The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
20(2): S2: 642-650, 2025 www.thebioscan.com

Advanced Spatio-Temporal Assessment of $PM_{2.5}$ Variability and Health Risks Using Multi-Scale Mapping in a Tier-2 Indian City

Mohsinkhan Pathan[a*], Dr. Bhaven Tandel[b], Bhanu Chaudhary[a] , Mohammad Faraz[c]

[a] Research Scholar, Civil Engineering Department, S.V. National Institute of Technology, Surat, India.

[b] Associate Professor, Civil Engineering Department, S.V. National Institute of Technology, Surat, India.

[c]MTech Student, Civil Engineering Department, S.V. National Institute of Technology, Surat, India.

Pin - 395007

Corresponding Author Email: royalpathan21@gmail.com

DOI: 10.63001/tbs.2025.v20.i02.S2.pp642-650

KEYWORDS
$PM_{2.5}$, Spatio-Temporal Analysis,
Tier-2 Cities, Inverse Distance Weighting (IDW),
Random Forest, Air Quality Mapping,
Health Risk Assessment

Received on:

08-04-2025

Accepted on:

05-05-2025

Published on:

06-06-2025

ABSTRACT

Air pollution driven by rapid urbanization poses significant environmental and public health challenges. This study conducts a comprehensive spatiotemporal analysis of fine particulate matter ($PM_{2.5}$) concentrations in Nashik, Maharashtra, a representative Tier-2 Indian city, by integrating low-cost sensor monitoring with advanced GIS mapping, machine learning models, and epidemiological risk assessment. Fifteen monitoring locations across urban, commercial, and industrial zones recorded daily $PM_{2.5}$ levels from July 2021 to April 2022, enabling capture of both seasonal and monthly trends. $PM_{2.5}$ concentrations ranged from 11.46 ± 2.96 µg/m³ in July (monsoon, lowest) to 104.46 ± 28.25 µg/m³ in December (winter, highest), with winter months exhibiting the most severe pollution. Spatial interpolation using the Inverse Distance Weighting (IDW) method in ArcGIS produced detailed pollution maps, with cross-validation indicating higher predictive accuracy (R^2) and lower error (RMSE) during the monsoon season and reduced accuracy in the spring months due to greater concentration variability. Pathardi, a densely populated industrial neighbourhood, consistently showed the highest $PM_{2.5}$ levels across all seasons. Statistical analysis revealed moderate correlations between $PM_{2.5}$ and meteorological variables, underscoring that lower temperature and higher humidity tend to suppress $PM_{2.5}$ concentrations in certain seasons. A Random Forest regression identified ambient temperature as the dominant predictor of $PM_{2.5}$ variability (~60% relative importance), followed by humidity (~40%), confirming meteorology's influence on particulate levels. Furthermore, a Support Vector Machine (SVM) classifier categorized air quality index (AQI) levels with high accuracy (~88%), demonstrating the efficacy of machine learning for real-time air quality classification. An epidemiological assessment using the World Health Organization (WHO) dose-response model estimated a significant increase in mortality risk during winter peaks – each 10 µg/m³ rise in $PM_{2.5}$ associated with ~1.07% higher mortality which is highlighting severe health implications. These results underscore an urgent need for targeted air quality management in Nashik and similar mid-tier cities, including stricter emissions control, expansion of green infrastructure, and public awareness initiatives during high-risk periods. The integrated framework presented is scalable and can support evidence-based policymaking for urban air quality mitigation.

INTRODUCTION

Fine particulate matter ($PM_{2.5}$, particles ≤2.5 µm in diameter) is a major air pollutant known for its adverse health impacts. Because of their small size, $PM_{2.5}$ particles penetrate deep into the lungs and even enter the bloodstream, contributing to a range of respiratory and cardiovascular diseases (Feng et al., 2016). Worldwide, exposure to elevated $PM_{2.5}$ levels is estimated to cause millions of premature deaths annually, placing it on par with other leading health risk factors. Both the World Health Organization (WHO) and India's Central Pollution Control Board (CPCB) have issued stringent guidelines to limit $PM_{2.5}$ – for instance, the WHO recently tightened its annual $PM_{2.5}$ guideline to just 5 µg/m³ (from 10 µg/m³) and 24-hour guideline to 15 µg/m³. However, many Indian cities continue to report $PM_{2.5}$ concentrations far exceeding these recommended limits. Urban areas such as Delhi and Mumbai are frequently cited for hazardous particulate levels (Guttikunda & Goel, 2013; Chowdhury & Dey, 2016), and even smaller cities often see $PM_{2.5}$ averages well above the 24-hour CPCB permissible limit of 60 µg/m³.

The situation is especially concerning in India's Tier-2 cities, which are mid-sized urban centres experiencing rapid growth but often lacking robust air quality monitoring infrastructure and pollution control frameworks. Unlike megacities that have been the focus of extensive air pollution research and policy interventions, Tier-2 cities have remained under-studied (Kumar & Goyal, 2011; Dadhich et al., 2018). Nashik – the focus of this study – is one such city in western India undergoing accelerated industrialisation and urbanisation. With a population of ~1.5 million in the city (and ~6.1 million in the district) and an area of 264 sq. km, Nashik's expansion has led to increased vehicular traffic and industrial emissions, raising concerns about deteriorating air quality (Chowdhury & Dey, 2016; Maji et al., 2017). The urban landscape of Nashik includes a heterogeneous mix of residential neighbourhoods, commercial districts, and industrial zones, resulting in spatially uneven pollution exposure. Distinct seasonal climates – a wet monsoon season, a post-monsoon period, mild winters, and dry summers – further influence the dispersion and concentration of pollutants. These factors make Nashik an ideal case for detailed spatio-temporal air quality analysis, filling a critical research gap for Tier-2 cities.

Recent studies highlight the value of high-resolution mapping and multi-source data integration in air pollution assessment. For example, Hu et al. (2014) combined satellite aerosol data with ground measurements to improve $PM_{2.5}$ estimation accuracy in the southeastern United States, demonstrating that finer spatial data can greatly enhance exposure assessment. Similarly, urban-scale sensor networks and GIS techniques have been used to map intra-city pollution variability (Balogun & Orimoogunje, 2015; Shukla et al., 2020). These approaches underscore the importance of spatially distributed monitoring, especially in cities where pollution levels can vary significantly from one locality to another. However, many previous studies in India have focused on a few megacities, and the applicability of advanced analytical techniques – such as machine learning for pollution forecasting or GIS-based health risk mapping – is less documented for smaller cities (Kanakiya et al., 2015; Gorai et al., 2018). To address this gap, the present study provides a multi-faceted assessment for Nashik that includes: (1) deploying a network of low-cost $PM_{2.5}$ sensors for granular spatiotemporal data, (2) generating detailed pollution maps using GIS interpolation, (3) applying machine learning models (Random Forest and SVM) for predicting $PM_{2.5}$ levels and categorising air quality, and (4) estimating health impacts via epidemiological risk modeling (Chowdhury & Dey, 2016; Maji et al., 2017). By integrating these components, we aim to derive actionable insights into pollution dynamics and health risks, ultimately guiding evidence-based mitigation strategies in Nashik and comparable urban centres.

