Author Responses/Comments (text in blue) to comments by Referees (text in black) (AC 2 to RC 2)

We thank the referees for their valuable comments which have greatly helped us to improve the manuscript. Please find below our point-by-point responses (in blue) after the referee comments (in black). The changes in the revised manuscript are written *in italic*.

Comment on egusphere-2022-1165

Anonymous Referee #2

The manuscript entitled: “Real-time measurements of NMVOCs in the central IGB, Lucknow, India: Source characterization and their role in O3 and SOA formation” by Jain et al. investigates air quality in Lucknow, India to understand how local and regional emissions contribute to ozone and SOA formation. They employed PMF to apportion NMVOCs to sources related to traffic, solid fuel combustion, volatile chemical products, and secondary formation. These factors were further investigated and found to be consistent with other key measurements such as organic PM PMF factors. Traffic and solid fuel combustion contributed most to SOA formation and OFP. Overall, this article fits the scope of the journal and addresses important questions regarding the sources of pollutionwhich drive air quality. I would recommend publication after major revisions with regardsto the comments below.

General Comments

Hydrocarbons were measured via PTR-TOF-MS in this study, but alkanes and small alkenes cannot be detected. It is important to be transparent regarding this fact as such species may represent significant fractions of the true total NMVOCs as well as potentiallyyour PMF factors’ relative abundance, OFP, and SOA yield. Acknowledgement and discussion of this limitation is necessary.

Reply: We agree with the reviewer that some small compounds, (C1-C4) alkanes and (C1-C4) alkenes, cannot be detected by PTR-ToF-MS. These compounds are substantially present in nature and contribute significant fractions to the true total NMVOCs concentrations. We added some discussion in section 3.2.

Line 466-472:

For the first time, we have included mass spectra of 170 NMVOCs from m/z 42.034 to m/z 197.216 in the PMF analysis. The three abundant NMVOCs (compounds below m/z 42), detected by PTR-ToF-MS, acetaldehyde, acetone, and acetic acid are not includedfrom PMF analysis. Including these NMVOCs in the PMF analysis, resulted in biased solutions where only these ions are well-explained. Additionally, few small alkanes and alkenes (C1-C4) *compounds, which are not detected by PTR-ToF-MS are excluded from PMF analysis. However, previous studies have found that these ions are minor contributors to SOA formation. Included compounds (above m/z 42) are major contributors to SOA formation and dominant markers of various sources.*

Key findings regarding factors' relative importance of OFP and SOA yield (and thus, the dominant source(s) of pollution) rely on some significant assumptions (unknown MIRs in Table S2, assumed SOA yields in Table S3). This limitation is briefly noted as "There are many NMVOCs species with unknown ozone and SOA yield values. More research on the section is needed." (lines 529 – 530). It is understandable that the analysis presented here uses the available information, but these limitations must be discussed in greater detail in relation to the findings themselves. Additionally, a sensitivity analysis regarding the unknown values is necessary to substantiate these findings. Do these assumptions make a significant impact or, if not, why?

Reply: We agree with the reviewer that it will bias the result, but the values of MIR of predominant VOCs, which are used as a tracer for different sources in PMF, were taken into account. Therefore, it will not affect the interpretation significantly.

Comparisons of the most abundant NMVOCs, PMF factors, and dominant contributors to OFR and SOA production with other studies of nearby major cities would help put these measurements and conclusions into a greater context.

Reply: We added the comparisons with the previous studies for each aspect as mentioned in different sections. We added one specific section “Comparison with other Indian and Asian cities” to highlight the importance of the study.

Line 680-720

Comparison with other Indian and Asian cities

Figure 11 represents mapped pie-chart to compare overall NMVOCs concentrations, and relative source contributions in different Asian and Indian cities. The earlier studies reported the total NMVOCs concentrations between 15-35 ppbv in different cities of China during winters(Wang *et al.*, 2016; Hui *et al.*, 2018; Yang *et al.*, 2018; Wang *et al.*, 2021). *The highest concentration of NMVOCs found in Wuhan city (~34.6 ppbv) with maximum contributions from alkanes and oxygenated VOCs* (Hui *et al.*, 2018). *The relative composition of sources of NMVOCs found in Wuhan was Industrial/Solvent usage (29.9%), followed by traffic (24.4%), fuel evaporation (23.87%), biomass burning (19.3%) and biogenic (2.5%). The urban site in Beijing reported maximum contribution from secondary VOCs (54.6%), followed by biomass burning (24.4%) and traffic (21%)* (Wang *et al.*, 2021) *while the rural site in Beijing had significant contributions from biomass burning (37%)* (Yang *et al.*, 2018). *Industrial and Traffic contributed similarly at the rural site in Beijing (~31.5%). The difference of source profiles and contributions in urban and rural areas inferred the need of different control strategies and policies in the country*(Zhang *et al.*, 2020). *It is found that vehicular emissions and biomass burning sources contributes to NMVOCs concentrations (average ~21.5 ppbv) overall 50%, and 41%, respectively during summers, in a land locked urban city, Lhasa, Tibet*(Guo *et al.*, 2022) *while Industrial/Solvent usage contributed 68% to NMVOCs (average ~33.7 ppbv) in Tokyo, Japan* (Fukusaki *et al.*, 2021a). *It is interesting to note that near the coastal region in Hong Kong, 63.7% and 13.5 % NMVOCs contributions (average ~9.8 ppbv) are related to biomass burning and ship emissions* (Tan *et al.*, 2021). *Despite various air pollution control strategies implemented for more than a decade, NMVOCs and O3 concentrations did not decrease at significant levels in Hong Kong* (Lyu *et al.*, 2017). *Previous study in Kathmandu*(Sarkar *et al.*, 2017)*, Nepal demonstrated that biomass co-fired brick kilns (29%) and traffic (28%) contributes to SOA production significantly. Other sources, such as Industrial/ Solvent-usage, biomass burning, and biogenic related emissions also dominated in the city.*

Earlier source apportionment studies over the NMVOCs mass spectra conducted in Indian cities are limited to two cities in upper IGB region, Delhi (full year) and Mohali (summer). Comparing the urban and sub-urban site of Delhi found that vehicular emissions are dominant at both sites, relatively less contributions to NMVOCs at sub-urban region (36%) as compared to urban region (57%). Throughout the year, traffic emissions dominated the NMVOCs concentration (31%), with comparable contributions from biomass burning (28%), and secondary formations (31%) overall in Delhi. Mohali is located upwind of Delhi city, with maximum contributions from biomass burning (47%), followed by traffic (25%), and secondary formations (16%). The industrial source contributed about 5%, and 12% to NMVOCs concentrations in Delhi and Mohali, respectively. While in the present study, it is found that the solid fuel combustion related emissions majorly (41.3%) contributed to NMVOCs concentrations in Lucknow, located in the central IGB region. The traffic-related emissions (23.5%) and secondary formations (18.6%) are relatively less contributing to NMVOCs as compared to upper IGB region cities (Delhi and Mohali). Moreover, the volatile chemical products emitted more during summer period in Lucknow than compared to Delhi and Mohali.

Solid fuel combustion sources aided concentrations of NMVOCs in both Mohali and Lucknow significantly. This may be due to both cities are located downwind of widespread area of agricultural fields. Both of these cities observed relatively less formations of secondary volatile organic compounds, suggested the dominance of fresh emissions than aged compounds in the air mass. Overall, the ambient concentrations of NMVOCs in Indian cities majorly influenced by emissions from solid fuel combustion, vehicular related emissions, secondary formations, and industrial sources. This suggested the need of control measures, and policies implemented for specific sources country-wide and specific to city.

