Rebuttal to the concerns raised by the reviewers for the paper titled “A Framework for Dynamic Hyper-local Source Apportionment using Low-cost Sensors for Real-time Policy Action”

(MS No.: egusphere-2025-5677)

Concerns of Reviewer 4:

The manuscript proposes an appealing and potentially impactful approach: using ML to train low-cost sensors against co-located reference instruments so they can reproduce reference-grade source apportionment. The concept is sound and worth pursuing, but in its current form the manuscript has several gaps in the input data, the ML validation and evaluation, and the reliability of the PMF, as well as internal inconsistencies. These must be addressed through major revisions before it is ready for publication.

Concern #1: Unreliable input predictors are retained without justification.

Not all of the LCAQ features used as model inputs are adequately calibrated, despite the text stating that all are used "after calibration" (l.211). Most clearly, there is no calibration equation for VOC: Eqn (1) provides calibration models for five species only (CO, NO_2, O_3, SO_2, $PM_{2.5}$), and VOC has neither a reference instrument nor a calibration model. Please clarify whether VOC is fed raw, state what physical quantity and units it represents, and justify retaining it.

In addition, the manuscript reports that ambient SO_2 is below the sensor detection limit at both sites (R^2 = 0.03 at Site-B, R^2 = -0.17 at Site-C) and that NO_2 calibrates poorly (R^2 = 0.22 at Site-B). Inputs with this little signal contribute noise rather than information, yet both are retained in the regression and assigned non-negligible coefficient magnitudes (Fig. 16). Please report a sensitivity analysis of model performance with and without SO_2, NO_2, and VOC, and state precisely how sub-detection-limit values were handled in the feature pipeline. If performance is comparable without them, retaining noisy predictors should be avoided; if performance drops, that dependence on unreliable inputs is itself a concern for field deployment.

Response: We sincerely thank the reviewer for this thoughtful and constructive comment regarding the selection of predictor variables used in the machine learning framework.

The reviewer correctly notes that the original manuscript did not clearly describe the treatment of the VOC measurements and did not sufficiently justify the inclusion of **SO_2** and **NO_2**, despite their comparatively weak calibration performance. In response, we have substantially revised both the predictor analysis and the accompanying discussion in the revised manuscript.

First, we clarify that the **VOC variable is not an absolute concentration measurement**. The deployed LCAQ sensor reports VOC as the **ratio of the instantaneous VOC concentration to the average VOC concentration over the preceding 24-hour period**. Consequently, the

VOC predictor is a **dimensionless index** rather than a calibrated concentration, and this has now been explicitly stated in the revised manuscript.

Second, following the reviewer's suggestion, we performed a comprehensive **feature ablation study** using the best-performing machine learning model (**Extra Trees**) to quantify the influence of the potentially unreliable predictors (**SO_2**, **NO_2**, and **VOC**) on source apportionment performance.

For **elemental source apportionment**, the feature ablation study showed that progressively removing these predictors resulted in comparable or slightly improved performance. Based on these results, **SO_2**, **NO_2**, and **VOC** have been excluded from the final elemental SA framework. The revised elemental model therefore uses only the more reliable predictors:

- CO
- O_3
- $PM_{2.5}$
- BC
- Temperature
- Relative Humidity
- Cyclical time features

For **organic source apportionment**, we additionally revised the target representation by replacing the original five-factor PMF solution with a chemically more interpretable **three-factor representation (BBOA, HOA and OOA)**. Feature ablation and feature importance analyses were subsequently performed for this revised framework. These analyses showed that although **SO_2**, **NO_2**, and **VOC** are not among the dominant predictors, they provide complementary information for predicting the organic aerosol composition. Consequently, these variables have been retained in the revised organic SA model.

To further improve transparency, the revised manuscript now includes **feature importance analysis** for both elemental and organic SA using the Extra Trees model. These results quantitatively demonstrate the relative contribution of each predictor and provide additional justification for the final predictor selection.

We believe these additional analyses directly address the reviewer's concerns regarding predictor reliability and provide a considerably more rigorous basis for the revised machine learning framework.

We sincerely thank the reviewer for this valuable suggestion, which has substantially strengthened both the predictor selection methodology and the interpretation of the machine learning models.

Changes to be made in the revised manuscript:

- Clarify that **VOC** is a **dimensionless 24-hour normalized index** rather than a calibrated concentration.
- Include feature ablation studies evaluating the influence of **SO_2**, **NO_2**, and **VOC**.
- Remove **SO_2**, **NO_2**, and **VOC** from the final **elemental** SA predictor set.
- Include **feature importance analysis** for both elemental and organic SA models.
- Expand the discussion explaining the rationale behind the final predictor selection.

Concern #2: PMF factor solutions require further validation and justification.

The authors justify the study design on the premise that the two sites are strongly contrasting: Site-B a traffic site, Site-C a background site (l.144-148) and they train separate per-site regression models on that basis (l.520). Yet, the PMF was run on a single pooled Site-B/Site-C matrix (l.290, l.561), which constrains each factor to a single profile across the whole run; only the time-resolved contributions differ by site. Pooling therefore cannot reveal whether a factor represents a different source at the two sites, only how active a shared profile is at each. Considering the policy relevance, even in the title, this ambiguity is consequential. Take the "Ferrous-Smelting" factor for example. The manuscript attributes it to "industrial sources or non-exhaust traffic emissions, such as brake or tire wear" (l.587-589), so the same Fe-rich signature could represent smelting at one site but brake/tire wear at the other. Examining the diurnal variability of each factor and its correlation with co-measured species would further help establish factor identity. These analyses, together with rotational and statistical diagnostics (Q/Q_exp, bootstrapping, scaled residuals), would demonstrate the robustness of the PMF solutions.

Response: We sincerely thank the reviewer for the encouraging assessment of the novelty and potential impact of the proposed framework, as well as for the constructive suggestions that have helped us significantly improve the methodological rigor and presentation of the revised manuscript.

We significantly increased the dataset from about 20 to 63 measurement days by adding Cycle-II datasets from CSIR-CIMAP and BBAU (collected from February to March 2024) and Kanpur measurements (from December 2025). The PMF analyses were repeated using this expanded dataset. The revised HR-ToF-AMS analysis resolved five factors: HOA, BBOA-I, BBOA-II, SVOOA, and LVOOA, while the Xact 625i analysis resolved seven factors: S-rich, Cl-rich, Dust, Non-Exhaust, SFC-I, SFC-II, and Pb-rich (Fig. 4.2, 4.3).