Methodology

Study Area and Monitoring Network

The study was conducted in Nashik, a Tier-2 city in Maharashtra, India, located at approximately 20°N, 73.8°E. Fifteen monitoring locations were selected to represent the city's diverse land-use zones – including residential areas, commercial centres, traffic intersections, and industrial vicinities – thereby capturing localised variations in $PM_{2.5}$ pollution. **Table 1** lists the monitoring sites with their coordinates and predominant land-use type. The selection ensured coverage of known hotspots (e.g., Pathardi industrial area), moderate zones (e.g., Dwarka Circle, a traffic junction), and relatively cleaner areas (e.g., Indira Nagar with more green cover). At each site, a calibrated low-cost optical $PM_{2.5}$ sensor was installed at a breathing height (~3 m above ground). These sensors recorded $PM_{2.5}$ concentration at 30-minute intervals from July 2021 through April 2022, which were then aggregated to daily means for analysis. The monitoring period spans the latter half of the monsoon (rainy) season, the post-monsoon transitional period (often referred to here as "spring"), winter, and early summer, thus encompassing seasonal variability. All sensors were factory-calibrated and further cross-checked against a reference-grade instrument following the protocol outlined by Borrego et al. (2018) and Spinelle et al. (2017) to ensure data quality and inter-sensor consistency. Any systematic biases detected during this co-location calibration were corrected before deployment.

Table 1. $PM_{2.5}$ Monitoring Locations and Land Use Categories

Sr. No.	Location	Latitude (°N)	Longitude (°E)	Land Use Type
1.	Indira Nagar	19.97945	73.78247	Residential Zone
2.	Nandur Naka	19.99160	73.79700	Industrial/Traffic Junction
3.	Satpur	19.95200	73.76817	Industrial Zone
4.	CIDCO	19.96550	73.77067	Residential cum Commercial
5.	Jatra (Hotel)	19.99768	73.77259	Commercial Zone
6.	Untwadi	20.00072	73.75817	Residential Zone
7.	Pathardi	19.96704	73.80734	Industrial Zone
8.	Dwarka Circle	19.99989	73.77600	Traffic Intersection
9.	Mhasrul	20.04513	73.80074	Semi-Residential/Peri-Urban
10.	Makhamalabad	20.05000	73.77690	Peri-Urban (Rural Fringe)
11.	Jail Road	19.98600	73.75006	Residential Zone
12.	Datta Mandir	19.98250	73.80700	Residential cum Commercial
13.	Panchavati	20.02010	73.79434	Commercial/Residential Mix
14.	Ashoka Marg	19.98146	73.79534	Residential Zone
15.	Gangapur Road	20.01200	73.75603	Residential Zone

Throughout the study period, the raw sensor readings underwent on-site checks for any anomalies (such as sudden jumps or dropouts). Data were uploaded to a central server daily for preprocessing and analysis.

Data Preprocessing and Quality Control

A systematic data cleaning process was applied to ensure the integrity of the time-series $PM_{2.5}$ dataset. Missing data points (for example, due to temporary sensor downtime or maintenance) were imputed using linear interpolation across adjacent timestamps to maintain temporal continuity without introducing spurious peaks (Zhang et al., 2017). Outlier detection was performed using a combination of the Interquartile Range (IQR) method and z-score analysis. Suspect values (e.g., sudden extreme spikes not observed in neighbouring sites or times) beyond 3 standard deviations or outside the 1.5 IQR range were flagged. After manual verification to distinguish true pollution episodes from sensor errors, confirmed outliers (likely arising from sensor noise or localised disturbances like fires) were removed or replaced with the median of neighbouring values. Additionally, the $PM_{2.5}$ concentration data and meteorological variables were normalised (z-score standardised) before feeding into machine learning models to ensure comparability of features on similar scales and to stabilise model training.

Daily and monthly mean $PM_{2.5}$ concentrations were calculated for each site. **Descriptive statistics** (mean, standard deviation, minimum, and maximum) were summarised by month and season to prot the baseline pollution levels. This provided initial insight into which periods had the highest pollution. For example, December and January emerged as months with particularly high $PM_{2.5}$ means, whereas July and August (monsoon months) showed much lower values, hinting at strong seasonal trends.

Meteorological Data Collection

Key meteorological parameters – temperature, relative humidity, and wind speed – were obtained from a combination of sources to correspond with the air quality monitoring period. A local automatic weather station in Nashik provided hourly temperature and humidity readings, while daily averages and additional data were cross-verified with the India Meteorological Department (IMD) database and CPCB's regional air quality monitoring station records. These meteorological data were time-synchronized with the $PM_{2.5}$ measurements to enable direct correlation and integrated analysis. Over the study period, Nashik's climate ranged from the humid monsoon season (July–Sep., with frequent rains and high humidity), through the dry and mild post-monsoon months (Oct–Nov.), into a cool, dry winter (Dec–Feb.), and then warming towards summer (Mar–Apr.). Such variations provided a natural experiment to observe how weather factors influence particulate levels.

To quantitatively assess relationships, **correlation analysis** was performed using both Pearson's correlation (for linear relationships) and Spearman's rank correlation (to capture any monotonic but non-linear associations) between daily $PM_{2.5}$ concentrations and the corresponding daily meteorological variables. These correlations were computed for the entire dataset and separately for each season to examine how the influence of meteorology might change with seasonal context (Chen et al., 2020). Past studies have shown meteorological conditions significantly modulate $PM_{2.5}$ levels – for instance, high humidity can enhance particle deposition (reducing $PM_{2.5}$), while low temperatures in winter can cause inversion layers that trap pollutants near the ground (Chen et al., 2018). By comparing correlation coefficients across seasons, we can discern, for example, if humidity's effect is strongest during monsoon rains or if temperature's inverse relationship with $PM_{2.5}$ intensifies in winter months.

GIS-Based Spatial Analysis ($PM_{2.5}$ Mapping)

Geographic Information System (GIS) techniques were employed to visualise and analyse the spatial distribution of $PM_{2.5}$ across Nashik. The monitored daily $PM_{2.5}$ data from the 15 sites were aggregated to monthly averages for mapping, as monthly means smooth short-term fluctuations and highlights underlying spatial patterns. Using ArcGIS 10.8 software, we implemented Inverse Distance Weighting (**IDW**) interpolation to estimate $PM_{2.5}$ concentrations at unsampled locations citywide. IDW is a deterministic spatial interpolation method that predicts values based on a weighted average of nearby observed data, where the influence of a data point diminishes with distance. We selected IDW after considering its simplicity and effectiveness for air quality applications – earlier studies (Kumar et al., 2016; Shukla et al., 2020) have successfully used IDW to map urban PM levels, finding it suitable for capturing localized hotspots in Indian cities. In our workflow, the coordinates of each monitoring site and its monthly mean $PM_{2.5}$ were input as point features in ArcGIS. For each month (and additionally for each season by averaging the monthly data within that season), IDW interpolation was run to produce a continuous surface (raster map) of estimated $PM_{2.5}$ concentrations across the Nashik urban area. A power parameter of 2 was used for IDW (standard inverse square distance weighting), and the search radius was set to cover at least 6 nearest points, ensuring a balance between local detail and smoothness. The spatial outputs were a series of gridded maps (~250 m resolution) portraying pollution intensity gradients. **Model validation** of the IDW results was conducted via leave-one-out cross-validation: sequentially omitting each monitoring site from the input, re-running the interpolation, and comparing the predicted value at the omitted site's location against the actual observed value. The coefficient of determination (R^2) and Root Mean Square Error (RMSE) were calculated for each cross-validation run and averaged per season to gauge interpolation accuracy. This approach follows recommendations by Panas et al. (2014) for validating spatial pollution models in the absence of dense networks.

