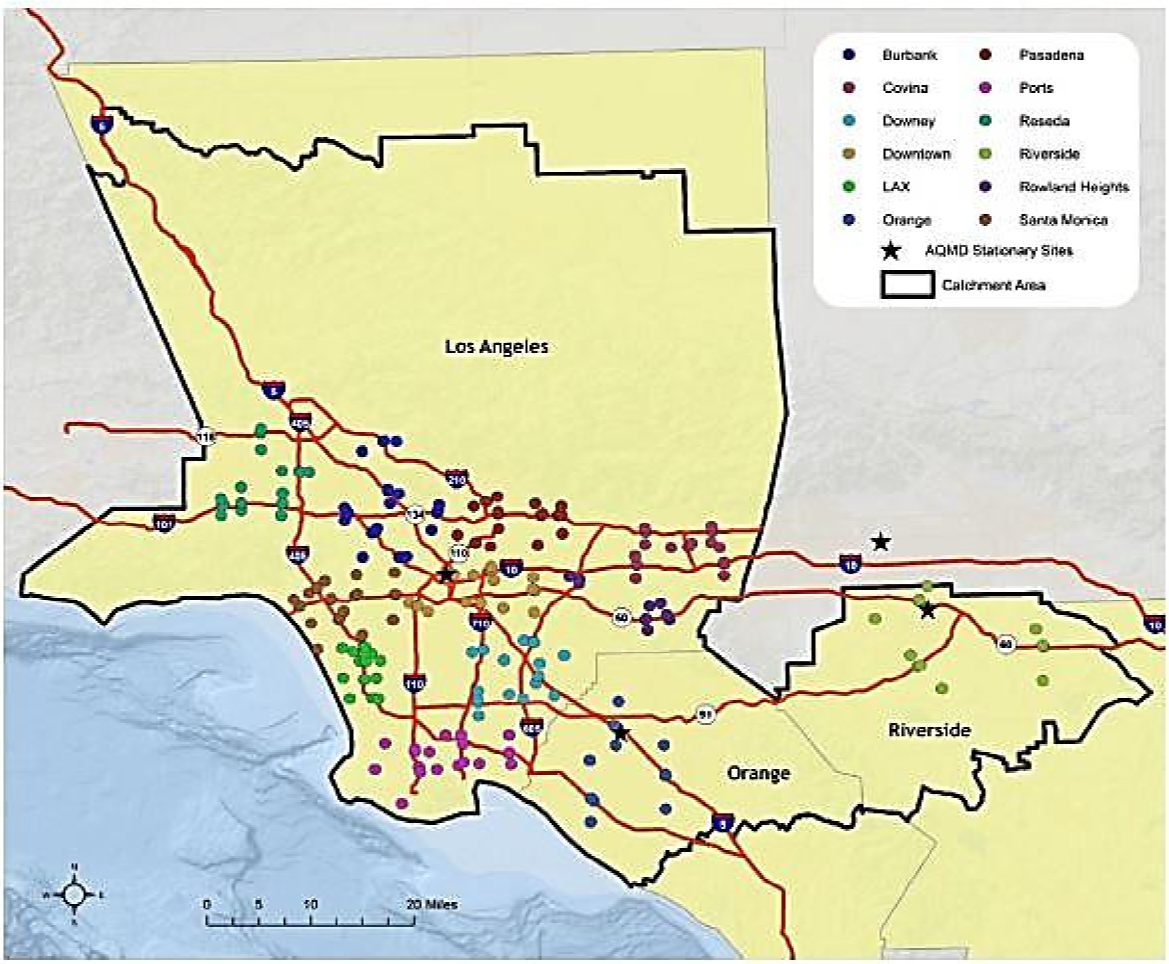


Figure 8. Map of 215 monitoring sites within 12 sampling clusters and four South Coast Air Quality Management District (AQMD) stationary monitoring sites

estimates by developing a LUR model for Los Angeles. LUR models were developed for UFP, Particulate matter ($PM_{2.5}$), and Black carbon (BC) in California. Characterization of the spatial variability for a well-known area across U.S. having high traffic volume and ambient air pollution was done **(Figure 8)**

The model developed based on short-term monitoring performed well as well as based on large number of monitoring locations. The results showed that UFP model can be used and applied for epidemiological studies in the long-period exposure, whereas $PM_{2.5}$ and BC models accomplished relatively less as compared to UFP model (R.R. Jones et al., 2020)

2.6.6. Title of the study: Effect of monitoring network design on land use regression models for estimating residential NO_2 concentration (H. Wu et al., 2017)

The objective of this study was to build and evaluate LUR models using different monitoring designs by having the concentration at the respective home addresses to check the prediction efficiency of LUR models. This study was done in the city of Edinburgh, east of Scotland, UK.

The locations of the monitoring sites and monthly model concentration is shown in figure. Increase in the number of monitoring sites gave better estimation in the resultant LUR model. Developing a LUR model based on higher number of monitoring locations improved the ability to estimate the pollutant concentrations **(Figure 9)**. Beyond a certain number of monitoring locations, efficiency was not significant. LUR models that were built on keeping in mind of both road network and residential area, provides better and more characterization of the pollutant concentration around the specific study (H. Wu et al., 2017)

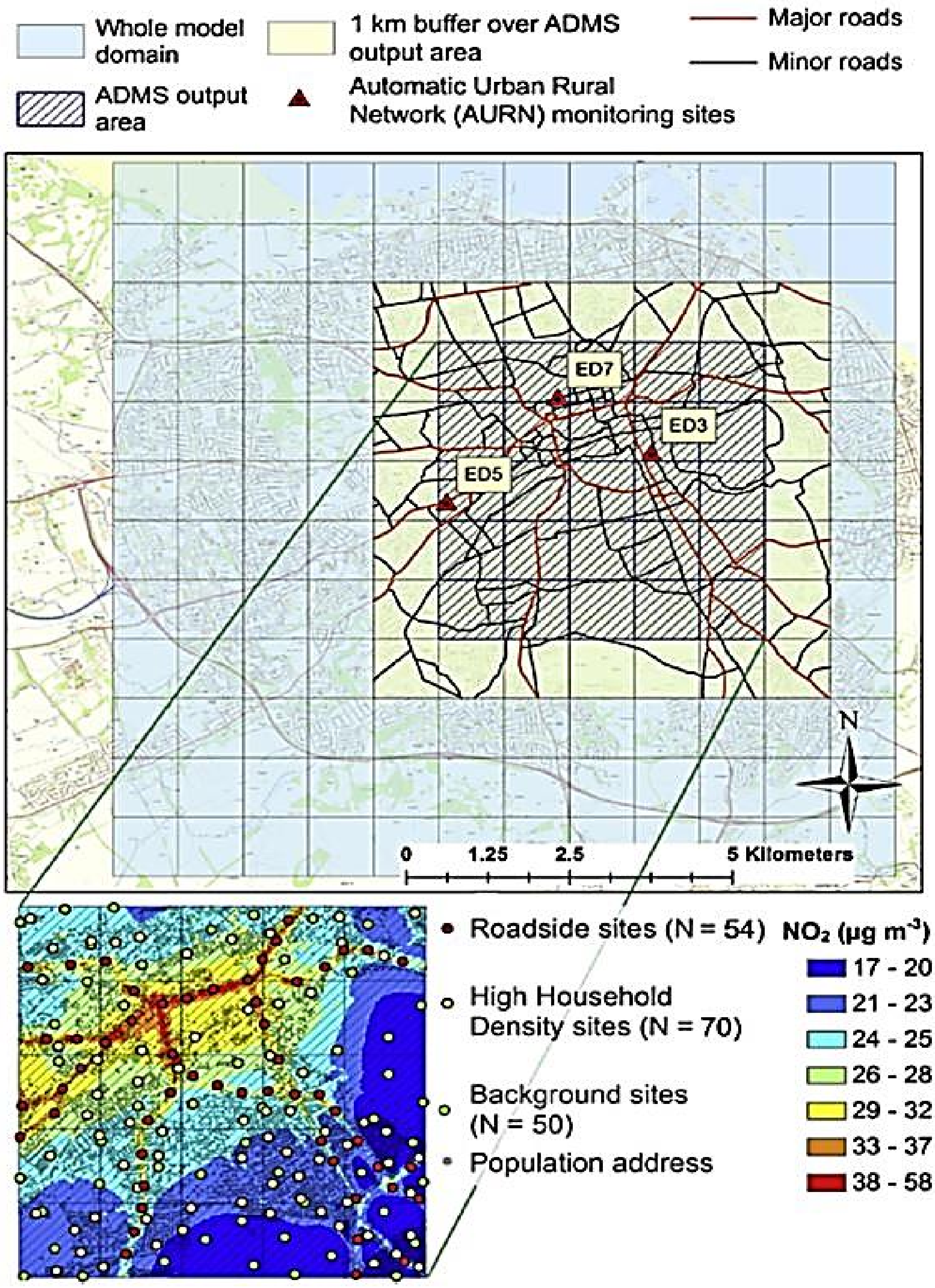


Figure 9. The modelling domain for the city of Edinburgh

2.6.7. Title of the study: Effect of sample number and location on accuracy of land use regression model in NO_2 prediction (J. Dong et al., 2021)

This study was performed to investigate the effect of number and monitoring locations in the development of LUR model and to improve the accuracy.

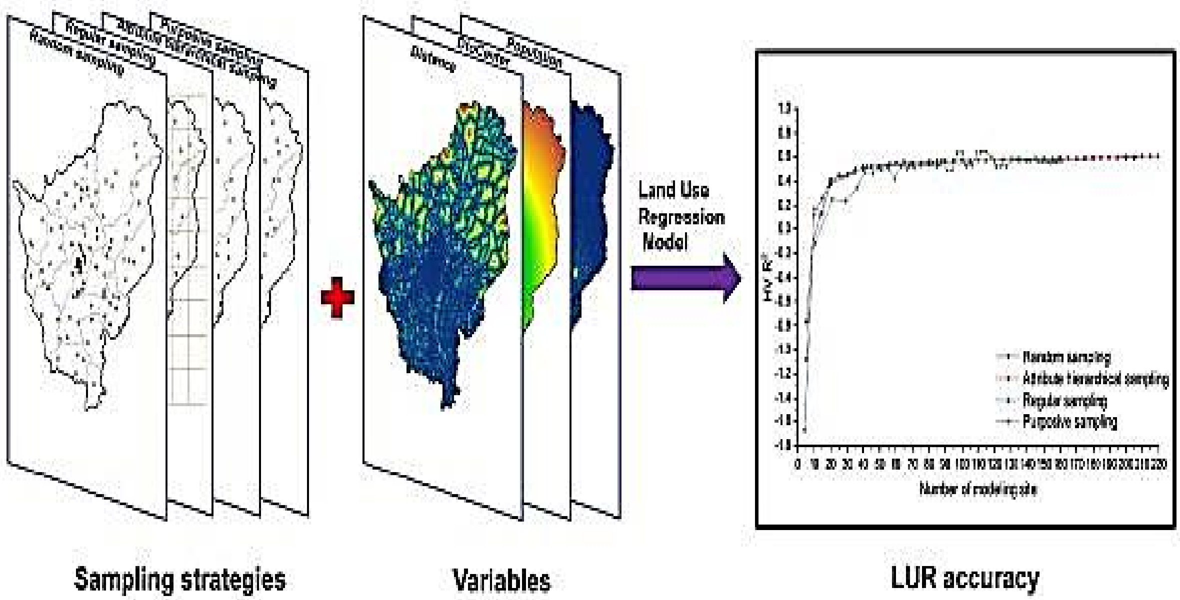


Figure 10. Sample strategies, variables and LUR accuracy

Different sampling strategies were proposed to investigate exposure assessment and air pollution control **(Figure 10)**. Four strategies for sampling were used to check the effects of locations and number of monitoring sites on the performance of model. with the increase in number of sites, R^2 and LOOCV R^2 was shown to be decreased. Different studies have also found the same trend. There were some limitations in this study, worth to be highlighted. This study just conducted in the chosen area, which needs further confirmation for this sampling strategies elsewhere. This study gives the reference to study the effects of monitoring sites and locations for the performance and accuracy for developing LUR model, which can be used for public health exposure assessment and to control air pollution. In future, more studies should be performed in different areas to explore the specific number of sites for designing LUR model and for different pollutants e.g., O_3 and $PM_{2.5}$ (J. Dong et al., 2021)

