Architecture
Image Studies

A Mathematical Model for Sustainable Urban Planning

Sarita Agarwal[1], Shivam Gupta[2], Ramashish Kumar[3], Vaibhav Kumar[4], Nikhil Yadav[5]

Abstract

Delhi has undergone rapid urbanization over the past three decades, transforming vast areas of vegetation, agricultural land, and open spaces into built-up urban fabric. During this period Delhi's air pollution has worsened dramatically: for example, modeled annual PM 2.5 levels rose by ~40% from 1998 to 2020, and recent reports identify New Delhi as the world's most polluted capital. In this study we have observed how the land of Delhi is utilized in different categories – water bodies, vegetation, forest, agriculture, built-up areas, and open land – related to long-term air quality trends in Delhi. Using satellite-derived land cover data and annual pollution records (e.g. PM 2.5 concentrations), we have applied correlation and regression techniques to quantify the influence of each land use type on ambient pollution levels over 1991–2021.

Introduction

New Delhi's unprecedented growth in population and built environment has been accompanied by a persistent air quality crisis. Today's Delhi metropolitan area (≈33 million people) is one of the fastest-growing megacities globally, and its residents routinely breathe some of the world's most polluted air. Regulatory monitors and satellite models indicate that Delhi's mean PM 2.5 (fine particulate matter) concentrations increased sharply over the 1990s and 2000s, peaking above 125 μg/m³ around 2010 – well above the WHO guideline [15] of 5 μg/m³. Recent analyses show that India as a whole rank near the top in global air pollution, with Delhi identified as the foulest capital city. For instance, the 2023 World Air Quality Report noted a ~10% rise in Delhi's PM 2.5 from the previous year, reaching a monthly average of 255 μg/m³ in November [7]. Such extreme pollution – driven by vehicle emissions, industrial sources, biomass burning and regional stubble fires – poses severe health and environmental risks across the city.

Underlying Delhi's air pollution trends is an equally dramatic transformation of land use and land cover. Remote sensing studies of the National Capital Territory (NCT) document that built-up urban area has expanded to become the dominant land use – comprising over 50% of Delhi's territory by 2020. In the same period (1991–2020), builtup area increased by roughly 271 km²[9], largely through conversion of vacant land, agricultural fields, parks, and even water bodies. According to Singh and Biswas (2022)[1] report that by 2020 Delhi's land cover was ~53% built, ~21% vegetation, ~25% vacant/open land, and only ~1% surface water. This means Delhi has lost much of its urban green and blue space to development. Such land cover change alters the city's microclimate and pollutant dynamics: impervious surfaces absorb and re-emit heat and reduce natural deposition of particulates, while reductions in vegetation and water area eliminate sinks that would otherwise remove some pollutants from the air.

In general, the literature suggests that urban greenery and water bodies mitigate air pollution, whereas dense built infrastructure tends to exacerbate it. For instance, studies in Delhi have found that personal PM 2.5 exposure during walking is lower in areas with higher tree cover and vegetation index (NDVI) in the warmer seasons. This indicates that pockets of urban green space can locally reduce particulate concentrations. Conversely, compact built environments increase travel distances and traffic congestion, leading to higher fuel consumption and emission of pollutants. Indeed, sustainable land-use planning frameworks emphasize that integrated city design – including mixed-use development and preservation of green infrastructure – is essential to control urban air pollution and

---
[1] Department of Mathematics, Acharya Narendra Dev College, University of Delhi INDIA
[2] Department of Mathematics, Acharya Narendra Dev College, University of Delhi INDIA.
[3] Department of Mathematics, Acharya Narendra Dev College, University of Delhi INDIA
[4] Department of Mathematics, Acharya Narendra Dev College, University of Delhi INDIA
[5] Department of Mathematics, Acharya Narendra Dev College, University of Delhi INDIA