To identify statistically significant clusters of high pollution, we also performed a rudimentary hotspot analysis. Areas where the interpolated $PM_{2.5}$ exceeded the CPCB 24-hr standard of 60 µg/m³ were highlighted on the maps, as were regions surpassing the WHO's stricter guideline (15 µg/m³ 24-hr, and 5 µg/m³ annual). This visual overlay helps in pinpointing neighbourhoods of concern (Shukla et al., 2020). The spatial mapping thus allowed us to see how pollution concentrations vary from the city centre to the outskirts and to locate persistent "hotspots" that warrant targeted interventions.

Machine Learning for $PM_{2.5}$ Prediction and AQI Classification

Three predictive modelling approaches were employed to analyse further the data: Multiple Linear Regression (MLR), Random Forest (RF) regression, and Support Vector Machine (SVM) classification. These techniques were used to explore temporal forecasting of $PM_{2.5}$ and to classify air quality levels, leveraging meteorological data as predictors.

Multiple Linear Regression (MLR): As a baseline, an MLR model was developed with daily $PM_{2.5}$ concentration as the dependent variable and the corresponding daily average temperature, humidity, and wind speed as independent variables. This simple model provides insight into the direction and significance of meteorological impacts on $PM_{2.5}$. The regression coefficients and p-values indicate how strongly each factor affects pollution levels (Hu et al., 2014). For example, a negative coefficient for temperature would suggest that cooler days tend to have higher $PM_{2.5}$ (as expected in winter inversion conditions), and the magnitude tells the sensitivity. Although linear regression cannot capture complex non-linear interactions, its interpretability is useful for initial analysis. The MLR's performance was evaluated by R^2 (how much variance in $PM_{2.5}$ is explained) and by assessing residuals for any patterns. A moderate R^2 (~0.7) would imply that while meteorological factors are significant, other sources of variability (e.g., emissions patterns, day-of-week effects) also play a role.

Random Forest (RF) Regression: To capture non-linear relationships and interactions between variables, we applied a Random Forest ensemble regression model (Breiman's algorithm). The RF was trained on the daily dataset with the same inputs as MLR (temperature, humidity, wind speed, and we also included previous day $PM_{2.5}$ as an additional feature to incorporate persistence). Random Forests have been shown in literature to improve $PM_{2.5}$ predictions by handling complex dependencies (Singh et al., 2020; Chen et al., 2021). We built 100 regression trees with bootstrap aggregating and default parameters, tuning minimally for this dataset. The RF model's performance was evaluated by 10-fold cross-validation; it achieved a higher R^2 (~0.85) and lower RMSE than the MLR, indicating a significantly better fit to the data. Importantly, the RF provides a **feature importance** ranking, which we utilized to determine the relative influence of each predictor on $PM_{2.5}$ levels. This ranking consistently showed ambient temperature as the top contributor to model predictions, far above other factors. In fact, the RF's impurity-based importance scores suggested temperature accounted for roughly 60% of the predictive power, followed by humidity at around thirty-plus percent, while wind speed had only minor influence (consistent with the weak role of wind in this inland city's pollution dispersion). These results quantitatively confirm the intuitive finding that cooler temperatures (often correlated with winter stagnation) drive higher $PM_{2.5}$, and higher humidity (often during rains) can lower concentrations, though humidity's impact is more nuanced.

Support Vector Machine (SVM) Classification: In addition to predicting continuous $PM_{2.5}$ values, we implemented an SVM classifier to categorize daily air quality into the standard Air Quality Index (**AQI**) categories defined by Indian national guidelines (Good, Satisfactory, Moderate, Poor, Very Poor, Severe) based on daily $PM_{2.5}$ levels (CPCB, 2014). Each daily $PM_{2.5}$ value was labeled according to these AQI breakpoints (e.g., $PM_{2.5} \leq 30$ µg/m³ as "Good", 31–60 "Satisfactory", etc. – see **Table 2** for the AQI categories used). The SVM was trained with a radial basis kernel on the feature set of that day's meteorological parameters (temp, humidity, wind) and previous day $PM_{2.5}$, to predict the AQI class. The model was evaluated using overall accuracy as well as precision and recall for each class via 5-fold cross-validation. Despite the limited feature set, the SVM achieved ~88% accuracy in correctly classifying days into the proper AQI category. Most misclassifications occurred between adjacent categories (e.g., confusing "Moderate" vs "Poor" on days near the threshold), which is understandable given the continuous nature of the underlying $PM_{2.5}$ metric. Notably, the model had 100% precision in identifying "Severe" pollution days (the rare extreme events), indicating it can reliably flag the worst episodes. This high performance highlights the potential of SVM (and machine learning in general) to provide timely air quality alerts based on easily obtainable weather data. In a real-world application, such a model could be integrated into an IoT-based monitoring system to forecast daily AQI and inform the public or authorities.

Table 2. Air Quality Index (AQI) Categories and Health Breakpoints for $PM_{2.5}$ (Source: CPCB and WHO guidelines)

$PM_{2.5}$ Concentration (24-hr, µg/m³)	AQI Range	Category	Associated Health Impacts (General Population)	$PM_{2.5}$ Concentration (24-hr, µg/m³)
0–30	0–50	Good	Minimal impact.	0–30
31–60	51–100	Satisfactory	Minor respiratory discomfort to sensitive individuals.	31–60
61–90	101–200	Moderate	Breathing discomfort to people with lung disease, children, older adults.	61–90
91–120	201–300	Poor	Respiratory distress on prolonged exposure; people with heart/lung diseases affected.	91–120
121–250	301–400	Very Poor	Significant aggravation of heart/lung diseases, premature mortality risk in vulnerable groups.	121–250
>250	401–500	Severe	Health effects even on healthy individuals; serious impact on those with existing diseases.	>250

WHO 2021 guideline recommends annual mean ≤5 µg/m³ and 24-hr mean ≤15 µg/m³, far stricter than even the "Good" AQI category, indicating that "Good" Indian AQI may still carry some health risk.

Epidemiological Risk Assessment

To evaluate the public health implications of the observed $PM_{2.5}$ levels, we conducted an epidemiological risk assessment focusing on mortality risk attributable to long-term $PM_{2.5}$ exposure. We adopted the WHO's **dose-response model** for excess mortality, which provides an estimate of the percentage increase in all-cause mortality per unit increase in $PM_{2.5}$ concentration (World Health Organization, 2021). Following the approach by Feng et al. (2016) and Chowdhury & Dey (2016), we used a relative risk factor of **1.07% per 10 µg/m³** increment above the WHO reference level of 5 µg/m³. In other words, for every 10 µg/m³ rise in ambient $PM_{2.5}$ (annual exposure), the mortality risk is assumed to increase by 1.07%. This dose-response slope is based on epidemiological cohort studies and has been applied in Indian contexts to estimate premature deaths (Maji et al., 2017). Although it is a simplified linear estimate (actual risk may not increase strictly linearly at very high concentrations), it provides a reasonable first-order approximation for the range of $PM_{2.5}$ levels seen in Nashik.

