AISHWARYA MISHRA
SANJAY KUMAR PALEI
and
NETAI CHANDRA KARMAKAR

Receptor modelling of air pollutants for source apportionment on coal-based anthropogenic activities

Identification of sources of air pollutants plays a predominant role in coal-based anthropogenic activities to manage the air quality. Coal mining and thermal power plants (TPPs) are two multi-activity-centred coal-based sources that affect the ambient air quality of the region. Among five air pollutants under consideration, the annual average concentration of PM_{10} and $PM_{2.5}$ exceeded the Indian Central Pollution Control Board (CPCB) prescribed annual standard limit, and the probability of exceedance of daily average concentration for PM_{10} was 80.3% and that of $PM_{2.5}$ was 60.7%. Therefore, in this paper, pollutants' mass concentration data has been used to quantify the proportionate contribution of the sources using Positive Matrix Factorization (PMF) analysis for five pollutants. In addition, correlation analysis and literature survey has been used to designate the sources. The PMF analysis revealed that 74.4% of PM_{10} are emitted from coal mining and its allied activities, and 25.6% by TPP; whereas TPPs contribute 57.6% of $PM_{2.5}$ emission and 42.4% is by coal mining and its allied activities. Source apportionment of pollutants can help policy-makers and company management to devise suitable short-term as well as long-term emission control measures to manage particulate pollutants.

Keywords: *Source apportionment; air pollutants; positive matrix factorization; coal mine; thermal power plants*

1. Introduction

Deteriorating air quality has become a global concern. Air pollution in developing countries like India can be primarily linked to anthropogenic sources like industrialization, urbanization and activities supporting these processes. Industrial settlements have mostly affected the residents in the urban areas, as it is evident from the studies conducted in some renowned cities worldwide: Karachi (Mansha et al. 2012), Delhi (Guttikunda and Goel 2013; Sharma et al. 2014, 2016; Jain et al. 2018), Xiamin (Wang et al. 2019b), Macao (Wang et al. 2019a), Beijing (Yu et al. 2019), Tianjin (Peng et al. 2019), and New York (Zhou et al. 2019). The high concentration of pollutants in a geographical region is mainly because of continuous emissions from the anthropogenic activities and poor atmospheric dispersion of pollutants. Exposure to a high concentration of pollutants in the atmosphere increases the risk of human morbidity and mortality (Wang et al. 2019a). It is reported that an increase in 10ìg/m^3 concentration of particulate matter may increase the overall mortality risk by 1% (Lippmann 1998), and it may increase 3-6% in mortality risk in association with the respiratory diseases (Ostro et al. 1999). This study is also corroborated by another study stating that long-term exposure to the particulate matter will cause a substantial reduction in the life expectancy of a person (Brunekreef and Holgate 2002). Therefore air pollutants should be managed and it is essential to identify their sources.

Since air pollutants are particulate and gaseous in nature, these remain airborne for many days and transport to long distances depending on wind velocity and they have direct and indirect impacts on inhabitants. Particulate matters are present in environment as suspended dust generated from fossil fuel combustion in power and heat generation process, and from residential sector and diesel vehicles (Putaud et al. 2004; Nelson 2007; Juda-rezler et al. 2011). Particles of size below 10µm are regarded as PM_{10} and that of below 2.5µm are called $PM_{2.5}$. The largest contributory factor in PM_{10} is often soil or dust ranging between 20% and 64%, while elemental contribution is main in $PM_{2.5}$ (Vargas et al. 2012). The hazardous impacts of air pollutants, its elemental composition and physical characterization, with temporal variation in the meteorological parameters and characteristics of the source profile will further help in the reliable estimation of origin of pollutants and their quantification. Few studies use advanced dispersion modelling like hybrid single-particle Lagrangian integrated trajectory to identify the source of the pollution (Querol et al. 2009; Draxler and Rolph 2012). However, high uncertainties are linked to the hybrid single-particle Lagrangian integrated trajectory model results (Givehchi et al. 2013). Many researchers have used multiple site data for a single matrix input for source apportionment (Givehchi et al. 2013; Jain et al. 2018). Data from a single

Messrs. Aishwarya Mishra, Sanjay Kumar Palei and Netai Chandra Karmakar, Indian Institute of Technology (Banaras Hindu University), Varanasi, India. Email: aishwaryam.rs.min16@itbhu.ac.in / skpalei.min@iitbhu.ac.in / karmakarnc.min@iitbhu.ac.in

monitoring site can also be used to predict the sources using conditional probability function using wind direction (Kim and Hopke 2012). The conditional bivariate probability function (CBPF) is used to analyze pollutant source contribution with varying wind direction as well as wind speed (Rai et al. 2016).

For identification of source of pollutants, receptor modelling has been widely used to designate and quantify the source of pollutants from the air quality monitoring data available at the monitoring site (Contini et al. 2010). Among the multivariate statistical methods available for source apportionment, PMF and principal component analysis (PCA) are predominantly preferred by the researchers. PCA uses the concentration of pollutants collected at the sampling site and prior information about the source is not required. If the sources of the same profile are present in the region, cluster analysis of the measured concentration is required to apply PCA. Unlike PCA, PMF does not rely on information from the correlation matrix but utilizes the concept of non-negativity of factors, data uncertainties and tracer element information for source identification (Cesari et al. 2016; Rai et al. 2016). In addition, PMF utilizes a point-by-point least square optimization scheme. Therefore, the profiles produced can be directly compared to the input matrix without transformation (Lee et al. 1999; Kim and Hopke 2012; Cesari et al. 2016). This distinctive advantage of PMF over PCA makes it a powerful alternative to traditional receptor models.

Industrial settlements in mineral rich regions are identified by multi-activity-centred industries and hence the domain of source apportionment study in such region should be industry-centred rather than being activity-centred. The source apportionment requires PMF analysis, correlation analysis and literature survey for its assessment. The present paper illustrates source apportionment through a case study on an industrial area located in the northern region of India, where coal mining and coal use are the predominant anthropogenic activities. The study has considered five air pollutants viz PM_{10}, $PM_{2.5}$, CO, NO_2, SO_2 for source apportionment. There are 11 coal mines and seven TPPs representing the anthropogenic activities in the region. These 11 coal mines produced coal over 97 MT in 2016-17 (Coal India Limited 2018). It is expected that coal mining activity and TPPs are major contributors to particulate matters (PM_{10} and $PM_{2.5}$) and gaseous pollutants, though many small scale industries have also been operating in this region. Data reported by CPCB monitoring site at this region showed that annual average concentration of PM_{10} was 148μg/m^3 and that of $PM_{2.5}$ was 55μg/m^3 during 2017-18, and these values exceeded the standard prescribed limit. Since exposure to higher concentration of pollutants especially PM_{10} and $PM_{2.5}$ impacts human health and increases risk to mortality and morbidity, it is important to identify the sources of the pollutants to strategically reduce the emission of pollutants. This paper attempts to identify and quantify the contribution of various sources of air pollution in the region using PMF receptor model so that attempts can be made by the policy-makers and respective company managements to reduce the generation of particulate matters at respective sources.

2. Methodology

The sequence of operation to identify the sources of pollution and the quantification of pollutants from them is represented as in Fig.1. The pollutant data is collected from CPCB and then classified into different pollutants. The missing values of concentration data and the uncertainty data were identified and quantified. Then, the seasonal variation study of the pollutants is done to identify if the emission pattern of different pollutants change temporally. If the temporal variation is there, then seasonal source apportionment is done, else whole data may be considered for source apportionment. The source apportionment is done using PMF analysis. Different factors prescribed by PMF analysis are then identified based on the correlation analysis and extensive literature survey. Correlation analysis helps in identifying the similarity in source of different pollutants, and identifies the tracer elements among them. This will help in designating a name to different sources.

