Plant Archives Vol. 26, No. 1, 2026 pp. 1993-2002 e-ISSN:2581-6063 (online), ISSN:0972-5210

Plant Archives

Journal homepage: http://www.plantarchives.org
DOI Url : https://doi.org/10.51470/PLANTARCHIVES.2026.v26.no.1.275

ASSESSMENT OF SEASONAL AIR QUALITY INDEX (AQI) AND POLLUTION HOTSPOTS ACROSS MONITORING STATIONS USING HEAT-MAP AND PRINCIPAL COMPONENT ANALYSIS IN HYDERABAD, TELANGANA, INDIA

B. Jambamma[1], K. Naganjali[2], Vimala Beera[3], Kanaka Shankar[4]* and Syed Mazar Ali[5]

[1]Department of Agriculture Engineering, Agricultural College, Aswaraopet-507301,
Professor Jayashankar Telangana Agricultural University, Telangana, India

[2]Department of Agronomy, Agricultural College, Aswaraopet-507301,
Professor Jayashankar Telangana Agricultural University, Telangana, India

[3]Department of Food and Industrial Microbiology, Regional Agricultural Research Station, Anakapalle-531001,
Acharya N.G. Ranga Agricultural University, Andhra Pradesh, India

[4]Agricultural Economist, TN-IAMP, MDPU, Chennai-600005, Tamil Nadu, India

[5]Agriculture Research Station, UAS, Pavagada-561202, Bangalore, Karnataka, India

*Corresponding author E-mail: kanaka.janu@rediffmail.com

(Date of Receiving : 24-01-2026; Date of Revision : 20-03-2026; Date of Acceptance : 02-04-2026)

ABSTRACT

This study investigates the spatiotemporal variations and drivers of air quality across 13 monitoring stations in Hyderabad, Telangana, throughout 2024. Utilizing a dataset of 4,758 observations from the Central Pollution Control Board (CPCB), the research employed two-way ANOVA, Tukey's HSD post-hoc tests, and Principal Component Analysis (PCA) to characterize pollution dynamics. ANOVA results revealed that seasonal meteorological changes were the primary drivers of Air Quality Index (AQI) fluctuations ($F = 442.99$, $p < 0.001$), with a significant station-season interaction confirming that seasonal impacts are site-specific. Winter emerged as the most polluted season, with Zoo Park (129.05) and Sanath Nagar (126.07) identified as major hotspots. In contrast, the monsoon season exhibited the best air quality due to the washout effect, particularly at Central University (AQI 38.22). PCA identified three principal components explaining 60.59% of the total variance, with PC1 highlighting nitrogenous pollutants (NO, NO_2, NH_3) from vehicular and industrial emissions as dominant factors. Correlation analysis further established strong links between nitrogen oxides, while wind speed showed a significant inverse relationship with pollutant stagnation. Although most pollutants remained within NAAQS limits, $PM_{2.5}$ levels (38.32 µg/m^3) are approaching the annual safety threshold, necessitating targeted urban emission management.

Keywords: Air Quality Index (AQI); Hyderabad; $PM_{2.5}$; PCA; Seasonal Variation; Pollution Hotspots; Two-Way ANOVA

Introduction

Breathing clean air is crucial for human health and overall well-being. The World Health Organization emphasizes that having access to air free from pollution is vital for maintaining life and a good quality of living. Alongside land and water, air is an essential resource that supports breathing, agriculture, and ecological stability. However, the rapid pace of urbanization and economic development has led to increased air pollution worldwide, resulting in environmental harm and notable declines in productivity and economic efficiency (Büke & Köne, 2016; Vaddiraju, 2019).

In India, air pollution is a significant public health issue, ranking as the third leading cause of death and reducing life expectancy by approximately 2.6 years (Reddy *et al.*, 2020). The problem has been exacerbated by rapid urbanization, which has grown at an annual rate of 2.30% from 1990 to 2013 (Büke & Köne, 2016). Despite some progress through fuel substitution, air quality in many urban and industrial regions continues to worsen, with pollutants like SO_2 and particulate matter (PM) frequently surpassing acceptable limits and posing serious health threats (Büke & Köne, 2016). Air pollution consists of a complex mix of gaseous and particulate pollutants that have severe health consequences. Sulfur dioxide (SO_2) can cause respiratory issues and worsen cardiovascular conditions, while nitrogen dioxide (NO_2) can irritate the lungs and increase vulnerability to infections. Ozone (O_3) can impair lung function and elevate the risk of asthma and lung cancer, and carbon monoxide (CO) can interfere with oxygen transport, reducing work capacity. Particulate matter ($PM_{2.5}$ and PM_{10}), particularly fine particles, can lead to diminished lung function, COPD, and a shorter lifespan. Among these, $PM_{2.5}$ is the most dangerous due to its ability to penetrate deeply into the lungs and bloodstream. Even minimal exposure can have severe health effects, especially for vulnerable groups (Chabalala *et al*., 2026). It is linked to carcinogenic effects and nearly 7 million deaths each year (Van der Walt *et al*., 2023; Xing *et al*., 2016). $PM_{2.5}$ originates from sources such as fossil fuel combustion, vehicle emissions, industrial activities, and natural events, and can be emitted directly or formed through atmospheric chemical reactions.

To effectively convey air quality conditions, the Air Quality Index (AQI), developed by the United States Environmental Protection Agency, is widely used as a standardized tool. It combines multiple pollutant concentrations-including CO, O_3, SO_2, NO_2, PM_{10}, and $PM_{2.5}$ into a single index, making it easier to interpret air quality status and related health risks (Reddy *et al*., 2020; Chauhan and Joshi, 2008).

Hyderabad, one of India's rapidly growing metropolitan areas, illustrates the escalating air pollution crisis. The city is facing severe air quality deterioration due to human activities such as industrialization, vehicle emissions, sewages, biomass, construction, and energy production (Vaddiraju, 2019). The reduction in green spaces has further worsened pollution, as vegetation naturally absorbs pollutants. Similar patterns are seen in cities like Delhi, where the AQI often falls into the "very poor" category due to high $PM_{2.5}$ levels (Jay *et al*., 2017). Long-term studies also show changing seasonal pollution patterns, reflecting shifts in emission sources and atmospheric processes (Mohan & Kandya, 2007).

