The impact of crop residue burning (CRB) on the diurnal and seasonal variability of the ozone and PM levels at a semi-urban site in the north-western Indo-Gangetic plain

Madhvi Rana[1], Susheel K Mittal[1,*], Gufran Beig[2] and Prashant Rana[3]

[1]*School of Chemistry and Biochemistry, Thapar University, Patiala 147 004, India.*
[2]*Indian Institute of Tropical Meteorology, Pashan, Pune 411 008, India.*
[3]*Department of Computer Science and Engineering, Thapar University, Patiala 147 004, India.*
Corresponding author. e-mail: smittal@thapar.edu

Ozone and particulate matter (PM), PM_{10} and $PM_{2.5}$, were monitored along with meteorological parameters at a semi-urban location, Patiala, in the north-western Indo-Gangetic plain from December 2013 to November 2014. The annual mean concentration levels of PM_{10}, $PM_{2.5}$ and ozone were recorded as 178 µg m^{-3}, 88 µg m^{-3} and 39 ppb, which also exceeded the national standards on a 24-h average basis, by 77.3% (335 days), 53.2% (338 days) and 20 days, respectively. High levels of PM (PM_{10} and $PM_{2.5}$) were observed in winter (58–381 and 42–270 µg m^{-3}) and in the post-monsoon (71–320 and 39–320 µg m^{-3}) season and a rise in the level of ozone was observed in summer (22–72 ppb) and in the post-monsoon season (23–73 ppb), respectively. The rate of ozone production was the highest during the post-harvest fire period (3.94 ppb O_3/h in May and 4.23 ppb O_3/h in November). A Pearson correlation study showed the strong dependency of PM and ozone on meteorological variables. Relative humidity has the highest ranking for ozone and PM_{10}, while wind speed has the lowest rank for ozone, PM_{10} and $PM_{2.5}$ in the order of factors affecting the level of pollutants. The generalised linear model and the neural network model (for ozone) and the random forest model (for PM) outperformed on the basis of performance indices and further cross-validation was done.

Keywords. Indo-Gangetic plain; ozone; PM10 and PM2.5; meteorology; correlation analysis; linear and nonlinear regression analysis.

1. Introduction

The Indo-Gangetic plain (IGP) ranging from 21.75°N, 74.25°E to 31.0°N, 91.50°E accounts for 21% of the land area of the Indian subcontinent and holds nearly 40% of the total population which gives rise to the worst air quality in the region (Badarinath *et al.* 2009). WHO reported the data of critically polluted cities, stretching from the west to the east of northern India (WHO 2018). Among the most significant contributors of air pollution such as automobiles, industries and domestic emissions, post-harvest crop residue burning (CRB) is an episodic contributor to the level of pollutants. Rice and wheat are the major crops, which generate a

large amount of leftover crop residue, intensified by the usage of mechanical combine harvester technology, which is further subjected to open field CRB due to the short duration between rice harvesting and wheat sowing, labour shortage, uncertainty of weather and no economical technologies are used for collecting leftover agricultural residues from the field (Jain *et al.* 2014; Kumar *et al.* 2016). The harvesting of crops is done in the months of April and October for wheat and rice, respectively, whereas the CRB from harvesting is done from mid-April to May (in the summer season) and from mid-October to mid-November (in the post-monsoon season) for the respective crop seasons. In these months, the level of pollutants is higher than the permissible limit, especially when the weather conditions are unfavourable for the dispersion of air pollutants. CRB, due to less than ideal combustion conditions, typically produces a thick cloud of smoke, clearly visible over the IGP region which contains micron-sized aerosols, trace gases (NO_x, CO and volatile organic compounds) and secondary species, such as ozone, peroxy-acetyl nitrate (PAN) and acid rain, can have a negative impact on human health, environment, climate and agricultural crop yield to a large extent (Galanter *et al.* 2000; Mittal *et al.* 2009; Awasthi *et al.* 2011; Agarwal *et al.* 2013; Kaskaoutis *et al.* 2014; Ashworth *et al.* 2015; Sinha *et al.* 2015).

Furthermore, the pollutant emissions are subject to alter on account of the changing meteorological factors playing a critical function in the transport, formation or accumulation of primary and secondary pollutants, depending on the geographical and topographical conditions of the region (Han *et al.* 2011; Yoo *et al.* 2011). Possible associations among the particulate matter (PM) and the secondary pollutant (ozone) with meteorological parameters can reveal the sources affecting the chemical generation and deposition of pollutants on the surface (Tarasova and Karpetchko 2003; Vardoulakis and Kassomenos 2008; Yoo *et al.* 2011) and thereby contribute to air pollution episodes during CRB periods (Mittal *et al.* 2009; Kumar *et al.* 2016). Hence, the assessment of the behaviour with PM and ozone levels is obligatory to figure out the observed trends and to analyse the controlling nature of meteorology, with regard to the pollution levels in the city.

For that reason, the air quality in cities has been statistically correlated with a combination of various meteorological factors (Banerjee *et al.* 2011;

Gunthe *et al.* 2016). Pearson correlation analysis was performed to determine the relationship between pollutants and meteorological variables, and further linear and nonlinear regression analyses were performed to determine the contribution of meteorological variables to air pollution and utilise them in air pollution prediction (Ghazali *et al.* 2009; Nishant and Kumar 2011; Dubey *et al.* 2014). Conclusively, to evaluate the model's performance, the predicted and monitored concentrations were statistically compared.

In this study, we present the measurements of surface ozone and PM (PM_{10} and $PM_{2.5}$) and meteorological variables (relative humidity (RH), total solar radiation (SR), average temperature (AT), rainfall, wind speed (WS) and wind direction) at Patiala, a semi-urban site, from December 2013 to November 2014. Based on data, temporal variations and interrelationships of criteria pollutants were characterised and the influence of the meteorological conditions was analysed. Correlation analysis determined the relationship between pollutants and meteorological variables. Furthermore, a ranking of meteorological parameters was done based on their influence on the level of pollutants. Regression analysis was done for the identification of the best prediction approach for ozone and PM with meteorology as inputs using 'R' software (Muezzinoglu 1987; Icaga and Sabah 2009).

2. Site description

The data used in this work were collected from continuous ambient air quality monitoring station situated at Thapar University, Patiala (30.35°N, 76.36°E, 250 m above mean sea level) with the expertise of the Indian Institute of Tropical Meteorology (IITM), Pune by Envirotech Online Equipments Pvt. Ltd., Hyderabad on a continuous basis under the MAPAN (modelling atmospheric pollution and networking) project. The station is installed on the second floor and the inlet system of analysers and samplers is installed at the rooftop (45 feet height) from the ground level. The university campus has a chain of buildings including residential houses and interlinked roads with trees planted on both sides and scattered grassy lawns. The observational site is shown in figure 1. As per the data of the 2011 census, Patiala city had a population of 406,192.

Patiala is situated in the north-western part of the IGP, close to the Shivalik hills in the east and
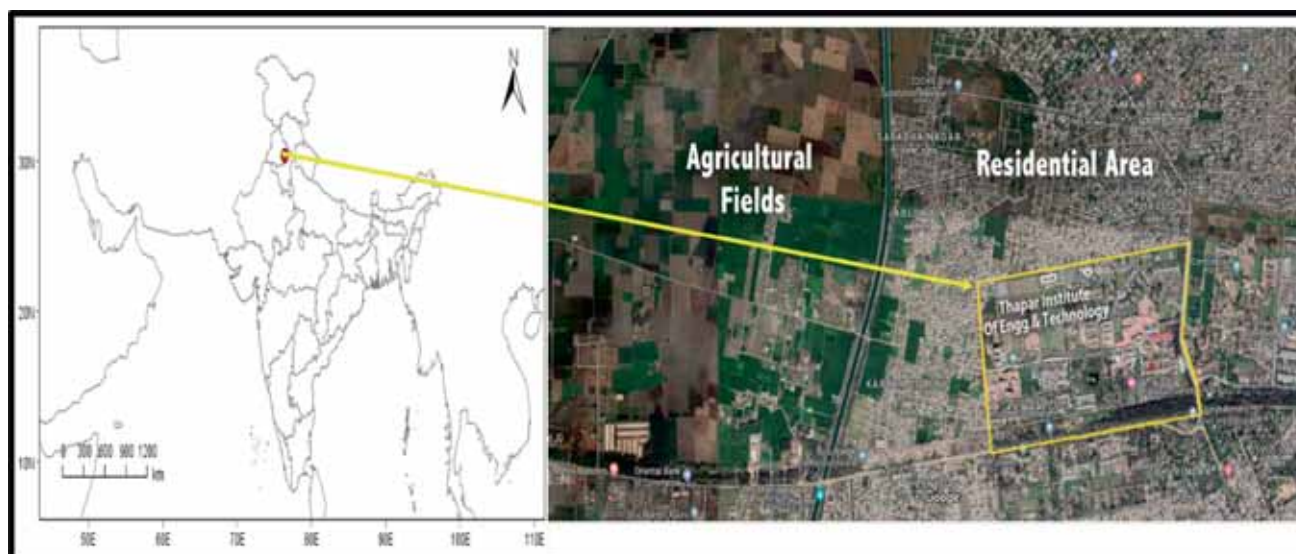
Figure 1. Geographical location of the continuous ambient air quality monitoring station situated at Thapar University, Patiala (India).
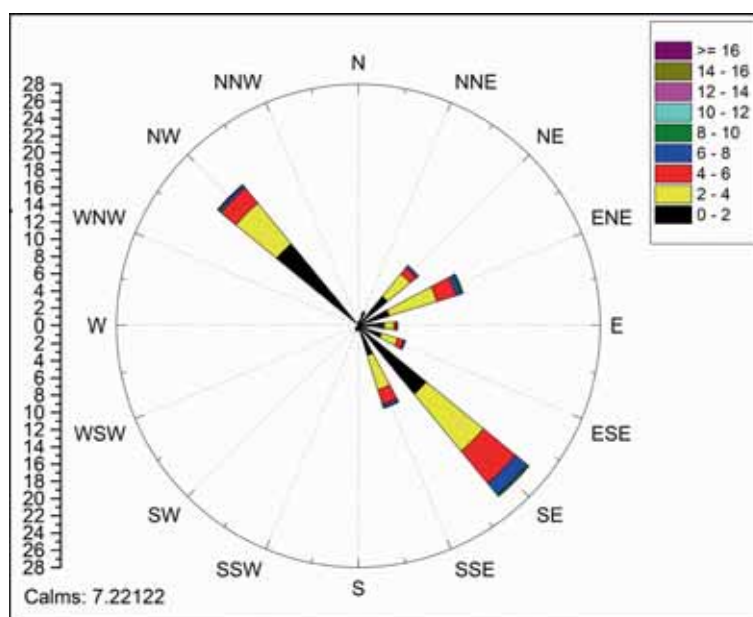

