Spatio-Temporal Analysis of Air Quality Disparities Across Urban and Rural India

Akshaya Srikrishna
Dept. of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
2022103065@student.annauniv.edu

Anagha Srikrishna
Dept. of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
2022103066@student.annauniv.edu

Krishnendu M. R.
Dept. of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
2022103081@student.annauniv.edu

Dr. Somu Renugadevi
Dept. of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
srenuga@annauniv.edu

***Abstract*—Air pollution is typically treated as an urban problem, with rural regions presumed cleaner. However, this assumption overlooks emerging pollution patterns in low-population areas. This analysis examines air quality in five pairs of urban-rural locations in India, calculating AQI using the Central Pollution Control Board framework. Data from September 2022 to January 2025 evaluated key pollutants ($PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3) through spatio-temporal decomposition, Random Forest Regression, and K-Means Clustering to identify pollution drivers and trends. Results reveal distinct seasonal patterns: $PM_{2.5}$ dominance in winter urban areas versus elevated ozone levels in rural summer conditions. We introduce a novel Pollution Disparity Index (PDI) to quantify urban-rural pollution differences, revealing significant disparities that challenge conventional assumptions. These findings demonstrate the critical need for region-specific air quality policies, enhanced satellite-based monitoring in rural areas, and predictive modeling frameworks that account for the complex spatial and temporal dynamics of air pollution across India's diverse landscape.**

***Index Terms*—Air Quality Index (AQI), Spatiotemporal analysis, Random forest, K-Means clustering, Pollution Disparity Index (PDI), Seasonal Air Pollution Trends, Air Pollution Monitoring, CPCB standards, Feature importance, Urban-rural air quality**

I. INTRODUCTION

Air pollution is a problem for the environment and public health, especially in fast-growing cities in India, where industrial growth, vehicle exhausts, and agricultural practices lead to air degradation. Although city centers are heavily affected by pollution caused by high population density and localized anthropogenic activities, rural areas are traditionally thought to enjoy cleaner air. However, this assumption ignores pollution spillover from neighboring cities, biomass burning, dust storms, and other local sources that play a role in rural air quality. Understanding these spatial and temporal variations is crucial for developing effective mitigation strategies tailored to specific regions.

To scientifically find these differences, this study investigates air quality trends in five sets of urban-rural locations in India based on standardized Air Quality Index (AQI) calculations following the guidelines of the Central Pollution Control Board (CPCB). AQI has been calculated based on a sub-index where the concentrations of the pollutants ($PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3) are translated into sub-indices following CPCB-established breakpoints and the maximum sub-index defines the final AQI. This method offers a uniform, comparable measure of pollution intensity between sites.

The research combines several analytical techniques to find the spatio-temporal pollution dynamics. Descriptive and comparative analysis emphasize pollution trends, while spatio-temporal decomposition methods determine long-term trends and episodic pollution peaks. Random Forest Regression identifies major drivers of AQI variation, providing information on pollutant contributions over time. Furthermore, K-Means clustering categorizes areas according to AQI severity, demonstrating clear urban-rural pollution patterns. To measure the disparity in urban-rural pollution, the work formulates the Pollution Disparity Index (PDI), presenting an objective indication of pollution spillover across densely located urban and rural areas ,an aspect often overlooked in traditional studies. By integrating machine learning with environmental analysis, the study bridges disciplinary gaps to enhance decision-making. With growing evidence of increasing pollutant levels in rural regions, this research highlights the need to extend air quality assessment beyond urban centers to ensure equitable and informed policy interventions.These methodologies facilitate a systematic estimation of the variation in pollution between various environments and seasons.

This research presents evidence to policy makers and environmental authorities, stressing the necessity of location-specific air quality management policies over policies that do not consider localized pollution sources and transport mechanisms. In addition, the research highlights the necessity of seasonally adaptive pollution control policies, as urban $PM_{2.5}$ concentrations are highest during winter, whereas rural ozone concentrations increase during summer.