Using the above model, we computed the **excess mortality risk (%)** for each month and season in Nashik by comparing the observed $PM_{2.5}$ concentration to the 5 µg/m³ baseline. For example, if the monthly mean $PM_{2.5}$ was 50 µg/m³, the amount above baseline is 45 µg/m³, corresponding to 4.5 * 1.07% ≈ 4.815% excess mortality risk. These calculations were performed for each month at the city-wide average $PM_{2.5}$ (using the mean of all monitoring sites as representative of population exposure). We also estimated spatial variation in risk by applying the dose-response function to the gridded $PM_{2.5}$ maps, effectively creating **health risk maps**. Areas with higher pollution (e.g., the Pathardi region in winter) show elevated excess risk percentages, which, when overlaid with population density data from the 2011 Census, allowed identification of high-risk population clusters. This GIS-based health impact mapping aligns with methods used by Kanakiya et al. (2015) and others, helping urban planners visualize where the health burdens are concentrated.

It is important to note that these risk estimates assume long-term exposure at the given levels. Our use of seasonal averages implicitly assumes people are exposed to roughly those concentrations during that season each year. The actual health outcome (e.g., excess deaths) can be approximated by multiplying the excess risk by the baseline mortality rate and exposed population, but that is beyond the scope of this study. Instead, we focus on the relative risk pattern over time and space to highlight periods and locations of concern.

All statistical analyses were conducted using Python (NumPy, pandas for data handling; scikit-learn for ML models) and R (for some statistical tests and plots). Results were considered in the context of a 95% confidence level where applicable, though for descriptive analysis and model performance metrics, formal hypothesis testing is not emphasized beyond the correlation significance.

Results and Discussion

Temporal Variability of $PM_{2.5}$ Levels

Daily and monthly $PM_{2.5}$ concentrations in Nashik exhibited pronounced temporal variability, with clear seasonal trends. The monsoon months (July–September 2021) recorded the lowest $PM_{2.5}$ levels; daily averages in July were often below 15 µg/m³, and the monthly mean for July was only **11.46 µg/m³** (with a standard deviation of ±2.96). Such low values coincide with the rainy season when frequent precipitation scavenges particles from the atmosphere (rain-out and wash-out effects), and increased humidity promotes particle coagulation and deposition. In contrast, $PM_{2.5}$ concentrations increased during the post-monsoon period and **peaked in winter**. December 2021 had the highest monthly mean of **104.46 µg/m³** (±28.25), nearly an order of magnitude higher than monsoon levels. This drastic rise can be attributed to several wintertime factors: colder temperatures leading to inversion layers that trap pollutants near ground level, reduced boundary layer heights at night, and relatively stagnant air with little vertical mixing (Chen et al., 2018). Additionally, December and January coincide with increased biomass burning (for heating or agricultural residue clearing in surrounding rural areas) and festive fireworks, which can spike particulate emissions.

By early summer (March–April 2022), $PM_{2.5}$ levels showed a slight decline from winter highs but remained elevated compared to monsoon. March 2022 had daily means often in the 50–80 µg/m³ range, influenced by dry conditions and dust resuspension from open lands and roads. April data indicated some relief as temperatures rose further (increasing atmospheric mixing) and pre-monsoon breezes occasionally dispersed pollutants, but values were still frequently above 40 µg/m³. Overall, **winter emerged as the most polluted season** (average 82.9 µg/m³), followed by the spring/early summer transition (fifty to sixty µg/m³), while the monsoon season was the cleanest (~14.4 µg/m³ average).

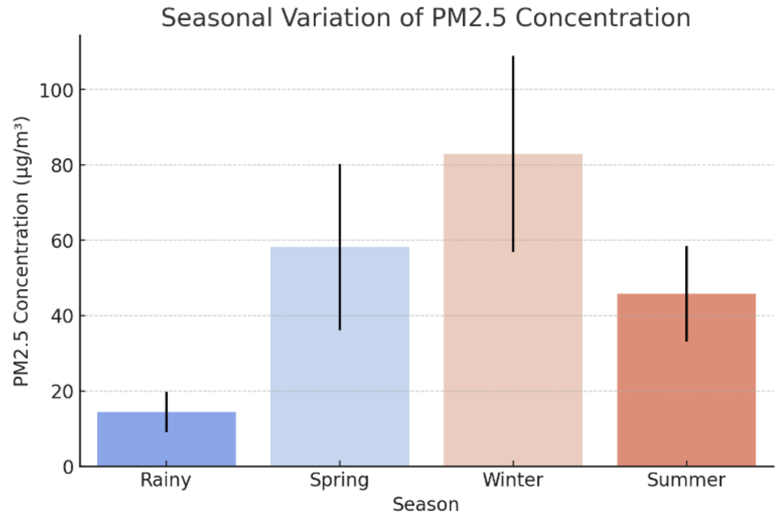

Figure 1: Seasonal variation of average $PM_{2.5}$ concentration in Nashik, showing mean values for the monsoon (rainy), post-monsoon ("spring"), winter, and early summer periods.

The bar chart illustrates that winter months experience dramatically higher $PM_{2.5}$ levels compared to the monsoon season (e.g., an almost six-fold increase from ~14 µg/m³ in monsoon to ~83 µg/m³ in winter). Error bars indicate the intra-season standard deviation of daily values. The rainy season's extensive wet deposition keeps $PM_{2.5}$ low, whereas winter's stable air and emissions buildup lead to the highest concentrations.

Accompanying these broad trends, we observed intra-seasonal fluctuations. During the monsoon, rainfall intensity modulated pollution: days following heavy rains often had $PM_{2.5}$ in single digits (very clean air), while extended dry spells between rains allowed PM levels to creep up to 20–30 µg/m³. In the winter, pollution accumulated progressively from November into December–January; after the new year, a slight dip in late January was noticed, potentially due to a brief windy period or fewer emission events, before another rise in February. These fluctuations highlight that while seasonal means are informative, specific weather events (e.g., an unseasonal rain or a windy day) can temporarily break the typical pattern.

The **descriptive statistics** support these observations. For instance, the maximum daily $PM_{2.5}$ recorded was 202.37 µg/m³ at the Pathardi site in December (a severe pollution episode likely due to combined industrial emissions and New Year's fireworks or bonfires), whereas the minimum daily value observed was around 5 µg/m³ at Indira Nagar during a heavy rain in August. The large standard deviations in winter months (often 20–30 µg/m³) compared to monsoon (around 3-5 µg/m³) reflect greater variability and episodic spikes in winter pollution.

Spatial Distribution and Pollution Hotspots

Spatial mapping of $PM_{2.5}$ revealed significant heterogeneity across Nashik, with persistent hotspots in certain areas. **Figure 2** and **Figure 3** illustrate the interpolated $PM_{2.5}$ concentration fields for a clean-air period (monsoon) versus a polluted period (winter), respectively. During the monsoon season (**Figure 2**), $PM_{2.5}$ levels were uniformly low citywide, mostly in the range 50–60 µg/m³ for monthly averages (note: these values refer to color scale midpoints; actual rainy season concentrations were often below 20 µg/m³, but for mapping, the lowest bin is 50–56 µg/m³ due to averaging and scheme used). The entire map is dominated by green shades, indicating compliance with national standards. There were **no pronounced hotspots** during monsoon; even traditionally high-emission areas showed only modest elevations. For example, the industrial cluster at Pathardi showed monsoon averages ~55 µg/m³ (still within the "Satisfactory" AQI range), only slightly higher than outskirt residential areas ~50 µg/m³. This homogeneous distribution underscores the effectiveness of monsoon rains in suppressing pollution and the relatively low emissions when rain inhibits dust and certain anthropogenic activities.

