International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Comparative Seasonal and Temporal Analysis of AQI in Noida and Agra

Harsh Vardhan, Aanurag Gangwar
Dept. of Civil Engineering Institute of Engineering and Technology
Lucknow, India

Abstract- — Air pollution continues to be one of the most important environmental problems in fast-growing parts of the Indo-Gangetic Plains (IGP), such as Agra and Noida, characterized by declining air quality levels. The current paper offers a comparative analysis of the Air Quality Index (AQI) temporal and seasonal trends in the cities of Agra and Noida based on the daily data from the Central Pollution Control Board (CPCB) for the period of 2021–2025. The descriptive statistical analysis demonstrates that AQI is higher and more variable in Noida than Agra, which implies that Noida is under higher levels of pollution. The time-series analysis reveals considerable AQI dynamics in the two cities, including peaks of the parameter under discussion during winter months. The results of the monthly and seasonal analysis also suggest strong seasonality, according to which AQI scores peak during winter and post-monsoon months, while monsoon months are associated with improved air quality. In terms of AQI categories, Noida witnesses more days classified as "Poor," "Very Poor," and "Severe," while Agra shows a larger number of "Moderate" and "Satisfactory" days. Finally, the autocorrelation analysis demonstrates a high level of AQI dependence on the time dimension in both cities. These differences have been observed due to variations in the sources of emission, population density, pollution transport dynamics, and weather conditions. In conclusion, the analysis shows that Noida is more heavily and variably polluted than Agra. This study offers valuable guidance for designing air quality management plans specific to regions and helps frame effective measures for pollution-prone urban areas.

Keywords: Air Quality Index (AQI); Seasonal Variability; Temporal Analysis; Indo-Gangetic Plain; Noida; Agra; CPCB Data.

I. INTRODUCTION

One of the pressing environmental concerns that the world faces is air pollution. This is especially true for fast-growing and industrializing cities. It is estimated that the Indo-Gangetic Plain (IGP) is among the most polluted zones in the world owing to factors like high population density, heavy industrialization, vehicle emission, and biomass burning(Gurjar, 2016; P. et al. Kumar, 2021). Air quality in urban zones in the IGP often falls drastically, resulting in several environmental, economic, and health effects. Exposure to higher particulate matter levels has been associated with respiratory and cardiovascular diseases, mortality, and reduced quality of life (Guttikunda & Goel, 2014).

The Central Pollution Control Board (CPCB) continuously monitors air pollution levels in India and publishes data on the same. Data collected from these sources can provide accurate and long-term information for evaluating air pollution trends in urban environments. Many researchers have utilized these datasets to assess air pollution trends in urban centers(Venkataraman, 2018). Some of the urban centers that are vital for air pollution assessment in India include Noida and Agra.

Air quality variability in urban regions is determined not only by the source of emissions but also by meteorological factors. Many studies indicate that there are distinct temporal and seasonal patterns in AQI. More specifically, seasonality takes a dominant position, as during winter months, air pollution peaks because of temperature inversion, weak winds, and poor atmospheric mixing. This leads to an accumulation of pollutants close to the ground level and causes serious pollution episodes.

In turn, monsoon seasons are characterized by better air quality since precipitation washes out pollutants. Moreover, post-monsoon seasons are marked by a rapid increase in air pollution because of regional biomass and agricultural residue burning (Venkataraman, 2018). Summer months are associated with moderate AQI values since increased temperature and wind speed contribute to the dispersion of pollutants. Therefore, the described characteristics show how significant it is to analyze temporal and seasonal trends related to AQI.

A variety of statistical and machine learning methodologies have been utilized in predicting and analyzing air quality. For instance, the use of traditional statistical models like ARIMA in predicting AQI values and pollutant concentration levels is common practice(Panda & Sameen, 2021). Nevertheless, such models cannot adequately represent the nonlinear nature of air pollution data. In this regard, the combination of statistical and machine learning models has proven beneficial in developing more accurate air quality forecasts (Diaz-Robles, 2008; Zhang et al., 2018).

In recent years, the emergence of machine learning techniques, particularly deep learning and ensemble models, has led to improved air quality indices forecasting (Iqbal & Mukherjee, 2025). Furthermore, the integration of machine learning and statistical techniques through hybrid and ensemble models has led to enhanced accuracy in predicting AQI by effectively capturing both linear and nonlinear associations (Sarkar & Gupta, 2022). Nonetheless, most of the literature has centered mainly on model development and forecast generation while paying scant attention to temporal and seasonal analyses of air quality index in neighboring urban areas.

In this regard, the current study is designed to conduct a comparison of AQI in Noida and Agra based on the CPCB data of the years 2021-2025. The main objective of the study is to analyze the variation in the pattern of air pollution, seasonal differences, and AQI category between both the cities. In addition to that, the research will try to determine the causative factors behind such variations. This will help in formulating appropriate policies in managing air quality in the Indo-Gangetic Plain.