Nonetheless, the Air Quality Index (AQI) and its related parameters show considerable variation across different locations and times, posing challenges for precise evaluation and forecasting. Due to the intricate and variable nature of air quality data, conventional deterministic methods often fall short in providing accurate predictions and analyses. As a result, there is increasing use of data-driven techniques, especially machine learning models such as Support Vector Machines (SVM), Random Forest, AdaBoost, and linear regression. Ensemble methods like stacking and AdaBoost offer improved predictions, while advanced algorithms such as XGBoost achieve high accuracy, with correlation coefficients exceeding 0.95 (Vignesh *et al*., 2023). In this scenario, combining statistical and machine learning methods provides a thorough framework for comprehending air quality dynamics. Principal Component Analysis (PCA) is used to identify variation patterns and relationships among genotypes and traits. It transforms correlated variables into a reduced set of uncorrelated principal components that retain most of the data variability. Each component represents a source of variation, and eigenvalues indicate the variance explained (Acharyya *et al*., 2025; Parsaniya *et al*., 2025).

This study employs a sequential approach using an unsupervised machine learning workflow, two-way ANOVA, and Principal Component Analysis (PCA) to pinpoint significant influencing factors and to visualize intricate structural patterns within high-dimensional environmental datasets.

Materials and Methods

Details of the Study Area

The present study was conducted in Hyderabad, a 400-year-old metropolitan city and the capital of Telangana, India, with a population of approximately 1.16 crore. The city is geographically located at 17°23′13″ N latitude and 78°29′30″ E longitude. A total of 13 Air Quality Index (AQI) monitoring stations distributed across the city were selected for this study to represent diverse AQI characteristics, across residential, industrial, and traffic-dominated regions (Figs. 1 and 2; Table 1).

Fig. 1: Study Area Map of Hyderabad Showing City Boundaries

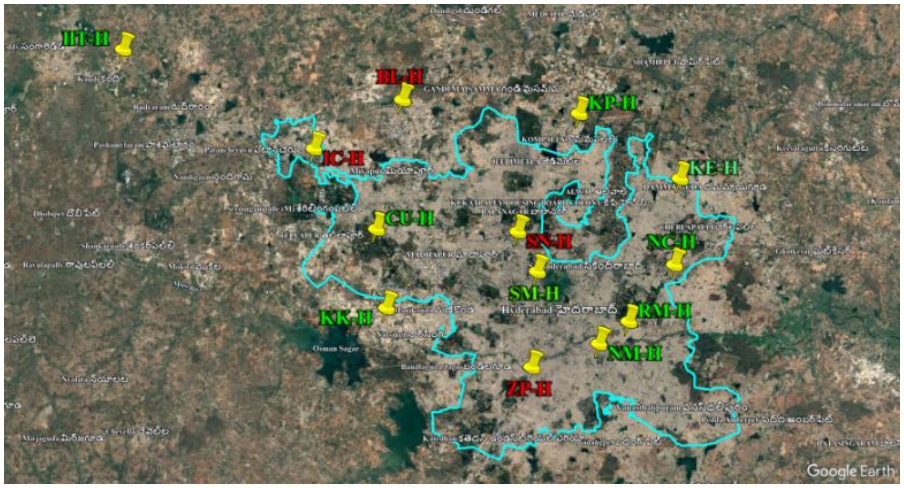


Fig. 2: Google Satellite Map of 13 AQI Stations (2024) in Hyderabad Showing Pinned Locations

Table 1: Hyderabad Monitoring Stations with Their Coordinates

Station name	Code	Longitude	Latitudes	Site Remarks
ICRISAT-Hyderabad	IC-H	17.51854492307882	78.27736274794941	Located in the Patancheru industrial belt; unique mix of heavy industry and vast semi-arid research lands.
New Malakpet-Hyderabad	NM-H	17.369442866080558	78.50583908147296	High-density residential area; proximity to the Government Printing Press and key power infrastructure.
Zoo Park-Hyderabad	ZP-H	17.351574957982844	78.45175686524067	Significant green cover (Nehru Zoo Park) and large water body (Mir Alam Tank); sensitive receptor zone.
Central University-Hyderabad	CU-H	17.457120362383616	78.3273560687526	HCU Campus; characterized by low-density built-up area, high vegetation, and internal seasonal lakes.
IIT-Hyderabad	IIT-H	17.594559506316813	78.12383402825915	Large open campus lands; situated along the high-traffic Nanded-Mumbai Highway (NH-65).
Nacharam-Hyderabad	NC-H	17.429894165763017	78.5661970027416	Heavy Industrial Development Area (IDA) presence; complex mix of residential pockets and local lakes.
Kompalli-Hyderabad	KP-H	17.54546623243523	78.48972056201909	Rapidly developing elite residential hub; significant traffic transit point with several suburban lakes.
Ramanthapuram-Hyderabad	RM-H	17.386456438934136	78.52965704754332	High-traffic residential zone; proximity to the Musi River and the city's major Wastewater Treatment Plant (STP).
Sanath Nagar -Hyderabad	SN-H	17.455214284077744	78.44056206841256	Industrial hub (chemical, pharmacy and manufacturing); high heavy-vehicle traffic and dense worker housing.
Somajiguda-Hyderabad	SM-H	17.424675995482545	78.45617592680028	Central Business District; characterized by high-rise office spaces, intense traffic, and proximity to Hussain Sagar.
Bollaram-Hyderabad	BL-H	17.55599625942321	78.34814170644397	Active Industrial Development Area (IDA); primarily manufacturing units interspersed with worker colonies.
Kapra ECIL-Hyderabad	KE-H	17.49581338340246	78.57061393787676	Dominated by the ECIL electronics complex; high-density residential area with significant peak-hour traffic.
Kokapet-Hyderabad	KK-H	17.396786916035104	78.33532784695673	"Neopolis" area; modern IT/Commercial high-rises, luxury residential developments, and multiple catchment lakes.