TABLE I: Pollutant Abbreviations and Their Full Forms

Abbreviation	Full Form
$PM_{2.5}$	Particulate Matter (diameter $\leq$ 2.5µm)
PM_{10}	Particulate Matter (diameter $\leq$ 10µm)
SO_2	Sulfur Dioxide
NO_2	Nitrogen Dioxide
O_3	Ozone
CO	Carbon Monoxide

II. Literature Review

Air pollution is an important environmental and public health issue, with extensive evidence in multiple studies of its negative impacts. The *World Health Organization* (WHO) estimates that air pollution is responsible for 6.7 million premature deaths every year, with a considerable portion happening in South Asia [1]. In India, air quality has become highly critical, and cities like Delhi, Mumbai, and Kolkata suffer from hazardous levels of air often due to road transport emissions, industrial processes, and construction materials dust [2]. The 2024 progress report on the *National Clean Air Programme* (NCAP) indicates that out of 227 cities with over 75% data availability in 2023, 78 NCAP cities and 118 non-NCAP cities exceeded the NAAQS for PM_{10} levels. [3].

While much research has been devoted to urban air pollution, rural areas are relatively understudied, even though increasing evidence shows that pollution spillover, agriculture, and local emissions all lead to affecting air quality in these areas. Research indicates that rural sources of air pollution—like biomass burning, dust storms, and pesticide application—can result in $PM_{2.5}$ and ozone concentrations equal to those found in urban areas [4]. A study done by Sujan Khanal et al. highlighted transboundary pollution and crop residue burning as major factors contributing to severe seasonal pollution episodes, particularly in North India [5].

There have been various studies of urban air pollution employing varied methodologies. AQI sensitivity in Tehran was investigated by Fotouhi et al., who found that pollutant-specific statistical parameters affect reported air quality, and maximum values tend to overestimate the severity of pollution [6]. Yadav and Toshniwal did a comparative study of the air quality in four large Indian cities, applying Principal Component Analysis (PCA) and time series models to discern seasonal patterns of major pollutants [7]. Likewise, Nandanwar and Chauhan examined improvements in air quality during India's COVID-19 lockdown and recorded a substantial decline in $PM_{2.5}$ and PM_{10} concentrations due to diminished vehicle and industrial activity [8].

As a result of wide-ranging urban research, very little work exists on rural air pollution. Borghi et al. stressed the necessity of monitoring rural-specific pollutants and their emissions, mainly from plowing, harvesting, and the use of pesticides, which emit particulate matter and ammonia [4]. Mota-Bertran et al. also investigated the impact of air pollution on mental health among children in rural Catalonia, correlating pollutant exposure with behavioral disorders, anxiety, and developmental difficulties [9]. Yet, most rural studies are based on sparse monitoring networks, constraining their capability to observe fine-scale spatio-temporal patterns.

Urban-rural pollution spillover is still an unexplored topic. Mohajeri et al. estimated indoor air pollution exposures, which found large urban-rural differences in air pollution. [10]. Likewise, a *Greenpeace India* report identified that 89% of seven major Indian cities' monitoring stations had been surpassing WHO air pollution standards, highlighting vehicular emissions as a prevailing NO_2 source [11]. These studies do not include comparative urban-rural studies examining pollution gradients in terms of geographical proximity.

Advanced modeling approaches have been utilized in recent studies to analyze air pollution. Multivariate clustering has been used by Rahman and Khatun to classify Asian cities by their amounts of pollution and to identify gross differences between East and South Asia [12]. Jagtap et al. analyzed deep learning models for predicting air quality, using CNN and RNN and contrasting them with standard statistical methods [13].

Improved rural monitoring has also been emphasized in some papers. Sharma et al. traced air quality trends in 336 Indian cities from 1987 to 2019 and showed that whereas levels of SO2 reduced, PM10 concentrations went up, most notably in the Indo-Gangetic Plain [14]. In addition, a *Down to Earth* report revealed that rural $PM_{2.5}$ concentrations in India averaged 46.4 µg/m³ in 2022—almost on par with urban averages—contradicting the notion that rural environments have cleaner air [15]. However, rural environments are still underrepresented in national monitoring, networks, and are constrained by limited data availability.

Although earlier studies have widely examined air pollution in urban and, to a lesser degree, rural areas, there is a significant paucity of holistic spatio-temporal analyses comparing both environments over time. The literature does not measure urban-rural pollution differences or analyze spillover impacts. Bridging these gaps is crucial for developing *region-specific air quality management policies* that consider urban emissions and rural sources of pollution.

III. Proposed Methodology

This study aims to quantify air quality disparities between urban and rural areas in India using statistical and machine learning techniques. Based on the literature reviewed, it was observed a significant lack of comprehensive spatio-temporal analyses comparing urban and rural environments over time. While it is commonly assumed that urban areas are more polluted, rural areas can also experience significant air pollution due to sources such as biomass burning, agricultural activities, and natural dust. To systematically investigate this issue, a methodology was developed by the following key research questions:

1) How do pollutant concentrations vary between urban and rural environments across different spatial and temporal scales?

