Supplementary Information (SI) for Environmental Science: Processes & Impacts.

Interplay between Oxidative Potential and Health Risk of $PM_{2.5}$-Bound Metals at a Site of Indo-Gangetic Plain—Exploring the Influence of Biomass Burning

Muskan Agarwal, Isha Goyal and Anita Lakhani*

Department of Chemistry, Faculty of Science, Dayalbagh Educational Institute, Dayalbagh, Agra-282005

Corresponding Author's e-mail: anita.lakhani01@gmail.com

Health Risk Assessment (HRA)

The daily intake of the chemical through oral ingestion (CDI^{ing}_{ij}, mg kg^{-1} day^{-1}), exposure concentration through inhalation (EC^{inh}_{ij}, μg m^{-3}), and skin absorption dose through dermal contact (DAD^{der}_{ij}, mg kg^{-1} day^{-1}) (Zhang et al., 2021 and USEPA, 2014) were calculated using the Eqs. (S1)–(S3) below:

$$CDI_{ij}^{ing} = \frac{C_{ij} \times IR_{ing} \times EF \times ED \times CF}{(BW \times AT)} \quad \text{(S1)}$$

$$EC_{ij}^{inh} = \frac{C_{ij} \times IR_{inh} \times ET \times EF \times ED}{(BW \times AT_n)} \quad \text{(S2)}$$

$$DAD_{ij}^{der} = \frac{C_{ij} \times SA \times AF \times ABS \times EF \times ED \times CF}{(BW \times AT)} \quad \text{(S3)}$$

The corresponding hazard quotient (HQ) was calculated as shown in Eqs. (S4)–(S7) and the carcinogenic risk (CR) of toxic metals through the three pathways were further evaluated based on Eqs. (S8)–(S11).

$$HQ_{ing} = \frac{CDI_{ij}^{ing}}{RfD_j^{ing}} \quad \text{(S4)}$$

$$HQ_{inh} = \frac{EC_{ij}^{inh}}{RfC_j^{inh} \times 1000} \quad \text{(S5)}$$

$$HQ_{der} = \frac{DAD_{ij}^{der}}{RfD_j^{der} \times ABS} \quad \text{(S6)}$$

$$Hazard\ index\ (HI_i) = \sum(HQ_{ij}) \quad \text{(S7)}$$

$$CR_{ing} = LCDI_{ij}^{ing} \times SF \quad \text{(S8)}$$

$$CR_{inh} = LEC_{ij}^{inh} \times IUR \quad \text{(S9)}$$

$$CR_{der} = LDAD_{ij}^{der} \times RfD_j^{der} \times ABS \quad \text{(S10)}$$

Carcinogenic risk $(CR_i) = \sum(CR_{ij})$ (S11)

RfD_o (the oral reference doses (mg kg^{-1} day^{-1})), RfC_i (inhalation reference concentration (mg m^{-3})), $ABSG_I$ (gastrointestinal absorption factor), SF_o (oral slope factor (mg kg^{-1} day^{-1})) and IUR (inhalation unit risk (μg m^{-3}). A hazard index (HI), i.e., HQ all added up, is used for assessing the chronic effects of non-carcinogenic risks. When both HQs ≤ 1 and HI ≤ 1, there is no significant risk of chronic effects. By contrast, HQs > 1 or HI > 1 indicate a possibility of the occurrence of chronic effects (Murari et al., 2020). The carcinogen risk (CR) value shows the chance of developing any type of cancer by an individual due to a lifetime exposure to carcinogenic metals, divided into five categories: as very low ($CR \leq 10^{-6}$), low ($10^{-6} \leq CR < 10^{-4}$), moderate ($10^{-4} \leq CR < 10^{-3}$), high ($10^{-3} \leq CR < 10^{-1}$), and very high ($CR \geq 10^{-1}$) (Roy et al., 2019; Zhang et al., 2021a).
It is imperative to emphasize that the toxicity of Cr is predominantly linked to its hexavalent Cr(VI) form and thus health risk was estimated this form of Cr, which is capable of absorption through the respiratory and gastrointestinal tracts, and, to some extent, through dermal contact. In contrast, the trivalent Cr(III) is characterized by minimal absorption across all pathways. Thus, the concentration of Cr(VI) was estimated by taking the one-sixth part of total concentration of Cr as the ratio of Cr(VI) to Cr(III) in the air is about 1 : 6.

Positive Matrix Factorization (PMF)

The PMF v5.0 model requires two input files: one of the measured concentrations of the species and another for the estimated uncertainty of the concentration (Sharma & Mandal, 2017). A speciated data set can be viewed as a data matrix X of i by j dimensions, in which i number of samples and j chemical species are measured. The aim of multivariate receptor modeling, for example, with PMF, is to identify a number of factors p, the species profile f of each source, and the amount of mass g contributed by each factor to each individual sample which is given as:

$$X_{ij} = \sum_{k=1}^{p} .g_{ik} f_{kj} + e_{ij}$$

where e_{ij} is the residual for each sample/species.

Results are constrained so that no sample can have a negative source contribution. PMF allows each data point to be individually weighed. This feature allows the analyst to adjust the influence of each data point, depending on the confidence in the measurement. For example, data below detection limit can be retained for use in the model, with the associated uncertainty adjusted so these data points have less influence on the solution than measurements above the detection limit. The PMF solution minimizes the object function Q, based upon these uncertainties (u) as follows:

$$Q = \sum_{i=1}^{n} . \sum_{j=1}^{m} . \left[\frac{X_{ij} - \sum_{k=1}^{p} .g_{ik} f_{kj}}{u_{ij}} \right]$$