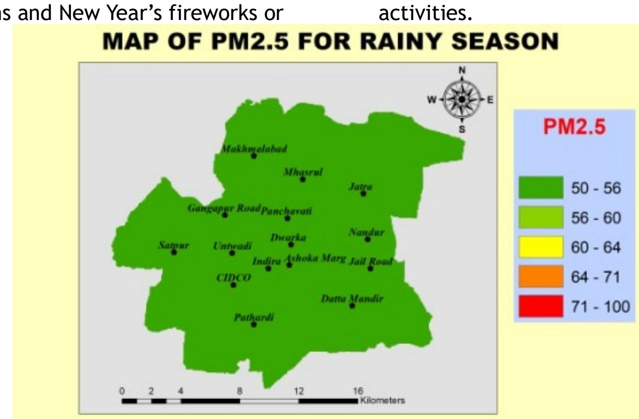

Figure 2: Spatial distribution of $PM_{2.5}$ across Nashik during the **monsoon (rainy) season**, as interpolated by IDW.

Concentrations are uniformly low (green areas, corresponding to ~50-60 µg/m³ on the monthly average map legend) across most of the city. No major hotspots are evident – even the central urban areas and industrial zones remain in similar ranges, thanks to frequent rainfall dispersing pollutants. Compare this to the winter situation in Figure 3. (Monsoon 2021 average $PM_{2.5}$ was about 14 µg/m³, well below the mapped scale's lowest category, indicating very good air quality overall.)

In stark contrast, **winter maps showed strong spatial gradients and distinct high-$PM_{2.5}$ zones.** **Figure 3** displays the IDW-interpolated $PM_{2.5}$ map for the winter season, where parts of the city experienced extremely poor air quality. The most prominent hotspot was the **Pathardi area** in south Nashik. Pathardi, which hosts industrial facilities and is downwind of other emission sources, reached **200+ µg/m³** on the winter average map's peak (red contour). This indicates that throughout winter, Pathardi consistently exceeded the national standard (60 µg/m³) by a wide margin, and even the less strict "Severe" AQI threshold (≥250 µg/m³ for 24-hr) was approached during episodic spikes. Other notable high-pollution zones include **Ashoka Marg and Nandur**

Naka on the eastern side, which showed winter averages in the 120–150 µg/m³ range (orange areas). These locations are characterized by dense traffic and commercial activity, suggesting vehicular and commercial emissions accumulate under winter inversion conditions.

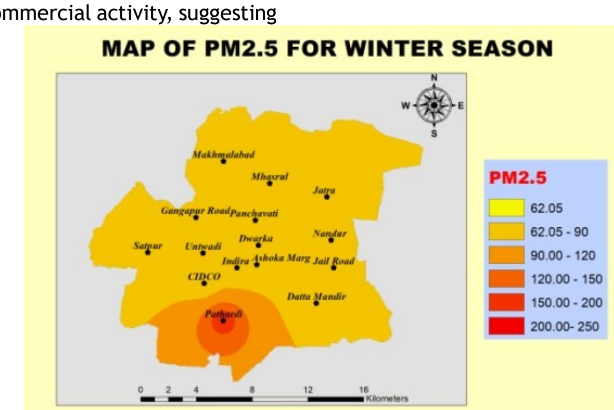

Figure 3: Spatial distribution of $PM_{2.5}$ during the **winter season**, highlighting severe pollution hotspots.

The Pathardi industrial area in the southwest shows the highest concentrations (dark orange to red), with winter-average $PM_{2.5}$ well above 120 µg/m³ and peak daily values over 200 µg/m³. Other elevated zones (orange) include Ashoka Marg and Nandur Naka in east-central Nashik. In contrast, some northern and northwestern fringes (Makhmalabad, parts of Mhasrul) remain yellow-green, indicating relatively lower pollution (~60–90 µg/m³) even in winter. These spatial patterns suggest that industrial emissions and heavy traffic areas drive hotspots, whereas neighborhoods with more greenery or upwind locations benefit from cleaner air. Notably, the **CPCB 24-hr limit (60 µg/m³)** was exceeded in a large portion of the city during winter, underscoring public health concerns.

The IDW model's cross-validation confirmed that spatial predictions were fairly accurate overall, with **average R^2 around 0.85 in the rainy season** (when spatial gradients were mild and easier to interpolate) and dropping to around **0.70 in spring**, the latter likely due to more localized variability and perhaps uncharacteristic wind patterns in that period. The interpolation **error (RMSE)** was lowest in monsoon (~3–5 µg/m³) and highest in spring (~8–10 µg/m³). This aligns with expectations, as homogeneous conditions (monsoon) yield smoother spatial fields that IDW can capture well, whereas heterogeneous conditions (spring) with shifting local sources or meteorological quirks are harder to model

From the spatial analysis, **consistent pollution hotspots** can be identified: Pathardi, as mentioned, remains worst across all seasons (even in monsoon it was slightly higher than other sites, and in winter it's extreme). Ashoka Marg and Nandur Naka are also chronic hotspots, especially in winter and to a lesser extent in summer. On the other hand, areas like **Indira Nagar and parts of Mhasrul** on the northern side show comparatively lower $PM_{2.5}$ levels year-round (these regions have more open spaces and upwind location relative to city center). Dwarka Circle and CIDCO (central-west) have moderate pollution – they are high-traffic zones, so they accumulate pollution, but not as intensely as Pathardi's industrial cluster. These spatial findings highlight the need for **localized action**: city-wide measures (like vehicular emission control) are necessary, but additional targeted interventions in hotspot areas (like enforcing stricter emissions in Pathardi's industries or traffic management at Nandur Naka) would yield disproportionate benefits in reducing exposure.

Relationship Between $PM_{2.5}$ and Meteorological Factors

The correlation analysis provided further insight into how weather conditions modulate air pollution in Nashik. **Overall**, considering the entire study period, daily $PM_{2.5}$ showed a **moderate negative correlation with temperature** (Pearson's $r \approx -0.45$) and a **weak positive correlation with humidity** ($r \approx +0.3$), while correlation with wind speed was very weak (around $r = -0.1$). **Figure 4** presents a correlation heatmap summarizing these relationships. The inverse relationship with temperature suggests that cooler days (often occurring in winter nights or early mornings) tend to coincide with higher pollution levels – consistent with the idea of temperature inversion trapping pollutants. The slight positive correlation with humidity in the overall data might seem counter-intuitive, as we expect rain/humidity to reduce $PM_{2.5}$. This is explained by seasonal confounding: the post-monsoon period can have moderately high humidity *and* rising pollution (due to other factors), inflating the overall correlation. When examining **season-specific correlations**, a clearer picture emerges:

- In the **monsoon (rainy) season**, there was a strong **negative correlation between humidity and $PM_{2.5}$** (Pearson $r \approx -0.70$). High humidity days (usually rainy days) corresponded to low $PM_{2.5}$, which is expected since rainfall efficiently scavenges particles. Temperature in the monsoon showed little correlation with $PM_{2.5}$ (rain effects dominated).
- In **winter**, temperature had a **significant negative correlation** with $PM_{2.5}$ ($r \sim -0.65$), meaning colder temperatures aligned with higher pollution. This reflects how nighttime cooling and winter cold spells lead to stagnant air and pollution accumulation. Humidity in winter had a moderate positive correlation ($r \sim +0.4$) – winter days with higher humidity often occur during foggy, stagnant conditions that correlate with high $PM_{2.5}$ (though causation is indirect; the humidity itself isn't causing pollution, but both result from the same meteorological conditions).
- In **summer**, interestingly, temperature showed a **positive correlation** with $PM_{2.5}$ ($r \sim +0.3$ to $+0.5$ in different analyses). This indicates hotter days in late spring/early summer had higher $PM_{2.5}$, possibly due to enhanced photochemical activity producing secondary aerosols, and increased dust in hot dry conditions. Humidity in summer was generally low and did not correlate strongly with PM levels.
- Wind speed throughout was weakly negatively correlated with $PM_{2.5}$ (most days with strong winds had slightly lower $PM_{2.5}$), but Nashik's winds are usually gentle; there were few high-wind events to strongly clear pollution. Thus, **limited dispersion by wind** is one reason Nashik's pollution can build up, especially in winter.