2) What dominant pollutants influence air quality trends in distinct regions?
3) How can the air quality disparities between urban and rural areas be quantitatively assessed?
4) What is the extent of urban-rural pollution disparities, and during which seasons are they most pronounced?

A. Data Collection and Study Design

This study analyzes urban–rural air quality disparities across India using pollutant and meteorological data collected from September 1, 2022, to January 31, 2025. Data was collected from the Open-Meteo API, which provides location-specific, high-resolution environmental readings through geospatial interpolation. For each location, hourly pollutant concentrations were retrieved and aggregated into daily average values. These daily values formed the basis for AQI computation in line with Central Pollution Control Board (CPCB) guidelines.

To ensure a balanced comparison, five geographically and climatically diverse city–village pairs were selected: Chennai–Koonimedu (Tamil Nadu), Delhi–Asawar (Delhi/UP border), Ahmedabad–Saputara (Gujarat), Kolkata–Bhalukhop (West Bengal), and Shillong–Mawlynnong (Meghalaya). Each rural site is characterized by a population under 10,000, offering a contrast to its urban counterpart in scale, infrastructure, and likely pollution sources.

These 10 locations were grouped into urban (cities) and rural (villages), allowing comparative analysis between the two classes during the study period [16].

TABLE II: Urban and Rural Locations Across States/UTs in India Considered for the Study

State/Union Territory	City (Urban)	Village (Rural)
Tamil Nadu	Chennai	Koonimedu
Delhi & Uttar Pradesh	Delhi	Asawar
Gujarat	Ahmedabad	Saputara
West Bengal	Kolkata	Bhalukhop
Meghalaya	Shillong	Mawlynnong

B. Preprocessing and Data Preparation

The raw dataset underwent the following preprocessing steps:

- **Unit Standardization:** CO concentrations were converted to mg/m³ per CPCB standards.
- **Outlier Detection and Removal:** Interquartile Range (IQR) and Z-score methods were used to detect sensor anomalies. IQR was applied to skewed distributions, while Z-score handled normally distributed features.
- **Missing Value Check:** The dataset was found complete.

Post-cleaning, over **212,160 hourly records** were processed and aggregated into 8,841 daily AQI values per location, ensuring high temporal resolution and analytical consistency.

C. Pollutant Selection and AQI Computation

Six air pollutants were examined : $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, and O_3. These pollutants were selected as they are the primary contributors to AQI as defined by CPCB. All data were retrieved in µg/m³; however, CO concentrations were converted to mg/m³ to align with AQI computation standards.

The daily Air Quality Index (AQI) for each location was computed using the CPCB sub-index method. This involved converting pollutant concentrations to individual sub-indices via piecewise linear interpolation between predefined breakpoint values. The final AQI per day was determined as the maximum of all sub-indices:

$$\text{AQI} = \max\left(I_{\text{PM}_{2.5}}, I_{\text{PM}_{10}}, I_{\text{NO}_2}, I_{\text{SO}_2}, I_{\text{CO}}, I_{\text{O}_3}\right)$$

D. Spatio-Temporal Trend Analysis

To analyze seasonal and long-term trends in pollution, AQI data were decomposed using seasonal-trend decomposition techniques. This separated each time series into:

- Long-term trends (multi-month/year patterns),
- Seasonal cycles (monthly or periodic variations),
- Residual anomalies (unexpected spikes).

Visualizations such as line plots and heatmaps highlighted recurring peak pollution periods and seasonal disparities across urban and rural clusters.

E. Feature Importance Analysis via Random Forests

Random Forest regression models were trained to evaluate the relative influence of each pollutant on AQI. Feature importance scores were constructed and analyzed to identify the dominant pollutants for each month and season. This provided insight into how the relevance of specific pollutants shifts over time in different environments. In particular, the analysis aimed to determine which pollutants—$PM_{2.5}$, PM_{10}, CO, SO_2,NO_2 and O_3—contribute more significantly to AQI levels in rural versus urban regions. This comparative approach helps in understanding the varying impact of pollutants across different geographical settings.

