



Identification of sources affecting fog formation using receptor modeling approaches and inventory estimates of sectoral emissions

Bhavin Mehta^a, Chandra Venkataraman^{a,*}, Mani Bhushan^a, Sachchida Nand Tripathi^b

^a Department of Chemical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

^b Department of Civil Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

ARTICLE INFO

Article history:

Received 25 July 2008

Received in revised form

24 November 2008

Accepted 26 November 2008

Keywords:

Positive matrix factorization (PMF)

Potential source contribution function (PSCF)

Kanpur (80°22'E, 26°26'N)

Thermal power plants

Brick kilns

Biofuel combustion

ABSTRACT

Positive matrix factorization (PMF) was used to identify factors affecting fog formation in Kanpur during the ISRO-GBP land campaign-II (LC-II) in December 2004. PMF predicted factors were validated by contrasting the emission strength of sources in the foggy and clear periods, using a combination of potential source contribution function (PSCF) analysis and quantitative emission inventory information. A time series aerosol chemical data set of 29 days and 12 species was decomposed to identify 4-factors: *Secondary species*, *Biomass burning*, *Dust* and *Sea salt*. PMF predicted particle mass with a satisfactory goodness-of-fit (slope of 0.83 ± 0.17 and R^2 of 0.8), and strong species within 11–12% relative standard deviation. Mean contributions of anthropogenic factors were significantly higher during the foggy period for *secondary species* (2.9 ± 0.3) and *biomass burning* (1.2 ± 0.09) compared to the clear period. Local sources contributing to aerosols that mediated fog events at Kanpur, based on emissions in a 200 km × 200 km area around Kanpur city were thermal power plants and transportation (SO₂) and biofuel combustion (BC and OM). Regional scale sources influencing emissions during the foggy period, in probable source regions identified by PSCF included thermal power plants, transportation, brick kilns and biofuel combustion. While biofuel combustion and transportation are distributed area sources, individual point sources include coal-fired thermal power plants located in Aligarh, Delhi, Ghaziabad, Jhansi, Kanpur, Rae Bareilly and Rupnagar and brick kilns located in Allahabad, Agra, Farrukhabad, Ghaziabad, Kanpur, Ludhiana, Lucknow and Rae Bareilly. Additionally, in the foggy period, large areas of probable source regions lay outside India, implying the significance of aerosol incursion from outside India.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Aerosol particles mediate the formation of fog in the atmosphere (Pandis et al., 1990; Seinfeld and Pandis, 1998) through the preferred heterogeneous nucleation of water vapor on pollution particles, at high RH or low supersaturation. Fog droplets further aid aerosol formation through aqueous-phase reactions of soluble gaseous precursors (e.g. SO₂ and H₂O₂), leading to higher aerosol concentrations of species like sulfate on fog abatement, which then nucleate subsequent fog–smog–fog cycles (Pandis et al., 1990). Fogs lead to atmospheric removal (by dissolution and deposition) of aerosol constituents for which aqueous-phase production reactions are not important (e.g. nitrate, chloride and ammonium).

Fog can occur if the wind is calm and the air is sufficiently moist, cool and descending (e.g. Pruppacher and Klett, 1997). Such conditions are prevalent in the Indo-Gangetic Plain (IGP, 21°75'–31°N, 74°25'–91°50'E) during winter months of November–February (average temperature 17 °C, RH 64% and maximum daily RH > 90%) as evidenced by campaign measurements (e.g. Ali et al., 2004). An increasing frequency of the occurrence of ground fog has been observed during winter in cities in the IGP, for example, foggy days in New Delhi during November–February averaged 43 during 2000–2002, 48 during 2003–2005 and 61 during 2006–2008, with visibility in the range of 0.1–2.3 km (based on airport data from www.wunderground.com). Fog formation in this region has been attributed to meteorological disturbances moving eastward from the west (Pasricha et al., 2003). A recent study, the Indian Space Research Organization–Geosphere Biosphere Programme land campaign-II (ISRO-GBP LC-II), examined the role of aerosols in fog formation during December 2004 (e.g. Tare et al., 2006; Tripathi et al., 2006). Aerosol physical, chemical and optical measurements

* Corresponding author. Tel.: +91 22 2576 7224; fax: +91 22 2572 6895.

E-mail address: chandra@iitb.ac.in (C. Venkataraman).

URL: <http://www.iitb.ac.in/~che/chandra>

were made simultaneously at seven sites in the IGP and at one outside the IGP. Findings include high concentrations of K^+ , NO_3^- , SO_4^{2-} , NH_4^+ and BC during fog events.

Quantitative aerosol source identification contributing to these fog events is yet to be made and has been undertaken in this work. Receptor modeling approaches such as positive matrix factorization (PMF) are effective tools in source identification for aerosols on urban to regional scales (e.g. Bhanuprasad et al., 2008; Pekney et al., 2006). PMF studies typically analyze large data sets of several hundred samples spanning a few years of measurements containing 25–30 chemical species to resolve emission source categories on temporal scales of weeks to seasons. As such large data sets are not available yet in the Indian region, recently, PMF was used to resolve a smaller number (six) of source categories using a relatively small data set, e.g. 23 species in 27 samples (Bhanuprasad et al., 2008). The uncertainty in PMF predictions of factor contributions (based on the relative standard deviation of species with high signal to noise ratio) was satisfactory (5–6%), giving confidence in the application of PMF to smaller data sets towards an understanding of regional scale source–receptor relationships in India.

In this work, we use PMF, PSCF and emissions inventory information to identify sources that affected fog formation in Kanpur during December 2004. We identify potential source regions of the factors during the fog event using PSCF, and combine it with emissions inventory information for source identification.

2. Modeling approach

2.1. Aerosol data set

2.1.1. Data description

Aerosol data were collected at Kanpur during a land campaign by the Indian Space Research Organization (ISRO-GBP LC-II). Kanpur city ($80^\circ 22'E$, $26^\circ 26'N$) is situated in the southern part of the Indo-Gangetic plain. Sampling was done at Indian Institute of Technology Kanpur (IITK), about 17 km away from Kanpur city, on the roof of a three storey building in the IITK campus at 12 m above ground level. Kanpur city and adjoining areas have a population of over 4 million with a density of 6800 persons per km^2 (Gupta, 2006). Foggy periods were identified visually (Tripathi et al., 2006), based on the

daily number of hours of fog. On this basis, 9–27 December (except 25 and 26 December) were classified as foggy weather and the remaining campaign period was classified as clear weather.