Figure 2. Wind rose plot showing the wind direction, WS and wind frequency in the study area for the period December 2013–November 2014.

the Thar desert in the south-west. It is bounded by the districts Fatehgarh Sahib and Sahibzada Ajit Singh Nagar (SAS Nagar), the town Mandi–Gobindgarh and the union territory of Chandigarh in the north; the districts of Sangrur (Punjab) in the west, Jind (Haryana) in the south and Ambala (Haryana) in the east. The IGP region experiences four dominant seasons each year: winter (December–February), summer (March–May), monsoon (June–September) and post-monsoon (October–November). The climate of the city is very hot in the summer and very cold in the winter. The region is generally dry and hot with a maximum temperature of 43°C during May. The monsoon season lasts for 3 months with an annual rainfall of 600 mm. January is the coldest month with a mean monthly minimum temperature of 2°C.

Wind roses for the corresponding study period have also been plotted in figure 2 and taken into consideration while analysing the data. These wind plots show the wind direction with a frequency
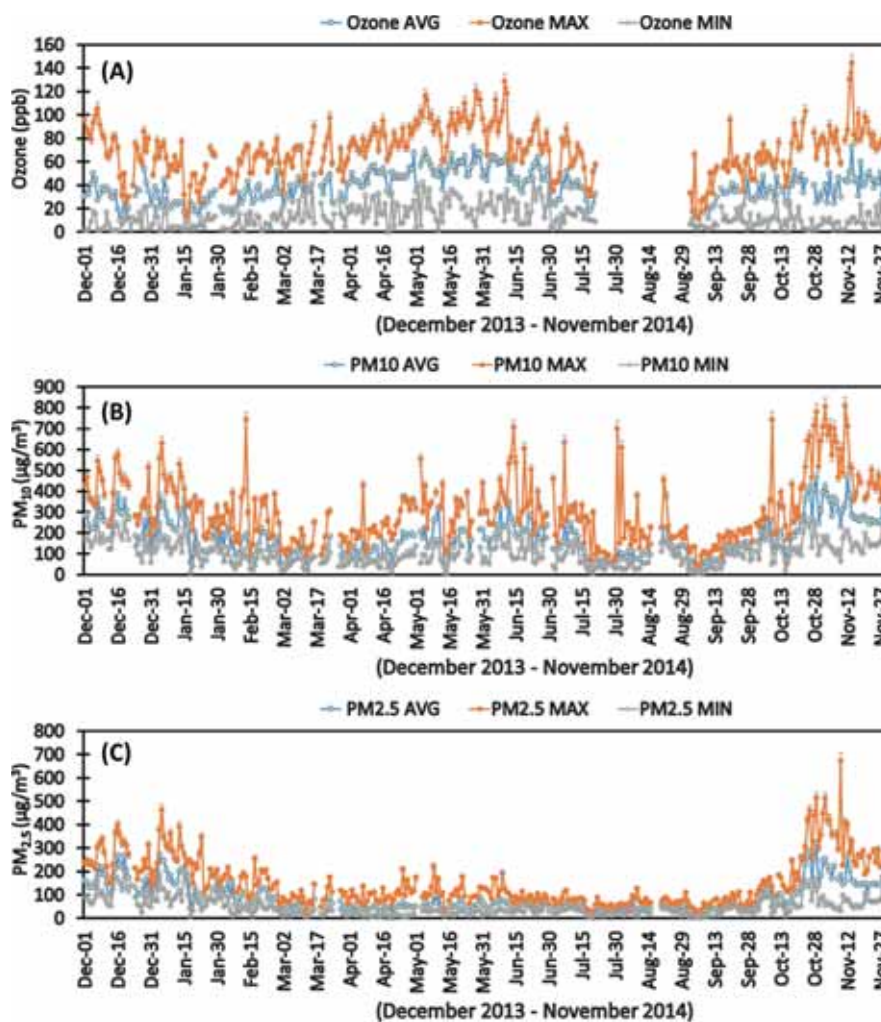
Figure 3. From December 2013 to November 2014: 24-h averaged daily maximum and daily minimum level of pollutants (ozone, PM_{10} and $PM_{2.5}$) and meteorological variables (RH, SR and daily AT).

of wind blowing along with the WS (m/s) based on hourly wind data. The length of each 'spoke' around the circle is related to the frequency of time that the wind blows from a particular direction. The y-axis represents the increasing percentage frequencies of each bin, emanating from zero at the centre. The fetch region ranges from NNW to SW. It can be seen that the most prevalent wind direction at the site is from the NW followed by NNW, WSW and SW.

Anthropogenic emissions of both trace gases and aerosols from industries, coal-power plant (Rajpura road) and vehicles from moderately heavy traffic roads on NH-64 (north-east), NH-1(north) and state highway no. 8 are present throughout the year in this region. The burning of agricultural waste in open fields for crop rotation is a common practice on a seasonal basis in the north-western region of the IGP (Gupta *et al.* 2004; Badarinath *et al.*

2006; Punia *et al.* 2008). Intense fire activities were observed during the post-monsoon season, from the second week of October until the end of November, and in summer, from the first week of May until mid-June.

3. Instrumental methodology

In this study, the daily air quality data of O3, PM_{10} and $PM_{2.5}$ and the surface meteorological variables (WS, wind direction, AT, SR and RH) for the 1-yr period from December 2013 to November 2014 were collected using the United States Environmental Protection Agency (USEPA) equivalent referenced methods and stored in a data acquisition system.

Ozone is measured using an analyser (Serinus 10 Ozone analyser, Ecotech) based on the non-dispersive ultraviolet (UV) absorption technology
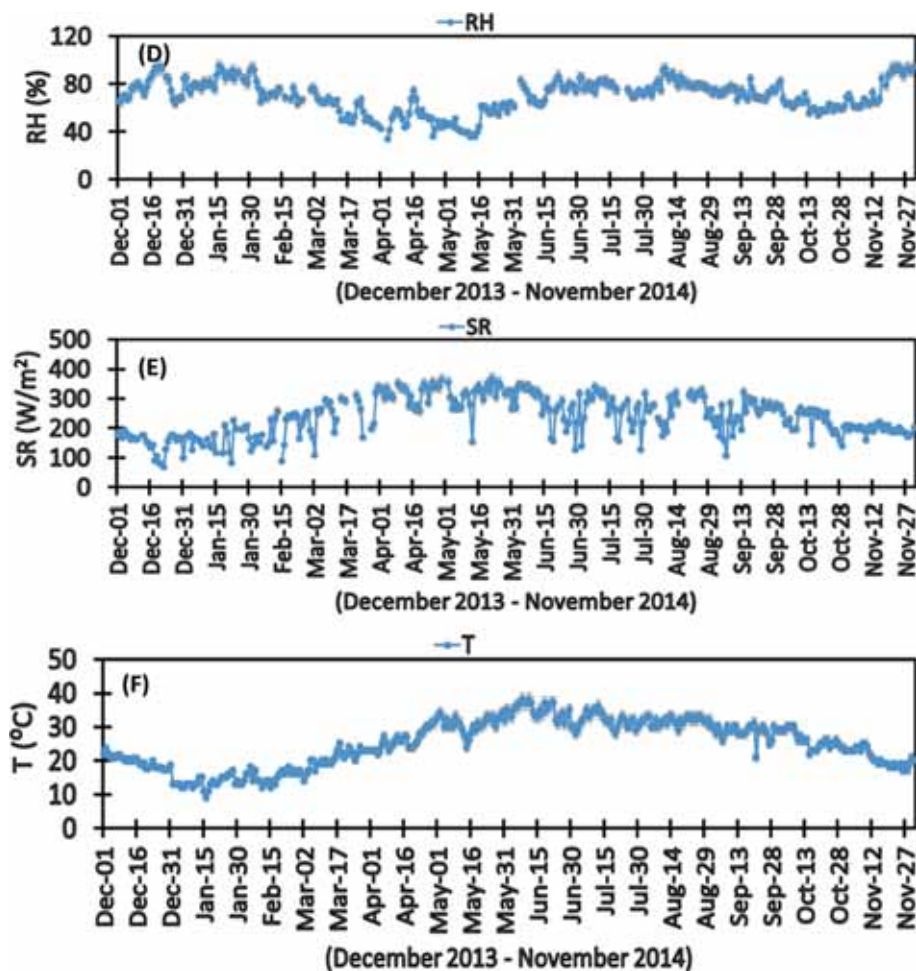
Figure 3. continued

to measure ozone with an accuracy and precision of ± 1 and 0.5 ppb, respectively, in the range of 0–20 ppm. The analyser incorporates corrections due to the changes in the temperature and pressure and can be referenced to $0°C, 20°C$ or $25°C$ at 1 atm. The daily zero/span checks were done using a dynamic gas calibrator (Ecotech Gas Cal 1000) in combination with zero air supply. The ozone levels were monitored at 5-min intervals and then averaged to calculate the hourly levels.

