Supplementary Information

Characterization, sources, and atmospheric transformation of few key short-lived climate pollutants (SLCPs) at a rural super-site in the Indo-Gangetic Plains (IGP) of India

Jai Prakash[a,b,#], Harsh Raj Mishra[a], Kalyan Mitra[c], Bhilok Chandra[a], Mattias Hallquist[a,c], Gazala Habib[b], Geetam Tiwari[b], Jan B. C. Pettersson[c], Johan Boman[a,c], Håkan Pleijel[a,d], Ravi Kant Pathak[a,c*]

[a]*Indo Gangetic Plains–Centre for Air Research and Education, RuriPara, Hamirpur, U.P., India*

[b]*Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi-110016 India*

[c]*Atmospheric Science, Department of Chemistry and Molecular Biology, University of Gothenburg, SE-41296 Gothenburg, Sweden*

[d]*Department of Biological and Environmental Science, University of Gothenburg, SE-41296 Gothenburg, Sweden*

Revised manuscript submitted to

Environmental Science: Atmospheres

15 March 2022

[#] Current affiliation: Aerosol and Air Quality Research Laboratory, Washington University in St. Louis, St. Louis, MO, USA

*Corresponding Author:

Email: ravikant@chem.gu.se Phone: + 46-766229095

List of Supplementary discussion

Text S1 Identification of local sources: CBPF analysis

To identify local sources of O_3, BC, and other SLCPs pollutants, a conditional bivariate probability function (CBPF) analysis was performed. The CBPF was generated based on measured SLCPs levels by coupling a conditional probability function (CPF) with wind speed as a third variable, distributing measured pollutant concentrations over a range of wind-sector bins. The CBPF is defined as follows:

$$CBPF_{\Delta\theta,\Delta u} = \frac{m_{\Delta\theta,\Delta u} \downarrow c \geq X}{n_{\Delta\theta,\Delta u}} \tag{S1}$$

where $m_{\Delta\theta,\Delta u} \downarrow c \geq X$ is the number of samples wind blowing to wind sector '$\Delta\theta$' with wind speed interval 'Δu' having concentration ‘c’ greater than a threshold value, and ‘X’; $n_{\Delta\theta,\Delta u}$ is the total number of samples observed in that wind direction for a given speed interval. The threshold criterion was chosen as the 50^{th} or 75^{th} percentile of the measured pollutant concentrations. CBPF analysis was performed using the openair R package (Carslaw and Ropkins, 2012) in R-studio (version 3.1.1; R Core Team, 2014). The method is described in more detail by Uria-Tellaetxe and Carslaw (2014).

Text S2 Back trajectories analysis: Identification of regional sources

Ninety-six hour backward air mass trajectories were retrieved for each day of the sampling year 2017 at intervals of 3 h using the Atmospheric Administration (NOAA) Air Resources Laboratory (ARL) Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) (Rolph et al., 2017) with global data integration system meteorological data as input with a spatial resolution of 1°×1° at an arrival height of 500 m above ground level. The openair R-package was used to download back-trajectories for the IGP-CARE site. The functions and code used to download monthly meteorological file from HYSPLIT PC model, merge 3-hourly back-trajectory end-point files for the IGP-CARE site were as described in the openair R-package manual. For distant sources, the potential source contribution function (PSCF) identifies the potential source region located within a gridded cell at latitude (i) and longitude (j) with probabilities of exceeding the threshold criterion calculated as:

$$PSCF_{ij} = \frac{m_{ij}}{n_{ij}} \tag{S2}$$

Here, n_{ij} is the total number of endpoints in cell i,j in the domain grid, and m_{ij} is the number of endpoints when the mass concentration exceeds the threshold value. However, PSCF values can be affected by grid cells and their endpoints which may be overestimated. Therefore, several authors suggest the use of a weighing function (W_{ij}) to avoid this issue.[1–3] The (W_{ij}) used here was:

$$W_{ij} = \begin{Bmatrix} P_c = 1.00 & when & n_{ij} > 2\,\bar{N} \\ P_c = 0.75 & when & \bar{N} > n_{ij} \le 2\,\bar{N} \\ P_c = 0.50 & when & 0.5\,\bar{N} > n_{ij} \le \bar{N} \\ P_c = 0.15 & when & n_{ij} \le 0.5\,\bar{N} \end{Bmatrix} \tag{S3}$$

where P_c is the probability corrected for the function W_{ij} and $\bar{N}$ is the average number of endpoints over the grid cells with one endpoint. As mentioned above, monthly meteorological files from the HYSPLIT PC model, the merging of 3-hourly back-trajectory end-point files for the receptor site, and all functions, and codes for estimating PSCF were as discussed in the openair R-package manual. Additionally, cluster analysis was also performed using the openair R-package [4]. Cluster analysis and potential source contribution function (PSCF) generation were also performed using the openair R-package.[5]

Supplementary Table

Table S1 Instruments deployed at the IGP-CARE for monitoring SLCPs and meteorological parameters.

Variable	Instruments	Time resolution	Detection limits
UVPM and BC	Met One (BC 1050)	1 min.	8 ng m^{-3}
O_3	Ecotech (O_3 (Serinus® 10 model)	1 sec.	0.5 ppbv
CO	Ecotech (CO (Serinus® 30 model)	1 sec.	40 ppbv
NO and NO_2	Ecotech (NO_X (Serinus® 40 model)	1 sec.	0.4 ppbv
Temperature		30 min.	0.1 °C
Relative humidity	AWS Davis Vantage Pro 2	30 min.	0.5 %
Wind speed		30 min.	0.1 m s^{-1}
Wind direction		30 min.	-
Solar radiation		30 min.	10 W m^{-2}

Table S2. Summary statistics of SLCPs and meteorological parameters measured during the day and night in 2017 over IGP-CARE.

Variables	Day					Night				
	N	Min	Max	Mean	Median	N	Min	Max	Mean	Median
Short-lived climate pollutants (SLCPs)										
BC concentration (μgm^{-3})	4046	0.2	13.8	2.6	2.0	3894	0.2	16.6	3.6	3.3
BrC concentration (μgm^{-3})	4046	0.0	12.9	1.1	1.2	3894	0.0	32.3	2.1	1.7
O3 (ppbv)	3889	0.2	207.3	40.3	41.1	3719	0.2	147.5	22.3	19.8
CO (ppbv)	2123	1.1	3656	430.2	347.7	2106	1.4	2740.6	489.7	420.0
NO (ppbv)	2792	0.0	9.5	0.2	0.13	2848	0.0	4.5	0.2	0.09
NO2 (ppbv)	2792	1.4	22.1	6.7	5.8	2848	1.2	21.2	6.4	5.8
NOX (ppbv)	2792	1.3	20.7	6.4	5.6	2848	1.1	20.8	6.2	5.7
Meteorological parameters										
Temperature (°C)	3788	7.8	46.3	29.8	30.1	3661	5.8	42.7	24.2	26.4
Relative Humidity (%)	3788	7.5	100	55.9	55.5	3661	13.5	99.0	68.3	71.6
Wind speed (ms^{-1})	3788	0.0	12.5	2.5	2.2	3661	0.0	11.4	1.2	0.5
Solar Radiation ($W\ m^{-2}$)	4046	67.5	651.7	140.7	105.3	3894	0.0	62.6	13.2	2.2
Planetary boundary layer height (m)	4745	10.8	4878.9	1053.2	808.0	4010	10.5	1441.5	133.0	80.4

Table S3 Criteria used to define episodic events for each SLCP at IGP-CARE

SLCPs variable	criteria for Episodic events (~95th Percentile of pollutants)
BC	7.4 μg m^{-3}
BrC	4.7 μg m^{-3}
O_3	70 ppbv
CO	1188 pbbv
NO	0.64 ppbv
NO_2	12.0 ppbv
NO_X	12.3 ppbv

Table S4 Correlation between BC and BrC in different episodic events for 24 h, day and night.

Events	Dates	BC vs BrC		
		24 hr (r^2)	Day-time (r^2)	Night-time (r^2)
E1	07 Jan	0.93	096	0.16
E2	12 Jan	0.96	0.99	0.92
	13 Jan	0.51	0.77	0.61
	14 Jan	0.85	0.98	0.66
	15 Jan	0.91	0.90	0.96
	16 Jan	0.63	0.53	0.45
	17 Jan	0.32	0.92	0.04
	18 Jan	0.78	-	0.78
	19 Jan	0.91	0.92	0.84
E3	21 Jan	0.95	0.84	0.98
	22 Jan	0.54	0.84	0.69
	23 Jan	0.96	0.81	0.97
E4	02 Feb	0.75	0.66	0.77
	03 Feb	0.78	0.65	0.89
	04 Feb	0.91	0.79	0.85
	05 Feb	0.88	0.90	0.91
	06 Feb	0.084	0.78	0.003
E5	13 Feb	0.78	0.90	0.60
	14 Feb	0.25	0.96	0.42
	15 Feb	0.86	0.76	0.82
	16 Feb	0.32	0.54	0.94
	17 Feb	0.36	0.99	0.92
E6	09 Mar	0.44	0.93	0.22
	10 Mar	0.87	0.83	0.77
E7	15 Apr	0.34	0.97	0.04
	16 Apr	0.28	0.35	0.82
	17 Apr	0.85	0.95	0.69
E8	20 May	0.77	0.87	0.88
	21 May	0.76	-	0.76
	22 May	0.23	0.76	0.036
E9	18 Oct	0.81	0.97	0.76
	19 Oct	0.89	0.92	0.97
E10	10 Nov	0.13	0.27	0.69
	11 Nov	0.26	0.18	0.92
E11	17 Nov	0.89	0.99	0.97
	18 Nov	0.56	0.94	0.86
	19 Nov	0.70	0.23	0.87
	20 Nov	0.75	0.99	0.44
E12	4 Dec	0.75	0.99	0.84
	5 Dec	0.37	0.64	0.10
E13	19 Dec	0.16	0.82	0.25
	20 Dec	0.10	0.28	0.08
	21 Dec	0.038	0.11	0.03

Table S5 Correlations of BC vs CO, and BrC vs CO in different episodic events for day, night, and 24 h.

