Meteorological Factors Correlation with Air Pollutants: A Case Study in Delhi

Meenakshi Malhotra* and Inderdeep Kaur Aulakh

***Abstract*—Air pollution is increasing at an alarming rate, and meteorological elements have a significant influence on the movement of major air pollutants. This article conducted a detailed investigation of air pollutants (SO_2, NO_2, NO, CO, O_3, $PM_{2.5}$, PM_{10}) and their interaction with meteorological factors in Delhi, India, from 2019 to 2021. We used correlation and partial correlation methods to examine the association. To better understand the relationship between air contaminants and meteorological characteristics, we controlled single and multiple parameters. We discovered that changing the variable causes a significant shift in the relationship between seasonal and regional attributes. We also explored the correlation between the Air Quality Index (AQI) and all of the air contaminants and climatic variables available for the study. Our research will help in the development of a better decision-making system based on the kind of meteorological parameters and air pollutants.**

***Index Terms*—Air pollutants, Air Quality Index (AQI), correlation, meteorological factors**

I. Introduction

Every person's dream is to breathe pure air. However, the fast rise of the economy has resulted in several major environmental issues, one of which is air pollution. Air includes a variety of gases, but the most common are SO_2, NO_2, CO, O_3, NO, NOx, and Particulate Matter (PM). It also carries meteorological fundamentals such as wind speed, wind direction, temperature, humidity, and sun radiation but not limited. $PM_{2.5}$ [1–3] and PM_{10} [4, 5] have gotten a lot of attention in recent decades because of their health implications and can induce hazardous diseases that can lead to death.

Air pollution has a variety of negative consequences, including the mortality of humans, animals, and plants. Human activities can contribute to air pollution such as traffic, industry emissions, waste, stubble burning, smoke, and so on, or it can also be caused by the impact of meteorological conditions on the air. As a result, air contaminants have varied growing and diminishing impacts depending on the season and place. It's just as vital to understanding the link between air pollutants and meteorological parameters as it is to regulate human behavior in terms of air pollution control.

Governments and agencies may create better policies with a deeper grasp of relationships. Why is it necessary to establish a link between air contaminants and meteorological factors in the study?

- Research into the relationship between meteorological conditions and air pollutants might aid in the awareness of the problem of air pollution and lead to the implementation of more effective pollution-reduction strategies.
- The research might be use to increase knowledge of the mechanisms that cause air pollution as well as the accuracy of forecasting systems and can be used as a guide for environmental policy decision-making.
- The research will help to improve the accuracy of air pollution forecasts using meteorological metrics that take seasonal and other aspects into account.

The primary motivation for this study demonstrates the necessity for current information on air pollution in cities.

The study's major goal is to give a complete data description of the factors that influence the AQI level of pollutants.

Seasonal and geographical differences in season, year, city, and regional areas were taken into account.

Before things grow worse, it's important to understand the pattern and process at work. It's important to keep track of the elements that influence it. Conditions or uncertainties should be taken into account hence understanding the pattern might lead to a more accurate and reliable forecasting system.

The rest of the paper is arranged as follows: Section II contains the previous work related to the correlation findings. Section III discusses the data and approach used in the study. The descriptive and statistical analysis are presented in Section IV. Section V contains a critical discussion of the novelty of our work. The observations are summarized in Section VI. Section VII concludes the paper.

II. Related Study

In research of Yang *et al*. [6], Six key pollutants ($PM_{2.5}$, PM_{10}, CO, SO, NO, O_3) and Meteorological factors (Temperature, Wind Speed and Precipitation, Sea level pressure, Boundary layer height) were considered for the study. The authors examined the association between numerous contaminants and climatic conditions while taking into account the various emission sources. The authors evaluated not just the association between pollution data and meteorological data, but also the relationship between pollutant concentration and emission sources. The authors noted that the association of pollutants and meteorological data relies on the season of that place, and they showed how control in emission sources may control the pollutants concentrations using the WRF-Chem (Weather Research and Forecasting using Chemistry) model. Qi *et al*. [7] analyzed that a single meteorological factor has a limited effect on pollutant concentration, so they combined the meteorological factors i.e., Temperature-Windspeed, Temperature-Pressure, and Humidity-Windspeed, and then identified their effect on

Manuscript received August 15, 2022; revised September 7, 2022; accepted October 9, 2022.

The authors are with the Department of Information Technology, University Institute of Engineering and Technology, Panjab University, Chandigarh, 160014 India.

*Correspondence: meenakshi.malhotra@outlook.com (M.M.)

doi: 10.18178/ijesd.2023.14.2.1420

pollutant concentration. By using the Fitted Relationship technique, they concluded that on the same pollutants, different meteorological factors showed different results and on the same meteorological conditions, different pollutants showed different results. Kayes *et al*. [8] considered seasonal fluctuations to identify the relationship between meteorological parameters and air pollutants. Both multilinear and non-linear regression models were used to identify the impact on air pollutants. By using both the models, the prediction results were same, but for gaseous pollutants both the models performed poor. Liu *et al.* [9] has revealed that concentration of air pollutants was affected by meteorological parameters at same level and the level of impact depends on the type of pollutant present. Seasonal and Regional characteristics were considered for the study with variation in trends (Warming, Cooling, Increasing, and Downward Trends). To identify the association between pollutant concentration and meteorological parameters, partial correlation analysis was done followed by multivariable linear regression and panel data model. Qiao *et al*. [10] has identified seasonal and regional variations of pollutants affected by meteorological factors. National-scale exploration was considered to examine the relationship between air pollution index and multiple meteorological parameters using partial correlation coefficient in addition to hierarchical cluster analysis. Battista and Vollaro [11], focused on statistics and cross-statistics techniques in time and space. Except for Particular Matter, remaining considered pollutants showed strong coupling with temperature, solar radiation, wind direction, and velocity using cross-correlation analysis. Yang *et al.* [12] focused on spatial and seasonal factors in the research. Season, year, city, regional scales, spatial and seasonal variations were analyzed. Authors after preprocessing the data, distributed the $PM_{2.5}$ concentration temporarily and spatially followed by applying multi-Scale correlation analysis. The correlation variation depending upon season and region showed the correlation results. Zhang *et al*. [13] has carried out a systematic analysis to identify the relationship. Three megacities of China were considered with their temporal and seasonal variation. Hou and Xu [14] integrated the generalized additive model with marginal effects for the analysis. In conclusion, the authors discovered that temperature, relative humidity, and visibility have a substantial relationship with $PM_{2.5}$ concentration, whereas O_3 has a strong relationship with temperature and relative humidity. Temperature and solar radiation have only a slight relationship with O_3. CO is significantly affected by atmospheric pressure and temperature. Other meteorological elements (Wind Speed, Visibility, Precipitation) have an impact on air quality as well, albeit to a lesser extent. The authors were unable to include PM_{10} and NO_2 due to a lack of data. Kliengchuay *et al*. [15] analyzed the association between climatic conditions and air pollution concentrations using R software. The authors concluded that PM_{10} and relative humidity have a negative association, whereas CO has a positive correlation with PM_{10}. The authors proposed that PM_{10} concentration was also impacted by various atmospheric elements such as seasonal change, daily temporal variation, and the available meteorological parameters. The authors noted that the approach might be utilized to increase knowledge of PM_{10} concentration patterns in Mae Hong Son, based on the research findings.

Manju *et al*. [16] applied a general linear model and multiple linear regression to analyze the data. A general linear model was used to study the regional and seasonal variation in pollution, and Pearson correlation was utilized to find the link between air pollutants (PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, O_3) and meteorological elements. According to the findings, CO and O_3 have a significant relationship with temperature and relative humidity. The authors suggested a thorough analysis to better understand the consequences of each pollutant and climatic component in the future.

He *et al*. [17] evaluated the temporal and geographical aspects of pollution in 31 Chinese provincial capital cities. The z-score approach was used to standardize the data, then principal component analysis and the k-means cluster method were merged. The wavelet artificial neural network model was used to investigate the relationship between air pollution and meteorological factors. Pollutant concentrations were shown to be favorably connected with temperature and negatively correlated with wind speed, according to the authors. The relationship between relative humidity and temperature revealed geographical differences. Each region, such as the north, northern China, and so on, has various correlation effects. The authors also noted that gas pollutants were more influenced than $PM_{2.5}$ and PM_{10}.

Fang *et al*. [18] employed multi-scale correlation analysis to determine the association between air quality and meteorological data, and the correlation coefficients were determined using the decomposed Intrinsic Mode Function and the decomposition of time series data was done using EEMD method. According to the authors, "by using correlation analysis on deconstructed modes, they might distinguish correlations on different scales without any extra filtering or smoothing." On both lower and higher frequency oscillations, a connection between PM and climatic parameters was discovered. By using the correlation analysis on decomposed Intrinsic Mode, the multiscale correlation analysis method was established to extract correlation on multiple scales. The study did not take into account other meteorological characteristics such as rainfall, radiation data, or onsite meteorological observations.

Deswal and Chandna [19] investigated the relationship between suspended particulate matter, respirable particulate matter, SO_2, and NO_2 employing meteorological data. The authors first investigated inter-site and intra-site variation in air quality, then the seasonal influence of varying meteorological conditions on air pollutant concentrations. Multiple linear regression was used to investigate the relationship between meteorological pollutants and air pollutants. The authors concluded that when precipitation, wet day frequency, and cloud cover diminish, thus does the diurnal temperature range. Though the amount of the influence of climatic circumstances varies per air pollutant. Cui *et al*. [20] employed multivariate statistical analysis that leads to normal distribution analysis utilizing SPSS software and Pearson correlation. The authors discovered the association by dividing the data into two periods: heating period (November-March) and non-heating period (April-October). The authors observed a significant shift in the association between climatic parameters impacting air pollution concerning the non-heating season. Wind speed, relative humidity, and cloud cover were linked with AQI during the heating season, whereas temperature, precipitation, wind speed, and relative humidity were key meteorological elements impacting air quality during the non-heating period.

To investigate the dynamic influence of climatic circumstances on air pollution, Zhang [21] used the vector autoregression model, the granger causality test, the impulse response function, and variance decomposition. The authors stated that air pollution in a city is self-aggregation, self-diffusion, and self-cumulative and that if the diffusion condition of air pollution worsens, air pollution would occur within three days. The authors found that climatic parameters such as severe wind speed, sunlight duration, and average humidity influence air pollution concentration and spatiotemporal dispersion. The authors suggested to add terrain features, regional pollution transport and transformation, residential heating, industrial pollution, and automobile exhaust emissions in the future.