Emission sources (Fig. 1) include biomass fuel (residential cooking and heating) and fossil fuel burning (mainly coal based thermal power plants and fertilizer plants, transportation, mining) (Reddy and Venkataraman, 2002), and seasonally variable forest and crop waste burning (Venkataraman et al., 2006). The vehicle fleet was estimated at 387,697 vehicles in 2002 and projected to increase by 30,000 every year (Gupta, 2006). Kanpur has over 310,100 cattle (BAHS, 2003), a source of ammonia emissions, with 140 g ammonia emissions reported per head of cattle per day (McGinn et al., 2007). Earlier work gives details of air sampling from December 1–29, 2004 (Tare et al., 2006; Tripathi et al., 2006) and results from size segregated aerosol samples collected for a period of 12 days. In this work, we use daily 8-h-average PM_{10} chemical composition of particles collected on glass fiber filters (Whatmann GF/A) using a high volume sampler operated at a flow rate of $0.7\text{--}1.1\text{ m}^3\text{ min}^{-1}$, from December 1–29, 2004.

Details of anion (Cl^- , NO_3^- and SO_4^{2-}) measurement using ion chromatography, elemental (Na, Mg, K, Ca, Fe, Al) measurement by atomic absorption spectrophotometry (AAS) and NH_4^+ ion measurement by indophenol blue method (Tare et al., 2006) are included in the Supplementary Online Data (Section A, Chemical Analysis Details). AAS measured Na, Mg, K and Ca were designated equal to water-soluble ionic forms by Tare et al. (2006) in earlier published work. Simultaneous measurements of black carbon (BC) were carried out continuously for 24 h using an aethalometer (Model AE-21-ER, Magee Scientific, USA) (Tripathi et al., 2006). The daily average ionic and elemental species in particle samples and aethalometer BC concentration used in the PMF analysis included NO_3^- , SO_4^{2-} , Cl^- , NH_4^+ , Na, Mg, K, Ca, Al, Fe and BC (Table 1). Chemical concentrations of all measured constituents were above detection limits. Three missing values of BC from December 22–24 were replaced by the geometric mean of measured values (Table 1). A time-series analysis (not shown), gave no extreme values leading to retention of all data for PMF analysis. A low Cl^-/Na^+ ratio indicates significant chloride depletion ranging $60 \pm 11\%$ averaged over the campaign, from atmospheric reactions of SO_2 or NO_x with sea-salt,

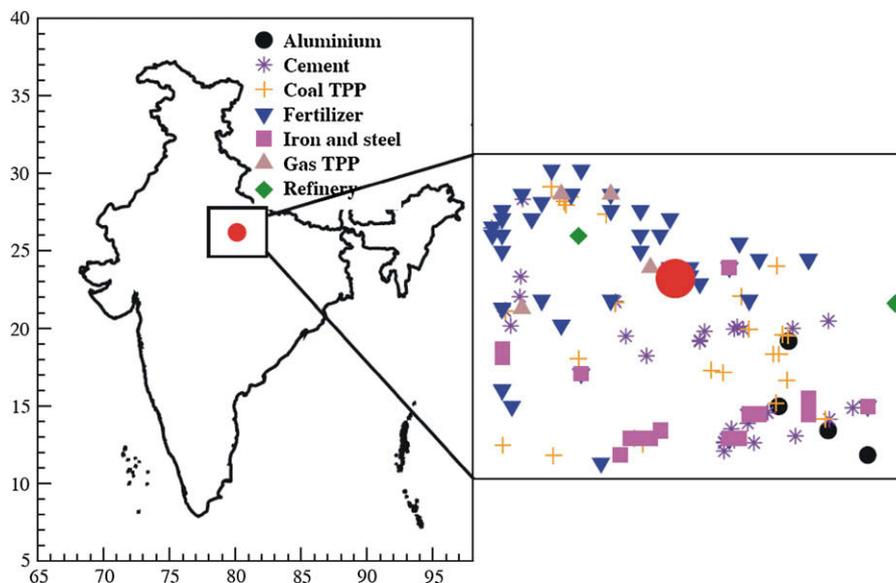


Fig. 1. Map of Kanpur and surrounding areas with location of large point emission sources.

Table 1
Summary of aerosol chemical measurements at Kanpur.

Species	Geometric mean ($\mu\text{g m}^{-3}$)	Arithmetic mean ($\mu\text{g m}^{-3}$)	Minimum ($\mu\text{g m}^{-3}$)	Maximum ($\mu\text{g m}^{-3}$)	MDL ^a (ng m^{-3})	Final p_j ^b (%)	Number of BDL values (%)	Number of missing values (%)	S/N
NO ₃ ⁻	15.9	16.9	7.9	29.1	8.7	14	0	0	3.6
SO ₄ ²⁻	14.8	15.6	6.5	27.8	12	30	0	0	1.7
Cl ⁻	2.9	3.0	1.4	4.7	3.1	20	0	0	2.5
NH ₄ ⁺	8.4	3.1	1.6	18.8	17	50	0	0	1
Na	4.2	4.3	2.7	6.2	30	14	0	0	3.5
Mg	0.2	0.2	0.05	0.4	8	10	0	0	4.5
K	5.1	5.2	2.9	6.9	14	20	0	0	2.5
Ca	1.6	1.7	0.95	3.95	1.4	22	0	0	2.3
Al	0.8	1.0	0.3	2.7	3.6	100	0	0	0.5
Fe	1.9	2.0	1.2	3.3	0.47	10	0	0	5
BC	10.0	10.6	4.8	19.8	200	200	0	3 (10)	0.25

^a Adapted from Kim et al. (2005a).

^b p_j values are taken as 2 times the values in Kim et al. (2005a) and higher p_j for NH₄⁺, SO₄²⁻, Al and BC.

indicating significant processing of sea-salt during its transport inland.