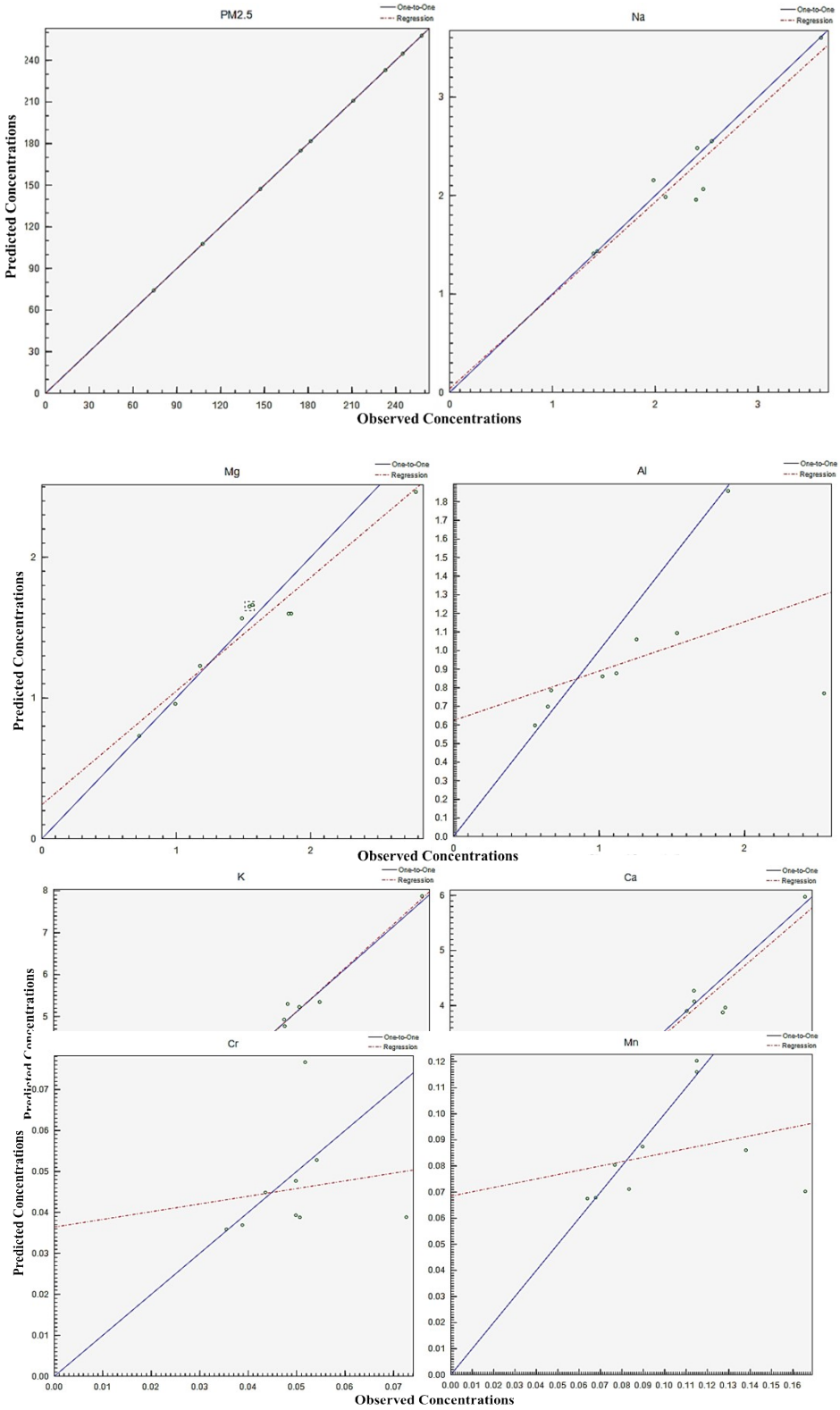
where X_{ij} are the measured concentration (in μg m^{-3}), u_{ij} are the estimated uncertainty (in μg m^{-3}), n is the number of samples, m is the number of species and p is the number of sources included

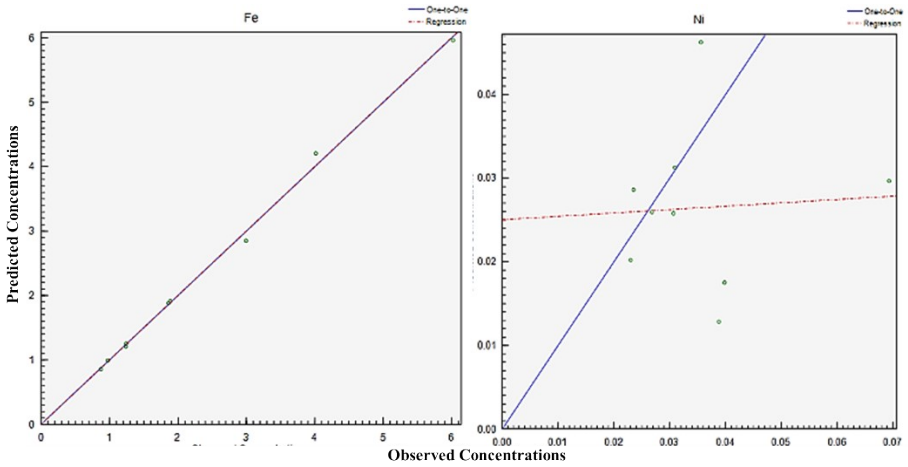
in the analysis. The detail descriptions of EPA PMF v.5.0 are described in its guide and EPA PMF User Guide (2008) (EPA PMF Guide, 2008). In this study, information on chemical properties of $PM_{2.5}$ samples has been used as input to the PMF model for the 14 metal species and DTTv studied. Categorization of quality of data was based on the signal to noise ratio (S/N) and the percentage of sample method detection limit (MDL). Those species which have S/N ≥ 2 were categorized as strong in data quality. Those with S/N between 0.2 and 2 were categorized as weak in quality. These species are not likely to provide enough variation in concentration and therefore contribute to the noise in the results. Those species with an S/N ratio below 0.2 are classified as bad values and were thus excluded from further analysis. In the present case, signal to noise ratio (S/N) estimated as N = 0.6 and the model performance in a base run showed coefficient of determination (R^2) (Table S1) between the modeled and experimental concentration of $PM_{2.5}$ and metals along with DTTv for each sample (Figure S1), respectively and most of the other chemical species are also well reconstructed.