Ma *et al.* [22] identified the association using the Pearson correlation coefficient. In terms of seasonal fluctuations, $PM_{2.5}$, PM_{10}, SO_2, and CO concentrations were highest in winter and lowest in summer. O_3 had the reverse seasonal variation. Air pollution was worst in winter and lightest in fall, according to the number of non-attainment days. $PM_{2.5}$ was the most significant contributor to air pollution, followed by NO_2, O_3, PM_{10}. Except for O_3, all of the pollutants were substantially positively linked due to emission sources. According to the authors, both emissions and climatic circumstances have an impact on air quality. $PM_{2.5}$ was most closely associated with wind speed. O_3 was most closely related to relative humidity and temperature. Giri *et al*. [23] employed the SPSS statistical software programme to determine the association. To demonstrate the statistical significance of differences, inferential statistical t-tests, ANOVAs, and the Tukey Honeslty Significantly Different (HSD) technique, as well as the Levene test, were used. During the correlation finding, the authors stated that wind did not blow pollutants out of the studied region, but rather delivered pollutants into the valley, showing no dilution impact. PM_{10} was favorably connected to wind speed, whereas atmospheric pressure was negatively related to precipitation and relative humidity. According to the authors, wind and relative humidity were the most critical meteorological characteristics impacting air quality behavior. The authors discovered that rainfall, humidity, and wind speed are the most important elements impacting PM_{10} concentration in the studied area by applying linear correlation. The direction of the wind did not influence the change in PM_{10} concentration. Wang and Ogawa [24] examined the relationship between $PM_{2.5}$ and meteorological data using linear and Spearman analyses. The SPSS statistical programme was utilized in the investigation. Temperature and $PM_{2.5}$ were substantially associated. For most months, $PM_{2.5}$ was substantially negatively connected with humidity, and for the rest of the months was favorably correlated with other parameters however with very low correlation coefficients. In the summer, the humidity was highly negatively linked with $PM_{2.5}$, but in the fall, relationships with all pollutants were negative. The correlation of $PM_{2.5}$ with each investigated meteorological parameter was established for each season, which was spring, summer, autumn, and winter. Temperature exhibited a negative link with $PM_{2.5}$, but precipitation had a positive correlation concerning wind direction. The authors discovered that the west wind might carry the greatest contaminants to the research location. Jayamurugan *et al*. [25] investigated the effect of relative humidity with temperature on ambient SO_2, NOx, RSPM, and SPM concentrations using regression analysis. According to the study's findings, both SO_2 and NOx were inversely connected with temperature and moderately and favorably correlated during the post-monsoon season. Except during the post-monsoon season, RSPM and SPM exhibited a positive connection with temperature. The results suggested that the effects of temperature on SO_2 & NOx concentration were considerably more effective in summer than in other seasons, while the correlation was found to be paradoxical in the case of particles. In all the seasons, negative association between humidity and particulates were found, but the moderate correlation was found only during the monsoon season.

In the study of Habeebullah *et al*. [26], for statistical analysis, the authors employed the R programming language. SO_2 and PM_{10} exhibited a poor relationship with NOx species, whereas O_3 had a substantial relationship with temperature but a negligible connection with NOx. Wind speed correlated positively with O_3 and PM_{10} and negatively with other pollutants. The pressure exhibited a modest relationship with most pollutants and a negative relationship with O_3 and PM_{10}. The relationship between relative humidity and O_3 was substantial. Temperature correlated negatively with relative humidity and positively with O_3. In terms of emission sources, rainfall was only slightly associated with various air contaminants.

III. Data and Method

A. Data Study

In recent years, the rapid growth of the country's economy also witnessed the pollution problem which has made India 3^{rd} on the list of most polluted countries in the world [27], and Delhi is one of the most polluted cities, followed by Lahore and Santiago [28]. However, depending on the geographical area covered by India, we can say that India is the most polluted among all. Delhi is India's capital state, located in the country's north-central region. Delhi, located around 160 km south of the Himalayas, is situated on the banks of the Yamuna River, which flows into the Ganga River. The neighboring urban areas and adjoining agricultural territories of Old Delhi and New Delhi make up the national capital territory. The state of Uttar Pradesh borders the area to the east, while Haryana borders the territory to the north, west, and south.

B. Data Collection

The dataset of air pollutant concentration and meteorological conditions in India from 1 March 2019 to 1 March 2021 was obtained from CPCB, Delhi. Seven air pollutants including SO_2, NO, NO_2, O_3, PM_{10}, $PM_{2.5}$, CO, and six meteorological Parameters including Wind Speed (WS), Wind Direction (WD), Solar Radiation (SR), Pressure (mentioned as BP), Atmospheric Temperature (AT), Relative

Humidity (RH) were collected from 27 monitoring station. Dataset availability for the study is as follows:

Data Set: Air pollution and meteorological data of Delhi, India 2019 to 2021.

URL: https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/data

C. Dataset Description

Several monitoring stations are located around the NCR to monitor air quality. The monitoring is carried out by the CPCB, DPCC, and SAFAR of IITM, Pune. NAMP of CPCB monitors areas like as Sarojini Nagar, Chandni Chowk, Mayapuri Industrial Area, Pitampura, Shahdara, Shahzada Bagh, Nizamuddin, Janakpuri, Siri Fort, and ITO. CAAQM monitors 11 areas including Anand Vihar, Civil Lines, DCE, Dilshad Garden, Dwarka, IGI Airport, ITO, Mandir Marg, Punjabi Bagh, R.K. Puram, and Shadipur, whereas DPCC monitors 6 locations including Civil Lines, Punjabi Bagh, Mandir Marg, Anand Vihar ISBT, IGI Airport, and R.K. Puram. From the collected data, NSIT Dwarka and Sirifont station data are maintained by CPCB, Pusa is maintained by DPCC & IMD, and the remaining monitoring stations data are maintained by DPCC. Apart from the CPCB and DPCC, SAFAR has eight monitoring stations located around Delhi to monitor the ambient air quality in real-time. The data collected from these stations is also utilized to calculate the national air quality index.

D. Data Preprocessing

Due to instrumental errors, measurement errors, data transmission errors, and in addition to some other factors, data for several days were missing. Delhi monitoring stations includes 40 stations but due to missing values in in pollutants as well as in meteorological conditions, only completed data stations were considered for evaluation. In the end, 27 stations named Alipur (Site 1), Anand Vihar (Site 2), Ashok Vihar (Site 3), Bawana (Site 4), Dr. karni Singh Shooting range (Site 5), DTU (Site 6), Dwarka (Site 7), Jhangir Puri (Site 8), Jawahar Lal Nehru stadium (Site 9), Major Dhyanchand National Stadium (Site 10) Mandir Marg (Site 11), Mundka (Site 12), Najafgarh (Site 13), Narela (Site 14), Nehru Nagar (Site 15), NSIT Dwarka (Site 16), Okhla (Site 17), Patpar Ganj (Site 18), Punjabi Bagh (Site 19), Pusa (Site 20), R K Puram (Site 21), Rohini (Site 22), Shadipur (Site 23), Shri Auro Bindo Marg (Site 24), Siri Font (Site 25), Sonia Vihar (Site 26), Wazirpur (Site 27) were taken into account for evaluation (see Fig. 1).

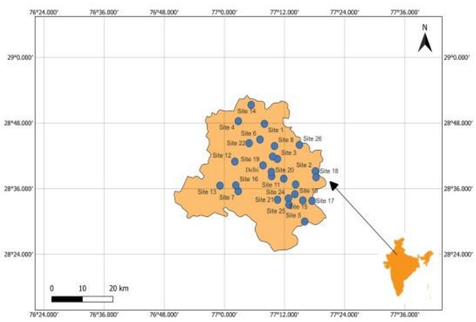


Fig. 1. Selected study location in Delhi, India (X-axis represents the longitude and Y-axis represents the latitude of the considered locations). The figure was generated using QGIS software.

Algorithm 1: Missing Data Handling

Input: A_{list} (D1, D2, ……., Dn , Dk) // training dataset
Output: Anew // the changed data set
Input: D ⟵ set of features with {D1, D2, D3 , ……., Dn, Dk }
where n is the set of features and Dk is the class feature.
A_{list} ⟵ training data set with i instances
A_{miss} ⟵ number of missing data {s1, s2, …..sm}, where m is the number of missing values.
Let T = {t1, t2…., tz} be the instances.
Consider a missing value from Amiss and note its value with A_{list}.
for each missing value, redo the process
Step 1: begin
Step 2: for j = 1 to n do begin
Sort the feature Dj in A along with Dk;
end;
Step 3: for i = 1 to z do begin
If (Ti = = NULL) // to check any missing data
A_{miss} = A_{list};
break;
end;
Step 4: for i = 1 to length(A_{miss}) do begin
append(A_{list}) = getFirstElement(A_{miss})
A_{miss} = ti − ti−1
end;
miss_data = abs(A_{miss} − set)
Step 5: redo Steps 3 and 4 until achieving the goal
Step 6: return Anew
Step 7: end;

Algorithm 2: Outlier Detection

Input: A= 4, outliers = 0, intervals = 0;
Step 1: begin
Step 2: If training_time = True:
intervals +=1
Step 3: Outlier_mean = mean_value (live_interval)
Outlier_std_dev = std_deviation_value (live_interval)
Step 4:
Outlier_max_limit= outlier_mean + outlier_std_dev *A
Outlier_min_limit = outlier_std_mean – outlier_std_dev * A
Step 5:
If outlier_max_limit < live interval < outlier_min_limit:
outlier +=1
Percent_outlier = outlier / intervals
Step 6: If Percent_outlier > 0.03;
A = A * 1.03
Step 7: Else
beacon_interval_cal_picked
Step 8: A = A*0.99
Step 9: If A <3;
A =3
If A >5;
A = 5
Step 10: end

E. Methodology

Our research design procedure is separated into four stages. The initial stage was data preparation as described above. We then did the descriptive analysis to have a better understanding of data, patterns of pollutants and meteorological factors, and seasonal variation impact on air pollutants. In addition to this, correlation and partial correlation analysis were performed on data, and for more depth understanding, one factor was controlled followed by controlling two factors to check if any changes in attribute(s) would be noticed or not. Lastly, AQI calculation was performed for each air pollutant and the relationship between AQI with air pollutants and meteorological factors was identified and visualized.

1) *Correlation*

Correlation or association is used to identify the connection between the two variables. The types of association can be seen in Fig. 2. The negative value indicates changing value in one variable oppositely affects another variable value. The more positive the value, the higher the correlation indicating changes in one variable value affect another variable value as well. If the relationship value is zero then there is no similarity between the two variables means one variable value changes does not bother the other variable value at all.