2.1.2. Prescription of measurement uncertainties

Uncertainties in the measurement data set were estimated (e.g. Pekney et al., 2006), as shown below:

$$S_{ij} = p_j X_{ij} + \frac{\text{MDL}_j}{3} \quad (1)$$

where, p_j is the uncertainty proportional parameter, X_{ij} the measured concentration, $i = 1-n$ is index for samples, $j = 1-m$ is index for species and MDL_{*j*} the method detection limit. PMF literature extensively suggests that the uncertainty assigned to a species must reflect both measurement uncertainty and the variation in source profiles over the measurement period (Kim et al., 2004, 2005a,b; Kim and Hopke, 2004; Begum et al., 2004). In the absence of both measurement uncertainty and source variability information for our campaign, we followed the usual practice in literature of selecting p_j 's from reported studies and systematically modifying them to obtain sensible solutions (Kim et al., 2005b; Kim and Hopke, 2004). Specifically, p_j values were chosen to be twice the p_j values obtained by the measurement network of U.S. Environmental Protection Agency's Speciation Trends Network (EPA-STN) (Kim et al., 2005a) for NO₃⁻, SO₄²⁻, Cl⁻, NH₄⁺, Na, Mg, K, Ca and Fe. When the aerosol mass is strongly influenced by a given species, most often sulfate, the composition of several factors may be dominated by that species, if the uncertainty assigned to it is too low (Kim et al., 2005a). Values of p_j for SO₄²⁻ and NH₄⁺ were increased in steps to reduce their loading on multiple factors, with final values of 30% and 50% above which better resolution was not obtained. Aethalometer black carbon measurements are known to have a larger uncertainty from shadowing and scattering effects (Kirchstetter and Novakov, 2007) leading to variation of a factor of 2 from thermal optical transmittance measurements. Hence, BC uncertainty proportional parameter was taken as 200%. The reported measurement of Al by XRF is known to have a smaller signal for Al than for other elements, leading to larger uncertainties in its measurement (Smiley, 2005). Earlier studies increased Al uncertainty to obtain realistic source profiles (Kim and Hopke, 2004) and hence the p_j value for Al was taken as 100%. In the absence of method detection limit values for this campaign, we matched measurement methods with those in the STN study (Kim et al., 2005a) to select appropriate values.

The signal to noise ratio (S/N) was used to adjust uncertainties. High S/N ratio greater than 2, identified NO₃⁻, Cl⁻, Na, Mg, K, Ca and

Fe as strong species, leaving their uncertainty as such. Moderate S/N ratio between 0.2 and 2 identified SO₄²⁻, Al, NH₄⁺ and BC as weak species (Table S1, Supplementary Online data), leading to 3 times increase in their uncertainty.

2.1.3. Mass balance closure on measurements

Mass balance closure showed gravimetric PM-10 mass was statistically much higher than the sum of measured species as reflected by slope of the regression line (slope of 2.48 ± 0.45 , R^2 of 0.83; Fig. S1, Supplementary Online data). Hence, PM-10 was included as an explicit species to provide independent information with a low weight (an uncertainty of 4 times its measurement) to prevent it disproportionately influencing resulting solutions (Pekney et al., 2006). The resulting PMF-2 generated source contributions and source profiles were rescaled by the PM mass concentration apportioned to the corresponding factor in order to get their values in $\mu\text{g}\mu\text{g}^{-1}$ (factor profile) and $\mu\text{g m}^{-3}$ (factor contribution) (Bhanuprasad et al., 2008).

2.2. Positive matrix factorization

Positive matrix factorization models, e.g. the two-dimensional model (PMF-2, Version 4.2) invert a matrix of time-series aerosol chemical concentration measurements using non-negativity constraints to predict source compositions and contributions (Paatero, 1997; Kim et al., 2004, 2005a; Kim and Hopke, 2004; Begum et al., 2004), prescribing individual weights for specific data points. A threshold outlier distance, typically taking a value between 2 and 4, is used to suppress the effect of outlier residuals in unduly distorting the PMF solution. A principal component analysis (PCA) based heuristic for specifying this distance (Bhanuprasad et al., 2008), used in the current work, identified no outliers in the data set leading to adoption of a value of 2 for the threshold distance.

2.3. Potential source contribution function (PSCF)

The PSCF uses back trajectories in combination with atmospheric pollutant concentrations or factor contributions identified by PMF (Hopke et al., 1995) to identify probable source locations or preferred atmospheric transport pathways from source to the receptor. The source domain is divided into a number of grid cells, typically $1^\circ \times 1^\circ$ (Kim and Hopke, 2004). The PSCF is the conditional probability of the event that a trajectory endpoint falling in the i th cell arrived at the receptor when the pollutant concentration exceeded a threshold value. A threshold criterion value of 50 percentile (median concentration) was used. The total probability

of material transfer from various heights is calculated from the conditional probabilities associated with each arrival height at the receptor (Hopke et al., 1995). Four arrival height levels of 10, 50, 500 and 1000 m were used in this study. Trajectories were computed using the HYSPLIT4 model and the archived meteorological data from National Oceanic Atmospheric Administration's (NOAA) Air Resources Laboratory (ARL) website (HYSPLIT4, 1997). Seven-day back trajectories at 2-h intervals (12 per day) were computed for each of the 29 daily samples at the four arrival heights. To correct for anomalously high PSCF values, resulting from low total number of trajectory endpoints in a given cell, a multiplicative weighting function (e.g. Bhanuprasad et al., 2008) is operated on the computed TPSCF for such cells.

3. Source identification using positive matrix factorization

3.1. Base-case solution selection

Ten random runs (corresponding to different starting points) were performed and the run with the minimum estimated Q (Kim and Hopke, 2004) value was retained for 2–6 factors. Estimated Q decreased with increasing number of factors (Fig. S2, Supplementary Online data) and its slope levelled off at 4-factors, with a small decrease on moving to a 5-factor solution. Solutions of 3, 4 and 5 factors were examined. The maximum individual column mean, IM, and the maximum individual column standard deviation IS, of the scaled residual matrix dropped sharply for solutions with 3 or more factors (Fig. S2, Supplementary Online data). Along with these indicators, the physical interpretation of factors in terms of likely number of source categories and model predictions for PM mass

were the main criteria for identifying useful solutions. The 3-factor solution had relatively poor factor resolution with mixed factors, such as *Secondary species + Biomass burning + Al*, *Sea salt* mixed with *Biomass burning + Fe + SO₄²⁻* and *Dust* with *BC*, giving poor correspondence with known aerosol sources. The 5-factor solution had insignificant loadings of most of the species in each factor, making it invalid. Additional PMF solutions excluding PM-10 as explicit species had highly mixed factors and negative coefficients on regression of factor contributions against PM mass leading to invalid solutions for 3-, 4- and 5-factors. An attempt to identify other valid solutions, by increasing the uncertainty on *BC* (by a factor of 4) and *SO₄²⁻* (by a factor of 2), was not successful.