achieve environmental sustainability. However, most existing air quality studies in Delhi have focused on emissions sources, meteorological factors, or policy interventions, with relatively little quantitative analysis of how aggregate land use patterns influence long-term pollution trends.
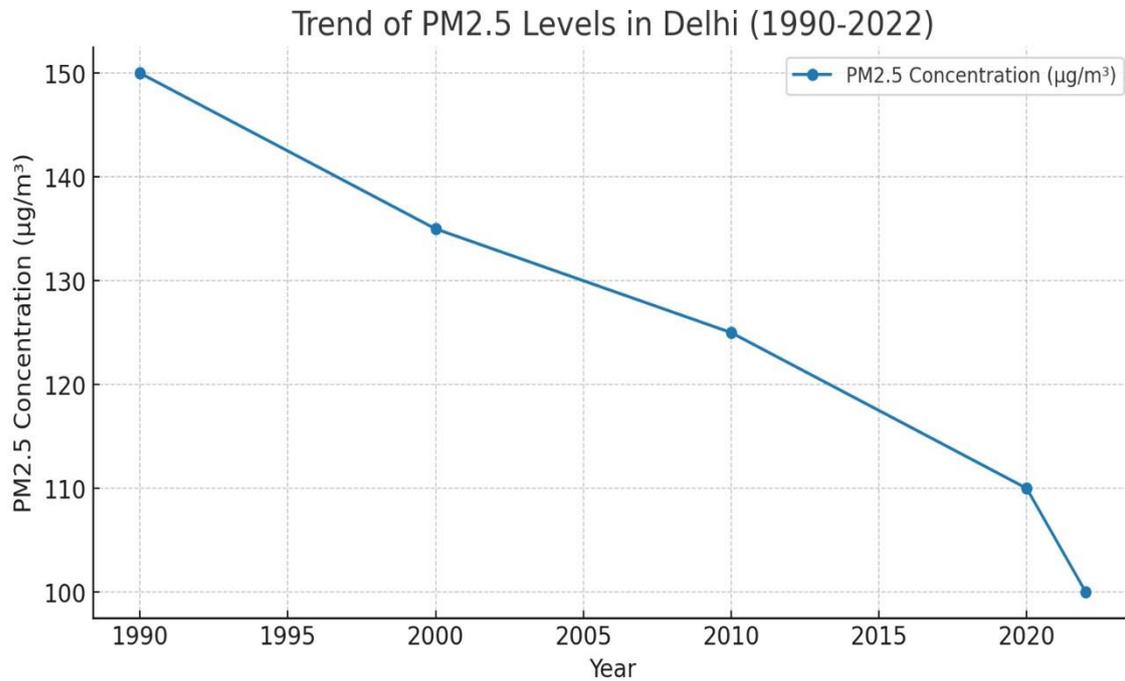
This research addresses that gap by statistically evaluating the relationships between multiple land use types and Delhi's air pollution levels from 1991 through 2021[14]. Specifically, we assemble time series of spatial extents (or proportions) for six land use categories – water bodies, vegetation cover, forest, agriculture, built-up areas, and open/vacant land – alongside annual measures of ambient pollution (e.g. $PM_{2.5}$ concentration) in the NCT of Delhi. Using multiple correlation and multivariate regression analyses, we quantify how year-to-year changes in each land use category are associated with changes in pollution. This approach allows us to isolate the contribution of each land type while controlling for interdependencies among them. By doing so, we can identify which land cover changes tend to exacerbate or alleviate air quality problems in Delhi.

Understanding these links is crucial for urban planning and environmental policy. If, for example, expanded urban vegetation or water features are shown to significantly reduce particulate pollution, city planners may prioritize green space conservation and wetland management as part of air quality improvement strategies. Conversely, demonstrating that uncontrolled expansion of impervious built-up or open dusty areas correlates with pollution spikes would support stricter land-use regulation. Ultimately, our goal is to inform sustainable land use planning in megacities: by empirically linking three decades of urban development patterns with air quality outcomes, we provide evidence that can guide efforts to balance Delhi's growth with the health of its environment.

Methodology

This study employs a structured approach to investigate the relationship between land use patterns and air quality index in Delhi from 1991 to 2021, focusing on six land use categories: water bodies, vegetation, forest, agriculture, built-up areas, and open/vacant land. The methodology consists of three main stages: literature review, data collection, and statistical analysis.

