

The Editor,
Atmospheric Chemistry and Physics

Dear Professor Laskin,

Greetings!

We wish to submit for your kind consideration our revised manuscript entitled **“Biomass burning sources control ambient particulate matter but traffic and industrial sources control volatile organic compound emissions and secondary pollutant formation during extreme pollution events in Delhi”** in which we provide a quantitative source apportionment of volatile organic compound and particulate matter emissions for Delhi.

We have carefully addressed all the reviewers' comments, and have made the required changes in our manuscript. We hope you find the revised manuscript acceptable for your esteemed journal.

All the authors declare that there is no conflict of interest in connection with this paper and as the corresponding author; I confirm that this work has been approved by all authors, is original and has not been published elsewhere nor, is it under consideration for publication in any other journal.

Yours sincerely,
Baerbel Sinha, on behalf of co-authors

Response to reviewer 1

Referee 1:

The manuscript presents a positive matrix factorization analysis of a consequent PTR-ToF-MS dataset, to which PM_{2.5} and PM₁₀ data were added. The general outline, scope and main conclusions are very clear. The results are interesting, and each of the 11 obtained factors is thoroughly described, backed up with external data and source profiles, and well explained.

We appreciate the referee for acknowledging the significance and content of the work, and for considering it of great interest to ACP readers.

However, I do feel that the methodology is not described enough. There should be more details on how the uncertainties were calculated. What were the uncertainties for each compound and their range?

We thank the referee for the suggestion. The overall uncertainty of each compound comprises of two components, the accuracy error and the precision error. The accuracy error was minimized with the help of 8 span calibrations using a certified calibration gas standard (Societa Italiana Acetilene E Derviat; S.I.A.D. S.p.A., Italy) that had 11 hydrocarbons at ~100 ppb, namely methanol, acetonitrile, acetone, isoprene, benzene, toluene, xylene, trimethylbenzene, and dichlorobenzene and trichlorobenzene. Additionally, to ensure accurate mass axis calibration for every acquired spectra, an internal standard namely 1,3-di-iodobenzene (C₆H₅I₂⁺) detected at m/z 330.848 and its fragment ion [C₆H₅I⁺] detected at m/z 204.943 were co-injected with ambient air. The technical details of these calibrations have been discussed in greater detail in the companion paper egosphere-2024-500 and calibration plots and transmission curve can be found there. However, for the purpose of PMF runs only the random uncertainty, that is the precision error, should be included as uncertainty in the PMF. If the systematic error is accidentally included the $Q_{\text{true}}/Q_{\text{theoretical}}$ ratio can drop below 1 even for a 3-factor solution. The precision error for each m/z was calculated from the observed count rate in counts per second (cps) using Poisson statistics. This is a routine way to report the precision error of measurements recorded by systems such as electron multipliers or multichannel plates. For entering the precision error into the PMF we used the average signal in cps of the m/z for the study period to calculate the average precision error in %. We have added the following text to lines 146ff in the main manuscript:

“The accuracy error was minimized by conducting a total of 8 span calibrations throughout the study period. The details of these calibrations can be found in Mishra et al., 2024. The precision error for each m/z listed in table S1, which needs to be included into the PMF model runs, was calculated from the average observed count rate in counts per second (cps) of each m/z with the help of Poisson statistics. The detection limit was determined as 2σ of the noise observed in clean zero air.”

We also have included the precision error and detection limit used in our model runs in the supplementary Table S1.

Also, more information is needed on the PMF approach of adding PM_{2.5} and PM₁₀ data to the VOC dataset.

While PMF has routinely been used to source apportion non-methane volatile organic compounds (NMVOCs) in the literature for a quite some time, several authors have recently pioneered the use of NM-VOC tracers in a PMF to source apportion greenhouse gases such as methane, CO₂ and N₂O (Guha, et al. 2015, Assan et al. 2018, Schulze et al. 2018) by making use of the fact that the VOCs source-fingerprints of many combustion sources are well constrained and understood. We now extend the use of this promising new technique towards source-apportionment of PM_{2.5} and PM₁₀. The PMF is a matrix decomposition factor analysis model that deconvolves a time series of measured species into a set of factors with fixed source fingerprints whose contributions to the input data set varies with time. This makes the model well suited to accommodate all chemical species co-emitted from the same combustion source as long as the emissions impacting the receptor site are fresh enough for the VOC fingerprint to be preserved. We have included the following text and new references in our method section and list of references, respectively:

“Several authors have recently pioneered the use of VOC tracers in a PMF to source apportion co-emitted greenhouse gasses such as methane, CO₂ and N₂O (Guha, et al. 2015, Assan et al. 2018, Schulze et al. 2023). Since the VOCs source-fingerprints of many combustion sources are well constrained and understood, we now extend the use of this promising new technique towards source-apportionment of co-emitted PM_{2.5} and PM₁₀. The PMF is a matrix decomposition factor analysis model that decomposes a time series of measured species into a set of factors with fixed source fingerprints whose contributions to the input data set varies with time. This makes the model well suited to accommodate all chemical species co-emitted from the same source.”