3.2. Constraining rotational freedom

PMF can lead to multiple valid solutions as below

$$X = (GT)(T^{-1}F) + E \quad (2)$$

where T is an invertible matrix such that GT and $T^{-1}F$ satisfy non-negativity constraints. Rotations (multiplication of G and F by T and T^{-1} , respectively) are prescribed in PMF-2 through a function f_{peak} (Pekney et al., 2006). The global minimum solution from 10 pseudorandom runs of the 4-factor solution was rotated using f_{peak} values from -10 to 10 (Fig. S3, Supplementary Online data). Solutions close to the edge ($f_{\text{peak}} = 5-7$) of the stable Q values had an anomalously high PM loading (95–100%) on a single factor (*Biomass burning + Secondary species*), leading to their rejection. Therefore, the base case 4-factor solution with $f_{\text{peak}} = 0$ was retained.

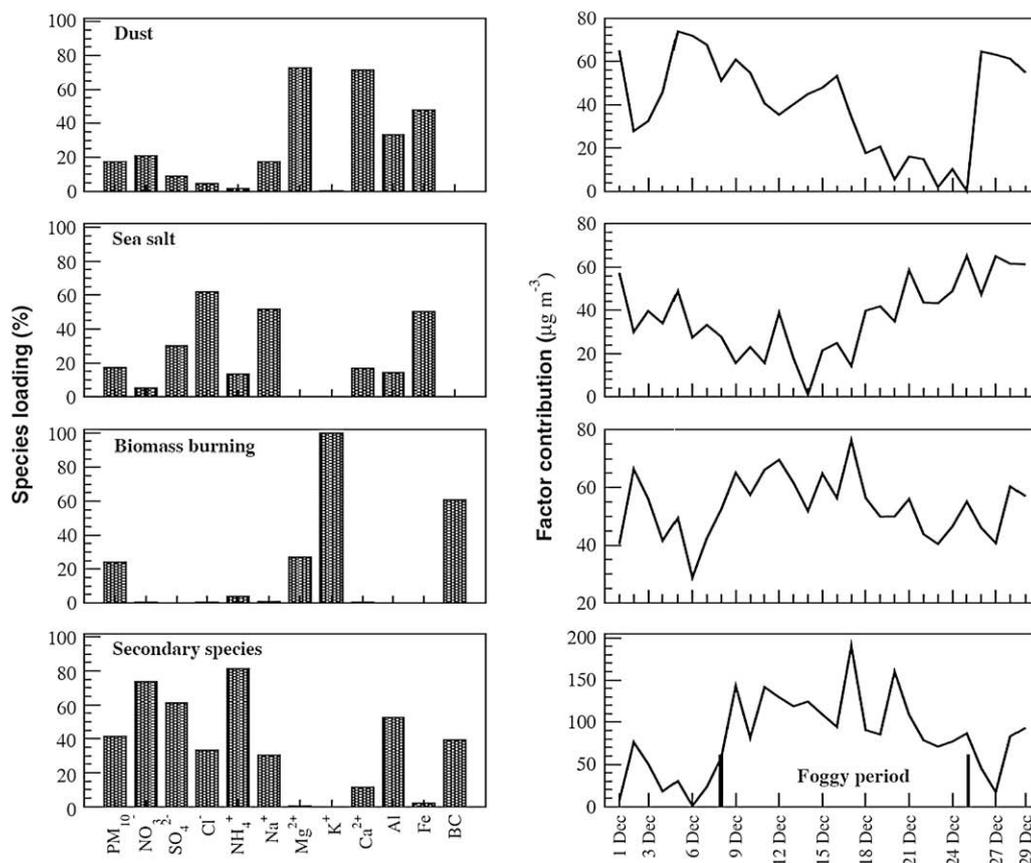


Fig. 2. Factor profiles and factor contribution time series for PMF predicted factors.

The base case 4-factor solution ($f_{\text{peak}} = 0$) had a *Dust* factor (Ca 55%, Fe 54%, Mg 49% loading), *Sea salt + Biomass burning + SO₄²⁻* factor (Na 54% and Cl⁻ 49% along with 31% SO₄²⁻, 30% Fe, 28% K and 24% BC loading), *Biomass burning* factor contaminated with *Dust* (52% Al, 42% Mg, 36% Ca, 32% K and 27% BC loading), *Secondary species + Biomass burning + chloride* factors (NH₄⁺ 68%, NO₃⁻ 56%, BC 42%, SO₄²⁻ 38%, K 27% and Cl⁻ 27% loading). The fitted straight line between predicted and measured PM had satisfactory goodness-of-fit (slope of 0.92 ± 0.17 and R^2 of 0.83). Additionally, 8 species were predicted with high (>0.7) R^2 values: NO₃⁻, Cl⁻, Mg, K, Ca²⁺, Na, NH₄⁺, and Fe, thereby indicating model's ability to adequately explain most of the measured mass.

When a factor (representing a source) is clearly "contaminated" by a species which is not expected from that source, the function f_{key} in PMF-2 is used to pull down these unrealistic species. Solutions were obtained by operating f_{key} only on single factors and then on two, or three factors to resolve mixed sources. The final selected solution had an f_{key} of 10 on K, BC on *Sea salt + Biomass burning + SO₄²⁻* factors, f_{key} of 10 on K on *Secondary species + Biomass burning* factors and f_{key} of 5 on K, BC on *Dust* factor. An f_{key} value of 5 on Al and Na and 3 on SO₄²⁻ on the *Biomass burning* factor was also used. The solution contained a *Dust* factor (Mg 73%, Ca 72%, Fe 48% and Al 33% loading), *Sea salt* factor (Cl⁻ 62%, Na 52%), *Biomass burning* factor (K 100%, BC 61%) and *Secondary species* factor (NH₄⁺ 81%, NO₃⁻ 74%, SO₄²⁻ 61%, BC-39% – Fig. 2). The predicted PM-10 loading on different factors was higher on *Secondary species* (39%) and *Biomass burning* (25%) followed by *Dust* (19%) and *Sea salt* factor (17%). This solution had satisfactory goodness-of-fit between predicted and measured PM-10 (slope of 0.83 ± 0.17 and R^2 of 0.8 – see Supplementary Online data Fig. S4; prediction of 7 species with R^2 values (>0.7): NO₃⁻, Cl⁻, Mg, K, Ca, NH₄⁺, Fe – Table S2). Regression of the measured species concentrations against PMF-2 estimated factor contributions for tracer species correlated well with *Dust*, *Biomass burning* and *Secondary species* factors. For example, Mg (R^2 : 0.95) and Ca (R^2 : 0.76) are tracer species for *Dust* while K (R^2 : 0.95) is a tracer for *Biomass burning* factor.