Line 77-102
Introduction
Only a few studies have observed and reported the ambient NMVOCs levels in Indian cities. These studies are mainly conducted in large Indian cities such as Delhi (Srivastava, Sengupta and Dutta, 2005; Hoque *et al.*, 2008; Garg, Gupta and Tyagi, 2019; Tripathi *et al.*, 2022), *Mumbai* (Srivastava, Joseph and Devotta, 2006), *Kolkata* (Chattopadhyay *et al.*, 1997; Majumdar, Mukherjeea and Sen, 2011), *Ahmedabad* (Sahu, Yadav and Pal, 2016; Sahu, Tripathi and Yadav, 2017; Tripathi and Sahu, 2020), *Udaipur* (Yadav *et al.*, 2019; Tripathi *et al.*, 2021), *and Mohali* (Sinha, Kumar and Sarkar, 2014) *.A previous study have presented the health risk assessments for ambient VOCs levels in Kolkata* (Majumdar (neé Som) *et al.*, 2008; Chauhan, Saini and Yadav, 2014)*. Most of these studies have examined only a few NMVOCs, mainly (BTEX), with less or no information related to their sources. Real-time characterization and source apportionment studies for NMVOCs in India are limited to national capital city of Delhi* (Wang *et al.*, 2020; Stewart *et al.*, 2021; Jain *et al.*, 2022)*, and Mohali* (Pallavi, Sinha and Sinha, 2019) *across different seasons and sites. Traffic emissions and solid fuel combustion are observed to be major contributors in both cities. Significant contributions from secondary VOCs are found in Delhi while solvent based industries contributed to NMVOCs in Mohali. It is necessary to understand the different source profiles and source contributions to ambient NMVOCs in different cities. The atmospheric interactions with radicals, and meteorology highly influence the concentrations of NMVOCs in the region. Recent source apportionment studies based on real-time measurements of non-refractory fine particulate matter using HR-ToF-AMS identified various sources present at different sites in Delhi*(Tobler *et al.*, 2020; Lalchandani *et al.*, 2021; Shukla *et al.*, 2021)*. These studies emphasized that it is essential to understand the variance of sources between day-to-night and different seasons. The major contributors to fine suspended particulate matters in the National Capital Region are the burning of crop residues in neighboring states and open burning of waste, as well as the increased construction activities, industrial expansion, thermal power plants, number of vehicles (two-wheelers and cars), and residential fuel use that result from an ever-increasing population. In addition, recent studies based on real-time measurements of NMVOCs using PTR-ToF-MS in Delhi* (Wang *et al.*, 2020; Jain *et al.*, 2022) *and Mohali* (Pallavi, Sinha and Sinha, 2019) *emphasized the importance of source characterization of NMVOCs simultaneously. Very few source apportionment studies highlighted the sources of NMVOCs present in other Asian cities* (Sarkar *et al.*, 2017; Hui *et al.*, 2018; Fukusaki *et al.*, 2021b; Tan *et al.*, 2021; Wang *et al.*, 2021)*. These studies highlighted that NMVOCs sources have substantial value in checking the secondary aerosols formation, and air quality.*

The goal of this work seems to be focused on the main contributors to air pollution, but there is relatively little discussion. Do any individual NMVOCs present a health risk based on your measurements and air quality standards? How frequently do ozone and PM concentrations exceed standards, and which NMVOC and AMS factors seem to drive these events?

Reply: Yes, some of the NMVOCs based on previous epidemiological studies possess great health-risk, which is discussed as given below. We also added the NAAQS standard lines into the timeseries and also added the VOCs species and their source information to the timeseries. It is observed that SFC1 and SFC2 are dominant factors during Jan-Feb. During this period, the PM2.5 exceeds most of the days than the NAAQS standard. PM2.5 exceeds standards more frequently than ozone. Approximately, 80% of the days during the whole study period, PM exceeds the NAAQS standards in the city. We added the discussion as follows in the respective sections.

Line 400-402

During the winter period, the PM2.5 exceeds most of the days than the NAAQS standard. PM2.5 exceed standards more frequently than ozone. Approximately, 80% of the days during the whole study period, PM2.5 exceeds the NAAQS standards in the city, as shown in Figure 2.

Line 417-422

The CPCB notified the annual National Ambient Air Quality Standards (NAAQS) Only for benzene as 5 (~1.6 ppbv). While WHO recommended no safe level of exposure of benzene. The mean mixing ratio of benzene during the study period found to be 2.9 ±1.9 ppbv which is 2 times higher than the standard guidelines. Prolonged exposure or high short-term exposure to benzene adversely affect the health of citizens of the city due to its haematotoxic, genotoxic and carcinogenic properties.

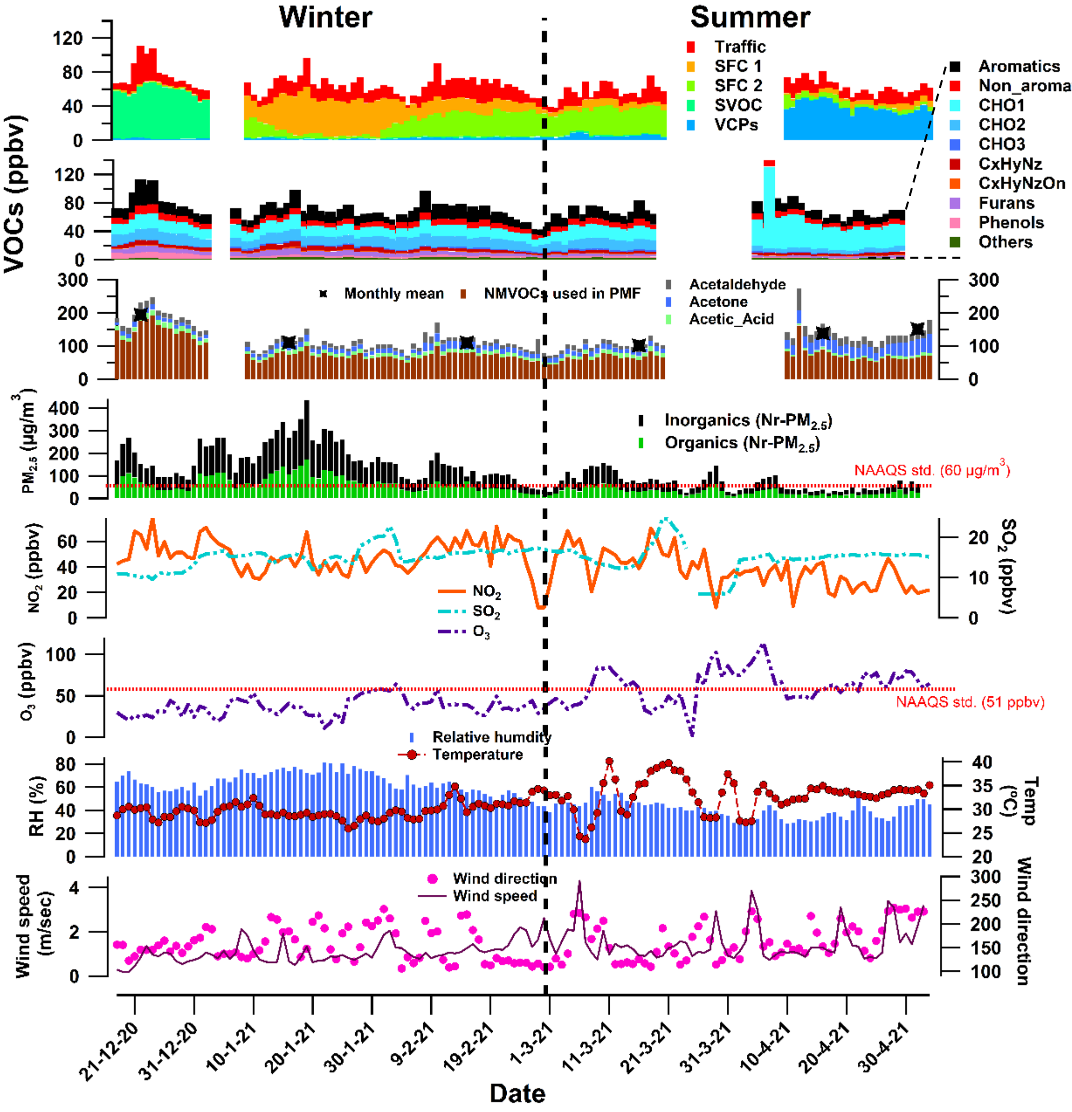


Figure 3: Daily averaged time series of acetaldehyde, acetone, and acetic acid, other NMVOCs, $PM_{2.5}$ and its organic fraction, NO_2, SO_2, O_3, temperature, relative humidity, and wind speed and direction

Specific Comments

Line 18: Notably PTR does not detect alkanes and small alkenes. I suggest specifying “measured” or “quantified” NMVOCs.

Reply: We agree that PTR does not detect alkanes and small alkenes. Therefore, we modified the sentences in the revised manuscript.

Line 18-20
About ~173 NMVOCs from m/z 31.018 to 197.216 were measured during the study period, including aromatics, non-aromatics, oxygenates, and nitrogen-containing compounds.

Lines 142 – 143: To calibrate the PTR-MS signals, “…a typical value of 2 x 10-9 cm3 s-1 ofthe proton transfer reaction rate coefficient…" was used. Was this for all NMVOCs you did not directly calibrate, or for all NMVOCs you did not have a literature value for? Clarification would be helpful.

Reply. Thanks for your suggestion. We have added a sentence in the revised manuscript.

Line 191-193

The reaction rates (k) of the ions were applied from the literature (Cappellin *et al.*, 2012)*. "A rate constant of 2 × 10−9 cm3 s−1 was assumed for all ions which reaction rates (k) were not available in the literature (Smith and Spanel, 2005).*

Given that these rate constants vary by a factor of ~2, how does the assumption of an average rate constant affect your calibrations and the rest of your analysis? A discussion on the resulting uncertainties is necessary.