CHAPTER 3

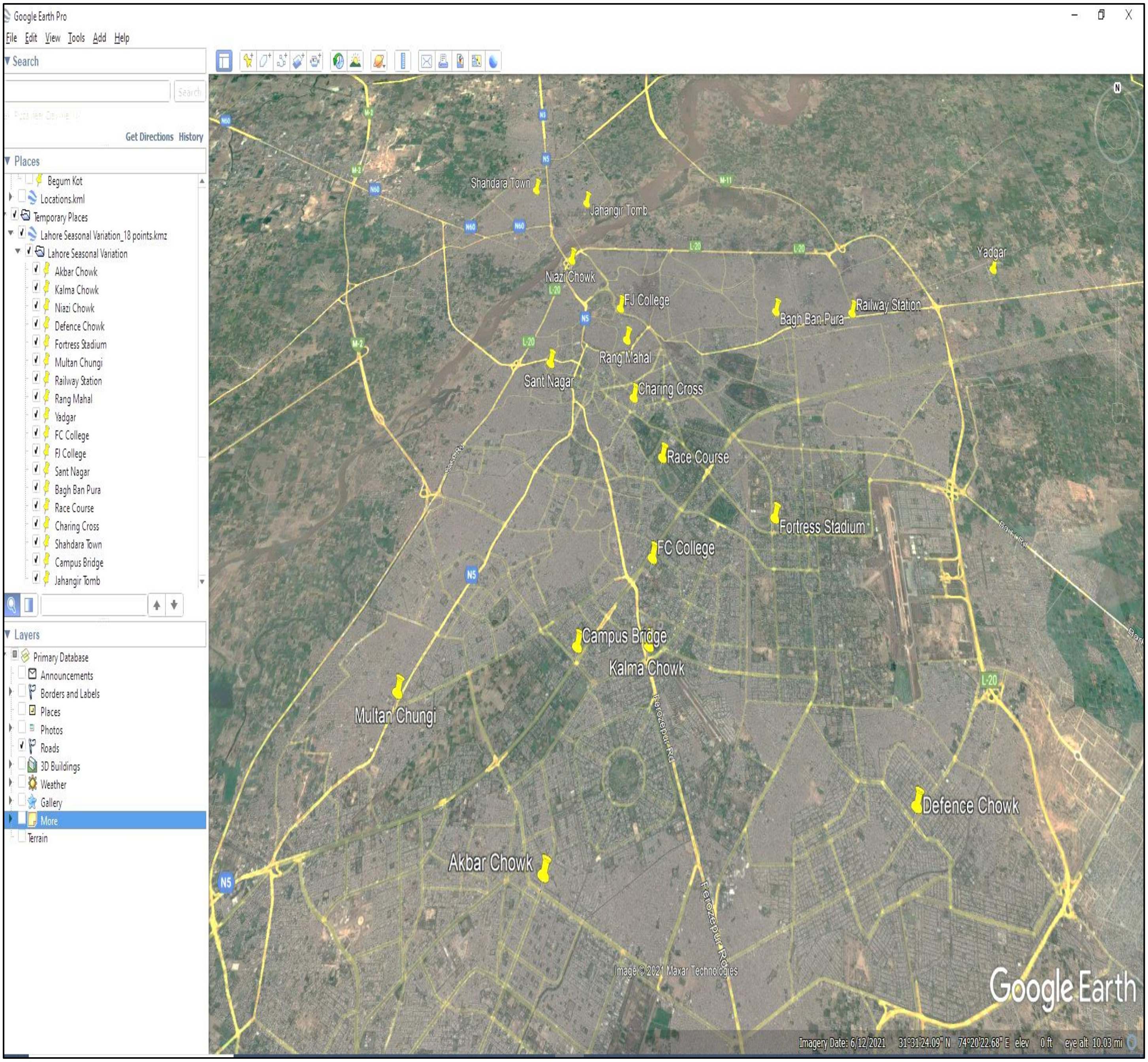
3. Materials and Methods

3.1. Study area

Lahore is the second largest city in Pakistan after Karachi and known as capital of Punjab province. It has an area of 1772 $Km^{2,}$ and population is 11.13 million based on 2017 census (Rana, I.A. et al. 2018). Lahore is in 74° 19' 45.75" longitude, 31° 34' 55.36" latitude, at an elevation of 217 m above sea level with a semi-arid climate and yearly precipitation of 628.8 mm. The average humidity is 39.8% and the main wind direction is north. The lowest and highest temperature in the city is between 19.8-40.4 °C with a yearly average of 30.16 °C (Shabbir, Y. et al. 2016). The country's oldest and longest road known as Grand Trunk (GT) road passes through the city, considered a source of pollution and causing to have higher level of air pollutants throughout the year. The study area and monitoring locations are shown **(Figure 11 A, B, C, D)**

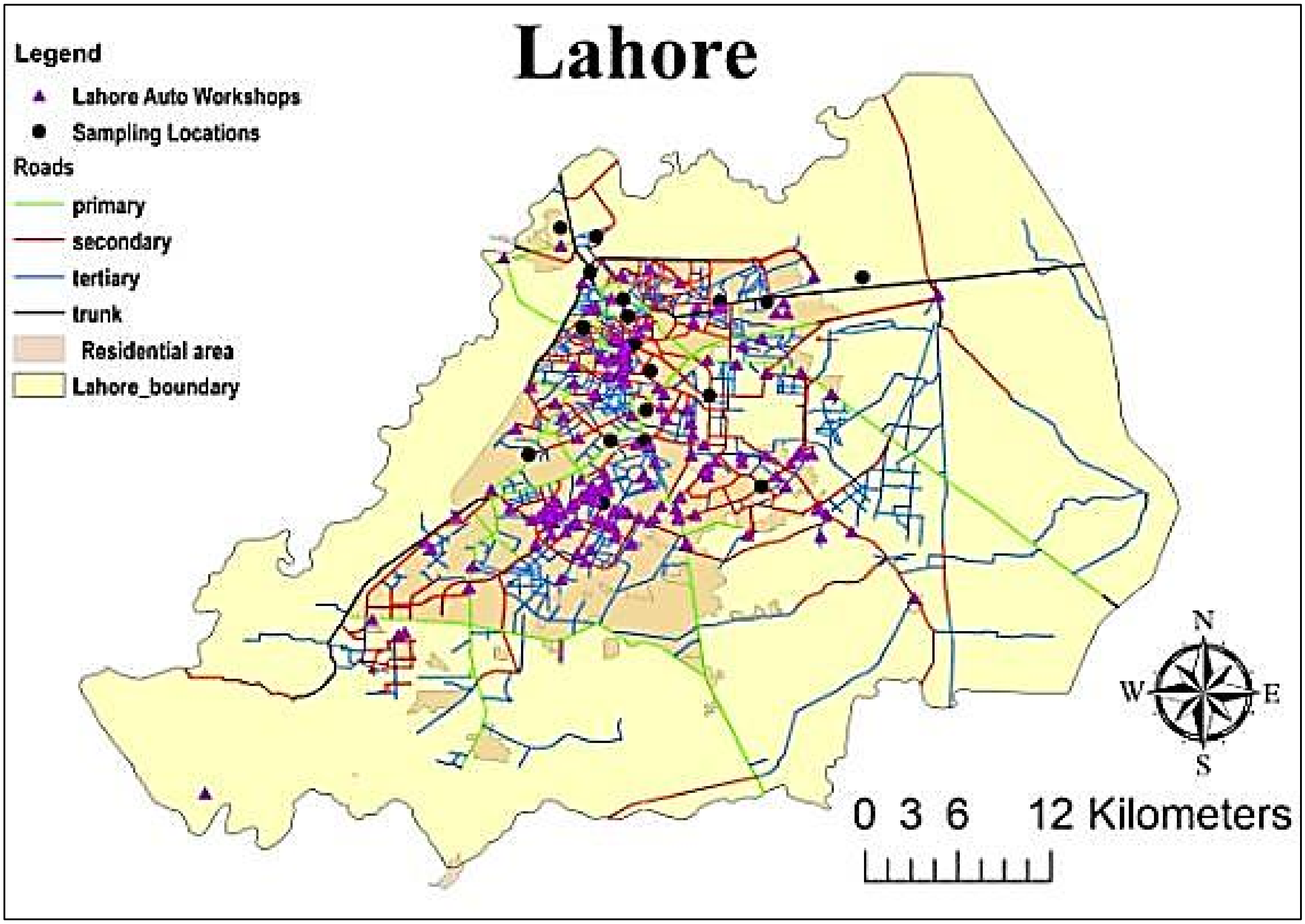
Fig. A showing the monitoring locations on the google earth, Fig. B showing the Lahore boundary, Fig. C showing the monitoring locations on the ArcGIS map with the addition of road network, residential area, and presence of auto workshops whereas Fig. D showing the map of Punjab (Pakistan) and Lahore.

A

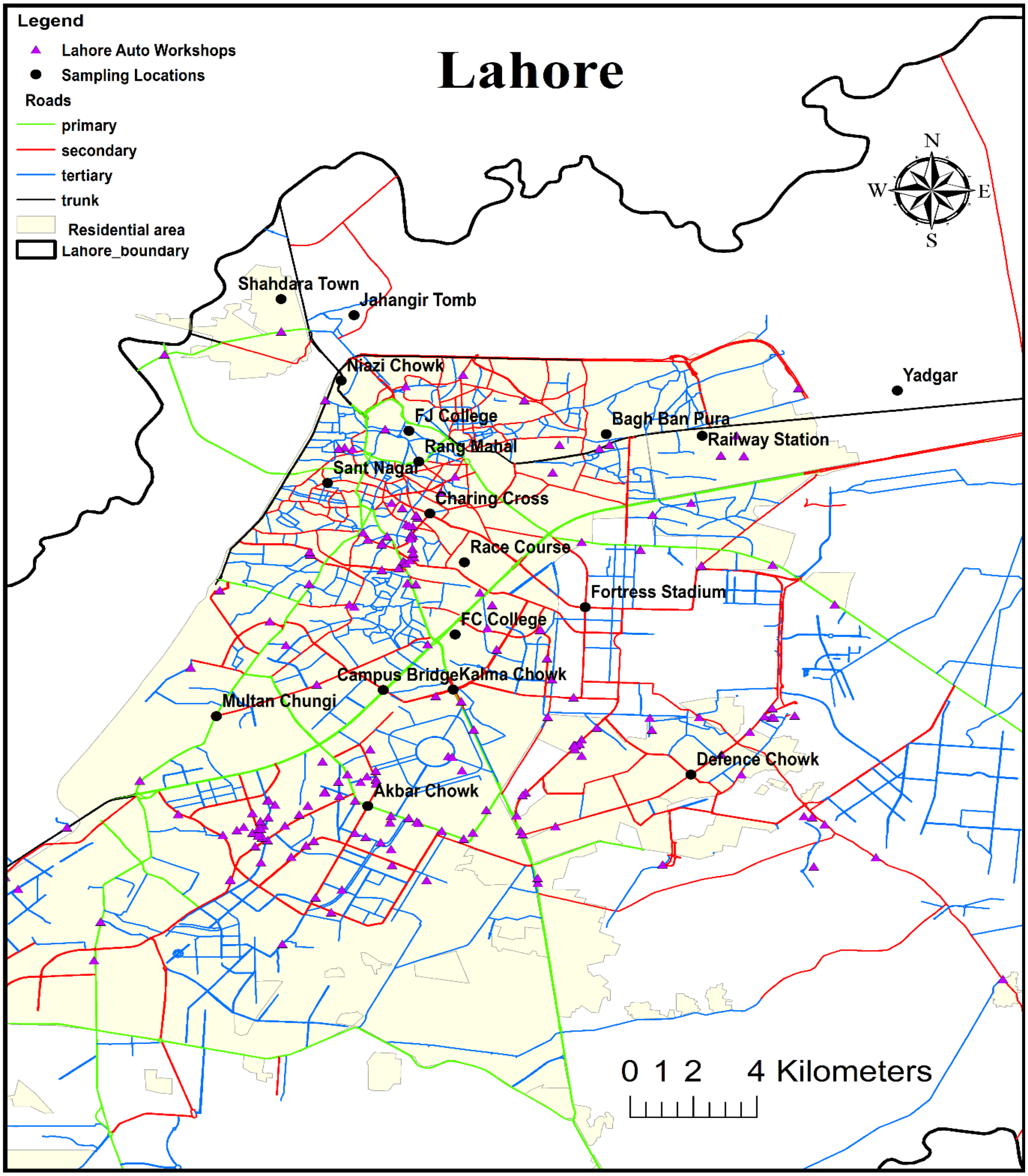


Google Earth Pro
File Edit View Tools Add Help
Search
Get Directions History
Places
Begum Kot
Locations.kml
Temporary Places
Lahore Seasonal Variation_18 points.kmz
Lahore Seasonal Variation
Akbar Chowk
Kalma Chowk
Niazi Chowk
Defence Chowk
Fortress Stadium
Multan Chungi
Railway Station
Rang Mahal
Yadgar
FC College
FJ College
Sant Nagar
Bagh Ban Pura
Race Course
Charing Cross
Shahdara Town
Campus Bridge
Jahangir Tomb
Layers
Primary Database
Announcements
Borders and Labels
Places
Photos
Roads
3D Buildings
Weather
Gallery
More
Terrain
Shahdara Town
Jahangir Tomb
Yadgar
Niazi Chowk
FJ College
Bagh Ban Pura
Railway Station
Rang Mahal
Sant Nagar
Charing Cross
Race Course
Fortress Stadium
FC College
Campus Bridge
Kalma Chowk
Multan Chungi
Defence Chowk
Akbar Chowk
Ferozepur Rd
Image © 2021 Maxar Technologies
Google Earth
Imagery Date: 6/12/2021 31°31'24.09" N 74°20'22.68" E elev 0 ft eye alt 10.03 mi

B


Lahore
Legend
Lahore Auto Workshops
Sampling Locations
Roads
primary
secondary
tertiary
trunk
Residential area
Lahore_boundary
N
W
E
S
0 3 6 12 Kilometers