The previously identified, highly source-specific **"Fe-smelting"** factor was not reproduced in the expanded analysis. Including additional sites and campaign periods increased the number of observations and introduced greater spatial, temporal, and source variability, thereby reducing the influence of short-duration or site-specific Fe-rich episodes on the PMF solution. Instead of forming a distinct factor, Fe was distributed mainly among the Dust and Non-Exhaust factors together with other crustal and traffic-related metals. This indicates that the earlier "Fe-smelting" assignment was insufficiently robust as a unique source and likely reflected mixed Fe-rich emissions in the shorter dataset.

A detailed procedure of optimum solution selection of the resolved factors and stability assessments, are provided in our related mobile-lab studies from the same sites by Sethi et al. (2026, under review) and Lakra et al. (2026). Therefore, only a concise summary of the factor characteristics, Q/Qexp, external correlation and bootstrap assessment has been included in the revised supplementary information of the manuscript, with appropriate references to these studies.

PMF input matrices and selection of the optimum solutions

For the **HR-ToF-AMS** analysis, the unit-mass-resolution (UMR) organic mass spectra were processed using SQUIRREL after removing spectra affected by instrument malfunction, calibration issues, or other invalid runs. The remaining data were averaged to a 30-minute resolution and used to construct the concentration and corresponding error matrices. The final AMS input matrix contained 2142 observations and 300 (m/z) variables. The large number of valid observations relative to the number of variables provided sufficient temporal variability and degrees of freedom for resolving the five-factor solution.

For the **Xact-625i** analysis, the 30-minute elemental concentration and uncertainty datasets were aligned across all included deployments. Elements with a detection limit below MDL were excluded. The final Xact input matrix contained 2424 observations and 21 elemental variables. The expanded multi-site dataset included substantially more observations than the original analysis and captured variations across sites, seasons, meteorological conditions, and source environments.

Solutions with different numbers of factors were examined for both datasets. The 5-factor AMS and 7-factor Xact solutions were selected based on changes in Q/Qexp, factor-profile interpretability, temporal behaviour, external-tracer relationships, and factor stability (Fig. R4.1). For the AMS analysis, the 5-factor solution was selected because we observed the maximum Q/Qexp relative change from 4 to 5 factors, whereas the 7-factor Xact solution shows the maximum Q/Qexp relative change from 6 to 7 factors. Adding further factors primarily splits physically meaningful factors, whereas solutions with fewer factors showed mixing of chemically distinct source profiles (Fig. R4.1).

Bootstrap assessment

To evaluate the stability of the PMF solutions, 1000 bootstrap resampling runs were performed in SoFi using criterion-based selection. The final five-factor AMS and seven-factor Xact solutions, with all factors unconstrained, were used as the base cases. Bootstrapped factors were mapped using an active Pearson correlation criterion ($R > 0.9$), and only valid, unmixed runs were retained. Overall, 857 runs (85.7%) for AMS and 923 runs (92.3%) for Xact were successfully mapped, indicating good factor stability and reproducibility.

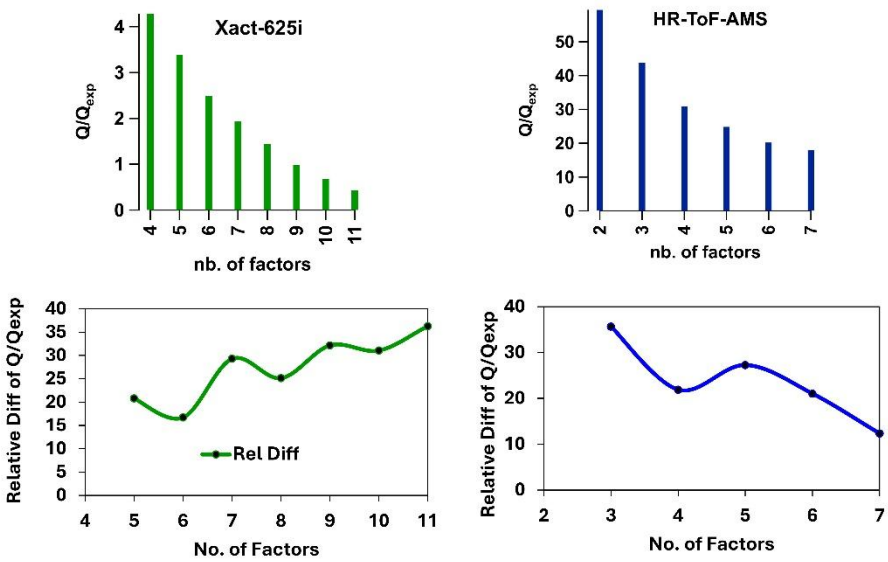


Figure R4.1: Q/Qexp and its relative difference (%) plot with respect to the number of factors for optimum PMF solution selection criteria of elemental SA (Xact-625i) and Organics SA (HR-ToF-AMS)

Brief interpretation of the HR-ToF-AMS factors:

HOA: As a primary organic aerosol (POA) factor, HOA was characterized by hydrocarbon fragments at m/z 41, 43, 55, 57, 69, and 71, together with the aromatic fragment at $m/z131$ and low contributions from oxygenated ions (Fig. R4.2). These fragments are associated with fresh traffic and fossil-fuel combustion emissions (Tobler et al., 2020). It's morning, and evening peaks were consistent with the traffic peak hour and the accumulation of primary emissions under a shallow boundary layer. HOA contributed 13% of the total organic aerosol and showed a good correlation with black carbon (BC) (R = 0.65), further supporting its association with vehicle exhaust (Shukla et al., 2021a) (Figs. R4.2, R4.4). HOA was highest at BBAU during Cycle-II, followed by Kanpur, consistent with its proximity to major traffic corridors and meteorological conditions that favoured pollutant accumulation. The high HOA in Kanpur was attributed to nearby commercial activity, traffic intersections, and frequent vehicular movement.