Data Sources and Measurements/Variables

Ambient air quality data were obtained from the Central Pollution Control Board (CPCB) under the National Air Monitoring Programme (NAMP), which monitors air quality across multiple locations using standardized procedures (CPCB, 2014).

Daily air quality data (AQI) for the year 2024 (366 days) were obtained for 13 monitoring stations. The dataset was categorized into seasons based on monthly distribution, with winter spanning November to February, summer from March to June, and monsoon from July to October. It included major criteria pollutants such as particulate matter ($PM_{2.5}$), nitrogen monoxide (NO), nitrogen dioxide (NO_2), ammonia (NH_3), sulfur dioxide (SO_2), carbon monoxide (CO), and ozone (O_3). The National Ambient Air Quality Standards (NAAQS) prescribed by the CPCB were used as reference limits (Table 2 & 3).

Table 2: Various Category of IND-AQI (National Air Quality Index, CPCB, 2014)

Sl. No.	AQI	Remark
1	0-50	Good
2	51-100	Satisfactory
3	101-200	Moderate
4	201-300	Poor
5	301-400	Very Poor
6	401-500	Severe

Source: https://cpcb.nic.in/openpdffile.php

Table 3: National ambient air quality standards in as per Environment (Protection) Act, 1986 in India

Meteo-rology	Pollutants units	Concentration of ambient air (in µg/m³)		
		Time Weighted Average	Industrial, Residential, Rural and other Areas	Ecologically Sensitive Area (notified by Central Government)
Temperature	°C	-	-	-
RH	%	-	-	-
WS	m/s	-	-	-
$PM_{2.5}$	µg/m³	Annual	40	40
NO	µg/m³	Annual	-	-
NO_2	µg/m³	Annual	40	30
NH_3	µg/m³	Annual	100	100
SO_2	µg/m³	Annual	50	20
CO	µg/m³	Annual	2000*	2000*
Ozone (O_3)	µg/m³	8 hours	100	100
AQI	-	-	-	-

Source: (Mohan & Kandya, 2007) RH=Relative humidity, WS=wind speed, *Converted mg/m3 to µg/m3

Statistical Analysis

This study utilized a sequential analytical approach integrating unsupervised machine learning and inferential statistics, processed *via* OriginPro 2026 (64-bit) SR1 (Fig. 3). Initially, a two-way ANOVA with Tukey's HSD post-hoc analysis was performed on AQI data from 13 stations (N= 4,758) to assess spatial and seasonal variability across Hyderabad for the year 2024. Following the identification of the four most severely impacted stations, a Principal Component Analysis **(**PCA**)** was conducted (N = 1,464) to evaluate the interactions between seven criteria pollutants ($PM_{2.5}$, NO, NO_2, NH_3, SO_2, CO, and O_3) and three climatic factors (AT, RH, and WS). This multivariate approach allowed for the identification of dominant environmental drivers and the visualization of high-dimensional pollution patterns across three distinct seasons.

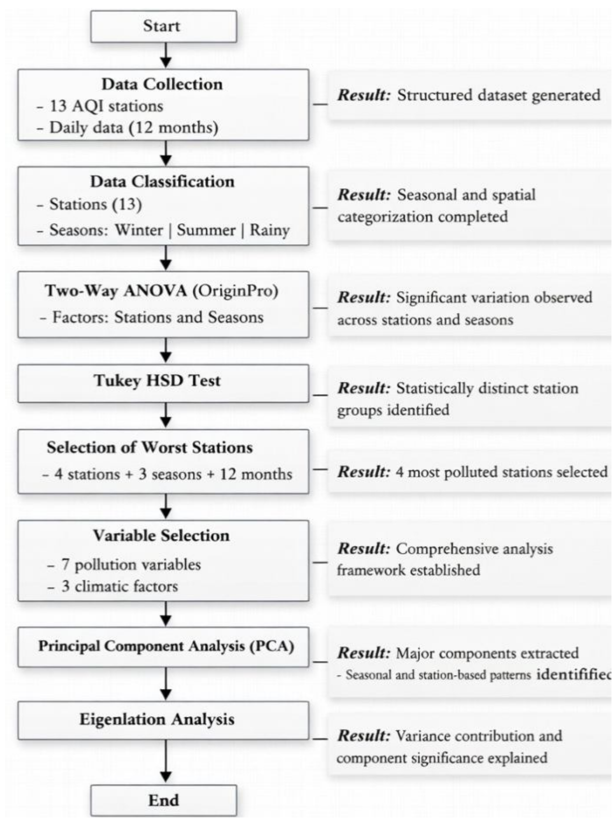


Fig. 3: Methodological Flowchart for AQI Analysis in Hyderabad

Results and Discussion

The study utilized daily air quality data from 13 AQI monitoring stations across three seasons in 2024, comprising a total of 4,758 observations obtained from the Central Pollution Control Board (CPCB) for Hyderabad, the capital of Telangana. Some missing daily AQI values in the dataset were imputed using the monthly median method (<10%) to minimize the influence of extreme values. The dataset was maintained at a daily scale and segregated seasonally into winter, summer, and monsoon periods for analysis. Station-wise and seasonal variations were examined using two-way ANOVA, and spatiotemporal patterns were visualized through heat map analysis. The results for different air pollutants across the selected monitoring locations are presented below.

Two-Way ANOVA of AQI Variation

The ANOVA results (Table 4) indicate that both main effects and their interaction were highly significant ($p < 0.001$). The season factor exhibited the largest effect (F = 442.99), suggesting that seasonal meteorological changes are the primary drivers of AQI fluctuations in the region. The significant Station effect (F = 35.67) confirms substantial spatial heterogeneity in air pollution across the 13 monitoring sites. Most importantly, the significant station x season interaction (F = 35.99, $p < 0.001$) demonstrates that seasonal impacts on air quality are station-specific.