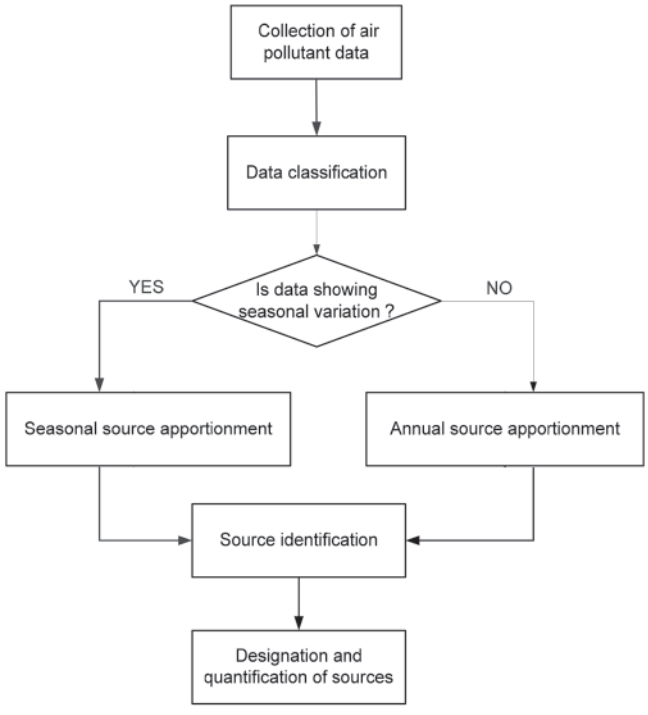

Fig.1 Flowchart of methodology followed

2.1 Source Apportionment using PMF

PMF is an advanced factor analysis technique to estimate the contribution of sources based on the work of Paatero and Tapper (Paatero and Tappert 1994). PMF uses realistic error estimates to weigh the data and imposes non-negativity constraints in the factor computational process. The quantitative contribution of pollutants to their source

obtained from PMF results form the basis to quantify the contribution of each source. The general receptor modelling problem can be stated in terms of the contribution from p-independent sources to all chemical species in a given sample (Hopke 2016), and it can be represented as follows:

$$e_{ij} = (x_{ij} - \sum_{k=1}^{p} f_{jk} g_{ki}) \quad \text{... (1)}$$

where f_{jk} is the factor profile (j^{th} species in the k^{th} factor) and g_{ki} is the factor contribution (of the k^{th} factor in the i^{th} sample), x_{ij} is the j^{th} species concentration measured in the i^{th} sample and e_{ij} is the residual associated with the j^{th} species concentration measured in the i^{th} sample, and p is the total number of independent sources. The corresponding matrix equation is:

$$X_{ij} = G_{ik} \times F_{kj} + E_{ij}$$

where X is an n×m data matrix with n measurements and m number of elements, E is an n×m matrix of residuals or errors. G is an n×p factor contribution matrix with p factors, and F is a p×m factor profile matrix. There are a potential infinite number of possible solutions to this bilinear factor analysis problem (rotations of G matrix and F matrix). To decrease rotational freedom, PMF uses non-negativity constraints over the factors. PMF provides a solution that minimizes an object function, Q(E), based upon uncertainties for each observation (Paatero and Tappert 1994; Paatero et al. 2014). This function is defined as:

$$Q(E) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\frac{x_{ij} - \sum_{k=1}^{p} f_{ik} g_{ki}}{s_{ij}} \right]^2 \quad \text{... (2)}$$

where s_{ij} is an uncertainty estimate in the j^{th} element measured in the i^{th} sample. The receptor modelling problem is then to minimize Q(E) with respect to G and F with the constraint that each of the elements of G and F is to be non-negative. This problem is solved iteratively as a weighted linear least squares problem (Paatero and Tappert 1994).The emission source identification is a difficult task (Paatero et al. 2002, 2005). Often the factors are identified by the users based on their intuition. The factors identified using PMF are often compared to profiles developed for the chemical mass balance method to identify likely contributors (Begum et al. 2007; Rizzo and Scheff 2007). Several researchers have concluded that multivariate statistical analysis techniques (e.g., PMF, PCA) are useful for identification and interpretation of emission sources, and to obtain apportionment quantifications to pollutants. Such multivariate analysis requires the emission profile of the sources.

However, in the event of data being available only from receptor, where no profile can be developed for likely contributors, literature survey can be used for source identification. PMF user manual recommends three types of variables such as strong, weak and bad variables in the PMF analysis (Reff et al. 2007; Norris et al. 2014). This designation of variables is based on the signal to noise ratio in such a way that variables with high uncertainty (noise) are weighted less than the variables with low uncertainty. When a species is labelled as weak, PMF triples the provided uncertainty. When it is labelled as bad, PMF excludes the species from the rest of the analysis. The variables with no concern with the source of emission are also labelled as bad. The seed number with minimum Q-value for a given number of factors is considered for further analysis. Detailed process flow of PMF for source apportionment is depicted in Fig.2 (Norris et al. 2014). The PMF analysis has been done using the EPA-PMF 5.0 software [35]. The PMF tool examines the G-space plot between different emission sources to find out the feasibility of the solution obtained so far. The feasible solution is shown by the independence of factors from one another. The G-space plot is represented as the scatter plot of the one factor versus another factor and it points in the plot converging to zero on both the axes of the factors under comparison shows that result is unique and optimum. Any deviation will result in greater rotational ambiguity. In this situation, the pair of factor matrix (G, F) should be transformed to another pair (G*, F*).

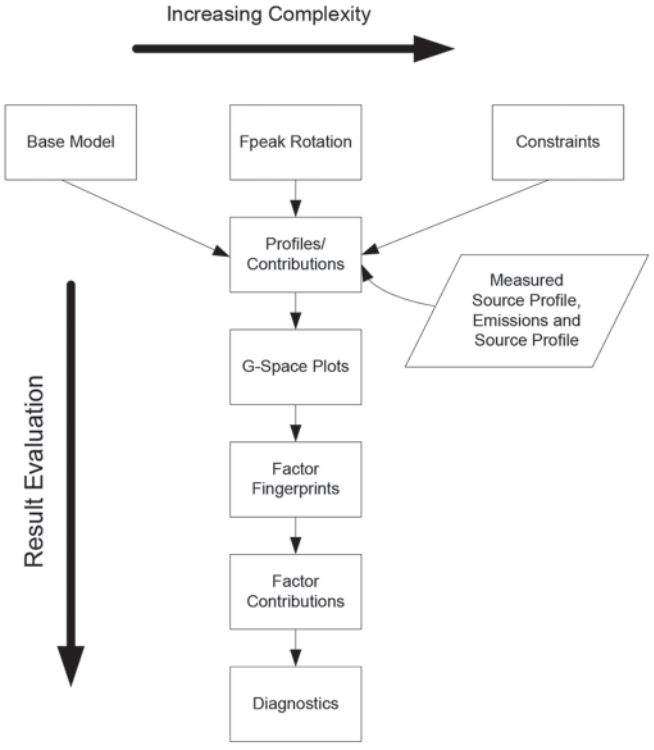

Fig.2 PMF result evaluation process

This is an elementary rotation operation on factor matrices. This rotation is represented by following equations:

$$g^*_{is} = g_{is} + r \times g_{iu} \qquad g^*_{iv} = g_{iv} \quad \text{... (3)}$$

$$f^*_{uj} = f_{uj} - r \times f_{sj} \qquad f^*_{wj} = f_{wj} \quad \text{... (4)}$$

where s and u are some random factor indices such that 1< s < p, 1 < u < p, and s is not equal to u. w denotes all index values from 1 to p except u. Index i represents all values from 1 to m. Index j represents all values from 1 to n.