BBOA-I: BBOA-I was also identified as a primary organic aerosol factor with low contribution from m/z44 (oxygenated ion) and contributed 12% of the total organic aerosol. Its mass spectrum was characterized by the biomass-burning marker ions at m/z 60 and 73, together with hydrocarbon and aromatic fragments (m/z 178, 202), indicating relatively fresh biomass and solid fuel combustion emissions (Goetz et al., 2022) (Fig. R4.2). Its diurnal profile showed high concentrations during the late evening, nighttime, and early morning, with a pronounced increase after approximately 19:00 and minimum concentrations during the afternoon. This pattern is consistent with nighttime biomass burning or domestic solid-fuel burning and the accumulation of primary emissions under reduced boundary-layer mixing. BBOA-I was highest at Kanpur, likely due to wintertime biomass and solid-fuel burning for domestic heating, followed by BBAU during Cycle I, where post-monsoon agricultural-residue burning may have contributed (Fig. R4.2). The lowest concentrations were observed at CSIR-

CIMAP, reflecting its background nature and limited nearby combustion sources. Its correlations with Chl (R = 0.57) and BC (R = 0.67) further supported its combustion-related origin (Lalchandani et al., 2022) (Fig. R4.4).

BBOA-II: BBOA-II contributed 14% of the total organic aerosol and was identified as a more oxidized biomass-burning factor. Its mass spectrum retained the characteristic biomass-burning fragments at m/z 60 and 73 but showed a stronger contribution at m/z 44, indicating atmospheric ageing and secondary organic aerosol (SOA) formation (Lalchandani et al., 2022). Its diurnal profile exhibited a morning peak around 09:00–10:00, following the earlier BBOA-I increment, suggesting oxidation of freshly emitted biomass-burning aerosol (Fig. R4.2). An increase during nighttime may be associated with oxidation by NO_3 radicals under stable atmospheric conditions (Lakra et al., 2024). BBOA-II was highest at Kanpur during winter, followed by BBAU during Cycle I and CSIR-CIMAP during Cycle II, reflecting differences in biomass-burning emissions and atmospheric processing across the sites.

SVOOA: SVOOA contributed 20% of the total organic aerosol and was characterized by oxygenated fragments at m/z 44, together with smaller contributions from less-oxidized hydrocarbon ions (Manchanda et al., 2021) (Fig. R4.2). This composition indicates partially oxidized and comparatively semi-volatile organic aerosol. Its concentration increased during the late morning, consistent with photochemical oxidation of locally emitted precursors under sunlight (Li et al., 2013). SVOOA was highest in Kanpur, possibly because stronger local wintertime emissions provided abundant precursors that underwent initial atmospheric ageing. It shows good correlation with nitrate (R=0.51), further suggesting a secondary nature (Fig. R4.4).

LVOOA: LVOOA was the dominant AMS factor, contributing 41% of the total organic aerosol. Its mass spectrum showed the strongest oxygenated signature, particularly a high relative abundance of $m/z44$, consistent with highly aged and low-volatility secondary organic aerosol (Li et al., 2013) (Fig. R4.2). Its concentration increased during the afternoon, following the rise in SVOOA, indicating continued photochemical ageing and transformation of semi-volatile into more oxidized, lower-volatility species. LVOOA was the largest contributor at each site and showed a comparatively higher concentration at CSIR-CIMAP than other factors, supporting the influence of regionally transported and aged aerosol at this background location. It shows a good correlation with sulfate (R=0.56), supporting its secondary nature (Guo et al., 2020) (Fig. R4.4).

Brief interpretation of the Xact 625i factors

S-rich: The S-rich factor contributed 32% of the total elemental mass and was dominated by S (85.7%), followed by K (12.2%) (Shukla et al., 2023) (Fig. R4.3). It shows a relatively flat diurnal profile, indicating a secondary source rather than a direct local emission. The moderate correlation with AMS sulfate (R = 0.51) further supports this. Sulfur released mainly as SO_2 from coal and other combustion activities can undergo gas-phase and aqueous-phase oxidation to form particulate sulfate. The associations with K (R = 0.53) likely reflect the combustion origin of the precursor emissions (Sharma and Mandal, 2023) (Fig. R4.4). S-rich was highest at Kanpur, followed by the Lucknow Cycle-II sites, and was lowest at CSIR-CIMAP during Cycle I. The higher Kanpur concentration likely reflects stronger combustion-related SO_2 emissions from the nearby Panki power plant and subsequent sulfate formation, whereas the lower concentration at CSIR-CIMAP is consistent with its background nature.

Cl-rich: The Cl-rich factor contributed 23% of the total elemental mass and was dominated by Cl (80.6%), followed by S (12.1%) and K (4.3%) (Fig. R4.3) (Lakra et al., 2026). It showed a pronounced morning peak, a strong daytime decline, and a slight evening increase, consistent with fresh combustion emissions followed by dilution and volatilization of particulate chloride. Its strong correlation with AMS chloride ($R = 0.82$) and K ($R = 0.70$), together with moderate correlations with Rb, Br, As, and Cu, supports contributions from biomass, solid-fuel, and waste-burning emissions (Shukla et al., 2021a) (Fig. R4.4). Cl-rich was highest at Kanpur, followed by the Lucknow Cycle-II sites, and lowest at CSIR-CIMAP during Cycle I. The elevated Kanpur concentration likely reflects stronger wintertime solid-fuel and waste-burning activities.

Dust: The Dust factor was dominated by Si (48.6%), Ca (24.2%), and Fe (19.9%), with smaller contributions from K and Ti and contributed 18% of the total elemental mass (Rai et al., 2020) (Fig. R4.3). Its gradual daytime increase and evening peak indicate dust resuspension associated with traffic. Strong correlations with Fe (R = 0.87), Mn (R = 0.63), and Cr (R = 0.51) suggest that the factor includes both crustal material and traffic-contaminated road dust (Lakra et al., 2026) (Fig. R4.4). Dust was highest at Kanpur, followed by BBAU during Cycle II, and lowest at CSIR-CIMAP during Cycle I. The higher concentrations at Kanpur and BBAU likely reflect higher traffic-induced resuspension.

SFC-I: The SFC-I factor was dominated by K (90.8%), followed by S (6.0%) (Fig. R2.4). Its morning and evening enhancements, together with strong correlations with K (R = 0.96), Rb (R = 0.81), BC (R = 0.54), and AMS Chl (R = 0.54), support emissions from biomass and other solid-fuel combustion (Rai et al., 2021) (Figs. R4.3, R4.4). It contributed 15% of the total elemental mass and was highest at Kanpur, followed by BBAU, likely due to wintertime domestic heating and local combustion.