F. K-Means Clustering of Pollution of Cities and Villages

Unsupervised clustering using the K-Means algorithm was applied to average AQI values across all locations during the study period. Although the dataset includes only 10 sites, K-Means was used to identify latent groupings in pollution severity, offering an objective, data-driven alternative to threshold-based classification.

This approach allows the algorithm to naturally group locations based on similarities in pollution profiles, independent of geographic or administrative boundaries. The resulting clusters reveal high-risk and moderate-risk zones across urban and rural areas, highlighting patterns that may not align with simple AQI banding. This helps guide targeted interventions and region-specific mitigation strategies.

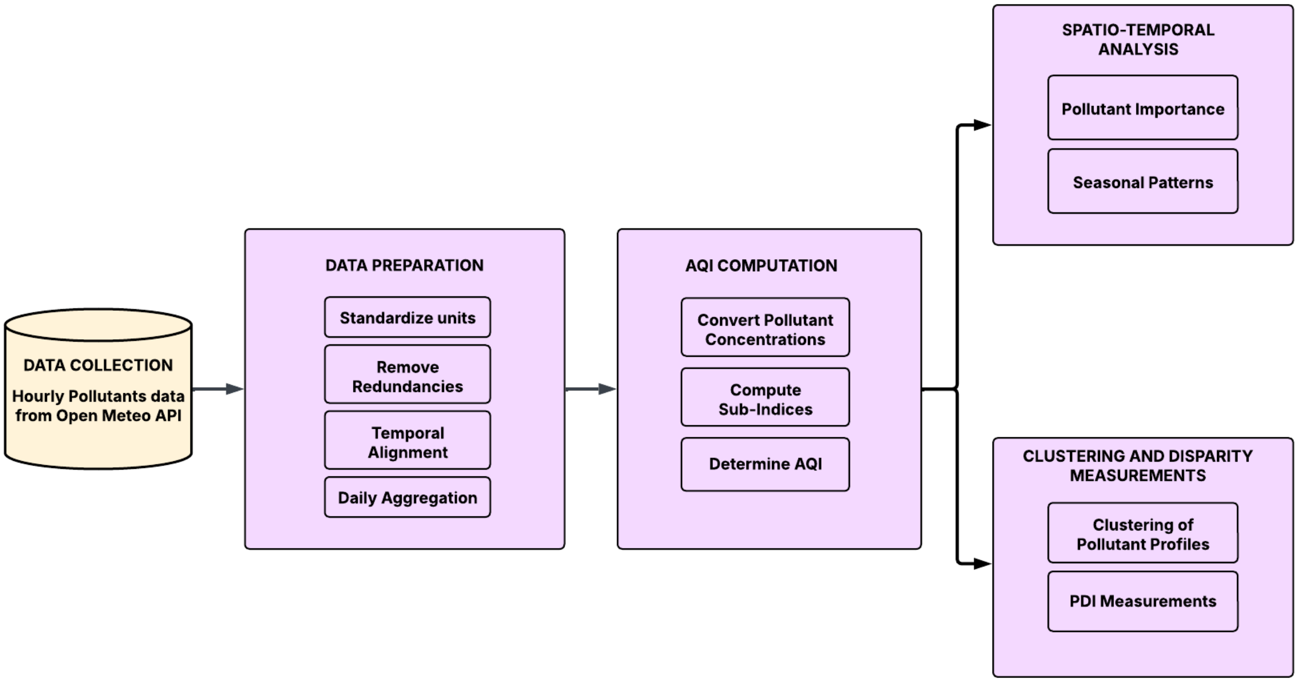

Fig. 1: Architecture Diagram

G. *Pollution Disparity Index (PDI)*

To quantify the degree of disparity between urban and rural air quality, a Pollution Disparity Index (PDI) was developed:

$$\text{PDI} = \left(\frac{\text{AQI}_{\text{urban}} - \text{AQI}_{\text{rural}}}{\text{AQI}_{\text{rural}}} \right) \times 100\%$$

Here, $\text{AQI}_{\text{urban}}$ and $\text{AQI}_{\text{rural}}$ refer to daily average AQI values aggregated across all city and village locations, respectively. A higher PDI indicates a greater urban-rural pollution gap. This metric was used to monitor disparities over time and across seasons, providing a compact yet interpretable representation of spatial inequality in air quality exposure.

Example Calculation

Assume the following daily average AQI values:

- Urban AQI: 180
- Rural AQI: 120

Using the formula:

$$\text{PDI} = \left(\frac{180 - 120}{120} \right) \times 100\% = \left(\frac{60}{120} \right) \times 100\% = 50\%$$

This indicates that the urban region has 50% higher air pollution compared to the rural counterpart on that day.