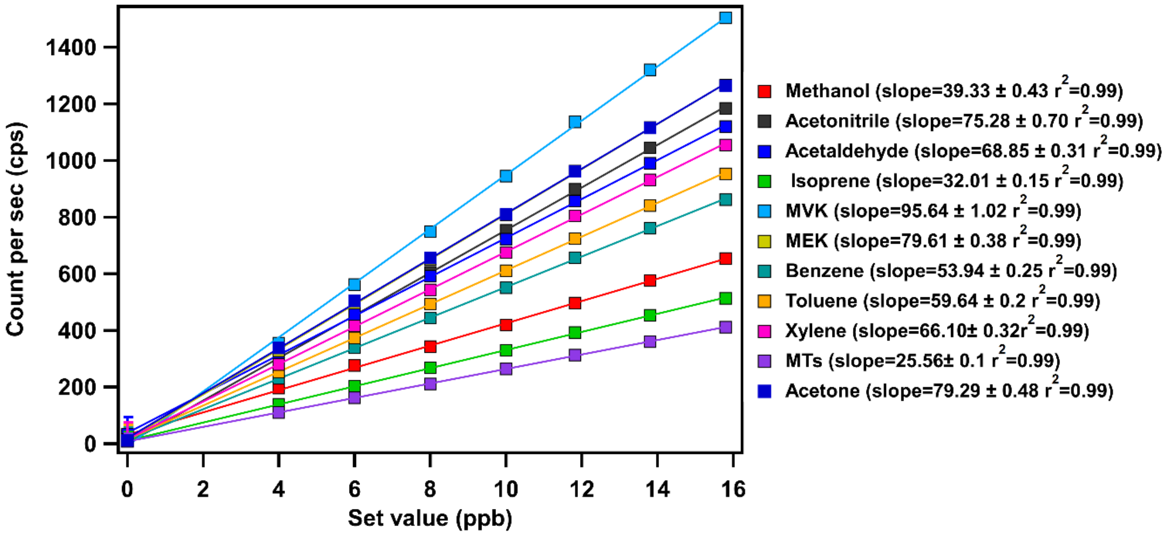
Reply: For the VOCs calibrated using the standard mixture, the overall uncertainties were in the range of 8%–13% in the calculations of the mixing ratios of VOCs including the uncertainties in the mass flow controllers (MFCs) of GCU and standard mixture (±5%–6%). It is also reported in our previous study Sahu & Saxena, 2015; Tripathi & Sahu, 2020. Several studies used k rate 2 × 10−9 cm3 s−1 for species whose reaction rates (k) are not available in the literature. Hansel et al. (1999) estimated the accuracy of the VOC mixing ratio measurements to 30%, mainly caused by the uncertainties of the reaction rate constants, which are up to ±20%. We have discussed these points in the revised version of the manuscript.

Line 193- 199

The overall uncertainties were in the range of 8%–13% in the calculations of the mixing ratios of VOCs which were present in the standard mixture. The cause of uncertainties in the calculation VOC mixing ratios includes the uncertainties in the mass flow controllers (MFCs) of GCU and standard mixture (±5%–6%). The reaction rates (k) of the ion were applied from the literature (Cappellin et al., 2012). A rate constant of 2 × 10−9 cm3 s−1 was assumed for all ions which reaction rates (k) were not available in the literature.(Hansel et al., 1999; Steinbacher et al., 2004) have reported up to 30% of uncertainty in the calculations of the mixing ratios of VOCs due to k reaction rate.

What were the sensitivities and limits of detection for your standards? Do estimated LODs have implications for your other measurements (e.g., are your measurements of tetradecane and others above the LOD)?

Reply. The sensitivities of the standards were presented in following Figure. The LOD of the standards in the range of 1 ppt to 45 ppt. Yes, the values below detection limit were removed from the data.



Was the instrument's transmission function investigated / were transmission correction applied?

Reply. Yes, transmission correction was applied.

Since calibrations were done in the beginning, middle, and end of the campaign, were the signals normalized to the reagent ion (as is typical for PTR-MS measurements) to account for relative humidity contributions to the water content in the reactor and general instrument variability?

Reply. Yes, we have already checked the variation of standard (Mentioned in the sensitivity figure) with humidity except few VOCs (Acetone, methanol) most of the VOCs do not show any significant variability with RH.

Were instrument background signals measured and applied? If so, how were they measured?

Reply. Yes, background signals were measured and applied. The detailed information was discussed in our previous studies. As per reviewer 1 suggestion we have included in the SI.

Line 194: There is some inconsistent use of "ions" and "m/z" alongside "mixing ratios." If calibrated NMVOCs were used in your PMF analysis, they are no longer ions. This also applies elsewhere in the manuscript.

Reply: We thank the reviewer for the comment. We checked thoroughly the manuscript and changed the "ions" to the "NMVOCs" or "NMVOC species".

Lines 270 – 271: "…more partitioning of the gas phase during summers relatively to winters." is somewhat ambiguous regarding the direction of partitioning during colder vs warmer months.

Reply: We thank the reviewer for the comment and we have removed the sentence. We also added more information in the section regarding the NMVOCs stats and their relative contributions in the section.

Line 335: It is unclear to me what "±3" refers to in the context of the scaled residuals.
Reply: When considering scaled residuals as a function of m/z (mass spectra profile) and time (timeseries), the data scatter around zero with the interquartile range almost always between "±3" throughout the entire year, is considered as a good quality of PMF solution. This means that the scaled residuals reasonable range for any m/z or for any particular time period is "±3" (Paatero and Hopke, 2003; Canonaco et al., 2021). Therefore, based on previous literature, we considered the ideal value of "±3" of scaled residuals for each data point when examining the scaled residuals as a function of timeseries for each PMF solution. After that, we finalized the optimum PMF solution. The timeseries graph of scaled residuals (supplementary Figure S4) suggests that the PMF solution with 5-factors is having less than "3" scaled residual values for any given data point. While solutions (2-4 factors) have a lot of data points with more than "3" value of scaled residual.

Line
Another important metric is the evaluation of scaled residuals as a function of timeseries and mass spectra. The scaled residuals ±3 for each data point in the time series are considered, which is *a evidence of a good quality PMF solution* (Paatero and Hopke, 2003; Canonaco *et al.*, 2021).

Lines 335 – 336: The text would suggest 3 – 7-factor solutions, but Figure S4 shows timeseries for 2 – 10-factor solutions. The diurnals in Figure S4 also include an 11-factor solution whereas the 11-factor timeseries is absent. Figure 4 includes 3 – 11-factor solutions. Please address these inconsistencies.
Reply: We addressed these inconsistencies. We examined the PMF solutions from 3-10 factors solutions. We changed the supplementary Figure and corrected the text also in the main section.

Revised supplementary:

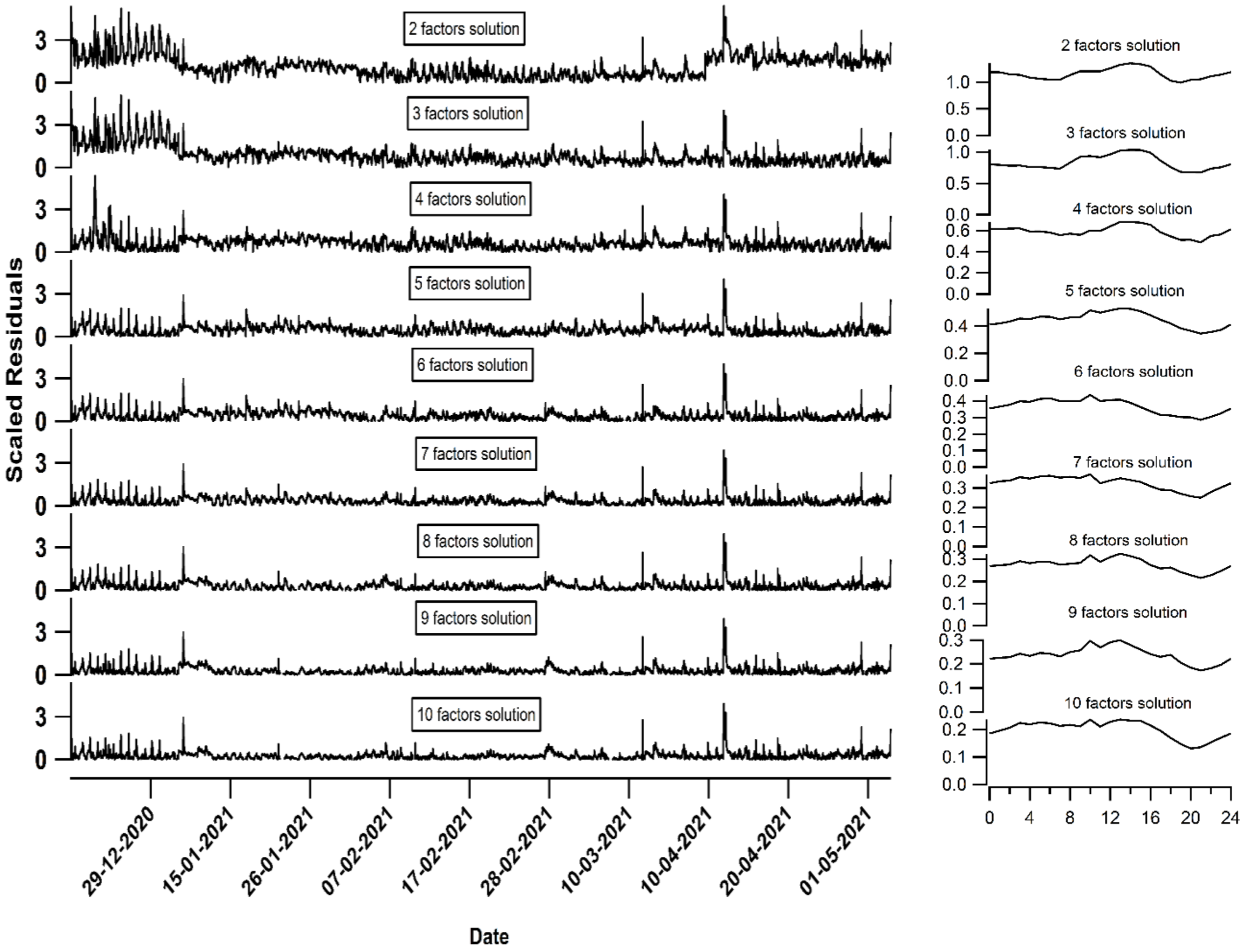


Figure S4: Timeseries and diurnal variation of the scaled residuals from 2-11 factors solution.

Line 339: Estimating from Figure 4, the percent change in Q/Qexp went from ~12% in the 4-factor solution to ~10% in the 5-factor solution to ~8% in the 6-factor solution which does not seem significantly different. Additionally, the total scaled residuals dropped significantly between the 5 and 6 factor solutions. From these parameters alone, one could argue that the 6-factor solution would be better. However, there are other metrics as described previously in the same paragraph that could be used to rule out the 6-factor solution as nonsensical. Please discuss further why the 5-factor solution was chosen as opposed to a 6-factor solution in relation to these other metrics.

Reply: The red line in the Figure 4 represents the % change in Q/Qexp while the black color line and grey shaded area represents the total residuals. The exact values for both parameters related to the respective factor is given below in a table. We have also added this in the supplementary Table. As the 3 to 5 factor solution have maximum % change in Q/Qexp, which is an important parameter. Previous studies have also reported that the solution having more % change in Q/Qexp values will be statistically relevant.

Another parameter, as pointed out, that the total scaled residuals for any given point should be less than 3*standard deviation. So, out of 3, 4 and 5 factor solution, 5-factor solution is having least total scaled residuals. We agree that 6 factor solution is having less total scaled residuals, but it also has less (11% only) d(Q/Qexp) value. After examining the profile, timeseries and durnal of 6 factor solution, one of the factors was split into two factors with similar profiles.

As pointed out in Figure S4: the timeseries of scaled residuals from 2-10 factor solution., the distinct structure of scaled residual in 4 factor solution timeseries during December while it is not present in the 5-factor solution. This implies that 5 factor solution is better. Moreover, there is no significant change in timeseries when going from 5 factor to 6 factor and so on.

We have also refined the results by constraining with SVOC factor profile. The bootstrap analysis is used to examine the uncertainty in the solution. It is observed that 5 factor solution is having uncertainty of 1% or less for individual each factor as described in the section 2.5. We have also added and restructured the section 2.5 and supplementary text ST1.

Factor	**% Change in Q/Qexp**	**Total scaled residuals**
3	20.22	142.33
4	19.37	104.94
5	18.58	81.81
6	11.29	68.91
7	9.89	62.71
8	12.09	55.39
9	8.17	52.78

Line 363: A short discussion of Figure S5 (as a whole and for each factor) would be helpful for readers such as myself with limited knowledge of this type of analysis. The slopes suggest an error of 1% or less, but the figure seems to tell us more than that. SFC2 seems to have 2 distinct lobes, one with higher concentrations and lower uncertainties (such that PMF latched onto these measurements to determine this factor) and a second lobe with lower concentrations and a higher slope/uncertainty. This effect is stronger for VCPs, and the SVOC factor has essentially no correlation between spread and concentration. Are these observations meaningful and, if so, what do they mean for the quality of the factors?

Reply: We added some discussion in the supplementary text ST2.

Supplementary Text (ST2):

The figure represents the analysis of mass error estimation. It is an additional module in advanced ME-2 engine based PMF. As mentioned in supplementary text ST1, we have performed bootstrapping analysis over the mass spectra of our input. Bootstrapping is a technique where; several replicates were generated using resampling strategies. This applied systematic technique aims to explore rotational ambiguity over a defined setting. This analysis gives confidence in extracting the environmentally reasonable PMF runs. The mass error estimation bears information on these PMF runs. For example, in the study, we did bootstrapping for 500 runs, which means 500 times input (mass spectra and time series) were resampled again to check if the same solution with same factors came or not. These factors are defined by same profile and diurnal/ normal cycle. The mass error estimation here suggests the quantifiable error over these 500 PMF runs, when comparing the spread (standard deviation) to its contribution (mean/median). This graph actually suggests the error distribution for each factor for the selected user-criteria based PMF runs. This distribution represented by color schemes could be according to the date (timeseries) or variables (averaged value in the factor). In the graphs, the spread is standard deviation over the mean value for each factor. Given that there are no time dependencies, the relative error can be expressed as a percentage using the slope of the linear fit. In the study, for each factor mass *error estimation shows the cumulative average error of 1% for 500 PMF runs with same input after doing bootstrap analysis. This means that the selected solution is very robust, and environmentally reasonable.*

Section 3.2.2: More discussion of the CWT analysis is necessary, including references to Figure 1. Do these results align with expectations?

Reply: we changed the structure of the section 3.2.2 in continuation to the comments by Referee #1. We added the discussion in each sub-section for each factor.

Line 500-502

The back trajectory analysis of the factor, (CWT graph), given in Supplementary Figure S3, shows the probable sources of traffic present near the sampling site.

Line 526-530

The city is surrounded by various agricultural fields, which generally involve open biomass burning activities. The back trajectory analysis of the factor also shows the probable sources in nearby areas, mainly coming from the west direction of the sampling site (supplementary Figure S3). This argues that this factor is also influenced by the aged biomass-burning plume, transported from sources located on the outskirts of the city and nearby districts.

Line 561-563

It may be interpreted that SFC 2 is influenced by fresh oxidation of primary biomass burning emissions. Moreover, the CWT plots as shown in supplementary Figure S3, no evidence of its long-range transport is present.

Line 600-611

This may be due to the influence of particular activity in near-by industries. A conglomerate of the industries is present in the southwest direction of the sampling site within and outside the city, as shown in Figure 1. The direction of the wind changes to the southwest during summers may have brought the high levels of naphthalene and its derivatives emitted from these industrial areas to the sampling site. The CWT graph also shows the strong influence of the source present in the southwest direction of the sampling site (supplementary Figure S3). A previous study has found that among the emitted OVOCs from sewage sludge, first-order OVOCs constituent ~60%, followed by high-order OVOCs (Haider *et al.*, 2022).*Interestingly, there are three sewage treatment plants located near the sampling site. They may have also influenced the concentrations of OVOCs at the sampling site. The influence of factor contribution during summertime is probably due to the increased production of naphthalene, formaldehyde, and ethanol from their local industrial sources and secondary formations at higher temperatures, as shown in the time series of the factors (supplementary Figure S8).*

There is abundant discussion of key NMVOCs in each factor and how these factors were identified. How do these factors compare to other studies (distribution of NMVOCs within each factor, most abundant NMVOCs, relative abundance of each factor compared to each other, etc.)

Section 3.4: This section briefly reports the results of the OFP and SOA production results, but requires more discussion to articulate the impact and significance of these findings. They should be compared to other studies for context. Limitations should be restated and interpreted with regards to the results.

Reply: We thank the reviewer for the suggestions. We improved the section by comparing the results of our study to the previous studies. We also added the discussion regarding the control strategies of ozone and SOA and compared with previous studies of Asian cities. We improved the section replying to comments #1 by Referee 1 (section 3.4), and our Reply is repeated here.

Line 618-621

Toluene, xylene, and isoprene were found to be the highest contributor in terms of OFP in other Asian cities including Guangzhou, and Beijing (Duan *et al.*, 2008; Zheng *et al.*, 2009; Zhu *et al.*, 2016; Zhan *et al.*, 2021). *In the previous study in Delhi, it has also been noticed that contributions of aromatics (xylene and toluene) have substantital effect to the ozone formation potential* (Tripathi *et al.*, 2022).

Line 634- 638

Previous studies have also found that aromatic hydrocarbons contributed more than 95% to the SOA formation potential in other Asian cities (Qin *et al.*, 2021; Zhan *et al.*, 2021).*It was observed that the sources related to vehicular emissions (diesel and petrol driven vehicles), paddy stubble fire, and garbage fire emissions were most contributing factors for ozone formation potential in Mohali* (Kumar *et al.*, 2020).