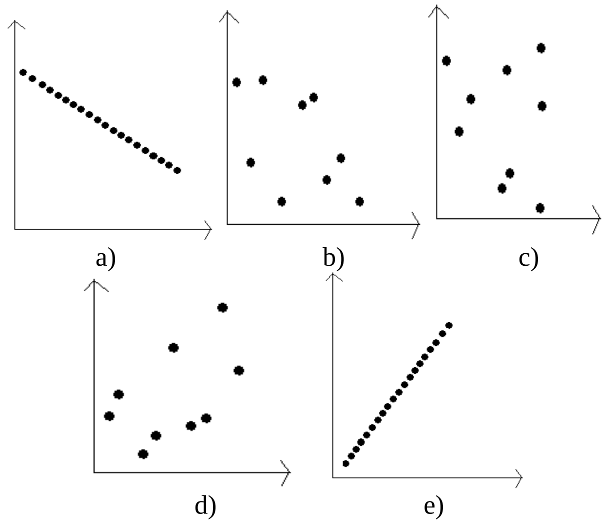


Fig. 2. a) Strong (−)ve association, b) Weak (−)ve association, c) No association, d) Weak (+)ve association, e) Strong (+)ve association.

The Pearson coefficient is used to measure the degree of two variables' linear connection. Pearson correlation aims to establish the best line of fit between two variables. As a result, it provides the distance between all of these data points and the line of best fit. When the value of 'r' is close to +1 or −1, it signifies that all of the observations are included on or closer to the line of best fit. Closer values of 'r' to '0', data points are closer to the line of greatest fit.

$$r = \sum(s_i - \bar{s})(n_i - \bar{n}) \Big/ \sqrt{\sum(s_i - \bar{s})^2 \sum(n_i - \bar{n})^2} \quad (1)$$

where, r = correlation coefficient, s_i = s variable sample, n_i = n variable sample, $\bar{s}$ = s variable mean value, $\bar{n}$ = n variable mean value.

Algorithm 3: Correlation Coefficient
Step 1: begin Step 2: Input: dataset (s1, s2,……,s_i), (n1, n2,……, n_i) Where i belongs to samples Output: r Step 3: calculate mean of the variables i.e $\bar{s}$, $\bar{n}$ Step 4: To determine the type of variable either positive or negative compute Pearson's correlation coefficient using equation 1. Step 5: end.

2) *Partial correlation*

It is used to describe the relation between two variables in the presence of controlling variables and before solving the below formula, one should determine the zero-order coefficient between all possible pairings of variables (between v1 and v2, v2 and v3, v3 and v1, and so on).

$$r_{v1v2.v3} = (r_{v1v2}) - ((r_{v1v3})(r_{v2v3})) \Big/ \left(\sqrt{1 - r^2_{v1v3}}\right)\left(\sqrt{1 - r^2_{v2v3}}\right) \quad (2)$$

where $r_{v1v2.v3}$ = Partial correlation coefficient between variable v1 and v3 by controlling variable v3.

Algorithm 4: Partial Correlation Coefficient
Step 1: begin Step 2: Input: v_k variables, r_{vivj} Output: r Step 3: To calculate the partial correlation, we must first calculate Pearson's correlation coefficient between all three variables. i.e. compute r_{v1v2}, r_{v1v3}, r_{v2v3} repeat the process for all the possible combinations Step 4: To determine the partial correlation coefficient, use equation 2. Step 5: end.

3) *AQI and its calculation*

Government agencies use the Air Quality Index (AQI) to assess and report air pollution levels to the public. As the AQI rises, a significant portion of the population may suffer serious health consequences. An air sensor and an air pollutant concentration over a predetermined average time are required for AQI measurement. The data is divided into ranges, each with its descriptor, color code, and public health advice. To know the six range classifications (Good, Satisfactory, Moderately Polluted, Poor, Very Poor, and Severe) and the related health implications, one should understand how the AQI works. Below Table I shows the AQI range and the health effects depending on that range.

The Indian CPCB and the US-EPA use identical formula to determine AQI. The AQI can be calculated independently for each pollutant (formula given below). To compute the AQI, at least three factors must be used, one of which must be PM_{10} or $PM_{2.5}$. 16 hours of data are required to construct sub-indices.

TABLE I: AQI RANGE, DESCRIPTION, AND COLOR

Air Quality Index Levels	Numerical Value	Meaning	Color
Good	0–50	Minimal consequence	Green
Satisfactory	51–100	Sensitive persons may experience mild breathing difficulties.	Yellow
Moderately Polluted	101–200	People with heart disease, and lung diseases like asthma may experience breathing difficulties.	Orange
Poor	201–300	People with cardiac issues may face difficulty in breathing as a result of prolonged exposure.	Red
Very Poor	301–400	Long-term exposure may cause respiratory disease. People with lung and heart disorders may feel the effects more strongly.	Purple
Severe	401–500	Even individuals can have respiratory effects, and lung/heart disease might have disastrous results. A small amount of exercise might have a detrimental impact on quality of life.	Maroon

$$Xp = [Qmax – Qmin / CPmax – CPmin] (Ap – CPmin) + Qmin \quad (3)$$

where, Xp = Pollutant Index, Ap = Abbreviated pollutant concentration, CPmax = Pollutant cut-off point i.e., greater or equal to Ap, CPmin = Pollutant cut-off point i.e., less than or equal to Ap, Qmax = AQI value equivalent to CPmax, Qmin = AQI value equivalent to CPmin.

Algorithm 5: AQI Calculation

Step 1: Begin

Step 2: Input: SO_2, NO_2, CO, O_3, NO, $PM_{2.5}$, PM_{10}

// an average 3 pollutants are required among all $PM_{2.5}$ or/and PM_{10} is necessary

Output: Xp.

Step 3: Calculate the sub-indices of each pollutant by using equation 3.

// 24-hour average value for air pollutants

//8 hours for CO and O_3

Step 4: Each air pollutant sub-index contributes to providing the status of current air quality.

Step 5: End.

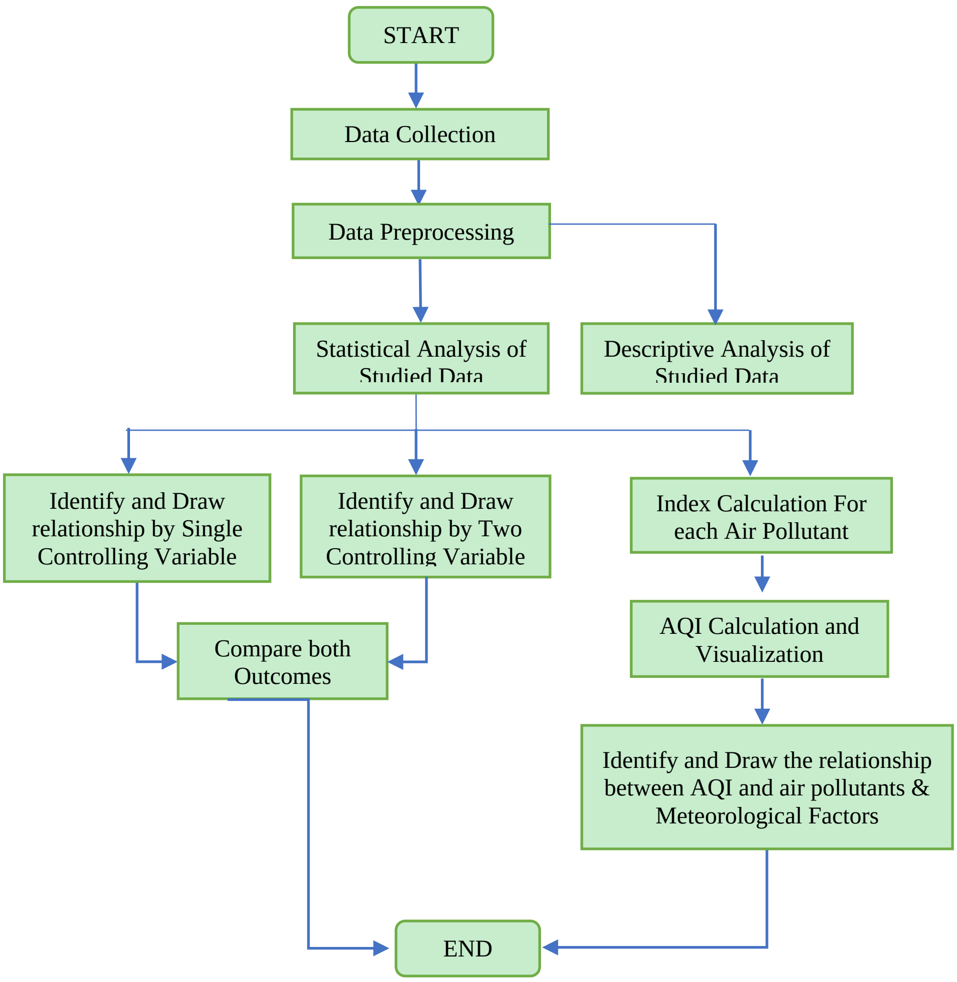


Fig. 3. Workflow of proposed work.