Events	Dates	BC vs CO (r^2)			BrC vs CO (r^2)		
		Day-time	Night-time	24-hr	Day-time	Night-time	24-hr
E1	07 Jan	-	-	-	-	-	-
E2	12 Jan	0.07	0.09	0.016	0.06	0.14	0.007
	13 Jan	0.87	0.92	0.91	0.87	0.70	0.67
	14 Jan	0.99	0.58	0.81	1.0	0.98	0.89
	15 Jan	0.22	0.59	0.63	-0.59	0.44	0.55
	16 Jan	0.65	0.57	0.70	0.97	0.02	0.25
	17 Jan	0.96	0.80	0.85	0.98	0.01	0.81
	18 Jan	-	-0.74	-0.74	-	-0.33	-0.33
	19 Jan	-	0.44	0.44	-	0.59	0.59
E3	21 Jan	-	-	-	-	-	-
	22 Jan	0.12	0.91	0.11	0.20	0.98	0.07
	23 Jan	-	0.4	0.40	-	0.01	0.008
E4	02 Feb	0.61	0.71	0.71	0.68	0.82	0.81
	03 Feb	0.60	0.50	0.58	0.66	0.24	0.51
	04 Feb	0.89	0.27	0.46	0.78	0.48	0.53
	05 Feb	0.87	0.42	0.67	0.91	0.56	0.64
	06 Feb	0.12	0.02	0.05	-0.51	-0.44	-0.55
E5	13 Feb	0.97	0.43	0.69	0.96	0.48	0.67
	14 Feb	0.23	-0.90	0.28	-0.25	0.65	0.13
	15 Feb	0.96	0.01	0.20	0.84	-0.01	0.06
	16 Feb	0.01	0.53	0.04	-0.26	0.58	0.01
	17 Feb	0.99	0.60	0.48	0.99	0.72	0.54
E6	09 Mar	-	-	-	-	-	-
	10 Mar	-	-	-	-	-	-
E7	15 Apr	0.91	0.14	0.01	0.95	0.05	0.06
	16 Apr	0.68	0.17	0.01	0.28	0.12	0.17
	17 Apr	0.01	0.85	0.01	0.01	0.72	0.01
E8	20 May	0.64	0.65	0.56	0.88	0.79	0.87
	21 May	-	0.88	0.89	-	0.97	0.97
	22 May	0.73	0.02	0.32	0.84	0.89	0.89
E9	18 Oct	0.49	0.01	0.11	0.55	0.29	0.24
	19 Oct	0.98	0.95	0.96	0.86	0.89	0.87
E10	10 Nov	-	-	-	-	-	-
	11 Nov	0.67	0.88	0.68	0.67	0.93	0.60
E11	17 Nov	0.78	0.92	0.67	0.70	0.93	0.47
	18 Nov	0.26	1.0	0.39	0.20	1.00	0.27
	19 Nov	-	-	-	-	-	-
	20 Nov	-	-0.99	-0.99	-	-0.98	-0.98
E12	4 Dec	-	-	-	-	-	-
	5 Dec	-0.97	0.45	-0.42	0.75	0.45	0.26
E13	19 Dec	0.58	0.76	0.74	0.51	0.48	0.37
	20 Dec	0.62	0.19	0.23	0.01	0.01	0.01
	21 Dec	0.11	0.11	0.01	0.14	0.68	0.11

Table S6 Rate of change of O_3 concentration (ppb h^{-1}) at the rural IGP-CARE site and other Indian sites reported in previous studies.

Sampling Sites	**Rate of change at (08:00-11:00 h) ppb h^{-1}**	**Rate of change at (17:00-19:00 h) ppb h^{-1}**	**References**
IGP-CARE*	**6.8**	**-5.3**	Present study
	9.3^{W}	-6.2^{W}	
	7.5^{S}	-5.3^{S}	
	3.3^{M}	-3.0^{M}	
	7.3^{A}	-7.1^{A}	
Delhi	4.9	-3.7	Verma et al.[6], Tiwari et al.[7] and Tyagi et al.[8]
Gadanki	4.6	-2.6	Ojha et al. [9]
Mohal	7.3	-5.9	Sharma et al.[10]
Anantpur	4.6	-2.5	Reddy et al.[11]
Ahmedabad	5.9	-6.4	Lal et al.[12]
Thumba	4.8	-2.6	Nair et al.[13]
Joharpur	4.5	-3.3	Debaje and Kakade (2006)[14]
Tranquebar	3.1	-2.8	Debaje et al. (2003)[15]

* Rural site; W=winter; S=Summer; M=monsoon; A=autumn

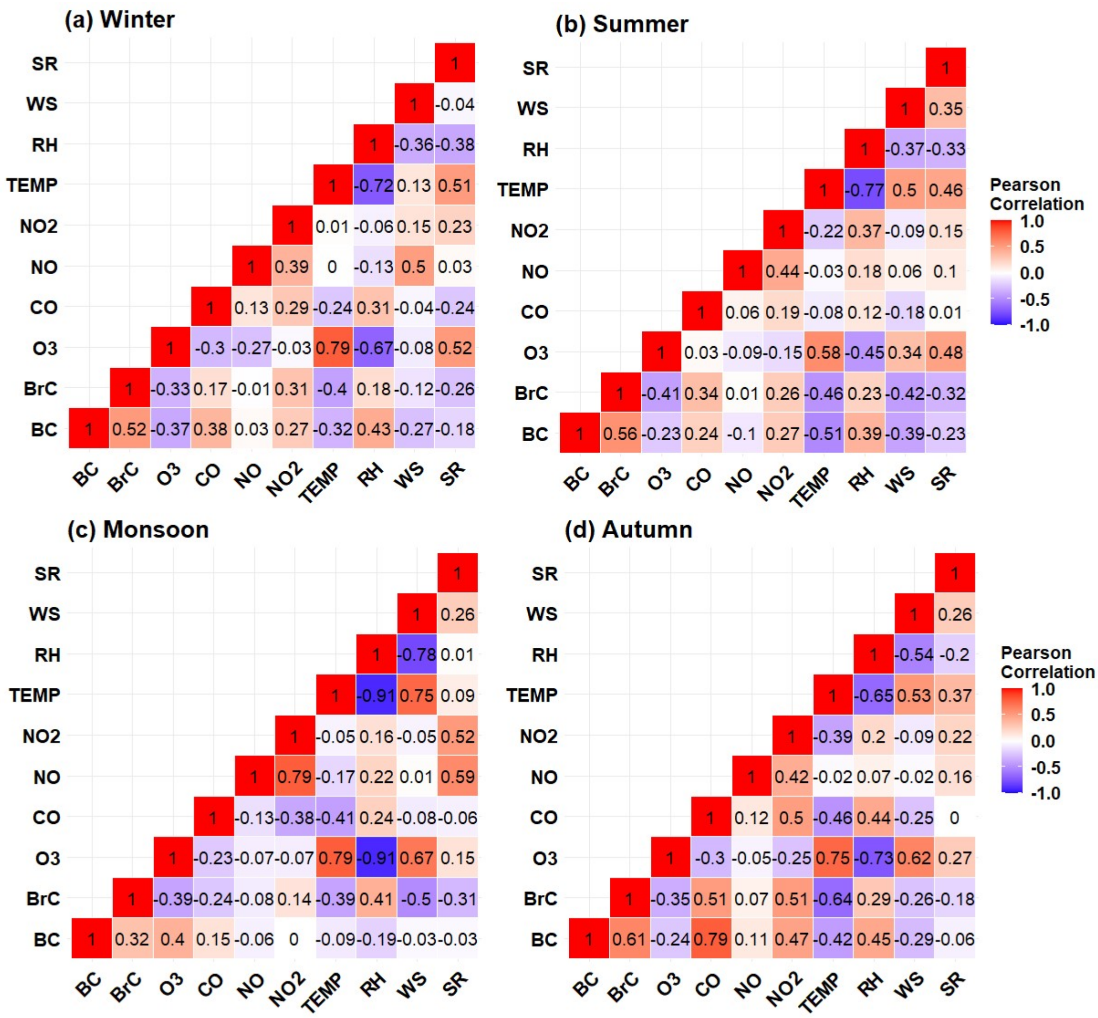

Fig S1. The Pearson's correlations amongst SLCPs and meteorological parameters in different seasons during the year 2017 at IGP-CARE.