The rotation matrix T, corresponding to Eq. (3) and Eq. (4), is almost a unit matrix of dimension $p \times p$ with the off-diagonal element t_{us} = r. This transformation rotates the factor matrices such that H^*F^* = GF. The rotation has been represented as in Eq. (5) and Eq. (6) with the condition that $g^*_{ih} > 0$, $f^*_{hj} > 0$. This non-negativity constraint limits the rotation so that under the condition of no allowed rotations, the unique solution is obtained.

$$G^* = GT \quad \ldots (5)$$

$$F^* = T^{-1}F \quad \ldots (6)$$

The complexity of rotational ambiguity is dealt with using EPA-PMF 5.0 software which has a tool having rotational parameter φ, called F_{PEAK}. A non-zero value of F_{PEAK} tries to impose rotation on the emerging solution during iteration. Positive value of φ tries to rotate the matrix using positive coefficient r and negative value of φ tries to rotate the matrix using negative coefficient r(Paatero et al. 2005). To get the feasible result, the domain of rotations in F_{PEAK} should be explored and the most appropriate scatter plot must be chosen for obtaining the final result.

3. Case study

3.1 Sampling Site Description

The sampling site of CPCB is located in northern India between latitudes 24°12'N and 23°47'N and longitudes 81°45' to 82°48'E, with an average elevation of about 383 m above mean sea level. The region is part of lower Gondwana sediments with depositional setting of coal seams occurring in Barakar and Raniganj formations. It caters the need of non-coking coal for TPPs of north and north-western India, and aluminium and chemical industries. It contains 11 mines arranged in W-pattern to the north of sampling station, and at this position, the measurements at the monitoring station are influenced by multiple sources. Therefore, the information provided is representative of the contributory effect of multiple emissive sources, typical of an industrial settlement. Seven TPPs are located to the east of sampling site in the region, and these TPPs are fed with coal mostly from these mines, and transported by different transportation systems to the TPPs to produce electricity that benefits several Indian states.

Various coal mining operations like drilling, blasting, loading, and transportation are the major sources of particulate emission to the ambient air. Similarly, on the part of TPP, coal combustion, transportation of fly ash, pollutants from boiler are some important contributors to air pollution. The emitted air pollutants are often stirred up and travel to longer distances with atmospheric wind and cause cardiovascular diseases to those living under the zone of influence. The fugitive emission from these industries is captured by the monitoring station, as depicted in windrose diagram in Fig.3, which is in downwind direction to the coal mines and TPPs.

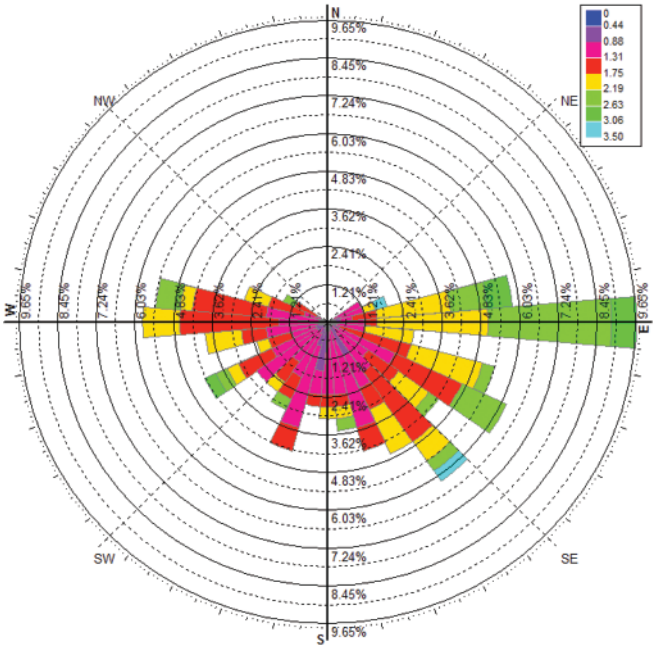

Fig.3 Windrose plot of the sampling site

3.2 Sampling Equipment and Data Extraction

The samplers installed at the monitoring station by CPCB measures concentration of 12 air pollutants (2018). The methods used for measurement of respective parameters are a combination of physical method, wet chemical method and continuous online method. The concentrations of PM_{10} and $PM_{2.5}$ have been estimated using gravimetric method. High volume samplers with automatic volumetric flow control calibrated at flow rate of 1132 l/min and glass filter of 8-inch by 10-inch size was used by CPCB to collect PM_{10} samples. It is reported that the filter is initially conditioned in a conditioning room within 20-30^{0}C and 40-50% relative humidity in airtight desiccators for getting proper results. For estimating $PM_{2.5}$ concentration, 47mm poly-tetraethylene filters were used. Gaseous pollutants like sulphur dioxide was measured using improved West and Gaeke method; Carbon monoxide by using chemical method, and nitrogen dioxide was measured using modified Jacob and Hochheiser method. Daily temperature, surface wind speed, wind direction and air pressure were also available from air monitoring site.

These measurements are updated daily on the website of CBCB and extracted online for a period ranging from December 2017 to November 2018 (2018). Since the uncertainty data is available only up to 29 September 2018, a total of 303 data have been considered for source

apportionment analysis in this paper. The captured data contains the concentration of air pollutants, and meteorological parameters e.g., ambient temperature, relative humidity, wind speed, wind direction, solar radiation, and rainfall. The error estimates of five major pollutants, such as PM_{10}, $PM_{2.5}$, sulphur dioxide, nitrogen dioxide and carbon monoxide, were also recorded from the dataset and considered source apportionment. The outliers in the pollutant concentration dataset have been removed by statistical analysis. Further, the missing values in the concentration data were replaced by the geometric mean of the available data, while uncertainty has been estimated as four times that of pollutant concentration, as suggested in many studies (EPA (PM F) 5.0 User Guide, 2014, Kim et al., 2018).

4. Results and discussions

4.1 Analytical Result of the Pollutant Samples

The descriptive statistics of pollutants concentration as recorded daily is presented in Table 1.

Fig.4 shows the average daily variation in the concentration of pollutants. It is observed that the average concentration of all five pollutants was more during the post monsoon season.