4) Implementation work

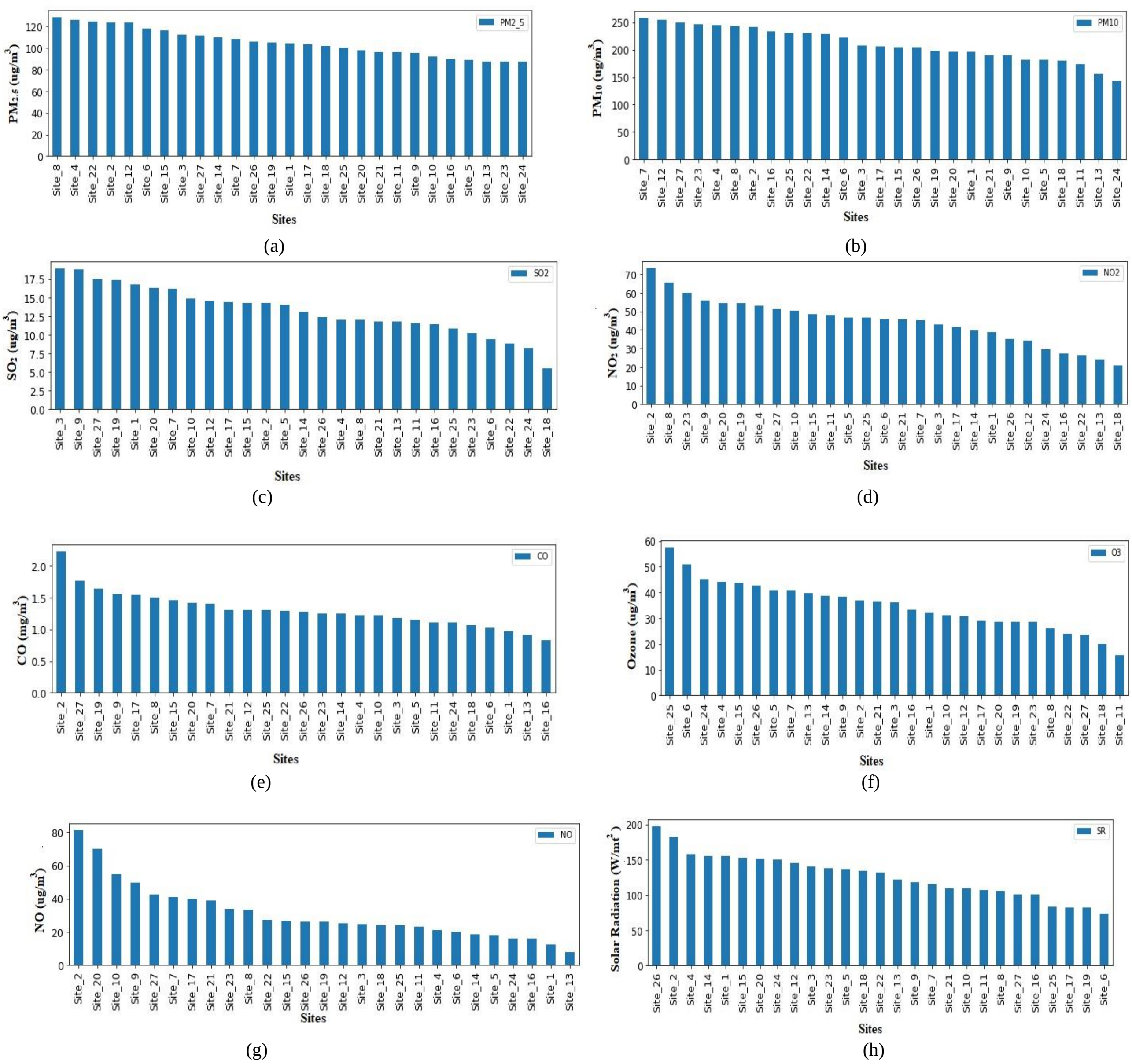
The workflow mentioned in Fig. 3 shows the steps taken for the work. The proposed method started by downloading the air pollutants concentration and meteorological data from the CPCB website. As the extracted data was having some missing and outlier elements which were dealt with in the data preprocessing stage. To have a better understanding of the data, descriptive analysis was performed. Statistical analysis was done by identifying the relationship between air pollutants and meteorological data. First, the relationship was identified by controlling a single variable followed by controlling two variables. Then a comparison between both the techniques was discussed for better understanding. Index calculation known as AQI calculation was also performed then the relationship between AQI with each air pollutant and meteorological factors was done.

IV. Data Analysis

A. Descriptive Analysis

1) Trend change analysis of pollutants and meteorological parameters over city

Fig. 4 shows that each city has its unique concentration of air contaminants and meteorological value. The greatest and lowest concentrations of each pollutant and a climatic component can be observed during a research period. For SO_2, Ashok Vihar (Site 3) has the highest concentration while Patparganj (site 18) has the lowest. The highest concentration of $PM_{2.5}$ was seen in Jahangirpuri (Site 8). Anand Vihar (Site 2) has the highest concentrations of CO, NO, and NO_2. Dwarka (Site 7) station had the highest PM_{10} concentration. Sirifort (Site 25) had the highest O_3 concentration. Apart from air pollutants concentration over different areas, Meteorological parameters were also utilized. Many stations noticed the maximum value of pressure and atmospheric temperature. Solar radiation was high in Sonia Vihar station (Site 26). Shri Aurobindo Marg (site 24) noticed the maximum value of wind speed.



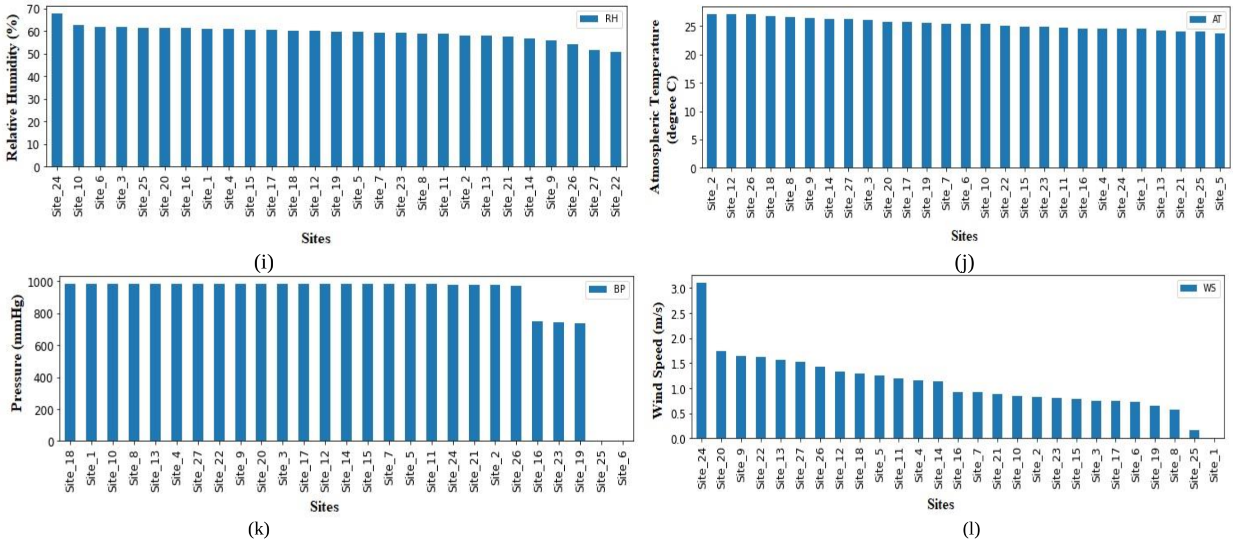


Fig. 4. Trend change of air pollutants and meteorological factors depending on the location.

2) Seasonal variation of pollutant parameters

Our data was categorized into four seasons: spring, summer, autumn, and winter. March to May is Spring, June to August is Summer, September to November is Autumn, and December to February is Winter. In Fig. 5, It can be seen that $PM_{2.5}$ in winter accounted was 45.1% of the total average concentration of $PM_{2.5}$ in Delhi whereas the Autumn, Spring, Summer accounted 26.1%, 18.4%, and 10.4% of the total respectively. CO concentration in winter was 33% followed by 27.8%, 19.8%, and 19.3% of total CO concentration in Delhi. O_3 concentration in spring was 36%, Where on the other hand Autumn, Summer, and Winter witnessed 24.5%, 20.6%, and 18.9% of the total concentration respectively. The high PM_{10} concentration was noticed in winter where the lowest was in summer and 28.5% in autumn, and 23.3% in spring. SO_2 concentration was high in spring, i.e., 32.6% followed by 24.1% in Autumn, 23.6% in winter, and 19.7% in summer. NO_2 concentration was witnessed high in winter, i.e., 32%, 28.2% in autumn, 23.4% in spring, and 16.4% in summer. In winter, NO concentration was noticed highest i.e., 49.6% of the total, and lowest in summer i.e., 8.99%. NO autumn concentration was 28.8% in autumn and 12.5% in spring. The high value of SR in spring accounted for 34.3% of the overall average value of Delhi, while summer, autumn, and winter accounted for 29.4%, 20.8%, and 15.5% of the total, respectively. In winters, BP value was 25.2%, with identical percent values in autumn and spring, i.e., 25% and 25%, with 24.8% of the total in summer. Summer had the highest AT value at 31.2% of the total, while spring, autumn, and winter had 27.6%, 26.4%, and 14.8% of the total value, respectively. In summer, WS witnessed highest percent value followed by spring, winter and autumn i.e., 28.8%, 21.8% and 20.2%. The highest percent value of RH was seen in winter, while the lowest was observed in spring, precisely 29.1%, 19%, 26.5% in summer, and 25.4% in autumn.

3) AQI visualization

As we collected data for the period of 2019 to 2021 from CPCB website. The data period has two categories: Pre-Covid data and Covid Period data. As shown in Fig. 6, India's AQI in 2019 ranged between 145 to 150, indicating a drop in 2020. The graph also shows a dramatic surge in AQI.

Fig. 5. Seasonal variation of air pollutants and meteorological factors.

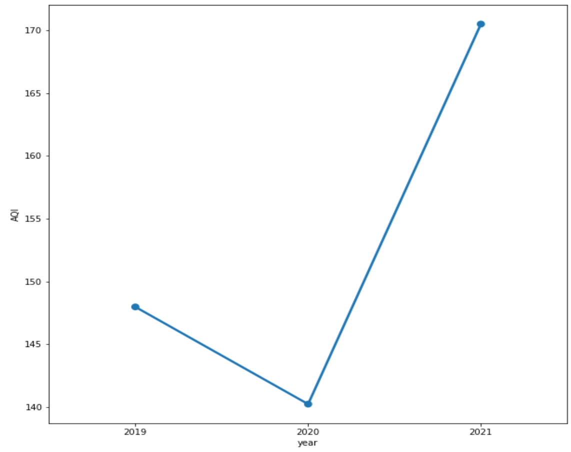


Fig. 6. Change in AQI over the study period.

B. Statistical Analysis

1) Correlation between air pollutants & meteorological factors

In the below-shown Fig. 7, we can see the relationship correlation map where darker in blue color, negative the relationship and darker the maroon, positive the association between pollutants and weather data. The relationship was quantified using spearman-rank correlation coefficient. The relationship between all pollutants and six weather data throughout the study period was calculated. In Delhi, $PM_{2.5}$, PM_{10}, CO, NO_2, and NO concentration was negatively correlated whereas SO_2, and O_3 were positively correlated with Atmospheric Temperature. $PM_{2.5}$, NO, CO, SO_2, and O_3 every pollutant except PM_{10} and NO_2 were positively correlated with BP. $PM_{2.5}$, PM_{10}, CO, NO_2, all the pollutants except SO_2 and O_3 were negatively correlated with Solar Radiation (SR), whereas on the other hand, all the pollutants except SO_2 were positive correlated with Wind Direction (WD), where SO_2 was not showing any relationship with WD. All pollutants except O_3 were negatively correlated with Wind Speed (WS). $PM_{2.5}$, CO, and NO showed a positive correlation with Relative Humidity (RH) whereas PM_{10}, NO_2, SO_2, and O_3 showed a negative correlation.