4. Interpretation of factors

4.1. Identification of source categories

The *Secondary species* factor had 81% NH₄⁺, 74% NO₃⁻, 61% SO₄²⁻, 52% Al and 39% BC loading (Fig. 2). Based on the relative mass concentration of different chemical species loading on a factor, termed "factor composition," the factor had a composition of 35% NO₃⁻, 25% SO₄²⁻, 21% NH₄⁺ and 11% BC. Satisfactory goodness-of-fit was obtained between measured tracer (NO₃⁻, NH₄⁺) concentrations with both their PMF-2 predicted concentrations and the corresponding factor contributions. Mean *Secondary species* factor contributions ($107 \mu\text{g m}^{-3}$) were significantly higher during the foggy period and lower ($36 \mu\text{g m}^{-3}$) during the clear period (Table 2), indicating that aerosol constituents of this factor are strongly linked to fog formation, as corroborated by earlier studies on the fog-smog-fog cycle (Pandis et al., 1990). This factor contained a significant loading of Al. There is a cluster of Al smelters to the south-east of Kanpur (Fig. 1), which could explain the association of industrially emitted Al along with sulphate from SO₂ emitted by coal combustion in the smelter.

The *Biomass burning* factor had 100% K, 61% BC, 27% Mg and 4% NH₄⁺ loadings. The factor had a composition of 53% BC and 43% K. The model predicted K well (slope of 1.05, R^2 of 0.95) and BC moderately well (slope of 0.73, R^2 of 0.63). The mean *Biomass burning* factor contributions were somewhat higher ($57 \mu\text{g m}^{-3}$) during the foggy period than ($47 \mu\text{g m}^{-3}$) the clear period (Table 2).

Table 2

Average PM concentrations from 4-factors predicted by PMF during the foggy and clear periods in Kanpur.

Source name	Average contribution			
	Foggy period		Clear period	
	$\mu\text{g m}^{-3}$	%	$\mu\text{g m}^{-3}$	%
Dust	32	14	57	31
Sea-salt	34	14	44	24
Biomass burning	57	25	47	25
Secondary species	107	47	36	20

The *Dust* factor had a high loading of crustal elements (73% Mg, 72% Ca, 48% Fe and 33% Al). The factor had a composition of 14% Ca, 11% Fe, 2% Mg and 2% Al, along with acidic species NO₃⁻ (44%) and SO₄²⁻ (16%). This is consistent with reported soil composition in the region and surface reactions of acidic species on alkaline soil constituents (Rastogi and Sarin, 2005). Satisfactory goodness-of-fit was obtained between measured tracer (Ca, Mg, Fe) concentrations with PMF-2 predicted concentrations of these species and the *Dust* factor contribution. Mean *Dust* factor contributions were lower ($32 \mu\text{g m}^{-3}$) during the foggy period (8–28 December except 26 and 27 December) and higher ($57 \mu\text{g m}^{-3}$) during the clear period (1–7, 26, 27 and 29 December, Table 2), consistent with the expectation that high dust concentrations were probably accompanied by greater dispersion, not conducive for fog formation.

The *Sea salt* factor had 62% Cl⁻ and 52% Na, 50% Fe and 30% SO₄²⁻ loadings. The factor had a composition of 18% Na, 15% Cl⁻ along with 38% SO₄²⁻, 11% NH₄⁺, 8% Fe and 7% NO₃⁻. The significant sulphate loading is consistent with the low Cl⁻/Na⁺ ratio in the observations from Cl⁻ depletion through atmospheric reactions of SO₂ or NO_x with sea-salt. An electroneutrality check on the factor composition supported these levels of sulphate and gave a cation to anion ratio of 1.08. Moderate goodness-of-fit was obtained between measured tracer (Na, Cl⁻) concentrations with PMF-2 predicted concentrations of these species and the *Sea salt* factor contribution. NaCl has been associated with dust in some locations in India (Rastogi and Sarin, 2005), potentially explaining a part of the loading of Fe on this factor. Additionally, an iron and steel plant near Kanpur and a cluster of plants south of Kanpur (Fig. 1) could be a source of both Fe and SO₂ emissions, contributing to the association of both Fe and sulphate with this factor. Iron is a primary pollutant while sulphate results by Cl⁻ depletion from atmospheric reactions of SO₂ with sea-salt discussed previously. Mean *Sea salt* factor contributions were also lower ($34 \mu\text{g m}^{-3}$) during the foggy period than ($44 \mu\text{g m}^{-3}$) the clear period similar to *Dust* (Table 2).

In view of the small data set available (29 samples and 12 species) we examined the relative standard deviation in the PMF-2 predictions of the mean factor contributions. Only strong species (S/N ratio greater than 2 in Table 1) were considered representative of the errors in PMF model prediction following Bhanuprasad et al. (2008). This was 11–12% for the foggy and clear periods (Table S3, Supplementary Online data) and indicates satisfactory model performance, even with the small data set.

4.2. Validation of factors using the emission strength of sources in the foggy and clear periods

As seen in the previous section, the *Secondary species* factor had a significantly larger (2.9 ± 0.3) mean concentration during the foggy versus the clear period. The biomass burning factor showed a moderately higher (1.2 ± 0.09) concentration during the foggy period (refer to Section D, Supplementary online data, for details of standard deviation calculation). PMF predicted factors were validated by contrasting the emission strength of sources in the foggy

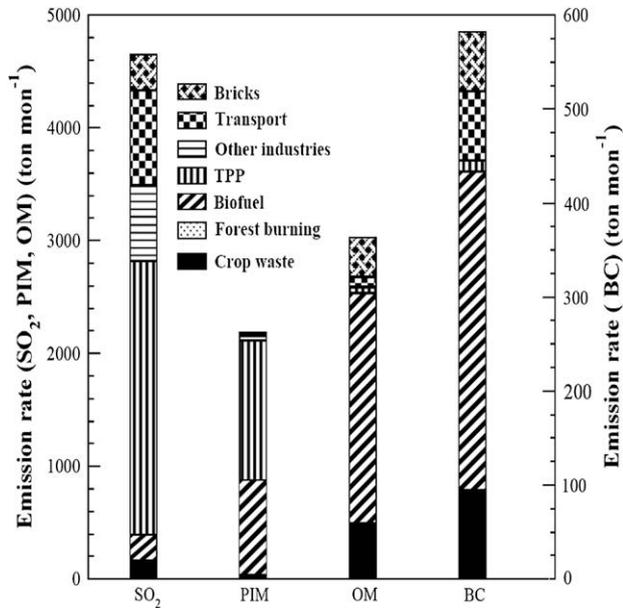


Fig. 3. Aggregated emission rate (tons mon⁻¹) of SO₂, particulate inorganic matter (PIM), organic matter (OM) and black carbon (BC), in a 200 km × 200 km area around Kanpur, resolved by emission sector.

and clear periods, using a combination of back trajectory modeling and quantitative emission inventory information. A part of the difference in mean concentration would be related to tight ground-level temperature inversions and a low mixing height during the foggy period. However, the differences in emissions from

contributing sources during the foggy and clear periods, especially those responsible for the *Secondary species* factor are important and are further examined in this section.