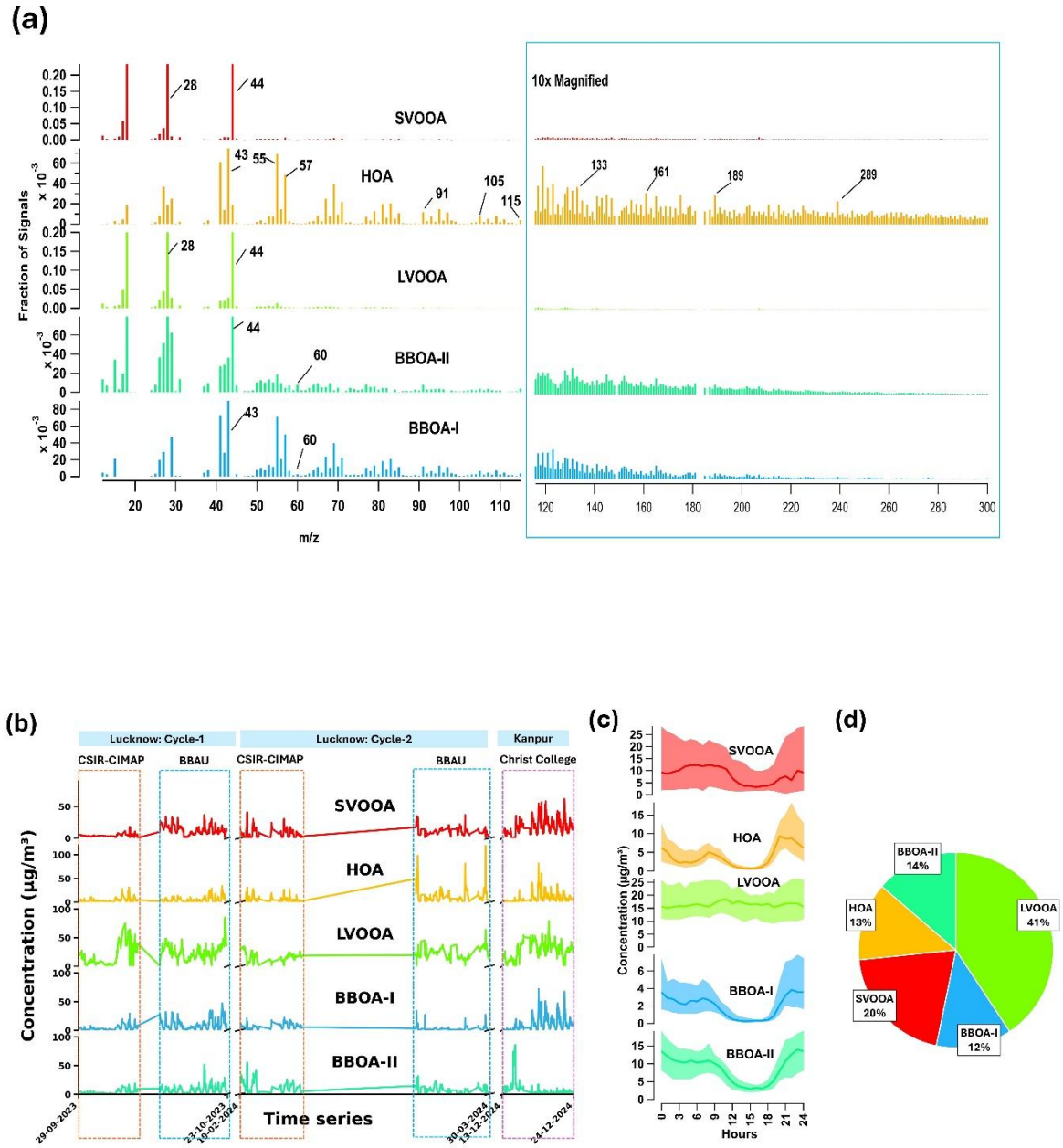
SFC-II: The SFC-II factor contributed the lowest with 3% and was dominated by Zn (72.5%), followed by S (17.1%), Fe (5.6%), and Si (4.4%) (Shukla et al., 2021b) (Fig. R4.3). It showed a modest morning enhancement and weaker associations with Fe (R = 0.39) and Mn (R = 0.31), indicating mixed Zn-rich emissions from solid fuels, waste burning, traffic, or industry (Fig. R4.4). Episodic peaks were most pronounced at BBAU, particularly during Cycle II, followed by Kanpur, whereas CSIR-CIMAP showed the lowest concentrations.

Non-Exhaust: The Non-Exhaust factor was dominated by Fe (43.9%), Si (28.6%), Mn (10.6%), and Ca (5.9%), with smaller contributions from Cu and Zn (Fig. R4.3). It contributed 6% of the total elemental mass. The diurnal variations show morning and evening increments, which were consistent with traffic activity and road-dust resuspension. Strong correlations with Fe (R=0.82), Mn (R=0.80), Cr (R=0.49), and BC (Lakra et al., 2026) (R=0.45) support contributions from brake wear, mechanical abrasion, and traffic-related resuspended dust (Fig. R4.4). It was highest at Kanpur, followed by BBAU during Cycle II, and lowest at CSIR-CIMAP during Cycle I, reflecting differences in traffic intensity and local road activity.

Pb-rich: The Pb-rich factor contributed 3% of the total elemental mass and was dominated by S (44.9%) and Pb (27.1%), followed by K (12.9%) and Cl (6.2%) (Fig. R4.3). Its strong correlations with As (R=0.82) and Se (R=0.65) indicate coal-combustion emissions (Li et al., 2012) (Fig. R4.4). The factor showed weak diurnal variation with occasional short-duration peaks, suggesting intermittent local sources. Pb-rich concentrations were highest at Kanpur due to the combustion of coal during the winter for heating purposes. It also reflects a stronger

influence from commercial, combustion, waste-burning, or metal-related activities at the more urban sites.

Overall, the revised factors are supported by their profiles, behavior, mass contributions, and relationships with external tracers. We avoided assigning factors to specific industries when marker species can come from multiple sources. The revised manuscript therefore describes them as broad PMF-resolved chemical or source-related factors rather than unique emission sources. Detailed factor profiles, diurnal variations, external correlations, and bootstrap results have been added to the Figures of the revised Supplementary Information.