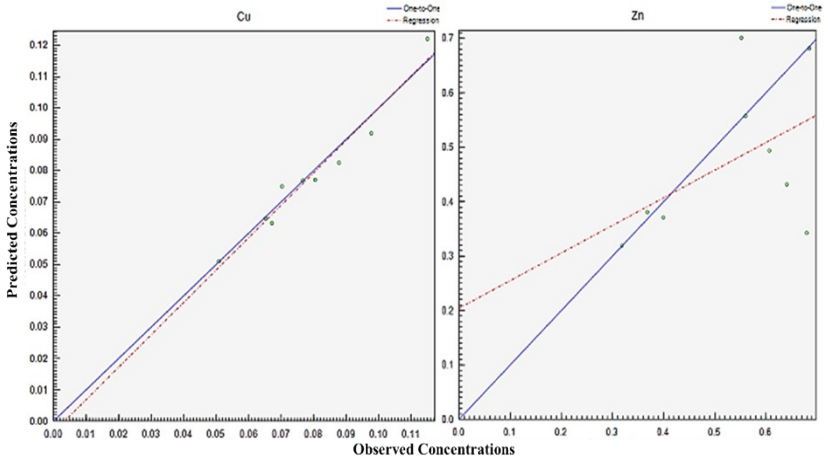
The PMF analysis was employed to unravel the contributions of diverse pollution sources to $PM_{2.5}$ mass within an area of Agra. The model fit parameter 'Q' was carefully assessed to ensure its placement within the FPEAK range, seeking the optimal factor resolution. Default robust mode was used to minimize the potential influence of extreme values on the solution. A meticulous trial-and-error approach was employed to systematically explore varying source numbers, pinpointing the optimal configuration. Subsequently, PMF was applied to the datasets using the selected factors, with close examination of the resultant Q value alterations. Notably, robust Q values represent those with minimized outlier impact, contrasting with true Q values, which lack outlier control. In this study, the remarkable proximity between robust and true Q values signified a reasonable model fit, even in the presence of outliers. Furthermore, to confirm attainment of a consistent global minimum and ensure equitable outlier fitting across random runs, a sufficiently narrow Q value range was verified. Analysis of the Q values (model fit parameter) variations revealed that a four-factor solution yielded the most robust and interpretable results.

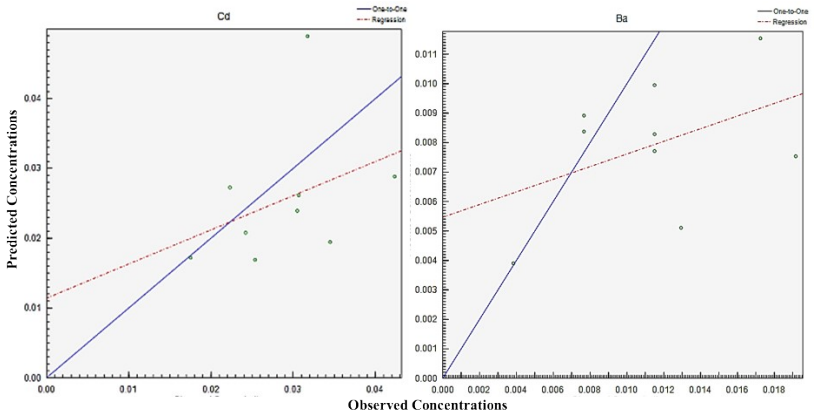
Table S1. Statistical Parameters observed during PMF Analysis

					KS Test	KS Test
Species	**Intercept**	**Slope**	**SE**	**r^2**	**Stat**	**P Value**
$PM_{2.5}$	0.00	1.00	0.00	1.00	0.18	0.93
Ca	0.14	0.94	0.32	0.93	0.21	0.82
Fe	0.00	1.00	0.09	1.00	0.20	0.87
Na	0.04	0.95	0.22	0.90	0.26	0.60
Mg	0.24	0.81	0.13	0.94	0.24	0.68
K	-0.12	1.02	0.22	0.98	0.23	0.71
Al	0.62	0.27	0.35	0.21	0.20	0.86
Ba	0.01	0.21	0.00	0.20	0.18	0.93
Cd	0.01	0.49	0.01	0.13	0.32	0.30
Cr	0.04	0.19	0.01	0.02	0.22	0.78
Cu	0.00	1.03	0.00	0.95	0.13	1.00
Mn	0.07	0.17	0.02	0.08	0.27	0.54
Ni	0.03	0.04	0.01	0.00	0.16	0.98
Pb	0.24	-0.05	0.08	0.00	0.17	0.97
Se	0.10	0.24	0.05	0.53	0.20	0.88
Zn	0.20	0.51	0.13	0.24	0.18	0.94
DTT_v	0.00	1.00	0.02	1.00	0.20	0.88


PM2.5
Na
Mg
Al
K
Ca
Cr
Mn
One-to-One
Regression
Predicted Concentrations
Observed Concentrations


Fe
Ni
One-to-One
Regression
Predicted Concentrations
Observed Concentrations


Cu
Zn
One-to-One
Regression
Predicted Concentrations
Observed Concentrations


Cd
Ba
One-to-One
Regression
Predicted Concentrations
Observed Concentrations

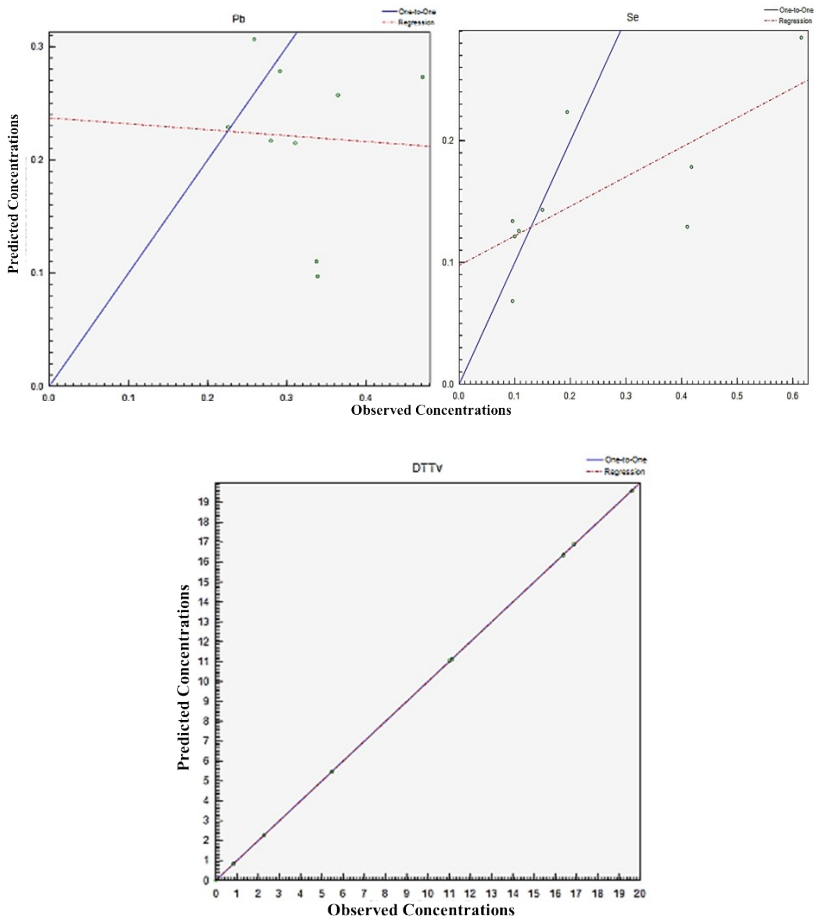


Figure S1. Linear Regression between Modeled/Predicted and Observed Concentrations of $PM_{2.5}$, Metals and DTTv