II. MATERIALS AND METHODOLOGY

Study Area

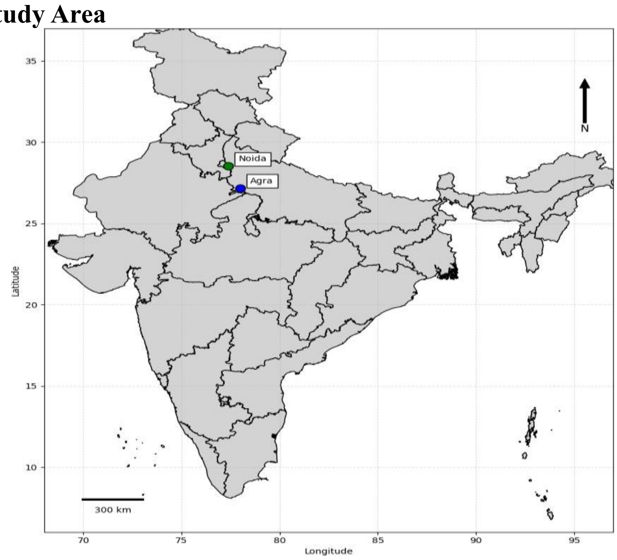

Figure 1: Geographic Locations of Agra and Noida

The current investigation focuses on two key cities in the Indo-Gangetic Plain (IGP): Agra and Noida, which are recognized as highly polluted regions due to rapid urbanization and intense anthropogenic activities. The IGP is characterized by high population density, extensive industrialization, and unfavorable meteorological conditions, making it one of the most polluted regions globally (P. et al. Kumar, 2021). The geographical locations of Agra and Noida within India are illustrated as shown in Figure 1.

Agra, located in the state of Uttar Pradesh (approximately 27.18°N latitude and 78.01°E longitude), is a medium-sized city with diverse emission sources. Major contributors to air pollution include vehicular emissions, industrial activities, construction dust, and emissions from small-scale industries and brick kilns. Additionally, the tourism sector, particularly around the Taj Trapezium Zone (TTZ), adds to local emissions. Agra is also significantly influenced by regional pollutant transport from nearby industrial and rural areas (Venkataraman, 2018).

Noida, situated in the National Capital Region (NCR) near Delhi (approximately 28.53°N latitude and 77.39°E longitude), is a rapidly developing urban center. It experiences high levels of pollution due to dense traffic, large-scale construction, industrial activities, and expanding infrastructure. Being part of the NCR, Noida is also strongly affected by transboundary pollution, especially during winter and post-monsoon periods (Guttikunda & Goel, 2014).

In both cities, seasonal variations in air quality are significant and primarily driven by meteorological factors. During winter, low wind speeds, temperature inversion, and reduced atmospheric mixing height lead to the accumulation of pollutants near the surface, resulting in severe pollution episodes. Conversely, during the monsoon season, rainfall and enhanced atmospheric dispersion reduce pollutant concentrations, leading to improved air quality. In the post-monsoon period, pollution levels increase sharply due to agricultural residue burning in neighboring regions (Tiwari, 2013).

Furthermore, both Agra and Noida are influenced by regional geographical features such as the Indo-Gangetic basin topography, which restricts pollutant dispersion and facilitates pollutant accumulation. Their proximity to major urban and industrial hubs further intensifies pollution levels.

The selection of Agra and Noida as case studies is based on their contrasting urban characteristics within the same geographical region. While Agra represents a moderately urbanized city with mixed industrial and tourism-related emissions, Noida represents a rapidly urbanizing metropolitan extension with high infrastructural growth. This contrast makes them ideal for comparative analysis of AQI variations across different temporal and seasonal scales.

Dataset

The data set adopted for this study was procured from the Central Pollution Control Board (CPCB), India. It should be noted that CPCB, India is the main authority overseeing air quality monitoring and management in the country. The reliability of the CPCB data is ensured by the use of consistent data on standardized high-resolution air quality obtained from CPCB's continuous ambient air quality monitoring station (CAAQMS). Data sets from CPCB have been extensively applied in scientific research into air pollution in terms of their wide temporal and spatial coverage (V. Kumar et al., 2025).

The data set covers daily readings of Air Quality Index (AQI) in the cities of Agra and Noida over a five-year period between January 2021 to December 2025. In case there are no AQI readings available, then the concentrations of pollutants, especially $PM_{2.5}$, were used to estimate AQI using CPCB breakpoints. The resulting data set has 1826 days of AQI values ranging from 1 January 2021 to 31 December 2025.

The dataset consists of the following variables: Date (date stamp of observation)

AQI value (either calculated or measured directly) Concentration of the following pollutants: $PM_{2.5}$, PM_{10}, NO_2, SO_2, O_3, NH_3, and CO

In order to preserve data consistency, timestamp data was transformed into the standardized datetime format, and data were arranged according to the daily time series. In case of missing data resulting from downtime periods or other data losses, linear interpolation was used to maintain time series continuity. Extreme values and outliers were kept in the dataset since they reflect actual pollution events and are important for accurate AQI analysis.