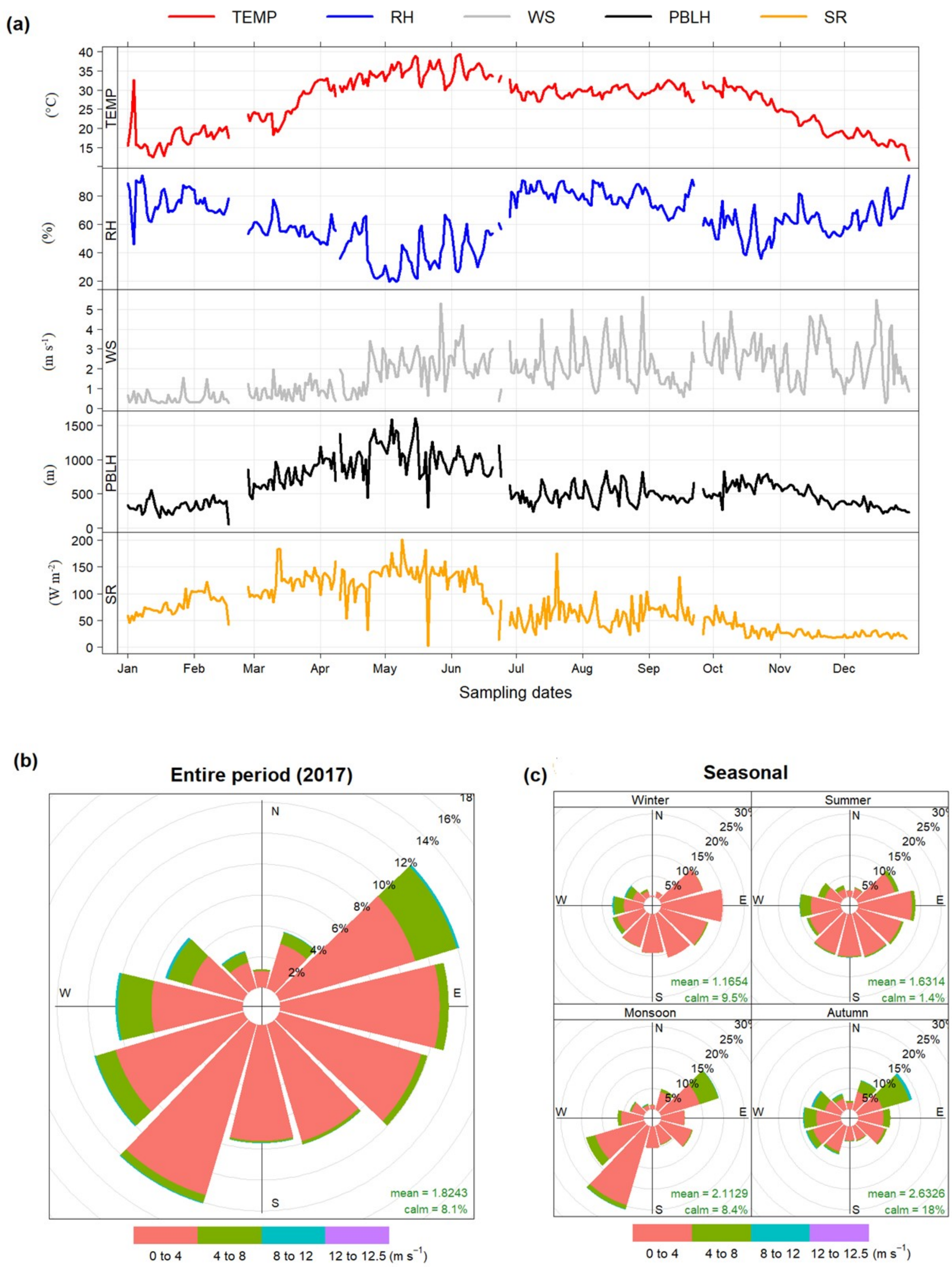

Fig S2. Time series of (a) meteorological parameters [temperature (TEMP), relative humidity (RH), wind speed (WS), planetary boundary layer height (m), and solar radiation (SR)], (b) wind rose plot for entire sampling period, and (c) seasonal wind rose plots at IGP-CARE during sampling year 2017.

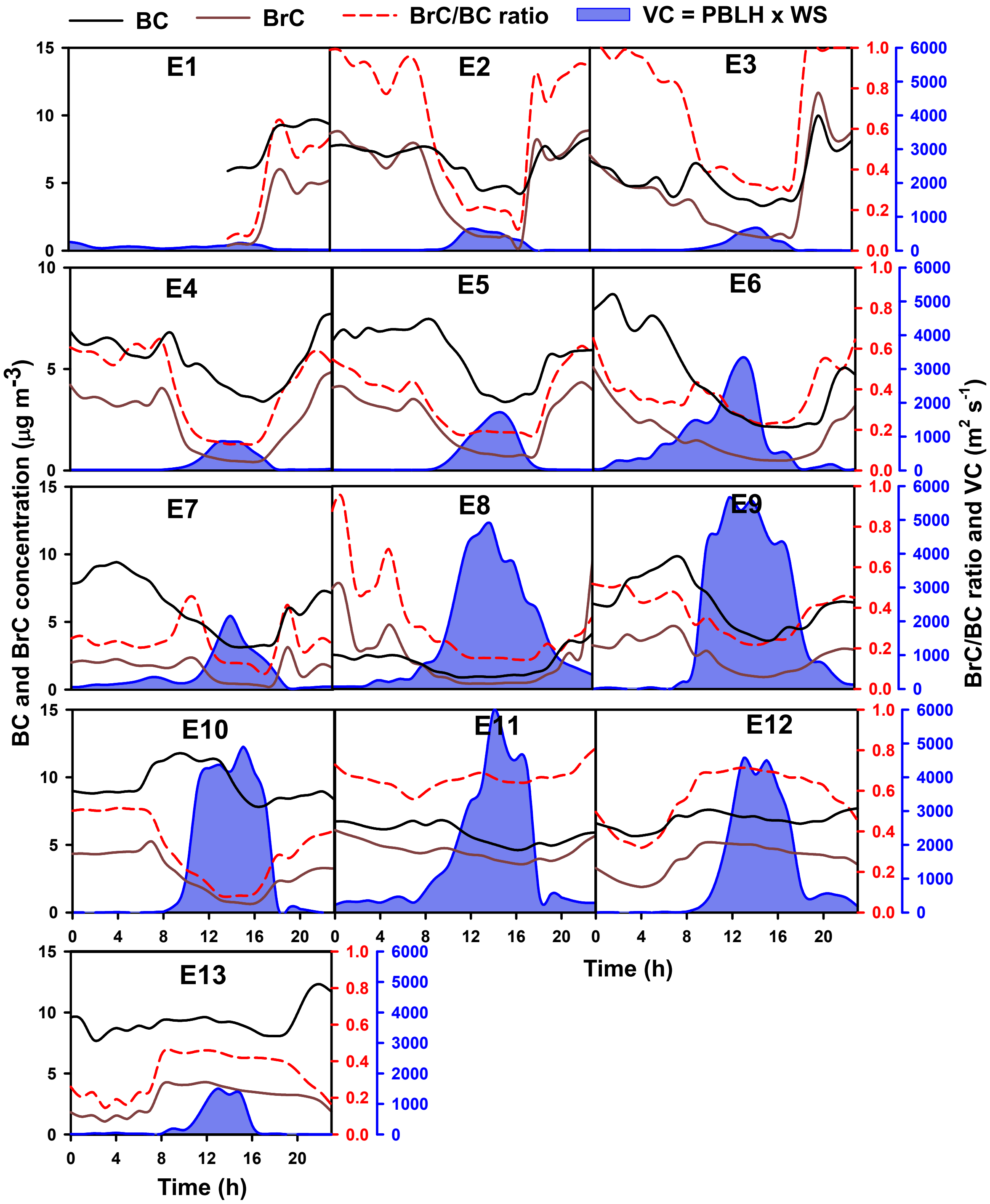

Fig S3. Diel variations of BC, BrC, and BrC/BC ratio with ventilation coefficient (VC) during thirteen episodic events at IGP-CARE in 2017. VC is the product of the planetary boundary layer height (PBLH) and wind speed (WS)] and is a measure of atmospheric conditions affecting pollutant dispersion.