Measurements of the PM (PM_{10} and $PM_{2.5}$) were carried out using a Met One Instruments Model BAM-1020 (Chung *et al.* 2011; Yadav *et al.* 2017). The filter tape (glass fibre filter) loading diagram on BAM-1020 is shown in figure SM-1. It measures and records air-borne particulate concentration levels using the principle of beta ray attenuation. The measurement analysers of PM_{10} and $PM_{2.5}$ conform to the USEPA standards as per FEM (EQPM-0798-122) and class
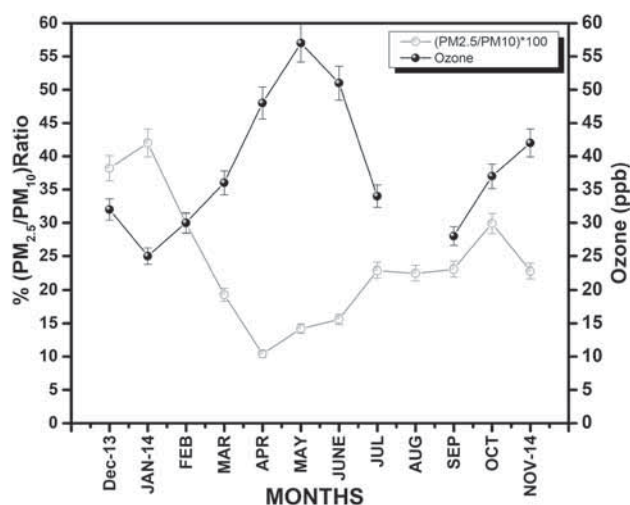

Figure 4. 24 Hour averaged, daily maximum and daily minimum levels of pollutants (Ozone, PM_{10} and PM2.5) and meteorological variables (relative humidity, solar radiation and daily average temperature) from December 2013 to November 2014

Table 1. *Hourly/monthly averaged ozone concentration observed during 00:00–05:00 and 12:00–17:00 h and its rate of change.*

Month	Average ozone concentration (ppb)		Rate of change of O_3/h
	00:00–05:00 h	12:00–17:00 h	
Dec-2013	20.0 ± 11.7	60.0 ± 20.4	3.33
Jan-2014	17.0 ± 7.8	47.2 ± 13.6	2.51
Feb	16.5 ± 9.8	51.1 ± 13.9	2.88
Mar	23.3 ± 7.2	57.2 ± 15.7	2.82
Apr	28.9 ± 8.4	72.2 ± 8.6	3.60
May	36.6 ± 10.5	83.9 ± 14.3	3.94
Jun	35.5 ± 9.9	72.9 ± 15.0	3.11
Jul	21.4 ± 6.9	50.2 ± 16.3	2.40
Aug	ER*	ER*	–
Sep	16.5 ± 7.4	38.6 ± 16.3	1.83
Oct	16.5 ± 7.6	61.0 ± 13.2	3.71
Nov-2014	23.3 ± 9.9	74.1 ± 11.3	4.23

*Error in data.

III FEM (EQPM-0308-170), respectively. The mass concentration was observed at every 5-min intervals and then stored in the data repository at a 1-h average. It performs an automatic check of its calibration (span check) and instrument drift caused by varying temperature, barometric pressure and RH on an hourly basis during the operation.

In addition to this, an automatic weather station model, ME-1310-AWS: Microcomm-ESD, UK was used to measure the meteorological parameters, given their measurement accuracy (WS (± 0.5 m/s), wind direction ($\pm 3°$), AT ($\pm 0.1°$C), SR (± 0.05 W/m^2) and RH ($\pm 1\%$)) at the same site (Ali *et al.* 2013). The maintenance and calibration protocols for the ozone analyser, PM monitor, meteorological sensors and gas calibration system are mentioned in detail in table SM-1. Average daily concentrations were calculated only if at least 18 of each day's hourly concentration values were available. The data points are within the range 89–98%. The ozone data for the month of August are not reported due to a technical error in the analyser. Data analysis used the IBM® SPSS® statistics version 20, 'R' software coupled with OriginPro 8.5 for graphical illustration.

4. Results and discussion

The severity of the impact of air pollution depends upon two variable factors, i.e., the ambient concentration of pollutants and their exposure time. Daily mean, maxima and minima between each of ozone, PM$_{10}$ and PM$_{2.5}$ levels and meteorological parameters (SR, AT, RH) drawn for each day from December 2013 to November 2014 are presented in figure 3(a–f), respectively. The daily averaged mean levels of the uncertainty of ozone, PM$_{10}$ and PM$_{2.5}$ ranged from 8.2 to 72.7 ppb (0.15%), 20.6 to 466.4 μg m^{-3} (0.06%) and 13.5 to 320.1 μg m^{-3} (0.14%), respectively, and the meteorological variables RH, SR, AT and WS varied from 33.6% to 94.8%, 68.0 to 366.1 W m^{-2}, 9°C to 38°C, and 0.1 to 7.9 m s^{-1}, respectively. The annual mean concentration levels of ozone, PM$_{10}$ and PM$_{2.5}$ were recorded as 39 ppb, 178 μg m^{-3} and 88 μg m^{-3}, respectively. Anthropogenic emissions from sources such as vehicles, domestic activities, road-dust, construction activity and industrial processes are released throughout the year, but the episodic event of post-harvest CRB of wheat and rice occur during April–May (summer) and October–November (post-monsoon), respectively. The intermediate period from December to February (winter) experiences fog formation, the manifestation of a shallower planetary boundary layer, emissions from biofuel burning and fossil-fuel combustion sources and moisture from Western disturbances.

Figure 4 shows the monthly variation in ozone levels, and the increase/decrease indicates the significant influence of anthropogenic activities, meteorological variations and related chemical transformations. The rise and fall in the ozone level directly correspond to the SR and ambient temperature exposure and inversely to the RH. The ozone level started increasing from January 2014 and the first maxima appeared in May (56.8

Table 2. *Comparison of ozone levels of ambient air at the Patiala site with those at some other locations in India.*

Location	Latitude (°)	Longitude (°)	Site type	Period of study (mm/yr)	Ozone (ppb)	References
Patiala	30.35°N	76.37°E	Semi-urban	December 2013–November 2014	38.52 ± 10.58	This study
Agra	27.18°N	78.02°E	Urban	January 2002–December 2002	$20.0 \pm *$	Satsangi *et al.* (2004)
Ahmedabad	23.0°N	72.60°E	Urban	January 2002–December 2002	26.27 ± 10	Sahu and Lal (2006)
Pune	18.53°N	73.86°E	Urban	June 2003–May 2004	30.9 ± 14	Beig *et al.* (2007)
Anantapur	14.68°N	77.60°E	Rural	January 2010–December 2010	40.7 ± 3.7	Reddy *et al.* (2013)
Kanpur	26.46°N	80.33°E	Urban	June 2009–May 2013	27.9 ± 17.8	Gaur *et al.* (2014)
Jabalpur	23.17°N	79.95°E	Urban	January 2014–December 2014	$34.0 \pm *$	Sagar *et al.* (2015)
Udaipur	26.58°N	73.68°E	Urban	April 2011–March 2012	$22.420 \pm *$	Yadav *et al.* (2016)

*Data not reported.

Table 3. *Summary of the number of day's exceedance/total number of sampling days (December 2013–November 2014) of the 24-h NAAQS value for the criteria pollutants.*

Month	PM_{10}	$PM_{2.5}$	Ozone
December	28/28	29/29	–
January	27/28	25/26	–
February	23/28	24/28	–
March	10/22	6/24	–
April	22/30	3/30	1/30
May	23/25	9/26	10/30
June	29/29	14/29	8/30
July	16/28	7/30	–
August	17/27	4/27	–
September	14/30	3/29	1/30
October	29/30	27/31	–
November	30/30	29/29	–

ppb), followed by decreases until September with a minimum value of 28.3 ppb in the monsoon season. The maximum monthly average of ozone concentration corresponds to minimum RH (34–70%) and high SR (153–366 W/m^2) and daily ambient temperature (24–34°C) along with the emissions of precursor gases (with moisture content less than 5%) from wheat crop residue burning (WCRB). The second highest monthly average of ozone concentration occurred in November (41.8 ppb), corresponding to moderate RH (54–71%), SR (159–219 W/m^2) and daily ambient temperature (17–25°C) along with the emissions of precursor gases (with moisture content 40–50%) from rice crop residue burning (RCRB). There is a noteworthy difference in the amplitude of the two maxima (May and November) observed during post-harvest CRB of wheat and rice crops, respectively, which indicates favourable meteorological conditions, i.e., high ambient temperature, SR and low RH for the photochemical production of ozone in summer than the post-monsoon season, during the RCRB

period. The daytime average ozone levels and rate of change in ozone were higher in May and November as compared to the rest of the year, as shown in table 1. The daily O_{3max}/O_{3min} ratio in the pollution index, for polluted sites, has values in excess of 10 while those lower than 1.40 for clean areas (Sharma *et al.* 2016). The index value of the order of 15.60 indicates that the study location has remarkable ozone pollution. The influence of short period emissions, due to CRB, along with favourable meteorology, on ozone production can be easily observed. A comparison of the level of ozone in the ambient air of the Patiala site with that of some other locations in India is reported in table 2.