Moreover, data collected at several monitoring stations for each city was aggregated in order to acquire an overall value of AQI for the city level. Daily averaging was applied during this process.

This prepared dataset forms a basis for further analysis of AQI trends and patterns in both Agra and Noida. Temporal and seasonal comparison and analysis of these cities is possible because of the use of consistent and comparable data sources. The details of monitoring stations used for AQI data collection in both cities are presented in Table 1.

S. NO.	CITY	STATION NAME	LOCATION TYPE	AGENCY
1	AGRA	SANJAY PLACE	COMMERCIAL	CPCB
2	AGRA	NUNHAI	INDUSTRIAL	CPCB
3	AGRA	SHAHJAHAN GARDEN	RESIDENTIAL	CPCB
4	NOIDA	SECTOR 62	INDUSTRIAL/TRAFFIC	CPCB
5	NOIDA	SECTOR 1	RESIDENTIAL/TRAFFIC	CPCB
6	NOIDA	SECTOR 116	RESIDENTIAL	CPCB
7	NOIDA	SECTOR 125	INSTITUTIONAL	CPCB

III. METHODOLOGY

The methodology adopted in this study involves a systematic framework for analyzing the temporal and seasonal variability of AQI in Agra and Noida. The approach includes data preprocessing, statistical analysis, temporal aggregation, seasonal classification, and autocorrelation assessment to understand the behavior and dynamics of air pollution in both cities

Data Preprocessing

The process of data preprocessing is essential in ensuring the precision, uniformity, and integrity of the data before any analytical processes take place. The air quality data collected from the Central Pollution Control Board (CPCB) included pollutant concentration measurements collected from different monitoring stations along with their respective timestamps.
First, the dataset underwent cleaning procedures in which the column names were standardized in order to use a consistent naming convention that will help in handling the data easily. Moreover, the timestamps were converted into a datetime object while indexing the data based on the date.

As the data collected was not uniform in terms of the time intervals in which they were recorded, it needed to be resampled to a common temporal resolution of one day using the daily mean value for each pollutant variable. This ensures uniform temporal data for the entire duration of the study period from January 1, 2021, to December 31, 2025.

If there were multiple monitoring stations within one city, the daily mean value of the pollutant concentration in all the stations was taken and averaged out.

Observations with missing data due to instrument failure or incomplete data transfer were taken care of through linear interpolation methods. Such observations were estimated by considering the adjacent data and maintaining continuity and trends in the time series. Furthermore, cases of forward and backward filling were performed to estimate missing data points.

Infinite and incorrect observations were first substituted with null observations. This would make sure that invalid values would not have any impact on our analysis.
Outlier observations along with extreme AQI values were deliberately kept in the dataset since they represent actual pollution events in the atmosphere such as smog, biomass

burnings, or unfavorable weather conditions. The removal of outlier and extreme values may result in a significant loss of information concerning pollution in our study.

The final step included the formation of a time series consisting of a daily sequence of data without any missing observations. Preprocessing environmental datasets is crucial for accurate results and reliable conclusions (Zhang et al., 2018)).

AQI Computation
The Air Quality Index (AQI) was obtained in this work according to the method recommended by the Central Pollution Control Board (CPCB), India. AQI is a composite index of overall air quality which is based on concentrations of several pollutants, like PM2.5, PM10, NO2, SO2, O3, NH3, and CO.

A sub-index for each individual pollutant was determined from the concentration ranges of breakpoint values of CPCB. Each sub-index indicates how much an individual pollutant contributes to the overall AQI value. The formula to compute the sub-index through linear interpolation between breakpoint values is given as follows:

continuous dataset of AQI collected from 1st January 2021 to 31st December 2025.

At first, the time-series plot of the daily AQI readings was constructed for studying the fluctuations in these readings along with identifying any cyclic trend in the data. Time-series analysis helps in studying short term trends like daily and weekly trends and long-term trends of the AQI over a certain period of time. The presence of sudden peaks in the time-series shows that there have been pollution episodes in that month.

For the temporal analysis of the AQI dataset, it is required to study the trend in monthly data. For this purpose, the AQI values were averaged over each month, thus forming a monthly dataset.

The temporal analysis further aids in separating systematic patterns from randomness in AQI values. Recurring patterns found in the time series imply that there are some environmental and anthropogenic factors influencing air quality on a recurrent basis instead of a random one.

Finally, through temporal analysis, the comparison can be drawn between pollution behaviors in Agra and Noida with

$$I = \frac{I_{hi} - I_{lo}}{BP_{hi} - BP_{lo}} (C_p - BP_{lo}) + I_{lo} \qquad (1)$$

Where:
- I = Sub-index for a pollutant
- Cp = Observed concentration of the pollutant
- BP {hi}, BP{lo} =

Breakpoint concentrations surrounding Cp
I{hi}, I{lo} = AQI values corresponding to the breakpoints regard to different features, such as peak intensity, number of events, and other trends in AQI.