Table 2 presents the trend of seasonality; higher concentration was observed during winter (January-March) and summer (April-June) seasons than the monsoon season (July-September). Similarly, the observed concentration values of the pollutants during the post monsoon month of December

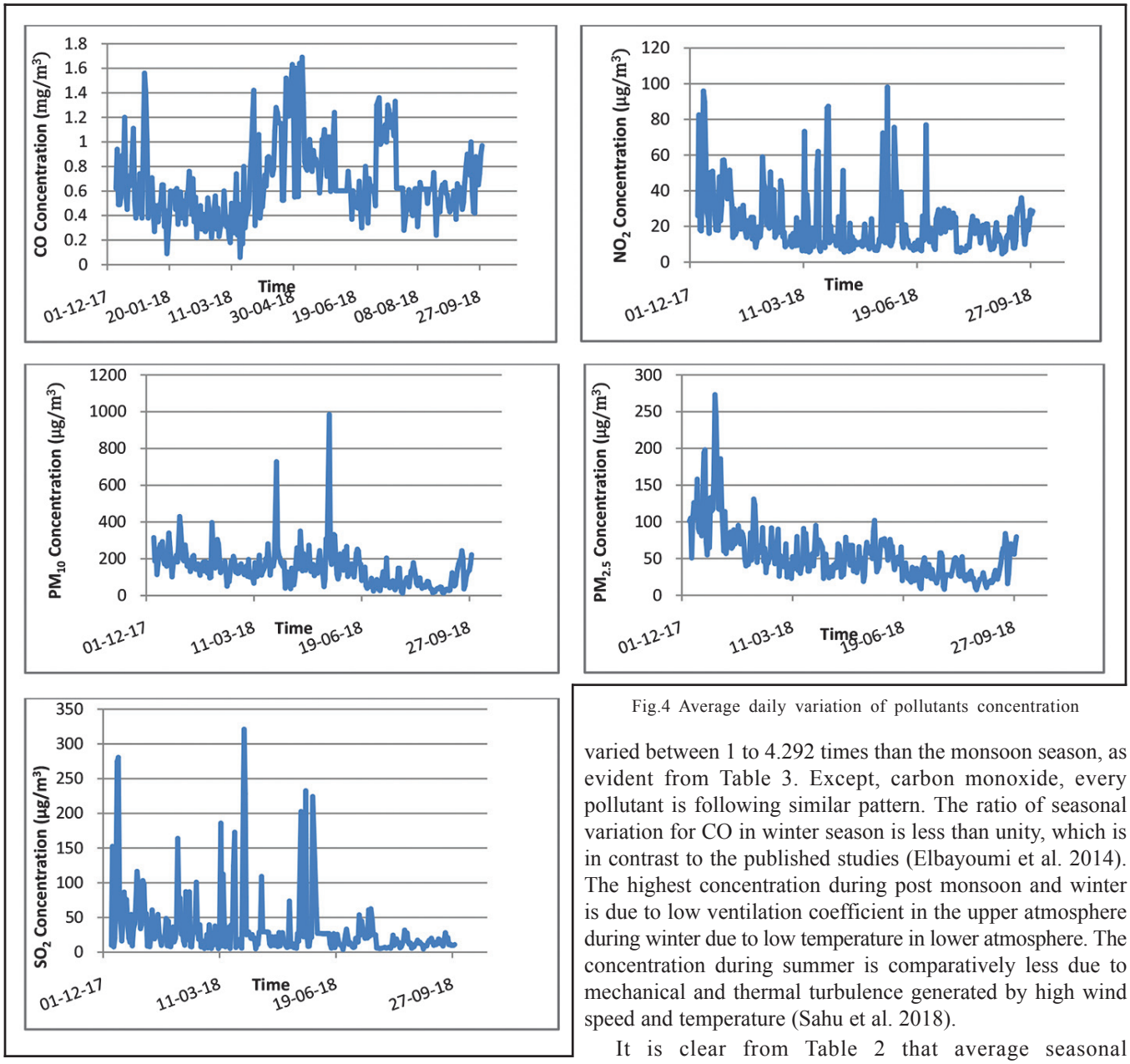

Fig.4 Average daily variation of pollutants concentration

varied between 1 to 4.292 times than the monsoon season, as evident from Table 3. Except, carbon monoxide, every pollutant is following similar pattern. The ratio of seasonal variation for CO in winter season is less than unity, which is in contrast to the published studies (Elbayoumi et al. 2014). The highest concentration during post monsoon and winter is due to low ventilation coefficient in the upper atmosphere during winter due to low temperature in lower atmosphere. The concentration during summer is comparatively less due to mechanical and thermal turbulence generated by high wind speed and temperature (Sahu et al. 2018).

It is clear from Table 2 that average seasonal

Table 1: Descriptive statistics of concentration of pollutants

Pollutants	Mean	St.dev	Range
PM_{10} (µg/m^3)	148.05	99.61	12.5-985
$PM_{2.5}$ (µg/m^3)	55.32	36.65	7.56-273
SO_2 (µg/m^3)	32.38	20.85	3.7-321
CO (mg/m^3)	0.67	0.31	0.06-1.69
NO_2 (µg/m^3)	21.32	16.98	4.66-98

concentration of PM_{10} and $PM_{2.5}$ exceed the prescribed Indian CPCB standard limit values. However, the average seasonal concentrations of NO_2, SO_2 and CO are below the prescribed value of CPCB in all seasons. Thus, instead of season-wise source apportionment, the source apportionment due to complete data set has been considered. The average annual concentration of pollutants and their annual limit as prescribed by CPCB is presented in Table 4.

It is revealed from Table 4 that the annual arithmetic mean of PM_{10} and $PM_{2.5}$ exceeds the CPCB annual limit by 2.467 and 1.383 times respectively with 80.3% as probability of exceedance of average daily exposure values for PM_{10} and 60.7% for $PM_{2.5}$ (Fig.5). This analysis further indicates that PM_{10} and $PM_{2.5}$ were the major components of air pollution and their concentration was high during post monsoon season and their values are reported to be 218.64±59.18 µg/m^3 and 118.81±50.51 µg/m^3 respectively (Table 2).

4.2 Results of Source Apportionment

The results obtained from Positive Matrix Factorization (EPA-PMF 5.0) model for the collected data of pollutants are shown in Figs.6 and 7. The results in Figs.6 and 7 are attributed to the F_{PEAK} value of $\varphi = -0.5$. For the two emission sources, Fig.8 shows most appropriate G-space plot implying the strong evidence of independence of the emission sources

Table 2: Seasonal concentration of pollutants in the region

Pollutants	Winter (Jan-March)	Summer (April-June)	Monsoon (July-Sept.)	Post Monsoon (Dec)
PM_{10} (µg/m^3)	171.08±64.68	174.01±121.31	79.91± 51.52	218.64±59.18
$PM_{2.5}$ (µg/m^3)	66.99±28.80	48.50±18.99	33.18±16.25	118.81±50.51
Sulphur Dioxide (µg/m^3)	36.25±23.44	33.92±19.00	16.49 ±9.00	70.78±53.00
Nitrogen Dioxide (µg/m^3)	22.36±14.40	18.80±10.77	17.48±7.80	40.91±23.23
Carbon Monoxide (mg/m^3)	0.48±0.23	0.83±0.28	0.69±0.20	0.72±0.29

Table 3: Ratio of seasonal variation of concentration of pollutants with respect to monsoon season

Pollutants	Winter (Jan-March)	Summer (April-June)	Post Monsoon (Dec)	Monsoon (July-Sept.)
PM_{10}	2.141	2.177	2.736	1
$PM_{2.5}$	2.019	1.461	3.580	1
Sulphur Dioxide	2.198	2.057	4.292	1
Nitrogen Dioxide	1.279	1.075	2.340	1
Carbon Monoxide	0.686	1.201	1.033	1