Tables

Table S2. t-test Analysis on Concentration of $PM_{2.5}$ and Meteorological Parameters during Haze and Non-Haze Period

Parameters	t-value	Significance (2-tailed)
PM	6.198	0.001
Temp	0.352	0.722
RH	7.077	0.001
WS	-0.524	0.620
WD	-0.074	0.943
SR	-1.453	0.167
Vis	-2.812	0.023

Significant at 95% confidence level ($p<0.05$)

Table S3. Mass Average Concentration of $PM_{2.5}$ and Mean Values of Meteorological Parameters during Haze and Non-Haze Period

Parameters	Haze Period			Non-Haze Period		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
$PM_{2.5}$ ($\mu g\ m^{-3}$)	225.6	29.9	190 -260	125.9	44.2	74- 182
Temperature (°C)	11.9	2.2	6.1-15.3	16.0	2.9	12.3-21.6
Relative Humidity (%)	89	2	90-94	66	11	54-82
Visibility (km)	1.4	0.5	0.5-1.6	10.2	1.7	4.6-8.1
Wind speed (m/s)	1.1	0.4	0.6-1.6	1.3	0.9	0.4-2.3
Wind Direction	N-NW			N-NW		

Table S4. Cluster Means of 24-hour Air Mass Trajectories during November, December and January

Months	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
November	12%	33%	27%	10%	6%	10%	3%
December	7%	12%	4%	25%	3%	32%	17%
January	28%	17%	13%	1%	24%	8%	9%

Table S4. Concentration of Metals during Haze and Non-Haze Period

Metals	Haze Period (µg m^{-3})	Non-Haze Period (µg m^{-3})
Ca	4.5±0.7	2.8±0.7
Fe	3.3±1.7	1.0±0.2
Na	2.6±0.5	1.8±0.4
Mg	1.9±0.5	1.1±0.2
K	5.6±1.2	4.1±0.76
Al	1.5±0.7	0.8±0.3
Ba	0.01±0.005	0.01±0.005
Cd	0.03±0.004	0.01±0.01
Cr	0.05±0.01	0.02±0.007
Cu	0.09±0.01	0.06±0.01
Mn	0.1±0.03	0.07±0.01
Ni	0.04±0.01	0.02±0.003
Pb	0.3±0.1	0.1±0.04
Se	0.3±0.1	0.2±0.1
Zn	0.6±0.1	0.5±0.1

Table S5. Linear Correlation Coefficient of Metals to DTTv Activity

Metals	Correlation coefficient (r)	
	Total Metal Fraction	Water-Soluble Fraction
Na	**0.91**	
Mg	**0.81**	
Al	0.17	
K	**0.90**	
Ca	0.17	
Cr	**0.81**	**0.81**
Mn	**0.59**	**0.67**
Fe	**0.65**	**0.66**
Ni	**0.70**	0.44
Cu	**0.78**	**0.65**
Zn	0.41	**0.50**
Cd	**0.72**	0.46
Ba	**0.74**	
Pb	0.24	0.26
Se	**0.57**	

Bold are the values having significant correlation with DTT_v

Table S6. The values used for all the parameters in Health Risk Assessment (HRA)

Factors	Notation	Unit	Values	
			Children	**Adults**
Ingestion Rate	IR_{ing}	mg/day	200	100
Exposure Frequency	EF	days/years	365	365
Exposure Duration	ED	Year	6	24
Unit Conversion Factor	CF	kg/mg	$1x10^{-6}$	$1x10^{-6}$
Inhalation Rate	IR_{inh}	m^3/day	5	20
Body Weight	BW	Kg	29	70
Average Lifetime	AT	Days	EDx365 (non-carcinogens)	EDx365 (non-carcinogens)
			70x365 (carcinogens)	70x365 (carcinogens)
Exposure Time	ET	h/day	24	24
Average Lifetime (n)	AT_n	Hours	EDx365x24 (non-carcinogens)	EDx365x24 (non-carcinogens)
			70x365x24 (carcinogens)	70x365x24 (carcinogens)
Skin Surface Area	SA	cm^3	2800	5700
Skin Adherence Factor	AF	mg/cm^2	0.2	0.07
Dermal Absorption Factor	DAF			
	$DAF_{(Pb)}$		0.1	0.1

$DAF_{(Cd)}$	0.001	0.001
$DAF_{(Cu)}$	0.01	0.01
$DAF_{(Mn)}$	0.01	0.01
$DAF_{(Zn)}$	0.01	0.01
$DAF_{(Al)}$	0.01	0.01
$DAF_{(Other\ metals)}$	0.01	0.01

Table S8. Parameters used to evaluate Carcinogenic and Non-carcinogenic Risk via Inhalation, Dermal and Ingestion Pathway

Parameter	Pb	Cd	Cu	Mn	Zn	Al	Cr	Ni	References
T_f	5	5			1				(Zhang et al., 2021b)
RfD_{oral}	3.50E-03	5.00E-04	4.00E-03	1.40E-01	3.00E-01	1	3.00E-03	1.10E-02	(Han et al., 2021; Mainka, 2021; Sakunkoo et al., 2022; US EPA, 1989; Zhang et al., 2021b)
RfC_{inh}	3.52E-03	1.00E-05	4.02E-03	5.00E-05	3.01E-01	5.00E-03	1.00E-04	2.00E-05	(Han et al., 2021; Mainka, 2021; US EPA, 2009; Zhang et al., 2021b)
GI_{ABS}	1	1.00E-03	1	4.00E-02	1		2.50E-02	4.00E-02	(Han et al., 2021; Mainka, 2021; Sakunkoo et al., 2022; US EPA, 2002; Zhang et al., 2021b)
SF_{oral}	8.50E-03						5.00E-01	9.10E-01	(Mainka, 2021; US EPA, 2004; Zhang et al.,

									2021b)
Inhalation UR (IUR)	1.20E-05	1.80E-03							(Sakunkoo et al., 2022; US EPA, 2009; Zhang et al., 2021b)