IV. Results and Key Findings

A. *Spatio-Temporal Analysis*

The Air Quality Index (AQI) data varies across locations, with Delhi (153.00), Asawar (147.23), and Calcutta (139.66) showing the highest pollution levels. In contrast, Ahmedabad (71.89), Chennai (82.12), Koonimedu (84.98), Mawlynnong (82.91), Saputara (87.04), and Shillong (82.91) have relatively lower AQI, indicating moderate air quality (Table III)

TABLE III: Location-wise AQI Values

Location	AQI
Ahmedabad	71.89
Asawar	147.23
Bhalukhop	90.45
Kolkata	139.66
Chennai	82.12
Delhi	153.00
Koonimedu	84.98
Mawlynnong	82.91
Saputara	87.04
Shillong	82.91

Urban areas, such as Delhi and Calcutta, have higher pollution due to vehicular emissions, industrial activities, and seasonal factors [17]. Meanwhile, cleaner air in regions like Shillong and Mawlynnong can be linked to lower industrialization and dense green cover [18].

The presence of coastal winds in Chennai and Koonimedu aids in pollutant dispersion, preventing excessive accumulation [19]. Additionally, Delhi's air quality worsens during winter due to temperature inversions that trap pollutants near the surface [20].

B. PDI-Disparity Measurements

Table IV compares pollution levels between city–village pairs, indicating the percentage difference in pollution (likely $PM_{2.5}$) between them. A positive PDI means the city is more polluted than its paired village, while a negative PDI shows the village has higher pollution levels. For instance, Kolkata is 35.23% more polluted than Bhalukhop, whereas Saputara shows 17.4% higher pollution than Ahmedabad. A PDI of 0% reflects equal pollution levels in both areas. These findings emphasize that pollution is not confined to urban areas alone, reinforcing the need for inclusive, region-specific air quality management strategies.

TABLE IV: Pairwise Location PDI

Location 1	**Location 2**	**PDI (%)**
Ahmedabad	Saputara	-17.40
Kolkata	Bhalukhop	35.23
Chennai	Koonimedu	-3.37
Delhi	Asawar	3.77
Shillong	Mawlynnong	0.00

C. Pollutants Importance Across Rural and Urban Areas

Figure 2 highlights the differences in pollutant contributions to AQI between urban and rural areas. Ozone (O_3) is highest in rural regions, while $PM_{2.5}$ and PM_{10} are significant in both. Other pollutants, such as CO, SO_2 and NO_2, have lower contributions. Urban areas generally have higher pollutant levels, except for ozone.

Ozone is elevated in rural areas due to reduced nitric oxide (NO) emissions, which typically break down ozone in cities. Winds transport ozone from urban sources, allowing it to collect in rural areas. Strong sunlight accelerates ozone formation, and fewer airborne pollutants mean less interference. Additionally, vegetation releases biogenic VOCs that react with transported NO_x, further increasing ozone [21], [22]. Unlike the protective ozone layer in the stratosphere, ground-level ozone is a pollutant affecting air quality. It can impact plant growth, reduce crop yields, and contribute to environmental stress in forests and ecosystems [21].

$PM_{2.5}$ and PM_{10} are crucial in air quality studies due to their widespread presence and environmental effects. $PM_{2.5}$ consists of fine particles that remain suspended in the air for long periods, affecting visibility and atmospheric composition. PM_{10}, being larger, contributes to dust and air pollution. Urban areas see higher $PM_{2.5}$ and PM_{10} levels from traffic, industry, and construction, while rural sources include agriculture, biomass burning, and dust storms. Monitoring these particles provides essential data for understanding pollution trends and developing air quality management strategies.