Table 4: Two-way ANOVA for AQI across 13 stations and 3 seasons in Hyderabad

Variable	**DF**	**Sum of squares**	**Mean square**	**F-value**	***p*-value**
Stations	12	300345.99597	25028.833	35.67284	<0.0001
Seasons	2	621620.01443	310810.00721	442.98808	<0.0001
Interaction	24	606063.45277	25252.64387	35.99183	<0.0001
Error	4719	3310952.3798	701.62161		

At the 0.05 level, the population means of Stations, Seasons & interaction are sign.diff.

Difference (HSD) post-hoc test was performed to enable pairwise comparisons and classify the data into homogeneous significance groups. This approach clearly identified the specific stations and seasons contributing most to air pollution variability. The post-hoc results (Table 5) indicate that the highest AQI levels occurred during winter, with Zoo Park (129.05) and Sanath Nagar (126.07) forming Group ‘a’ ($p < 0.05$), showing significantly higher pollution than all other station–season combinations. The next most affected group included ICRISAT (110.54) and Bollaram (105.18), categorized under Group ‘b’ ($p < 0.05$), indicating significantly elevated AQI but lower than Group ‘a’. In contrast, the lowest AQI was recorded at Central University during the rainy season (38.22), which was placed in Group ‘o’, representing the least polluted condition among all observations.

Table 5: Mean AQI across different stations and seasons in Hyderabad with Tukey HSD

Monitoring Station	**Winter**	**Summer**	**Rainy**
ZP-H (Zoo Park)	129.05^{a}	84.77^{cd}	52.15^{mn}
SN-H (Sanath Nagar)	126.07^{a}	80.79^{cdef}	57.02^{lmn}
IC-H (ICRISAT)	110.54^{b}	$70.22^{efghijk}$	48.32^{no}
BL-H (Bollaram)	105.18^{b}	82.28^{cde}	58.71^{klmn}
KP-H (Kokapet)	87.50^{c}	81.17^{cdef}	$75.27^{cdefghi}$
CU-H (Central University)	83.51^{cd}	64.93^{hijklm}	38.22^{o}
NC-H (Nacharam)	81.02^{cdef}	$75.83^{cdefghi}$	$69.61^{efghijkl}$
RM-H (Ramanthapuram)	77.02^{cdefgh}	79.15^{cdef}	78.17^{cdefg}
SM-H (Somajiguda)	78.32^{cdefg}	$75.68^{cdefghi}$	$68.10^{fghijkl}$
KE-H (Kapra ECIL)	78.23^{cdefg}	61.43^{jklmn}	$69.20^{efghijkl}$
NM-H (New Malakpet)	$69.25^{efghijkl}$	77.66^{cdefgh}	$69.46^{efghijkl}$
IIT-H (IIT)	$73.60^{defghij}$	64.82^{hijklm}	58.65^{klmn}
KK-H (Kompalli)	$69.85^{efghijkl}$	65.40^{ghijkl}	63.20^{ijklm}

Note: Means followed by the same superscript letter are not significantly different (p > 0.05) according to Tukey’s HSD test. Values are rounded to two decimal places.

Spatiotemporal Visualization of AQI (Heat Map Analysis)

The heat map (Fig. 4) provides a comprehensive visualization of the AQI distribution across 13 monitoring stations during three distinct seasons. The color gradient, ranging from deep blue (AQI ≈ 38) to vibrant red (AQI ≈ 129.5), illustrates the following key findings.

Temporal analysis of the AQI data revealed a consistent seasonal trend in the order of Winter >Summer >Rainy, representing a transition from peak pollution levels to significantly improved air quality.

During winter, air quality at these hotspots falls in the “Moderate” category. According to CPCB-guidelines (Table 2), this level is not satisfactory and may cause breathing discomfort, especially in people with asthma, respiratory, or cardiovascular conditions. Two primary hotspots, Zoo Park (ZP-H) and Sanath Nagar (SN-H) emerged with peak mean AQI values exceeding 125, placing them in the highest statistical tier (Tukey Group 'a' Table 5). These were closely followed by ICRISAT (IC-H) and Bollaram (BL-H), which also surpassed the 100-AQI threshold (Group 'b'). Elevated winter pollution is driven by temperature inversion, where cold air traps pollutants near the surface and limits dispersion. Combined with low wind speeds in Hyderabad, this leads to pollutant buildup and poor air quality.

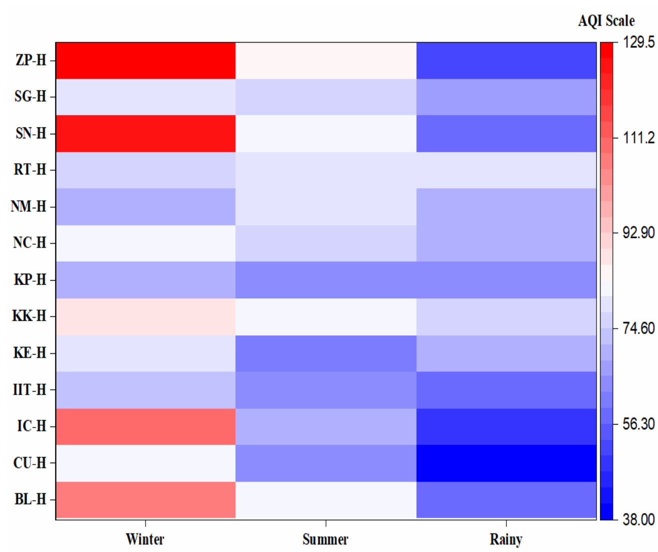


Fig. 4 : Seasonal distribution of mean AQI values across different monitoring sites in Hyderabad.

Most summer AQI values (≈61–85) fall within the “Satisfactory” category, showing clear improvement compared to winter. However, as per CPCB-guidelines, this range may still cause minor breathing discomfort to sensitive individuals. The persistence of this yellow-coded “Satisfactory” zone is attributed to dust re-suspension and increased formation of secondary pollutants such as ground-level ozone. Notably, stations like Zoo Park (84.77), Bollaram (82.28) and Kokapet (81.17) maintain relatively higher AQI levels even in summer as indicated in Table 5, indicating the influence of consistent local emission sources, such as industrial activities, that are less affected by seasonal variation.