Figures

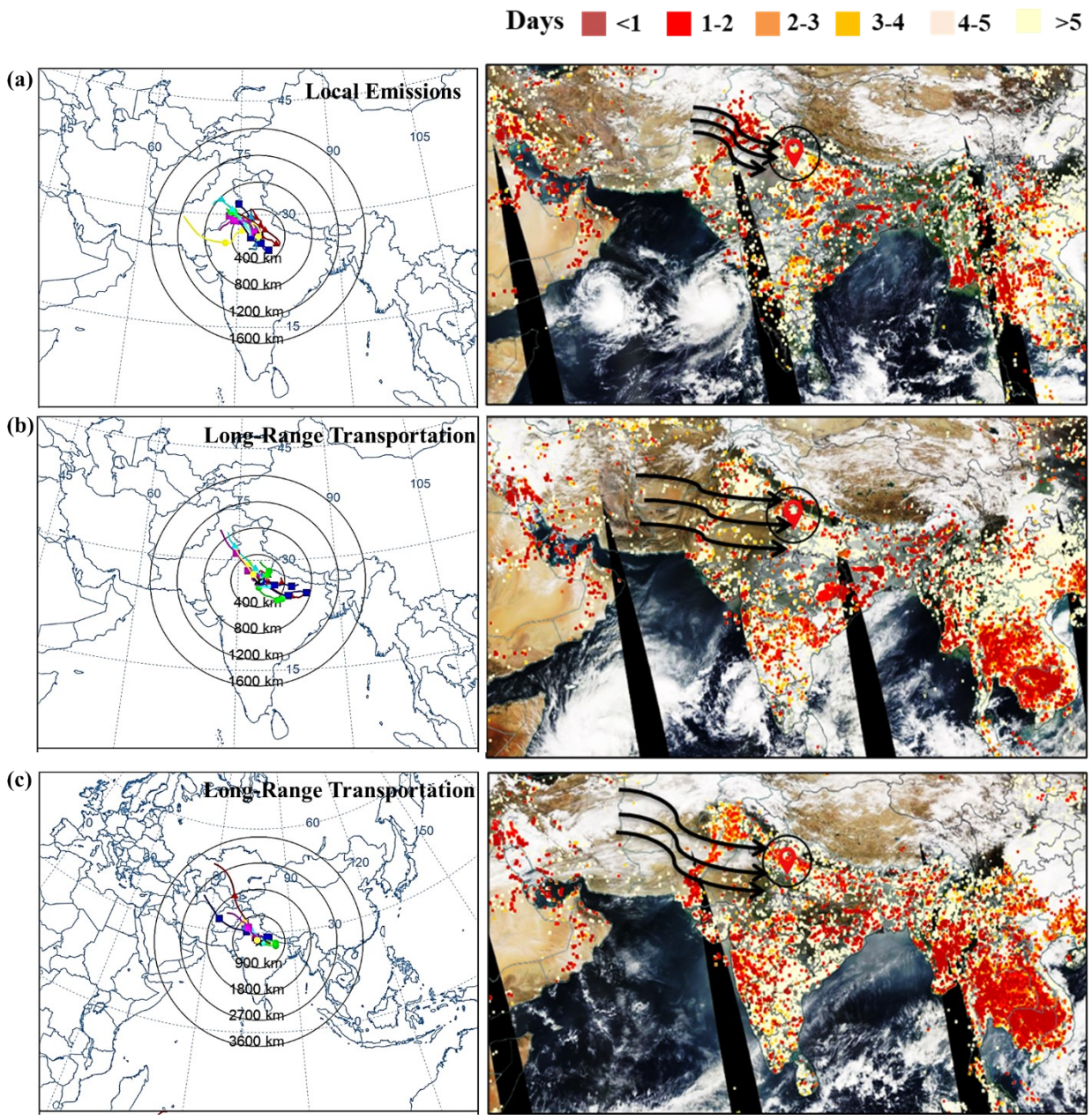


Figure S2. HYSPLIT-derived Air Mass Backward Trajectories and MODIS Aqua True Color Imagery overlaid with MODIS fire counts for (a) November (b) December (c) January 2019 (major wind direction at receptor site are shown by arrows)