In conclusion, temporal analysis is an essential part of air quality study because it is the first stage to consider when exploring the temporal evolution of the phenomenon under

I = Sub-index for a pollutant
Cp = Observed concentration of the pollutant
BP {hi}, BP{lo} = Breakpoint concentrations surrounding Cp
I{hi}, I{lo} = AQI values corresponding to the breakpoints
Breakpoints for each pollutant used in the analysis were determined according to the guidelines proposed by the CPCB. This methodology is based on dividing pollutant concentration levels into six classes, including Good (0-50), Satisfactory (51-100), Moderate (101-200), Poor (201-300), Very Poor (301-400), and Severe (401-500).

After calculating the sub-index for each measured pollutant on a particular day, an AQI value for that day was established by identifying the largest sub-index out of all pollutants considered in the analysis.

In other words, the most critical pollutant affecting air quality was chosen through this AQI calculation technique. AQI time series were computed using a programming language, Python. In this case, pollutant concentration data were transformed into breakpoints and sub-indices using the interpolation function.

CPCB-based AQI computation is a widely popularized method in the context of air quality studies due to its high effectiveness in communicating air pollution levels (Kumar et al., 2020).

Temporal Analysis

The temporal analysis was carried out to find the variations in AQI with respect to time, to find trends, cycles, seasonality and other outliers related to pollution levels in Agra and Noida. This was done on the basis of the consideration.

Seasonal Analysis

Analysis during different seasons was carried out to analyze the changes in AQI due to varying weather conditions and the impact of various meteorological factors on air quality at both Agra and Noida. As the level of air pollution is largely impacted by variations in seasons, data was sorted into four broad categories of seasons based on Indian climate: winter (Dec-Feb), summer (Mar-May), monsoon (Jun-Sep) and post monsoon (Oct-Nov).

Data regarding AQI levels per day was classified into the above seasons and seasonal average AQI was calculated for both cities. It provides an understanding of seasonal AQI variations along with determining which season had the maximum deterioration of air quality and vice versa. Seasonal bar graphs were plotted to compare the AQI between Agra and Noida across four seasons. These visualizations highlight differences in pollution intensity and provide insights into how seasonal factors influence AQI variability.

Seasonal classification is significant in recognizing major pollution periods, such as winter and post-monsoon seasons, where the level of pollution is usually high because of adverse meteorological factors, including temperature inversion, low wind velocity, and decreased atmospheric mixing. The monsoon season, on the other hand, shows lower AQI readings because of rain-based pollutant washing out.

Comparison of seasonal data between the two cities of Agra and Noida is essential in realizing how local sources of emissions, as well as regional factors, influence air pollution during comparable climatic seasons.

In general, seasonal analysis is very important in detecting major periods of pollution and can help in formulating pollution reduction measures.

AQI Distribution and Category Analysis

The AQI values were examined to get information about their variability, average levels, and extremes that reflect pollution events in Agra and Noida. Several statistical visualizations including box plots have been used to analyze the AQI values distribution and identify outliers reflecting pollution episodes. Box plots display medians, interquartile ranges, and outliers providing an opportunity for comparison of AQI variability. High AQI values interquartile ranges and many high outliers reflect great variability and pollution events in Agra and Noida.

To make the interpretation easier, the AQI values were additionally grouped according to the CPCB criteria in six categories: Good (AQI < 50), Satisfactory (AQI = 51–100), Moderate (AQI = 101–200), Poor (AQI = 201–300), Very Poor (AQI = 301–400), and Severe (AQI > 400).

The distribution of the AQI Categories was performed to show the percentage distribution of each category. It gives information about the distribution of different levels of air pollution and also allows comparing Agra and Noida in terms of the occurrence of different air pollution categories.

Distribution Category plots helped to demonstrate the above-mentioned percentage distributions. If there are more cases when the air pollution levels fall under “Poor”, “Very Poor”, and “Severe” air pollution categories, it means that people are exposed to bad air quality for a larger number of days.

Overall, combining AQI Distribution and Air Pollution Classification allowed making a conclusion concerning the variation and intensity of air quality conditions. At the same time, it also makes it possible to determine cities where pollution intensity is higher and more extreme episodes happen.

Autocorrelation Analysis (ACF)

To analyze temporal dependency and persistence of AQI readings of Agra and Noida, autocorrelation analysis was conducted. Autocorrelation Function (ACF) analyzes the

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

correlation between two AQI values at various lags and shows how current AQI readings are dependent on past readings.
In this study, the ACF was calculated based on daily AQI readings collected between 2021 and 2025. ACF charts reveal a series of correlations between AQI values and lag times, which can be used to detect the degree of short-term and long-term temporal dependency.

Highly persistent values of autocorrelation at lower lags, especially lag_1, reveal that there is short-term persistence in AQI readings in both locations. Short-term persistence refers to the fact that current AQI values depend on the AQI value of the previous day, and is attributed to the fact that pollutants remain in the atmosphere over several consecutive days due to poor dispersion.

Slow decay of the ACF plot also means that AQI values have some long-term temporal dependency, meaning that such values are structured over time and are not random.