Literature Review

The research began with an extensive review of existing literature on urban land use change, its environmental impacts, and the role of green and blue infrastructure in mitigating air pollution. Peer-reviewed journal articles, conference proceedings, and government reports were examined to understand the theoretical framework, identify relevant variables, and select appropriate statistical techniques. Special emphasis was placed on studies related to Delhi and other megacities facing similar urbanization and air quality challenges.

Data Collection

Data for land use and air pollution were collected from authoritative and credible government sources, ensuring reliability and consistency.

Land use data were obtained from official agencies such as the Delhi Development Authority (DDA), Census of India, and remote sensing databases maintained by national space research institutions. These datasets provided yearly or periodic records of land use distribution across the defined categories.

Air pollution data—with a focus on particulate matter concentrations (PM $_{2.5}$)—were sourced from the Central Pollution Control Board (CPCB), Delhi Pollution Control Committee (DPCC), and publicly available national air quality monitoring portals. Wherever available, satellite-derived air quality estimates were used to fill temporal gaps.

All datasets were processed to maintain uniform spatial boundaries for the National Capital Territory (NCT) of Delhi and standardized to cover the period 1991–2021.

Summary of Approach

By integrating insights from the literature with 30 years of empirical data and robust statistical methods, this study establishes a quantitative understanding of how land use changes influence air quality in Delhi. The methodology ensures that findings are both evidence-based and contextually relevant for urban environmental policy-making.

This research addresses that gap by statistically evaluating the relationships between multiple land use types and Delhi's air pollution levels from 1991 through 2021. Specifically, we assemble time series of spatial extents (or proportions) for six land use categories – water bodies, vegetation cover, forest, agriculture, built-up areas, and open/vacant land – alongside annual measures of ambient pollution (e.g. $PM_{2.5}$ concentration) in the NCT of Delhi. Using simple correlation and multiple correlation, we quantify how year-to-year changes in each land use category are associated with changes in pollution. This approach allows us to isolate the contribution of each land type while controlling for interdependencies among them. By doing so, we can identify which land cover changes tend to exacerbate or alleviate air quality problems in Delhi.

Understanding these links is crucial for urban planning and environmental policy. If, for example, expanded urban vegetation or water features are shown to significantly reduce particulate pollution, city planners may prioritize green space conservation and wetland management as part of air quality improvement strategies. Conversely, demonstrating that uncontrolled expansion of impervious built-up or open dusty areas correlates with pollution spikes would support stricter land-use regulation. Ultimately, our goal is to inform sustainable land use planning in megacities: by empirically linking three decades of urban development patterns with air quality outcomes, we provide evidence that can guide efforts to balance Delhi's growth with the health of its environment.

Mathematical Analysis

To assess the relationship between land use patterns and air pollution, the study will employ both simple correlation analysis and multiple correlation analysis.

Simple correlation and multiple corelations are used to measure the strength and direction of the linear relationship between each individual land use category and air pollution levels while multiple correlation analysis is applied to evaluate the combined effect of all land use variables on air pollution, allowing for the control of interdependencies among them.

The correlation coefficients determine whether changes in specifies types are associated with an increase in pollution levels. Statistical components carries out using standard statistical packages to ensure accuracy and readings.

For mathematical analysis, we need the data which has been collected from various websites. Some of them are official like government portals and some non-official websites like newspapers, articles etc.

Land use pattern of Delhi

Delhi Land Use Data (1991 - 2021)