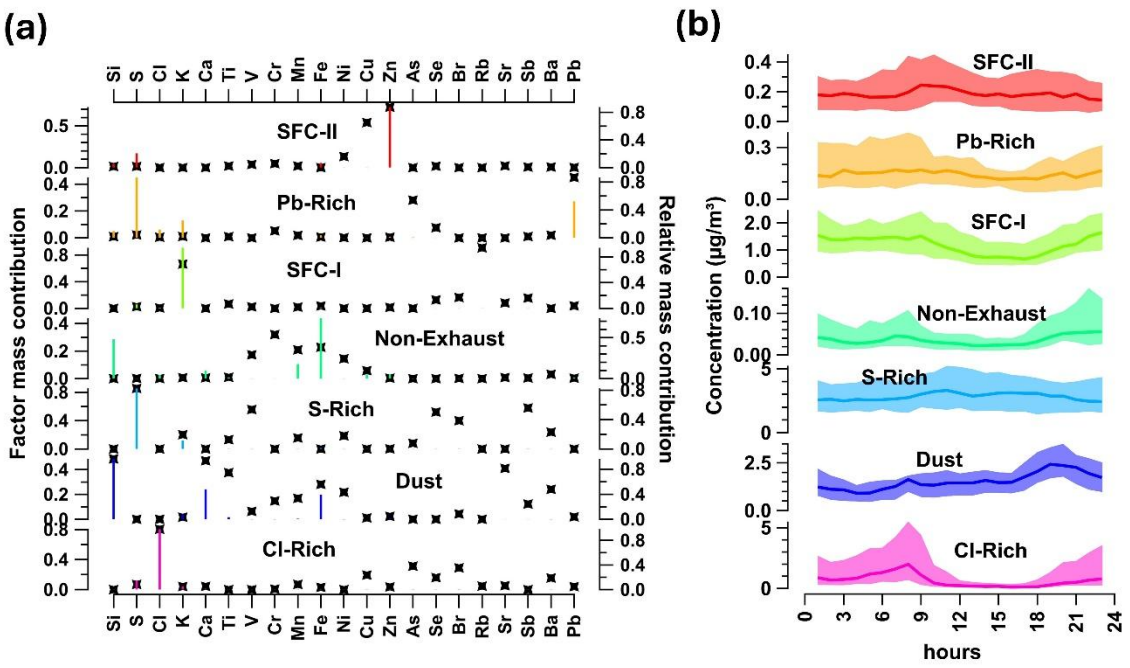


Figure R2.3. (a) m/z spectra for the different organic factors during the measurement period, (b) Time series plots (c) Diurnal (d) Fractional contribution for the different PMF resolved factors of organic aerosol.

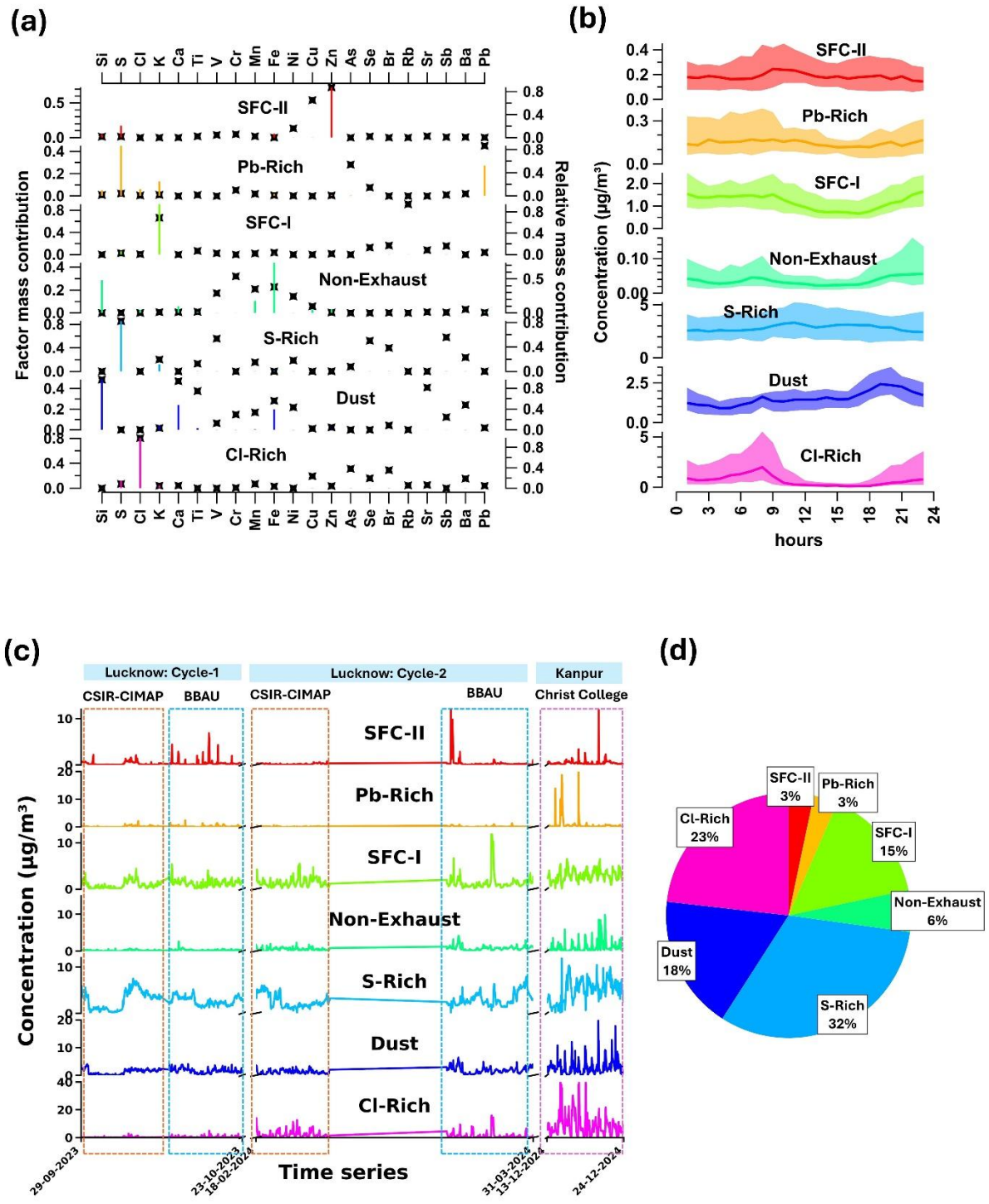


Figure R2.4. (a) Spectra profile for the different elemental factors during the measurement period, (b) Diurnal, (c) Time series plots, (d) Fractional contribution for the different PMF resolved factors of Trace elements.