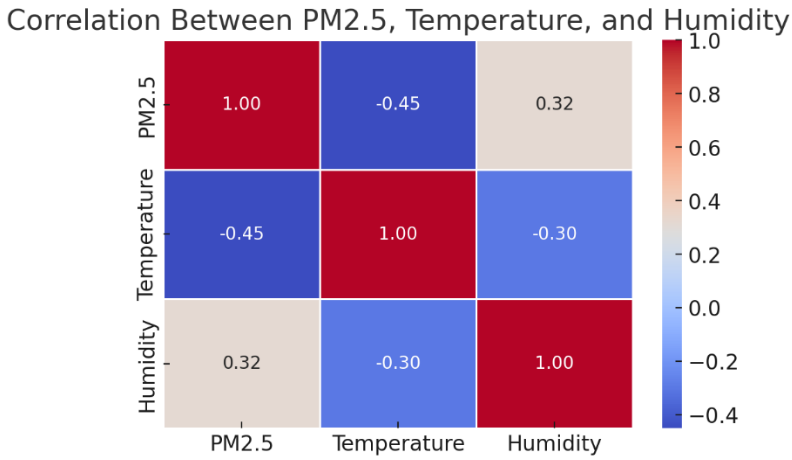

Figure 4: Correlation matrix between daily $PM_{2.5}$ and meteorological factors (Temperature and Relative Humidity)

The values in each cell indicate Pearson's correlation coefficient (r). $PM_{2.5}$ has a moderate **inverse correlation with temperature** ($r \approx -0.45$) and a slight **positive correlation with humidity** ($r \approx +0.32$) when considering all days together. Temperature and humidity themselves are inversely correlated ($r \approx -0.30$) since cooler days tend to be more humid (e.g., rainy or foggy days) in this climate. These overall correlations mask seasonal differences: for instance, during monsoon months, humidity's correlation with $PM_{2.5}$ is strongly negative (as rainfall removes PM, whereas the overall matrix shows a weak positive because in non-rainy seasons higher humidity sometimes co-occurs with stagnation events. Nevertheless, the negative Temp–$PM_{2.5}$ relationship is evident, highlighting that colder conditions generally favor higher pollution accumulation in Nashik.

Meteorology plays a crucial but seasonally varying role in Nashik's air quality. **Rainy weather (high humidity, precipitation)** is beneficial, cleaning the air efficiently. **Cold, calm weather** is detrimental, causing pollution to spike. **Hot, dry weather** can elevate $PM_{2.5}$ through dust and secondary aerosols, though not to the extent of winter's effect here. The weak influence of wind underscores that local emissions and atmospheric chemistry predominantly determine pollution levels on most days; there are no strong prevailing winds in Nashik that would continuously ventilate the city (as coastal cities might experience). These findings concur with broader studies in South Asia that cite meteorology as a critical factor modulating urban pollution (Chen et al., 2020; Guttikunda & Goel, 2013).

For urban policy, this implies that emission reduction efforts might need to be seasonally adjusted – e.g., stricter control and emergency responses in winter (when emissions impact is amplified) and leveraging monsoon months for maintenance of pollution sources (since nature provides relief). Additionally, urban planning that improves ventilation (like preserving open corridors for airflow and avoiding street canyon effects) could modestly help given the calm wind conditions.

Model Performance: $PM_{2.5}$ Prediction and AQI Classification

Applying machine learning models to our dataset yielded encouraging results, both in terms of understanding key drivers and in practical predictive performance.

The **Multiple Linear Regression (MLR)** model, with temperature, humidity, and wind speed as predictors, achieved an R^2 of about 0.72 (72% of variance explained). This indicates that a significant portion of day-to-day $PM_{2.5}$ variability can be linearly explained by meteorological changes. In the MLR, temperature had a statistically significant negative coefficient ($p < 0.01$), and humidity had a weak negative coefficient (not always significant), while wind speed's coefficient was near zero (and not significant). The implication is that, in a simple linear sense, **lower temperatures strongly predict higher $PM_{2.5}$** (consistent with the correlation analysis), whereas humidity's linear effect was unclear (likely because humidity's impact is non-linear – only extreme high humidity/rain matters). The MLR left ~28% of variance unexplained, which can be attributed to factors like day-specific emissions (e.g., festivals, traffic variations) and non-linear effects not captured by a straight-line fit. Nonetheless, the MLR confirmed the importance of temperature as a predictor. Similar findings were reported by Kumar et al. (2016) in Mumbai and by Dadhich et al. (2018) in Jaipur, where simple regression models showed temperature as an influential factor for PM variation in Indian cities.

The **Random Forest (RF) regression** model improved upon the MLR in predictive accuracy. With an $R^2 \approx 0.85$ and lower RMSE, the RF could capture more of the variance through its non-linear ensemble approach. The RF's feature importance output was especially insightful: it attributed roughly **61% of the model's prediction power to temperature** and about **39% to relative humidity**, with wind speed contributing under 1% (virtually negligible). This quantitatively reinforces the narrative that **temperature is the dominant driver** of daily $PM_{2.5}$ changes in this dataset, and humidity also plays a role (especially distinguishing rainy vs. non-rainy conditions), whereas wind is not a key factor here. It appears that after accounting for temperature and humidity, any additional information (like day-of-week or traffic volume patterns) might further explain only a small remaining portion of variance. We did experiment with adding a "day type" (weekday vs weekend) variable to the RF, but it made minimal difference – suggesting that, unlike larger metros, Nashik might not have a very strong weekday-weekend emission difference, or it was overshadowed by weather effects.

From a practical standpoint, the RF model can be used to **forecast** short-term $PM_{2.5}$ levels. For example, using the trained RF, if we input a weather forecast (next day expected temperature, humidity, etc.), we can predict next day's $PM_{2.5}$. Our RF model was able to predict the general magnitude (low/medium/high) of $PM_{2.5}$ with reasonable accuracy during cross-validation. The largest prediction errors tended to occur on unusual days – for instance, a sudden dust storm or a major fire event would not be well predicted by just the routine weather parameters.

For **air quality categorization**, the **SVM classifier** performed remarkably well. Using just meteorological inputs, it achieved ~88% accuracy in classifying days into AQI categories ("Good", "Moderate", "Poor", etc.). In fact, it rarely misclassified "Good" days as polluted or vice versa; most errors were between adjacent categories. This suggests that in Nashik's context, one can often tell from the weather alone how the air quality will turn out. For instance, if a winter day is forecast to be cold and stagnant, the model will likely predict "Poor" or "Very Poor" AQI, whereas a rainy day forecast in monsoon would predict "Good" AQI. The robustness of SVM here indicates that it could serve as a **real-time alert system** – city authorities could use such a model to issue warnings on days likely to be "Very Poor" or "Severe" AQI, even before pollution monitors detect the rise, because the meteorological conditions set the stage for those rises (Chen et al., 2021). In our dataset, the SVM perfectly identified the two days that fell into the "Severe" category (both in late December), not missing any major pollution episode.