Assan, S., Vogel, F.R., Gros, V., Baudic, A., Stauffer, J., Ciais, P.: Can we separate industrial CH₄ emission sources from atmospheric observations? - A test case for carbon isotopes, PMF and enhanced APCA, Atmospheric Environment, 187, 317-327, <https://doi.org/10.1016/j.atmosenv.2018.05.004>, 2018.

Guha, A., Gentner, D. R., Weber, R. J., Provencal, R., and Goldstein, A. H.: Source apportionment of methane and nitrous oxide in California's San Joaquin Valley at CalNex 2010 via positive matrix factorization, *Atmos. Chem. Phys.*, 15, 12043–12063, <https://doi.org/10.5194/acp-15-12043-2015>, 2015.

Schulze, B., Ward, R. X., Pfannerstill, E. Y., Zhu, Q., Arata, C., Place, B., Nussbaumer, C., Wooldridge, P., Woods, R., Bucholtz, A., Cohen, R. C., Goldstein, A. H., Wennberg, P. O., Seinfeld, J. H.: Methane Emissions from Dairy Operations in California's San Joaquin Valley Evaluated Using Airborne Flux Measurements, *Environ. Sci. Technol.* 2023, 57, 48, 19519–19531, <https://doi.org/10.1021/acs.est.3c03940>, 2023.

What were the steps leading to the solution (how many runs, how was the base case chosen if these runs gave different solutions.

The following text was added to the manuscript:

“The model was initiated for 20 base runs with the recommended block size of 379, and the run with the lowest Q_{robust} and Q_{true} was chosen for further analysis and display in Figure 2.”

Were there any challenges with this approach?

No there are no challenges with this approach. Once the number of factors in the PMF approaches the true number of major sources, the PMF output becomes very stable with minimal differences between the different base runs and even minimal differences in the factor time series and percentage of VOCs explained by individual factors.

In the abstract you mention “our novel source apportionment method”, but it is not very clear in the paper how novel or different it is.

The approach of source-apportioning PM sources with the help of high-time resolution measurements and better understood VOC tracers instead of highly fragmented AMS mass spectra or low time resolution offline aerosol samples is novel. We have revised following text in line 45ff to highlight the novelty:

“Several authors have recently pioneered the use of VOC tracers in a PMF to source apportion co-emitted greenhouse gasses such as methane, CO₂ and N₂O (Guha, et al. 2015, Assan et al. 2018, Schulze et al. 2023). Since the VOCs source-fingerprints of many combustion sources are well constrained and understood, we now extend the use of this promising new technique towards source-apportionment of co-emitted PM_{2.5} and PM₁₀. The PMF is a matrix decomposition factor analysis model that decomposes a time series of measured species into a set of factors with fixed source fingerprints whose contributions to the input data set varies with time. This makes the model well suited to accommodate all chemical species co-emitted from the same source.”

Also, you mention that the factors are stable in the bootstrap repetitions; however, the uncertainties of the model in Figure 3 seem quite important.

As noted in the figure caption, we have plotted the 2 σ uncertainty of the model in Figure 3. Hence the error bars may look-worse than they are because many authors typically report 1 σ error bars or fail to include any error bars in their factor profiles. When the PMF model actually behaves in an unstable manner, it is typical to see uncertainties in excess of 100% of the mass assigned, due to factor swapping during bootstrap runs. The average uncertainty of VOCs that are present with a loading of >1 $\mu\text{g m}^{-3}$ in factor profile in our PMF runs is only 25%. The largest uncertainty bars belonged to source fingerprints which in some cases have a large vehicle-to-vehicle and fire-to-fire variability of the emission factor for certain compounds. However, upon introspection in response to the reviewer's comment, we note that it may be more appropriate to report the error of the mean of the emission factor, rather than the vehicle-to-vehicle and fire-to-fire variability as uncertainty in this figure, and we have updated the figure accordingly:

emissions have their own independent factor profile in the PMF, the amount of PM attributed to paddy residue burning and the VOC source fingerprint of the source become stable in the PMF solution. The amount attributed to residential heating and waste burning stabilizes after a separate factor for cooking emissions opens up in the 9-factor solution. The following text was inserted into line 174 to make this clearer:

“Until the PMF opens distinct factors for the industrial OVOC emissions in the 7-factor solution, the partitioning between paddy residue burning and heating and waste burning $PM_{2.5}$ and PM_{10} emissions in the model remains unstable, because these sources with their strong OVOC emissions are most agreeable to accommodating additional OVOC sources in their fingerprint at the expense of explaining the $PM_{2.5}$ and PM_{10} emissions. Once the industrial OVOC emissions have their own factor, this split becomes stable. The amount of PM attributed to residential heating and waste burning stabilizes after a separate factor for cooking emissions opens up in the 9-factor solution.”

Do you have other information to back up the factors' stability (i.e., low time-series correlations between the factors)?