				(in %)
LULC Class	1991	2001	2011	2021
Water body	1.2	1.1	1.61	1.36
Vegetation	12.55	10.38	11.06	7.35
Forest	10.19	7.45	8.86	17.73
Agriculture	35.18	33.45	29.53	26.31
Built-up	17.51	28.07	31.73	38.01
Open space	23.37	19.55	17.21	9.24
AQI	363	152	220	257
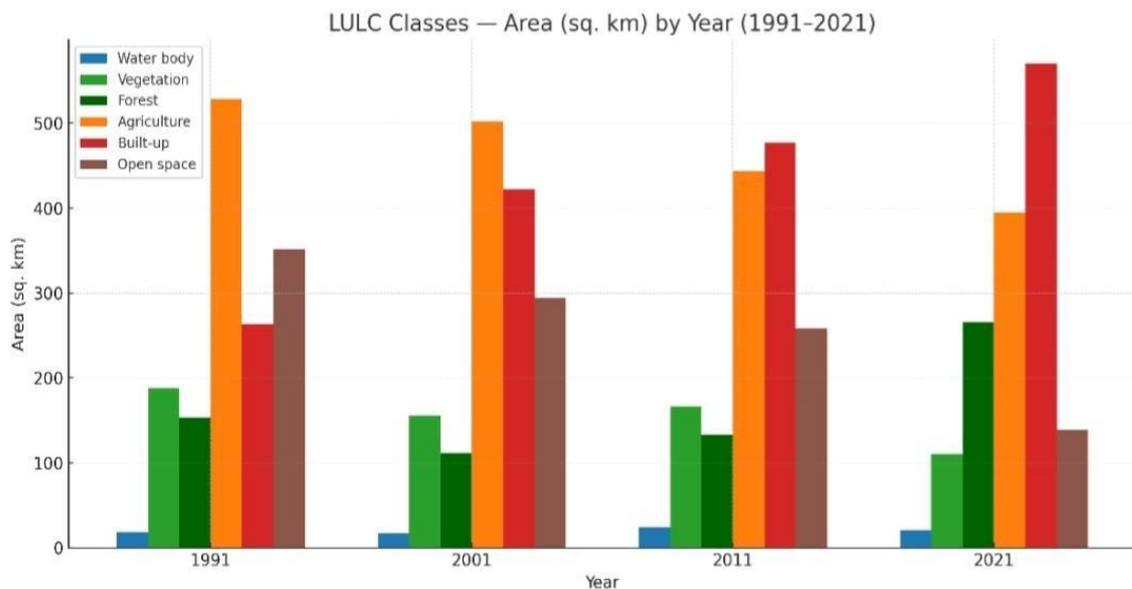
Mathematical Analysis of AQI (Delhi: 1991 - 2021)

AQI(Y)	Water body(X_1)	Vegetation(X_2)	Forest(X_3)	Agriculture(X_4)	Builtup(X_5)	Open space(X_6)
363	1.2	12.55	10.19	35.18	17.51	23.37
152	1.1	10.38	7.45	33.45	28.07	19.55
220	1.61	11.06	8.86	29.53	31.73	17.21
257	1.36	7.35	17.73	26.31	38.01	9.24

Correlation Analysis

Correlation quantifies how strong and in which direction the two variables are related. It is being measured using the correlation coefficient which goes to the extent of negative one and positive one.

The correlation coefficient, written as "r," is as follows:

$$r_{XY} = \frac{n(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{\sqrt{[n\sum X_i^2 - (\sum X_i)^2][n\sum Y_i^2 - (\sum Y_i)^2]}}$$

where

- n is sample size

- X and Y are the individual sample points indexed with *i*.

The correlation coefficient, denoted by *"r"*, ranges between -1 and 1.

CORRELATION		
r_{YX1} = -0.00748	r_{YX3} = 0.30353093	r_{YX5} = -0.540341299
r_{YX2} = 0.35151	r_{YX4} = 0.23214489	r_{YX6} = 0.261073777
r_{X1X2} = -0.1266	r_{X3X4} = -0.7395786	r_{X5X6} = -0.940520385

Correlation Coefficient between:

- AQI and Water Body is -0.00748.
- AQI and Vegetation is 0.35151.
- Water Body and Vegetation is -0.12666.

Correlation coefficient between:

- AQI and Forest is 0.30353093.
- AQI and Agriculture is 0. 23214489.
- Forest and Agriculture is -0.7395786.

Correlation coefficient between:

- AQI and Built-up is -0.540341299
- AQI and Open Space is 0.261073777.
- Built-up and Open Space is -0.940520385.

Multiple Correlation Coefficient:

The multiple correlation coefficient, often denoted as R, measures the strength of the relationship between a dependent variable and two or more independent variables. It indicates how well the independent variables, taken together, can predict the dependent variable using a linear model. Essentially, it's an extension of the simple correlation coefficient, which only handles two variables.

$$R_{x.yz} = \sqrt{\frac{r_{xy}^2 + r_{yz}^2 - 2r_{xy}r_{xz}r_{yz}}{1 - r_{yz}^2}}$$

Where, r_{xy} is the correlation coefficient between variable X_1 and X_2.

r_{yz} is the correlation coefficient between variable X_2 and X_3.

r_{xz} is the correlation coefficient between variable X_1 and X_3.