Besides trace gases, atmospheric PM is another species found in the atmosphere as a mixture of solid and liquid phases suspended in the air. Mass concentration variation patterns of both PM_{10} and $PM_{2.5}$ were similar except for their magnitudes throughout the study period which exhibited distinct seasonal variations. Figure 3(b and c)
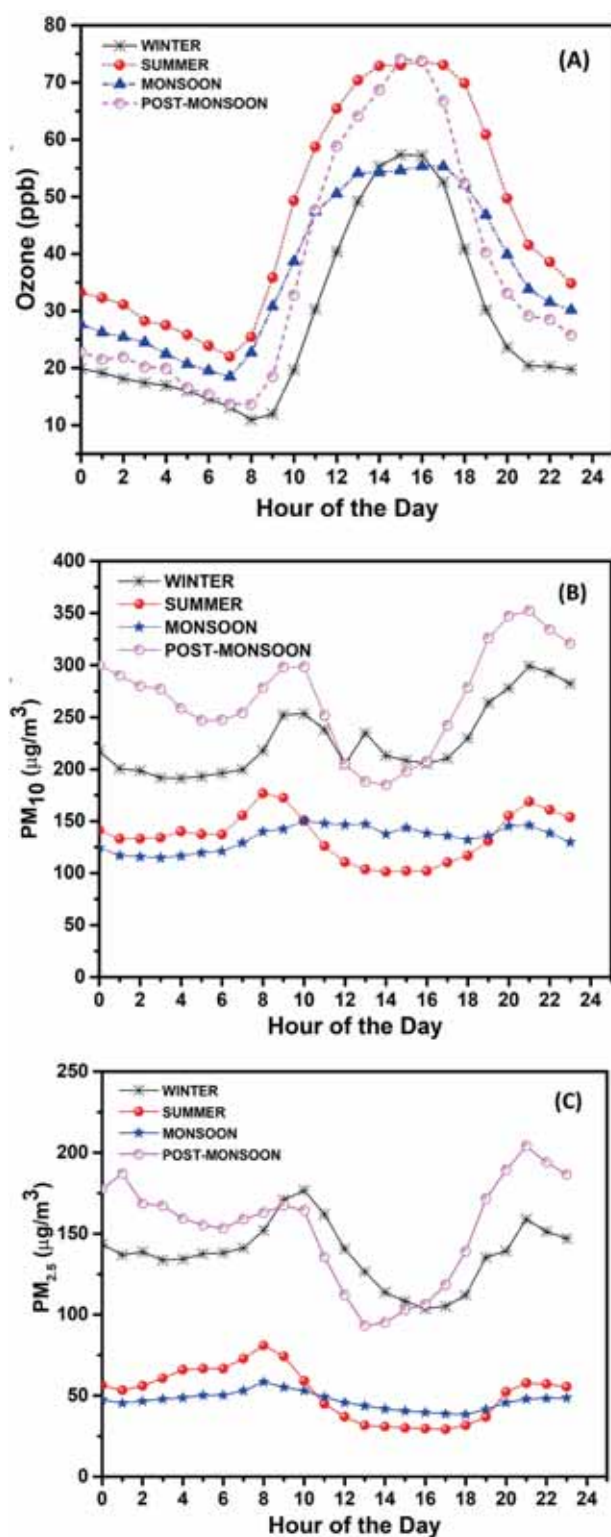
Figure 5. Seasonal (Winter, Summer, Monsoon and Post-monsoon) diurnal variations of O_3, PM_{10} and $PM_{2.5}$ levels during the study period December 2013–November 2014 measured at Patiala.

shows the daily variations in PM_{10} and $PM_{2.5}$, respectively. The levels of both PM_{10} and $PM_{2.5}$ are significantly higher than their daily mean

standard limit. The elevated concentration of PMs results in poor regional and global air quality, along with a significant impact on climatic processes including precipitation and cloud formation, and contributes to the earth radiation budget (Wang *et al.* 2011).

Regulatory bodies for the control of pollution in India, Ministry of Environment, Forest and Climate Change have set national standards for the criteria pollutants ozone, PM_{10} and $PM_{2.5}$ (National Ambient Air Quality Standards (NAAQS), as 60 ppb, 100 µg m^{-3} and 60 µg m^{-3}, respectively, on a 24-h average basis). As per our observations, the percentage of days of exceedances on a 24-h average basis is 77.3% and 53.2% for PM_{10} and $PM_{2.5}$ and 20 days for ozone, respectively. The summary of the data is given in table 3.

For more detailed information about the temporal variation of the emission sources in the area, a monthly variation of the $PM_{2.5}/PM_{10}$ ratios and their impact on ozone concentrations along with error bars with one standard deviation are presented in figure 4. The ratio varies more greatly at the study site and indicates more complex and changing PM sources (Blanco-Becerra *et al.* 2015). The proportion of fine particles in PM_{10} also shows great variability ranges from 10% to 42% with an average value of 25% on a monthly basis. The proportions of $PM_{2.5}$ contained within PM_{10} were almost double in the RCRB period than in the WCRB period, due to incomplete combustion occurring at a low-temperature range (10–25°C) and limited dispersion due to calm wind conditions. The pollutants released during RCRB remained entrapped and showed an after-effect along with a rise in combustion due to bio-fuel and fossil-fuel heating which led to the maximum ratio during winter. The secondary aerosol formation may also contribute to a fine fraction of PM leading to an increase in the $PM_{2.5}/PM_{10}$ ratio in the months of winter and the post-monsoon season (Raja *et al.* 2010). The increase in the $PM_{2.5}/PM_{10}$ ratio corresponds to the decrease in the monthly averaged ozone levels as shown in figure 4. The possible effect of fine PM on the photochemical oxidant cycle eventually acts as a filter for ozone (Jacob 2000; Sharma *et al.* 2016).

4.1 *Diurnal seasonal variability of ozone, PM_{10} and $PM_{2.5}$*

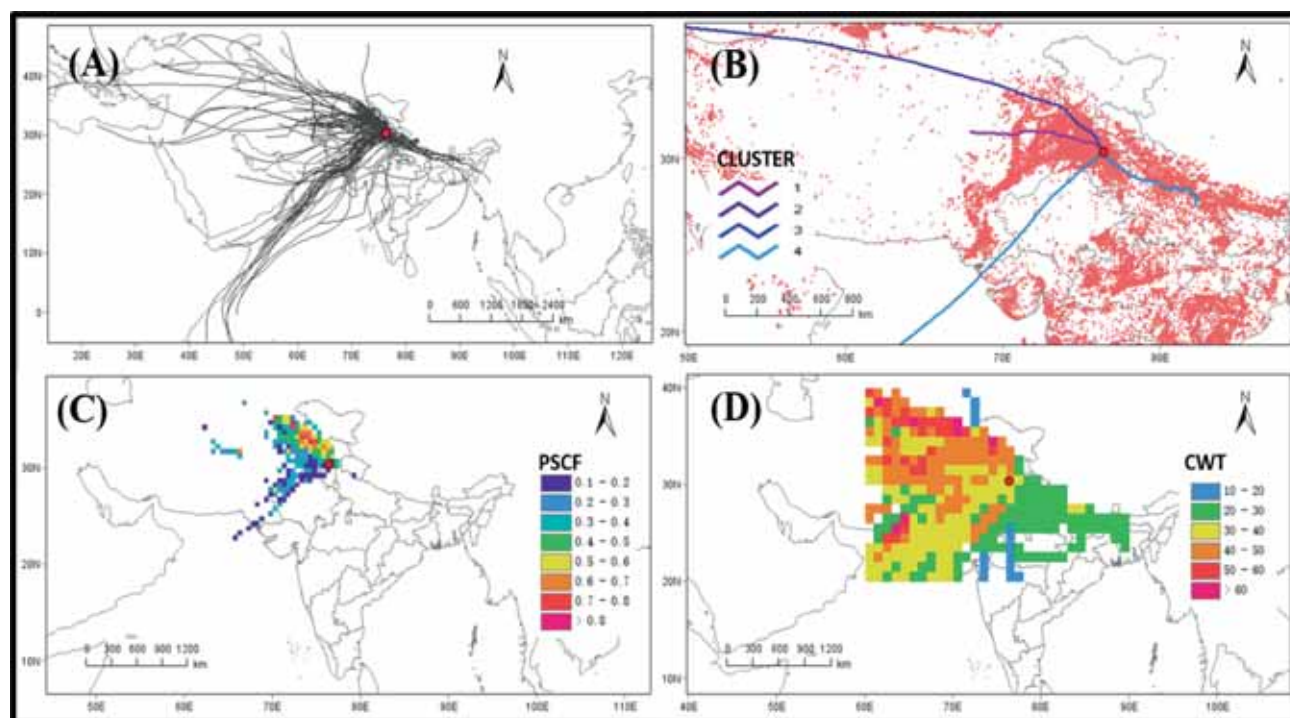PMs and ozone pollution periods are not concurrent on the seasonal timescale, i.e., higher PM

Figure 6. (**a**) Backward trajectories at a 500-m level over Patiala for the study period, (**b**) clustered trajectory analysis, (**c**) PSCF analysis of ozone and (**d**) CWT analysis of surface ozone during December 2013–November 2014.

concentrations during winter and higher ozone values are exhibited during the summer season. But on a diurnal scale, both exhibit time-dependent, day-to-day variations that are important to get the information on local pollution conditions. Diurnal seasonal variations of ozone, PM_{10} and $PM_{2.5}$ for the study period (December 2013–November 2014) are shown in figure 5. The hourly average values are further averaged on the diurnal scale for each month and for a particular season to get the diurnal variations.

For the diurnal ozone profile (figure 5a), a steep minimum is observed during the early morning around (06:00–08:00 h, IST) in all seasons of the year. After 08:00 h, a rise in ozone concentration is observed and hits maxima around the range of 14:00–16:00 h, IST and starts decreasing after late afternoon or evening. However, significant ozone concentration is found throughout the study period with changes in amplitude during different seasons. A prominent and sharper increase in daytime maxima of ozone was observed during the post-monsoon and winter months. This maximum was broader during the summer and monsoon months. The prolongation of the peak (resulting in broader maxima) is a prominent feature of ozone variations during summer followed by the monsoon season due to the longer days supported by favourable

meteorological conditions and vertical mixing from the ozone-rich upper layer until sunset (Reddy *et al.* 2012). The highest rate of increase of dO_3/dt during summer following the post-monsoon season indicated the influence of the increased emission of precursor gases, especially from the post-harvest CRB of wheat and rice, respectively.