During the rainy season, only Central University (38.22) and ICRISAT (48.32) achieved a “Good” AQI rating. Other stations (≈52–78) remained in the “Satisfactory” range despite monsoon rains. The dominance of blue shades confirms the washout effect, with the lowest AQI at Central University.

Most stations shift from red/pink in winter to blue in the rainy season, while RM-H (Ramanthapuram), SM-H (Somajiguda), NM-H (New Malakpet) and KK-H (Kompalli) remain in the light-blue/white range (79.15-63.20) across seasons, indicating minimal variation.

Principal Component Analysis (PCA)

A PCA was performed on 7 pollutants and 3 climatic factors (N=1464) across 4 major air polluted stations, as per AQI, in Hyderabad, Telangana state. The next sections explain how to find important pollutants using statistical methods. They focus on Eigenvalues, Variance, and Factor Loadings. To study how variables relate to each other and change over time, Correlation analysis and Seasonal-Station Biplots are used. A Pie chart shows the yearly distribution of pollutant mass in Hyderabad, as explained below.

Eigenvalue, Variance and Factor Loadings

The PCA of the correlation matrix showed that the first three principal components (PCs) have eigenvalues over 1. Together, they explain 60.59% of the total variance in air quality data (Table 6).

The first component (PC1) explains 29.28% of the variance (Eigenvalue: 2.92). As indicated in Table 7, PC1 is primarily affected by nitrogen pollutants, with notable values for NO (0.498), NO_2 (0.496), and NH_3 (0.478), representing the impact of vehicular and industrial emissions.

The second component (PC2) explains 17.03% of the variance (Eigenvalue: 1.70). It has a strong positive link with SO_2 (0.581) and Ozone (0.379), and a strong negative link with Wind Speed (WS) (-0.606). This indicates that lower wind speeds lead to the stagnation and accumulation of secondary pollutants and industrial markers.

Table 6: Eigenvalue Analysis and Variance

Sl. No.	Eigen Value	Percentage of Variance	Cumulative Variance
1	2.92814	29.28%	29.28%
2	1.70292	17.03%	46.31%
3	1.42834	14.28%	60.59%
4	0.92041	9.20%	69.80%
5	0.83231	8.32%	78.12%
6	0.71897	7.19%	85.31%
7	0.60512	6.05%	91.36%
8	0.49867	4.99%	96.35%
9	0.22785	2.28%	98.63%
10	0.13726	1.37%	100.00%

The third component (PC3) explains 14.28% of the variance, making the total over 60%. Even though PC3 is not shown in the 2D Biplot, it helps explain more about secondary pollutants.

The factor loadings (Table 7) show which variables affect air quality the most. PC1 is mainly affected by high values for NO, NO_2, and NH_3, showing the main impact from city combustion sources. PC2 is marked by a strong negative value for Wind Speed (-0.606) and a positive value for SO_2 (0.581), illustrating the significant impact of meteorological stagnation on pollutant concentrations.

Table 7: Factor Loadings for the first two Principal Components

Variables Pollutants (μg/m^3)	PC1	PC2	Dominant Source/Factor
$PM_{2.5}$	0.203	-0.079	Mixed Urban Aerosols
NO	**0.498**	-0.202	Vehicular Combustion
NO_2	**0.496**	-0.189	Traffic Emissions
NH_3	**0.478**	0.005	Industrial/Agricultural
SO_2	0.105	**0.581**	Industrial Stagnation
CO	0.233	0.206	Incomplete Combustion
Ozone (O_3)	0.29	0.379	Photochemical Smog
Climatic Factors			
Ambient Temp (AT)	-0.228	0.151	Thermal Influence
Rel. Humidity (RH)	-0.177	-0.037	Moisture Content
Wind Speed (WS)	-0.027	-0.**606**	Pollutant Dispersion
Eigenvalue	**2.92**	**1.7**	
Variance Explained (%)	**29.28**	**17.03**	
Cumulative (%)	**29.28**	**46.31**	

Correlation Analysis of Pollutants and Meteorological Factors

The correlation matrix identifies significant positive relationships between NO and NO_2 ($r = 0.831$), and NO and NH_3 ($r = 0.735$), reinforcing the idea of a shared emission source, likely vehicular combustion. Interestingly, climatic factors showed an inverse relationship with pollutants. Ambient temperature (AT) and relative humidity (RH) were negatively correlated with $PM_{2.5}$ (r = -0.19 and -0.24 respectively), suggesting that higher humidity or temperatures may play a role in pollutant dispersion or deposition in the Hyderabad region

Table 8 : Inter-variable correlation matrix for multivariate statistical analysis.

Variable	$PM_{2.5}$	NO	NO_2	NH_3	SO_2	CO	O_3	AT	RH	WS
$PM_{2.5}$	1	0.1282	0.1847	0.1807	-0.0047	0.2485	0.0186	-0.1980	-0.2428	0.0514
NO	0.1282	1	**0.8317**	**0.7353**	-0.0365	0.1114	0.2955	-0.2167	-0.0875	0.1038
NO_2	0.1847	**0.8317**	1	**0.6672**	-0.0337	0.2002	0.2601	-0.1919	-0.1453	0.0838
NH_3	0.1807	**0.7353**	**0.6672**	1	0.3051	0.1727	0.2414	-0.1593	-0.0476	-0.0189
SO_2	-0.0047	-0.0365	-0.0337	0.3051	1	0.1087	0.3128	0.1433	0.0022	-0.3792
CO	0.2485	0.1114	0.2002	0.1727	0.1087	1	0.1794	-0.2281	-0.2306	-0.1972
O_3	0.0186	0.2955	0.2601	0.2414	0.3128	0.1794	1	-0.1915	-0.2129	-0.2934
AT	-0.1980	-0.2167	-0.1919	-0.1593	0.1433	-0.2281	-0.1915	1	0.1696	-0.0889
RH	-0.2428	-0.0875	-0.1453	-0.0476	0.0022	-0.2306	-0.2129	0.1696	1	-0.0180
WS	0.0514	0.1038	0.0838	-0.0189	-0.3792	-0.1972	-0.2934	-0.0889	-0.0180	1

Seasonal and Stations Distribution (Scree & Biplot Analysis)

The Scree plot displays the percentage of variance captured by each eigenvalue across the principal components (Fig. 5a & b). By retaining components with eigenvalues greater than one (Kaiser's criterion), we successfully captured the highest degree of variability across different seasons and monitoring stations. This approach facilitated the identification of primary pollutants and provided a statistical basis for categorizing distinct source groups.