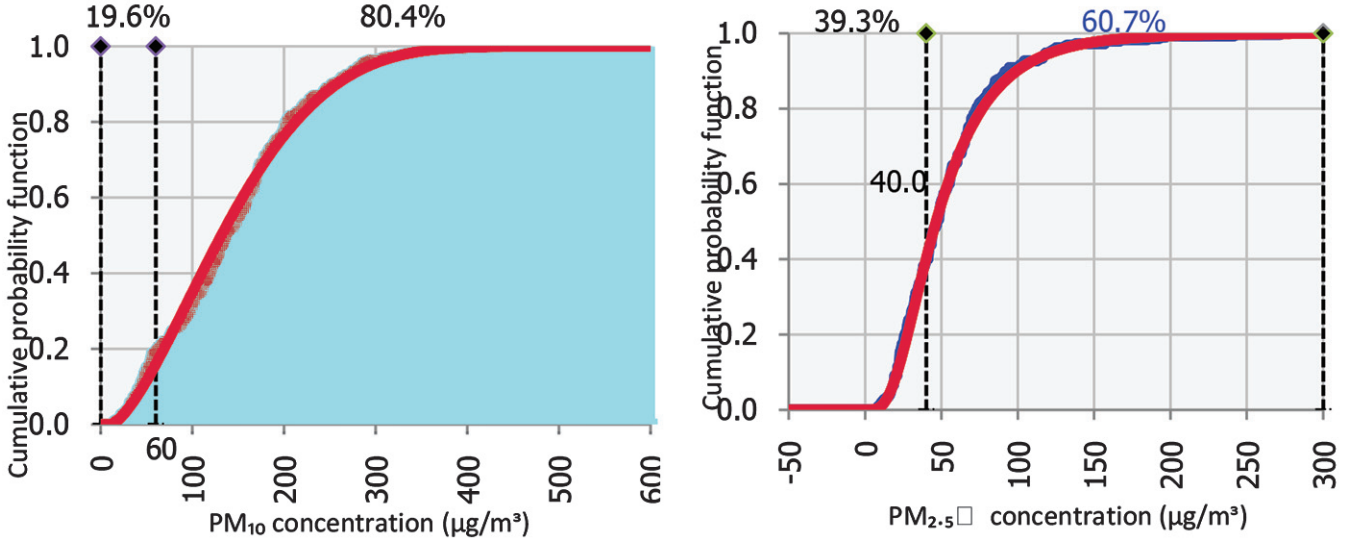

Fig.5 CDF plot of PM_{10} and $PM_{2.5}$ mass concentration showing their exceedance from annual limit prescribed by CPCB

from each other. The inter-species correlations between the pollutants are presented in Table 5. The mass concentrations of different constituents showed significant correlations with one another. For this dataset, Spearman correlation is followed because the data set did not follow bivariate normal distribution (Fig.9). The particulate matter PM_{10} and $PM_{2.5}$ showed strong correlation between them, which is suggestive of the similarity of their sources. The sulphurdioxide and nitrogen dioxide have also shown strong correlation between themselves. This result has been verified by many studies

TABLE 4: ANNUAL PROPORTIONAL REPRESENTATION OF POLLUTANTS WITH RESPECT TO THEIR ANNUAL STANDARD LIMIT

Pollutants	Annual average	Annual standard limit	Ratio
PM_{10} ($\mu g/m^3$)	148.047	60	2.467
$PM_{2.5}$ ($\mu g/m^3$)	55.321	40	1.383
Sulphur dioxide ($\mu g/m^3$)	32.380	50	0.648
Carbon monoxide (mg/m^3)	0.206	5	0.041
Nitrogen dioxide ($\mu g/m^3$)	21.320	40	0.533