4.2.1. Influence of proximate/local sources

Aerosol and SO₂ emission fluxes were calculated in 200 km × 200 km box around Kanpur (79.5–81.25°E and 25.5–27.25°N) aggregated from sources in our 25 km × 25 km India emissions inventory (Reddy and Venkataraman, 2002; Venkataraman et al., 2006). The inventory is on a monthly mean temporal scale and therefore cannot diagnose differences in emissions from local sources between the foggy and clear periods. We therefore identify dominant local sources based on the monthly mean emission flux of four pollutants: SO₂, particulate inorganic matter (PIM) which includes water-soluble ions and mineral matter like fly-ash, particulate organic matter (OM), and black carbon (BC). Among these pollutants, the atmospheric reactions of SO₂ to sulfate are known to affect fog formation (Pandis et al., 1990). PIM would contain soluble inorganic ions (Cl⁻, NO₃⁻, SO₄²⁻, NH₄⁺) present in combustion aerosols (e.g. Habib et al., 2008), which could also mediate water uptake and fog formation. The role of organic matter in water uptake by aerosol particles is still uncertain. Secondary organic matter from atmospheric reactions of hydrocarbons is believed to increase hygroscopicity beyond ionic content (Saxena et al., 1995), but primary organic matter, e.g. from forest biomass burning (Carrico et al., 2005) is believed to inhibit hygroscopic growth. BC may be associated with other hygroscopic compounds and provide condensation nuclei to mediate fog formation.

Among sources that affect the *Secondary species* factor, coal based thermal power plants (TPPs), transportation and other large

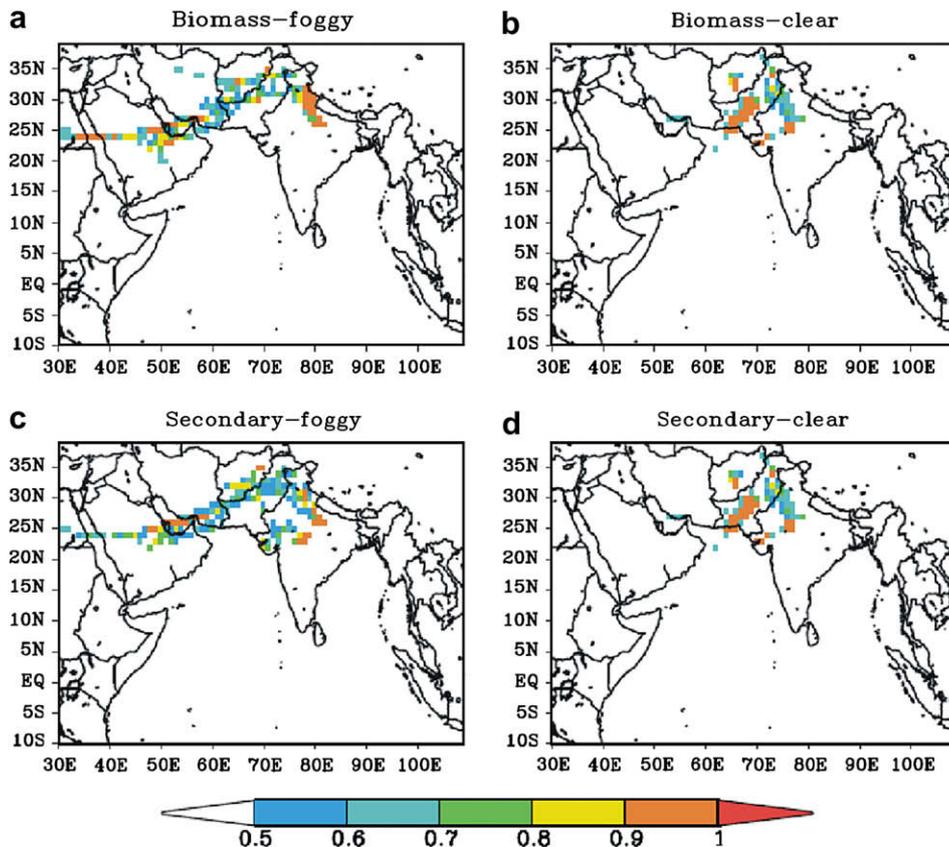


Fig. 4. Probable source regions from TPSCF analysis during the foggy and clear periods for the biomass burning (a and b) and secondary species (c and d) factors.

point sources (LPSs) were dominant accounting, respectively, for 2428, 841 and 675 ton mon^{-1} of SO_2 and 1240, 23 and 40 ton mon^{-1} of PIM, respectively (Fig. 3). There are 2 coal based thermal power plants, 5 fertilizer plants, 1 cement plant and 1 iron and steel plant, close to Kanpur. Transportation (diesel powered vehicles) influences BC emissions (74 ton mon^{-1}) and brick kilns influence OM emissions (349 ton mon^{-1}). Among source categories that affect the *Biomass burning* factor, biofuel combustion and crop waste burning were significant, respectively, emitting 2049 ton mon^{-1} and 483 ton mon^{-1} of OM and 339 and 94 ton mon^{-1} of BC. In summary, local sources of significance are thermal power plants to SO_2 and PIM emissions, biofuel combustion to PIM, BC and OM emissions, transportation to SO_2 , PIM and BC emissions and crop waste burning and brick kilns to OM and BC emissions.

4.2.2. Influence of sources in PSCF identified source regions in the foggy and clear periods

The TPSCF (Section 2.3) identifies probable source regions and transport pathways that affect the predicted factor contributions during a given time period. Distinct potential source regions were identified by TPSCF for the foggy and clear periods for the *Biomass burning* and *Secondary species* factors. Areas with PSCF between 0.5 and 1.0 in each period were designated as probable source regions during that period (Fig. 4). We aggregate the emission fluxes in the probable source region for each period (foggy and clear) for each source to estimate its relative emission strength. Median or 50th percentile values of PMF predicted factor contribution were used as the criterion value for the PSCF calculations.