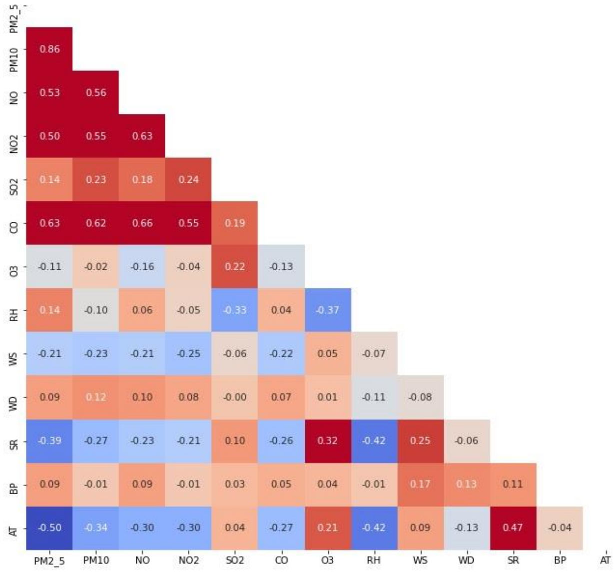


Fig. 7. Correlation map to show correlated data.

2) Partial correlation between air pollutants and weather data

In Fig. 8, the partial correlation between every meteorological factor with other meteorological factors can be noticed. The Figure states, the darker the color, the positive the relationship will be and the lighter in color, the negative the relationship will be. It can be seen that $PM_{2.5}$, NO, NO_2, and SO_2 were negatively correlated whereas PM_{10}, CO, and O_3 were positively correlated with Atmospheric Temperature (AT). $PM_{2.5}$, NO, SO_2, CO, and O_3 were positively correlated with BP whereas PM_{10} and NO_2, had a negative correlation. NO, SO_2 and O_3 showed a correlation with SR but showed a negative correlation in the case of $PM_{2.5}$, PM_{10}, CO, and NO_2. WD showed a positive correlation with PM_{10}, NO, and O_3 but a negative correlation with $PM_{2.5}$, NO_2, SO_2, and CO. every pollutant was a negative correlation with Wind Speed (WS). In Relative Humidity (RH) relationship with pollutants $PM_{2.5}$, NO and CO showed positive correlation and negative correlation with PM_{10}, NO_2, SO_2, and O_3.

3) Controlling single meteorological factor

Controlling one pollutant variable can impact differently on another pollutant factor. The value of covariance will differ with pollutant variables while controlling a single pollutant variable as compared with direct correlation, the resultant relationship of pollutant variables changes accordingly. We have experimented with this technique to study the impact on pollutant variables (see Table II).

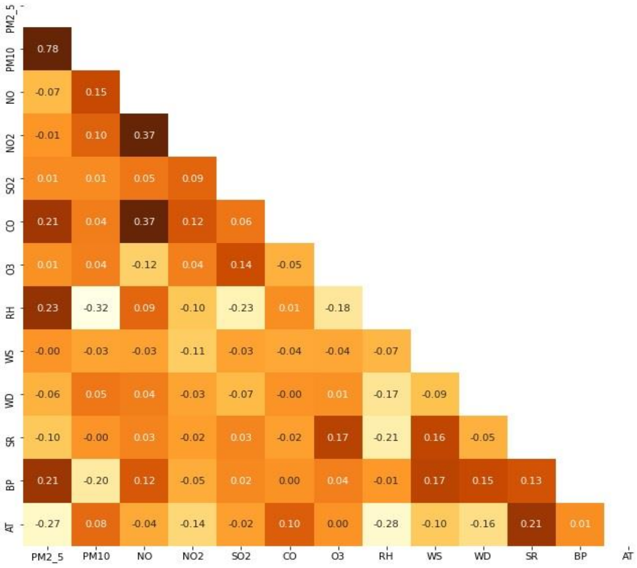


Fig. 8. Correlation map to show partial correlated data.

TABLE II: SINGLE CONTROLLING FACTOR EFFECT ON THE RELATIONSHIP

Meteorological Factors As Controlling Variable		**Air Pollutants**						
Var 1	**Var 2**	**$PM_{2.5}$**	**PM_{10}**	**NO**	**NO_2**	**SO_2**	**CO**	**O3**
WS	**WD**	0.048	0.060	0.052	0.065	0.004	0.053	0.003
	BP	0.082	0.065	0.076	0.087	0.005	0.079	0.016
	AT	0.285	0.161	0.132	0.143	0.006	0.116	0.046
	SR	0.171	0.104	0.080	0.088	0.015	0.093	0.107
	RH	0.062	0.064	0.047	0.067	0.115	0.051	0.139
WD	**BP**	0.015	0.015	0.016	0.018	0.000	0.009	0.003
	AT	0.259	0.127	0.099	0.092	0.002	0.076	0.047
	SR	0.163	0.089	0.066	0.052	0.011	0.074	0.105
	RH	0.031	0.022	0.016	0.007	0.109	0.008	0.140
	WS	0.066	0.094	0.054	0.022	-0.004	0.055	0.087
BP	**AT**	0.264	0.124	0.103	0.101	0.003	0.073	0.057
	SR	0.178	0.082	0.080	0.057	0.010	0.085	0.179
	RH	0.028	0.011	0.007	0.010	0.116	0.002	0.163
	WD	0.077	-0.023	0.043	0.042	0.026	-0.020	0.072
	WS	0.133	0.035	0.11	0.02	0.036	0.043	0.142
AT	**SR**	0.293	0.138	0.108	0.099	0.011	0.100	0.109
	RH	0.264	0.193	0.100	0.126	0.127	0.081	0.150
	WS	-0.495	-0.326	-0.26	0.207	0.046	-0.29	-0.297
	WD	-0.496	-0.325	-0.263	0.214	0.041	-0.292	-0.294
	BP	-0.5	-0.339	-0.259	0.231	0.037	-0.295	-0.309
SR	**RH**	0.159	0.135	0.060	0.071	0.109	0.076	0.174
	WS	-0.359	-0.228	-0.206	0.322	0.106	-0.16	-0.183
	BP	-0.408	-0.272	-0.273	0.414	0.086	-0.21	-0.265
	AT	-0.206	-0.137	-0.154	0.257	0.095	-0.087	-0.106
	WD	-0.39	-0.267	-0.255	0.322	0.103	-0.21	-0.227
RH	**WS**	0.136	-0.11	0.028	-0.365	-0.333	-0.067	0.051
	WD	0.154	-0.084	0.051	-0.369	-0.329	-0.038	0.073
	BP	0.14	-0.098	0.032	-0.398	-0.328	-0.063	0.063
	AT	-0.083	-0.275	-0.079	-0.316	-0.34	-0.195	-0.075
	SR	-0.025	-0.24	-0.074	-0.273	-0.314	-0.152	-0.04

Note: Var 1 is the Controlling Factor and the Relationship between Var 2 and Air Pollutants is Identified

- Windspeed as Controlling Factor: When Windspeed was controlled, then the effect of wind direction on NO_2 was higher compared to all and lighter on O_3. BP was affecting NO_2 much and less on SO_2. AT was affecting $PM_{2.5}$ much but SO_2 was very less. SR was affecting SO_2 very less but $PM_{2.5}$ on a big scale. RH affected NO very less but O_3 in a big amount.
- Wind Direction as Controlling Factor: When Wind direction was controlled, then BP affected NO_2 quite much but very less SO_2. AT affected $PM_{2.5}$ much but SO_2 less. SR affected $PM_{2.5}$ much but SO_2 on less scale. RH affected PM_{10} much but affected NO_2 very less.
- Pressure as Controlling Factor: When Pressure was controlled to identify the relation between a single meteorological factor and each air pollutant then AT affected $PM_{2.5}$ much but SO_2 very less. SR affected O_3 much but SO_2 very less. RH affected O_3 much but CO very less. WD affected $PM_{2.5}$ much and PM_{10} & CO_2 negatively. WS affected O_3 much and NO_2 very less.
- AT as controlling Factor: SR affected $PM_{2.5}$ much and SO_2 very less. In RH, the effect was seen much as on $PM_{2.5}$ and less on CO. WS, WD and BP affected NO_2 much and SO_2 very less. WS, BP, AT, and WD affected NO_2 and SO_2 much and negatively affected the remaining pollutants.
- SR as Controlling Factor: RH affected $PM_{2.5}$ much and NO very less. WS, BP, AT and WD affected NO_2, and SO_2 in a positive way and the remaining pollutants were negatively affected by the same.
- RH as Controlling Factor: WS, WD, and BP positively affected $PM_{2.5}$ and the remaining pollutants were negatively affected. AT and SR were affecting all the considered pollutants negatively.

4) Controlling two meteorological factors

We can see the relationship of single meteorological factor with each air pollutant by controlling single variable in previous section. To further understand the connection, most possible combinations of meteorological factors were made and then after controlling those combinations, the relationship between each air pollutant with each meteorological factor was identified.

a) Combination of wind speed with other possible variables

The maximum possible combinations of wind speed with other meteorological variables were WS-WD, WS-BP, WS-AT, WS-SR, and WS-RH (see Table III).

- Wind Speed-Wind Direction: After controlling WS and WD, the BP effect on $PM_{2.5}$ was positively high and negatively on O_3. AT affected O_3 and SO_2 positively, whereas the remaining were negatively affected. RH affected $PM_{2.5}$, CO and NO positively and PM_{10}, O_3, SO_2, and NO_2 were negatively affected. SR was positively correlated with O_3, and SO_2 whereas the remaining were negatively correlated.
- Wind Speed-Pressure: WD was negatively correlated with O_3 and positively correlated with the remaining ones. AT was positively correlated with O_3 & SO_2 and the remaining were negatively correlated. RH noticed a negative correlation with every air pollutant except for NO. SR was having positive correlation except for NO. SR was having a positive correlation with O_3 and SO_2. And for the remaining one's negative correlation was noticed.
- Wind Speed-Atmospheric Temperature: After controlling WS with AT, WD and BP noticed a positive correlation with every single considered air pollutant. Where on the other hand, RH noticed a negative correlation with every air pollutant. SR was positively correlated with O_3, and SO_2 and was neutral with NO_2.
- Wind Speed-Solar Radiation: WD and BP noticed a positive correlation with every single pollutant. AT noticed a negative correlation with every pollutant except O_3 and RH noticed negative correlation with every pollutant except for NO
- Wind Speed- Relative Humidity: WD noticed a positive correlation with every pollutant except O_3 and SO_2. BP noticed negative correlation with O_3 and positive correlation with remaining pollutants. AT and SR were affecting all the pollutants negatively except O_3.