Mass concentration variation patterns of both PM_{10} and $PM_{2.5}$ were similar except for their magnitudes throughout the study period and exhibited distinct seasonal variations. The diurnal patterns of PM_{10} and $PM_{2.5}$ show sharp peaks in the morning (07:00–10:00 h, IST) and evening (17:00–21:00 h, IST) as shown in figure 5(b and c). In each season, the lowest concentration was observed during afternoon hours (12:00–16:00 h, IST). High levels of concentration of particulates were observed in winter and post-monsoon seasons followed by the summer and monsoon seasons. The mass concentration of PM_{10} varied between 58 and 381 $\mu g\ m^{-3}$ in winter followed by a significant dip in summer (31–317 $\mu g\ m^{-3}$) and in the monsoon season (20–346 $\mu g\ m^{-3}$). Likewise, the mass concentration of $PM_{2.5}$ was markedly higher in winter (42–270 $\mu g\ m^{-3}$) followed by lower levels in summer (14–98 $\mu g\ m^{-3}$) and in the monsoon season (13–76 $\mu g\ m^{-3}$). Thereafter, the mass concentration rose in the post-monsoon season
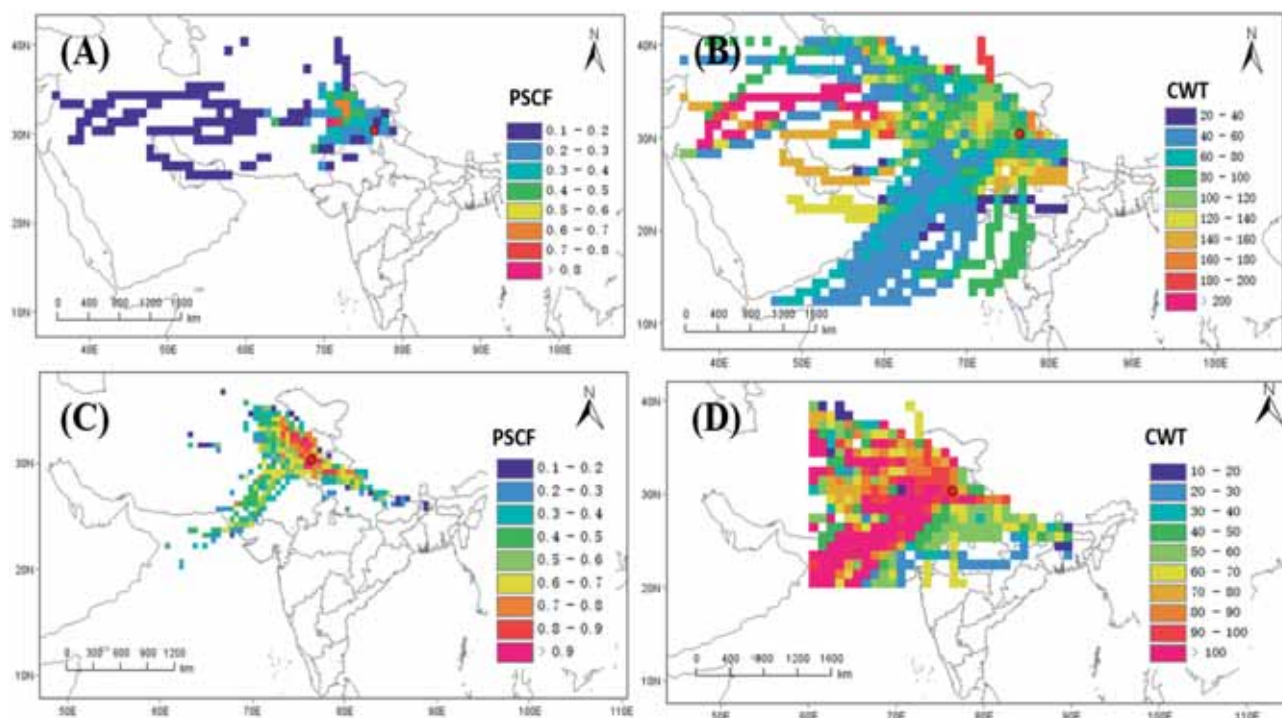
Figure 7. Potential source contribution function (PSCF) and concentration weighted trajectory (CWT) analysis of PM$_{10}$ (**a, b**) and PM$_{2.5}$ (**c, d**) during December 2013–November 2014.

(71–320 µg m^{-3}) for PM$_{10}$ and (39–320 µg m^{-3}) for PM$_{2.5}$. The morning and evening peaks of PM$_{2.5}$ and PM$_{10}$ coincide with the highest local emissions from anthropogenic activities associated with the rush hour. In spite of emissions from vehicular sources at night, the elevated concentrations of PMs are due to emissions from the incomplete RCRB (post-monsoon) and fuel combustion of wood and coal for heating (winter season), followed by stagnant wind flow leading to the accumulation of pollutants. In the monsoon season, the transport of cleaner air (prevailing SW winds) and frequent rainfalls lead to a washout of particulate pollution.

The backward trajectory plot showed that the realistic motion of air parcels and their transport pathways was calculated at 500 m above ground level (AGL) for the sampling duration and have been presented in figure 6(a). These groups of trajectories have been clustered together into distinct groups (clusters) based on the transport speed and direction, eventually yielding clusters with similar length and curvature to represent the major directions of the air mass offering a better conception of the source regions (Li *et al.* 2012). Out of these four clusters, the major percentage contribution was from cluster 1 (56%), followed by cluster

3 (17%), as shown in figure 6(b). Hence, the major influx of gaseous and particulate pollutants was from regional continental land masses to reach the receptor site and was not significantly influenced by long-range transport (figure 7). Along with air clusters, moderate resolution imaging spectroradiometer (MODIS) satellite observations of fire counts in the proximity of the study area during the study period are shown in figure 6(b).

Furthermore, potential source contribution function (PSCF) and concentration weighted trajectory (CWT) analysis was performed to highlight the source potentials of the studied geographical region and also to allocate their discrete contribution to the measured concentrations at the receptor site (Hopke 2003; Wang *et al.* 2009; Cheng *et al.* 2013). It was done using trajectories statistics software, TrajStat (version 1.2.2.6). Figures 6(c and d) and 7(a–d) show the PSCF and CWT analyses for ozone, PM$_{10}$, and PM$_{2.5}$, respectively. The probable source grids span the densely populated and polluted north-west regions, including parts of India as Punjab, Haryana and Rajasthan and parts of Pakistan which contain key oil fields, refineries and major thermal power plants (Naja *et al.* 2014; Sharma *et al.* 2016). Moreover, the enormous amounts of the burning of agricultural crop
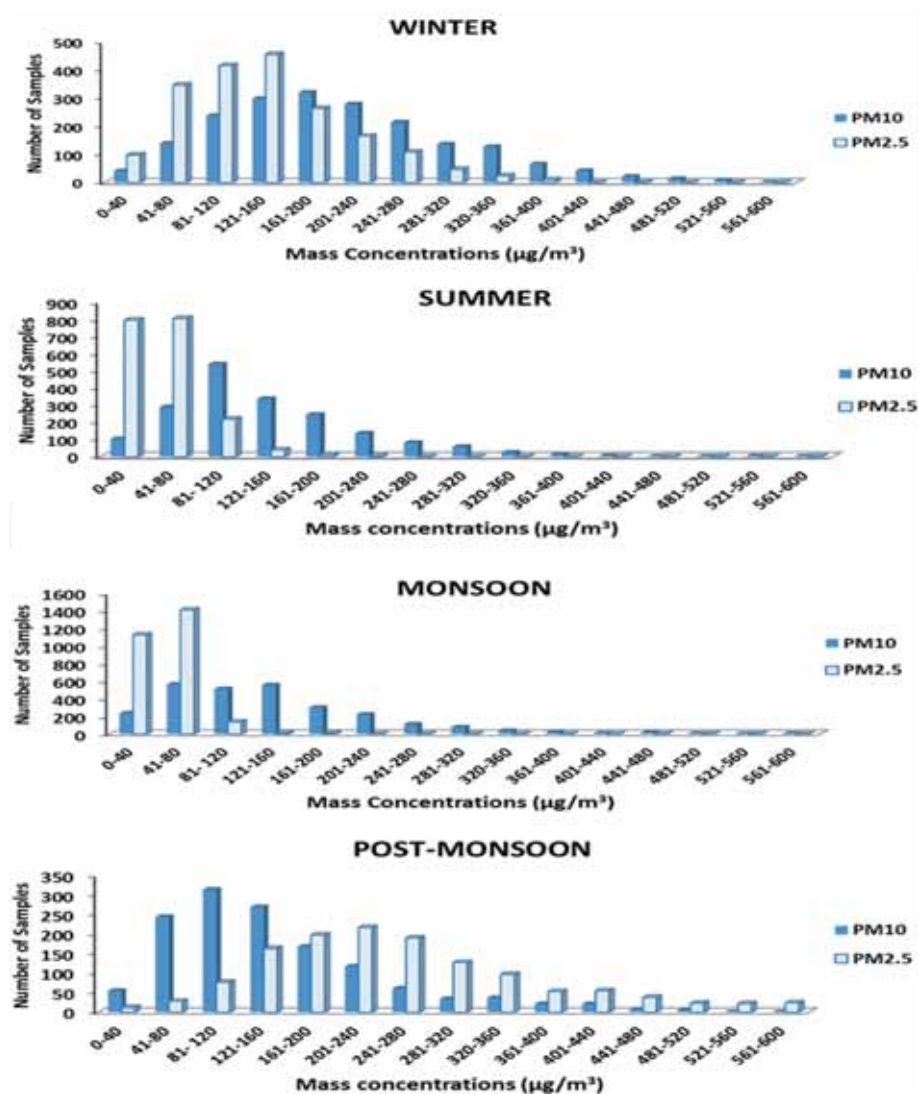
Figure 8. Frequency distributions of PM_{10} and $PM_{2.5}$ for different seasons during the study period.