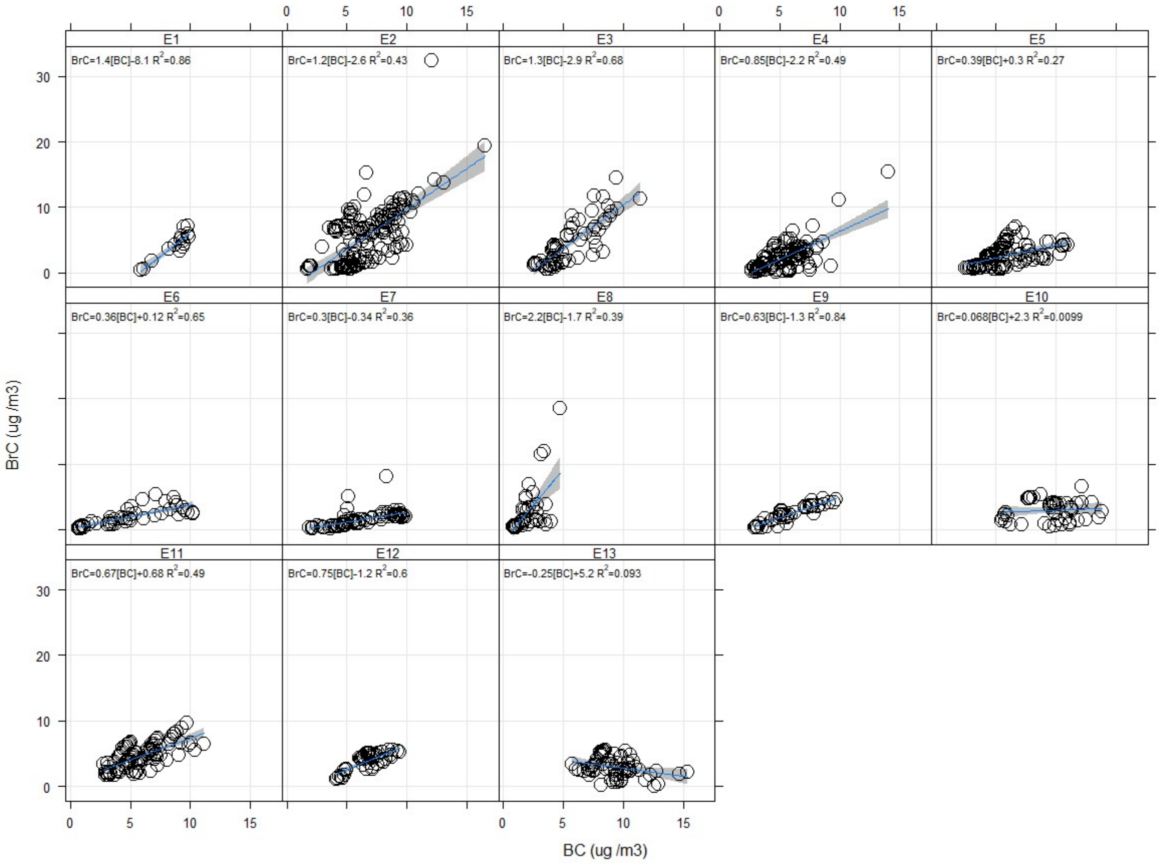

Fig S4. Scatter plots showing the relationships between BC and BrC during thirteen episodic events at IGP-CARE during the sampling year 2017.

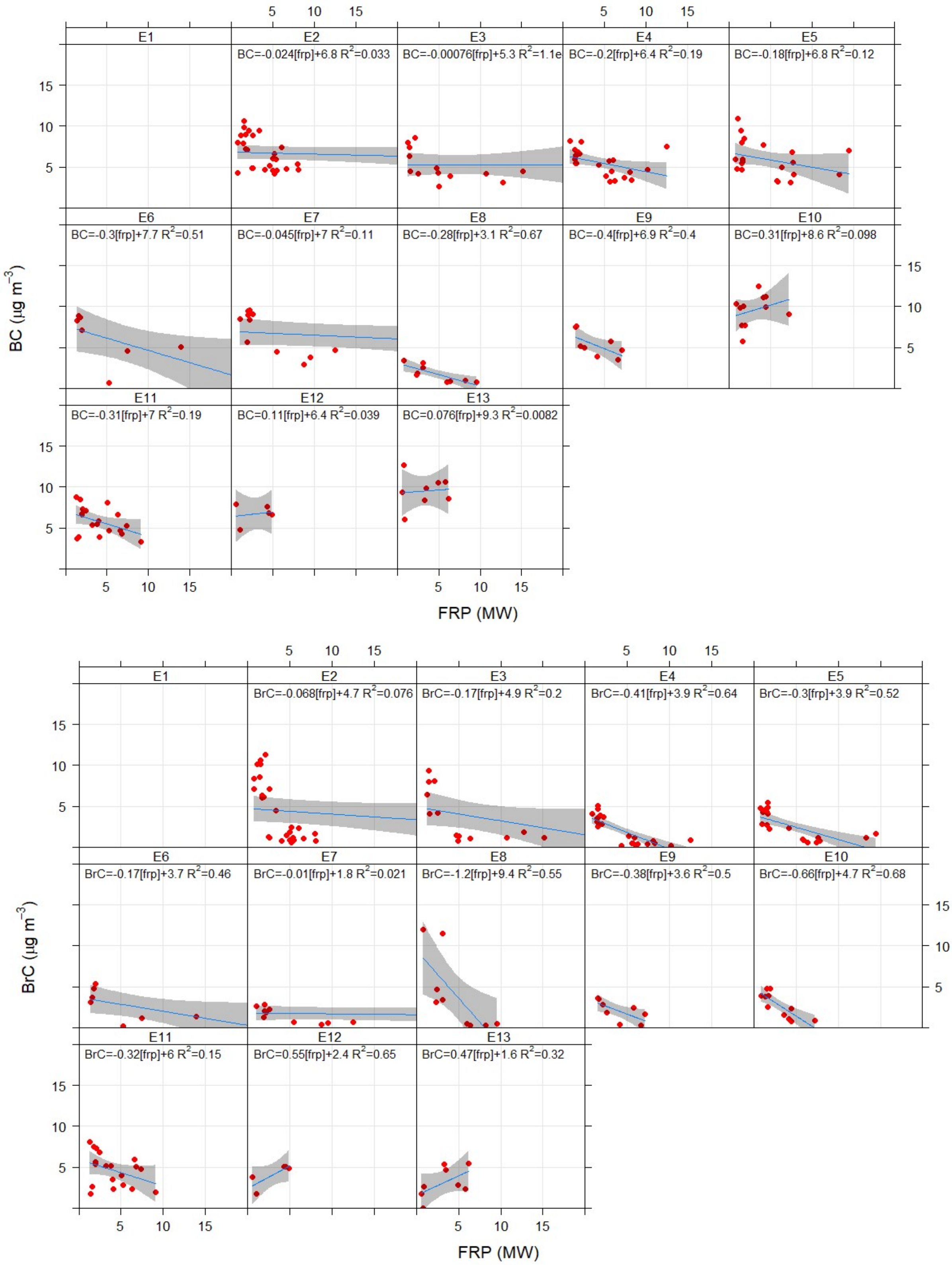

Fig S5. Correlations between BC and BrC with Fire Radiative Power (FRP) thirteen episodic events at IGP-CARE during the sampling year 2017.