Line 645-654

These values estimate the potential for ozone and SOA formation and do not indicate the actual yields of ozone and SOA. This estimation method represents the complex behavior of NMVOCs, NOx and solar radiation for producing tropospheric ozone and SOA. There are many NMVOCs species with unknown ozone and SOA yield values. It is needed to understand the chemical fates and pathways of many NMVOCs by mimicking real-time atmosphere in smog-chamber studies or through computational modelling studies. More research on this section is needed. Nonetheless, other parameters, including solar radiation and concentration of oxides of nitrogen, also play a key role in the formation of ozone in the troposphere. In reality, OFP and SOA do not provide complete information about how VOCs influence O3 and organic aerosol chemistry zone formation in Lucknow is more sensitive to NMVOCs concentrations than NOx, similar to other Asian cities.

Line 654-679

So, Decreasing the VOCs/NOx ratio would also help reduce the secondary pollutants (O3 and SOA). It is observed that vehicular emissions were the main source of aromatics (benzene, toluene, xylene). Therefore, vehicular emission control strategies should be implemented to reduce aromatic (BTEX). Stringent implementation of policies and fuel-efficient standards related to vehicular emissions in Japan and South Korea have primarily improved the air quality (13-17% reduction in NMVOCs) (Wang et al., 2014). In the present study, one of the key observations was that toluene is the main contributor to SOA and ozone production potential. This illustrates that targeting other sources of some NMVOCs (toluene and xylene) will also enhance its control. For example, paint solvents (source of ethylbenzene and xylene) and printing products (source of toluene) were targeted in a city, Hong Kong, where the VOC content of 172 types of consumer products was prescribed by the respective government (Lyu et al., 2017). In the present study, other potential contributor species are methyl cyclohexene (for ozone) and naphthalene (for SOA). These compounds are related to volatile chemical products, as found in the PMF analysis in Lucknow. This infers stringent policies related to solvent-based industries such as textile, automobile, paints, and disinfectants are needed. Regulation and control of NMVOCs content in manufacturing and use of solvent-based *products such as pants, disinfectants, fungicides, and insecticides should also be implemented. In China, various industries implemented end-of-pipe measures to control NMVOCs, such as refineries, plant oil extraction, gasoline storage and service stations, pharmacies, and crude oil storage and distribution (Wang et al., 2014) It is also estimated that China's end-of-pipe technologies and new energy-saving policies would help decrease about one-third of NMVOC emissions (Zhang et al., 2020). Phenols and Furans were observed as one of the highest contributors to SOA formation potential related to solid fuel combustion. This suggests controlling fossil fuel usage for residential energy and crop-residue burning in the fields within and around the city Lucknow. Firewood burning during the heating period and domestic in-fields straw burning have substantially reduced emissions from biomass burning in China (Wu et al., 2020) . (Derwent et al., 2007)reported that reactivity-based VOC control measures might be more effective than mass-based regulations in controlling ozone and secondary organic aerosol formation. The present study also suggests that the reduction of VOC, especially from vehicular emission is needed for the Abatement of ozone and SOA formation in urban areas.*

Lines 513 – 514: When breaking down the OFP (and later SOA) contributions for each factor, one average value is presented for the full measurement period. Do these distributions (and thus dominant factor) vary significantly between seasons (e.g., winter, late winter, and summer as in Figure S1)?
Reply: We added the supplementary figure S3 representing the diurnals from two different seasons, winter and summer for each factor. Each season is divided according to the defined period by IMD, Indian Meteorological Department as winter covering from December to February and summer covering from March to May. As it is observed from the two diurnals, the pattern of variation of sources from day-to-night does not change much between the seasons. However, their relative contributions change significantly. Therefore, We did not divide the period for measuring OFP and SOA, contributions of factors etc. for different seasons. These distributions did not vary between seasons. Only their absolute values differ but overall, their distribution and relative % remain similar.

Figure 5: What is the difference between the bars and dots in the profiles? For the SVOC profile, what do the grey bars represent?
Also, I would suggest creating a new supplementary figure with the factors' diurnals for each season.
Reply: We added the difference between the bars and dots in the profiles in the figure caption itself to be self-explanatory. Also, We added the diurnal for two seasons with the factors in the supplementary Figure 3. These figures show diurnal pattern remains similar for two seasons. However, only difference in their absolute values of NMVOCs concentrations.

Line 496-497
The diurnal pattern is compared between two seasons, winters and summers, also shows similar pattern in supplementary Figure S7.

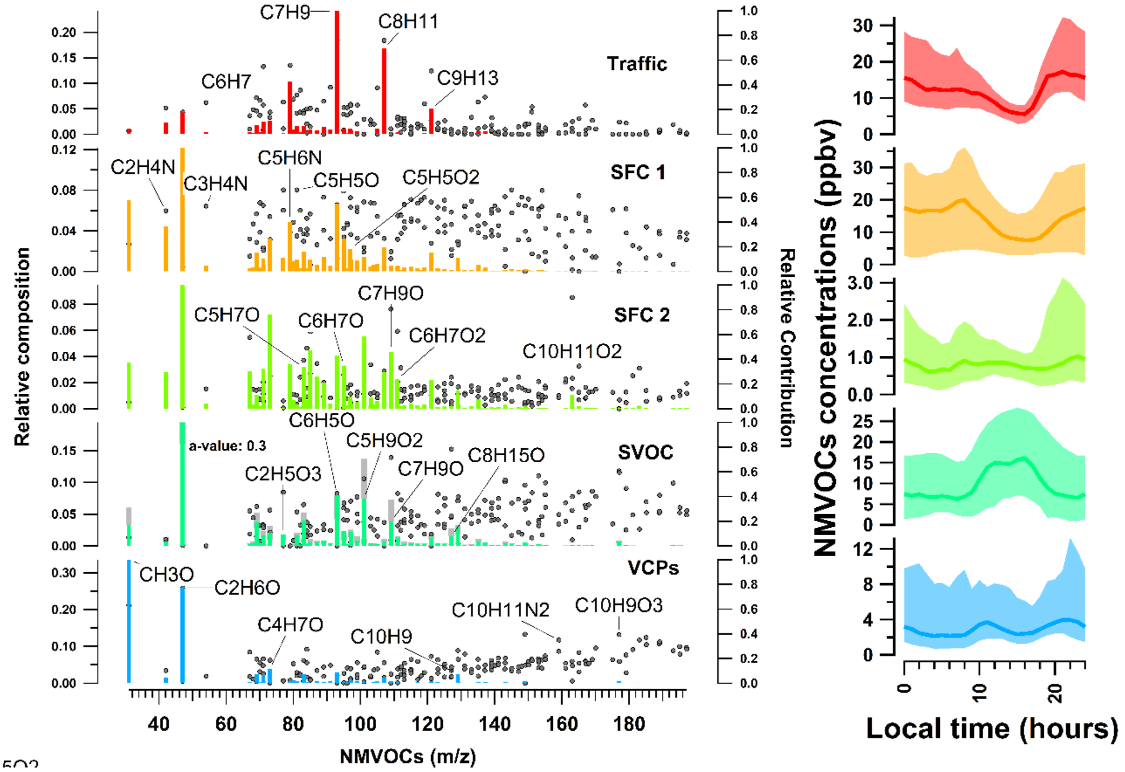


Figure 5: Profile and diurnal variation of individual factors of selected 5-factor solution after PMF analysis at Lucknow for the whole study period. In (a), the left axis represents the relative composition of each factor, given by the vertical bars. The sum of all the bars at different m/z for each factor is 1, and the right axis represents the relative contribution of each factor to a given m/z, shown as grey dots. The grey bars in the SVOC factor represents the degree of constraint on the known source profile and time series. In (b) the middle dark line represents the median of the diurnal while the shaded region represents the interquartile ranges from 25-75th percentiles.