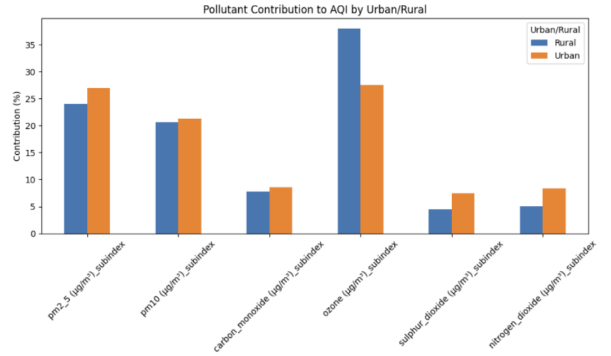

Fig. 2: Pollution Contribution to AQI for Urban VS Rural

D. K-Means Based Clustering of Rural and Urban Areas

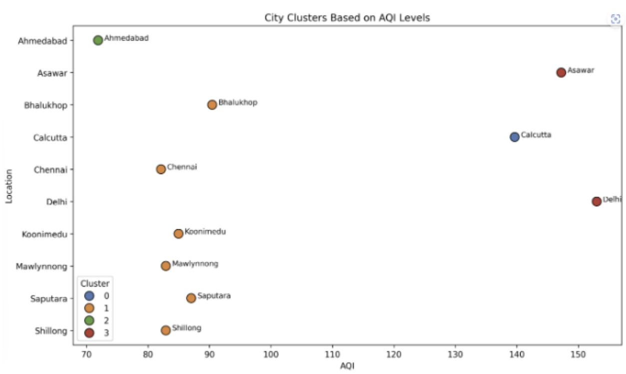

Fig. 3: K-Means categorized clusters

Figure 3 shows the spatial grouping of cities by their Air Quality Index (AQI) values with the K-Means algorithm. The chart classifies the cities into four clusters, each represented by a different color, to facilitate a comparative examination of air quality rates.

Cluster 0, which includes Calcutta, has a moderately high AQI level. Cluster 1, with cities like Bhalukhop, Chennai, Koonimedu, Mawlynnong, Saputara, and Shillong, reflects comparatively lower pollution areas. Cluster 2, which represents Ahmedabad, is the area with the lowest AQI, reflecting comparatively good air quality. Cluster 3, with very polluted cities like Delhi and Asawar, reflects critical air quality conditions.

E. Monthly AQI Trends

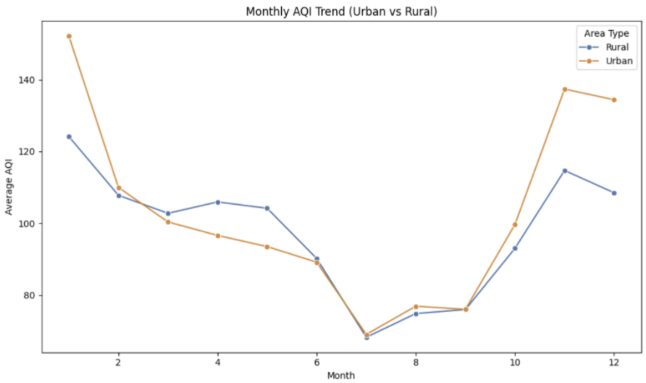

Fig. 4: Monthy AQI Trend(Urban Vs Rural)

Figure 4 presents the monthly Air Quality Index (AQI) trends for urban and rural areas, revealing significant seasonal variations in air pollution levels. Throughout the year, urban AQI remains consistently higher than rural AQI, underscoring the impact of anthropogenic emissions in densely populated regions. Both regions exhibit a declining trend in AQI from January to July, followed by a sharp increase from October to December.

The mid-year dip, most pronounced during July, coincides with the monsoon period, when enhanced rainfall serves as a natural scavenger, eliminating atmospheric pollutants via wet deposition. Strong surface winds and atmospheric turbulence also enhance pollutant dispersion, resulting in better air quality [23]. On the other hand, the steep increase in AQI during December and January can be explained by winter meteorological conditions, such as temperature inversions that accumulate pollutants at the surface [24]. Further, rising emissions due to biomass burning, domestic fuel, transportation activity, and industrial activities aggravate pollution levels. Agricultural residue burning and the depositing of particulate matter in rural areas also witness higher AQI in winter months [25].

F. Monthly Pollutant Variations Across Rural and Urban Areas

Figure 5 illustrates the monthly variation in pollutant importance in urban areas, highlighting how different pollutants influence air quality throughout the year. $PM_{2.5}$ has the highest impact during winter months, likely due to increased emissions from vehicles, industries, and lower atmospheric dispersion, while ozone (O_3) peaks in summer, driven by higher temperatures and photochemical reactions [21]. PM_{10} fluctuates, with a notable rise in late summer, possibly due to construction activities and dust. Other pollutants like carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2) remain relatively low in importance.

Compared to rural areas, urban regions show a stronger influence of PM_{10}, emphasizing the need for targeted pollution control strategies to address seasonal variations effectively.