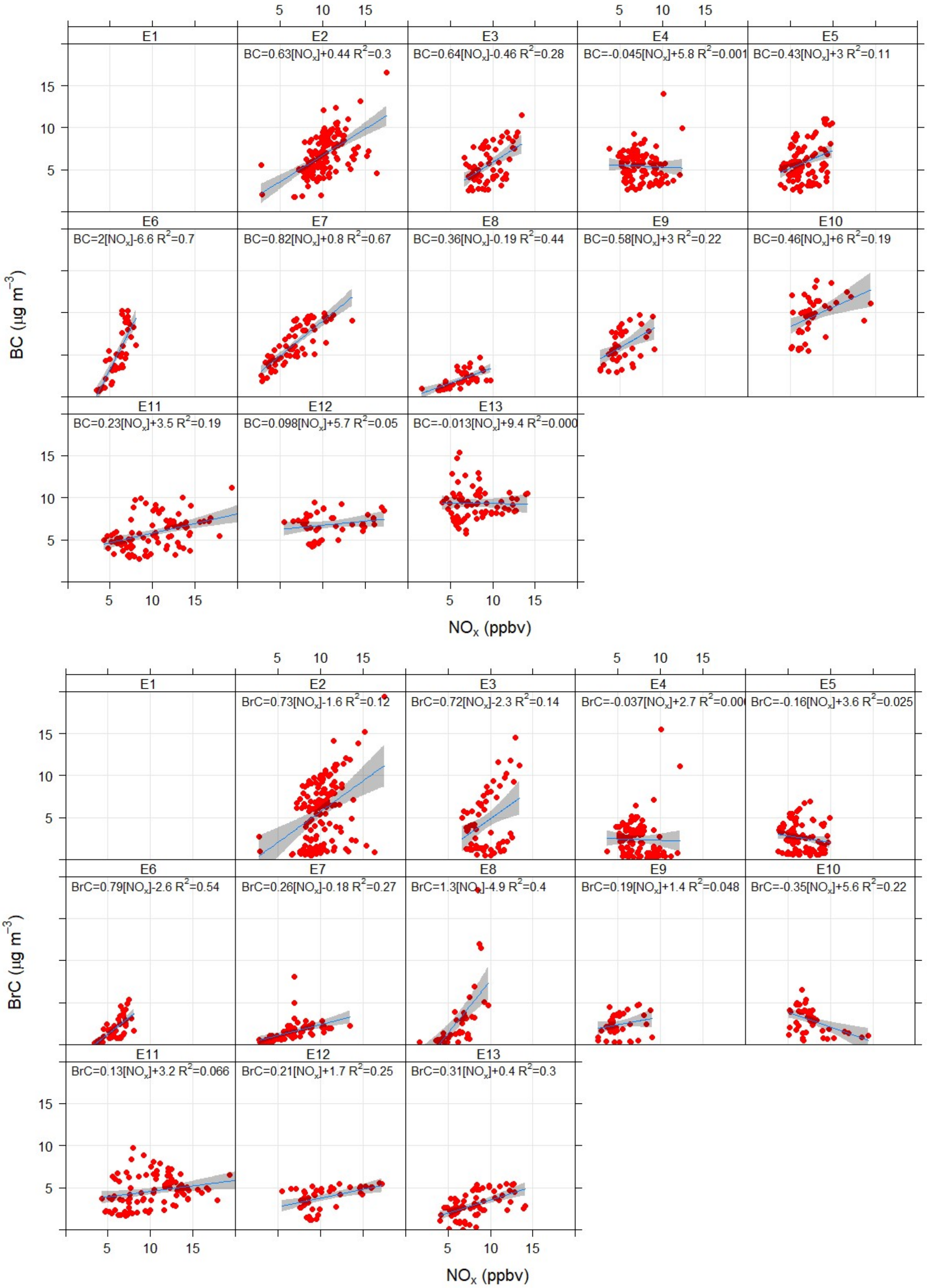

Fig S6. Scatter plots showing the relationships between BC and NO_X (upper panel) and BrC and NO_X (bottom panel) during thirteen episodic events at IGP-CARE during the sampling year 2017.

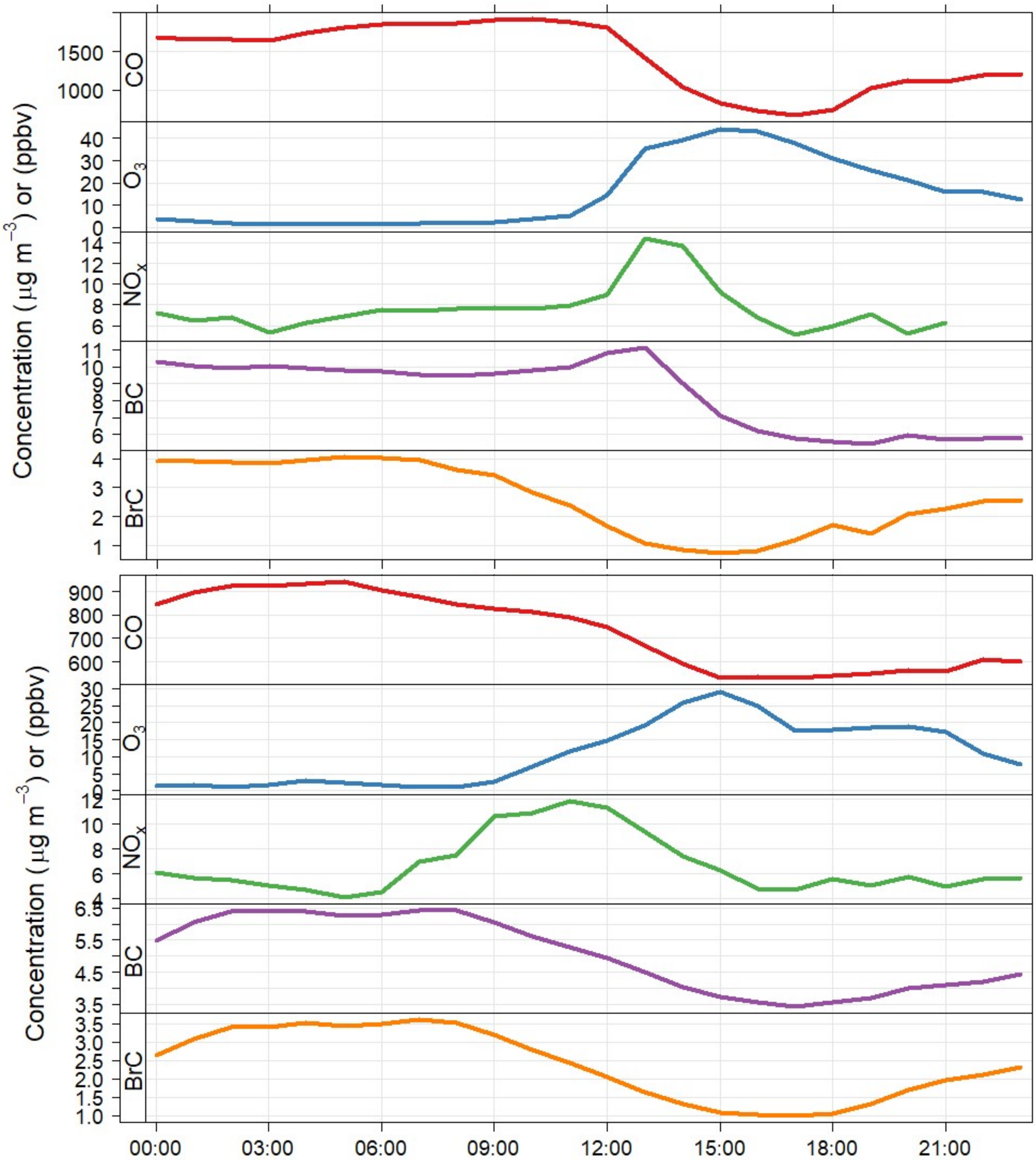

Fig S7. Time series of SLCFs concentration for the event E10 (10 and 11 November 2017).

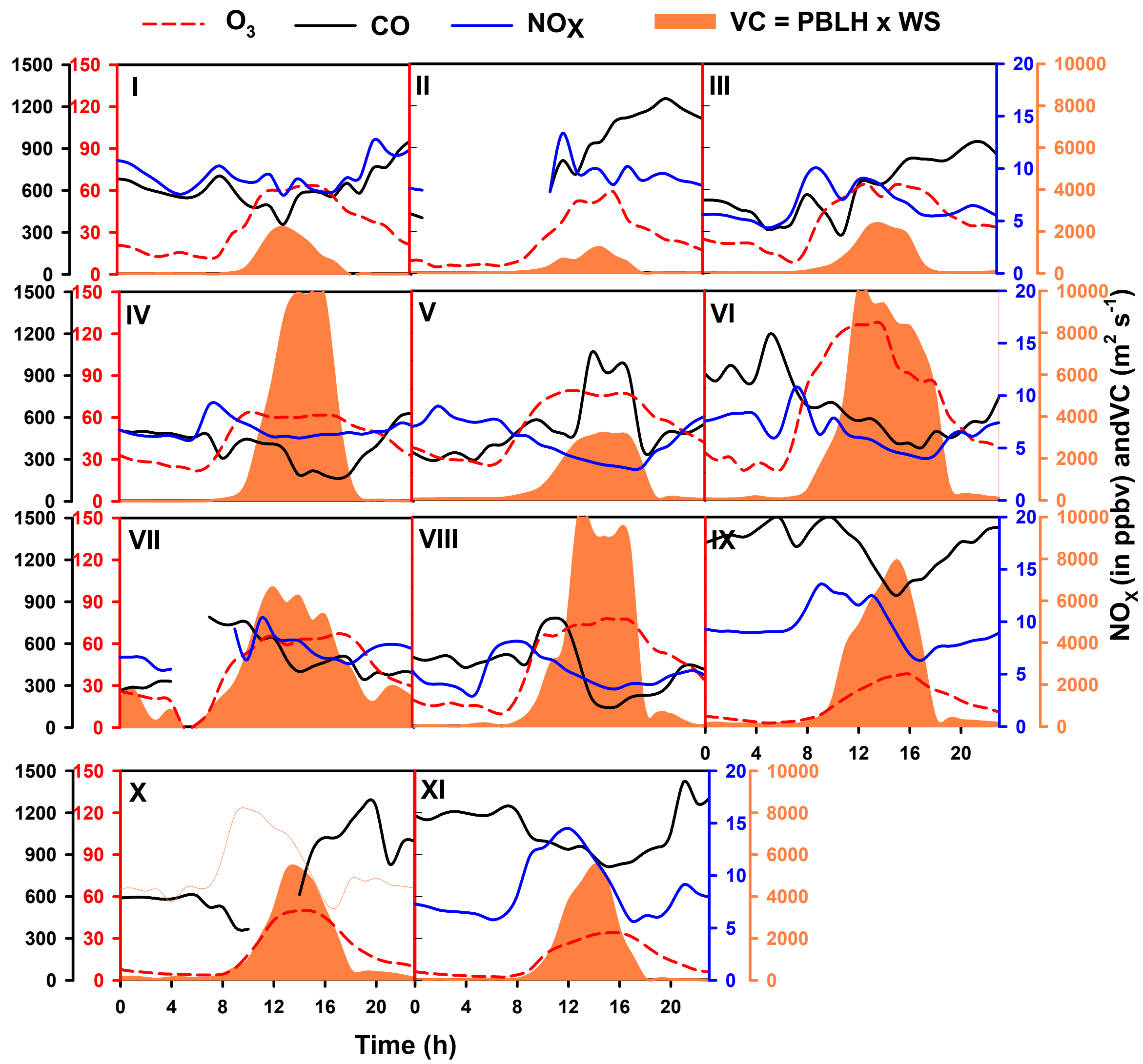

Fig. S8. Diel variation of O_3, CO, and NO_X concentrations with the ventilation coefficient (VC) during eleven episodic events at IGP-CARE in the sampling year 2017. VC is the product of the planetary boundary layer height (PBLH) and wind speed (WS)] and is a measure of atmospheric conditions and dispersion of pollutants.