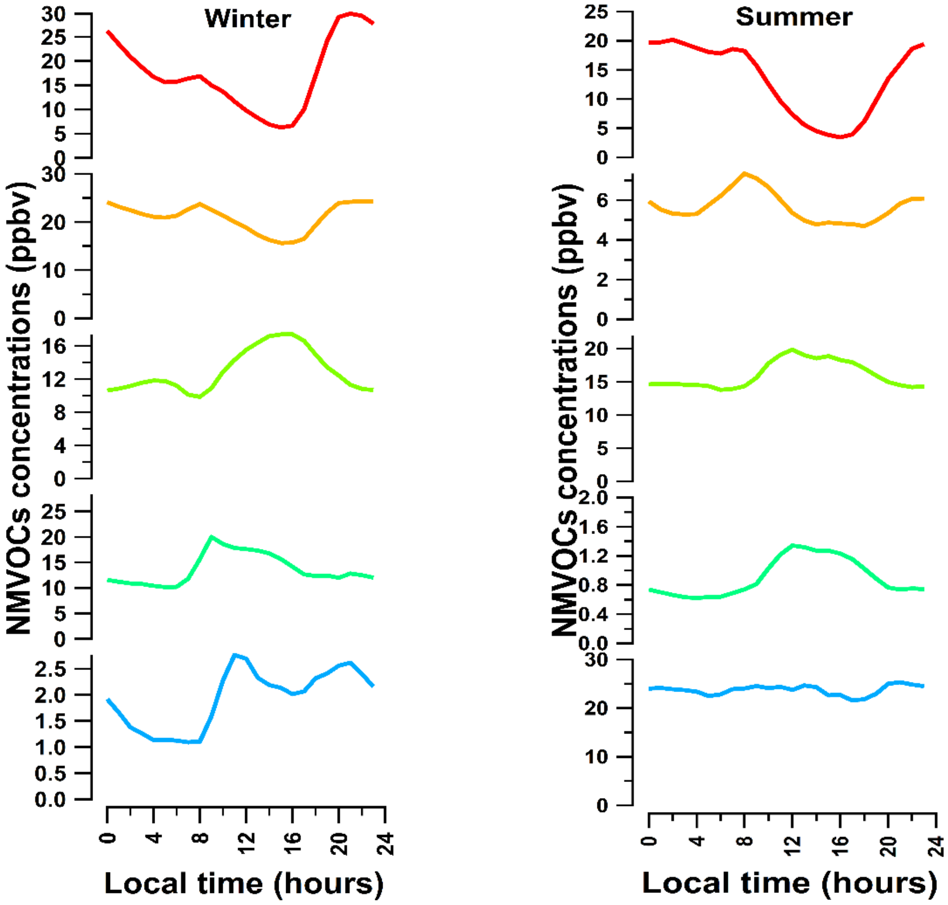


Figure S3: Diurnal plots of different factors from selected 5-factors solution for different seasons, Winter (Dec-Feb) and Summer (March-May)

Figure 7a-b: It may be easier to understand this plot with species names as opposed to formulas (or both names and formulas together). Species are already associated with these formulas in Table S1 and “Ethanol” is already provided as a specific compound inthis figure. “Unknown” species should be left as formulas.

Reply: We have modified the Figure and added the names and formula for each species in the figure. There are no unknown species.

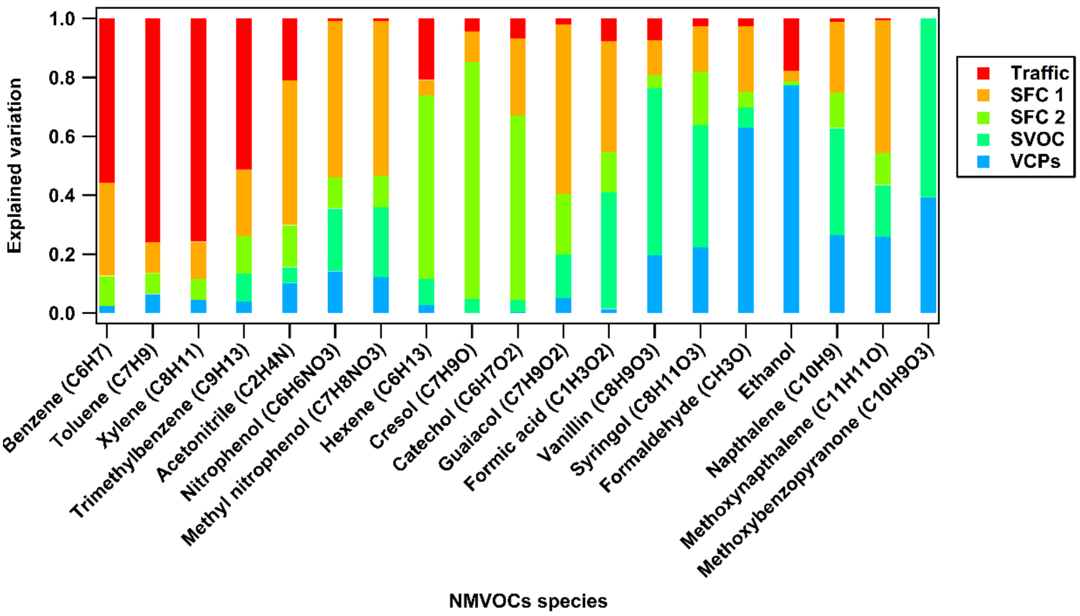


Figure 10: I believe the third value in each box of (b) and (c) refers to the SOA yield mass concentration and mixing ratios, respectively. Clarification is necessary in the caption. Also, if I am correct, why is the corresponding OFP value not included?

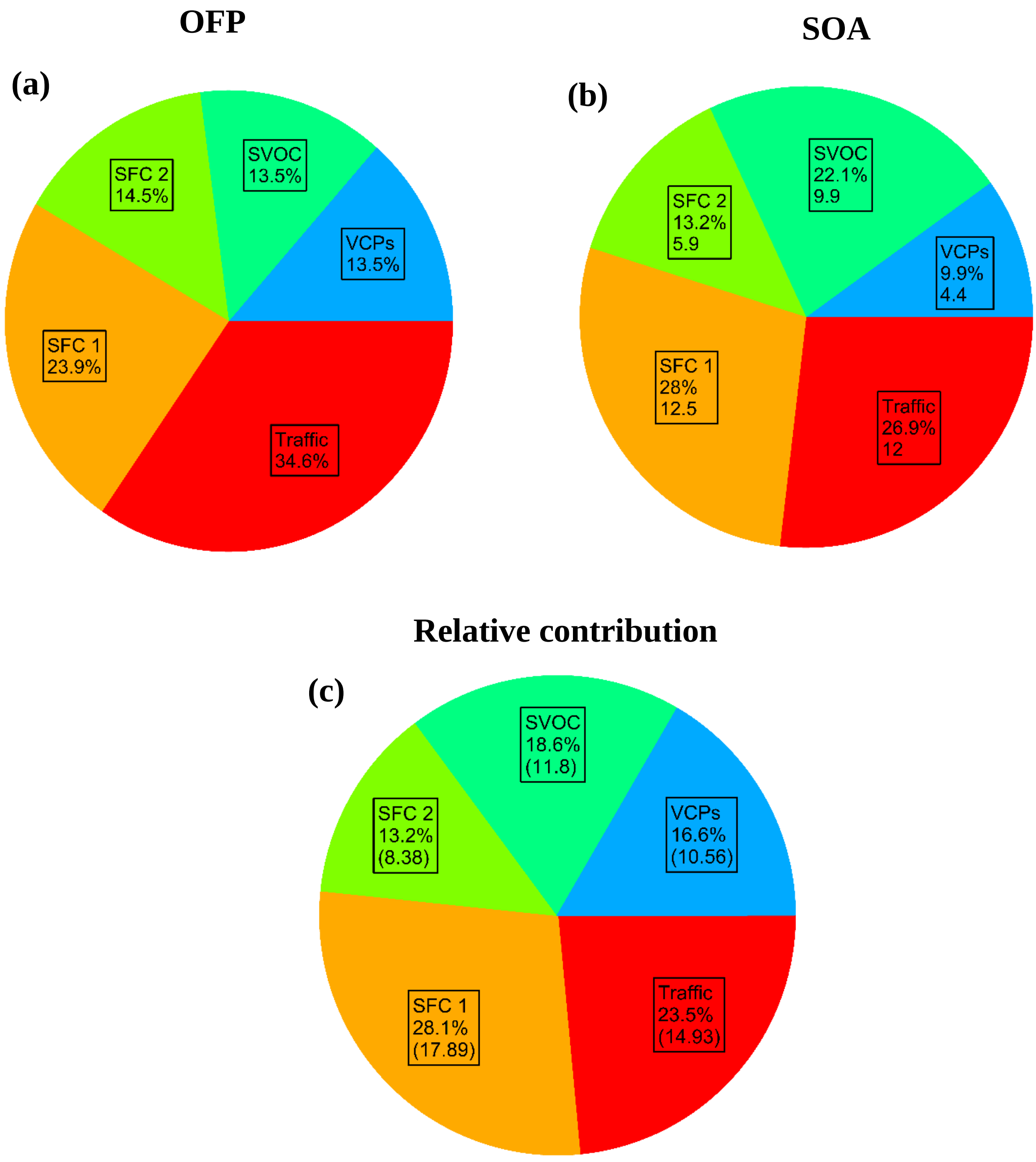


Figure 10: Distribution in percentage (%) of individual factors to (a) Ozone formation potential (OFP), (b) SOA formation, (c) Relative contribution. The bottom absolute values (in brackets) for (b) and (c) are the SOA yield mass concentration (µg/m3) and average mixing ratios (ppbv)

Technical Corrections

Line 101: Please define “MSMEs”.