Comparative analysis of the autocorrelation function for Agra and Noida enables determination of dissimilarities in terms of persistence and temporality of the pollutants. The more autocorrelation, the more persistent is the process, and less variable the dispersion is going to be. In general, the autocorrelation function proves that there is a clear dependence on time in AQI in these two cities.

Tools and Software
The analysis presented in this study was carried out using the Python programming language, which provides a robust and flexible platform for data processing, statistical analysis, and visualization. Python is widely used in environmental research due to its extensive ecosystem of libraries and its capability to handle large time-series datasets efficiently.

Several specialized Python libraries were utilized to perform different stages of the analysis. The Pandas library was used for data handling, preprocessing, and time-series operations, including data cleaning, resampling, aggregation, and interpolation. NumPy was employed for numerical computations and efficient handling of array- based operations.

For data visualization, the Matplotlib library was used to generate high-quality graphical outputs, including time series plots, seasonal bar charts, box plots, and category distribution graphs. All figures were produced with high- resolution settings (600 DPI) and formatted using consistent styling to ensure publication-quality presentation.

Time-series analysis, particularly autocorrelation analysis, was performed using the Statsmodels library, which provides reliable statistical tools for computing the Autocorrelation Function (ACF) and analyzing temporal dependencies.

The entire workflow was executed in a Python-based development environment, ensuring reproducibility and consistency of results. The use of open-source tools enhances transparency and allows for easy replication and extension of the study in future research.

Overall, the integration of these tools provided a comprehensive and efficient framework for analyzing AQI variability and conducting comparative assessments between Agra and Noida.

IV. RESULTS & DISCUSSION

AQI Time Series Analysis
The analysis of the temporal variability of the AQI for Agra and Noida was based on the data obtained from 1 January 2021 to 31 December 2025. The time series (Fig. 2) presents the change in AQI values over the analysis period and allows one to reveal the peculiarities in their dynamics and periodicity.

As can be seen from the data, the AQI values have an obvious temporal variability and demonstrate significant changes over time. At the same time, Noida's AQI index values are constantly higher than those of Agra, which indicates worse air pollution in the city.

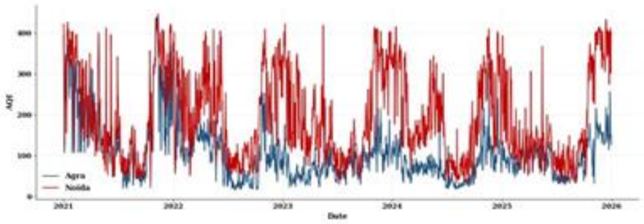

Figure 2: Temporal Variation of AQI in Agra and Noida (2021–2025)

In Noida, the AQI often exceeds the values of 300, achieving peaks at more than 400, which corresponds to "Very Poor" and "Severe" classes of air quality. In Agra, AQI levels are always low; its peaks are in the range from 200 to 250, but there are some exceedances of this range, too.

It is possible to observe a clear cyclic pattern in the time series under consideration, which is expressed in the regularity of

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

obtaining maximum values in the winter period and low values in the summer one. The peaks occurring in close succession testify about the accumulation of pollutants under adverse weather conditions.

Additionally, the time series reveals the occurrence of extreme pollution episodes, particularly in Noida, where sharp spikes indicate sudden increases in pollutant concentration. These episodes are likely associated with factors such as increased emissions, regional pollution transport, and adverse atmospheric conditions.

Overall, the temporal analysis demonstrates that while both cities experience significant AQI fluctuations, Noida consistently exhibits higher pollution intensity and greater variability compared to Agra, highlighting the need for more stringent air quality management strategies in the region.

Monthly Variation Analysis

The seasonal variations in AQI for the cities of Agra and Noida were evaluated using the monthly averages of AQI during the years 2021–2025. These results are depicted in Fig. 3 and show the seasonal trend of AQI variation.

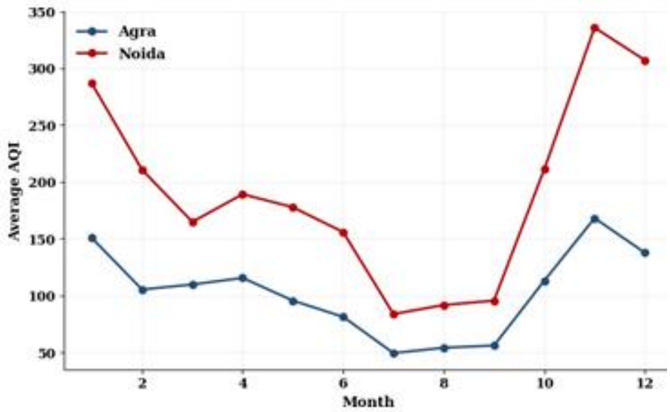

Figure 3: Monthly Variation of Average AQI in Agra and Noida

As evident from the analysis, there is an evident and constant pattern of monthly variation in AQI. The AQI values are at their maximum levels during winter seasons (Nov to Jan) and minimum during monsoon seasons (July to Aug) in both the cities.