C


Legend
Lahore Auto Workshops
Sampling Locations
Roads
primary
secondary
tertiary
trunk
Residential area
Lahore_boundary
Lahore
N
S
E
W
Shahdara Town
Jahangir Tomb
Niazi Chowk
FJ College
Rang Mahal
Sant Nagar
Bagh Ban Pura
Railway Station
Yadgar
Charing Cross
Race Course
Fortress Stadium
FC College
Campus Bridge
Kalma Chowk
Multan Chungi
Defence Chowk
Akbar Chowk
0 1 2 4 Kilometers

D

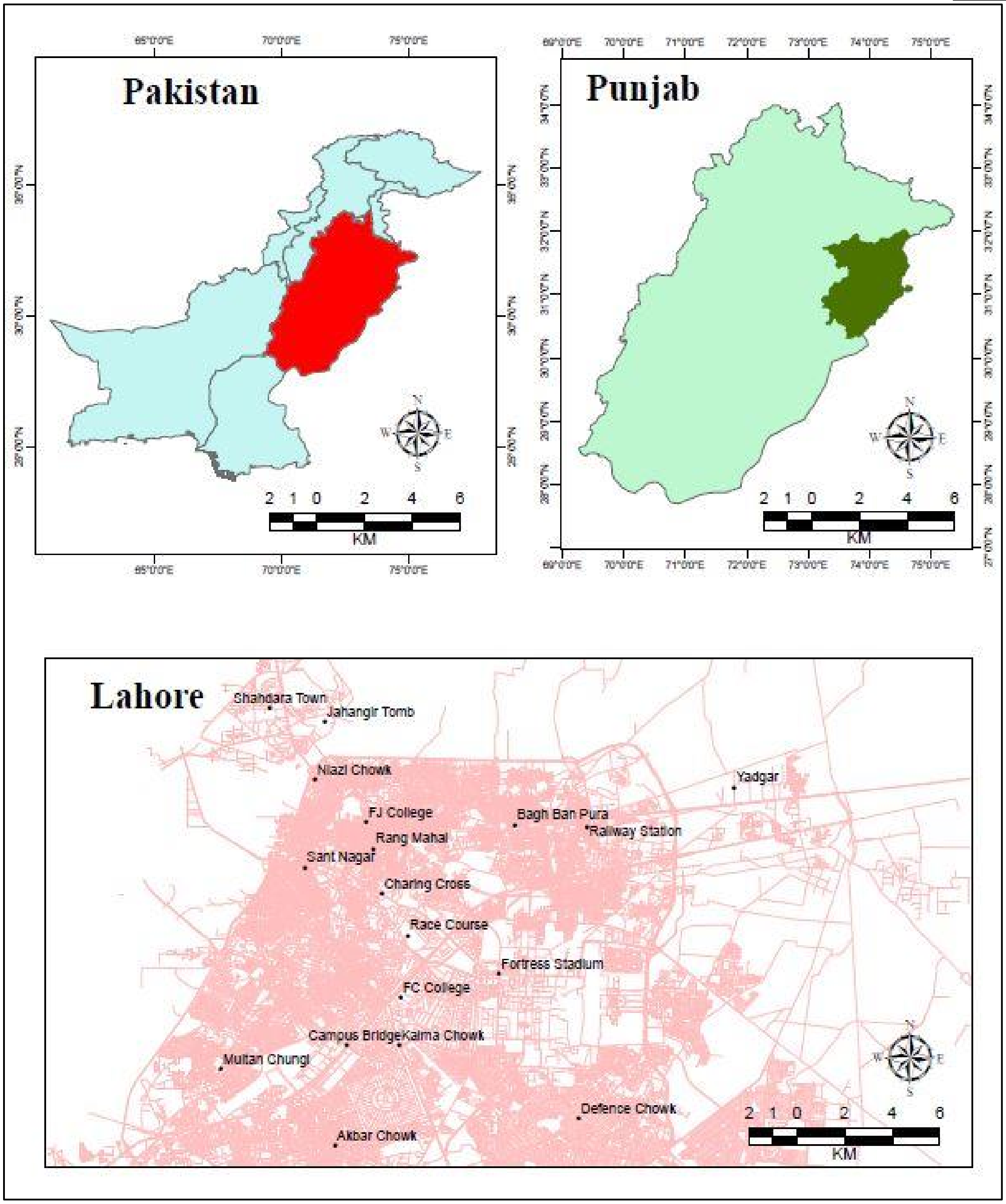


Figure 11 (A, B, C, D). Monitoring locations and study area

3.2. Air pollution data

The seasonal average concentration data of Nitrogen dioxide (NO_2) were obtained from a previous study (Mirza, I.A. et al. 2013). 15 field surveys were conducted in the metropolitan areas of Lahore covering 49 monitoring locations to represent the situation for NO_2 on the periodic basis and the samples were collected, analyzed, and the results were managed. The concentration of NO_2 was estimated using the diffusion tubes only once, and the sampling time was bi-weekly, recording the start and end time of the exposure. The sampling campaign was conducted during the year of 2006 from April to November, dividing the sampling work into three phases (pre-monsoon, monsoon, and post-monsoon). From the available data, 18 different monitoring stations and their respective concentration data were gathered. The mean, minimum and maximum concentration of pollutant against each periodic campaign was noted and shown **(Table 5 & 6)**

Table 5. Seasonal measured concentrations of NO_2

Pollutant	**Time period**	**Sites**	**Measured concentration (ppb)** [a]
NO_2	Pre-monsoon	18	68.34 (7.29) [47.74; 77.48]
	Monsoon		73.91 (7.18) [53.66; 83.70]
	Post-monsoon		79.01 (7.46) [60.19; 91.75]

Table 6. Monitoring Locations and respective data

Location Name	NO_2 Values (ppb)		
	Pre-monsoon	Monsoon	Post-monsoon
AKBAR CHOWK	67.91	73.88	77.70
KALMA CHOWK	74.09	79.87	86.09
NIAZI CHOWK	76.34	81.87	87.62
DEFENCE CHOWK	68.17	73.90	79.82
FORTRESS STADIUM	58.26	64.12	70.12
MULTAN CHUNGI	71.31	77.16	82.85
RAILWAY STATION	71.57	78.14	84.00
RANG MAHAL	77.48	83.70	91.75
YADGAR	73.89	80.09	87.34
FC COLLEGE	68.39	73.80	79.00
FJ COLLEGE	63.48	68.81	73.28
SANT NAGAR	70.38	76.01	78.23
BAGH BAN PURA	70.4	74.90	77.04
RACE-COURSE	59.27	65.74	72.55
CHARING CROSS	73.23	75.01	75.96
SHAHDARA TOWN	67.09	72.80	75.69
CAMPUS BRIDGE	71.23	77.09	83.09
JAHANGIR TOMB	47.74	53.66	60.19

3.3. Potential predictor variables

For the development of LUR model, selection and availability of predictor variables data is of critical importance because they represent the various features of the city i.e., representation of traffic, land use and geography etc.

In this study, 22 potentially predictor variables were selected because of the availability of their data because mostly the variables were having zero values and they were excluded in the selection procedure. Geographic information systems (GIS) shapefiles were obtained from the local department (Urban Unit), and GIS analysis for the extraction of potential predictor variables were performed using ArcGIS ver. 10.2, following the ESCAPE (Beelen, R. et al. 2010) procedure. The buffer radii were selected based on the previous studies ranging from 25, 50, 100, 300, 500 and 1000 m for road network, and ranges from 100, 300, 500 and 1000 m for land use data. As the NO_2 pollutant is related with traffic densities and the concentration can change for the distance of 50m or less, that is the reason why the radii for road network were started from 25m. All the predictor variables values were extracted using these buffer radii **(Table 7)**

The selected predictor variables were grouped into 4 main categories: road length, land use, distance variable, geographic location of monitoring stations, and divided into 18 sub-categories **(Table 7)**. Road length describes the length of different types of roads (primary, secondary, tertiary, and trunk) and the values were calculated in the buffers of 25, 50, 100, 300, 500 and 1000 m, whereas land use represents the 9 types of land use classification (Residential, Commercial, and educational area etc.), and the area of each specific classification were calculated in the buffers of 100, 300, 500 and 1000 m around these monitoring stations. Hospital area mostly represented in the buffer area of 1000m, due to less availability of main hospitals within the other radii distance.

Usage of generators facility for the generation of electricity in the hospitals also caused air pollution concentration to be rise. Further, the distance variable describes the distance of each station to the nearby local specific variable (vehicle maintenance workshop). Finally, the geographic information's include the elevation of monitoring stations in meters above the sea level and longitude and latitude respectively.

Table 7. Brief description of potential predictor variables and 'a priori' definitions

Variable name (unit)	**Variable type**	**Symbol used**	**Buffer size (m)**	**Input rationale**	**A priori[a]**
Road length (m)	Length of:		25, 50, 100, 300, 500, 1000	Road length is directly associated with traffic density, causing pollutant emission from traffic	
	Primary road	PR			+
	Secondary road	SR			+
	Tertiary road	TR			+
	Residential road	RR			+
	Trunk road	TRR			+
Land use (m^2)	Area of:		100, 300, 500, 1000	Different classification of land use generates different variety of pollutants or transportation activities	
	Residential area	LURA			+
	Commercial area	LUCA			+
	Educational area	LUEA			+
	Hospital area	LUHA			+
	Park area	LUPA			-
	Recreational area	LURCA			-
	Religious area	LUREA			+

	Health facility area	LUHFA			+
	Industrial area	LUIA			+
Distance variable (m)	Workshop distance to monitoring locations	DV	25, 50, 100, 300, 500, 1000	Vehicle maintenance workshops emits pollutant emissions	+
Geographic location (m, decimal degrees)	Elevation	Elev.	NA		NA
	Latitude	Lat.	NA		NA
	Longitude	Long.	NA		NA

3.4. Specific local data survey (SLDS)

For the collection of local specific data, such as vehicle maintenance workshops (VMW), a survey was conducted around the monitoring stations. VMW data was considered as a possible indication of pollution source in this study because mostly the workshops present near the road or even near the vicinity of the residential area, which can be a possible source for the pollution exposure to the community living around those areas. It was difficult to get the data of VMW from any department, therefore a survey was conducted to get the knowledge about the presence of VMW around the monitoring locations to be used as a potential predictor variable in this study.

The locations of the VMW were manually geocoded, noted the longitude and latitude with a GPS device, and then imported to google earth. Survey was helpful in determining the locations of the

VMW. There were 198 VMW that were selected in this study because that were present at the time of sampling year **(Figure 12)**. Inverse distance and inverse squared distance were calculated for distance variable (Beelen, R. et al. 2010)


Coca-Cola
ZIAFAT E MAKKAH
Travel & Tours


SUZUKI
AGC AUTOMOTIVE


HOTEL NEW GARDEN
Air Conditioned Rooms
ہوٹل نیو گارڈن
ائیر کنڈیشن روم
Bund Road Opp: Hascol Petrol Pump, Near Abdullah Travel Lahore.
ہوٹل نیو گارڈن


HONDA
12500


HONDA
STYLE COMPUTERS

Figure 12. Workshops in the study area.

3.5. Land use regression model development

The average concentrations of seasonal variations, the predictor variables including GIS parameters, and local specific data for eighteen monitoring stations were used for the development of LUR model. The average concentration was used as dependent variable, whereas all other variables were as independent. Linear regression models were established using forward step regression approach (Beelen, R. et al. 2013). Following the European Study of Cohorts for Air Pollution Effects (ESCAPE) methodology, a priori definition was assigned to each variable type, based on the assumption of air pollution dynamics, e.g., increase in the length of road would increase the traffic emissions (positive direction), while park area would decrease the concentration because of the absorption effect of plants (negative direction). Positive direction of effect was given to the local specific source e.g., VMW, considering as a point pollution source (Saucy, A. et al. 2018).

For the development of model, univariate regression analysis was performed with all potential predictor variables. The factor provided the maximum adjusted explained variance (R^2), was selected keeping in mind that the direction of effect was as pre-defined. The left-over variables were included in the model based on the four criteria: (1) the direction of effect of the variable was as pre-defined; (2) predictor variables already present in the model, having their original direction of effect; (3) the p values of all the variables were below 0.1; (4) the variation inflation factor (VIF) of the predictor variables were considered acceptable if less than 5 (Beelen, R. et al. 2010). To evaluate the influence of each season, models were developed for each season using measured concentrations for each sampling campaign (i.e., pre-monsoon, monsoon, and post-monsoon).

3.6. Model performance evaluation and mapping

To evaluate the performance of the developed model, leave-one-out cross validation (LOOCV) method was used, in which each site was consecutively omitted from the model, and the model was developed using n – 1 site, n being the total number of sites (Saraswat, A. et al. 2013). The process was repeated for each site, and the mean difference obtained from linear regression between measured and predicted values for the left-out sites based on adjusted R^2 and root mean square error (RMSE) was used to access the overall fit. This was determined using JMP software version 13.2.1 (Japan), using the ESCAPE procedure (Beelen, R. et al. 2010).