Reply: We have added the full form to the text- micro, small and medium enterprises

Line 236: I believe this should be “Equation 5.”

Reply: we have made the required changes.

Line 344: Please define “SVOC” (currently defined later in line 368).

Reply: we have made the required changes.

Line 345: Please define “random values” / “a”.

Reply: The “a” value is actually a scalar value which varies from 0 to 1. This constraining method can be applied using ME-2 Engine for PMF analysis. In the study, we have used the setting of “a” values starting from 0 to 1 with difference of 0.1 (delta a= 0.1). This means that “a” value would be 0, 0.1, 0.2, 0.3, 0.4, …. till 1. This a- value approach is used to constrain the elements of F matrix (factor profiles) and/ or g matrix (factor time series) as shown in the equation below:

$$f_{j,\, solution} = f_j \pm a \cdot f_j$$

$$g_{i,\, solution} = g_i \pm a \cdot g_i$$

Figure 4: there is no clear link between the traces and their corresponding axes. Add infoto the caption, e.g., “…total summed scaled residuals (red trace)…” or change right axis color to red (similar to Figure 3).

Reply: We have made the required changes. In Figure 4, the left axis represents the % change in Q/Qexp which should be red in color. While the total scaled residuals represent black line, filled with grey color. Please find the changed figure below.

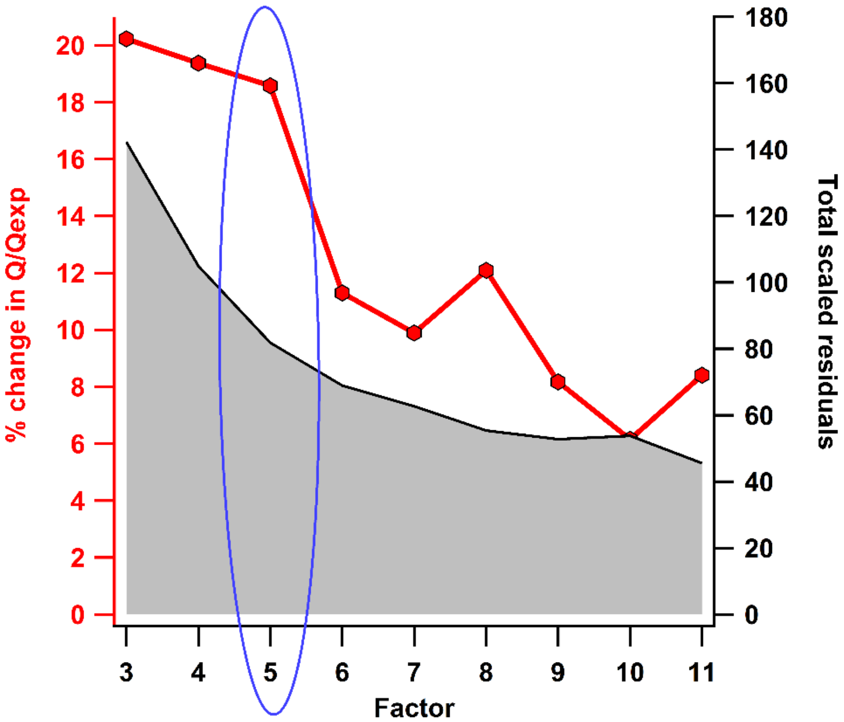


Figure 10: Assuming the bottom values for (b) and (c) are the SOA yield mass concentration and average mixing ratios, respectively, please include units

Reply: We have edited the caption of the Figure.

Figure 10: Distribution in percentage (%) of individual factors to (a) Ozone formation potential (OFP), (b) SOA formation, (c) Relative contribution. The bottom absolute values (in brackets) for (b) and (c) are the SOA yield mass concentration (µg/m^3) and average mixing ratios (ppbv)

Figure S1: Please specify the units of wind speed in the legend.

Reply: We have added the units in the caption of Figure S1.

Figure S1: (a) Wind rose plots showing wind speed (m/sec) and wind direction in different seasons, winter and summer at the sampling site.

Figure S5: The color schemes seem off. The SVOC factor is colored according to mean concentration (color gradient from left to right), but the others don't follow this same convention. Renaming the color bar label (currently just "ppbv" where both x- and y-axes are in ppbv) could help with interpretation.

Reply: We added the discussing explain the graph and mass error estimation in supplementary text ST2. This distribution represented by color schemes could be according to the date (timeseries) or variables (averaged value in the factor). In this study, we have chosen to represent the color scheme according to the averaged concentration of each factor.

Supplementary Text (ST2):
Mass error estimation for bootstrap analysis:
The figure represents the analysis of mass error estimation. It is an additional module in advanced ME-2 engine based PMF. As mentioned in supplementary text ST1, we have performed bootstrapping analysis over the mass spectra of our input. Bootstrapping is a technique where several replicates were generated using resampling strategies. This applied systematic technique aims to explore rotational ambiguity over a defined setting. This analysis gives confidence in extracting the environmentally reasonable PMF runs. The mass error estimation bears information on these PMF runs. For example, in the study, we did bootstrap for 500 runs, which means 500 times input (mass spectra and time series) were resampled again to check if the same solution with same factors came or not. These factors are defined by same profile and diurnal/ normal cycle. The mass error estimation here suggests the quantifiable error over these 500 PMF runs, when comparing the spread (standard deviation) to its contribution (mean/median). The graph (Supplementary Figure S5) suggests the error distribution for each factor for the selected user-criteria based PMF runs. This distribution represented by color schemes could be according to the date (timeseries) or variables (averaged value in the factor). In the graphs, the spread is standard deviation over the mean value for each factor. Given that there are no time dependencies, the relative error can be expressed as a percentage using the slope of the linear fit. In the study, for each factor mass error estimation shows the cumulative average error of 1% for 500 PMF runs with same input after doing bootstrap analysis. This means that the selected solution is very robust, and environmentally reasonable.

Figure S6: Please define “T/B ratio.” It does not seem to be mentioned in the main text.

Reply: We have added the required information in the caption itself of Figure S6.

Figure S6: Timeseries of high-resolution data for showing particular peaks of industrial events. The T/B ratio represents the ratio of concentrations of toluene/ benzene for the period.

References:

Canonaco, F. et al. (2021) ‘A new method for long-term source apportionment with time-dependent factor profiles and uncertainty assessment using SoFi Pro: Application to 1 year of organic aerosol data’, Atmospheric Measurement Techniques, 14(2), pp. 923–943. doi: 10.5194/amt-14-923-2021.

Cappellin, L. et al. (2012) ‘On quantitative determination of volatile organic compound concentrations using proton transfer reaction time-of-flight mass spectrometry’, Environmental Science and Technology, 46(4), pp. 2283–2290. doi: 10.1021/es203985t.

Chattopadhyay, G. et al. (1997) ‘Determination of benzene, toluene and xylene in ambient air of calcutta for three years during winter’, Environmental Technology (United Kingdom), 18(2), pp. 211–218. doi: 10.1080/09593331808616529.

Chauhan, S. K., Saini, N. and Yadav, V. B. (2014) ‘Recent Trends of Volatile Organic Compounds in Ambient Air & Its Health Impacts : a Review’, International Journal For Technological Research In Engineering, 1(8), pp. 667–678.

Duan, J. et al. (2008) ‘Concentration, sources and ozone formation potential of volatile *organic compounds (VOCs) during ozone episode in Beijing', Atmospheric Research, 88(1), pp. 25–35. doi: 10.1016/j.atmosres.2007.09.004.*

Fukusaki, Y. et al. (2021a) 'Source region identification and source apportionment of volatile organic compounds in the Tokyo Bay coastal area, Japan', Atmospheric Environment: X, 9, p. 100103. doi: 10.1016/j.aeaoa.2021.100103.

Fukusaki, Y. et al. (2021b) 'Source region identification and source apportionment of volatile organic compounds in the Tokyo Bay coastal area, Japan', Atmospheric Environment: X, 9, p. 100103. doi: 10.1016/j.aeaoa.2021.100103.

Garg, A., Gupta, N. C. and Tyagi, S. K. (2019) 'Study of seasonal and spatial variability among benzene, toluene, and p-Xylene (BTp-X) in ambient air of Delhi, India', Pollution, 5(1), pp. 135–146. doi: 10.22059/poll.2018.260934.469.

Guo, S. et al. (2022) 'Volatile organic compounds in urban Lhasa: variations, sources, and potential risks', Frontiers in Environmental Science, 10(August), pp. 1–16. doi: 10.3389/fenvs.2022.941100.

Haider, K. M. et al. (2022) 'Chemical identification and quantification of volatile organic compounds emitted by sewage sludge', Science of the Total Environment, 838(May). doi: 10.1016/j.scitotenv.2022.155948.

Hansel, A. et al. (1999) 'Proton-transfer-reaction mass spectrometry (PTR-MS): On-line monitoring of volatile organic compounds at volume mixing ratios of a few pptv', Plasma Sources Science and Technology, 8(2), pp. 332–336. doi: 10.1088/0963-0252/8/2/314.