MULTIPLE CORRELATION		
$R_{Y.X1X2}$ = 0.353489	$R_{Y.X3X4}$ = 0.74323351	$R_{Y.X5X6}$ = 0.906142268

Multiple correlation coefficient between AQI, Water body and Vegetation is 0.353489.

Linear Regression Model [4]:

$$Y = a + b_1 X_1 + b_2 X_2$$

where Y is AQI, X_1 is Water body, X_2 is Vegetation.

$$b_1 = \frac{r_{YX_1} - r_{YX_2} r_{X_1 X_2}}{1 - (r_{X_1 X_2})^2} \frac{\sigma_Y}{\sigma_{X1}}$$

$$b_2 = \frac{r_{YX_2} - r_{YX_1} r_{X_1 X_2}}{1 - (r_{X_1 X_2})^2} \frac{\sigma_Y}{\sigma_{X_2}}$$

$$a = \bar{Y} - b_1 \bar{X_1} - b_2 \overline{X_2}$$

$\bar{Y}$ is mean, $\overline{X_1}$ is mean, $\overline{X_2}$ is mean.

r_{YX_1} = correlation coefficient between Y and X_1.

r_{YX_2} = correlation coefficient between Y and X_2

$r_{X_1 X_2}$ = correlation coefficient between X_1 and X_2.

σ_Y = Standard deviation of Y.

σ_{X_1} = Standard deviation of X$_1$.

σ_{X_2} = Standard deviation of X$_2$.

$$Y = 79.9260 + 14.9159 X_1 + 14.3611 X_2$$

Multiple correlation coefficient between AQI, Forest and Agriculture is 0.74323351.

Linear Regression Model:

$$Y = a + b_3 X_3 + b_4 X_4$$

$$Y = 20.1572 X_3 + 22.3090 X_4 - 669.0900$$

where Y is AQI, X_3 is Forest, X_4 is Agriculture.

$$b_3 = \frac{r_{YX_3} - r_{YX_4} r_{X_3 X_4}}{1 - (r_{X_3 X_4})^2} \frac{\sigma_Y}{\sigma_{X_3}}$$

$$b_4 = \frac{r_{YX_4} - r_{YX_3} r_{X_3 X_4}}{1 - (r_{X_3 X_4})^2} \frac{\sigma_Y}{\sigma_{X_4}}$$

$$a = \bar{Y} - b_3 \bar{X_3} - b_4 \overline{X_4}$$

$\bar{Y}$ is mean, $\overline{X_3}$ is mean, $\overline{X_4}$ is mean

r_{YX_3} = correlation coefficient between Y and X_3

r_{YX_4} = correlation coefficient between Y and X_4

σ_Y = Standard deviation of Y.

σ_{X_3} = Standard deviation of X3.

σ_{X_4} = Standard deviation of X4.

Multiple correlation coefficient between AQI, built – up and Open Space is 0.906142268.

Linear Regression Model

$$Y = 1551.8194 - 26.2046X_5 - 31.6182X_6$$

where Y is AQI, X_5 is Built-up, X_6 is Open Space.

$$b_5 = \frac{r_{YX_5} - r_{YX_6}\, r_{X_5X_6}}{1 - (r_{X_5X_6})^2}\, \frac{\sigma_Y}{\sigma_{X_5}}$$

$$b_4 = \frac{r_{YX_6} - r_{YX_5}\, r_{X_5X_6}}{1 - (r_{X_5X_6})^2}\, \frac{\sigma_Y}{\sigma_{X_6}}$$

$$a = \bar{Y} - b_5\bar{X}_5 - b_6\overline{X_6}$$

r_{YX_5} = correlation coefficient between Y and X_5

r_{YX_6} = correlation coefficient between Y and X_6

σ_Y = Standard deviation of Y.

σ_{X_5} = Standard deviation of X5.

σ_{X_6} = Standard deviation of X6.