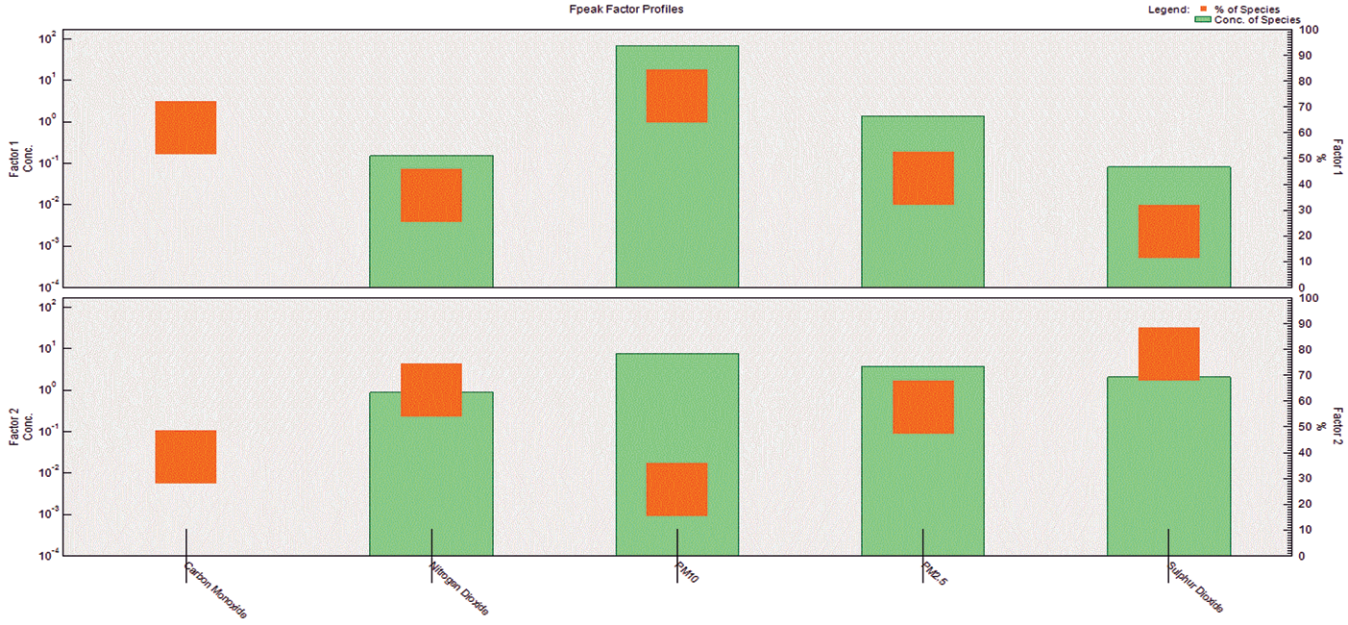

Fig.6 Source-wise distribution of pollutants

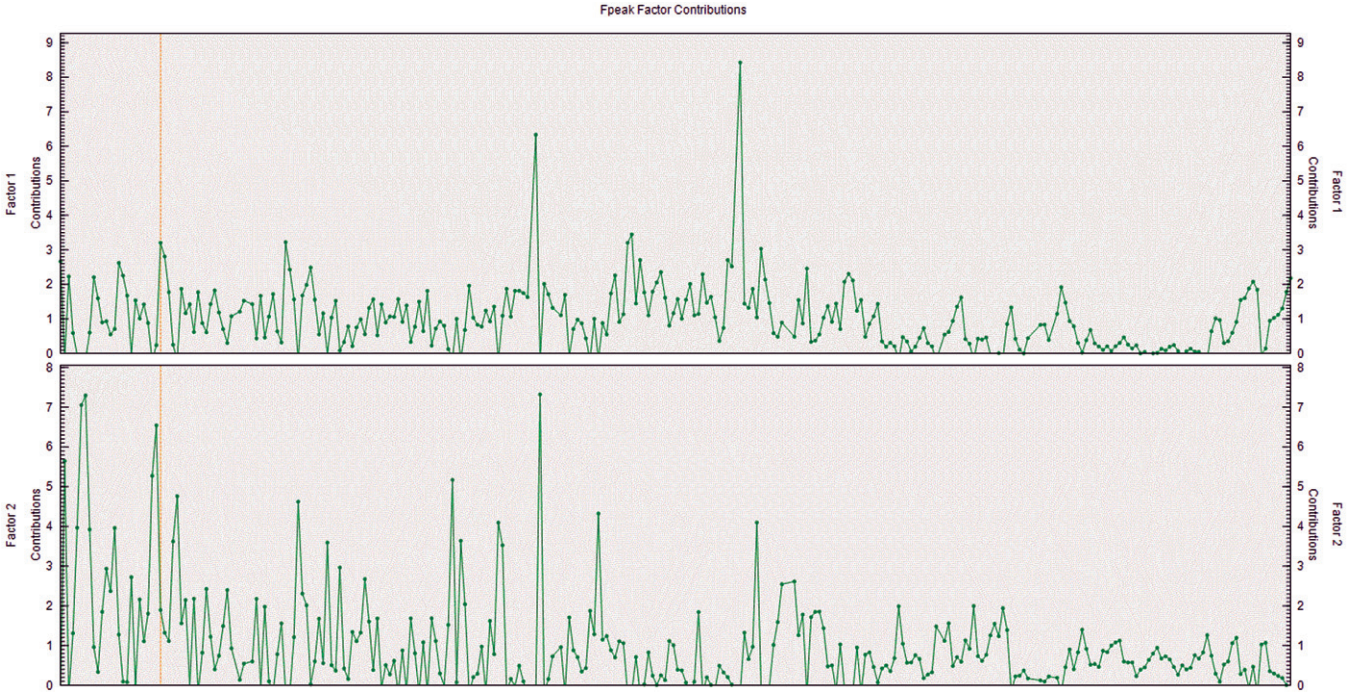

Fig.7 Temporal representation of mass concentration of pollutants from different sources

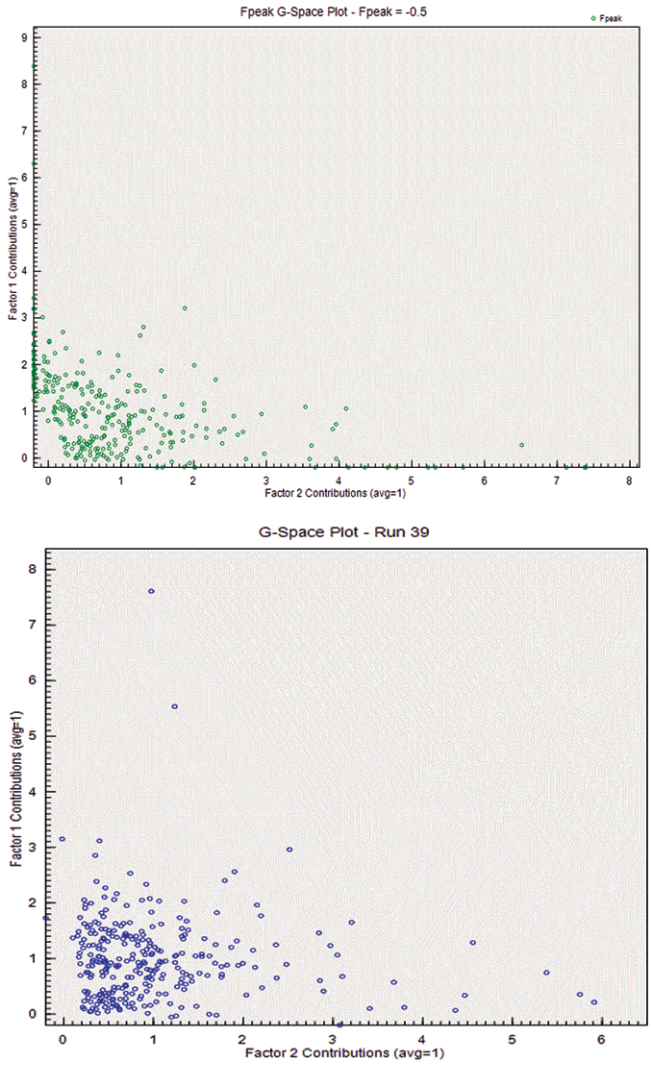

Fig.8 (a) G-space plot from rotational tool of EPA-PMF with F_{PEAK} value = –0.5, (b) G-space plot with base run model of EPA-PMF

(Kebin et al. 2002; Volchyn and Haponych 2014).

$PM_{2.5}$ shows moderate correlation with NO_2 and SO_2, which indicates that though these are emitted from same sources but they are not belonging to similar activities within the source. The PM_{10} shows comparatively weaker correlation with SO_2 and NO_2 which indicates different sources from which these are emitted.

In the industrial region, where opencast mines operate, mining is the major contributor to coarse material in the atmosphere (Vargas et al. 2012; Sahu et al. 2018; Yadav et al. 2019). Thus, factor 1 with highest loading of PM_{10} is attributed to coal mining and its allied activities. The high emission of PM_{10} and $PM_{2.5}$ is attributed to dust due to mining operations, haul road dust, public road side dust because of transport vehicles and coal dust that constantly remain in atmosphere due to mining related activities.

The aforementioned factor also suggests the presence of NO_2 and SO_2, and it may be attributed to vehicular emission due to mining related activities (Kumar et al. 2001).

Table 5: Spearman correlation coefficient for various pollutants

	CO	NO_2	PM_{10}	$PM_{2.5}$	SO_2
CO	1.00	0.15 (0.019)*	0.088 (0.14)	0.078 (0.19)	0.15 (0.01)*
NO_2		1.00	0.34(0)*	0.54(0)*	0.67(0)*
PM_{10}			1.00	0.68(0)*	0.33(0)*
$PM_{2.5}$				1.00	0.51(0)*
SO_2					1.00

*Correlation is significant at the level of 0.05.

A study also suggests that vehicular emission contain 30-40% of NO_x and 10% of SO_2 as gaseous pollutants (Rastogi et al. 2017). The factor 2 with high loading of SO_2 can be attributed to TPPs (Kumar et al. 2001; Volchyn and Haponych 2014). SO_2 is mainly emitted from combustion of coal to produce electricity. Such high emission is due to difficulty of removing sulphur from indigenous coal by physical cleaning methods (Mallik et al. 2019). NO_2 is also emitted from TPP due to reaction of N_2 with O_2 at high temperature in the boiler of TPPs. Moderate amount of PM_{10} and $PM_{2.5}$ are also contributed by TPP mainly in the form of flue gases and fly ash from the chimney of TPPs.

The second factor comprising maximum SO_2 and NO_2 emissions, and 57.6% of PM 2.5 emission reveals TPP as their common source of emission. This group has higher loading percentage of SO_2 because of combustion of sulphur rich coal (0.5% to 1% sulphur content) found in the region to produce electricity. The second factor comprising maximum SO_2 and NO_2 emissions, and 57.6% of $PM_{2.5}$ emission reveals TPP as their common source of emission. This group has higher loading percentage of SO_2 because of combustion of sulphur rich coal (0.5% to 1% sulphur content) found in the region to produce electricity. Source apportionment results (Table 6) show the significant contribution of coal mining towards pollution due to PM_{10} and $PM_{2.5}$. The percentage contribution of coal mining and allied activities due to CO is also significant but its overall contribution lies below the standard limit prescribed by CPCB. The maximum contribution of power plant to pollution is due to SO_2 emission (78.1%), while due to PM_{10} and $PM_{2.5}$ it is 25.6% and 57.6% respectively. The monitoring site may also be affected by distant sources of pollution; therefore pollutant contribution from distant sources under similar meteorological conditions should be evaluated.