During the foggy period, TPSCF plots (Fig. 4) for the factors of anthropogenic origin, the *Biomass burning* and *Secondary species* factors, both had probable source regions in the northwestern IGP (Indian states of Uttar Pradesh, Utaranchal, Haryana, Punjab and Himachal Pradesh) and in parts of west-Asia, i.e., Pakistan, Iran, UAE and Oman. In addition, the *Secondary species* factor had probable source regions to the south-west of Kanpur in the Indian states of Rajasthan, Gujarat and Madhya Pradesh (Fig. 4). During the clear period, probable source regions for both factors lay to the north and north-west of Kanpur.

We contrast the mean emission fluxes from different sources located in these probable source regions during foggy versus clear period to identify significant sources. Thermal power plants, other

LPS, transportation and brick kilns located in probable source regions had significantly larger emission fluxes of SO_2 during the foggy versus the clear period (factor of 1.93, 3.53, 2.35 and 3.69, respectively). Individual point sources likely to have contributed to SO_2 emissions during the foggy period are coal-fired thermal power plants located in Kanpur, Ghaziabad, Aligarh, Jhansi, Rupnagar and Delhi, iron and steel plants located in Kanpur and Bulandshahar and fertilizer plants located in Kanpur, Ghaziabad, Buduan, Bulandshahar, Firozabad. Black carbon could be associated with fog formation, if coated with hygroscopic aerosol species. The emission flux of BC and OM from brick kilns in probable source areas was substantially larger during the foggy period (Fig. 5). These include brick kilns located in Kanpur, Ghaziabad, Agra and Ludhiana districts. Since brick kilns use a mixture of locally available biomass and powdered coal, the BC and OM emissions from this source may also be associated with SO_2 emissions (from the coal) and therefore have sufficient hygroscopicity to be potential fog nuclei. Large areas of probable source regions for both *Secondary species* and *Biomass burning*, lay outside India (Pakistan, Iran, Oman and UAE) in the foggy period, while few probable source regions lay outside India during the clear period. This implies the significance of aerosol transport from outside India to fog formation in the IGP. However, lack of detailed emission inventory information for these regions does not permit further analysis of sources which may have contributed to the fog events.

5. Conclusions

PMF identified factors were analyzed to explain influence of source categories on fog formation in Kanpur during the ISRO-GBP land campaign-II (LC-II) in December 2004. PMF predicted factors were validated by comparing the emission strength of sources during the foggy and clear periods, using a combination of PSCF analysis and quantitative emission inventory information. The time series of 29 days and 12 species was decomposed to identify 4-factors: *Secondary species*, *Biomass burning*, *Dust* and *Sea salt*.

The ratios of mean contribution during the foggy to the clear period were significantly larger than unity for the anthropogenic *Secondary species* (2.9 ± 0.3) and *Biomass burning* (1.2 ± 0.09) factors. The larger predominance of the secondary species factor than the biomass burning factor during fog events is somewhat surprising. Significant emission fluxes from biomass burning sources in the Indo-Gangetic plain are estimated in emission inventories (e.g. Venkataraman et al., 2006). The predominance of secondary species during fog events therefore suggests significant secondary aerosol formation through aqueous-phase reactions in fog droplets. Further understanding of aerosol composition before, during and after fog events is needed to investigate such mechanisms. Since the *Secondary species* factor had high loadings of sulfate, nitrate and ammonium, sources of these aerosol precursors (thermal power and fertilizer plants, transport and cattle rearing) would have a large bearing on fog formation.

Based on their relative emission flux strength, thermal power plants and transportation emissions, on both local and regional scales, brick kilns on a regional scale and local biofuel combustion are the likely contributors to aerosols that mediated fog events at Kanpur during the ISRO-GBP land campaign. Biofuel combustion and transportation are distributed area sources. Contributing point sources include coal-fired thermal power plants located in Aligarh, Delhi, Ghaziabad, Jhansi, Kanpur, Rae Bareli and Rupnagar and brick kilns located in Allahabad, Agra, Farrukhabad, Ghaziabad, Kanpur, Ludhiana, Lucknow and Rae Bareli. This identifies industrial and transport emissions of SO_2 and NO_x and cattle-related emissions of NH_3 , as a relevant factor in fog formation. A more complete inventory especially for NH_3 from industrial sources and cattle

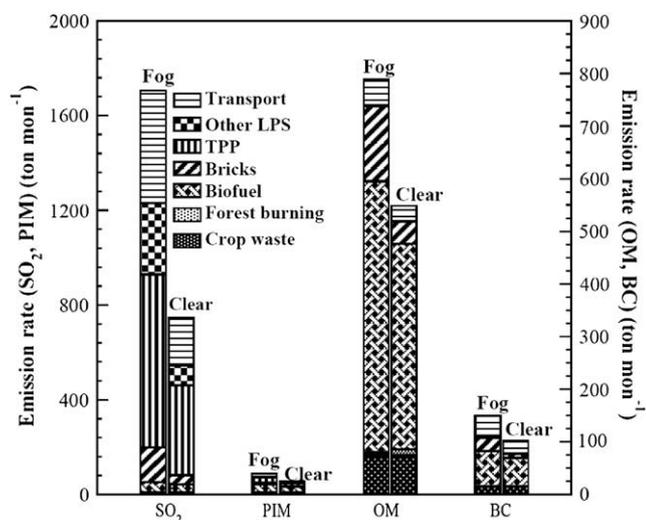


Fig. 5. Aggregated emission rate (tons mon^{-1}) of SO_2 , particulate inorganic matter (PIM), organic matter (OM) and black carbon (BC), from sources in TPSCF identified source regions during the foggy (labeled "fog") and clear (labeled "clear") periods.

rearing activities is needed to understand sources in more detail. Large areas of probable source regions for both *Secondary species* and *Biomass burning* factors in foggy period, lay outside India (Pakistan, Iran, Oman and UAE), suggesting the significance of regional scale aerosol transport from outside India to fog formation in the Indo-Gangetic Plain.

Acknowledgments

We thank the Indian Space Research Organisation's Geosphere Biosphere Programme (ISRO-GBP), Bengaluru, India, for enabling this collaborative work and Ribu Cherian for assistance with emissions data processing and figure preparation.

Appendix A. Supplementary information

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.atmosenv.2008.11.041.