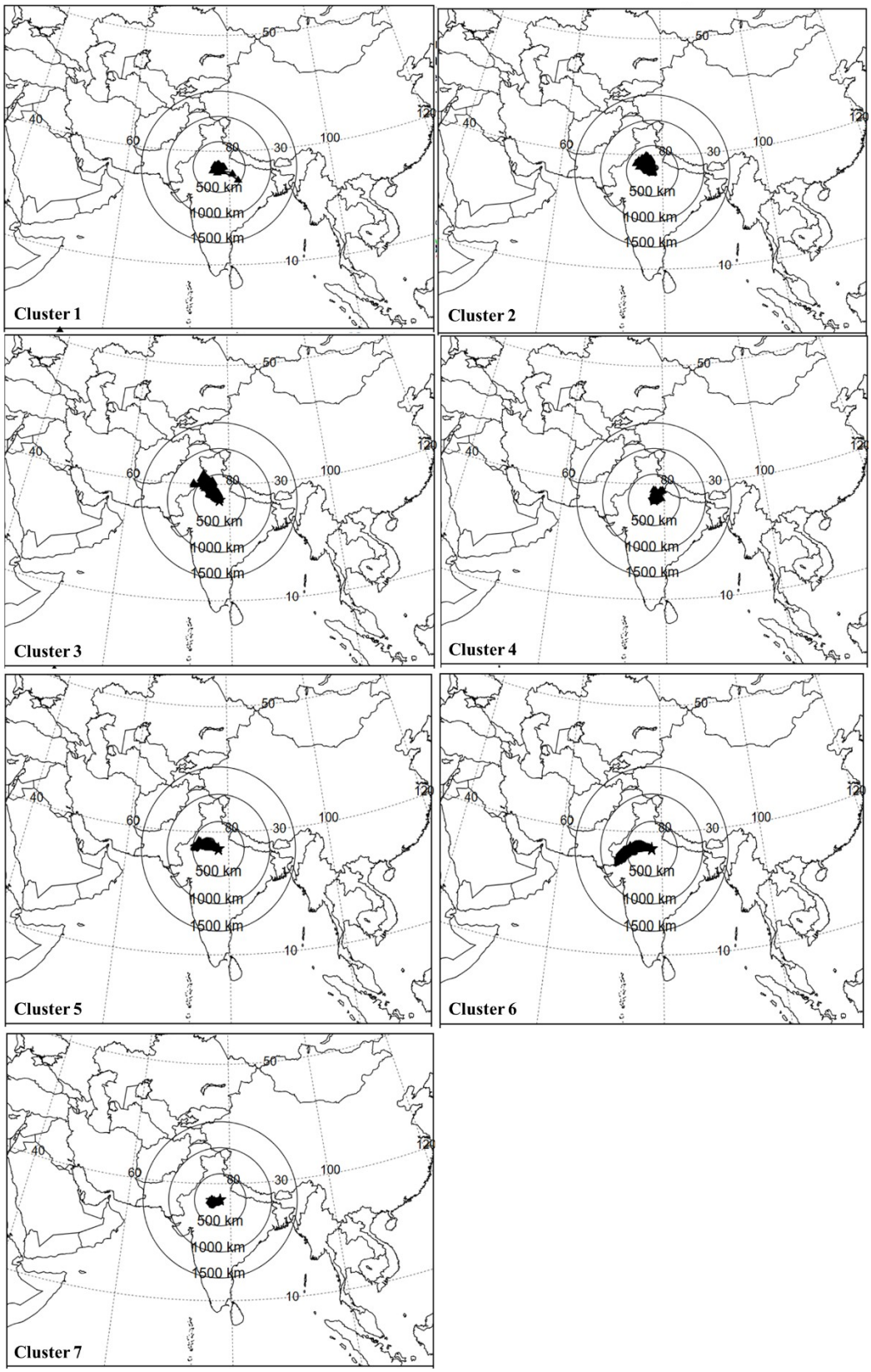


Figure S3. Cluster plots for 24-hour trajectories for the month November 2019

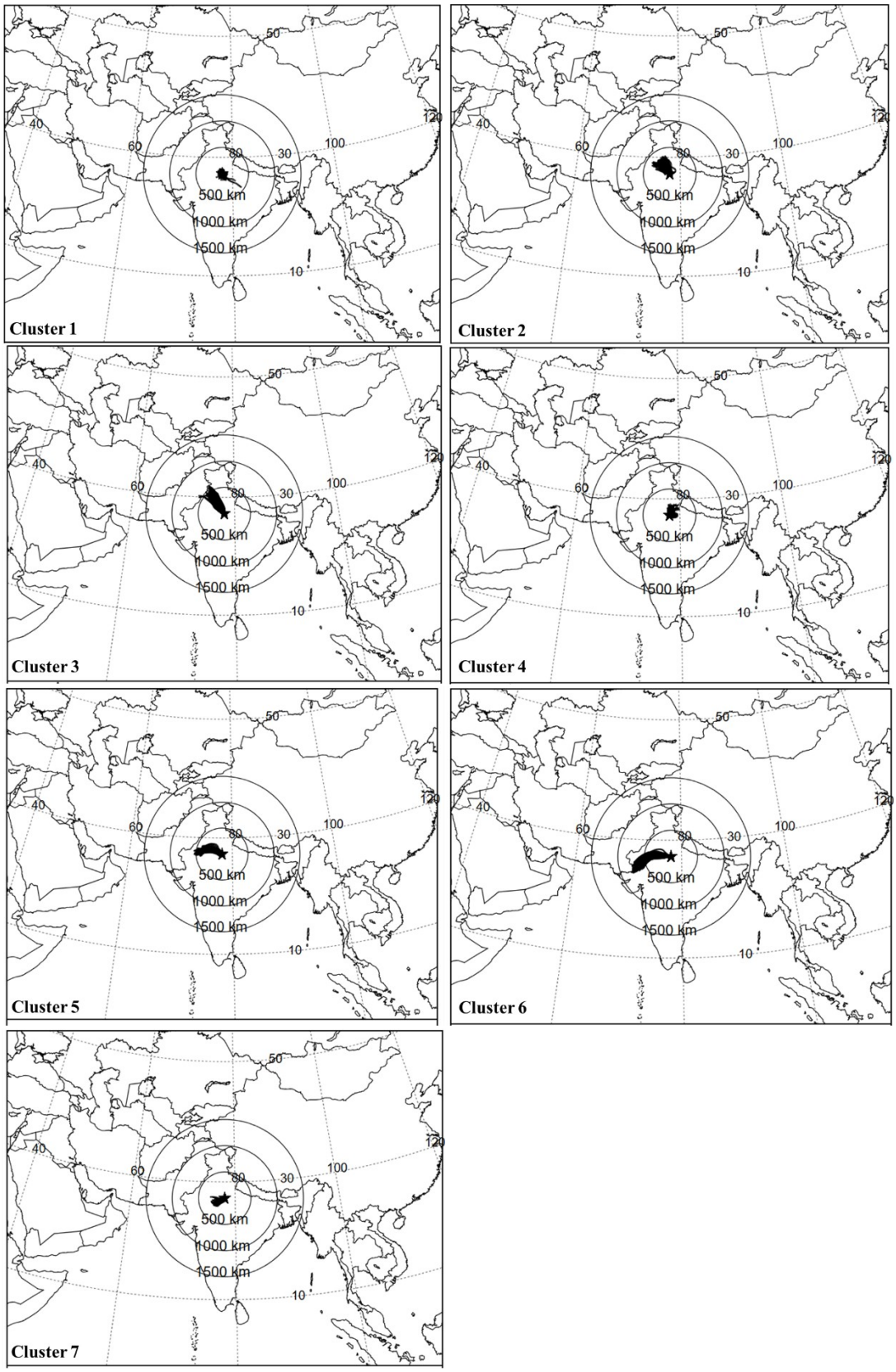


Figure S4. Cluster plots for 24-hour trajectories for the month December 2019