Table 4. *Bivariate Pearson correlation coefficient matrix for the 24-h averaged daily levels of pollutants and meteorological variables.*

Correlation parameters	Ozone (ppb)	PM_{10} (ppb)	$PM_{2.5}$ (ppb)	Avg. temp (°C)	Solar radiation (W/m^2)	RH (%)	WS (m/s)
Ozone	1	0.093	−0.216	0.475**	0.594**	−0.814**	0.283**
PM_{10}		1	0.866**	−0.073	−0.319**	−0.190**	−0.376**
$PM_{2.5}$			1	−0.452**	−0.594**	−0.207**	−0.489**
Avg. temp				1	0.623**	−0.583**	0.214**
Solar radiation					1	−0.640**	0.392**
Relative humidity						1	−0.305**
WS							1

*Correlation is significant at the 0.05 level (two-tailed).
**Correlation is significant at the 0.01 level (two-tailed).

residues in these fertile farmlands is a critical issue (Jain *et al.* 2014) that has also been adjudged a contributing factor (Badarinath *et al.* 2009) to the high values of gaseous and particulate emissions.

Thus, the air masses flowing from these regions are bound to bring pollutant-laden gases to the receptor site and in due course generate secondary pollutants such as ozone.
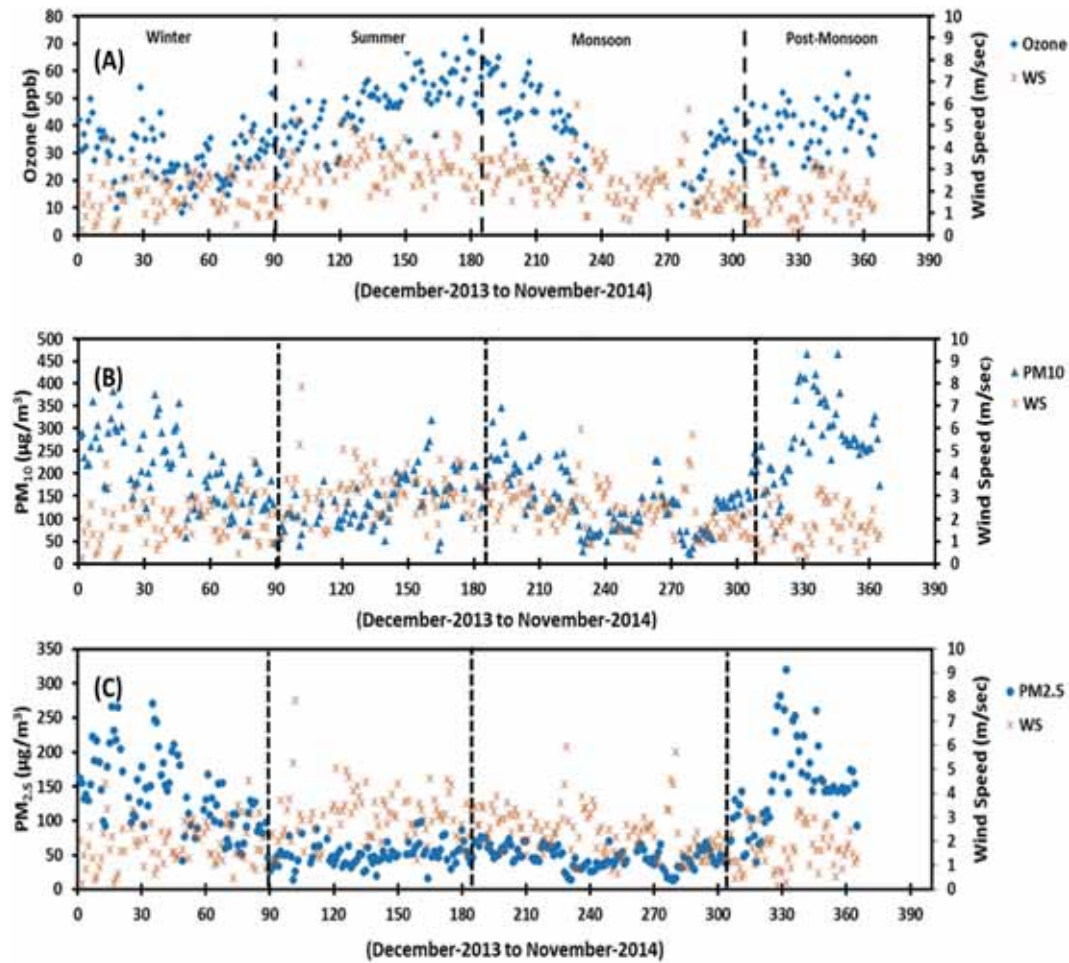
Figure 9. Seasonal variations in daily average levels of ozone, PM_{10} and $PM_{2.5}$ along with wind speed during the study period (December 2013 to November-2014) measured at Patiala.

Table 5. *Performance comparison of the linear and nonlinear methods in the prediction of ozone, PM_{10} and $PM_{2.5}$ with the testing dataset.*

Models	Ozone				PM_{10}				$PM_{2.5}$			
	RMSE	Corr.	R^2	Acc%	RMSE	Corr.	R^2	Acc%	RMSE	Corr.	R^2	Acc%
DT	7.01	0.78	0.61	66	27.97	0.91	0.82	53	14.44	0.94	0.89	67
RF	6.45	0.84	0.71	68	17.94	0.97	0.94	59	10.85	0.97	0.93	66
GLM	5.82	0.87	0.75	74	18.50	0.96	0.92	49	12.16	0.96	0.93	71
NNM	5.82	0.87	0.75	74	18.86	0.96	0.92	46	12.14	0.96	0.93	78

4.2 *Frequency distribution of PM_{10} and $PM_{2.5}$*

The frequency distribution of the mass concentrations of PM_{10} and $PM_{2.5}$ for different seasons is shown in figure 8. The hourly PM mass concentrations are divided into 15 classes with a bin size of 40 μg m^{-3} ranging from 0 to 600 μg m^{-3}. At higher mass ranges, the frequencies decreased gradually: 81–120 and 161–200 μg m^{-3} for PM_{10} and 201–240 and 121–160 μg m^{-3} for $PM_{2.5}$

during the post-monsoon and the winter season, respectively. The major anthropogenic sources of PMs are the incomplete combustion of bio and fossil fuels. Other sources include agricultural residual burning, municipal solid waste treatment, construction activities and emission from paved and unpaved roads. In the post-monsoon and the winter season, the frequencies of both PM_{10} and $PM_{2.5}$ were towards higher mass concentration ranges but during the summer and the monsoon season,
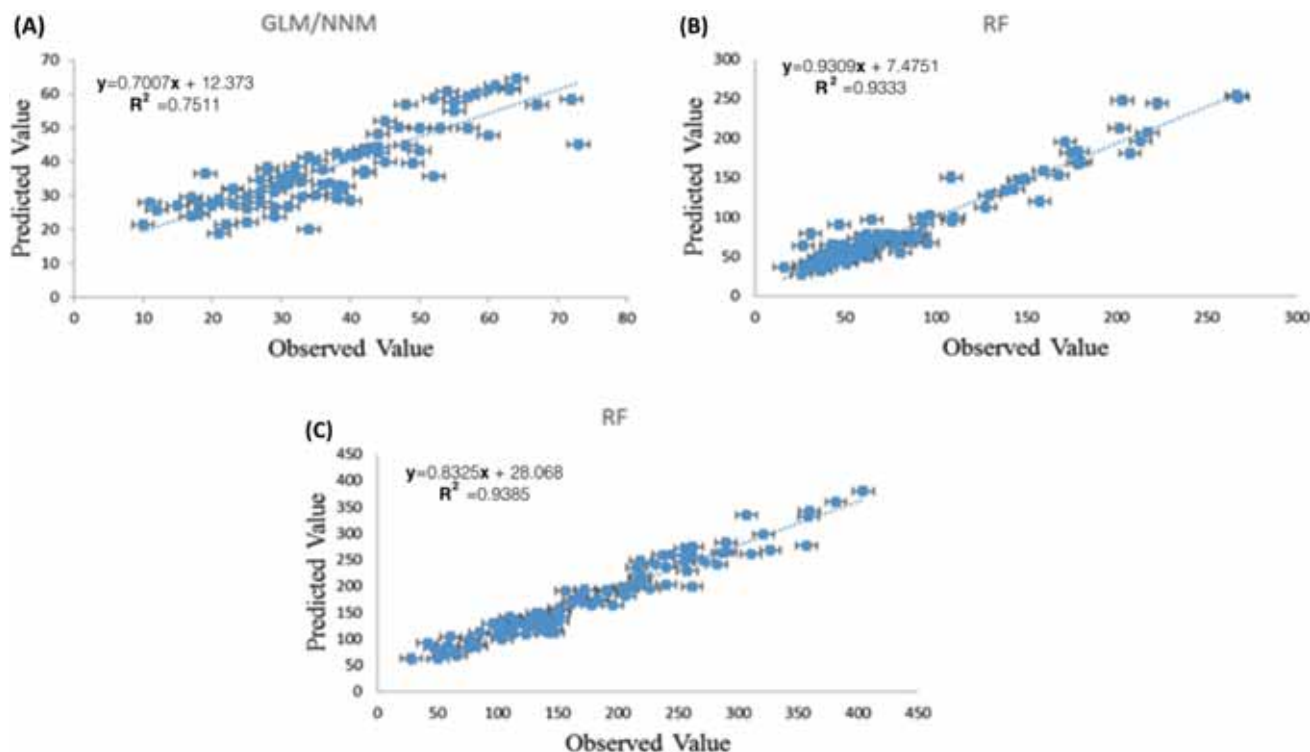
Figure 10. Predicted versus observed values of (**a**) ozone (ppb) by the generalised linear model/neural network model, (**b**) $PM_{2.5}$ (μg m^{-3}) by the random forest model and (**c**) PM_{10} (μg m^{-3}) by the random forest model.

peak frequencies shifted towards lower mass concentrations. This feature is caused by significant fine fraction contribution from the post-harvest RCRB period post-monsoon, followed by a limited dispersion of pollutants due to the calm weather conditions in the winter season. In the post-monsoon season, the mass fraction for $PM_{2.5}$ was twice than that of the PM_{10} for higher ranges, i.e., from 400 to 600 μg m^{-3}. In the winter season, the prominent PM fraction is PM_{10} over $PM_{2.5}$, due to biofuel and fossil fuel burning for heating at a lower temperature and calm wind conditions leading to the trapment of pollutants. The sharp decrease in mass range was observed in the summer and monsoon seasons, and peaks were observed in the lower mass fraction ranges up to 80 μg m^{-3}.