Seasonal- Specific Pollutant Profiles (Biplot Analysis)

The PCA biplot (Fig. 6) demonstrates clear seasonal clustering. In rainy season, samples were tightly clustered near the origin, with slight extension into the negative PC1 and PC2 regions, indicating low pollutant influence and minimal variability. This compact, near-circular distribution reflects uniform and cleaner air conditions, primarily owing to wet deposition and enhanced atmospheric dispersion. In summer, the samples were moderately clustered around the origin, with a limited spread toward the positive PC1 and slightly negative PC2 regions. The distribution was mainly with $PM_{2.5}$ and SO_2 vectors, indicating their dominance in the dataset. There was minimal spread into the extreme positive or negative zones, showing controlled variability. The shape was compact to mildly elliptical, reflecting moderate pollution levels and better dispersion than in winter. During the winter season (yellow triangles), the samples exhibited a broad spatial dispersion across both the positive and negative regions of PC1, with a pronounced extension into the quadrant defined by positive PC1 and negative PC2. This scattered distribution signifies high environmental variability and a dominant influence from pollutants, most notably NO_2, NO, NH_3, and CO-which, which align strongly with the positive PC1 axis. The elongated, non-uniform shape of the winter cluster reflects a period of maximum pollution load and atmospheric instability. This pattern is mainly due to stagnant conditions (e.g., temperature inversions) that limit dispersion and trap high concentrations of pollutants near the surface.

The PCA biplot clearly shows a transition from compact (rainy) → moderate (summer) → highly dispersed (winter) distributions, indicating increasing pollution intensity and variability across seasons.

Station-Specific Pollutant Profiles (Biplot Analysis)

The PCA biplot illustrates the distribution of monitoring stations along the first two principal components, where PC1 (29.3%) and PC2 (17.0%) together explain 46.3% of the total variance. A clear separation of stations is observed along the PC1 axis, which represents the pollution gradient. Stations such as SN-H (yellow triangles) are widely dispersed toward the positive PC1 and negative PC2 region, indicating higher pollutant loadings, particularly influenced by NO, NO_2, and NH_3, as evidenced by the direction of these vectors. The elongated spread of SN-H reflects greater variability and elevated pollution levels. In contrast, stations such as ZP-H (green inverted triangles) are clustered in the lower-left quadrant, exhibiting negative loadings on both PC1 and PC2. Their proximity to meteorological vectors like WS (wind speed) and RH (relative humidity) indicates lower pollution levels and a stronger influence of dispersion processes. Stations BL-H (blue squares) and IC-H (red circles) are mostly concentrated around the central region (near origin) with slight extension toward the positive PC2 axis, suggesting moderate pollution levels and association with SO_2, O_3, and CO. Their relatively compact clustering indicates less variability compared to SN-H stations.

The PCA results were evaluated against the National Ambient Air Quality Standards (NAAQS) under the Environment (Protection) Act, 1986 (**(**Table 2**)**. While PC1 identifies nitrogenous species as the primary pollutants, the descriptive statistics indicate that mean concentrations of NO_2 (26.17 µg/m^3) and NH_3 (12.90 µg/m^3) currently remain below the annual limits for residential and industrial areas (40 µg/m^{3}and 100 µg/m^3, respectively). However, the $PM_{2.5}$ mean of 38.32 µg/m^3 is approaching the annual safety threshold of 40 µg/m^3, signaling a high risk for public health.

The pie chart (Fig. 3) illustrates the percentage contribution of each pollutant, derived from the descriptive statistics presented in Table 9 based on mean concentrations (µg/m³). The results indicate that CO is the dominant pollutant, contributing the largest fraction (86.3%) to the total pollutant load, with a standardized mean concentration of 805.656 µg/m³.

Along with CO, other major pollutants such as $PM_{2.5}$ (38.32 µg/m³) and NO_2 (26.17 µg/m³) significantly contribute to air quality deterioration due to their higher toxicity and proximity to permissible limits. Furthermore, the high standard deviation observed for $PM_{2.5}$ (44.20) and CO (359.499) suggests substantial temporal variability in pollutant levels over the 365-day study period, likely influenced by seasonal variations and localized anthropogenic activities.

Table 9: Descriptive statistics of air pollutants and climatic factors across 12 months in Hyderabad city

Pollutants	N	Mean	Standard Deviation
$PM_{2.5}$	1464	38.320	44.204
NO	1464	10.991	18.595
NO_2	1464	26.173	24.099
NH_3	1464	12.902	10.176
SO_2	1464	10.737	8.615
CO	1464	805.656	359.499
Ozone (O_3)	1464	28.316	16.612

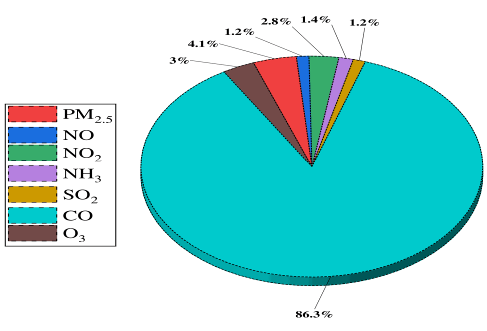


Fig. 9 : The Pie chart of air pollutants for total mass concentration distribution across the 366-day study period in Hyderabad city

Various studies conducted across India and globally have extensively analyzed air quality index (AQI), pollutant concentrations, and the influence of climatic factors on air pollution. These studies consistently indicate that pollutant levels are highest during winter, followed by summer and monsoon, with relatively minor spatial variation across locations (Sengani *et al.*, 2024; Jay *et al.*, 2017; Mohan & Kandya, 2007; Pavón *et al*., 2025). Seasonal variation in air quality is largely attributed to meteorological conditions such as temperature inversions, which trap pollutants near the ground (Chabalala *et al*., 2026).