For the evaluation of spatial autocorrelation, Moran's I was calculated on the residuals of the final model and the results were explained by Z-score values. Grid dimensions of 500 m * 500 m were used, and 8172 grids were created by ArcGIS ver. 10.2 and the spatial distribution of predicted NO_2 concentration were done by using developed models on the relevant grids.

CHAPTER 4

4. Results

4.1. LUR models

Based on the data input i.e., seasonal variation of NO_2 concentration data and 22 predictor variables, three final LUR models for seasonal variation (Pre-monsoon, monsoon, post-monsoon) of NO_2 were developed and their results were presented in **(Table 8,10,12)**. The adjusted R^2 and overall fit of LOOCV of LUR models were described in **(Table 15)**.

4.1.1. Pre-monsoon model

In the final model, three significant factors were identified, including length of tertiary road within 50-m buffer (TR_50), area of residential within 100-m buffer (LURA_100), and distance variable within 300-m (DV_300) buffer. Tertiary road was found to be the primary influencing factor, as NO_2 is traffic-related air pollutant, and the concentration was higher near the roads and residential area due to the presence of roads around the residential area and people move frequently to do their errands. All the three influencing predictor variables were found to be positively associated, showing increase in the concentration of NO_2 **(Table 8)**

VIF were less than 5. The adjusted R^2 and RMSE was 0.70 and 4.35 ppb. The LOOCV R^2 and LOOCV RMSE was 0.60 and 6.11 ppb respectively. Results of residual spatial autocorrelation analysis were presented and showed that the Z-score was 0.70, the pattern appears to be random which shows the consistency with the hypothesis of spatial error independence **(Table 16 & Figure 14A)**.

Table 8. Summary of final land-use regression model predicting NO_2

Time period	Variable	Co-efficient	t	VIF
Pre-monsoon	Intercept	57.71	28.95	-
	Tertiary road length within 50 m	2.1×10^{-1}	2.25	1.12
	Residential area within 100 m	4×10^{-4}	2.50	1.16
	Distance variable within 300 m	1.3×10^{-3}	2.26	1.05

Equation obtained from the model:

$$Y = (2.1 \times 10^{-1} * X_1) + (4 \times 10^{-4} * X_2) + (1.3 \times 10^{-3} * X_3) + 57.7146 \quad \ldots\ldots\ldots (1)$$

X_1 [m] = Length of Tertiary Road within 50m buffer

X_2 [m^2] = Residential Area within 100m buffer

X_3 [m] = Distance variable within 300m buffer

Predicted concentration was calculated based on the equation (1), which was obtained through the final LUR model for pre-monsoon season **(Table 9)**.

Table 9. Measured vs Predicted concentration (Pre-monsoon)

Sr.#	Location Name	Measured NO_2 (ppb)	Predicted NO_2 (ppb)
1	AKBAR CHOWK	67.91	66.54
2	KALMA CHOWK	74.09	69.34
3	NIAZI CHOWK	76.34	71.11
4	DEFENCE CHOWK	68.17	57.71
5	FORTRESS STADIUM	58.26	57.71
6	MULTAN CHUNGI	71.31	68.84
7	RAILWAY STATION	71.57	57.71
8	RANG MAHAL	77.48	78.56
9	YADGAR	73.89	73.49
10	FC COLLEGE	68.39	70.28
11	FJ COLLEGE	63.48	57.71
12	SANT NAGAR	70.38	68.07
13	BAGH BAN PURA	70.4	58.19
14	RACE-COURSE	59.27	57.71
15	CHARING CROSS	73.23	73.65
16	SHAHDARA TOWN	67.09	70.28
17	CAMPUS BRIDGE	71.23	68.11
18	JAHANGIR TOMB	47.74	57.71

4.1.2. Monsoon model

Four influencing factors were entered in the final model, including tertiary road length within 100-m buffer (TR_100m), residential area within 100-m buffer (RA_100m), hospital area within 1000-m buffer (LUHA_1000m), and distance variable within 300-m buffer (DV_300m). Buffer radii of the tertiary road increased from 50m to 100m in this season as compared to pre-monsoon, and the increase in road length signifies the higher concentration in NO_2. Hospital area (LUHA_1000m) entered in the final model, also showed another factor for the source of pollution. Other parameters (RA_100m) and (DV_300m) were same in both seasons.

All the four influencing predictor variables were found to be positively associated, showing in the increase in the concentration of NO_2. VIF were less than 5. **(Table 10)**

The adjusted R^2 and RMSE was 0.71 and 4.09 ppb. The LOOCV R^2 and LOOCV RMSE was 0.50 and 6.19 ppb respectively. Results of residual spatial autocorrelation analysis were presented and showed that the Z-score was 0.11, the pattern appears to be random, showing the consistency with the hypothesis of spatial error independence **(Table 16 & Figure 14B)**.

Table 10. Summary of final land-use regression model predicting NO_2

Time period	Variable	Co-efficient	t	VIF
Monsoon	Intercept	60.07	25.18	-
	Tertiary road length within 100 m	8.08×10^{-2}	1.71	1.31
	Residential area within 100 m	5×10^{-4}	2.94	1.45
	Hospital area within 1000 m	5×10^{-5}	0.90	1.15
	Distance variable within 300 m	1.9×10^{-3}	3.26	1.27

Equation obtained from the model is:

$$Y = (8.08 \times 10^{-2} * X_1) + (5 \times 10^{-4} * X_2) + (5 \times 10^{-5} * X_3) + (1.9 \times 10^{-3} * X_4) + 60.0727 \quad (2)$$

X_1 [m] = Length of Tertiary Road within 100m buffer

X_2 [m] = Residential area within 100m buffer

X_3 [m] = Hospital area within 1000m buffer

X_4 [m] = Distance variable within 300m buffer

Predicted concentration was calculated based on the equation (2), which was obtained through the final LUR model for monsoon season **(Table 11)**.

Table 11. Measured vs predicted concentration (Monsoon)

Sr.#	**Site Name**	**Measured NO_2 (ppb)**	**Predicted NO_2 (ppb)**
1	AKBAR CHOWK	73.88	71.22
2	KALMA CHOWK	79.87	74.74
3	NIAZI CHOWK	81.87	77.00
4	DEFENCE CHOWK	73.90	60.07
5	FORTRESS STADIUM	64.12	60.07
6	MULTAN CHUNGI	77.16	77.36
7	RAILWAY STATION	78.14	72.33
8	RANG MAHAL	83.70	84.30
9	YADGAR	80.09	83.12
10	FC COLLEGE	73.80	75.78
11	FJ COLLEGE	68.81	71.64
12	SANT NAGAR	76.01	75.43
13	BAGH BAN PURA	74.90	60.78
14	RACE-COURSE	65.74	66.59
15	CHARING CROSS	75.01	74.74
16	SHAHDARA TOWN	72.80	75.78
17	CAMPUS BRIDGE	77.09	73.07
18	JAHANGIR TOMB	53.66	60.07

4.1.3. Post-monsoon model

In the final model, four influencing factors were identified, including secondary road length within 1000-m buffer (SR_1000m), length of tertiary road within 300-m buffer (TR_300m), area of residential within 100-m buffer (LURA_100m), and the distance variable within 300-m buffer (DV_300m). As comparison to the other two season model, Tertiary Road buffer radii was high (300m), influencing in the concentration of pollutant. In addition, Secondary Road was another factor in this season found to be different from other two seasons and cause in the concentration of pollutant. Other parameters (RA_100m) and (DV_300m) were same in all the seasons.

All the four influencing factors were found to be positively associated with the NO_2 concentration. VIF were less than 5 **(Table 12)**

The adjusted R^2 and RMSE was 0.77and 3.87 ppb. The LOOCV R^2 and LOOCV RMSE was 0.57 and 6.34 ppb respectively. Results of residual spatial autocorrelation analysis were presented and showed that the Z-score was -0.29, the pattern appears to be random, showing the consistency with the hypothesis of spatial error independence **(Table 16 & Figure 14C)**.

Table 12. Summary of final land-use regression model predicting NO_2

Time period	Variable	Co-efficient	t	VIF
Post-monsoon	Intercept	63.00	20.81	-
	Secondary road length within 1000 m	7×10^{-4}	0.96	1.17
	Tertiary road length within 300 m	8.8×10^{-3}	0.96	1.15
	Residential area within 100 m	5×10^{-4}	2.38	1.28
	Distance variable within 300 m	2.2×10^{-3}	3.26	1.22

Equation obtained from the model is:

$$Y = (7 \times 10^{-4} * X_1) + (8.8 \times 10^{-3} * X_2) + (5 \times 10^{-4} * X_3) + (2.2 \times 10^{-3} * X_4) + 63.0076 \ldots (3)$$

X_1 [m] = Length of Secondary Road within 1000m buffer

X_2 [m] = Length of Tertiary Road within 300m buffer

X_3 [m^2] = Residential area within 100m buffer

X_4 [m] = Distance variable within 300m buffer

Predicted concentration was calculated based on the equation (3), which was obtained through the final LUR model for monsoon season **(Table 13)**.

Table 13. Measured vs predicted concentration (post-monsoon)

Sr.#	**Site Name**	**Measured NO_2 (ppb)**	**Predicted NO_2 (ppb)**
1	AKBAR CHOWK	77.70	78.24
2	KALMA CHOWK	86.09	87.03
3	NIAZI CHOWK	87.62	81.79
4	DEFENCE CHOWK	79.82	68.65
5	FORTRESS STADIUM	70.12	71.30
6	MULTAN CHUNGI	82.85	81.63
7	RAILWAY STATION	84.00	64.01
8	RANG MAHAL	91.75	89.70
9	YADGAR	87.34	87.30
10	FC COLLEGE	79.00	81.23
11	FJ COLLEGE	73.28	74.67
12	SANT NAGAR	78.23	77.90
13	BAGH BAN PURA	77.04	70.12
14	RACE-COURSE	72.55	65.91
15	CHARING CROSS	75.96	77.41
16	SHAHDARA TOWN	75.69	78.71
17	CAMPUS BRIDGE	83.09	80.16
18	JAHANGIR TOMB	60.19	64.39

4.2. Variation of measured NO_2 concentration

The seasonal variation (Pre-monsoon, Monsoon, Post-monsoon) of measured Nitrogen dioxide concentration was shown in **(Table 14)**. The descriptive statistics of concentration which showed that the lowest concentration (47.74 ppb) was observed in summer (Pre-monsoon), and the highest concentration (91.75 ppb) was seen in winter days (post-monsoon). The scatter plots of measured vs predicted values of NO_2 were shown in the **(Figure 13)**

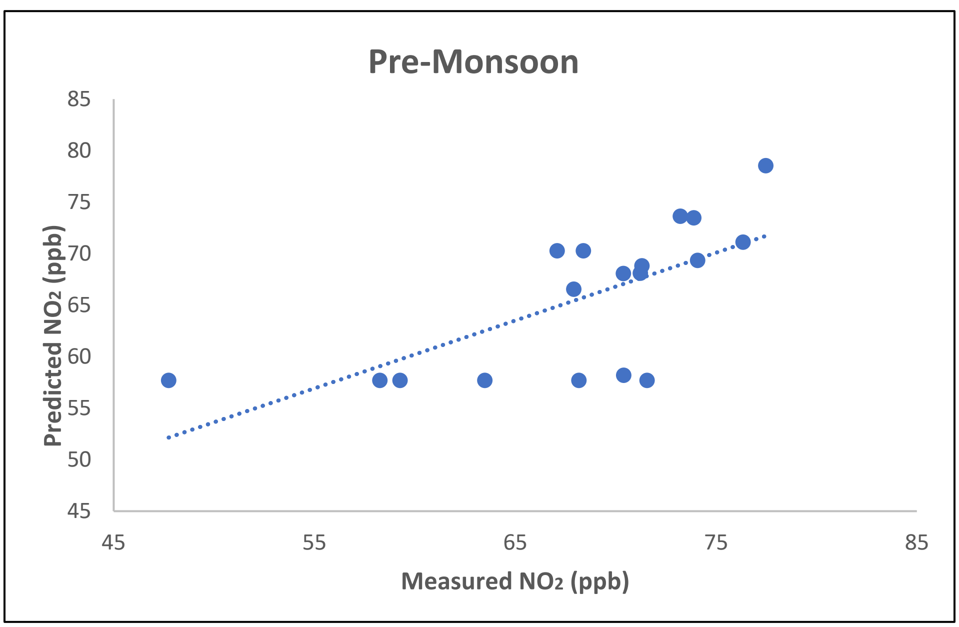
4.3. Variation of predicted NO_2 concentration

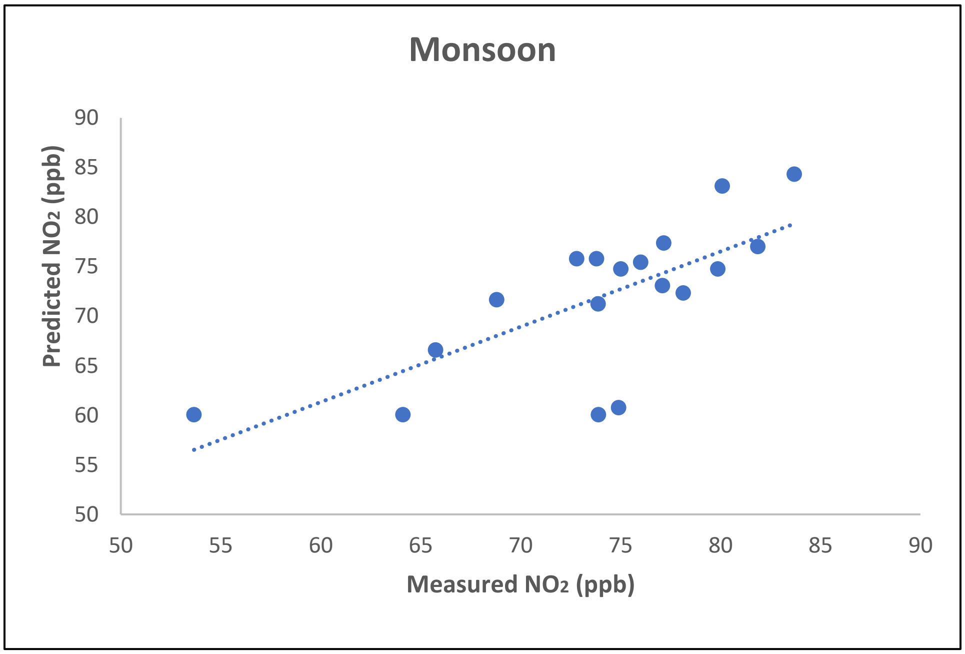
The predicted concentration was calculated for each season from the final model variables. The seasonal variation (Pre-monsoon, Monsoon, Post-monsoon) of predicted Nitrogen dioxide concentration was shown in **(Table 14)**. Similar trend was shown in the measured concentration as comparison to predicted concentration, showed that the lowest concentration (57.71 ppb) was observed in summer (Pre-monsoon), and the highest concentration (89.70 ppb) was seen in winter days (post-monsoon).