Source apportionment results (Table 6) show the significant contribution of coal mining towards pollution due to PM_{10} and $PM_{2.5}$. The percentage contribution of coal mining and allied activities due to CO is also significant but its overall contribution lies below the standard limit prescribed by CPCB. The maximum contribution of power plant to pollution is due to SO_2 emission (78.1%), while due to PM_{10} and $PM_{2.5}$ it is 25.6% and 57.6% respectively. The monitoring

Table 6: Relative contribution of identified sources of pollutants

Sources	Anthropogenic Activities	Pollutants				
		% PM_{10}	% $PM_{2.5}$	% CO	% NO_2	% SO_2
Factor 1	Coal Mining and allied activities	74.4	42.4	61.8	35.6	21.9
Factor 2	TPP	25.6	57.6	38.2	64.4	78.1

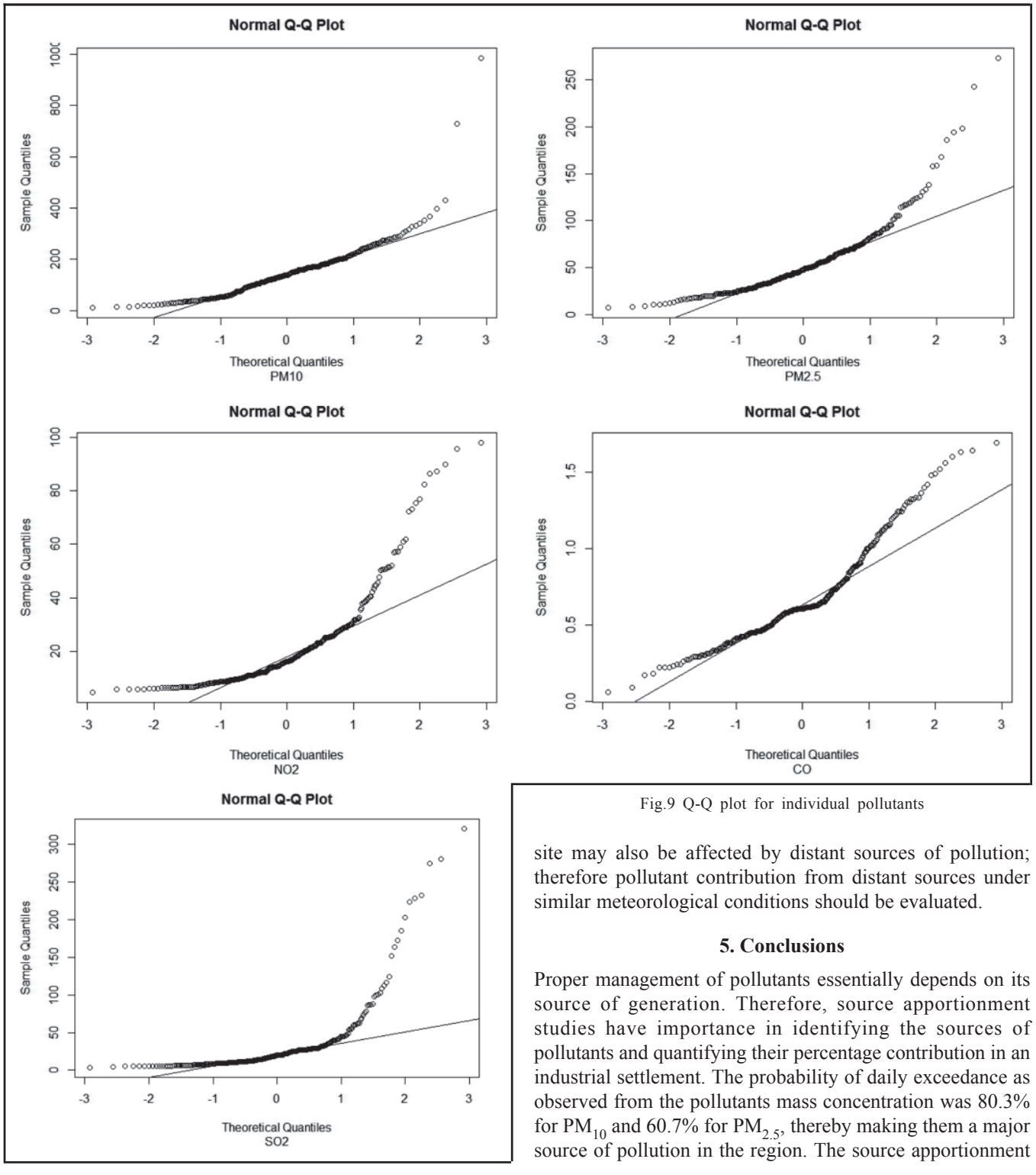

Fig.9 Q-Q plot for individual pollutants

site may also be affected by distant sources of pollution; therefore pollutant contribution from distant sources under similar meteorological conditions should be evaluated.

5. Conclusions

Proper management of pollutants essentially depends on its source of generation. Therefore, source apportionment studies have importance in identifying the sources of pollutants and quantifying their percentage contribution in an industrial settlement. The probability of daily exceedance as observed from the pollutants mass concentration was 80.3% for PM_{10} and 60.7% for $PM_{2.5}$, thereby making them a major source of pollution in the region. The source apportionment

results indicate that coal mining and its allied activities are major contributors of PM_{10} while TPPs are major contributors for $PM_{2.5}$ concentration in the atmosphere. While studying seasonal variations of pollutants, the emissions recorded at monitoring station are many folds higher in post monsoon, winter and summer season with respect to monsoon season. One of the reasons for the high concentration during the post monsoon and winter season is the low ventilation coefficient in the upper atmosphere. Thus, to alleviate the air pollution problem due to PM_{10} and $PM_{2.5}$, coal mining and its allied activities need to take precautionary measures to control the particulate matter generation. On the other hand, TPPs should use more sophisticated technologies to reduce $PM_{2.5}$ emissions on their part. Further this work can be extended for profiling the zone of influence of particulate matter in the coal-based anthropogenic region to understand and estimate the possible health impacts of particulate matter on human beings living in and around coal-based industrial zones. The present paper is expected to help the policy makers and company management to design a better air quality management strategy.

Acknowledgment

The authors are thankful to the Central Pollution Control Board of India for sharing their daily air quality monitoring data online and permitting us to use them. Authors also thankfully acknowledge the US-EPA for providing open license software EPA-PMF-5.0.