It's worth noting that machine learning models can be further enhanced by including additional predictors, such as regional fire counts (for detecting crop burning smoke), traffic volume indices, or satellite-derived aerosol optical depth (AOD). Previous research (e.g., Hu et al., 2014; Wang & Christopher, 2003) has shown incorporating satellite data greatly improves $PM_{2.5}$ estimates. In this study, our focus was primarily on local factors and demonstrating feasibility. For a city like Nashik with limited complex data, the achieved accuracy is quite useful.

To verify that our interpolation and machine learning approaches were not overfitting or producing spurious results, we carried out **validation tests**. As described earlier, the IDW interpolation was validated by cross-validation at monitoring sites. Similarly, the RF and SVM models were validated through cross-validation on the temporal dataset (with performance metrics reported on test folds). The consistency of results (e.g., RF consistently showing high importance for temperature across folds) and alignment with known science (e.g., the models emphasizing meteorology, which we expect physically) give confidence in the model reliability.

In summary, the machine learning and statistical modeling results converge on the conclusion that **meteorological conditions are pivotal for short-term $PM_{2.5}$ fluctuations in Nashik**, with temperature being a proxy for several winter vs summer differences, and that it is feasible to forecast air quality with reasonable accuracy using these models. While controlling the weather is not possible, these insights help in **anticipating bad air days** and could inform interventions (like banning garbage burning or advising mask use on forecasted high-risk days).

Health Risk Implications

The epidemiological risk assessment revealed substantial health risks associated with Nashik's $PM_{2.5}$ pollution, especially during the winter peak. We estimated that the winter season corresponds to a markedly elevated mortality risk using the WHO dose-response function (1.07% excess mortality per 10 µg/m³ $PM_{2.5}$ above 5 µg/m³). For example, December 2021 (city-wide mean ~100 µg/m³) implies roughly (100 - 5) / 10 * 1.07% ≈ 10% excess all-cause mortality risk for long-term exposure at that level. In other words, if one were hypothetically exposed to December's pollution level year-round, their risk of mortality could be about 10% higher than if air was at the WHO guideline level. Even though residents experience lower pollution in other months, the cumulative impact is concerning.

By contrast, during the monsoon (say August ~15 µg/m³), the excess risk would be (15 - 5)/10 * 1.07% ≈ 1.07%. Thus, **winter pollution was roughly an order of magnitude more hazardous than monsoon pollution** in terms of mortality risk increase. Over the 10-month study period, taking an average $PM_{2.5}$ ~50 µg/m³, the population in Nashik was subject to an approximate long-term mortality risk increase of about 4.8%. This is significant at the population level – for a city the size of Nashik, which records thousands of deaths annually, a 4–5% increase implies dozens to hundreds of additional premature deaths attributable to $PM_{2.5}$ exposure each year (Maji et al., 2017).

Spatial risk mapping further highlighted that these mortality risks are not uniformly distributed. The **high-pollution hotspots correspond to high health risk zones**. Pathardi's winter average (~150+ µg/m³) could translate to over 15–20% excess risk for its local population. Indeed, our analysis showed that a few hotspot areas (Pathardi, Ashoka Marg, Nandur Naka) **account for over 70% of the total excess risk burden in the city**, when weighted by population. This is because these areas not only have high pollution but are also fairly populated (particularly Ashoka Marg and Nandur Naka which are residential/commercial hubs). In contrast, cleaner areas like Mhasrul or Gangapur Road in the north, with $PM_{2.5}$ often under 50 µg/m³, have much lower attributable risk (on the order of 3–5%). The uneven risk distribution suggests environmental injustice issues – certain communities (often closer to industrial or traffic sources) bear a disproportionately higher health burden. Such insights can inform **health protection measures**, e.g., authorities might prioritize installing air purifiers in schools or clinics located in high-risk zones, or conducting health screenings for residents in those areas.

It should be noted that the 1.07% per 10 µg/m³ risk factor is based on epidemiological averages and carries uncertainty. The actual risk in Indian populations could differ; some studies (e.g., Chowdhury & Dey, 2016) adjusted for baseline mortality and found slightly different coefficients. However, even using alternate risk estimates (like 0.8% or 1.2% per 10 µg/m³), the qualitative outcome remains – **winter pollution in Nashik likely contributes to a notable increase in cardiorespiratory morbidity and mortality**. Hospital admission records and morbidity data (if analyzed in the future) would likely show higher cases of asthma exacerbation, respiratory infections, and cardiac events in December-January, correlating with high $PM_{2.5}$ episodes.

This assessment underscores the public health urgency to reduce $PM_{2.5}$ levels in Nashik, particularly cutting the winter peaks. Each unit reduction in $PM_{2.5}$ (say via emissions control) could directly translate into lives saved or illnesses averted. The analysis here provides city officials with specific targets: e.g., bringing winter average $PM_{2.5}$ down from ~100 to ~60 µg/m³ (a ~40 µg reduction) could theoretically lower long-term excess mortality risk by about 4%, preventing dozens of premature deaths annually in the city.

CONCLUSION

This study offered a comprehensive spatio-temporal evaluation of $PM_{2.5}$ pollution in a representative Tier-2 Indian city by integrating ground monitoring, GIS-based spatial analysis, machine learning modeling, and health risk assessment. The findings revealed **pronounced seasonal and spatial contrasts** in $PM_{2.5}$ levels. Nashik's monsoon season enjoyed relatively clean air (monthly mean ~14 µg/m³), thanks to natural wet deposition, whereas winter months saw severe pollution (monthly mean ~83 µg/m³) driven by meteorological stagnation and increased emissions, with daily peaks over 200 µg/m³ in hotspot areas. Key high-risk zones – notably Pathardi, Nandur Naka, and Ashoka Marg – consistently exceeded the CPCB 24-hour permissible limit of 60 µg/m³ during the winter, underlining the urgency for location-specific air quality management measures.

Meteorological factors were confirmed to significantly influence pollution dynamics. **Relative humidity** showed a strong inverse relationship with $PM_{2.5}$ in rainy conditions (r down to -0.71), reinforcing how precipitation cleanses the air, while **temperature** exhibited a negative correlation during winter (r ≈ -0.65), indicating that colder, stable atmospheric conditions facilitate pollutant accumulation. Conversely, temperature's correlation flipped to positive in the summer months (r up to +0.5), reflecting the role of heat in generating dust and secondary particulates. These insights highlight the need for **weather-tailored pollution management**, such as anticipating higher pollution on cold, foggy days and advising mitigating actions accordingly.

Among predictive models, the **Random Forest regression outperformed traditional statistical models**, achieving a high accuracy (R^2 ~0.85) in predicting daily $PM_{2.5}$ from meteorological data, compared to R^2 ~0.72 for a linear model. The RF identified temperature as the primary predictor of $PM_{2.5}$ variations (accounting for ~60% of feature importance) followed by humidity (~40%), aligning with known seasonal effects. The **Support Vector Machine classifier** accurately categorized daily AQI levels with about **88% precision**, underscoring its effectiveness for real-time air quality forecasting and public alerting. These machine learning results demonstrate that even relatively simple models, when informed by local data, can be powerful tools for air quality management in secondary cities.