Table 14. Measured and predicted NO_2 levels

Time period	Measured NO_2 (ppb)		Predicted NO_2 (ppb)	
	Mean (S.D) [a]	Range [b]	Mean (S.D) [a]	Range [b]
Pre-monsoon	68.34 (7.29)	[47.74; 77.48]	65.71 (7.00)	[57.71; 78.56]
Monsoon	73.91 (7.18)	[53.66; 83.70]	71.89 (7.54)	[60.07; 84.30]
Post-monsoon	79.01 (7.46)	[60.19; 91.75]	76.67 (7.85)	[64.02; 89.70]

Abbreviations: [a] Mean (Standard deviation), [b] [min; max]





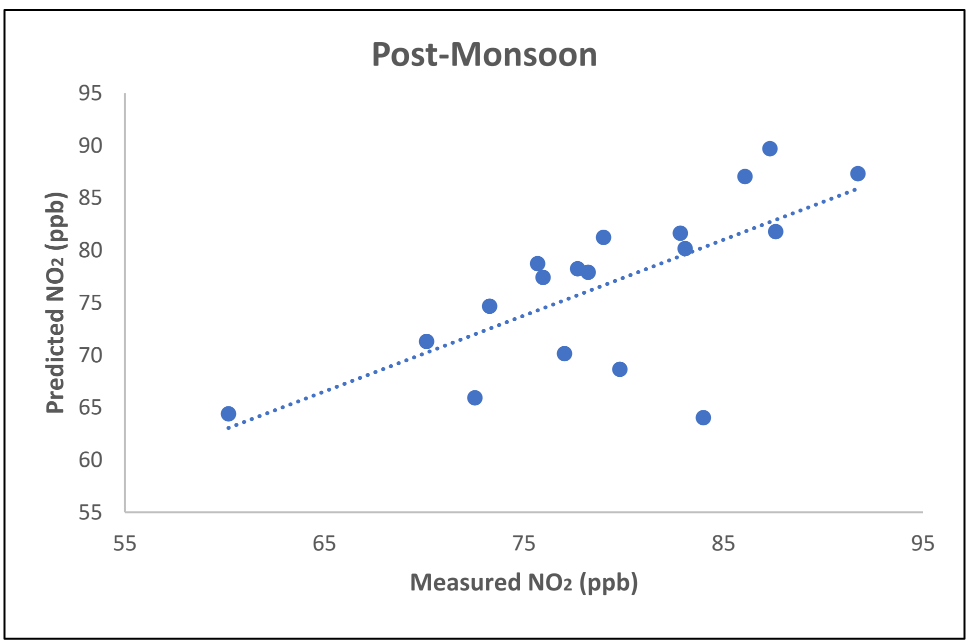


Figure 13. **Scatter plots showing the measured vs predicted values of NO_2 (Pre-monsoon, Monsoon, and Post-monsoon)**

LUR models showed moderate to good variance for all the season. Adjusted explained variance of the LUR models was highest for post-monsoon (77%), followed by monsoon (71%), and was lowest for pre-monsoon (70%). As for the overall fit of LOOCV, the R^2 was highest for pre-monsoon (61%), followed by post-monsoon (57%), and was the lowest for monsoon (50%).

Table 15. Adjusted R^2 of LUR model and overall fit of LOOCV

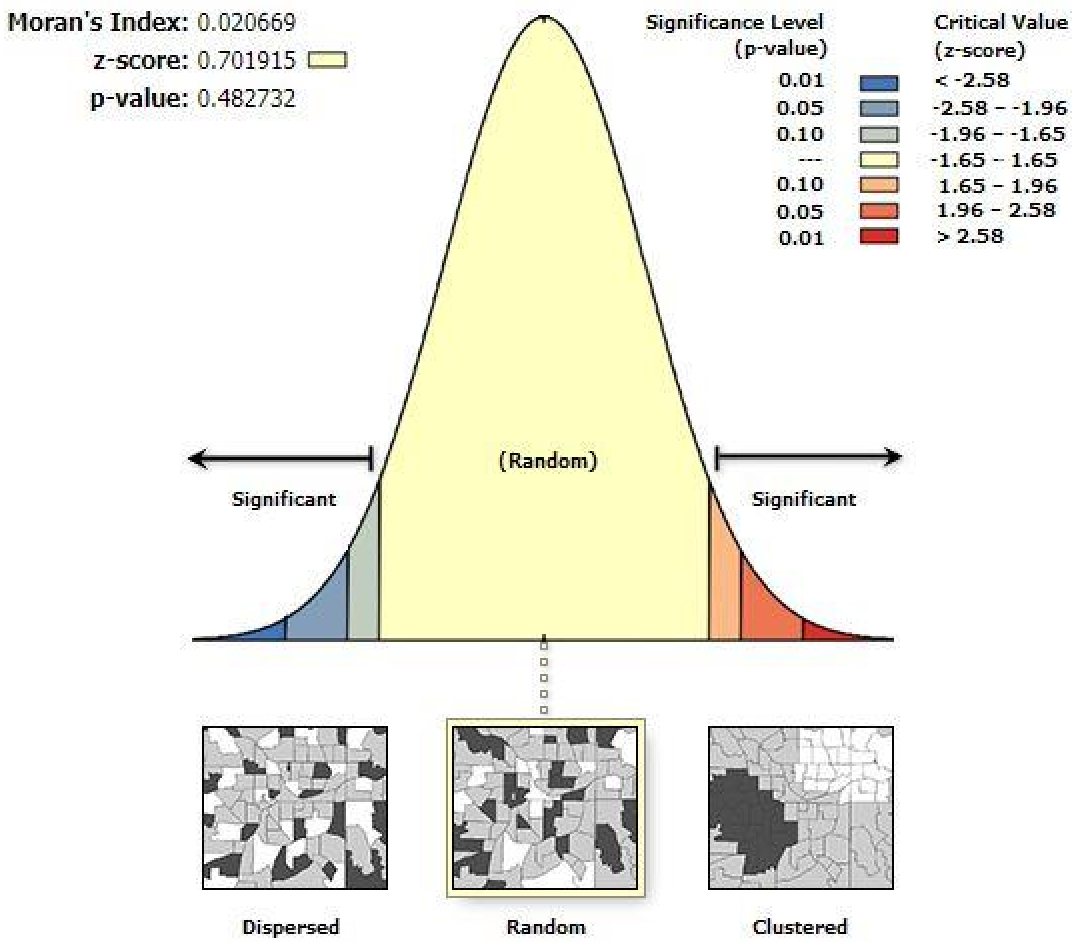
Time period	**Model performance evaluation**		**LOOCV**	
	Adj-R^2	RMSE (ppb)	R^2	RMSE (ppb)
Pre-monsoon	0.7	4.35	0.6	6.11
Monsoon	0.71	4.09	0.5	6.19
Post-monsoon	0.77	3.87	0.57	6.34

Abbreviations: RMSE, root mean square error, LOOCV, leave-one-out-cross-validation

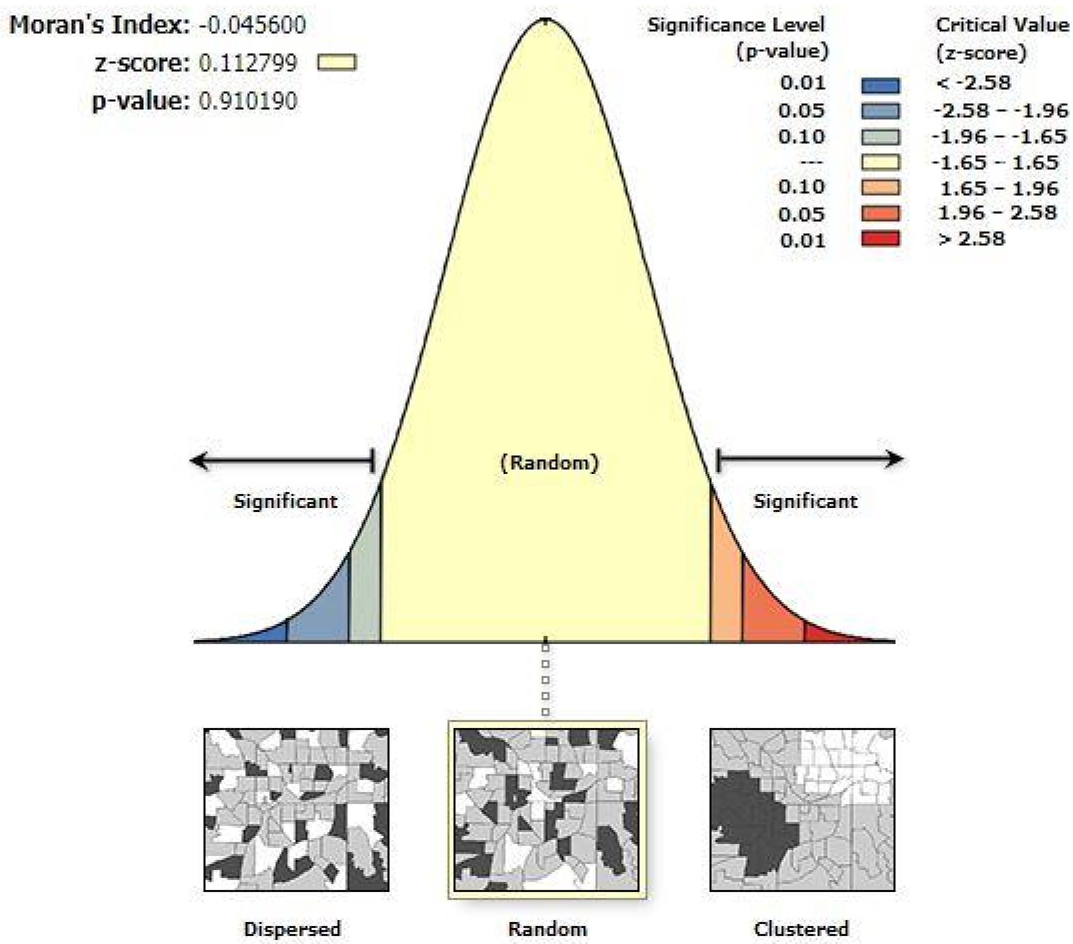
Table 16. Spatial autocorrelation results of LUR model residuals

Time Period	**Moran's Index**	**z-score**	**p-value**
Pre-monsoon	0.02	0.70	0.48
Monsoon	-0.04	0.11	0.91
Post-monsoon	-0.09	-0.29	0.77

A


Moran's Index: 0.020669
z-score: 0.701915
p-value: 0.482732
Significance Level (p-value)
Critical Value (z-score)
0.01 < -2.58
0.05 -2.58 - -1.96
0.10 -1.96 - -1.65
--- -1.65 - 1.65
0.10 1.65 - 1.96
0.05 1.96 - 2.58
0.01 > 2.58
Significant
(Random)
Significant
Dispersed
Random
Clustered

B


Moran's Index: -0.045600
z-score: 0.112799
p-value: 0.910190
Significance Level (p-value)
Critical Value (z-score)
0.01 < -2.58
0.05 -2.58 - -1.96
0.10 -1.96 - -1.65
--- -1.65 - 1.65
0.10 1.65 - 1.96
0.05 1.96 - 2.58
0.01 > 2.58
(Random)
Significant
Significant
Dispersed
Random
Clustered