.

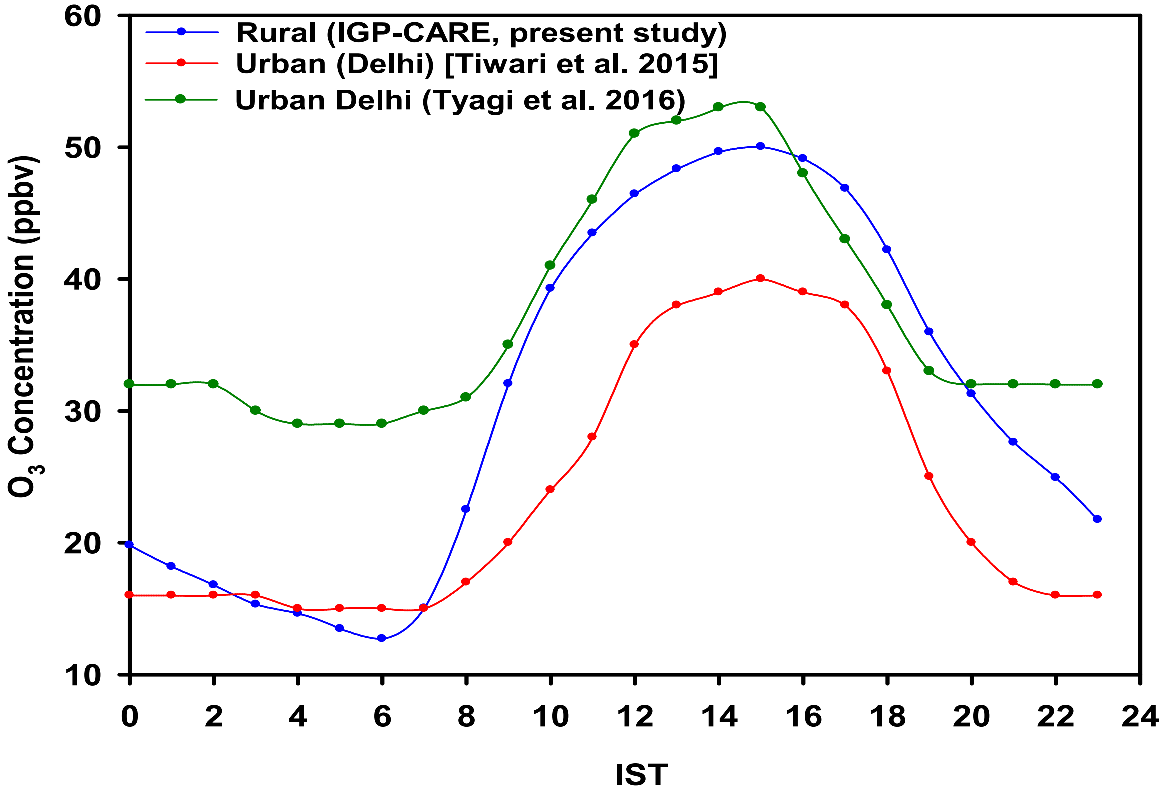

Fig. S9. Comparison of diurnal variation in O_3 at (a) IGP-CARE (rural site, present study) and an urban site in Delhi. (b,c) Diurnal variation of O_3 at an urban site as reported by Tiwari et al.[7] and Tyagi et al. (2016)[8] for the year 2014.

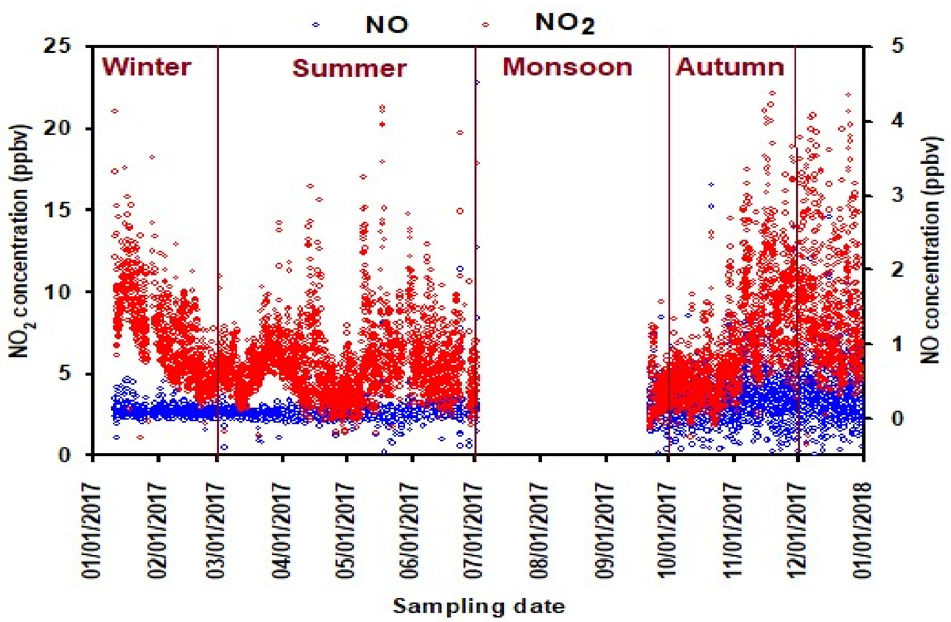

Fig. S10. 24-h NO and NO_2 concentrations at the IGP-CARE site during the sampling period.

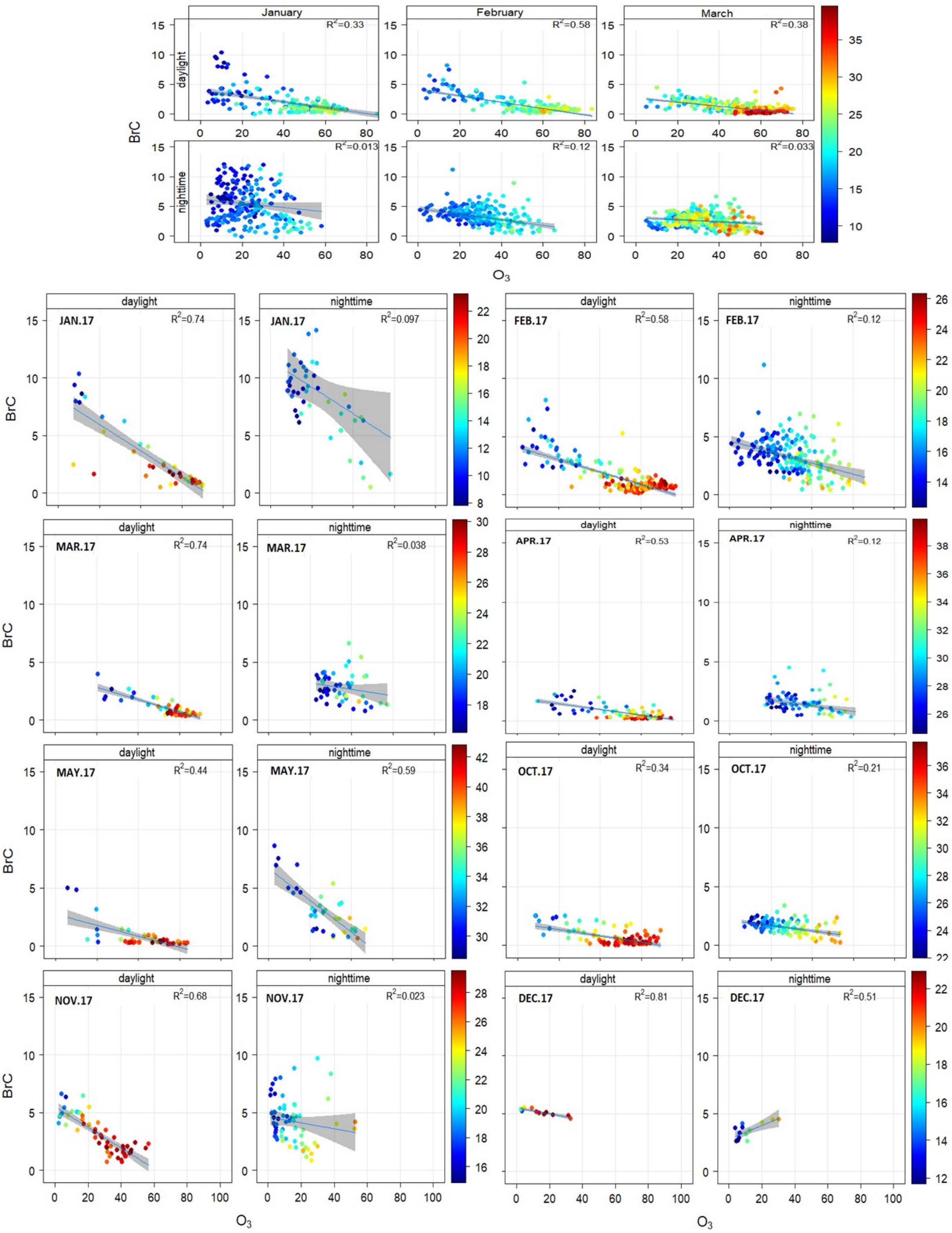

Fig. S11. Daytime and night-time correlation of BC and BrC with O_3 as functions of temperature (shown using a color scale) for January, February (winter), March, April, and May (summer) and October and November (autumn) over IGP-CARE.