For Noida, the AQI values reach the maximum level of about 335–340 during November and stay at high levels (~305) in December and January, suggesting severe pollution during winters. In Agra, the AQI values are comparatively lower, reaching ~165–170 during November and ~135–150 during December-January.

From winters to the monsoon season, there is a decreasing trend of AQI. The minimum AQI values are recorded during July (~85 for Noida, ~50 for Agra) and August (~90 for Noida, ~55 for Agra).

Following the monsoon period, AQI values begin to rise again during September and October, eventually leading to sharp increases in November, particularly in Noida. This rapid rise is likely associated with post-monsoon agricultural residue burning and stable atmospheric conditions.

Throughout all months, Noida consistently records higher AQI values than Agra, indicating a persistent difference in pollution levels between the two cities. The gap between the cities becomes more pronounced during high pollution periods, particularly in winter.

Overall, the monthly analysis confirms the presence of strong periodic behavior in AQI, driven by seasonal meteorological conditions and emission patterns, with Noida experiencing significantly higher pollution intensity compared to Agra.

Seasonal Analysis

One of the key characteristics of AQI behavior in Agra and Noida is seasonal variability. Seasonal average values of AQI were calculated based on data classified according to four climatic seasons (winter, summer, monsoon, post- monsoon).

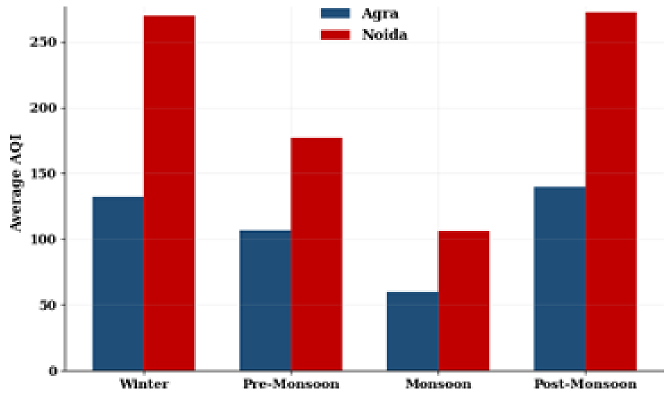

Figure 4: Seasonal Variation of Average AQI in Agra and Noida

As seen from the results, the seasonal behavior of AQI is clearly manifested in both cities where the maximum levels of this parameter are observed in winter. For instance, in Noida, AQI during winter months ranges within about 270- 275, while in Agra, AQI is somewhat lower and ranges within about 130-135. Such high values of the AQI are mainly due to unfavorable meteorological conditions, including temperature inversion, low wind speed, and weak atmospheric turbulence.

AQI levels in post-monsoon season continue to be at high levels, especially for Noida (~270). However, in Agra, AQI drops slightly compared to winter season to ~140. This steep increase in AQI right after monsoon is mostly associated with agricultural residues burning in nearby areas as well as due to stable meteorological conditions.

As for summer season, AQI levels in this season are rather moderate in both cities, ranging within 175-180 in Noida and 105-110 in Agra. This is related to favorable weather conditions characterized by warm weather and high wind speed.

The AQIs have been found to be lower during monsoons in both cities with Noida at about 105 and Agra at about 60. This marked decrease in AQIs during this particular season is mostly because of precipitation-driven wet deposition processes that result in effective removal of suspended particles from the atmosphere.

Noida is also consistently seen to have high AQIs compared to Agra, thus showing higher levels of pollution. This variation in the levels is seen especially during winter and post-monsoon seasons.

In conclusion, the analysis shows that there are meteorologically significant differences between Noida and Agra's AQIs depending on the different seasons. This means that air pollution in both cities is season dependent and more so during winter and post-monsoon seasons.

AQI Distribution Analysis

The analysis of the distribution of AQI values in Agra and Noida was carried out using box plots (Fig. 5).

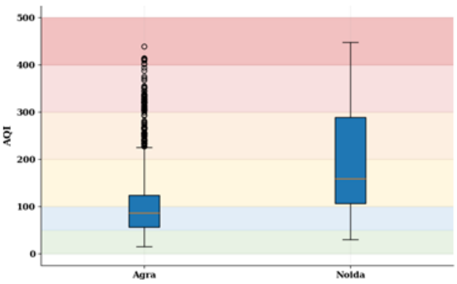

Figure 5: Statistical Distribution of AQI with CPCB Pollution Categories in Agra and Noida

Based on the results of the analysis using the box plot, the difference in the distribution of the AQI in the two cities is clearly noticeable. So, in Noida, a noticeably higher median value is observed, which ranges from 150 to 170, while the median in Agra is only 80-90, meaning that Agra had a relatively better state of air pollution.

The dispersion of the interquartile range, which reflects the scatter of the middle half of the data, is also greater in Noida, ranging from 110 to 290, while in Agra, it is 60-120, which means that AQI fluctuates more in Noida.