4.3 *Correlation analysis of ozone, PM_{10}, and $PM_{2.5}$ with meteorological parameters*

In order to analyse the relationship between ozone, PM_{10} and $PM_{2.5}$ with meteorological parameters, the relationship between individual pollutants and meteorological variables was determined by the Pearson correlation coefficient (table 4). Daily variations indicated that the criteria pollutants were significantly influenced by meteorological conditions.

Trends in ozone levels show an exact inverse image of RH (figure 3d), which can be further justified considering the strong inverse correlation between the two parameters. The positive correlation ($r = 0.475$) between ozone and the ambient temperature is due to the fact that the sun radiations control the temperature, and hence, the photolysis efficiency will be high, leading to the generation of ozone from its precursors such as hydrocarbons, NO$_x$, etc. A negative correlation ($r = -0.814$) between the surface ozone and humidity lies in the fact that with an increase in humidity levels, the precursor gases are washed away from the environment, resulting in a decrease in the production of tropospheric ozone concentration, the major photochemical path for the removal of ozone.

Higher levels of humidity slow down the photochemical process due to its association with greater cloud abundance, atmospheric instability and low incoming SR (Saini *et al.* 2014). Also, the surface ozone is depleted through the deposition of its molecules on water droplets (Londhe *et al.* 2008). The positive correlation ($r = 0.283$) between ozone and WS indicates ozone transport. The increase of

WS implies the increasing transport of air, thus the influence of WS on primary and secondary pollutants is shown in figure 9. In the case of primary pollutants, it acts as a diluting agent, while ozone concentration increases due to its transport from other regions (Dragan *et al.* 2008). The significant negative correlation between PM_{10}, $PM_{2.5}$ and WS levels indicates the dilution of these air pollutants. The extent of correlation is high for $PM_{2.5}$ ($r = -0.489$) as compared to PM_{10} ($r = -0.376$); hence, the impact of WS is more on $PM_{2.5}$ rather than the large-sized PM_{10}. The decrease in SR levels and AT is associated with stagnant atmospheric conditions and an increase in PM levels (Barmpadimos *et al.* 2012; Zhao *et al.* 2014). PM_{10} and $PM_{2.5}$ also exhibit a high degree of significant positive correlation ($r = 0.866$).

Since some of the considered meteorological variables (features) have higher importance than others in the prediction, ranking (feature importance) was done according to the weightage with the use of 'R' software. It is found that RH has the highest ranking for ozone and PM_{10} while WS has the lowest rank in all three cases.

4.4 *Linear and nonlinear prediction models: performance and validation*

Here, the generalised linear model (GLM) and nonlinear models (decision tree (DT), random forest (RF) and neural network (NN) models) were applied to the criteria pollutants (ozone, PM_{10} and $PM_{2.5}$) by using the input of four meteorological variables (RH, SR, AT and WS), and their hourly predictions, performance on the basis of RMSE, correlation coefficient, R^2 and accuracy were evaluated using 'R' open source software. A detailed description of the used models is available (Chambers 1977; Quinlan 1986; Riedmmiller and Braun 1993; Liaw and Wiener 2002), and although there are various ways to measure the performance of the prediction model, some are more suitable than others, depending upon the application considered.

The distribution of data in the training–testing experiment was optimised to 70% and 30%, respectively, for all methods. Table 5 shows the comparative performance of all methods in the prediction of ozone, PM_{10} and $PM_{2.5}$ on the basis of RMSE, correlation, R^2 and accuracy using 1000 iterations. The performance result shows that the GLM and NNM methods are considered more efficient than other methods for ozone and the RF method is efficient for PM_{10} and $PM_{2.5}$. Figure 10 shows the

scatter plots of the predicted versus the observed values of ozone, PM_{10} and $PM_{2.5}$ for the outperformed model. The 10-fold cross-validation is used to measure the robustness of the GLM and NNM models for the ozone and RF model for PM. Cross-validation results show the uniform performance on all model evaluation parameters shown in figure SM-2. The linear (GLM) and the nonlinear model (NNM) resulted in very good performance for ozone as compared to the PM levels by the nonlinear model (RF). The ozone levels are reasonably explained even in the absence of precursor emission levels.

5. Conclusions

Diurnal and seasonal variations of ozone and PMs indicate a critical rise in the level of pollutants in the area due to CRB events, along with the decisive role of meteorological factors in the formation and buildup of primary and secondary pollutants. Statistical analyses of data on the interdependence of meteorological parameters indicate that the rate of ozone production was the highest during the post-harvest fire period (3.94 ppb of O_3/h in May and 4.23 ppb of O_3/h in November) leading to a remarkable ozone pollution index (15.60) in the area. The proportions of $PM_{2.5}$ contained within PM_{10} were almost double in the RCRB period than in the WCRB period, due to incomplete combustion occurring at a low temperature range (10–25°C) and limited dispersion due to calm wind conditions. The higher ratios of $PM_{2.5}/PM_{10}$ can be attributed to the prevalent particle pollution from anthropogenic sources such as agricultural and biofuel burning, vehicular exhaust, industrial process and construction activities, and smaller ratios can be attributed to particle pollution from natural sources, transport of aged pollutants and windblown dust. Days of exceedances of PM_{10}, $PM_{2.5}$ and O_3 levels from those of the standard levels of NAAQS exceeded by 77.3%, 53.2% and 20 days during the study period.

For O_3 and PM_{10}, RH has the highest influence, while WS has the lowest rank in all three cases. Numerical simulation with a linear model (GLM) and nonlinear models (DT, RF and NNM) using meteorology as the input shows that the GLM and NNM methods outperform other methods with regard to hourly ozone predictions while RF outperforms other methods with regard to the hourly predictions of PM_{10} and $PM_{2.5}$. A 10-fold

cross-validation shows the uniform performance on all model evaluation parameters. To assess the seriousness of the increase in pollution levels due to episodic events and to take further preventive measures, short-term forecasting of air quality is needed.

Acknowledgements

References

Agarwal R, Awasthi A, Singh N, Mittal S K and Gupta P K 2013 Epidemiological study on healthy subjects affected by agriculture crop-residue burning episodes and its relation with their pulmonary function tests; *Int. J. Environ. Health Res.* **23** 281–295.

Ali K, Chate D, Beig G, Srinivas R, Parkhi N and Satpute T 2013 Spatial temporal variation and deposition of fine and coarse particles during the common wealth games in Delhi; *Aerosol. Air Qual. Res.* **13**(**2**) 748–755.

Ashworth K, Wild O, Eller A S D and Hewitt C N 2015 Impact of biofuel poplar cultivation on ground-level ozone and premature human mortality depends on cultivar selection and planting location; *Environ. Sci. Technol.* **49**(**14**) 8566–8575.

Awasthi A, Agarwal R, Mittal S K, Singh N, Singh K and Gupta P K 2011 Study of size and mass distribution of particulate matter due to crop residue burning with seasonal variation in the rural area of Punjab, India; *J. Environ. Monit.* **13** 1073–1081.

Badarinath K V S, Chand T R K and Prasad V K 2006 Agriculture crop residue burning in the Indo-Gangetic Plains – A study using IRS-P6 AWiFS Satellite Data; *Curr. Sci.* **91**(**8**) 1085–1089.

Badarinath K V S, Kharol S K, Sharma A R and Prasad V K 2009 Analysis of aerosol and carbon monoxide characteristics over Arabian Sea during crop residue burning period in the Indo-of plains using multi-satellite remote sensing datasets; *J. Atmos. Sol.-Terr. Phys.* **71** 1267–1276.

Banerjee T, Singh S B and Srivastva R K 2011 Development and performance evaluation of statistical models correlating air pollutants and meteorological variables at Pantnagar, India; *Atmos. Res.* **99**(**3–4**) 505–517.

Barmpadimos J K, Oderbolz D, Hueglin C and Prevot A S H 2012 One decade of parallel fine (PM2.5) and coarse (PM10–PM2.5) particulate matter measurements in Europe: Trends and variability; *Atmos. Chem. Phys.* **12**(**7**) 3189–3203.

Beig G, Gunthe S and Jadhav D B 2007 Simultaneous measurements of ozone and its precursors on a diurnal scale at a semi urban site in India; *J. Atmos. Chem.* **57**(**3**) 239–253.

Blanco-Becerra L C, Gafaro-Rojas A I and Rojas-Roa N Y 2015 Influence of precipitation scavenging on the PM2.5/PM10 ratio at the Kennedy locality of Bogota, Colombia; *Rev. Fac. Ing. Univ. Antioquia* **76** 58–65.