TABLE III: TWO CONTROLLING FACTORS EFFECTS ON THE RELATIONSHIP (CONSIDERED WIND SPEED AS ONE CONTROLLING VARIABLE WITH THE COMBINATION OF OTHERS)

Meteorological Factors As Controlling Variable			Air Pollutants					
Var 1	Var 2	Var	$PM_{2.5}$	PM_{10}	CO	O_3	SO_2	NO_2	NO
WS	WD	BP	0.461	0.287	0.233	−0.063	0.102	0.216	0.331
		AT	−0.526	−0.307	−0.246	0.199	0.017	−0.284	−0.369
		RH	0.025	−0.171	0.046	−0.421	−0.354	−0.107	0.108
		SR	−0.328	−0.192	-0.184	0.336	0.105	−0.137	−0.218
WS	BP	WD	0.084	0.101	0.046	−0.012	0.001	0.074	0.11
		AT	−0.393	−0.205	−0.139	0.186	0.076	−0.204	−0.273
		RH	−0.066	−0.24	−0.01	−0.423	−0.387	−0.174	0.049
		SR	−0.263	−0.14	−0.155	0.381	0.118	−0.101	−0.175
WS	AT	WD	0.075	0.096	0.028	0.028	0.007	0.046	0.084
		BP	0.296	0.182	0.152	0.032	0.127	0.111	0.199
		RH	−0.265	−0.358	−0.07	−0.378	−0.378	−0.265	−0.071
		SR	−0.1	−0.05	−0.077	0.28	0.11	0	−0.048
WS	SR	WD	0.121	0.124	0.051	0.015	0.008	0.073	0.117
		BP	0.437	0.272	0.208	0.021	0.127	0.205	0.313
		AT	−0.453	−0.26	−0.188	0.044	−0.038	−0.257	−0.319
		RH	−0.143	−0.292	−0.04	−0.329	−0.341	−0.189	0.004
WS	RH	WD	0.125	0.107	0.06	−0.05	−0.04	0.063	0.133
		BP	0.478	0.335	0.242	−0.015	0.161	0.251	0.342
		AT	−0.578	−0.434	−0.257	0.034	−0.147	−0.37	−0.374
		SR	−0.356	−0.298	−0.186	0.199	−0.045	−0.204	−0.2

b) Combination of wind direction with other possible variables

The maximum possible combinations of wind speed with other meteorological variables were WD-BP, WD-AT, WD-SR, and WD-RH (see Table IV).

TABLE IV: TWO CONTROLLING FACTORS EFFECT ON THE RELATIONSHIP (CONSIDERED WIND DIRECTION AS ONE CONTROLLING VARIABLE WITH THE COMBINATION OF OTHERS)

Meteorological Factors As Controlling Variable			**Air Pollutants**						
Var 1	**Var 2**	**Var**	**$PM_{2.5}$**	**PM_{10}**	**CO**	**O_3**	**SO_2**	**NO_2**	**NO**
WD	BP	WS	−0.401	−0.346	−0.389	0.08	-0.11	−0.417	−0.433
		AT	−0.43	−0.253	−0.208	0.198	0.053	−0.263	−0.326
		RH	−0.008	−0.179	0.047	−0.433	-0.374	−0.097	0.106
		SR	−0.349	−0.228	−0.265	0.384	0.091	−0.211	−0.291
WD	AT	WS	−0.257	−0.258	−0.265	−0.034	−0.05	−0.322	−0.314
		BP	0.221	0.115	0.066	0.044	0.118	0.036	0.093
		RH	−0.236	−0.326	−0.044	−0.377	−0.38	−0.223	−0.037
		SR	−0.166	−0.122	−0.16	0.257	0.105	−0.093	−0.15
WD	SR	WS	−0.197	−0.226	−0.23	−0.117	−0.079	−0.3	−0.275
		BP	0.38	0.214	0.134	0.018	0.121	0.137	0.217
		AT	−0.437	−0.241	−0.17	0.047	−0.048	−0.237	−0.285
		RH	−0.134	−0.283	−0.04	−0.33	−0.339	−0.176	0.004
WD	RH	WS	−0.293	−0.311	−0.288	−0.056	−0.093	−0.361	−0.336
		BP	0.413	0.276	0.172	0.006	0.167	0.187	0.247
		AT	−0.576	−0.437	−0.268	0.019	−0.17	−0.376	−0.37
		SR	−0.405	−0.359	−0.262	0.168	−0.064	−0.284	−0.283

- Wind Direction-Pressure: WS noticed a negative correlation with every pollutant except O_3. AT noticed a negative correlation with O_3 & SO_2 and a positive correlation with the remaining ones.
- Wind Direction-Atmospheric Temperature*:* WS noticed a negative correlation with all the considered air pollutants. BP noticed a positive correlation with all the pollutants. SR noticed a positive correlation with O_3 & SO_2 and a negative for the remaining pollutants.
- Wind Direction-Solar Radiation: WS noticed negative correlation with all the pollutants where on the other hand BP noticed positive correlation with every pollutant. AT noticed negative correlation with every pollutant except for O_3. RH on the other side, noticed negative correlation with every pollutant except NO. Wind Direction- Relative
- Humidity: WS noticed negative correlation and BP noticed positive correlation with every pollutant. AT and SR noticed negative correlation with every pollutant except O_3.

c) Combination of pressure with other possible variables

The maximum possible combinations of wind speed with other meteorological variables were BP-AT, BP-SR, and BP-RH (see Table V).

- Pressure-Atmospheric Temperature*:* WS noticed a negative correlation with every pollutant except O_3. WD noticed a positive correlation except for O_3. RH noticed a negative correlation with every pollutant. SR noticed a negative correlation with every pollutant except O_3 and SO_2.
- Pressure-Solar Radiation: WS and RH noticed a negative correlation with every single pollutant. WD noticed a negative correlation with O_3 and a positive correlation with remaining pollutants. AT noticed a negative correlation with every pollutant except O_3 and SO_2.
- Pressure-Relative Humidity: WS noticed a negative correlation with every pollutant except O_3. WD noticed a positive correlation with every pollutant except O_3 and SO_2. AT noticed a negative correlation with every pollutant except O_3. SR noticed a negative correlation with every pollutant except O_3. Here O_3 was having opposite relation condition in comparison to other pollutant behavior.

TABLE V: TWO CONTROLLING FACTORS EFFECT ON THE RELATIONSHIP (CONSIDERED PRESSURE AS ONE CONTROLLING VARIABLE WITH THE COMBINATION OF OTHERS)

Meteorological Factors As Controlling Variable			**Air Pollutants**						
Var 1	**Var 2**	**Var**	**$PM_{2.5}$**	**PM_{10}**	**CO**	**O_3**	**SO_2**	**NO_2**	**NO**
BP	AT	WS	−0.351	−0.31	−0.362	0.039	−0.124	−0.384	−0.395
		WD	0.084	0.106	0.055	−0.006	0.006	0.073	0.113
		RH	−0.238	−0.335	−0.053	−0.386	−0.383	−0.243	−0.053
		SR	−0.182	−0.125	−0.193	0.336	0.075	−0.101	−0.162
BP	SR	WS	−0.333	−0.302	−0.341	−0.041	−0.14	−0.381	−0.384
		WD	0.103	0.118	0.065	−0.013	0.004	0.086	0.128
		AT	−0.325	−0.176	−0.101	0.022	0.011	−0.194	−0.229
		RH	-0.19	−0.319	−0.08	−0.321	−0.367	−0.215	−0.035
BP	RH	WS	−0.407	−0.378	−0.39	0.037	−0.163	−0.437	−0.43
		WD	0.1	0.095	0.073	−0.084	−0.052	0.073	0.141
		AT	−0.483	−0.374	−0.215	0.033	−0.113	−0.341	−0.324
		SR	−0.391	−0.341	−0.274	0.256	−0.069	−0.28	−0.28

d) Combination of atmospheric temperature with other possible variables

The maximum possible combinations of wind speed with other meteorological variables were AT-SR and AT-RH (see Table VI).

TABLE VI: TWO CONTROLLING FACTORS EFFECT ON THE RELATIONSHIP (CONSIDERED ATMOSPHERIC TEMPERATURE AS ONE CONTROLLING VARIABLE WITH THE COMBINATION OF OTHERS)

Meteorological Factors As Controlling Variable			Air Pollutants						
Var 1	**Var 2**	**Var**	**$PM_{2.5}$**	**PM_{10}**	**CO**	**O_3**	**SO_2**	**NO_2**	**NO**
AT	SR	WS	−0.227	−0.241	−0.238	−0.118	−0.08	−0.314	−0.296
		WD	0.099	0.12	0.051	0.028	0.009	0.067	0.111
		BP	0.245	0.136	0.086	0.024	0.116	0.051	0.117
		RH	−0.308	−0.389	−0.1	−0.329	−0.363	−0.267	−0.101
AT	RH	WS	−0.28	−0.294	−0.27	−0.055	−0.07	−0.345	−0.322
		WD	0.054	0.061	0.041	−0.043	−0.064	0.026	0.1
		BP	0.215	0.101	0.068	0.007	0.09	0.022	0.101
		SR	−0.248	−0.234	−0.181	0.175	0.004	−0.165	−0.17

- Atmospheric Temperature-Solar Radiation*:* WS and RH noticed a negative correlation with everyone whereas, on the other side, WD and BP noticed a positive correlation with all pollutants.
- Atmospheric Temperature-Relative Humidity: WS noticed a negative correlation and BP noticed a positive correlation with every pollutant. WD noticed a positive correlation for every pollutant except O_3 and SO_2. SR noticed a negative correlation with everyone except O_3 and SO_2.

e) Combination of solar radiation with other possible variables

The maximum possible combinations of wind speed with other meteorological variables were SR-RH (see Table VII).

- Solar Radiation-Relative Humidity: WS and AT noticed a negative correlation with every pollutant. WD noticed a positive correlation with every considered pollutant except O_3 and SO_2. BP noticed a positive correlation with every pollutant.