In addition, the upper range is noticeably higher in Noida, with frequent values exceeding 400, meaning that “Severe” pollution occurs in Noida, while in Agra, such values rarely exceed 250, but some peaks are still visible.

Outliers found in both the cities indicate that there were episodic occurrences of pollution, especially when pollution levels were high. Nevertheless, such outliers are more common and severe in Noida, as observed earlier, confirming the higher levels of pollution intensity in the city.

From the findings presented above, it is clear that Noida has higher average AQI levels than those recorded in Agra. Also, the higher variance and frequency of outliers confirm higher pollution severity in Noida.

AQI Category Distribution

Distribution of AQI values in relation to the standard AQI classifications was also evaluated to examine the extent to which air pollution conditions occur in Agra and Noida. This classification uses CPCB-defined AQI categorizations as "Good" (0–50), "Satisfactory" (51–100), "Moderate" (101–200), "Poor" (201–300), "Very Poor" (301–400), and "Severe" (>400). The percentage distribution in these categories is depicted in Fig. 6.

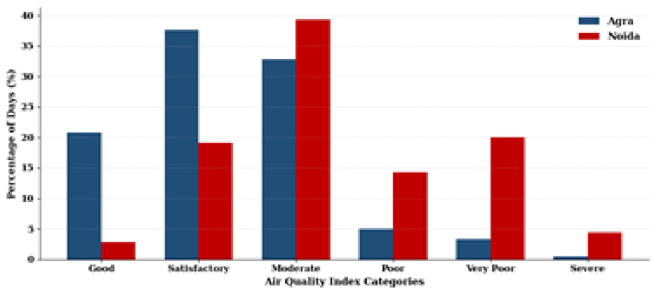

Figure 6: Distribution of AQI Categories in Agra and Noida

Based on the results, there is a noticeable difference between the air quality conditions that exist in Agra and Noida. Agra has relatively more occurrences of days categorized as cleaner air, with "Good", "Satisfactory", and "Moderate" accounting for about 22%, 36%, and 31% respectively. Conversely, the number of air days classified as "Good" and "Satisfactory" in Noida is noticeably lower at 10% and 19% respectively.

However, Noida tends to have higher air pollution days than Agra, especially those classified in the "Moderate", "Poor", "Very Poor", and "Severe" categories. "Moderate" and "Poor" account for 34% and 20% while "Very Poor" and "Severe" stand at 14% and 3-4% respectively in Noida.

A key observation is that more than 70% of days in Noida fall under Moderate to Very Poor categories, whereas Agra experiences a relatively higher share of cleaner air conditions. This indicates that residents of Noida are exposed to more frequent and severe air pollution levels compared to Agra.
The presence of "Severe" category days in Noida further highlights the occurrence of extreme pollution events, which are relatively rare in Agra. These findings are consistent with earlier results from time series, seasonal, and distribution analyses, reinforcing the conclusion that Noida exhibits higher pollution intensity.

Overall, the AQI category distribution provides a practical representation of air quality conditions and confirms that

Noida experiences more frequent unhealthy air quality conditions than Agra, emphasizing the need for targeted pollution mitigation strategies.

Autocorrelation Analysis (ACF)
The autocorrelation test was used to evaluate the degree of dependency and persistence of AQI in Agra and Noida. The Autocorrelation Function (ACF) plot (Fig. 7) provides the degree of correlation of AQI value with its own past observation at various time intervals.

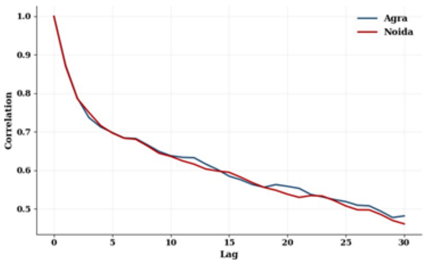

Figure 7: Autocorrelation Function (ACF) of AQI for Agra and Noida

The findings show the existence of strong positive autocorrelation at low lags, especially lag_1 where the correlation coefficient value ranges between 0.85–0.90 for both locations. Therefore, AQI values in a particular day depend significantly on the previous AQI level. It shows the consistency and persistence of pollution in the atmosphere.
Additionally, ACF plots show a relatively slow decay and significant correlations up to lag 20–25. Thus, there exists a strong temporal pattern which demonstrates that the AQI values are not random and are systematic over time.

Finally, comparison of ACF plots in both cities shows a slightly higher autocorrelation in Noida across all lags. The implication is that there is greater persistence of pollutants in the atmosphere than in Agra.

The observed autocorrelation pattern is consistent with earlier findings from temporal and seasonal analyses, where prolonged periods of high AQI were observed, particularly during winter months. The persistence of pollution over consecutive days can be attributed to continuous emissions combined with unfavorable meteorological conditions such as low wind speed and temperature inversion.

Overall, the ACF analysis confirms that AQI in both cities exhibits significant temporal dependence and persistence, with Noida showing relatively stronger autocorrelation. These findings highlight the importance of time-series approaches in understanding air quality behavior and support the observed trends in earlier sections.