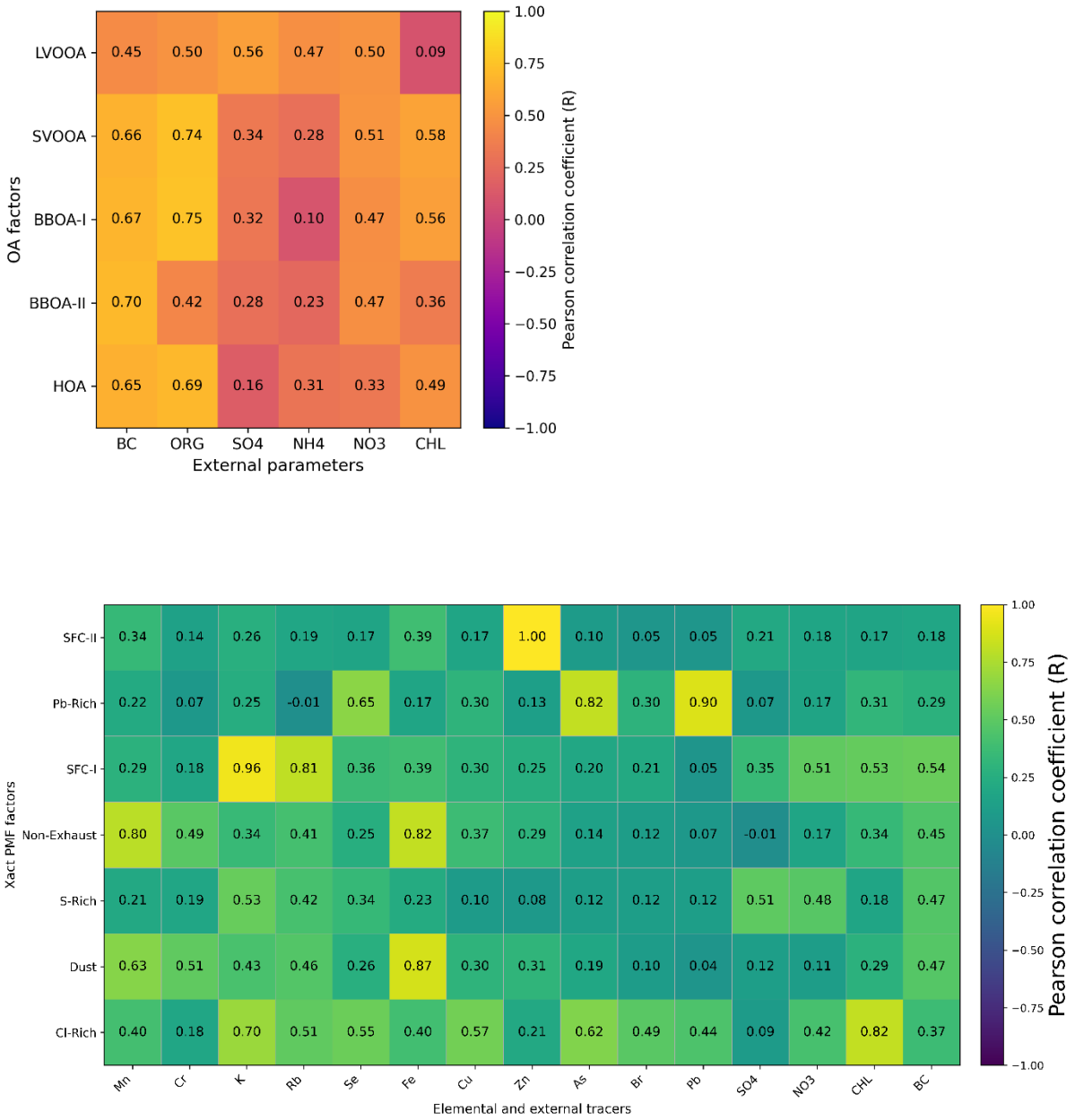


Figure R2.5. Heatmap between (a) PMF resolved OA sources and external parameters, and (b) PMF resolved elemental sources and external parameters.

Changes to be made in the revised manuscript:

- All the details of RTSA have been revised in the appendix of the manuscript, including new and modified figures.

Concern #3: The model's high metric scores largely reflect source dominance, not predictive skill.

The SA is dominated by a single source in every case (LVOOA ~49% at Site-B and ~58% at Site-C; S-rich ~35-52% in the elemental case). A perfect NDCG for identifying the single top source at Site-C simply reflects that one source is almost always the largest. The component-wise MAE is likewise misleading when read in absolute terms: recast relative to each source's mean abundance, the minor factors show no skill with Pb-rich at Site-B has MAE 2.12% against a mean of 1.70%, and Cl-rich at Site-C has MAE 2.41% against 1.30%, i.e. the error equals or exceeds the quantity being predicted. The apparent success of the model therefore rests almost entirely on tracking the one dominant source, while the figures (pie charts and 0-100% time series) make the minor-source failures visually negligible. The model should be benchmarked against naïve baselines (e.g., always predicting the training-mean composition, or carrying the previous window forward). Only the improvement over such baselines shows whether the model has learned source dynamics rather than reproducing the average. This concern compounds the uncertainty issue below: the reported scores are both inflated (source dominance) and imprecise (small test set), so at present neither their magnitude nor the differences between models can be reliably interpreted.

Response: We sincerely thank the reviewer for this thoughtful and insightful comment regarding the interpretation of the reported evaluation metrics.

We agree that, in datasets where one or two source categories dominate the overall composition, ranking-based metrics such as **Normalized Discounted Cumulative Gain (NDCG)** should not be interpreted in isolation. Accordingly, in the revised manuscript we have expanded the performance evaluation to include the complementary metrics **Mean Absolute Error (MAE)** and **Spearman Rank Order Correlation Coefficient (SROCC)** for all machine learning models.

While NDCG primarily evaluates the correctness of the ranked source contributions, **SROCC provides an independent assessment of the monotonic agreement between the predicted and reference source contributions**. A reasonably high SROCC therefore indicates that the proposed model can preserve the relative temporal ordering of source contributions with respect to the reference PMF-derived ground truth, rather than simply reproducing a fixed average source composition. We have accordingly revised the discussion throughout the manuscript to interpret the model performance jointly using **MAE, SROCC and NDCG**, instead of relying on any single evaluation metric.