C

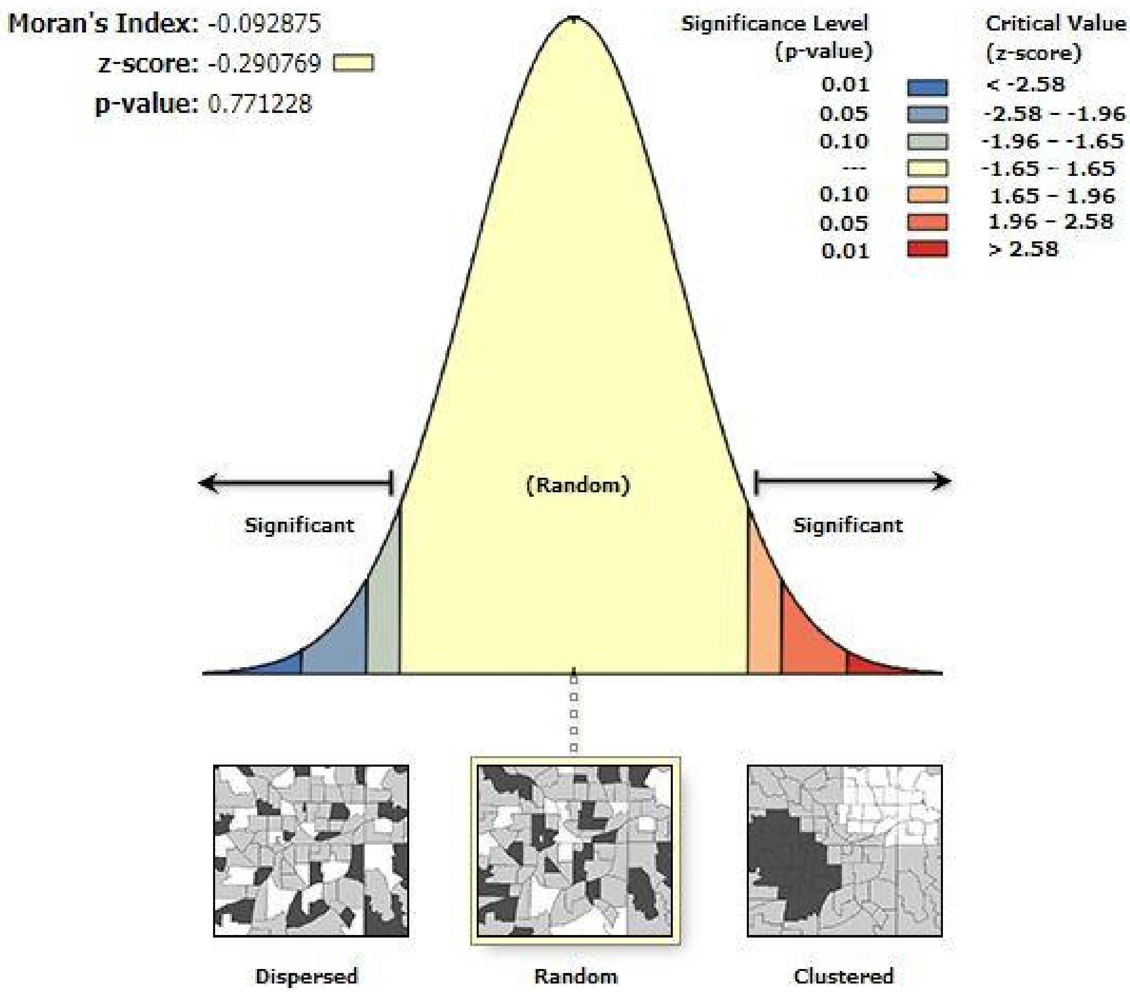


Figure 14. Spatial autocorrelation of the LUR models (A is pre-monsoon, B is monsoon, and C is post-monsoon)

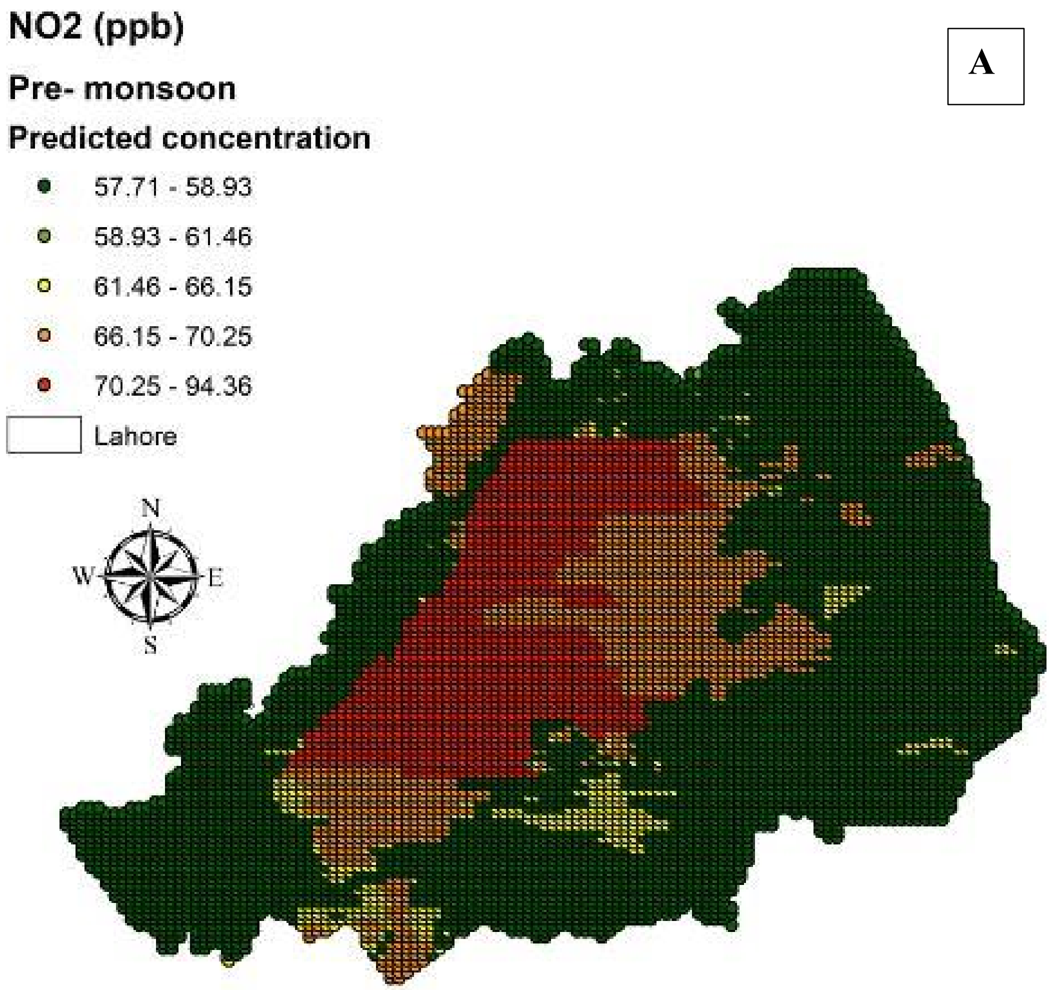
4.4. Spatial characteristics and regression maps of predicted NO_2

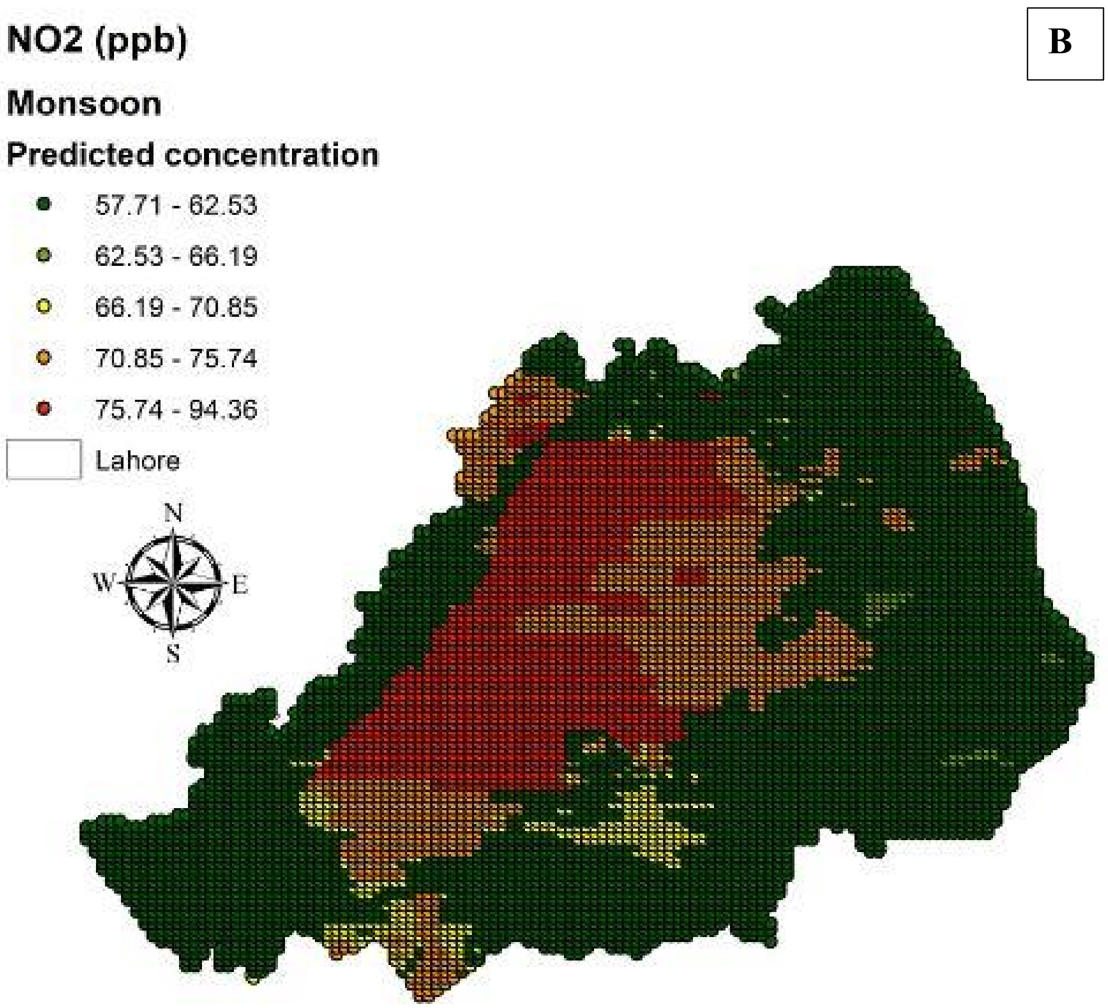
The maps of seasonal variation of predicted NO_2 were done using 8172 uniformly distributed grids for the Lahore city **(Figure 15A, 15B, 15C)**. Higher concentration was observed in the center of the city, where mostly the residential area is located because people prefer to live where they can access everything i.e., shopping malls, Hospitals etc. Roads around the residential area is high, and people move frequently using their own private transportation i.e., motorcycle, car etc., and this contributes to large emissions of NO_2 because of vehicle exhaust. Distance variable also contributed as a factor in emissions, because VMW mostly present near the roads and residential area, representing another reason for the higher concentration **(Figure 15A)**

Regression map for the monsoon season **(Figure 15B)** showed that the concentration of NO_2 was also higher in the center of the city. The reason was the presence of residential area, high roads density and of VMW. In addition, Hospital area also showed another source of pollution in this season. The reason was the use of fuel-based generators to produce electricity, because during monsoon season due to heavy rainfall, electricity is down due to tripping of the feeder which distribute and provide electricity. Therefore, the concentration of NO_2 was higher as compared to pre-monsoon season due to increase in traffic load and of hospital area.

Highest concentration was observed in the post-monsoon season **(Figure 15C)**, because of the higher density of roads, residential area and VMW. In addition, Secondary Road showed another source of increase in the concentration, because of the increase in the number of vehicles and their emissions cause higher concentration of pollution. During the winter season (post-monsoon), people use fuel-based generators for the combustion and heating purposes in their homes. High roads density (Tertiary and secondary roads), heating emissions from the residential area increase the concentration of pollutants, and that was the reason for the highest concentration observed in post-monsoon season.

Highest concentration can be seen around the center of the city almost during all the seasons, indicating road traffic as a source of pollution while, lower concentration can be seen around the sub-urban areas where the road network is not so strong, and having open spaces and agricultural land.