TABLE VII: TWO CONTROLLING FACTORS EFFECT ON RELATIONSHIP (CONSIDERED SOLAR RADIATION AS ONE CONTROLLING VARIABLE WITH THE COMBINATION OF OTHERS)

Meteorological Factors As Controlling Variable			Air Pollutants						
Var 1	**Var 2**	**Var**	**$PM_{2.5}$**	**PM_{10}**	**CO**	**O_3**	**SO_2**	**NO_2**	**NO**
SR	RH	WS	−0.202	−0.234	−0.233	−0.112	−0.071	−0.304	−0.283
		WD	0.121	0.107	0.065	−0.028	−0.038	0.068	0.139
		BP	0.407	0.258	0.149	0.031	0.149	0.164	0.238
		AT	−0.508	−0.359	−0.197	−0.044	−0.15	−0.309	−0.311

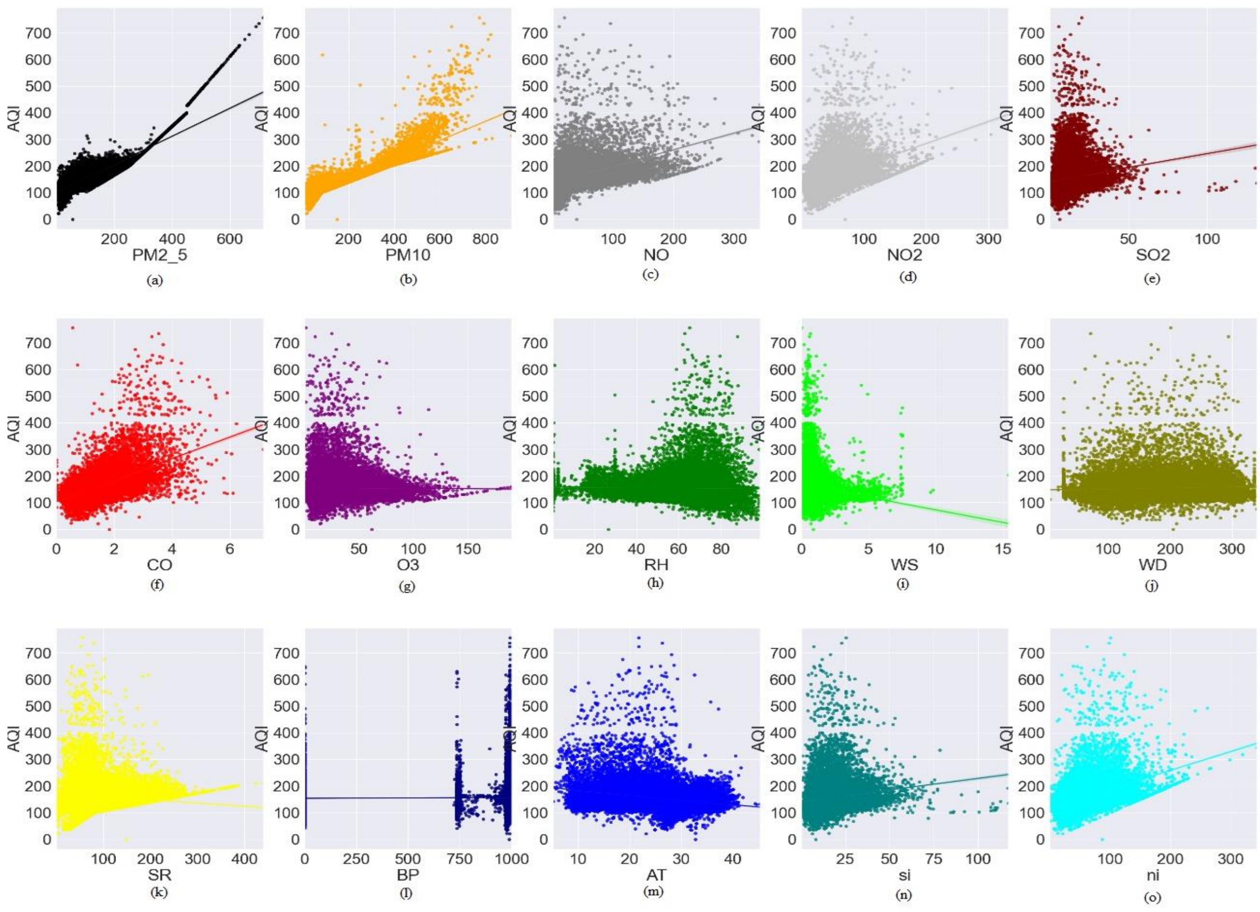


Fig. 9. Correlation of each pollutant with AQI (X-axis represents pollutant factors and Y-axis represents AQI value).

f) AQI relation with air pollutants and meteorological factors

From Fig. 9, it can be noticed that AQI has a positive correlation with $PM_{2.5}$, PM_{10}, NO, NO_2, SO_2, CO, SR, si, and ni. Whereas on the other hand, it has a negative correlation with AT, WS is slightly near the zero value, i.e, neutral relation was identified for O_3, RH, BP, and WD.

V. Critical Discussion

Meteorological circumstances, without a doubt, have a significant impact on the development of air pollution and variations in pollutant concentration. A detailed analysis of the association between meteorological conditions and air pollutant concentrations is required and serves as the foundation to develop air pollution prevention and control strategies. Even though previous researches have attempted to investigate the link between air pollution and meteorological circumstances, there is still a lack of comprehensive investigations that take into account major air contaminants and meteorological characteristics.

The below-mentioned table (Table VIII) clearly shows that certain studies included only some air pollutants and meteorological factors while ignoring the others (the below table contains the parameters in the columns considered in our study thus the existing studies may have included other parameters for the analysis except the parameters shown in the table). In our investigation, we attempted to take into account crucial contaminants and climatic characteristics that helped us comprehend the association.

The partial correlation coefficient between meteorological characteristics and air pollutants was established in Reference [9] by adjusting one meteorological parameter at a time. When wind speed was held constant, the authors indicated that atmospheric pressure was positively correlated with all pollutants except CO and SO_2. However, in our study, all pollutants were positively correlated with pressure under the same conditions (see Table II). When relative humidity was held constant, the authors found a negative correlation with CO, NO_2, $PM_{2.5}$, and PM_{10} with wind speed in Reference [9], but when the same relationship was identified with relative humidity as the controlling variable in our study, it was negative with PM_{10}, NO_2, SO_2, and CO but positive with $PM_{2.5}$ and O_3. So, to better investigate and comprehend the mechanism of pollutants' interaction with climatic factors, we controlled two variables and discovered the association, as can be shown in the table (Table III – Table VII) and the observations section. The results obtained in our work show similar trends as in previous studies, thus verifying the correctness of our approach. As an enhancement, the current work found some new insights into the correlation between meteorological parameters and air pollutants.

VI. Observations

Meteorological variables have a great impact on the concentration of air pollutants. As in Section IV, at the same time, two meteorological variables were controlled. If we consider the case where WS, WD, and AT are considered together then we can notice in Table III, when WS-WD was controlled then AT was affecting all the pollutants negatively except O_3 and SO_2. When WS-AT was controlled then WD was affecting all the pollutants positively in Table III. Controlling AT and WD together was affecting all the pollutants negatively by WS in Table IV.

When we considered the WS, WD, and RH then controlling WS and WD, RH affected the pollutants negatively except $PM_{2.5}$, CO, and NO in Table III. WS and RH when controlled then WD was affecting O_3, and SO_2 negatively but for remaining pollutants, it was affecting positive in Table III. When controlling RH and WD, WS was affecting all the pollutants negatively in Table IV.

For the combination of WS, WD, and SR. when WS and WD were controlled, SR was affecting all the pollutants negatively except O_3 and SO_2 in Table III. WS and SR when controlled were affected positively by all the pollutants by WD in Table III. SR and WD when controlled then WS was affecting all the pollutants negatively in Table IV.

For the combination of WS, WD, and BP, when WS and WD were controlled, BP was affecting all the pollutants positively except for O_3 in Table III. When BP and WD were controlled, the effect was negative on all the pollutants except O_3 by WS in Table IV. Controlling BP and WS, WD was affecting all the pollutants positively except O_3 in Table III. In this case, we noticed that O_3 was the opposite of the effect made by meteorological factors after controlling the variables.

In the case of WS, WD, and AT, when WS and WD were controlled, AT affected all the pollutants negatively except O_3 and SO_2 in Table III. When WS and AT were controlled, WD was affecting all the pollutants positively in Table III. When AT and WD were controlled, WS was affecting all the pollutants negatively in Table IV.

From the discussion, it is clear that not only does the meteorological factor affects the concentration fluctuation of air pollutants but the presence or absence of variable also affect the pollutant concentration differently. Understanding this relationship can help society to have better decision-making and to have an improved alert system.

TABLE VIII: Air pollutants and Meteorological Parameters Considered by Previous vs Current Work (Tick Indicates the Inclusion of the Parameter)

Reference	Air Pollutants							Meteorological Parameters					
	$PM_{2.5}$	PM_{10}	SO_2	NO	NO_2	CO	O_3	Wind Speed	Wind Direction	Solar Radiation	Pressure	Atmospheric Temperature	Relative Humidity
[6]	✓	✓	✓		✓	✓	✓	✓			✓	✓	
[7]	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓
[8]	✓	✓	✓			✓	✓					✓	✓
[9]	✓	✓	✓		✓	✓	✓	✓			✓	✓	✓
[10]		✓	✓		✓			✓			✓	✓	✓
[11]	✓	✓	✓	✓	✓	✓	✓		✓	✓		✓	
[12]	✓							✓				✓	✓
[13]	✓	✓	✓		✓	✓	✓	✓	✓				✓
[14]	✓		✓			✓	✓					✓	✓
[15]	✓	✓						✓	✓			✓	✓
[16]	✓	✓	✓		✓	✓	✓	✓				✓	✓
[17]	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓
[18]	✓	✓					✓	✓	✓		✓	✓	✓
[19]			✓		✓			✓				✓	✓

[20]	✓	✓	✓		✓	✓	✓	✓				✓	✓
[21]	✓	✓	✓		✓	✓	✓	✓			✓	✓	✓
[22]	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓
[23]		✓						✓	✓		✓	✓	✓
[24]	✓							✓	✓			✓	✓
[25]			✓					✓	✓			✓	✓
[26]		✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
Our Study	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

VII. Conclusion

This paper has analyzed the relationship between air pollutant concentrations ($PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3) with weather data (Temperature, Humidity, Wind Speed, WD, BP, SR) in Delhi, India from 01 march 2019 to 01 march 2021. The present model is easier to comprehend and the outcomes are useful and dependable. The study has found that the effect of one meteorological factor is different on various air pollutants and the influence of a combination of meteorological factors on air pollutants is different. It can also be noticed that depending on the type of meteorological factors the effect on the same pollutant can be the same or different. We focused on controlling factors by controlling single as well as a combination of two meteorological factors. When we identified the relation by controlling one single meteorological factor, we noticed only positive correlations or near to zero when we further drill down the approach by controlling two meteorological factors, we noticed a negative correlation also. So, by controlling more than one single factor, one can understand the relationship in detail. But combining more and more meteorological parameters can make the system complex and can also contribute to increase the computational time. Every single and the combination of meteorological factors has a significant impact on the relationship. This research will help in making a decision system and will help to identify to take the proper action while noticing any drastic change in the meteorological factors or on-air pollutants or maybe both.