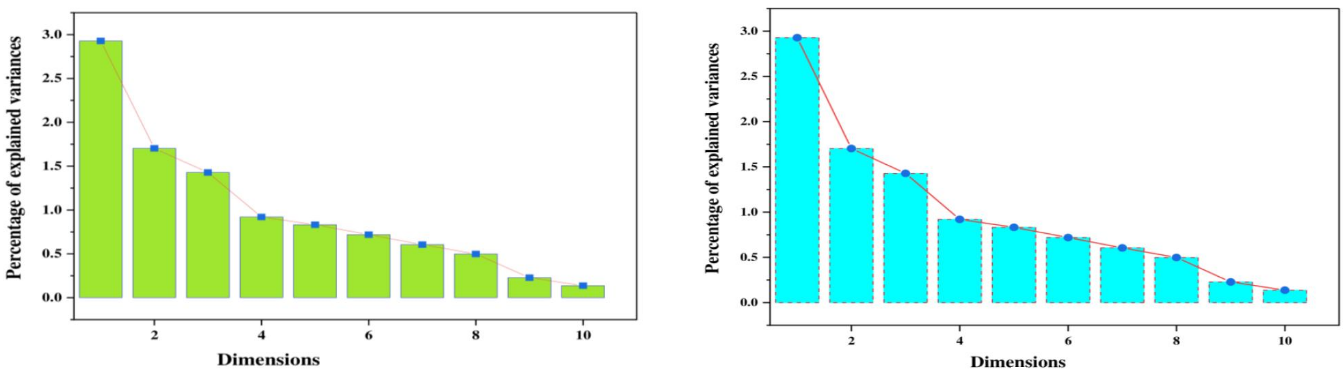
In India, regional assessments highlight varying pollution levels. At the coastal region of Gandhidham–Kachchh, Gujarat, PM_{10} (118–227 µg/m³) and $PM_{2.5}$ (47–82 µg/m³) frequently exceeded National Ambient Air Quality Standards (NAAQS), while AQI ranged from "Good" to "Satisfactory," with occasional "Moderate" levels (Sengani *et al.*, 2024). In contrast, major urban centers such as Delhi (Dwarka, R.K. Puram, Punjabi Bagh, and Anand Vihar) consistently recorded "Very Poor" AQI, with $PM_{2.5}$ identified as the critical pollutant exceeding permissible limits throughout the year. Similar trends have been observed in cities like Hyderabad (Jay *et al*., 2017; Reddy *et al*., 2020; Vaddiraju, 2019). Additionally, in Haridwar (Shivalik Nagar and SIDCUL), extremely high SPM concentrations exceeded NAAQS limits, leading to "Heavy" pollution levels due to vehicular emissions, industrial activities, and poor infrastructure (Chauhan & Joshi, 2008).

Globally, similar concerns persist. In Turkey, analysis of 20 monitoring stations across urban and industrial regions showed exceedances of EU standards for SO_2 (4 stations), NO_2 (11 stations), and PM_{10} (18 stations), indicating severe particulate pollution (Büke & Köne, 2016). Furthermore, studies from Switzerland and the Gauteng province of South Africa emphasized the growing role of advanced modelling techniques, where spatiotemporal graph neural networks (ST-GNNs) effectively forecast $PM_{2.5}$ concentrations using meteorological and remote sensing hyper-spectral data (Chabalala *et al*., 2026).

Recent advancements also highlight the application of machine learning and statistical techniques in air quality analysis. Models such as Support Vector Machines (SVM) and XGBoost have demonstrated superior performance in predicting pollution-related factors, with SVM achieving classification accuracy of approximately 83% (Pavón *et al.*, 2025). Additionally, Principal Component Analysis (PCA) has proven effective in simplifying complex environmental datasets by identifying dominant pollution sources, with the first three principal components explaining a significant proportion of total variability (Acharyya *et al*., 2025; Parsaniya *et al*., 2025). Temperature and humidity were found to have minimal direct influence on pollution levels, acting as secondary variables (Pavón *et al*., 2025).

Conclusion

The study concludes that air quality in Hyderabad is characterized by pronounced seasonal and spatial heterogeneity, driven by the complex interplay between fluctuating meteorological conditions and localized anthropogenic activities. A distinct transition from "Moderate" pollution levels in winter to "Good/Satisfactory" conditions during the rainy underscores the decisive role of atmospheric dispersion and thermal inversions in modulating ground-level concentrations. High-resolution statistical analysis identified Zoo Park, Sanath Nagar, ICRISAT and Bollaram as persistent pollution hotspots with significantly elevated AQI loads, while Principal Component Analysis (PCA) isolated nitrogenous compounds from vehicular and industrial combustion as the primary drivers of the city's pollution profile. Furthermore, although Carbon Monoxide (CO) constitutes the largest mass fraction of total pollutants, the proximity of $PM_{2.5}$



a b

Fig.5: Scree plots illustrating the percentage of explained variance for each principal component (dimension). The "elbow" indicates that the first three components are sufficient to capture the majority of the air quality data variability across **(a)** Season & **(b)** Stations in Hyderabad.

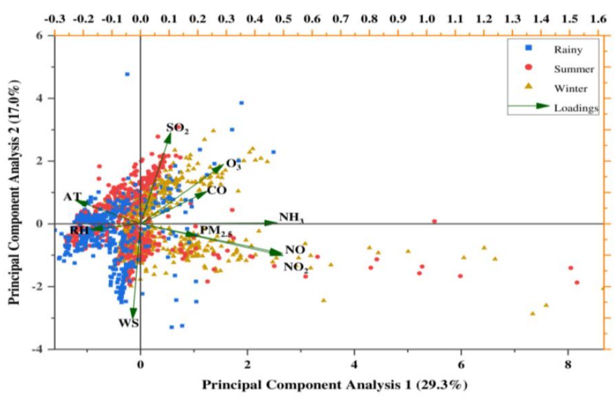


Fig. 6: The Biplot of air quality parameters for PC1 and PC2 across three distinct seasons (Rainy, Summer, and Winter) in Hyderabad city.