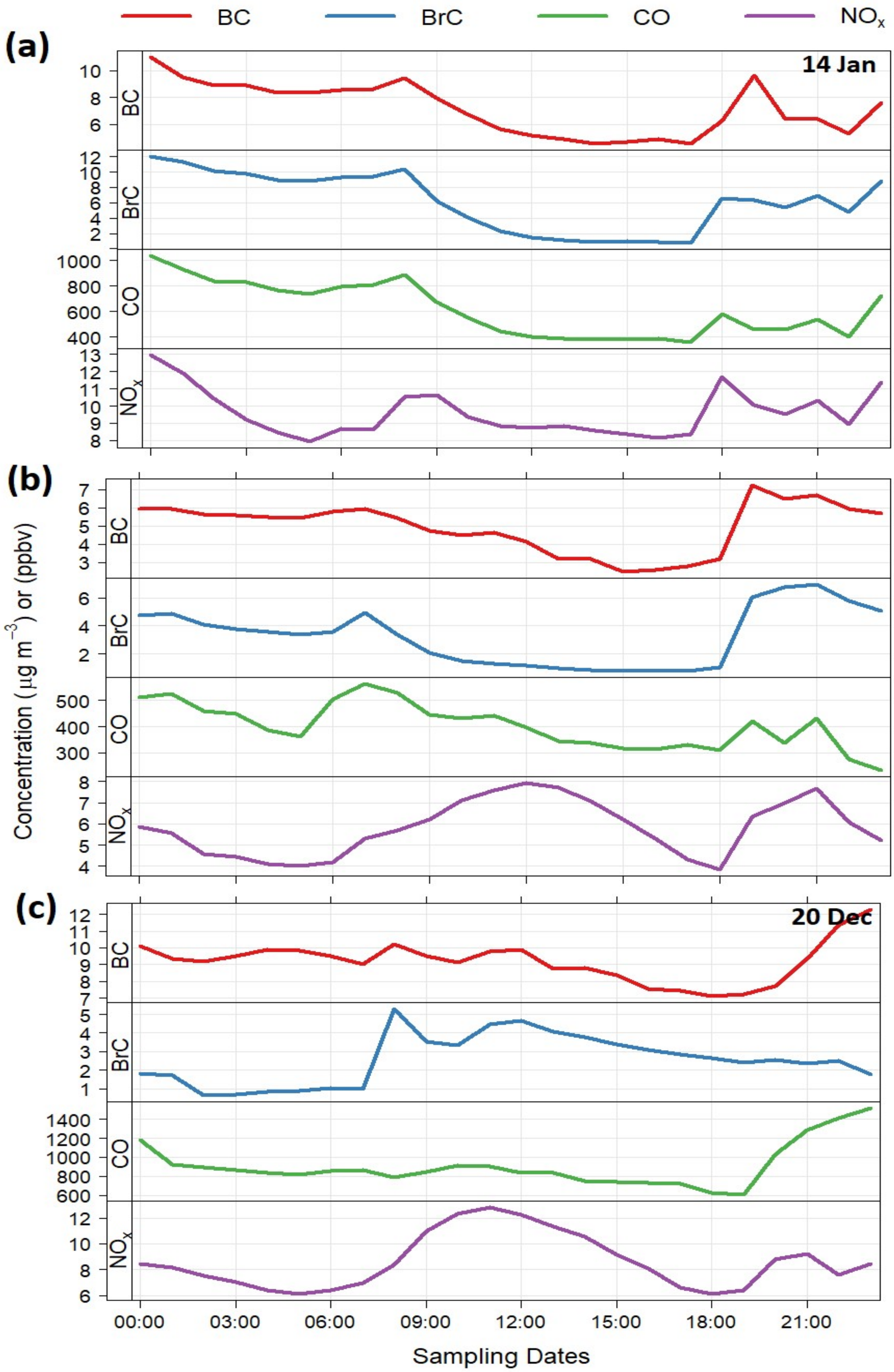

Fig. S12. Time Series plots of SLCPs concentration in few events to show typical evening and morning peaks of biomass burning.

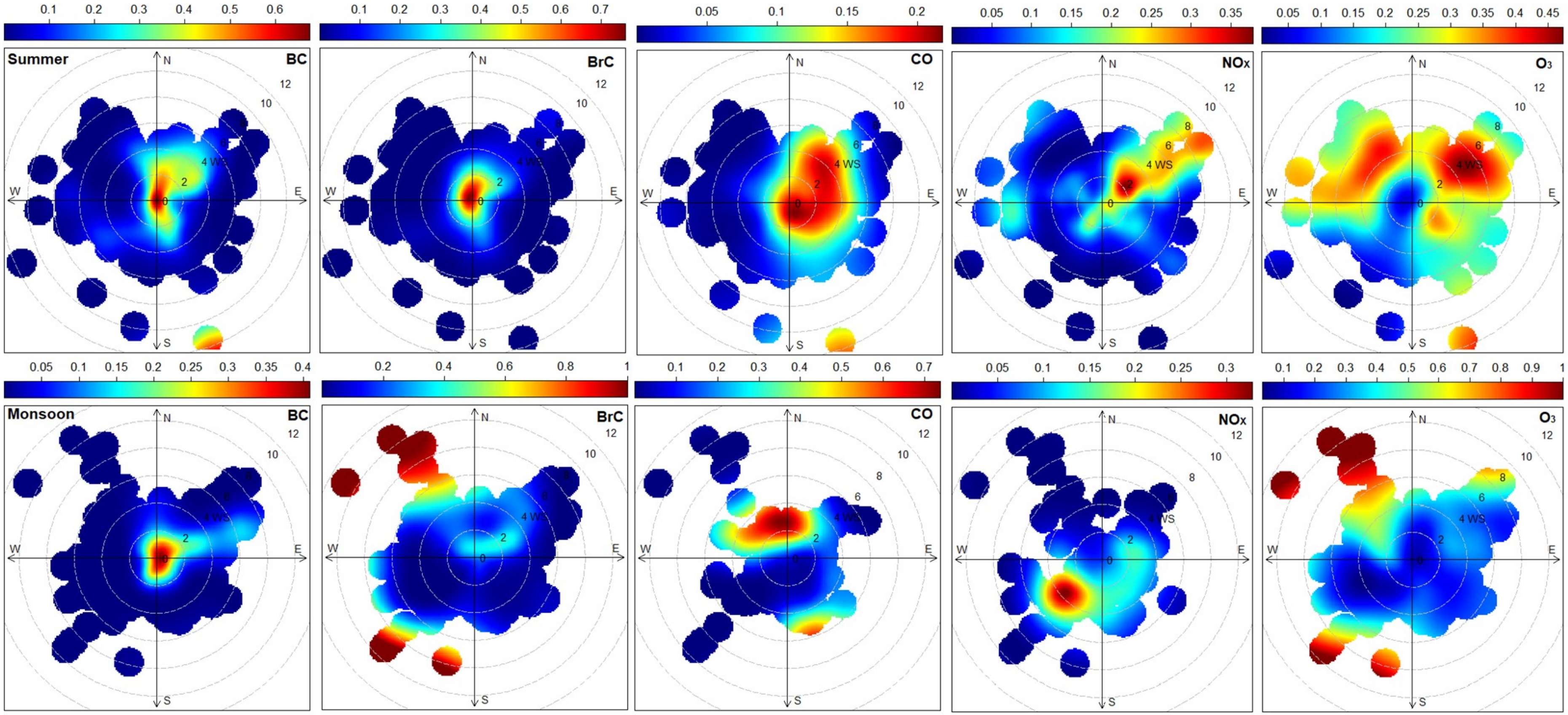

Fig S13. CBPF plots of SLCPs pollutants in summer (top panel) and monsoon (bottom panel) season over IGP-CARE during year 2017. Threshold criteria was chosen at 75th percentile of SLCPs pollutants.