NO2 (ppb)
Pre- monsoon
Predicted concentration
57.71 - 58.93
58.93 - 61.46
61.46 - 66.15
66.15 - 70.25
70.25 - 94.36
Lahore
N
W
E
S
A


NO2 (ppb)
Monsoon
Predicted concentration
57.71 - 62.53
62.53 - 66.19
66.19 - 70.85
70.85 - 75.74
75.74 - 94.36
Lahore
N
W
E
S
B

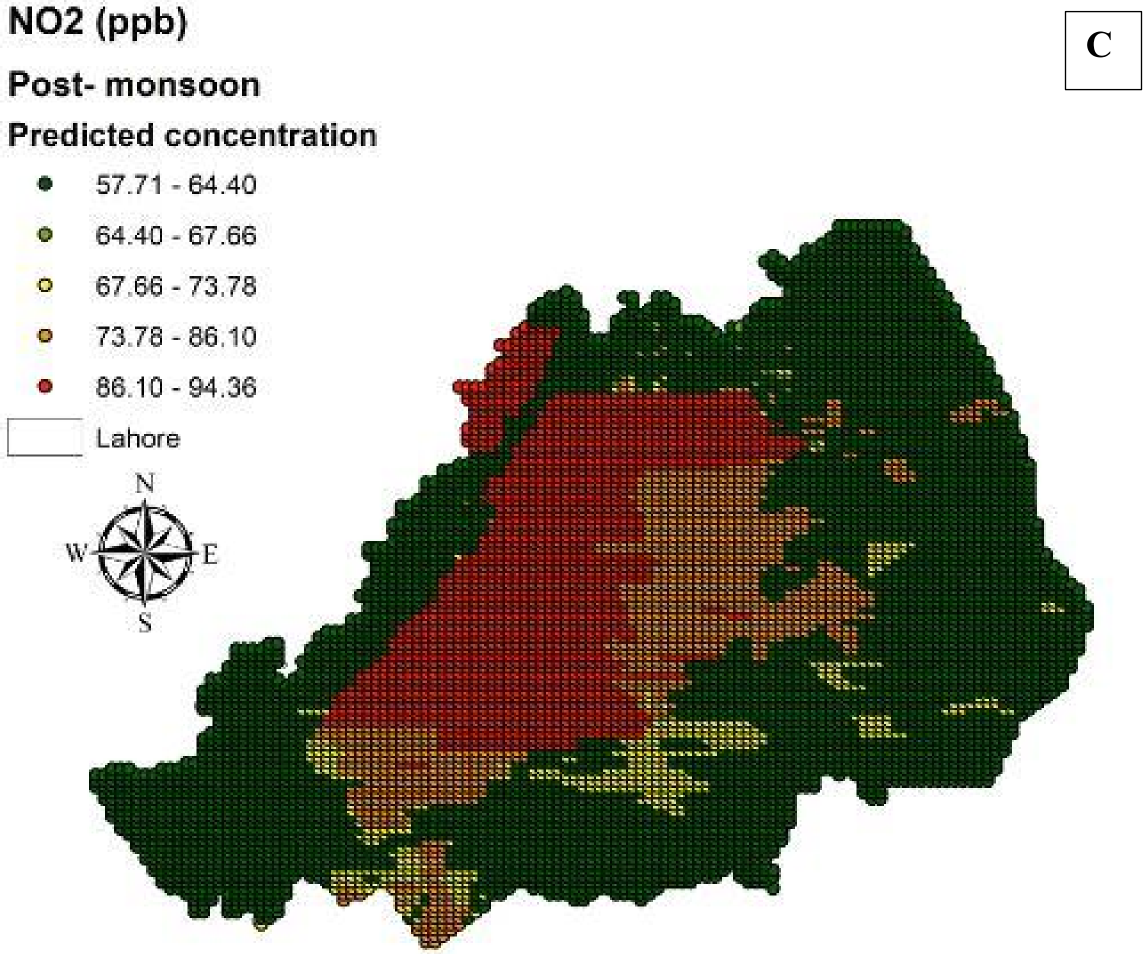


Figure 15. Regression mapping of predicted concentration (A is pre-monsoon, B is monsoon, C is post-monsoon respectively)

CHAPTER 5

5. Discussion

Land use regression model has been applied in the developed countries, but still there is lack of application of LUR model in developing countries (Meng, X. et al. 2015). To the best of our knowledge, this is the first attempt to apply LUR model in Pakistan's urban area setting for a city, although LUR has been used in Pakistan at a national level for ambient $PM_{2.5}$ exposure (Shi, Y. et al. 2020). LUR models were developed for seasonal variation (pre-monsoon, monsoon, and post-monsoon) mean concentration of NO_2 pollutant, based on the collected data of 18 monitoring locations in the Lahore city, Pakistan. The final developed LUR models performed well, showing the reliability with high accuracy and spatial heterogeneity.

Previous LUR models in the literature have showed the values of R^2 ranging from lowest (0.41) in the startup model to highest (0.73) in the final model, achieving an R^2 of 0.68 for winter model and 0.59 for summer model (Muttoo, S. et al. 2018). The study conducted in Xian, China reported that the value of R^2 was greater than 0.8, indicating that the heating season had the best simulation effect (Liu, Z. et al. 2019). The study conducted in Nanjing; China reported the R^2 value of 0.7 for NO_2 model (Huang, L. et al. 2017). The LUR models developed in this study, shows that the values of Adj R^2 (0.7 – 0.77) **(Table 14)**, like other studies conducted in the literature, thus indicating that it is feasible to develop these models for developing countries and can use for exposure assessment studies. The values of the model R^2 was close to the LOOCV R^2, showing the robustness of the LUR models for all the seasons (Meng, X. et al. 2015).

Table 14. Comparison with other studies

This Study			Other Studies		
Reference	**Variables**	**Performance ($Adj\text{-}R^2$)**	**Reference**	**Variables**	**Performance ($Adj\text{-}R^2$)**
S. Pervez et al., 2021	Road Length (Tertiary ad Secondary Road) Residential Area Distance to workshop	0.7 – 0.77	Z. Liu et al., 2019	Residential Area Industrial Emissions	0.8
			L. Huang et al., 2017	Residential Area Road Length	0.7
			A. Saucy et al., 2018	Major Roads Distance to bus stops	0.76

Data collected from manually surveying the study area, proved helpful to increase the performance of LUR models. Such type of survey can provide us the valuable specific feature of potential predictor variables in the study area, that would improve the overall fit of LUR model. In this study, vehicle maintenance workshops data were collected by doing the manual survey of the study area, which can be a source of NO_2, thus highlighting the importance of culture or site-specific land use classes (Wu, C.D. et al. 2017; Smargiassi, A. et al. 2012). Distance to vehicle maintenance workshop entered in the final models of all the seasons, found to be as one of the influencing factors for the source contribution to NO_2.

Different potential predictor variables entered in the final developed models, showing that the different factors have influence on the pollutant concentrations, although same predictor variables were used to develop the LUR models **(Table 15)**. The road network (road length) was the influencing factor, as NO_2 being known as the traffic-based air pollutant, showed the positive association with the road length factor in the final LUR models for all the seasons. The residential area seems to be another effective factor, showing a positive association with the concentration of NO_2. The reason is that people in the household uses the fuel-based generators for the electricity and combustion processes, and because people travel frequently around residential area, causing increase in the concentration of NO_2 due to vehicle exhaust emissions (Martin, A.N. et al. 2016). Distance variable also shows the contribution in all the models, highlighting the importance of local automobile workshops in the emission of NO_2. Hospital area showing the contribution of NO_2 from generators facility used in the vicinity of the hospital area, and local automobile workshop.

Tertiary road was found to be the primary influencing factor, NO_2 being traffic-related air pollutant, and the concentration was high due to the presence of roads around the residential area and people move frequently to do their errands. VMW was also found to be significant factor in the emissions.

Buffer radii of the tertiary road increased from 50 m to 100 m in monsoon season, and the increase in road length signifies the higher concentration in NO_2. Hospital area entered in the final model, also showed another factor for the source of increase in concentration. Higher concentration was observed in this season as compared to pre-monsoon.

As comparison to other two season models, Tertiary Road buffer radii was higher (300m), influencing in the concentration of pollutant. In addition, secondary road was another factor in this season found to be different from other two seasons and cause in the concentration of pollutant. High roads density (Tertiary and secondary roads), heating emissions from the residential area increase the concentration, that was the reason, highest concentration was observed in this season **(Table 15)**.

Table 15. Difference in variables for each season

Pre-monsoon	Monsoon	Post-monsoon
TR_50 m Tertiary road length within 50m	**TR_100 m** Tertiary road length within 100m	**TR_300 m** Tertiary road length within 300m
LURA_100 m Residential area within 100m	**LURA_100 m** Residential area within 100m	**LURA_100 m** Residential area within 100m
DV_300 m Distance variable within 300m	**DV_300 m** Distance variable within 300m	**DV_300 m** Distance variable within 300m
	LUHA_1000 m Hospital area within 1000m	**SR_1000 m** Secondary road length within 1000m

Electricity generation and its demand is a serious issue in developing countries, especially in Pakistan, which leads to the usage of fuel-based generators to produce electricity, in the households and hospitals, causing an increase in the pollutant concentration. The residential area also contributes as a source of NO_2 pollution, due to the fact of burning of fossil fuels in the households, especially during the heating season (post-monsoon) resulting in the highest concentration during all the seasons. Specific local data survey (SLDS) was helpful in determining another factor, which was vehicle maintenance workshops, surrounded around the roads and residence area, cause of increase in the concentration of pollutant due to maintenance of vehicles.

The significant predictor variables identified in this study are like other studies conducted previously, showing the length of roads and residential land area are common influencing factor (Madsen, C. et al. 2011; Aguilera, I. et al. 2008). The study conducted in Taipei city, Taiwan also reflected that the high concentrations attributed to road length, as one of the influencing factors and to the dense road network (Lee, J.H. et al. 2014). Another study, also showed one of the influencing factors were major roads and traffic influence, included in the models for heating and non-heating seasons (Chen, L. et al. 2010). The study conducted in Nanjing, China indicated that the residential area within 100m and 5000m buffer, entered in the final model, proved to be a significant predictor variable (Huang, L. et al. 2017). The developed models based on the influencing predictor variables in this study comparable with the other studies conducted previously, supporting the LUR models in the urban settings of Lahore, Pakistan.

Among the three seasons, predicted concentration regression maps showed the similar spatial characteristics with high concentration in the city center, where the residential land area is high due to the population. Next to residential area land feature, the maximum concentration can be observed around the road network, which is mainly distributed in the center and to the west part of the city. The vehicle maintenance workshops also surrounded around the roads, and nearby the residential area and hospital area, showing the public living to nearby those areas could have negative effect on their health, due to exposure to the pollutants. Since, NO_2 pollutant concentration relates to traffic intensity and residential area, so this spatial distribution is reasonable.

6. Conclusions and recommendations

Land use regression models for seasonal variation of NO_2 were developed based on different predictor variables for the first time in a specific city of Pakistan. The final generated models showed the precision of the final models which can be related with the output of R^2 values. The specific local data survey was useful in determining another local source of pollution in the city of Lahore, which was vehicle maintenance workshops. Inclusion of VMW in the final models of all the seasons showed the importance of this variable. This study shown that the higher concentration can be seen around the areas surrounded by roads, residential areas, and vehicle maintenance workshops, whereas lower concentration was observed around the parts of city, having not strong road network. People living nearby such areas are more affected, and exposed to air pollutants, and can increase the possibility of bad health outcomes. The developed models would more contribute to the broader use of LUR models in Pakistan, and if connected with the people's health indicators, this research can be helpful in providing support to provide epidemiological based studies in future.

Different studies used various variables to develop LUR models and every new variable have significance in the emission of pollutant. It is recommended to use the general survey to collect the data of specific city. Monitoring locations also have significance in the accuracy of the model. Number of monitoring sites and their locations should cover the specific area.

7. Limitations and Future directions:

There were some limitations in this study.

1. The selected measurements sites may not be able to capture the pollutant concentration distribution across the whole study area because these sites were mostly in the city area and therefor lack the coverage resolution of rural area, which can be considered in the future studies.
2. Although NO_2 concentration related to the industrial emissions, but due to limited availability of industrial data, did not enter the final model, which can be considered by using the specific local data survey to capture the industrial sites.
3. Data of health indicators i.e., body weight, breathing rate etc. is helpful in performing health risk assessment studies due to these pollutants, but due to unavailability of data, these types of studies can be performed in the future using developed models to have a better and clear understanding of health risk assessment and can be used to have possible mitigation and adaptations measures.
4. As these models were developed for the specific year of the monitored pollutant concentration, the limitation is that they can perform well within the certain period. As situation of air pollution and concentration of pollutants is changing with the cities, it is necessary to develop these types of models for the specific year and certain period to have the clear knowledge of air pollution spatial distribution and exposure assessment.

Future studies can be performed by incorporating different new type of air pollution sources present within the cities. For example, in the cities of Pakistan, local three-wheeler mode of transport is used, which is very cheap and convenient, but the problem lies is the emissions. The name is known as "Auto-rickshaw" or "Qingqi" **(Figure 16)**

It is pertinent to mention here that the number of autorickshaw in Pakistan is not known officially because of the unclear differences in the motorcycle and autorickshaw registration. Some of these autorickshaws made in local workshops, and maybe never got registration as officially three-wheeler operating motorcycle (26th World Road Congress).

Figure 16. Auto-rickshaw or Qingqi (A, B)

7.1. Regulatory framework and policy implications:

Air pollution in Pakistan is worst in the world, and it causes severe damages to human health, environment, and the economy. Pakistan is a developing country with increase in population, urbanization, and energy use. Urban centers are becoming the agglomeration of pollution sources which creates air pollution problems. In this scenario, it is better to have the policy implementations for the reduction of air pollution issue. First, monitoring of pollutants is a necessary step. After that, development of cost effective and reliable exposure assessment methods/models/techniques should be developed to have the spatial distributions of different pollutants within the cities. The model developed in this study can be a step towards achieving the potentials for developing new models. Furthermore, this model can be applied for other cities too, and can be incorporated with the addition of different data sources such as remote sensing and/or satellite observations. With respect to health implications studies, this research can provide a suitable modelling option to access the health risk assessment studies, so it is important to have the necessary data availability within the regulatory framework, which can be included in the policy and institutional capacity.