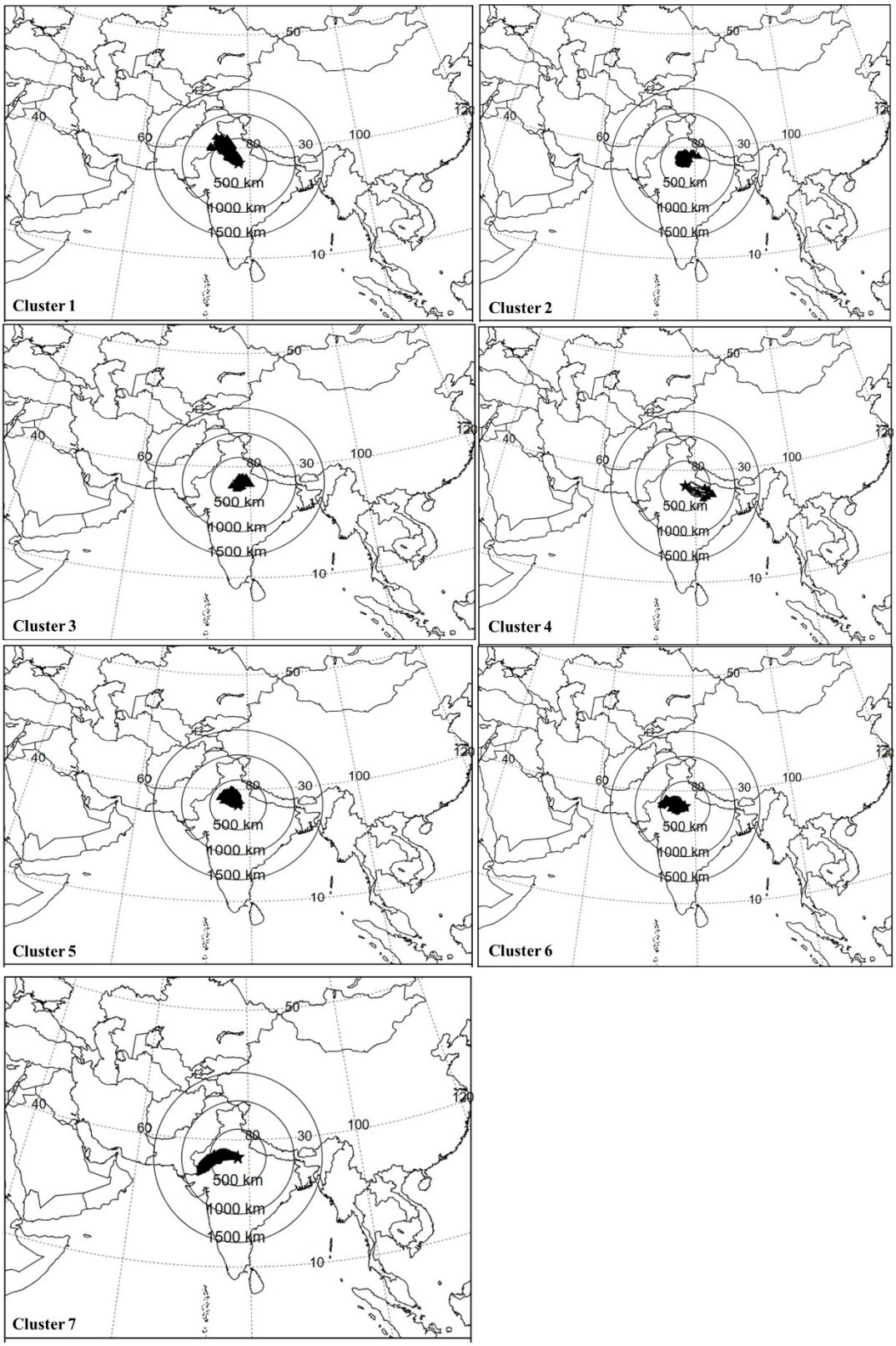


Figure S5. Cluster plots for 24-hour trajectories for the month January 2019

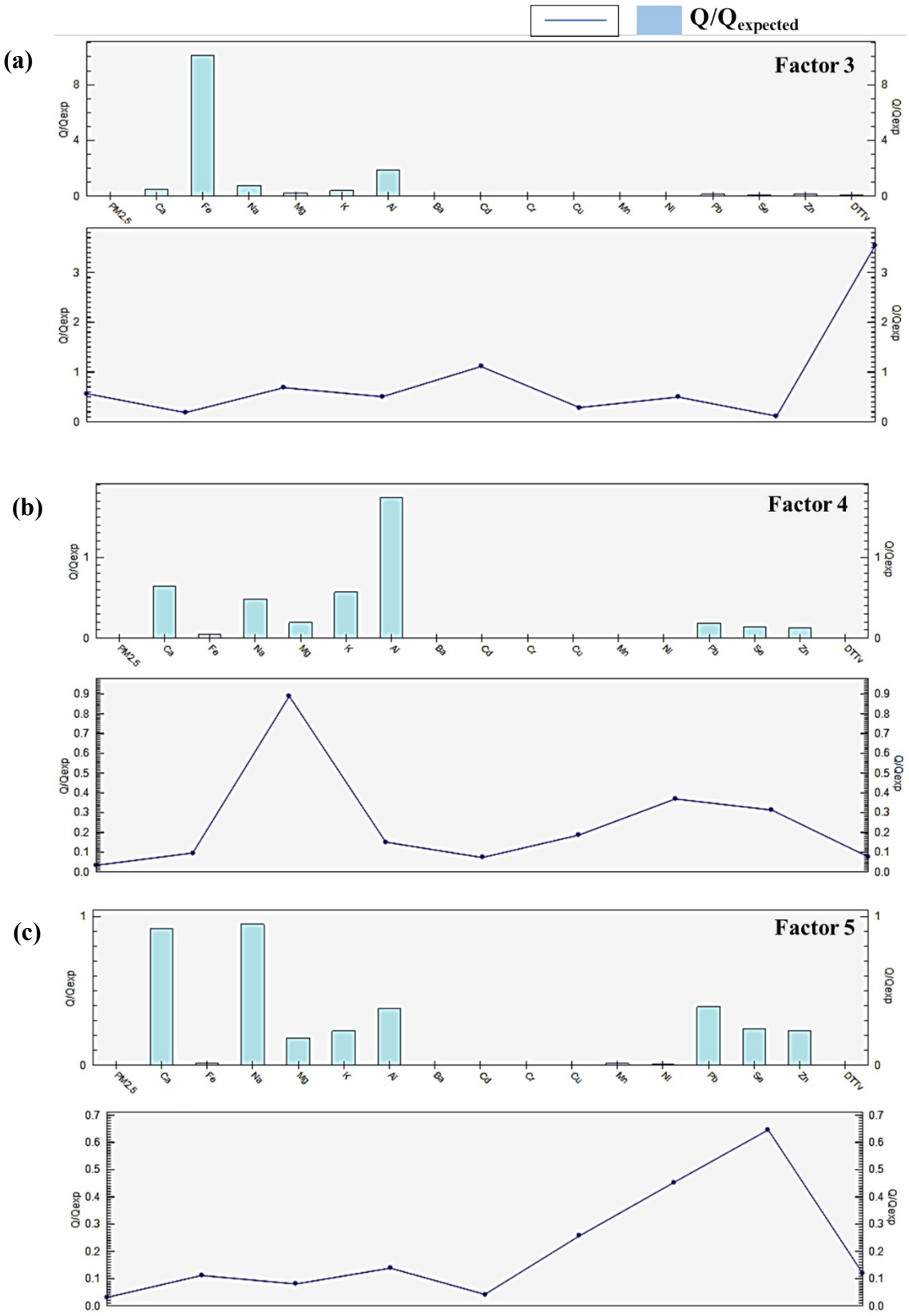


Figure S6. The $Q/Q_{expected}$ graphs for (a) Factor 3 (b) Factor 4 and (c) Factor 5