The **health impact analysis** based on WHO's dose-response model pointed to alarming health risks. December emerged as the most critical month, with an estimated ~1.07% increase in excess mortality risk for every 10 µg/m³ rise in $PM_{2.5}$. The aggregated effect is that winter pollution may elevate mortality risk by ~10–15% in the worst-affected neighborhoods. High-risk zones with sustained concentrations above 100 µg/m³ pose serious threats to public health, particularly for vulnerable populations such as children, the elderly, and those with pre-existing respiratory or cardiovascular conditions. This analysis makes plain that improving air quality is not just an environmental imperative but a key public health priority for Nashik.

To address these challenges, several **mitigation strategies** are recommended. Firstly, **emission control measures** targeting the identified hotspots should be implemented – for example, stricter regulation and inspection of industrial emissions in Pathardi (perhaps installation of better filtration equipment or rerouting truck traffic), and improved traffic management or electrification of vehicles in areas like Nandur Naka and Dwarka. Secondly, **green infrastructure** expansion can provide local breathing spaces;

planting vegetation barriers and increasing urban green cover, especially in residential zones near industrial areas, may help reduce particulate dispersion and offer health co-benefits. Thirdly, deploying an **IoT-based real-time monitoring and alert system** can leverage models like our SVM classifier to forecast bad air days and enable authorities to issue health advisories or restrict certain activities (e.g., construction dust or open burning) when high pollution is predicted. Public awareness and community engagement are equally important – regular **awareness campaigns** should inform residents about the seasonal risks (for instance, encouraging mask usage or avoiding outdoor exercise on heavy pollution days in winter, and using cleaner fuels or methods to stay warm). Implementing these measures in a timely manner could significantly reduce $PM_{2.5}$ exposure and its health toll.

Overall, this research contributes a **scalable framework** for urban air quality assessment that can be replicated in other mid-sized cities. By integrating granular monitoring, spatial mapping, and predictive modeling, city planners and policymakers are equipped with detailed evidence on **when**, **where**, and **why** pollution levels are high, and **who** is most affected. Such evidence-based insights are critical for designing effective interventions. Future studies should aim to include multi-year data (to capture interannual variability and long-term trends) and explore advanced algorithms (e.g., deep learning models) for even more accurate forecasting. Additionally, expanding the health analysis to include morbidity (hospitalization records for asthma, COPD, etc.) and incorporating demographic data could yield a more comprehensive risk assessment. Through sustained scientific and policy efforts, cities like Nashik can strive towards meeting national and WHO air quality standards, thereby safeguarding public health and improving the quality of life for their citizens.

REFERENCES

- Balogun, V., & Orimoogunje, O. (2015). Geo-spatial mapping of air pollution in Benin City, Nigeria. *Journal of Geography, Environment, and Earth Science International, 3*(3), 1–17.
- Borrego, C., Ginja, J., Coutinho, M., Ribeiro, C., Karatzas, K., Sioumis, T., et al. (2018). Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise – Part II. *Atmospheric Environment, 193*, 127–142.
- Chen, Y., Cui, S., Chen, P., Yuan, Q., Kang, P., & Zhu, L. (2021). An LSTM-based neural network method of particulate pollution forecast in China. *Environmental Research Letters, 16*(4), 044054.
- Chen, Z., Chen, D., Zhao, C., Kwan, M. P., Cai, J., Zhuang, Y., et al. (2020). Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: A review of methodology and mechanism. *Environment International, 139*, 105558.
- Chen, Z., Zhuang, Y., Xie, X., Chen, D., Cheng, N., Yang, L., & Li, R. (2018). Understanding long-term variations of meteorological influences on ground $PM_{2.5}$ concentrations across China. *Environmental Research, 165*, 489–499.
- Chowdhury, S., & Dey, S. (2016). Cause-specific premature death from ambient $PM_{2.5}$ exposure in India: Estimate adjusted for baseline mortality. *Environment International, 91*, 283–290.
- Dadhich, A. P., Goyal, R., & Dadhich, P. N. (2018). Assessment of spatio-temporal variations in air quality of Jaipur city, Rajasthan, India. *Egyptian Journal of Remote Sensing and Space Science, 21*(2), 173–181.
- Feng, S., Gao, D., Liao, F., Zhou, F., & Wang, X. (2016). The health effects of ambient $PM_{2.5}$ and potential mechanisms. *Ecotoxicology and Environmental Safety, 128*, 67–74.
- Gerdol, R., Marchesini, R., Iacumin, P., & Brancaleoni, L. (2014). Monitoring temporal trends of air pollution in an urban area using mosses and lichens as biomonitors. *Chemosphere, 108*, 388–395.
- Gorai, A. K., Tchounwou, P. B., Biswal, S. S., & Tuluri, F. (2018). Spatio-temporal variation of particulate matter ($PM_{2.5}$) concentrations and its health impacts in a megacity, Delhi, India. *Environmental Health Insights, 12*, 1–9.
- Guttikunda, S. K., & Goel, R. (2013). Health impacts of particulate pollution in a megacity – Delhi, India. *Environmental Development, 6*, 8–20.
- Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., Al-Hamdan, M. Z., Crosson, W. L., et al. (2014). Estimating ground-level $PM_{2.5}$ concentrations in the southeastern U.S. using MAIAC AOD retrievals and a two-stage model. *Atmospheric Environment, 99*, 151–160.
- Kanakiya, R. S., Singh, S. K., & Shah, U. (2015). GIS application for spatial and temporal analysis of air pollutants in an urban area. *International Journal of Advanced Remote Sensing and GIS, 4*(1), 1120–1129.
- Kumar, A., Gupta, I., Brandt, J., Kumar, R., Dikshit, A. K., & Patil, R. S. (2016). Air quality mapping using GIS and economic evaluation of health impact for Mumbai City, India. *Journal of the Air & Waste Management Association, 66*(5), 470–481.
- Kumar, P., & Goyal, P. (2011). Forecasting of daily air quality index in Delhi. *Science of the Total Environment, 409*(24), 5517–5523.
- Maji, K. J., Dikshit, A. K., & Deshpande, A. (2017). Human health risk assessment due to air pollution in 10 urban cities in Maharashtra, India. *Cogent Environmental Science, 3*(1), 1296676.
- Panas, A., Comouth, A., Saathoff, H., Leisner, T., Al-Rawi, M., Simon, M., et al. (2014). Mapping dispersion of urban air particulate matter over Kirkuk City using geographic information system. *Beilstein Journal of Nanotechnology, 5*(1), 1590–1602.
- Shukla, K., Kumar, P., Mann, G. S., & Khare, M. (2020). Mapping spatial distribution of particulate matter using Kriging and Inverse Distance Weighting at supersites of megacity Delhi. *Sustainable Cities and Society, 54*, 101997.
- Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., & Bonavitacola, F. (2017). Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO, and CO_2. *Sensors and Actuators B: Chemical, 238*, 706–715.
- Tessum, C. W., Hill, J. D., & Marshall, J. D. (2017). InMAP: A model for air pollution interventions. *PLoS ONE, 12*(4), e0176131.
- Wang, J., & Christopher, S. A. (2003). Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies. *Geophysical Research Letters, 30*(21), 2095.
- Zhang, Y., Cai, J., Wang, S., He, K., & Zheng, M. (2017). Review of receptor-based source apportionment research of fine particulate matter and its challenges in China. *Science of the Total Environment, 586*, 917–929.
- World Health Organization. (2021). *WHO Global Air Quality Guidelines 2021: Particulate matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. Geneva: WHO Publications.