References:

Almetwally, A.A., M. Bin-Jumah, and A.A. Allam, Ambient air pollution and its influence on human health and welfare: An overview. Environmental Science and Pollution Research, 2020. **27**(20): p. 24815-24830.

Abrar, A., W. Sundas, F. Perveen and M. Habib (2014). "Air quality monitoring of some gaseous pollutants at selected points in Gullberg II, Lahore, Pakistan". International Research Journal of Environmental Sciences. Vol **3**: 38-47.

Aziz, A., and I. U. Bajwa. 2004. "Energy and Pollution Control Opportunities for Lahore." Urban Transport 16: 751–60.

Bourdrel, T., et al., Cardiovascular effects of air pollution. Archives of cardiovascular diseases, 2017. **110**(11): p. 634-642.

Beelen, R., G. Hoek, D. Vienneau, M. Eeftens, K. Dimakopoulou, X. Pedeli, M.-Y. Tsai, N. Künzli, T. Schikowski et al. (2013). "Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe–the ESCAPE project". Atmospheric Environment. Vol 72: Pages 10-23.

Block, M. L., A. Elder, R. L. Auten, S. D. Bilbo, H. Chen, J.-C. Chen, D. A. Cory-Slechta, D. Costa, D. Diaz-Sanchez et al. (2012). "The outdoor air pollution and brain health workshop". Neurotoxicology. Vol **33**(5): Pages 972-984.

Briggs, D. J., S. Collins, P. Elliott, P. Fischer, S. Kingham, E. Lebret, K. Pryl, H. Van Reeuwijk, K. Smallbone et al. (1997). "Mapping urban air pollution using GIS: a regression-based approach". International Journal of Geographic Information Science, 11(7): 699-718.

Briggs, D. J. J. o. T. and P. A. Environmental Health (2005). "The role of GIS: coping with space (and time) in air pollution exposure assessment". Journal of Toxicology and Environmental health Part A **68**(13-14): 1243-1261.

Brunekreef, B. and S. T. Holgate (2002). "Air pollution and health". Lancet **360**(9341): 1233-1242.

Chen, Z., M. T. Salam, S. P. Eckel, C. V. Breton and F. D. Gilliland (2015). "Chronic effects of air pollution on respiratory health in Southern California children: findings from the Southern California Children's Health Study". Journal of Thoracic Disease **7**(1): 46-58.

Ciccone G, Forastiere F, Agabiti N., 1998. Road traffic and adverse respiratory effects in children. SIDRIA Collaborative Group. Occup Environ Med 55, 771–778.

Colbeck, I., Z. Nasir, and Z. Ali (2009). "The state of ambient air quality in Pakistan-a review." Environmental science and pollution research international **17**: 49-63.

Colbeck, I., Z. A. Nasir, Z. Ali, and S. Ahmad (2010). "Nitrogen dioxide and household fuel use in Pakistan". Science of Total Environment **409**(2): 357-363.

Colbeck, I., Z. A. Nasir, Z. Ali (2010). "The state of ambient air quality in Pakistan—a review". Environmental Science and Pollution Research 17(1): 49-63.

Cohen, A.J., et al., Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet, 2017. **389**(10082): p. 1907-1918.

Cohen, A. J., Ross Anderson, H., Ostro, B., Pandey, K. D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope, A., Romieu, I., Samet, J.M., Smith, K. (2005). The global burden of disease due to outdoor air pollution. Journal of Toxicology and Environmental Health, Part A, 68(13-14), 1301-1307.

Cerro, J.C., et al., Variability of air pollutants, and PM composition and sources at a regional background site in the Balearic Islands: Review of western Mediterranean phenomenology from a 3-year study. Science of The Total Environment, 2020. **717**: p. 137177.

Colvile R.N., Hutchison E.J., Mindell J.S. and Warren R.F., The transport sector as a source of air pollution, Atmos. Environ, 35(9), 1537-1565 (2011)

Dhital, S., and D. Rupakheti, Bibliometric analysis of global research on air pollution and human health: 1998–2017. Environmental Science and Pollution Research, 2019. **26**(13): p. 13103-13114.

Feigin, V.L., et al., Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet Neurology, 2016. **15**(9): p. 913-924.

Gilbert, N. L., M. S. Goldberg, B. Beckerman, J. R. Brook, M. Jerrett (2005). "Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model". Journal of the Air and Waste Management and Association **55**(8): 1059-1063.

Gu, H., et al., Air pollution risks human mental health: an implication of two-stages least squares estimation of interaction effects. Environmental Science and Pollution Research, 2020. **27**(2): p. 2036-2043.

Grennfelt, P., et al., Acid rain and air pollution: 50 years of progress in environmental science and policy. Ambio, 2020. **49**(4): p. 849-864.

Grineski, S.E., T.W. Collins, and D.X. Morales, Asian Americans, and disproportionate exposure to carcinogenic hazardous air pollutants: A national study. Social Science & Medicine, 2017. **185**: p. 71-80.

Hoek, G., R. Beelen, K. De Hoogh, D. Vienneau, J. Gulliver, P. Fischer, and D. J. A. e. Briggs (2008). "A review of land-use regression models to assess spatial variation of outdoor air pollution". Atmospheric Environment **42**(33): 7561-7578.

Hang, J., et al., The influence of street layouts and viaduct settings on daily carbon monoxide exposure and intake fraction in idealized urban canyons. Environmental pollution, 2017. **220**: p. 72-86.

Ilyas, S. Z., S. M. Nasir, and S. M. Raza (2008). "Air pollution problems and diseases caused by hazardous gases in Quetta, Pakistan". Journal of Applied Sciences and Environmental Management **12**(1).

Jamison, D. T., J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, A. Mills and P. Musgrove (2006). Disease control priorities in developing countries, The World Bank.

Johnson, M., V. Isakov, J. Touma, S. Mukerjee et al. (2010). "Evaluation of land-use regression models used to predict air quality concentrations in an urban area". Atmospheric Environment 44(30): 3660-3668.

Janssen, N.A., et al., Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM_{10} and $PM_{2.5}$. Environmental health perspectives, 2011. **119**(12): p. 1691-1699.

Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. Environmental pollution, 151(2), 362-367.

Kim, Y., and J.-M. Guldmann (2015). "Land-use regression panel models of NO_2 concentrations in Seoul, Korea". Atmospheric Environment **107**: 364-373.

Kinney, P.L., Interactions of climate change, air pollution, and human health. Current environmental health reports, 2018. **5**(1): p. 179-186.

Kunzli, N., & Tager, I. B. (2005). Air pollution: from lung to heart. Swiss Med Wkly, 135(47-48), 697-702.

Lee, J.-H., C.-F. Wu, G. Hoek, K. de Hoogh, R. Beelen, B. Brunekreef et al. (2014). "Land use regression models for estimating individual NOx and NO_2 exposures in a metropolis with a high density of traffic roads and population". Science of the Total Environment 472: 1163-1171.

Liao, N.S., et al., Particulate air pollution and risk of cardiovascular events among adults with a history of stroke or acute myocardial infarction. Journal of the American Heart Association, 2021. **10**(10): p. e019758.

Lodhi, A, B. Ghauri, M.R. Khan, S. Rahman, S. Shafique (2009). Particulate matter (PM2.5) concentration and source apportionment in Lahore, Journal of Brazilian Chemical Society, Vol. 20, No. 10.

Madsen, C., U. Gehring, S. E. Håberg, P. Nafstad, K. Meliefste, W. Nystad, K. C. L. Carlsen and B. Brunekreef (2011). "Comparison of land-use regression models for predicting spatial NOx contrasts over a three-year period in Oslo, Norway". Atmospheric Environment **45**(21): 3576-3583.

Mage, D., Ozolins, G., Peterson, P., Webster, A., Orthofer, R., Vandeweered, V., et al. (1996). Urban air pollution in megacities of the world. Atmospheric Environment, 30, 681–686.

Mehmood, T., Z. Ali, N. Noor, S. Sidra, Z. A. Nasir, and I. Colbeck (2016). Measurement of NO_2 indoor and outdoor concentrations in selected public schools of Lahore using passive sampler, Pakistan Agricultural Scientists Forum.

Meng, X., L. Chen, J. Cai, B. Zou, C.-F. Wu, Q. Fu, Y. Zhang, Y. Liu, and H. Kan (2015). "A land use regression model for estimating the NO2 concentration in Shanghai, China". Environmental Research **137**: 308-315.

Ma, Y., et al., Initial simulated acid rain impacts reactive oxygen species metabolism and photosynthetic abilities in Cinnamonum camphora undergoing high temperature. Industrial Crops and Products, 2019. **135**: p. 352-361.

Majid H., Madl P., Alam K. (2012). Ambient air quality with emphasis on roadside junctions in metropolitan cities of Pakistan and its potential health effects. Health, 3, 79-85.

Morrison-Saunders, A., Advanced introduction to environmental impact assessment. 2018: Edward Elgar Publishing.

Pakistan Millenium Development Goals Report (2005) Government of Pakistan Planning Commission. Islamabad, Pakistan

Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (2013). World Population Prospects: The 2012 Revision. New York: United Nations

Pak EPA/JICA (2001). Three Cities Investigation of Air and Water Quality with Analytical Comments (Lahore, Rawalpindi, and Islamabad). Pakistan Environmental Protection Agency and Japan International Cooperation Agency.

Niaz, Y., Zhou, J., Iqbal, M., Nasir, A., & Dong, B. (2015). Ambient air quality evaluation: a comparative study in China and Pakistan. Polish Journal of Environmental Studies, 24(4).

W.H.O (2016). World health statistics 2016: monitoring health for the SDGs. Sustainable Development Goals, World Health Organization.

Olanipekun, J.A. and J.O. Babatunde, Towards Reducing the Burden of Global Environmental Related Health Problems in the 21st Century. Journal of Education and Practice, 2016. **7**(32): p. 57-64.

Pak EPA, Pakistan Environmental Protection Agency (2005). State of the Environment Report, Draft.

Pope III, C. A., D. W. Dockery (2006). "Health effects of fine particulate air pollution: lines that connect". Journal of Air and Waste Management Association **56**(6): 709-742.

Pope III, C.A., M. Ezzati, and D.W. Dockery, Fine-particulate air pollution and life expectancy in the United States. New England Journal of Medicine, 2009. **360**(4): p. 376-386.

Pope III, C.A., Mortality effects of longer-term exposures to fine particulate air pollution: review of recent epidemiological evidence. Inhalation toxicology, 2007. **19**(sup1): p. 33-38.

Rao, M., L. A. George, T. N. Rosenstiel, V. Shandas and A. Dinno (2014). "Assessing the relationship among urban trees, nitrogen dioxide, and respiratory health". Environmental Pollution **194**: 96-104.

Rojas, N.Y., et al., $PM_{2.5}$ emissions, concentrations, and air quality index during the COVID-19 lockdown. Environmental Pollution (Barking, Essex: 1987), 2021. **272**: p. 115973.

Ramachandran, G., Adgate, J.L., Pratt, G.C., Sexton, K., 2003. Characterizing indoor and outdoor 15-minute average PM2.5 concentrations in urban neighborhoods. Aerosol Science and Technology. 37, 33–45.

Sanchez-Triana, E., S. Enriquez, J. Afzal, A. Nakagawa, and A. S. Khan (2014). Cleaning Pakistan's air: policy options to address the cost of outdoor air pollution, The World Bank.

Starkey, P., Batool, Z. (2019). Qingqis in Pakistan: public transport strategies that consider social, economic, traffic safety and environmental issues. In: Proceedings of the 26th PIARC World Road Congress, 6-10 October 2019, Abu Dhabi, United Arab Emirates.

Sahsuvaroglu, T., Arain, A., Kanaroglou, P., et al., 2006. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. Journal of Air Waste Management Association 56, 12.

Sughis et al. 2012. Blood pressure and particulate air pollution in schoolchildren of Lahore, Pakistan. BMC Public Health 12:378.

Thron, R.W., 1996. Direct and indirect exposure to air pollution. Otolaryngol. Head Neck Surg 114, 281-285.

Facts Sheet, W.H.O (2014). "Ambient (outdoor) air quality and health".

Trasande, L., G. D. Thurston (2005). "The role of air pollution in asthma and other pediatric morbidities". The Journal of allergy and clinical immunology **115**(4): 689-699.

Wang, W., Wang, T., 1995. On the origin and the trend of acid precipitation in China. Water Air Soil Pollut. 85, 2295‐2300

Xie, R., et al., (2019) The effect of traffic density on smog pollution: Evidence from Chinese cities. Technological Forecasting and Social Change. **144**: p. 421-427.

Yu, B., Chen, F., Shirazi, Z., 2017. Impact of biomass fires on tropospheric nitrogen dioxide in South and Southeast Asia for the years from 2005 to 2014. Appl. Geography. 86, 92–101.