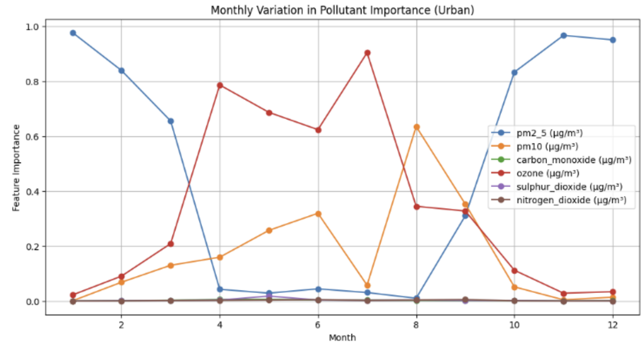

Fig. 5: Monthly Pollutant Variations - Urban

Figure 6 illustrates the monthly variation in pollutant importance for rural areas, emphasizing the influence of different pollutants on air quality throughout the year. Notably, ozone (O_3) and $PM_{2.5}$ exhibit the most significant variations, with ozone concentrations peaking from March to July, while $PM_{2.5}$ levels dominate in January and towards the end of the year.

Other air pollutants, such as PM_{10}, carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2), have relatively lower significance with minor changes over time. The predominance of ozone in warmer months indicates a high influence of photochemical reactions induced by sunlight and temperature [21], [22], while the increase in PM2.5 in colder months can be explained by enhanced biomass burning, decreased atmospheric dispersion, and stagnant weather patterns.

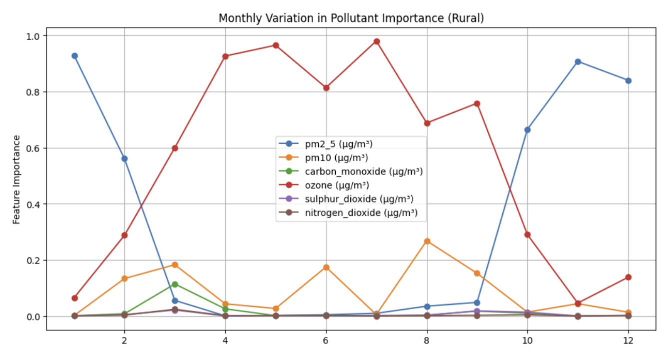

Fig. 6: Monthly Pollutant Variations - Rural

V. Conclusion

The current research presents a comprehensive analysis of air quality trends in urban and rural areas, reflecting the evolving nature of pollution patterns over time. Through spatiotemporal analysis, the study reveals clear seasonal differences in pollutant levels and highlights the complex interaction between environmental conditions and emission sources.

Results indicate that the highest concentrations of $PM_{2.5}$ in urban regions occur during winter, primarily due to transport, industrial processes, and unfavorable atmospheric conditions. Conversely, elevated ozone levels in rural areas during summer are linked to agricultural activities and photochemical reactions. The penetration of urban effluents into nearby villages and settlements demonstrates that pollution in one region can significantly impact adjacent areas.

These findings underscore the need to move beyond urban-centric mitigation strategies. Identifying pollution clusters across urban and rural settings highlights the importance of inclusive and region-specific air quality management. Corrective actions should be tailored to local characteristics, taking into account seasonal variation and spatial spillover, to ensure more effective and equitable pollution control.

VI. Future Scope

To improve spatial and temporal resolution in air quality assessment, future work should integrate satellite data sources (e.g., Sentinel-5P, MODIS) with ground-based sensor networks, especially in under-monitored rural regions. Real-time monitoring can be enhanced using low-cost IoT-based sensors connected to centralized dashboards for continuous tracking and early alerts.

Forecasting future AQI trends can be advanced by training machine learning models such as LSTM or XGBoost on combined meteorological, pollutant, and mobility data, enabling proactive air quality management.

Policy interventions like odd-even vehicle rules or industrial restrictions can be evaluated through scenario simulations using AQI and traffic datasets to guide data-driven decisions. Finally, expanding the study to more regions with diverse geographic and climatic profiles will offer deeper insights into local pollution patterns and seasonal variations.

References

[1] World Health Organization, "Ambient (outdoor) air pollution," 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-%28outdoor%29-air-quality-and-health

[2] Dr Bhola Ram Gurjar, "Air pollution in india: Major issues and challenges," 2021.