Text S3 Cluster analysis using back trajectories

The seasonal cluster analysis is presented in Fig. S9 of the SI. During winter, the shortest and slow-moving cluster C1 originating from the north contributed 29.1% of the group of trajectories, the C2 cluster from Delhi and Punjab contributed 27.6%, and the moving C3 and C4 clusters originating from the northwest (including Iran, Afghanistan, and Pakistan; C3+C4) contributed 43.3% of the ensemble of trajectories governing regional/super-regional transport. The summer season is highly dominated by air mass trajectories moving from the northwest of IGP region (C2+C1; 56.0% of ensemble of trajectories) that originated from Delhi, Punjab, Haryana Afghanistan, and Pakistan. About 24% of trajectories are clustered in the east in close proximity of one-another, indicating the influence of local air masses (Fig. S9 in the SI). C4 is fast-moving cluster (20.2% of ensemble of trajectories) originating from Gujarat in the southwest and moving across Rajasthan, transporting significant amounts of dust to the receptor site. In general, the trajectory cluster pattern for monsoon season originates from the Arabian sea (C3+C4; 56.1% of ensemble of trajectories), and Bay of Bengal (C1; 19.2%), and other cluster (C2) originates from local air masses, accounting for 24.7% of trajectories (Fig. S9 in the SI). During autumn season, the dominant transport patterns were short-moving local air masses recirculating, but some regional trajectory clusters were observed. Clusters C1, C2, and C3 were slow-moving air masses originating in the immediate northwestern and eastern regions of the IGP and contributed 90% of trajectories. Overall, the cluster analysis explained significant amounts of the seasonal variability of local and regional/super-regional air masses. The cluster analysis clarified the comparative importance of regional and interannual sources by quantifying differences between the clusters in terms of SLCPs concentration and % of each cluster.

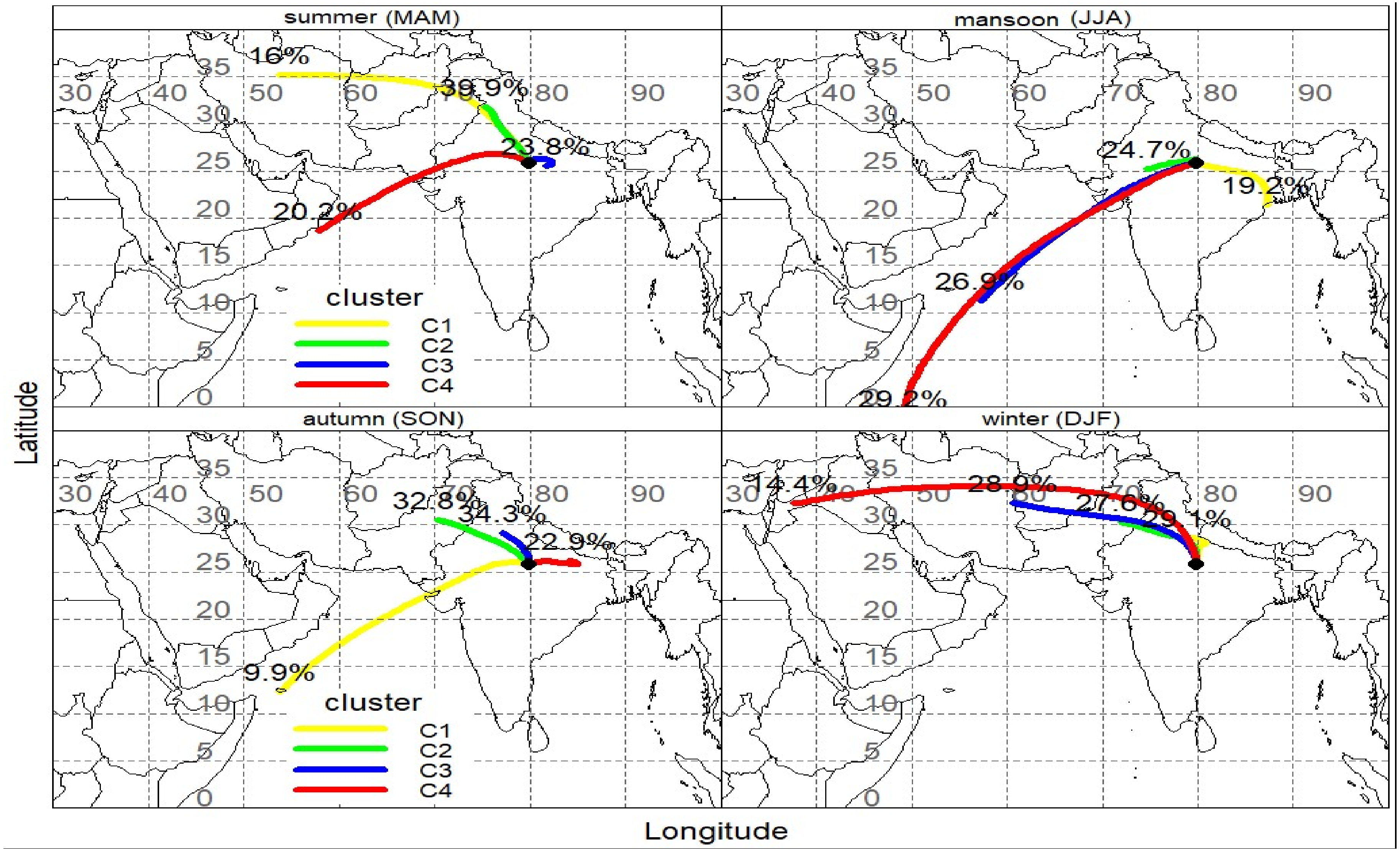

Fig S14. Trajectory cluster analysis using 4-days air mass back-trajectories over the IGP-CARE site at 500 m ABL in four different seasons. The black circle denotes IGP-CARE (receptor site, rural site of IGP-region).

References

1 B. A. Begum, S. K. Biswas and P. K. Hopke, Key issues in controlling air pollutants in Dhaka, Bangladesh, *Atmos. Environ.*, 2011, **45**, 7705–7713.

2 P. K. Hopke, K. Ito, T. Mar, W. F. Christensen, D. J. Eatough, R. C. Henry, E. Kim, F. Laden, R. Lall, T. V Larson and G. D. Thurston, PM source apportionment and health effects : 1 . Intercomparison of source apportionment results, 2006, 275–286.

3 A. V. Polissar, P. K. Hopke and R. L. Poirot, Atmospheric aerosol over Vermont: Chemical composition and sources, *Environ. Sci. Technol.*, 2001, **35**, 4604–4621.

4 D. Carslaw and K. Ropkins, The openair manual - an R package for air quality data analysis, *Environ. Model. Softw.*, 2012, **27**, 52–61.

5 D. C. Carslaw and K. Ropkins, Openair - An r package for air quality data analysis, *Environ. Model. Softw.*, 2012, **27–28**, 52–61.

6 N. Verma, A. Lakhani and K. Maharaj Kumari, High ozone episodes at a semi-urban site in India: Photochemical generation and transport, *Atmos. Res.*, 2017, **197**, 232–243.

7 S. Tiwari, A. Dahiya and N. Kumar, Investigation into relationships among NO, NO2, NOx, O3, and CO at an urban background site in Delhi, India, *Atmos. Res.*, 2015, **157**, 119–126.

8 S. Tyagi, S. Tiwari, A. Mishra, P. K. Hopke, S. D. Attri, A. K. Srivastava and D. S. Bisht, Spatial variability of concentrations of gaseous pollutants across the National Capital Region of Delhi, India, *Atmos. Pollut. Res.*, 2016, **7**, 808–816.

9 N. Ojha, M. Naja, K. P. Singh, T. Sarangi, R. Kumar, S. Lal, M. G. Lawrence, T. M. Butler and H. C. Chandola, Variabilities in ozone at a semi-urban site in the Indo-Gangetic Plain region: Association with the meteorology and regional processes, *J. Geophys. Res. Atmos.*, 2012, **117**, 1–19.

10 P. Sharma, P. Sharma, S. Jain and P. Kumar, An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi, *Atmos. Environ.*, 2013, **70**, 7–17.

11 B. Suresh Kumar Reddy, K. Raghavendra Kumar, G. Balakrishnaiah, K. Rama Gopal, R. R. Reddy, L. S. S. Reddy, Y. Nazeer Ahammed, K. Narasimhulu, K. Krishna Moorthy and S. Suresh Babu, Potential source regions contributing to seasonal variations of black carbon aerosols over Anantapur in southeast India, *Aerosol Air Qual. Res.*, 2012, **12**, 340–354.

12 S. Lal, M. Naja and B. H. Subbaraya, Seasonal variations in surface ozone and its precursors over an urban site in India, *Atmos. Environ.*, 2000, **34**, 2713–2724.

13 P. R. Nair, D. Chand, S. Lal, K. S. Modh, M. Naja, K. Parameswaran, S. Ravindran and S. Venkataramani, Temporal variations in surface ozone at Thumba (8.6°N, 77°E)- a tropical coastal site in India, *Atmos. Environ.*, 2002, **36**, 603–610.

14 S. B. Debaje and A. D. Kakade, Weekend Ozone Effect over Rural and Urban Site in India, *Aerosol Air Qual. Res.*, 2006, **6**, 322–333.

15 S. B. Debaje, S. J. Jeyakumar, K. Ganesan, D. B. Jadhav and P. Seetaramayya, Surface

ozone measurements at tropical rural coastal station Tranquebar, India, *Atmos. Environ.*, 2003, **37**, 4911–4916.