Chambers J M 1977 Computational methods for data analysis; *Appl. Stat.* **1**(**2**) 1–10.

Cheng I, Zhang L, Blanchard P, Dalziel J and Tordon R 2013 Concentration-weighted trajectory approach to identifying potential sources of speciated atmospheric mercury at an urban coastal site in Nova Scotia, Canada; *Atmos. Chem. Phys.* **13**(**12**) 6031–6048.

Chung A, Chang D P Y, Kleeman M J, Perry K D, Cahill T A and Dutcher D 2011 Comparison of real-time instrument used to monitor air-borne particulate matter; *J. Air Waste Manag.* **51**(**1**) 109–120.

Dragan M M, Dragan A M, Anka J, Lazar L and Zoran M 2008 Determination of O3, NO2, SO2, CO, and PM10 measured in Belgrade urban area; *Environ. Monit. Assess.* **145**(**1–3**) 349–359.

Dubey J, Singla V, Maharaj K and Lakhani A 2014 Influence of meteorology on surface ozone, NO$_2$ and fine particulate matter at a tropical site in India; *Asian Reson.* **4**(**1**) 82–88.

Galanter M, Levy H and Carmichael G R 2000 Impacts of biomass burning on tropospheric CO, NO$_X$ and O$_3$; *J. Geophys. Res.* **105** 6633–6653.

Gaur A, Tripathi S N, Kanawade V P, Tare V and Shukla S P 2014 Four-year measurements of trace gases (SO2, NOx, CO, and O3) at an urban location, Kanpur, in Northern India; *J. Atmos. Chem.* **71**(**4**) 283–301.

Ghazali N A, Ramli N A, Yahaya A S, Yusof N F F, Sansuddin N and Madhoun W A 2009 Transformation of nitrogen dioxide in to ozone and prediction of ozone concentration using multiple linear regression techniques; *Environ. Monit. Assess.* **165**(**1–4**) 475–489.

Gunthe S S, Beig G and Sahu L K 2016 Study of relationship between daily maxima in ozone and temperature in an urban site of India; *Curr. Sci.* **110**(**10**) 1994–1999.

Gupta P K, Sahai S, Singh N, Dixit C K, Singh D P and Sharma C 2004 Residue burning in rice-wheat system cropping system: Causes and implications; *Curr. Sci.* **87**(**12**) 1713–1717.

Han S, Bian H, Feng Y, Liu A, Li X and Zeng F 2011 Analysis of the relationship between O$_3$, NO and NO$_2$ in Tianjin, China; *Aerosol. Air Qual. Res.* **11** 128–139.

Hopke P K 2003 Recent developments in receptor modeling; *J. Chemom.* **17** 255–265.

Icaga Y and Sabah E 2009 Statistical analysis of air pollutants and meteorological parameters in Afyon, Turkey; *Environ. Model. Assess.* **14**(**2**) 259–266.

Jacob D J 2000 Heterogeneous chemistry and tropospheric ozone; *Atmos. Environ.* **34** 2131–2159.

Jain N, Bhatia A and Pathak H 2014 Emissions of air pollutants from crop residue burning in India; *Aerosol. Air Qual. Res.* **14** 422–430.

Kaskaoutis D G, Kumar S, Sharma D, Singh R P, Kharol S K, Sharma M, Singh A K, Singh S, Singh A and Singh D 2014 Effect of crop residue burning on aerosol properties, plume characteristics, and long-range transport over northern India; *J. Geophys. Res. Atmos.* **119** 5424–5444.

Kumar V, Sarkar C and Sinha V 2016 Influence of post-harvest crop residue fires on surface ozone mixing ratios in the N.W. IGP analyzed using 2 years of continuous *in situ* trace gas measurements; *J. Geophys. Res. Atmos.* **121**(**7**) 3619–3633.

Li M, Huang X, Zhu L, Li J, Song Y, Cai X and Xie S 2012 Analysis of the transport pathways and potential sources of PM10 in Shanghai based on three methods; *Sci. Total Environ.* **414** 525–534.

Liaw A and Wiener M 2002 Classification and regression by random forest; *R. News* **2**(**3**) 18–22.

Londhe A L, Jadhav D B, Buchunde P S and Kartha M J 2008 Surface ozone variability in the urban and nearby rural locations of tropical India; *Curr. Sci.* **95**(**12**) 1724–1729.

Mittal S K, Singh N, Agarwal R, Awasthi A and Gupta P K 2009 Ambient air quality during wheat and rice crop stubble burning episodes in Patiala; *Atmos. Environ.* **43**(**2**) 238–244.

Muezzinoglu A 1987 *Growth of environmental technology as foreseen by the recent administrative and legislative developments*; Martinus Nijhoff Publishers, Dordrecht.

Naja M, Mallik C, Sarangi T, Sheel V and Lal S 2014 SO2 measurements at a high altitude site in the central Himalayas: Role of regional transport; *Atmos. Environ.* **99** 392–402.

Nishant T and Kumar S M K 2011 Diurnal variation of surface ozone with meteorological parameters at Kannur, India; *Adv. Appl. Sci. Res.* **2**(**3**) 407–417.

Punia M, Nautiyal V P and Kant K 2008 Identifying biomass burned patches of agricultural residue using satellite remote sensing data; *Curr. Sci.* **94**(**9**) 1185–1190.

Quinlan J R 1986 Introduction of decision trees; *Mach. Learn.* **1**(**1**) 81–106.

Raja S, Biswas K F, Husain L and Hopke P K 2010 Source apportionment of the atmospheric aerosol in Lahore, Pakistan; *Water Air Soil Pollut.* **208**(**1–4**) 43–57.

Reddy K K, Naja M, Ojha N, Mahesh P and Lal S 2012 Influences of the boundary layer evolution on surface ozone variations at a tropical rural site in India; *J. Earth Syst. Sci.* **121** 911–922.

Reddy B S, Kumar K R, Balakrishnaiah G, Gopal K R, Reddy R R and Sivakumar V 2013 Analysis of diurnal and seasonal behavior of surface ozone and its precursors (NO$_X$) at a semi-arid rural site in southern India; *Aerosol Air Qual. Res.* **12**(**1**) 1081–1094.

Riedmmiller M and Braun H 1993 A direct adaptive method for faster back propagation learning: The RPROP algorithm; In: *Proceedings of the IEEE International Conference on Neural Networks, San Francisco, 28 March–1 April* 1993, pp. 586–591, https://doi.org/10.1109/ICNN.1993.298623.

Sagar K, Srivastava R K and Sarkar S 2015 Diurnal and seasonal variabilityof PM10 with surface ozone concentration at Jabalpur; *Brit. J. Environ. Sci.* **3**(**3**) 62–70.

Sahu L K and Lal S 2006 Distribution of C2–C$_5$ NMHCs and related trace gases at a tropical urban site in India; *Atmos. Environ.* **40** 880–891.

Saini R, Singh P, Awasthi B B, Kumar K and Taneja A 2014 Ozone distributions and urban air quality during summer in Agra – A world heritage site; *Atmos. Pollut. Res.* **5**(**3**) 796–804.

Satsangi G, Lakhani A, Kulshrestha P R and Taneja A 2004 Seasonal and diurnal variations of surface ozone and preliminary analysis of exceedance of its critical levels at a semi-arid site in India; *J. Atmos. Chem.* **47** 271–286.

Sharma A, Mandal T K, Sharma S K, Shukla D K and Singh S 2016 Relationship of surface ozone with its precursors, particulate matter and meteorology over Delhi; *J. Atmos. Chem.* **74**(**4**) 451–474.

Sinha B, Singh K, Sangwan Y, Maurya V K, Sarkar C and Chandra B P 2015 Assessment of crop yield losses in Punjab and Haryana using 2 year of continuous *in situ* ozone measurement; *Atmos. Chem. Phys.* **15**(**16**) 9555–9576.

Tarasova O A and Karpetchko A Y 2003 Accounting for local meteorological effects in the ozone time-series of Lovozero (Kola Peninsula); *Atmos. Chem. Phys.* **3**(**4**) 941–949.

Vardoulakis S and Kassomenos P 2008 Sources and factors affecting PM10 levels in two European cities: Implications for local air quality management; *Atmos. Environ.* **42**(**17**) 3949–3963.

Wang M, Ghan S, Ovchinnikov M, Liu X, Easter R, Kassianov E, Qian Y and Morrison H 2011 Aerosol indirect effects in multi-scale aerosol-climate model PNNL-MMF; *Atmos. Chem. Phys.* **11** 5431–5455.

Wang Y Q, Zhang X Y and Draxler R R 2009 Trajstat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data; *Environ. Modell. Softw.* **24**(**8**) 938–939.

World health statistics 2018: monitoring health for the SDGs, sustainable development goals. Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO.

Yadav R, Sahu L K, Beig G and Jaffrey S N A 2016 Role of long-range transport and local meteorology in seasonal variation of surface ozone and its precursors at an urban site; *Atmos. Res.* **176–177** 96–107.

Yadav R, Sahu L K, Beig G, Tripathi N and Jaffrey S N A 2017 Ambient particulate matter and carbon monoxide at an urban site of India: Influence of anthropogenic emissions and dust storms; *Environ. Pollut.* **225** 291–303.

Yoo H J, Kim J, Yi S M and Zoh D K 2011 Analysis of black carbon, particulate matter, and gaseous pollutants in an industrial area in Korea; *Atmos. Environ.* **45**(**40**) 7698–7704.

Zhao C X, Wang Y Q, Wang Y J, Zhang H L and Zhao B Q 2014 Temporal and spatial distribution of PM2.5 and PM10 pollution status and the correlation of particulate matters and meteorological factors during winter and spring in Beijing; *Environ. Sci.* **35**(**2**) 418–427.

Corresponding editor: Suresh Babu