In response to the reviewer's suggestion, we will further strengthen the revised manuscript by including a comparison with a **naïve baseline predictor** that assigns the average source composition observed in the training dataset to every test sample. The performance of this baseline will be compared with the proposed machine learning framework using the same evaluation metrics. This additional comparison will provide a direct quantitative assessment of the extent to which the proposed models learn meaningful variations in source contributions beyond the average source composition.

Furthermore, the original five-factor organic source apportionment framework has been replaced by a chemically more interpretable **three-factor representation (BBOA, HOA and**

OOA). This revision reduces ambiguity among highly correlated organic sub-factors and provides a more stable and interpretable prediction problem. The revised manuscript also includes feature importance and feature ablation analyses, thereby providing additional evidence regarding model behaviour beyond the reported performance metrics.

We thank the reviewer for this valuable suggestion, which has led to a more balanced interpretation of the evaluation metrics and a more rigorous assessment of the proposed framework.

Changes to be made in the revised manuscript:

- Present the model performance using the complementary metrics **MAE**, **SROCC**, and **NDCG**, together with a balanced discussion of their respective interpretations.
- Include a comparison with a **naïve baseline predictor** based on the average source composition observed in the training data.
- Replace the original five-factor organic SA framework with the revised **three-factor (BBOA–HOA–OOA)** representation.
- Include feature importance and feature ablation analyses to provide additional insight into the predictive behaviour of the proposed models.

Concern #4: The reported performance of the selected model is not quantified with uncertainty.

With N_test ~ 46-75, the headline metrics for the chosen linear-regression model (and the abstract's summary claims of MAE < 5%, SROCC > 0.75, NDCG > 0.85) are point estimates from a few dozen test points, with wide and unreported uncertainty. Please report these metrics with bootstrap confidence intervals (resampling test timestamps) so the reliability of the reported performance can be assessed.

Response: We sincerely thank the reviewer for this valuable suggestion regarding the statistical interpretation of the reported performance metrics.

We agree that reporting only point estimates of **MAE**, **SROCC**, and **NDCG** does not fully characterize the uncertainty associated with model performance, particularly when the evaluation is performed on relatively small test datasets.

In the revised manuscript, the original short-duration site-specific experiments have been replaced by a substantially more comprehensive evaluation framework comprising independent **cross-city, cross-site, and cross-season transfer-learning experiments**. In particular, the principal evaluation now employs the complete **Kanpur** dataset as an independent external test set, resulting in a considerably larger number of test observations than those available in the originally submitted manuscript. Consequently, the revised evaluation provides a more reliable estimate of model performance than the original two-day test-period experiments.

Following the reviewer's suggestion, we will further strengthen the statistical evaluation by estimating **95% bootstrap confidence intervals** for the principal performance metrics (**MAE**, **SROCC**, and **NDCG**) through repeated resampling of the independent test dataset. These

confidence intervals will accompany the corresponding point estimates in the revised manuscript and will provide a quantitative measure of the uncertainty associated with the reported model performance.

We believe that the combination of substantially larger independent test datasets together with bootstrap confidence intervals will provide a more rigorous and statistically informative assessment of the proposed framework.

We sincerely thank the reviewer for this constructive suggestion, which has improved the statistical presentation of the revised results.

Changes to be made in the revised manuscript:

- Replace the original short-duration experiments with the revised transfer-learning evaluation using substantially larger independent test datasets.
- Report **95% bootstrap confidence intervals** for **MAE**, **SROCC**, and **NDCG**.
- Expand the discussion to include the statistical uncertainty associated with the reported performance metrics.

Concern #5: Technical corrections:

- *S3, Fig. 5: correct BC units and values. VOC: units*
- *S3, l.302 vs l.565: reconcile the elemental PMF factor list. Remove or restore "Fireworks".*
- *S4, Eqn (5): add the square root to SAE.*
- *S9: While the full 22-element PMF input list is evident in Fig. A3, state it explicitly in the text as well.*

Response: All the above recommended changes will be included in the revised manuscript.

References:

Chow, J.C., 1995. Measurement methods to determine compliance with ambient air quality standards for suspended particles. J. Air Waste Manag. Assoc. 45, 320–382. https://doi.org/10.1080/10473289.1995.10467369

Davison, A.C., Hinkley, D. V., 1997. Bootstrap Methods and Their Application. Biometrics 54, 795. https://doi.org/10.2307/3109789

Decarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Anal. Chem. 78, 8281–8289. https://doi.org/10.1021/ac061249n

DeCarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Anal. Chem. 78, 8281–8289. https://doi.org/10.1021/ac061249n

Goetz, J.D., Giordano, M.R., Stockwell, C.E., Bhave, P. V, Puppala, P.S., Panday, A.K.,

Jayarathne, T., Stone, E.A., Yokelson, R.J., Decarlo, P.F., 2022. Aerosol Mass Spectral Profiles from NAMaSTE Field-Sampled South Asian Combustion Sources. Environ. Sci. Technol. https://doi.org/10.1021/acsearthspacechem.2c00173

Guo, J., Zhou, S., Cai, M., Zhao, J., Song, W., Zhao, W., Hu, W., Sun, Y., He, Y., Yang, C., Xu, X., Zhang, Z., Cheng, P., Fan, Q., Hang, J., Fan, S., Wang, Xinming, Wang, Xuemei, 2020. Characterization of submicron particles by time-of-flight aerosol chemical speciation monitor (ToF-ACSM) during wintertime: Aerosol composition, sources, and chemical processes in Guangzhou, China. Atmos. Chem. Phys. 20, 7595–7615. https://doi.org/10.5194/acp-20-7595-2020

Jayne, J.T., Leard, D.C., Zhang, X., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an aerosol mass spectrometer for size and composition analysis of submicron particles. Aerosol Sci. Technol. 33, 49–70. https://doi.org/10.1080/027868200410840

Kadachi, A.N., Al-eshaikh, M.A., 2012. Limits of detection in XRF spectroscopy. X-Ray Spectrom. 350–354. https://doi.org/10.1002/xrs.2412

Lakra, A., Kumar, A., Sethi, D., Sekhar, H., Jain, V., Tripathi, S.N., 2026. High-resolution source apportionment and health risks of PM 2 . 5 -bound trace elements across a major Indian city : Seasonal and diurnal insights from a multi-site campaign using a mobile laboratory platform. Atmos. Environ. 365, 121675. https://doi.org/10.1016/j.atmosenv.2025.121675