The drawn conclusions of our study suggested that before predicting any air pollutant it is suggested to understand the relationship between air pollutants and meteorological factors when the relationship between air pollutants with single meteorological factors was identified, the results were different as compared to the relationship between a combination of two meteorological pollutants with air pollutants. The occurrence of pollutants concentration was different depending upon seasonal variations. Future research should include more variables to address the issues more efficiently. Furthermore, the combination of meteorological factors can be controlled and identification of the relationship between the two can be done but ignorance to the system complexity can lead to the difficulty.

Appendix

Abbreviations
AT: Atmospheric Temperature
BP: Pressure
CAAQM: Continuous Ambient Air Quality Monitoring
CPCB: Central Pollution Control Board
DPCC: Delhi Pollution Control Committee
EEMD: Ensemble Empirical Mode Decomposition
IITM: Indian Institute of Tropical Meteorology
IMF: Intrinsic Mode Functions
NAMP: National Air Quality Monitoring Programme
NCR: National Capital Region
ni: Nitrogen dioxide Index value
PM: Particulate Matter
RH: Relative Humidity
RSPM: Respirable Suspended Particulate Matter
si: Sulfur dioxide Index value
SPM: Suspended Particulate Matter
SR: Solar Radiation
WD: Wind Direction
WS: Wind Speed

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Meenakshi Malhotra performed the data collection, data analysis and wrote the paper. Prof. Inderdeep Kaur Aulakh analyzed, reviewed, and revised the draft into the final paper. Both the authors had contributed to this work and approved the final paper.

Acknowledgment

The authors would like to acknowledge the Central Pollution Control Board (CPCB), Delhi for providing the air quality and meteorological data online.

References

[1] Z. Ma, X. Meng, C. Chen, B. Chao, C. Zhang, and W. Li, "Short-term effects of different PM2.5 ranges on daily all-cause mortality in Jinan, China," *Scitific Reports 2022 121*, vol. 12, no. 1, pp. 1–8, Apr. 2022, doi: 10.1038/s41598-022-09057-4.

[2] Z. Yang *et al.*, "Health effects of Long-term exposure to ambient PM2.5 in Asia-Pacific: A systematic review of cohort studies," *Current Environment Health Reports*, vol. 9, no. 2, pp. 130–151, Jun. 2022, doi: 10.1007/S40572-022-00344-W/FIGURES/5.

[3] A. Tuladhar, P. Manandhar, and K. L. Shrestha, "Assessment of health impact of PM2.5 exposure by using WRF-Chem model in Kathmandu Valley, Nepal," *Frontiers Sustainable Cities*, vol. 3, p. 42, Jun. 2021, doi: 10.3389/FRSC.2021.672428/BIBTEX.

[4] B. Mahapatra, M. Walia, W. R. Avis, and N. Saggurti, "Effect of exposure to PM10 on child health: evidence based on a large-scale survey from 184 cities in India," *BMJ Global Health*, vol. 5, no. 8, Aug. 2020, doi: 10.1136/BMJGH-2020-002597.

[5] S. Sasmita, D. B. Kumar, and B. Priyadharshini, "Assessment of sources and health impacts of PM10 in an urban environment over eastern coastal plain of India," *Environmental Challenges*, vol. 7, p. 100457, Apr. 2022, doi: 10.1016/J.ENVC.2022.100457.

[6] J. Yang, Z. Ji, S. Kang, Q. Zhang, X. Chen, and S. Y. Lee, "Spatiotemporal variations of air pollutants in western China and their relationship to meteorological factors and emission sources," *Environmental Pollution*, vol. 254, p. 112952, Nov. 2019, doi: 10.1016/J.ENVPOL.2019.07.120.

[7] X. Qi, G. Mei, S. Cuomo, C. Liu, and N. Xu, "Data analysis and mining of the correlations between meteorological conditions and air quality: A case study in Beijing," *Internet of Things*, vol. 14, 100127, Jun. 2021, doi: 10.1016/J.IOT.2019.100127.

[8] I. Kayes, S. A. Shahriar, K. Hasan, M. Akhter, M. M. Kabir, and M. A. Salam, "The relationships between meteorological parameters and air pollutants in an urban environment," *Global Journalof Environmental Science amd Management*, vol. 5, no. 3, pp. 265–278, 2019, doi: 10.22034/gjesm.2019.03.01.

[9] Y. Liu, Y. Zhou, and J. Lu, "Exploring the relationship between air pollution and meteorological conditions in China under environmental

governance," *Scientific Reports 2020 101*, vol. 10, no. 1, pp. 1–11, Sep. 2020, doi: 10.1038/s41598-020-71338-7.

[10] Z. Qiao, F. Wu, X. Xu, J. Yang, and L. Liu, "Mechanism of spatiotemporal air quality response to meteorological parameters: A national-scale analysis in China," *Sustainability 2019*, vol. 11, no. 14, p. 3957, Jul. 2019, doi: 10.3390/SU11143957.

[11] G. Battista and R. Vollaro, "Correlation between air pollution and weather data in urban areas: Assessment of the city of Rome (Italy) as spatially and temporally independent regarding pollutants," *Atmospheric Environment*, vol. 165, pp. 240–247, Sep. 2017, doi: 10.1016/J.ATMOSENV.2017.06.050.

[12] Q. Yang, Q. Yuan, T. Li, H. Shen, and L. Zhang, "The relationships between PM2.5 and meteorological factors in China: Seasonal and regional variations," *International Journal of Environmental Research and Public Health*, vol. 14, no. 12, 2017, doi: 10.3390/ijerph14121510.

[13] H. Zhang, Y. Wang, J. Hu, Q. Ying, and X. M. Hu, "Relationships between meteorological parameters and criteria air pollutants in three megacities in China," *Environmental Research*, vol. 140, pp. 242–254, Jul. 2015, doi: 10.1016/J.ENVRES.2015.04.004.

[14] K. Hou and X. Xu, "Evaluation of the influence between local meteorology and air quality in Beijing using generalized additive models," *Atmosphere 2022*, vol. 13, no. 1, p. 24, Dec. 2021, doi: 10.3390/ATMOS13010024.

[15] W. Kliengchuay, A. C. Meeyai, S. Worakhunpiset, and K. Tantrakarnapa, "Relationships between meteorological parameters and particulate matter in Mae Hong Son Province, Thailand," *International Journal of Environmental Research and Public Health*, vol. 15, no. 12, Dec. 2018, doi: 10.3390/IJERPH15122801.

[16] Manju, A., K. Kalaiselvi, V. Dhananjayan, M. Palanivel, G. S. Banupriya, M. H. Vidhya, K. Panjakumar, and B. Ravichandran., "Spatio-seasonal variation in ambient air pollutants and influence of meteorological factors in Coimbatore, Southern India," *Air Quality, Atmosphere & Health*, vol. 11, no. 10, pp. 1179–1189, Dec. 2018, doi: 10.1007/S11869-018-0617-X/TABLES/7.

[17] J. He, S. Gong, Y. Yu, L. Yu, L. Wu, H. Mao, C. Song, S. Zhao, H. Liu, X. Li, and R. Li., "Air pollution characteristics and their relation to meteorological conditions during 2014-2015 in major Chinese cities," *Environtal Pollution*, vol. 223, pp. 484–496, 2017, doi: 10.1016/J.ENVPOL.2017.01.050.

[18] X. Fang, Q. Fan, H. Li, Z. Liao, J. Xie, and S. Fan, "Multi-scale correlations between air quality and meteorology in the Guangdong−Hong Kong−Macau Greater Bay Area of China during 2015–2017," *Atmospheric Environment*, vol. 191, pp. 463–477, Oct. 2018, doi: 10.1016/J.ATMOSENV.2018.08.018.

[19] S. Deswal and P. Chandna, "Effects of meteorological conditions on concentration of air pollutants in Delhi," *International Journal of Environment and Pollution*, vol. 42, no. 1–3, pp. 58–67, 2010, doi: 10.1504/IJEP.2010.034226.

[20] H. Cui, R. Ma, and F. Gao, "Relationship between meteorological factors and diffusion of atmospheric pollutants," *Chemical Engineering Transactions.*, vol. 71, pp. 1417–4122, 2018, doi: 10.3303/CET1871237.

[21] Y. Zhang, "Dynamic effect analysis of meteorological conditions on air pollution: A case study from Beijing," *Science of Total Environment*, vol. 684, pp. 178–185, Sep. 2019, doi: 10.1016/J.SCITOTENV.2019.05.360.

[22] T. Ma, F. Duan, K. He, Y. Qin, D. Tong, G. Geng, X. Liu, H. Li, S. Yang, S. Ye, and B. Xu, "Air pollution characteristics and their relationship with emissions and meteorology in the yangtze river delta region during 2014-2016," *Journal of Environmental Science (China)*, vol. 83, pp. 8–20, Sep. 2019, doi: 10.1016/J.JES.2019.02.031.

[23] D. Giri, V. Krishna Murthy, and P. R. Adhikary, "The influence of meteorological conditions on PM10 concentrations in Kathmandu Valley," *International Journal of Environmental Research.*, vol. 2, no. 1, pp. 49–60, 2008.

[24] J. Wang and S. Ogawa, "Effects of meteorological conditions on PM2.5 concentrations in Nagasaki, Japan," *International Journal of Environmental Research and Public Health*, vol. 12, no. 8, pp. 9089–9101, 2015, doi: 10.3390/ijerph120809089.

[25] R. Jayamurugan, B. Kumaravel, S. Palanivelraja, and M. P. Chockalingam, "Influence of temperature, relative humidity and seasonal variability on ambient air quality in a coastal urban area," *International Journal of Atmospheric Sciences.*, vol. 2013, pp. 1–7, Dec. 2013, doi: 10.1155/2013/264046.

[26] T. M. Habeebullah, S. Munir, A. H. A. A. Awad, E. A. Morsy, A. R. Seroji, and A. M. F. Mohammed, "The interaction between air quality and meteorological factors in an arid environment of Makkah, Saudi Arabia," *International Journal of Environmental Science and Development*, vol. 6, no. 8, pp. 576–580, 2015, doi: 10.7763/IJESD.2015.V6.660.

[27] Most polluted countries 2022. *World Population Review*. [Online]. Available: https://worldpopulationreview.com

[28] Air quality and pollution city ranking. (2022). [Online]. Available: https://www.iqair.com/in-en/world-air-quality-ranking.