Hoque, R. R. et al. (2008) 'Spatial and temporal variation of BTEX in the urban atmosphere of Delhi, India', Science of the Total Environment, 392(1), pp. 30–40. doi: 10.1016/j.scitotenv.2007.08.036.

Hui, L. et al. (2018) 'Characteristics, source apportionment and contribution of VOCs to ozone formation in Wuhan, Central China', Atmospheric Environment, 192(February), pp. 55–71. doi: 10.1016/j.atmosenv.2018.08.042.

Jain, V. et al. (2022) 'Seasonal variability and source apportionment of non-methane VOCs using PTR-TOF-MS measurements in Delhi , India', Atmospheric Environment, 283(April), p. 119163. doi: 10.1016/j.atmosenv.2022.119163.

Kumar, A. et al. (2020) 'Non-methane hydrocarbon (NMHC) fingerprints of major urban and agricultural emission sources for use in source apportionment studies', Atmospheric Chemistry and Physics, 20(20), pp. 12133–12152. doi: 10.5194/acp-20-12133-2020.

Lalchandani, V. et al. (2021) 'Real-time characterization and source apportionment of fine particulate matter in the Delhi megacity area during late winter', Science of the Total Environment, 770. doi: 10.1016/j.scitotenv.2021.145324.

Lyu, X. P. et al. (2017) 'Evaluation of the effectiveness of air pollution control measures in Hong Kong', Environmental Pollution, 220, pp. 87–94. doi: 10.1016/j.envpol.2016.09.025.

Majumdar (neé Som), D. et al. (2008) 'Source apportionment of VOCs at the petrol pumps in Kolkata, India; exposure of workers and assessment of associated health risk', Transportation Research Part D: Transport and Environment, 13(8), pp. 524–530. doi: 10.1016/j.trd.2008.09.011.

Majumdar, D., Mukherjeea, A. K. and Sen, S. (2011) 'BTEX in Ambient Air of a Metropolitan

City', Journal of Environmental Protection, 02(01), pp. 11–20. doi: 10.4236/jep.2011.21002.

Paatero, P. and Hopke, P. K. (2003) 'Discarding or downweighting high-noise variables in factor analytic models', Analytica Chimica Acta, 490(1–2), pp. 277–289. doi: 10.1016/S0003-2670(02)01643-4.

Pallavi, Sinha, B. and Sinha, V. (2019) 'Source apportionment of volatile organic compounds in the northwest Indo-Gangetic Plain using a positive matrix factorization model', Atmospheric Chemistry and Physics, 19(24), pp. 15467–15482. doi: 10.5194/acp-19-15467-2019.

Qin, J. et al. (2021) 'Source apportionment of VOCs in a typical medium-sized city in North China Plain and implications on control policy', Journal of Environmental Sciences, 107, pp. 26–37. doi: 10.1016/j.jes.2020.10.005.

Sahu, L. K., Tripathi, N. and Yadav, R. (2017) 'Contribution of biogenic and photochemical sources to ambient VOCs during winter to summer transition at a semi-arid urban site in India', Environmental Pollution, 229, pp. 595–606. doi: 10.1016/j.envpol.2017.06.091.

Sahu, L. K., Yadav, R. and Pal, D. (2016) 'Source identification of VOCs at an urban site of western India: Effect of marathon events and anthropogenic emissions', Journal of Geophysical Research: Atmospheres RESEARCH, 121, pp. 2416–2433. doi: 10.1002/2015JD024454.

Sarkar, C. et al. (2017) 'Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization', Atmospheric Chemistry and Physics, 17(13), pp. 8129–8156. doi: 10.5194/acp-17-8129-2017.

Shukla, A. K. et al. (2021) 'Real-time quantification and source apportionment of fine particulate matter including organics and elements in Delhi during summertime', Atmospheric Environment, 261(November 2020). doi: 10.1016/j.atmosenv.2021.118598.

Sinha, V., Kumar, V. and Sarkar, C. (2014) 'Chemical composition of pre-monsoon air in the Indo-Gangetic Plain measured using a new air quality facility and PTR-MS: High surface ozone and strong influence of biomass burning', Atmospheric Chemistry and Physics, 14(12), pp. 5921–5941. doi: 10.5194/acp-14-5921-2014.

Smith, D. and Spanel, P. (2005) 'Selected ion flow tube mass spectrometry SIFT-MS for on-line trace gas analysis', Mass Spectrometry Reviews, 24, pp. 661–700. doi: 10.1002/mas.20033.

Srivastava, A., Joseph, A. E. and Devotta, S. (2006) 'Volatile organic compounds in ambient air of Mumbai - India', Atmospheric Environment, 40(5), pp. 892–903. doi: 10.1016/j.atmosenv.2005.10.045.

Srivastava, A., Sengupta, B. and Dutta, S. A. (2005) 'Source apportionment of ambient VOCs in Delhi City', Science of the Total Environment, 343(1–3), pp. 207–220. doi: 10.1016/j.scitotenv.2004.10.008.

Steinbacher, M. et al. (2004) 'Performance characteristics of a proton-transfer-reaction mass spectrometer (PTR-MS) derived from laboratory and field measurements', International Journal of Mass Spectrometry, 239(2–3), pp. 117–128. doi: 10.1016/j.ijms.2004.07.015.

Stewart, G. J. et al. (2021) 'Sources of non-methane hydrocarbons in surface air in Delhi, India', Faraday Discussions, 226, pp. 409–431. doi: 10.1039/d0fd00087f.

Tan, Y. et al. (2021) 'Characteristics and source apportionment of volatile organic compounds

(VOCs) at a coastal site in Hong Kong', Science of the Total Environment, 777, p. 146241. doi: 10.1016/j.scitotenv.2021.146241.

Tobler, A. et al. (2020) 'Chemical characterization of PM2.5 and source apportionment of organic aerosol in New Delhi, India', Science of the Total Environment, 745, pp. 1–12. doi: 10.1016/j.scitotenv.2020.140924.

Tripathi, N. et al. (2021) 'Ambient air characteristics of biogenic volatile organic compounds at a tropical evergreen forest site in Central Western Ghats of India', Journal of Atmospheric Chemistry, 78(2), pp. 139–159. doi: 10.1007/s10874-021-09415-y.

Tripathi, N. et al. (2022) 'Characteristics of VOC composition at urban and suburban sites of New Delhi, India in winter', Journal of Geophysical Research: Atmospheres, pp. 1–28. doi: 10.1029/2021jd035342.

Tripathi, N. and Sahu, L. (2020) 'Chemosphere Emissions and atmospheric concentrations of a -pinene at an urban site of India : Role of changes in meteorology', Chemosphere, 256, p. 127071. doi: 10.1016/j.chemosphere.2020.127071.

Wang, G. et al. (2016) 'Characteristics and source apportionment of VOCs in the suburban area of Beijing , China', Atmospheric Pollution Research, 7(4), pp. 711–724. doi: 10.1016/j.apr.2016.03.006.

Wang, L. et al. (2020) 'Source characterization of volatile organic compounds measured by PTR-ToF-MS in Delhi, India', Atmospheric Chemistry and Physics, (February), pp. 1–27. doi: 10.5194/acp-2020-11.

Wang, L. et al. (2021) 'Characteristics of wintertime VOCs in urban Beijing: Composition and source apportionment', Atmospheric Environment: X, 9. doi: 10.1016/j.aeaoa.2020.100100.

Yadav, M. et al. (2019) 'Source apportionment of particulate matter, gaseous pollutants, and volatile organic compounds in a future smart city of India', Urban Climate, 28(April), p. 100470. doi: 10.1016/j.uclim.2019.100470.

Yang, W. et al. (2018) 'Volatile organic compounds at a rural site in Beijing: Influence of temporary emission control and wintertime heating', Atmospheric Chemistry and Physics, 18(17), pp. 12663–12682. doi: 10.5194/acp-18-12663-2018.

Zhan, J. et al. (2021) 'Ozone and SOA formation potential based on photochemical loss of VOCs during the Beijing summer', Environmental Pollution, 285(February), p. 117444. doi: 10.1016/j.envpol.2021.117444.

Zhang, F. et al. (2020) 'Estimation of abatement potentials and costs of air pollution emissions in China', Journal of Environmental Management, 260(January), p. 110069. doi: 10.1016/j.jenvman.2020.110069.

Zheng, J. et al. (2009) 'Speciated VOC emission inventory and spatial patterns of ozone formation potential in the Pearl River Delta, China', Environmental Science and Technology, 43(22), pp. 8580–8586. doi: 10.1021/es901688e.

Zhu, Y. et al. (2016) 'Characteristics of ambient volatile organic compounds and the influence of biomass burning at a rural site in Northern China during summer 2013', Atmospheric Environment, 124, pp. 156–165. doi: 10.1016/j.atmosenv.2015.08.097.