References

- Ali, K., Momin, G.A., Tiwari, S., Safai, P.D., Chate, D.M., Rao, P.S.P., 2004. Fog and precipitation chemistry at Delhi, North India. *Atmos. Environ.* 38, 4215–4222.
- BAHS, 2003. Basic Animal Husbandry Statistics, 17th Livestock Census Data. Department of Animal Husbandry Dairying and Fisheries, Government of India, New Delhi.
- Begum, B.A., Kim, E., Biswas, S.K., Hopke, P.K., 2004. Investigation of sources of atmospheric aerosol at urban and semi-urban areas in Bangladesh. *Atmos. Environ.* 38, 3025–3038.
- Bhanuprasad, S.G., Venkataraman, C., Bhushan, M., 2008. Positive matrix factorization and trajectory modeling for source identification: a new look at Indian Ocean Experiment ship observations. *Atmos. Environ.* 42, 4836–4852.
- Carrico, C.M., Kreidenweis, S.M., Malm, W.C., Day, D.E., Lee, T., Carrillo, J., McMeeking, G.R., Collett Jr., J.L., 2005. Hygroscopic growth behavior of a carbon-dominated aerosol in Yosemite national park. *Atmos. Environ.* 39, 1393–1404.
- Gupta, U., 2006. Valuation of Urban Air Pollution: a Case Study of Kanpur City in India. South Asian Network for Development and Environmental Economics. Available from: <http://www.sandeeonline.org/publications/working_papers/wp17/working_paper_17.pdf>.
- Habib, G., Venkataraman, C., Bond, T.C., Schauer, J.J., 2008. Climate relevant properties of primary particles from the combustion of biofuels. *Environ. Sci. Technol.* 42 (23), 8829–8834.
- HYSPLIT4 (Hybrid Single-Particle Lagrangian Integrated Trajectory), Model, 1997. NOAA Air Resources Laboratory, Silver Spring, MD. Available from: <<http://www.arl.noaa.gov/ready/open/hysplit4.html>>.
- Kim, E., Hopke, P.K., 2004. Improving source identification of fine particles in a rural northeastern U.S. area utilizing temperature-resolved carbon fractions. *J. Geophys. Res.* 109, D09204. doi:10.1029/2003JD004199.
- Kim, E., Hopke, P.K., Edgerton, E.S., 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. *Atmos. Environ.* 38, 3349–3362.
- Kim, E., Hopke, P.K., Qin, Y., 2005a. Estimation of organic carbon blank values and error structures of the speciation trends network data for source apportionment. *J. Air Waste Manage. Assoc.* 55, 1190–1199.
- Kim, E., Hopke, P.K., Kenski, D.M., Koerber, M., 2005b. Sources of fine particles in a rural midwestern U.S. area. *Environ. Sci. Technol.* 39, 4953–4960.
- Kirchstetter, T.W., Novakov, T., 2007. Controlled generation of black carbon particles from a diffusion flame and applications in evaluating black carbon measurement methods. *Atmos. Environ.* 41, 1874–1888.
- McGinn, S.M., Flesch, T.K., Crenna, B.P., Beauchemin, K.A., Coates, T., 2007. Quantifying ammonia emissions from a cattle feedlot using a dispersion model. *J. Environ. Qual.* 36, 1585–1590.
- Paatero, P., 1997. Least squares formulation of robust, non-negative factor analysis. *Chemometr. Intell. Lab. Syst.* 37, 23–35.
- Pandis, S.N., Seinfeld, J., Pilinis, C., 1990. The smog–fog–smog cycle and acid deposition. *J. Geophys. Res.* 95 (D11), 18489–18500.
- Pasricha, P.K., Gera, B.S., Shastri, S., Maini, H.K., Ghosh, A.B., Tiwari, M.K., Garg, S.C., 2003. Role of the water vapor greenhouse effect in the forecasting of the fog occurrence. *Boundary Layer Meteorol.* 107 (2), 469–482.
- Pekney, N.J., Davidson, C.I., Robinson, A., Zhou, L., Hopke, P.K., Eatough, D., Rogge, W.F., 2006. Major source categories for PM_{2.5} in Pittsburgh using PMF and UNMIX. *Aerosol Sci. Technol.* 40, 910–924.
- Pruppacher, H.R., Klett, J.D., 1997. In: *Microphysics of Clouds and Precipitation*, second revision ed. Kluwer Academic Publishers, pp. 10–15.
- Rastogi, N., Sarin, M.M., 2005. Long-term characterization of ionic species in aerosols from urban and high-altitude sites in western India: role of mineral dust and anthropogenic sources. *Atmos. Environ.* 39, 5541–5554.
- Reddy, M.S., Venkataraman, C., 2002. Inventory of aerosol and sulfur dioxide emissions from India: I fossil fuel combustion. *Atmos. Environ.* 36, 677–697.
- Saxena, P., Hildermann, L.M., McMurry, P.H., Seinfeld, J.H., 1995. Organics alter hygroscopic behavior of atmospheric particles. *J. Geophys. Res.* 100 (D9), 18755–18770.
- Seinfeld, J.H., Pandis, S.N., 1998. *Atmospheric Chemistry and Physics*, vol. 15. Wiley-Interscience, 793–801 pp.
- Smiley, J., 2005. Experimental Inter-comparison of Speciation Laboratories. Technical Memorandum. U.S. Environmental Protection Agency. Available from: <<http://www.epa.gov/ttn/amtic/files/ambient/pm25/spec/multilab06.pdf>>.
- Tare, V., Tripathi, S.N., Chinnam, N., Srivastava, A.K., Dey, S., Manar, M., Kanawade, V.P., Agarwal, A., Kishore, S., Lal, R.B., Sharma, M., 2006. Measurements of atmospheric parameters during Indian Space Research Organization Geosphere Biosphere Program land campaign II at a typical location in the Ganga basin: 2. Chemical properties. *J. Geophys. Res.* 111, D23210. doi:10.1029/2006JD007279.
- Tripathi, S.N., Tare, V., Chinnam, N., Srivastava, A.K., Dey, S., Agarwal, A., Kishore, S., Lal, R.B., Manar, M., Kanawade, V.P., Chauhan, S.S.S., Sharma, M., Reddy, R.R., Rama Gopal, K., Narasimhulu, K., Siva Sankara Reddy, L., Gupta, S., Lal, S., 2006. Measurements of atmospheric parameters during Indian Space Research Organization Geosphere Biosphere Programme land campaign II at a typical location in the Ganga basin: 1. Physical and optical properties. *J. Geophys. Res.* 111, D23209. doi:10.1029/2006JD007278.
- Venkataraman, C., Habib, G., Kadamba, D., Shrivastava, M., Leon, J.F., Crouzille, B., Boucher, O., Streets, D.G., 2006. Emissions from open biomass burning in India: integrating the inventory approach with high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) active-fire and land cover data. *Global Biogeochem. Cycles* 20, GB2013. doi:10.1029/2005GB002547.