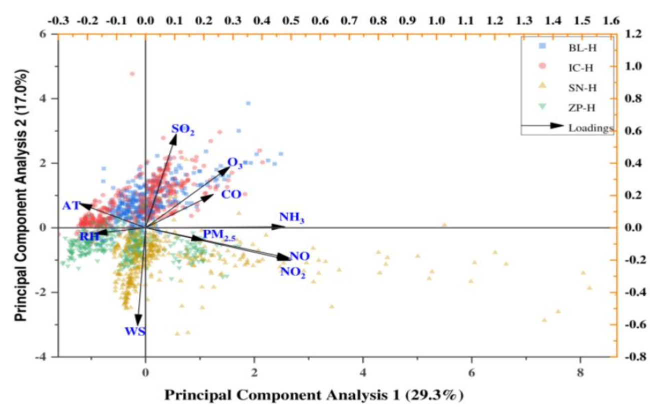


Fig. 7: The Biplot of air quality parameters for PC1 and PC2 across four worst-affected monitoring stations (BL-H, IC-H, SN-H, and ZP-H) in Hyderabad city.

Reference

Acharyya, S., Panda, S. and Sharma, D. (2025). Assessment of genetic diversity using PCA for yield attributing and grain quality traits in aromatic short-grain rice (*Oryza sativa* L.) mutants along with their parental lines. *Plant Archives*, **25(2)**. https://doi.org/10.51470/PLANTARCHIVES.2025.v25.no.2.196

Büke, T. and Köne, A. Ç. (2016). Assessing air quality in Turkey: A proposed air quality index. *Sustainability*, **8(1)**, 73. https://doi.org/10.3390/su8010073

Central Pollution Control Board (CPCB) (2014). Guidelines for manual sampling & analyses (Guidelines for the measurement of ambient air pollutants, Vol. I). National Ambient Air Quality Series (NAAQMS/36/2012–13).

Chabalala, V., Rudolph, C., Mosala, K., Nkadimeng, E.K., Mosomane, C., Mathaha, T. and Basu, P. (2026). Spatiotemporal graph neural networks for PM2.5 concentration forecasting. *Air*, **4(1)**, 2. https://doi.org/10.3390/air4010002

Chauhan, A. and Joshi, P.C. (2008). Air quality index and its variations in Haridwar. *Journal of Environment and Biosciences*, **22(2)**, 171–176.

Mohan, M. and Kandya, A. (2007). An analysis of the annual and seasonal trends of air quality index of Delhi. *Environmental Monitoring and Assessment*, **131(1)**, 267–277. https://doi.org/10.1007/s10661-006-9471-4

Parsaniya, T.A., Patel, R.N., Gami, R.A., Kumar, S., Kugashiya, K.G., Acharya, A. and Donga, A.R. (2025). Multivariate statistical approaches for investigating diversity among sorghum (*Sorghum bicolor* L. Moench) hybrids. *Plant Archives*, **25(2)**. https://doi.org/10.51470/PLANTARCHIVES.2025.v25.no.2.288

Patel, J.S., Salvi, H.U. and Patel, N.R. (2017). Ambient air quality assessment using air quality index of Delhi. *Kalpa Publications in Civil Engineering*, 124–131. https://doi.org/10.29007/xxxx *(check exact DOI if required)*

Pavón, D.M., Rodríguez-Sufuentes, S., Aguado, A., González-Colom, R., Gómez-López, A., Kristian, A., Badyda, A., Kepa, P., Pérez, L. and Fermoso, J. (2025). Modelling the presence of smokers in households for future policy and advisory applications. *Air*, **3(4)**, 27. https://doi.org/10.3390/air3040027

Reddy, S., Verma, P. and Waghulade, M. (2020). Air quality index A study to assess the air quality. *International Research Journal of Engineering and Technology (IRJET)*, **7(2)**, 3363–3367.

Sengani, P.D., Chandra, R. and Agarwal, G. (2024). Ambient air quality assessment through air quality index and air quality health index for Eastern Coastal Region, Gandhidham, Gujarat, India. *International Journal of Environment and Climate Change*, **14(4)**, 692–707. https://doi.org/10.9734/ijecc/2024/v14i44028

Soleimani, F., Dobaradaran, S., De-La-Torre, G.E., Schmidt, T.C. and Saeedi, R. (2022). Content of toxic components of cigarette, cigarette smoke vs cigarette butts: A comprehensive systematic review. *Science of the Total Environment*, **813**, 152667. https://doi.org/10.1016/j.scitotenv.2021.152667

Vaddiraju, S.C. (2019). Mapping of air pollution using GIS: A case study of Hyderabad. *International Journal of Research in Advent Technology*, **7(5)**, 482–487. https://doi.org/10.32622/ijrat.752019328

Van der Walt, P.A., Garl, R.M., Burger, R.P. and Naidoo, M. (2023). Impacts of coal-fired power plants on aerosol particles in the Highveld (Master's thesis). North-West University, Potchefstroom, South Africa. http://hdl.handle.net/10394/20503

Vignesh, P.P., Jiang, J.H. and Kishore, P. (2023). Predicting PM2.5 concentrations across USA using machine learning. *Earth and Space Science*, **10**, e2023EA002911. https://doi.org/10.1029/2023EA002911

Xing, Y.F., Xu, Y.H., Shi, M.H. and Lian, Y.X. (2016). The impact of PM2.5 on the human respiratory system. *Journal of Thoracic Disease*, **8**, E69. https://doi.org/10.3978/j.issn.2072-1439.2016.01.19

OpenAI. (2026). *ChatGPT (GPT-5.3)*. https://chat.openai.com