Multiple correlation coefficient with all variables:

We find out the multiple correlation of AQI(Air Quality Index) with water body, vegetation, forest, agriculture, built-up space and open space for creating linear regression model. So, we have used least square method given in Friedberg ,Insel and Spence[5]

Multiple Regression Model:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6$$

$$b = ((X^TX)^{-1}X^T)Y$$

$$X = \begin{pmatrix} 1 & 1.2 & 12.55 & 10.19 & 35.18 & 17.51 & 23.37 \\ 1 & 1.1 & 10.38 & 7.45 & 33.45 & 28.07 & 19.55 \\ 1 & 1.51 & 11.06 & 8.86 & 29.53 & 31.73 & 17.21 \\ 1 & 1.36 & 7.35 & 17.73 & 25.17 & 38.01 & 9.24 \end{pmatrix}$$

$$X^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1.2 & 1.1 & 1.51 & 1.36 \\ 12.55 & 10.38 & 11.06 & 7.35 \\ 10.19 & 7.45 & 8.86 & 17.73 \\ 35.18 & 33.45 & 29.53 & 25.17 \\ 17.51 & 28.07 & 31.73 & 38.01 \\ 23.37 & 19.55 & 17.21 & 9.24 \end{pmatrix}$$

$$X^T X = \begin{pmatrix} 4 & 5.17 & 41.34 & 44.23 & 123.33 & 115.32 & 69.37 \\ 5.17 & 6.77 & 53.17 & 57.91 & 157.83 & 151.49 & 88.10 \\ 41.34 & 53.17 & 441.59 & 433.52 & 1300.32 & 1141.42 & 754.47 \\ 44.23 & 57.91 & 433.52 & 552.19 & 1315.58 & 1342.59 & 700.09 \\ 123.33 & 157.83 & 1300.32 & 1315.58 & 3862.08 & 3448.64 & 2216.88 \\ 115.32 & 151.49 & 1141.42 & 1342.59 & 3448.64 & 3546.07 & 1855.26 \\ 69.37 & 88.10 & 754.47 & 700.09 & 2216.88 & 1855.26 & 1309.92 \end{pmatrix}$$

$$(X^T X)^{-1} X = \begin{pmatrix} -1.25 & -0.375 & -1 & -0.25 \\ 0.5 & -3.125 & 3.1875 & -0.1875 \\ -0.093 & 0.125 & -0.203 & -0.03906 \\ 0.068 & -0.041 & -0.107 & 0.072 \\ 0.035 & -0.000 & -0.005 & -0.009 \\ -0.0625 & 0.099 & -0.0312 & 0.004 \\ 0.0312 & 0.039 & 0.0234 & -0.007 \end{pmatrix}$$

$$Y = \begin{pmatrix} 363 \\ 152 \\ 220 \\ 257 \end{pmatrix}$$

$$b = \begin{pmatrix} -795 \\ 359.56 \\ -69.75 \\ 13.51 \\ 8.81 \\ -13.16 \\ 20.42 \end{pmatrix}$$

$y = 359.56\ x_1 - 69.75\ x_2 + 13.51\ x_3 + 8.81\ x_4 - 13.16 x_5 + 20.42\ x_6 - 795$

Suggestions

On the basis of above model, a comprehensive table is constructed corresponding to the AQI range of 0–50, systematically presenting the recommended land area requirements across diverse land-use categories, while simultaneously providing a structured framework that supports sustainable urban planning, informed decision-making, and efficient resource allocation.

AQI*	50	40	30	20	10	0
Water body	5.041047	5.015086	4.989126	4.963165	4.937204	4.911244
Vegetation	21.33401	21.34443	21.35485	21.36527	21.3757	21.38612
Forest	18.17416	18.17753	18.1809	18.18426	18.18763	18.191

Agriculture	18.35254	18.3563	18.36006	18.36383	18.36759	18.27136
Buit-up	19.18633	19.19196	19.19758	19.20321	19.20884	19.21446
Open space	17.91192	17.9147	17.91748	17.92026	17.92304	17.92582

*AQI—Air Quality Index

Comparison of land use/land cover distribution under current AQI (257, year 2021) and suggested optimal AQI (50).