References

1. Begum B.A., Biswas S.K., Hopke P.K. (2007): Source Apportionment of Air Particulate Matter by Chemical Mass Balance (CMB) and Comparison with Positive Matrix Factorization (PMF) Model. Aerosol Air Qual Res 7:446–468. https://doi.org/10.4209/aaqr.2006.10.0021
2. Brunekreef B, Holgate S.T. (2002): Air pollution and health. Lancet 360:1233–1242. https://doi.org/10.1016/S0140-6736(02)11274-8
3. Cesari D, Amato F, Pandolfi M, et al (2016): An inter-comparison of PM 10 source apportionment using PCA and PMF receptor models in three European sites. *Environ Sci Pollut Res* 15133–15148. https://doi.org/10.1007/s11356-016-6599-z
4. Coal India Limited (2018): Annual Report 2017-18, Northern Coalfields Limited, Coal India Limited
5. Contini D, Genga A, Cesari D, et al (2010): Characterisation and source apportionment of PM10 in an urban background site in Lecce. Atmos Res 95:40–54. https://doi.org/10.1016/j.atmosres.2009.07.010
6. Draxler RR, Rolph GD (2012): Evaluation of the Transfer Coefficient Matrix (TCM) approach to model the atmospheric radionuclide air concentrations from Fukushima. *J Geophys Res* 117:1–10. https://doi.org/10.1029/2011JD017205
7. Elbayoumi M, Ramli N.A., Md Yusof NFF, Madhoun W Al (2014): The effect of seasonal variation on indoor and outdoor carbon monoxide concentrations in Eastern Mediterranean climate. *Atmos Pollut Res* 5:315–324. https://doi.org/10.5094/APR.2014.037
8. Givehchi R, Arhami M, Tajrishy M (2013): Contribution of the Middle Eastern dust source areas to PM 10 levels in urban receptors/: Case study of Tehran , Iran. *Atmos Environ* 75:287–295. https://doi.org/10.1016/j.atmosenv.2013.04.039
9. Guttikunda S.K., Goel R (2013): Health impacts of particulate pollution in a megacity — Delhi , India. *Environ Dev* 6:8–20. https://doi.org/doi.org/10.1016/j.envdev.2012.12.002
10. Hopke P.K. (2016): Review of receptor modelling methods for source apportionment Review of receptor modelling methods for source apportionment. *J Air Waste Manage Assoc* 2247:237–259. https://doi.org/10.1080/10962247.2016.1140693
11. Jain S, Sharma S.K', Mandal T.K., Saxena M (2018): Source apportionment of PM 10 in Delhi , India using PCA/APCS , UNMIX and PMF. *Particuology* 37:107–118. https://doi.org/doi.org/10.1016/j.partic.2017.05.009
12. Juda-rezler K, Reizer M, Oudinet J (2011): Determination and analysis of PM 10 source apportionment during episodes of air pollution in Central Eastern European urban areas/: The case of wintertime 2006. *Atmos Environ* 45:6557–6566. https://doi.org/10.1016/j.atmosenv.2011.08.020
13. Kebin H, Qiang Z, Hong H (2002): Types and amounts of vehicular emissions. In: Li Q (ed) Point sources of pollution: Local effects and it's control. UNESCO, pp 64–84
14. Kim E, Hopke PK (2012): Source Identifications of Airborne Fine Particles Using Positive Matrix Factorization and U . S . Environmental Protection Agency Positive Matrix Factorization Source Identifications of Airborne Fine Particles Using Positive Matrix Factorization and *U . S. J Air Waste Manage Assoc* 2247:811–819. https://doi.org/10.3155/1047-3289.57.7.811
15. Kim S, Kim T, Yi S, Heo J (2018): Source apportionment of PM 2.5 using positive matrix factorization (PMF) at a rural site in Korea. *J Environ Manage* 214:325–334. https://doi.org/doi.org/10.1016/j.jenvman.2018.03.027
16. Kumar A.V., Patil R.S., Nambi K.S.V. (2001): Source apportionment of suspended particulate matter at two traffic junctions in Mumbai , India. *Atmos Environ*

35:4245–4251. https://doi.org/10.1016/S1352-2310(01)00258-8

17. Lee E, Chan C.K., Paatero P (1999): Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmos Environ* 33:3201–3212. https://doi.org/10.1016/S1352-2310(99)00113-2
18. Lippmann M (1998): The 1997 US EPA Standards for Particulate Matter and Ozone. *R Soc Chem* 10:75–99. https://doi.org/doi.org/10.1039/9781847550095-00075
19. Mallik C, Mahapatra P.S., Kumar P, et al (2019): Influence of regional emissions on SO2 concentrations over Bhubaneswar, a capital city in eastern India downwind of the Indian SO2 hotspots. *Atmos Environ* 209:220–232. https://doi.org/10.1016/j.atmosenv.2019.04.006
20. Mansha M, Ghauri B, Rahman S, Amman A (2012): Characterization and source apportionment of ambient air particulate matter (PM2.5) in Karachi. *Sci Total Environ* 425:176–183. https://doi.org/10.1016/j.scitotenv.2011.10.056
21. Nelson P.F. (2007): Trace Metal Emissions in Fine Particles from Coal Combustion. *Energy & Fuels* 21:477–484. https://doi.org/10.1021/ef060405q
22. Norris G, Duvall R, Brown S, Bai S (2014): EPA Positive Matrix Factorization (PM F) 5.0 Fundamentals and User Guide. U.S. Environmental Protection Agency Office of Research and Development Washington, DC 20460
23. Ostro B.D., Hurley S, Lipsett M.J. (1999): Air pollution and daily mortality in the Coachella valley, california: A study of PM10 dominated by coarse particles. *Environ Re*s 81:231–238. https://doi.org/10.1006/enrs.1999.3978
24. Paatero P, Eberly S, Brown S.G., Norris G.A. (2014): Methods for estimating uncertainty in factor analytic solutions. *Atmos Meas Tech* 781–797. https://doi.org/10.5194/amt-7-781-2014
25. Paatero P, Hopke P.K., Begum B.A., Biswas S.K. (2005): A graphical diagnostic method for assessing the rotation in factor analytical models of atmospheric pollution. *Atmos Environ* 39:193–201. https://doi.org/10.1016/j.atmosenv.2004.08.018
26. Paatero P, Hopke P.K., Song XH, Ramadan Z (2002): Understanding and controlling rotations in factor analytic models. Chemom Intell Lab Syst 60:253–264. https://doi.org/10.1016/S0169-7439(01)00200-3
27. Paatero P, Tappert U (1994): Positive Matrix Factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5:111–126. https://doi.org/10.1002/env.3170050203
28. Peng X, Liu X, Shi X, et al (2019): Source apportionment using receptor model based on aerosol mass spectra and 1 h resolution chemical dataset in Tianjin , *China. Atmos Environ* 198:387–397. https://doi.org/doi.org/10.1016/j.atmosenv.2018.11.018
29. Putaud J, Raes F, Dingenen R Van, et al (2004): A European aerosol phenomenology – 2/: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. *Atmos Environ* 38:2579–2595. https://doi.org/10.1016/j.atmosenv.2004.01.041
30. Querol X, Pey J, Pandolfi M, et al (2009): African dust contributions to mean ambient PM 10 mass-levels across the Mediterranean Basin. *Atmos Environ* 43:4266–4277. https://doi.org/10.1016/j.atmosenv.2009.06.013
31. Rai P, Chakraborty A, Kumar A, Gupta T (2016): Composition and source apportionment of PM 1 at urban site Kanpur in India using PMF coupled with CBPF. *Atmos Res* 179:506–520. https://doi.org/doi.org/10.1016/j.atmosres. 2016.04.015
32. Rastogi A, Rajan A V., Mukherjee M (2017): A Review of Vehicular Pollution and Control Measures in India. In: Advances in Health and Environment Safety. *Springer*, Singapore, pp 237–245
33. Reff A, Eberly SI, Bhave P V, et al (2007): Receptor Modelling of Ambient Particulate Matter Data Using Positive Matrix Factorization/: Review of Existing Methods. *J Air Waste Manag* 57:146–154. https://doi.org/10.1080/10473289.2007.10465319
34. Rizzo M.J., Scheff P.A. (2007): Fine particulate source apportionment using data from the USEPA speciation trends network in Chicago, Illinois/: Comparison of two source apportionment models. *Atmos Environ* 41:6276–6288. https://doi.org/10.1016/j.atmosenv.2007.03.055
35. Sahu S.P., Patra A.K., Kolluru SSR (2018): Spatial and temporal variation of respirable particles around a surface coal mine in India. *Atmos Pollut Res* 9:662–679. https://doi.org/10.1016/j.apr.2018.01.010
36. Sharma S.K., Mandal T.K., Saxena M, et al (2014): Source apportionment of PM10 by using positive matrix factorization at an urban site of Delhi, India. *Urban Clim* 10:656–670. https://doi.org/doi.org/10.1016/j.uclim.2013.11.002
37. Sharma S.K., Sharma A, Saxena M, et al (2016): Chemical characterization and source apportionment of aerosol at an urban area of Central Delhi , India. Atmos Pollut Res 7:110–121. https://doi.org/doi.org/10.1016/j.apr.2015.08.002

(Continued on page 311)