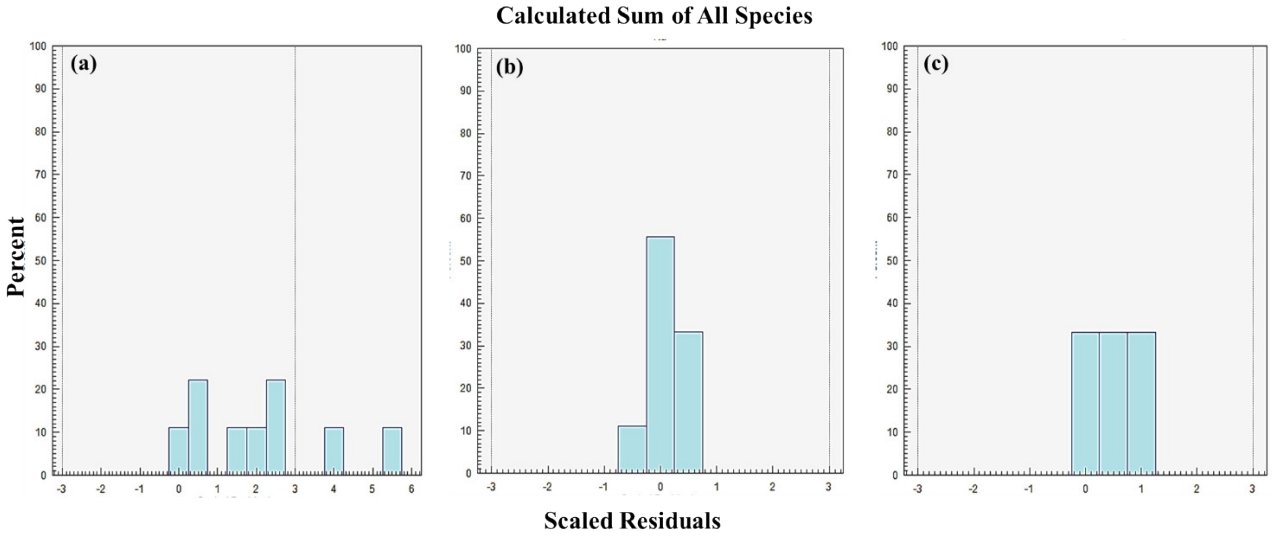


Figure S7. Residual Analysis for (a) Factor 3, (b) Factor 4 and (c) Factor 5

References

EPA PMF Guide, 2008. (2008). EPA Positive Matrix Factorization (PMF) 3.0 Fundamentals & User Guide. *Cir.Nii.Ac.Jp*, 81.

Han, J., Lee, S., Mammadov, Z., Kim, M., Mammadov, G., & Ro, H. M. (2021). Source apportionment and human health risk assessment of trace metals and metalloids in surface soils of the Mugan Plain, the Republic of Azerbaijan. *Environmental Pollution*, *290*. https://doi.org/10.1016/j.envpol.2021.118058

Mainka, A. (2021). Children health risk assessment of metals in total suspended particulate matter (Tsp) and pm1 in kindergartens during winter and spring seasons. *Atmosphere*, *12*(9), 1096. https://doi.org/10.3390/atmos12091096

Murari, V., Singh, N., Ranjan, R., Singh, R. S., & Banerjee, T. (2020). Source apportionment and health risk assessment of airborne particulates over central Indo-Gangetic Plain. *Chemosphere*, *257*. https://doi.org/10.1016/j.chemosphere.2020.127145

Roy, D., Singh, G., & Seo, Y. C. (2019). Carcinogenic and non-carcinogenic risks from PM10-and PM2.5-Bound metals in a critically polluted coal mining area. *Atmospheric Pollution Research*, *10*(6), 1964–1975. https://doi.org/10.1016/j.apr.2019.09.002

Sakunkoo, P., Thonglua, T., Sangkham, S., Jirapornkul, C., Limmongkon, Y., Daduang, S., Tessiri, T., Rayubkul, J., Thongtip, S., Maneenin, N., & Pimonsree, S. (2022). Human health risk assessment of PM2.5-bound heavy metal of anthropogenic sources in the Khon Kaen Province of Northeast Thailand. *Heliyon*, *8*(6), e09572. https://doi.org/10.1016/j.heliyon.2022.e09572

Sharma, S. K., & Mandal, T. K. (2017). Chemical composition of fine mode particulate matter (PM2.5) in an urban area of Delhi, India and its source apportionment. *Urban Climate*, *21*, 106–122. https://doi.org/10.1016/j.uclim.2017.05.009

US EPA. (2004). *Risk Assessment Guidance for Superfund (RAGS): Part E | US EPA*. https://www.epa.gov/risk/risk-assessment-guidance-superfund-rags-part-e

US EPA. (1989). *Risk Assessment: Guidance for Superfund Volume 1 Human Health Evaluation Manual (Part A)*. http://www.epa.gov/swerrims/riskassessment/risk_superfund.html

US EPA. (2002). *Risk Assessment: Guidance for Superfund Volume I: Human Heatlh Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment) Final*. http://www.epa.gov/oswer/riskassessment/

US EPA. (2009). *RISK ASSESSMENT GUIDANCE FOR SUPERFUND VOLUME I: HUMAN HEALTH EVALUATION MANUAL (RAGS) PART F, SUPPLEMENTAL GUIDANCE FOR INHALATION RISK ASSESSMENT*. https://www.epa.gov/risk/risk-assessment-guidance-superfund-rags-part-f

Zhang, X., Eto, Y., & Aikawa, M. (2021a). Risk assessment and management of PM2.5-bound heavy metals in the urban area of Kitakyushu, Japan. *Science of the Total Environment*, *795*. https://doi.org/10.1016/j.scitotenv.2021.148748

Zhang, X., Eto, Y., & Aikawa, M. (2021b). Risk assessment and management of PM2.5-bound heavy metals in the urban area of Kitakyushu, Japan. *Science of the Total Environment*, *795*. https://doi.org/10.1016/j.scitotenv.2021.148748