AQI	257	50
Water body	1.36	5.041047
Vegetation	7.35	21.33401
Forest	17.73	18.17416
Agriculture	26.31	18.35254
Built-up	38.01	19.18633
Open space	9.24	17.91192

Conclusion

Although correlation shows AQI relates proportionally to vegetation and water bodies which should not be the case because vegetation cleans up the atmosphere but vegetation cover of Delhi consists of many foreign invasive species which are not very beneficial for air quality.

Our analysis reveals that increases in urban built-up and open/vacant land are strongly associated with higher pollutant concentrations, whereas greater coverage of vegetation, forest, and surface water correlate with lower pollution levels. These statistical findings align with expectations that impervious, developed land tends to generate and trap more emissions, while green and blue spaces (trees, parks, water bodies) help remove airborne particulates.

The results underscore the importance of sustainable land use planning for Delhi's environmental health. Preserving and expanding urban green and blue infrastructure could mitigate pollution as the city grows. This study quantifies the land use–air quality relationship over 30 years, offering insights for urban planners to balance development with cleaner air.

Our model is useful for planners to prioritize green space in regulating urban expansion. The aim of this project is to balance Delhi's growth with environmental health through sustainable land use planning. Our results show how that AQI is highly correlated to build- up which is sensible because construction causes release of particulate matter which are major pollutants of air.

References

[1]   Bhadwal, S., & Kumar, M. (2025). Assessing Urban Green Infrastructure Transformation in Delhi (1991–2021): A Landscape Ecology and Remote Sensing Approach.

[2]   Central Pollution Control Board (CPCB). (2021). Annual Report on Ambient Air Quality in Delhi. Government of India.

[3]   Delhi Pollution Control Committee (DPCC). (2021). Monitoring of $PM_{2.5}$ Levels in Delhi: 1991–2021 Data. DPCC, NCT of Delhi.

[4]   Devore, J. L. Probability and Statistics (Ninth Edition)

[5]   Friedberg, S. H., Insel, A. J., & Spence, L. E. (n.d.). Linear Algebra (5th ed.).

[6]   Guttikunda, S.K. et al. (2023), Air Pollution in Delhi: Review from 1990 to 2022. MDPI Sustainability, 15(5), 4209. https://doi.org/10.3390/su15054209

[7]   IQAir. (2024), World Air Quality Report 2023: Delhi Among the Most Polluted Cities Globally.

[8]   Ministry of Environment, Forest and Climate Change (MoEFCC)(2018), Green Infrastructure and Air Quality: Policy Recommendations for Indian Cities.

[9]   NASA / NRSC. (2021), Land Cover Classification of National Capital Territory of Delhi, 1991–2020. Remote Sensing Dataset.

[10]   Planning Department, Delhi Development Authority (DDA). (2019),Land Use Data for NCT, Delhi.

[11]   Shah, R., & Kumar, N. (2019), Built Environment, Urban Heat, and Air Quality in Indian Megacities. Urban Climate, 29, 100498.

[12]   Singh, A., & Biswas, S. (2022), Urban Land Use Change and Environmental Impacts in Delhi (1991–2020). Journal of Urban Studies, 15(3), 55–72.

[13]   Tiwari, P., et al. (2020), Personal Exposure and Tree Cover Density: A

[14]   Comparison Across Delhi Neighbourhoods. Environmental Health Perspectives, 128(11).

[15]   Urban Emissions Info. (2023), What is Polluting Delhi's Air? Review from 1990 to 2022. [Infographic presentation].

[16]   WHO. (2021), Air Quality Guidelines – Global Update 2021. WHO Regional Office for Southeast Asia.

[17]   World Bank. Ten Years of Progress and Challenges in Urban Air Quality Management in India 1993–2002. www.worldbank.org/in

[18]   World Air Quality Report 2023. (2024). Delhi Among the Most Polluted Cities Globally. IQAir.

[19]   . https://cpcb.nic.in/upload/national-air-quality-index/AQI-Calculator.xls

[20]   https://en.m.wikipedia.org/wiki/Air_pollution_in_Delhi

[21]   https://en.wikipedia.org/wiki/Pearson_correlation_coefficient