V. CONCLUSION

The study carried out a comparative analysis of AQI variability in Agra and Noida with CPCB data from 1 January 2021 to 31 December 2025 (1826 observations). It was found that there are distinct patterns of air pollution for these two regions, and that Noida suffers from comparatively higher values of AQI as well as AQI variability.

Temporal analysis demonstrated that AQI values in Noida regularly exceeded 300-400, while in Agra they stayed under 250, which meant better conditions of air pollution in the latter region. According to monthly analysis, there were AQI peaks in November-January, with Noida AQI value being approximately 335-340 in November, while in Agra it dropped to 165-170. The lowest AQI values were obtained during monsoon months, ranging around ~85-90 in Noida and ~50-55 in Agra.

Regarding seasons, it was identified that winter was the most polluted time, when AQI averaged 270-275 in Noida and 130-135 in Agra, whereas monsoon had the lowest values, which ranged around ~105 in Noida and ~60 in Agra. According to distribution analysis, the median of AQI values was higher in Noida (~150-170) than in Agra (~80- 90), and there was much greater variation and AQI extremes >400 in Noida than in Agra.

AQI category analysis showed that more than 70% of days in Noida fall under Moderate to Very Poor categories, whereas Agra experiences a higher proportion of cleaner air conditions, with nearly 58% of days classified as Good or Satisfactory. Additionally, Noida recorded 3–4% Severe pollution days, while such extreme conditions were nearly negligible in Agra.

Autocorrelation analysis further confirmed strong temporal dependence, with high lag_1 correlation values (~0.85–0.90) and a slow decay pattern, indicating persistent pollution conditions, particularly in Noida.

Overall, the study demonstrates that Noida experiences higher pollution intensity, greater variability, and more frequent extreme events compared to Agra. These differences are attributed to higher urban emissions, regional pollution transport, and unfavorable meteorological conditions. The findings emphasize the need for city-specific air quality management strategies, particularly for rapidly urbanizing regions like Noida, to mitigate pollution and improve public health outcomes.

REFERENCES

1. Diaz-Robles, L. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter. Atmospheric Environment, 42, 8334–8345. https://doi.org/10.1016/j.atmosenv.2008.07020
2. Gurjar, B. R. et al. (2016). Human health risks in megacities due to air pollution. Atmospheric Environment, 142, 394–402.
3. Guttikunda, S. K., & Goel, R. (2014). Health impacts of particulate pollution in a megacity—Delhi, India. Environmental Development, 6, 8–20.
4. He, K., Huo, H., & Zhang, Q. (2002). Urban air pollution in China: Current status and characteristics. Atmospheric Environment, 36(1), 1–13.
5. Iqbal, A., & Mukherjee, N. (2025). A Systematic Review and Comparative Study of Machine Learning Techniques for Air Quality Prediction. Water, Air, and Soil Pollution, 236(12). https://doi.org/10.1007/s11270-025-08410- 3
6. Jiang, D., Zhang, Y., Hu, X., Zeng, Y., Tan, J., & Shao, D. (2004). Progress in developing an ANN model for air pollution forecasting. Atmospheric Environment, 38(40), 7055–7064.
7. Kumar, P. et al. (2021). Air quality in the Indo-Gangetic Plain: A review. Atmospheric Environment, 253, 118376.
8. Kumar, V., Aggarwal, V., Singla, S., Rana, S. K., Kumar, D., & Kanauzia, R. (2025). Large-Scale AQI Forecasting Using LSTM and XGBoost: A Comparative Study Over Delhi NCR. 2025 2nd International Conference on Computational Intelligence and Computing Applications (ICCICA), 307–312. https://doi.org/10.1109/ICCICA67008.202 5.11337632
9. Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality. Nature, 525, 367–371.
10. Panda, P. R., & Sameen, B. S. L. S. (2021). ARIMA modeling of air pollution in major Indian cities. Environmental Science and Pollution Research, 28, 39250–39263.
11. Sarkar, N., & Gupta, R. (2022). Air Quality Index prediction using an effective hybrid deep learning model. Environmental Pollution, 315. https://doi.org/10.1016/j.envpol.2022.1204 04
12. Sharma, M., & Dikshit, O. (2016). Comprehensive study on air pollution and greenhouse gases in Delhi. TERI Report.
13. Tiwari, S. et al. (2013). Variation of air pollutants in Delhi. Atmospheric Research, 132, 52–63.
14. Venkataraman, C. et al. (2018). Source influence on air quality in India. Current Science, 115, 101–110.
15. World Health Organization. (2021). WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. WHO.
16. Zhang, Y., Bocquet, M., & Mallet, V. (2018). A hybrid ARIMA and machine learning model for air quality prediction. Atmospheric Environment, 180, 1–10. https://doi.org/10.1016/j.atmosenv.2018.02.032
17. Zou, B., You, J., Lin, Y., Duan, X., Zhao, X., Fang, X., & Huang, B. (2015). Air pollution intervention and life-saving effect in China. Environment International, 76, 1–8. https://doi.org/10.1016/j.envint.2014.11.01 7