[3] Manoj Kumar N and Sunil Dahiya, "Tracing the hazy air 2024," 2024.

[4] F. Borghi, A. Spinazzè, N. De Nardis, S. Straccini, S. Rovelli, G. Fanti, D. Oxoli, A. Cattaneo, D. M. Cavallo, and M. A. Brovelli, "Studies on air pollution and air quality in rural and agricultural environments: A systematic review," *Environments*, vol. 10, no. 12, 2023. [Online]. Available: https://www.mdpi.com/2076-3298/10/12/208

[5] S. Khanal, R. Pokhrel, B. Pokharel, S. Becker, B. Giri, L. Adhikari, and M. LaPlante, "An episode of transboundary air pollution in the central himalayas during agricultural residue burning season in north india," *Atmospheric Pollution Research*, vol. 13, p. 101270, 11 2021.

[6] S. Fotouhi, M. H. Shirali-Shahreza, and A. Mohammadpour, "A sensitivity analysis on the air quality index based on the pollutants concentrations in tehran," in *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2018, pp. 333–338.

[7] A. Yadav and D. Toshniwal, "Extracting patterns and variations in air quality of four tier i cities in india," in *2017 IEEE World Congress on Services (SERVICES)*, 2017, pp. 17–20.

[8] H. Nandanwar and A. Chauhan, "Comparative study of impact of covid-19 lockdown on aqi parameters across urban india," in *2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, 2021, pp. 1–6.

[9] A. Mota-Bertran, G. Coenders, P. Plaja, M. Saez, and M. A. Barceló, "Air pollution and children's mental health in rural areas: compositional spatio-temporal model," *Scientific Reports*, vol. 14, no. 1, p. 19363, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-70024-2

[10] N. Mohajeri, S.-C. Hsu, J. Milner, J. Taylor, G. Kiesewetter, A. Gudmundsson, H. Kennard, I. Hamilton, and M. Davies, "Urban–rural disparity in global estimation of pm2·5 household air pollution and its attributable health burden," *The Lancet Planetary Health*, vol. 7, no. 8, pp. e660–e672, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S254251962300133X

[11] Greenpeace India, *NO Pollution and Health Risks in Seven Major Indian Cities*. Greenpeace, 2024.

[12] A. Rahman and M. Khatun, "Multivariate analysis of urban air pollution: Clustering and patterns across major asian cities," 11 2024, pp. 462–467.

[13] S. S. Jagtap, G. Indhumathi, R. Bharathy, S. N. Priyadharshini, T. Sajitha Sharon, and N. S. Saravana Priya, "Smart analysis of air pollutants in urban areas," in *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, 2024, pp. 60–65.

[14] G. Sharma, M. Gupta, P. Gargava, and S. H. Kota, "Mapping air quality trends across 336 cities in india: Insights from three decades of monitoring (1987–2019)," *Environment International*, vol. 191, p. 108979, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160412024005658

[15] J. Basu, "Air pollution impacts villages and cities almost equally but pollution control funds only for urban india, shows analysis," *Down to Earth*, 2023.

[16] "Dataset - preprocessed and feature engineered." [Online]. Available: https://github.com/KrishnenduMR/ProjectAir/blob/main/data/featureddata/featured_daily.csv

[17] Mausumi Basu, "The great smog of delhi," 2023.

[18] United Nations Environment Programme, "Climate impacts of air pollution," 2023.

[19] Online Desk, "Delhi's pollution crisis: Why bengaluru and chennai have 'good' air quality," *New Indian Express*, 2024.

[20] Tanushree Ganguly, "Understanding north india's pollution crisis," 2018.

[21] U.S. Environmental Protection Agency, *Ground-level Ozone Basics*. U.S. Environmental Protection Agency, 2024.

[22] IRCEL-CELINE, "Why are ozone concentrations higher in rural areas than in cities?" *IRCEL-CELINE*.

[23] Petrimed, *Monsoon and Air Pollution*. Petrimed CA, 2023.

[24] Jim Zhang, *Quick Take: Why is Air Pollution Worse in the Winter?* Global Health Institute Duke, 2023.

[25] R. Lan, S. D. Eastham, T. Liu, L. K. Norford, and S. R. H. Barrett, "Air quality impacts of crop residue burning in india and mitigation alternatives," *Nature Communications*, vol. 13, no. 1, p. 6537, 2022. [Online]. Available: https://doi.org/10.1038/s41467-022-34093-z