Lakra, A., Shukla, A.K., Bhowmik, H.S., Yadav, A.K., Jain, V., Murari, V., Gaddamidi, S., Lalchandani, V., Tripathi, S.N., 2024. Comparative analysis of winter composite-PM2.5 in Central Indo Gangetic Plain cities: Combined organic and inorganic source apportionment and characterization, with a focus on the photochemical age effect on secondary organic aerosol formation. Atmos. Environ. 338, 120827. https://doi.org/10.1016/j.atmosenv.2024.120827

Lalchandani, V., Srivastava, D., Dave, J., Mishra, S., Tripathi, N., Shukla, A.K., Sahu, R., Thamban, N.M., Gaddamidi, S., Dixit, K., Ganguly, D., Tiwari, S., Srivastava, A.K., Sahu, L., Rastogi, N., Gargava, P., Tripathi, S.N., 2022. Effect of Biomass Burning on PM2.5 Composition and Secondary Aerosol Formation During Post-Monsoon and Winter Haze Episodes in Delhi. J. Geophys. Res. Atmos. 127. https://doi.org/10.1029/2021JD035232

Li, Q., Cheng, H., Zhou, T., Lin, C., Guo, S., 2012. The estimated atmospheric lead emissions in China, 1990-2009. Atmos. Environ. 60, 1–8. https://doi.org/10.1016/j.atmosenv.2012.06.025

Li, Y.J., Lee, B.Y.L., Yu, J.Z., Ng, N.L., Chan, C.K., 2013. Evaluating the degree of oxygenation of organic aerosol during foggy and hazy days in Hong Kong using high-resolution time-of-flight aerosol mass spectrometry (HR-ToF-AMS). Atmos. Chem. Phys. 13, 8739–8753. https://doi.org/10.5194/acp-13-8739-2013

Manchanda, C., Kumar, M., Singh, V., Faisal, M., Hazarika, N., Shukla, A., Lalchandani, V., Goel, V., Thamban, N., Ganguly, D., Tripathi, S.N., 2021. Variation in chemical composition and sources of PM2.5 during the COVID-19 lockdown in Delhi. Environ. Int. 153, 106541. https://doi.org/10.1016/j.envint.2021.106541

Mukta, T.A., Hoque, M.M., Sarker, E., Hossain, N., 2021. Seasonal variations of gaseous air

pollutants (SO 2 , NO 2 , O 3 , CO) and particulates (PM 2 . 5 , PM 10) in Gazipur : an industrial city in Bangladesh. Adv. Environ. Technol. 4, 195–209. https://doi.org/10.22104/AET.2021.4890.1320

Rai, P., Furger, M., El Haddad, I., Kumar, V., Wang, L., Singh, A., Dixit, K., Bhattu, D., Petit, J.E., Ganguly, D., Rastogi, N., Baltensperger, U., Tripathi, S.N., Slowik, J.G., Prévôt, A.S.H., 2020. Real-time measurement and source apportionment of elements in Delhi's atmosphere. Sci. Total Environ. 742. https://doi.org/10.1016/j.scitotenv.2020.140332

Rai, P., Slowik, J.G., Furger, M., El Haddad, I., Visser, S., Tong, Y., Singh, A., Wehrle, G., Kumar, V., Tobler, A.K., Bhattu, D., Wang, L., Ganguly, D., Rastogi, N., Huang, R.J., Necki, J., Cao, J., Tripathi, S.N., Baltensperger, U., Prevot, A.S.H., 2021. Highly time-resolved measurements of element concentrations in PM10 and PM2.5: Comparison of Delhi, Beijing, London, and Krakow. Atmos. Chem. Phys. 21, 717–730. https://doi.org/10.5194/acp-21-717-2021

Sethi, D., Lakra, A., Kumar, A., Chakraborty, S., Sekhar, H., Jain, V., Tripathi, S.N., 2026. Hyperlocal air quality monitoring and source apportionment of non- refractory $PM_{2.5}$ at three urban sites using stationary van - based measurements : A Lucknow case study. Atmos. Chem. Phys. 1–41. https://doi.org/10.5194/egusphere-2026-595

Sharma, S.K., Mandal, T.K., 2023. Elemental Composition and Sources of Fine Particulate Matter (PM2.5) in Delhi, India. Bull. Environ. Contam. Toxicol. 110, 1–8. https://doi.org/10.1007/s00128-023-03707-7

Shukla, A.K., Lalchandani, V., Bhattu, D., Dave, J.S., Rai, P., Thamban, N.M., Mishra, S., Gaddamidi, S., Tripathi, N., Vats, P., Rastogi, N., Sahu, L., Ganguly, D., Kumar, M., Singh, V., Gargava, P., Tripathi, S.N., 2021a. Real-time quantification and source apportionment of fine particulate matter including organics and elements in Delhi during summertime. Atmos. Environ. 261, 118598. https://doi.org/10.1016/j.atmosenv.2021.118598

Shukla, A.K., Lalchandani, V., Bhattu, D., Dave, J.S., Rai, P., Thamban, N.M., Mishra, S., Gaddamidi, S., Tripathi, N., Vats, P., Rastogi, N., Sahu, L., Ganguly, D., Kumar, M., Singh, V., Gargava, P., Tripathi, S.N., 2021b. Real-time quantification and source apportionment of fine particulate matter including organics and elements in Delhi during summertime. Atmos. Environ. 261. https://doi.org/10.1016/j.atmosenv.2021.118598

Shukla, A.K., Tripathi, S.N., Canonaco, F., Lalchandani, V., Sahu, R., Srivastava, D., Dave, J., Thamban, N.M., Gaddamidi, S., Sahu, L., Kumar, M., Singh, V., Rastogi, N., 2023. Spatio-temporal variation of C-PM2.5 (composition based PM2.5) sources using PMF*PMF (double-PMF) and single-combined PMF technique on real-time non-refractory, BC and elemental measurements during post-monsoon and winter at two sites in Delhi, India. Atmos. Environ. 293, 119456. https://doi.org/10.1016/j.atmosenv.2022.119456

Tobler, A., Bhattu, D., Canonaco, F., Lalchandani, V., Shukla, A., Thamban, N.M., Mishra, S., Srivastava, A.K., Bisht, D.S., Tiwari, S., Singh, S., Močnik, G., Baltensperger, U., Tripathi, S.N., Slowik, J.G., Prévôt, A.S.H., 2020. Chemical characterization of PM2.5 and source apportionment of organic aerosol in New Delhi, India. Sci. Total Environ. 745. https://doi.org/10.1016/j.scitotenv.2020.140924