Yes. Out of 55 possible factor pairs 51 factor pairs have an $R < 0.5$ and 49 $R < 0.4$ while 4 have an R between 0.5 and 0.6. No pair displays an $R > 0.6$. We also have additional support to back up that these two factors are distinct and real. The intensity of the paddy burning factor correlates with the same day fire counts in the 24-h fetch region ($R = 0.8$) and the burning decreases by the time wheat sowing is almost complete. The heating and waste disposal factor keeps on increasing proportional to the increase in heating demand towards the onset of winters and it shows an $R = 0.8$ with the 24-h averaged heating demand. Their time series correlation displays an $R = 0.5$ on account of both activities being negligible in monsoon and active in post monsoon season and both being most active in the early evening hours. The time series correlation of hourly averaged data is not necessarily a highly diagnostic tool that can be used in isolation to identify whether or not factors are genuine, as R values in the range of 0.5-0.6 can be accomplished merely because two sources share the same diurnal pattern such as high concentration values at night when emissions mix into a shallow nocturnal boundary layer and lower values during the day when the boundary layer is well mixed. This is particularly true if one of the two sources is as ubiquitous as traffic. The highest R values for any factor pairs in our 11-factor solution occur for the correlation between the industrial source and two traffic sources. It displays $R = 0.59$ and 0.56 with 4-wheelers and 2-wheelers, respectively. This happens despite the fact that the industrial source emissions primarily reach the receptor from what appears to be point sources located in the wind sector SE to SW of the site, while both 4-wheeler and 2-wheeler emissions reach the site every night and from all wind directions. 4-wheeler and 2-wheeler also show $R = 0.51$ with each other because both type of emission occur simultaneously on the same roads. The following text was modified in line 180:

“Therefore, the 11-factor solution, which showed $R < 0.6$ for all possible factor pairs, was analyzed further.”

How do the scaled residuals change when increasing the number of factors or between different runs?

The scaled residual outside the -3σ to $+3\sigma$ range decreases in an exponential pattern with the increase in the number of factors. Since this is a large dataset, their number is still large (10^3 observations) in the 11-factor solution. This information has been added to Fig. S5 and the figure is referenced as in the text as follows:

“Figure S5 shows how the $Q_{\text{true}}/Q_{\text{theoretical}}$ ratio and $Q_{\text{robust}}/Q_{\text{theoretical}}$, and scaled residuals beyond 3 standard deviations drop exponentially when the number of factors increases. It can be seen that initially the $Q_{\text{true}}/Q_{\text{theoretical}}$ ratio drops faster than $Q_{\text{robust}}/Q_{\text{theoretical}}$ ratio on account of additional major plumes being better explained with each additional factor. However, with the increase from 11 to 12 factors both drop in a parallel fashion indicating that the point of diminishing returns has been reached.”

The comparison of the PMF output with emission inventory results needs more justification. If I understand correctly, PMF results are concentrations and seem to be directly compared to emissions, which are in different quantities and on different scales.

While evaluating the percentage contribution of different sources to the burden of specific pollutants such as $PM_{2.5}$ over a fetch region that is reasonable and related to the atmospheric lifetime of the pollutant in question, the comparison can be considered valid. After all, the lifetime for any given VOC such as benzene is independent of its source. Hence the percentage share each source contributes to the measured burden at a site should be proportional to the percentage share the different sources within the fetch region contribute to emissions, provided that the emissions are correctly represented in the emission inventory and the fetch region is chosen suitably small to ensure that emissions from a source within the fetch region can reach the receptor without significant loss. In this study, we are not comparing the absolute concentrations of the PMF and emission inventories, but rather a relative percentage contribution of sources to the total burden. This approach has been routinely used at many other sites of the world (e.g. Buzcu-Guven and Fraser, 2008 <https://doi.org/10.1016/j.atmosenv.2008.02.025>, Morino et al. 2011, <https://doi.org/10.1029/2010JD014762> Li et. al., 2019 <https://doi.org/10.5194/acp-19-5905-2019>; Qin et al., 2022 <https://doi.org/10.1007/s11356-022-19145-7>,) to compare PMF outputs with emission

inventories. The reason why absolute concentrations are also sometimes brought into the discussion is that at times the look of pie charts can be deceptive, as is the case e.g. for industrial PM_{2.5} emissions. Both the EDGAR and the REAS inventory have almost identical industrial PM_{2.5} emissions in the inventory, yet the pie charts look visibly different, because the larger energy sector emissions and the presence of agricultural burning emissions in the EDGAR inventory visually shrink the size of that “pie slice” compared to how it looks like for the REAS inventory. Looking at the absolute numbers helps to resolve which inventory is more likely to be wrong and more importantly for which source. To address the reviewer’s concerns, we have now reworded section 2.5 in the Materials and Methods section and have added more justification and references. It now reads as follows:

“The observational data was grouped according to the predominant airflow into a south-westerly, north-westerly, and south-easterly group, and the fetch region from which air masses would reach the receptor site within 24 h was determined for each group separately spanning latitude 21–31°N and longitude 72–82°E, latitude 28–32°N and longitude 72–80°E and latitude 25–30°N and longitude 75–88°E, respectively, for the three flow regimes. Two gridded emission inventories namely the Emission Database for Global Atmospheric Research (EDGARv6.1) for the year 2018 (Crippa et al., 2022), and the Regional Emission inventory in Asia (REAS v3.2.1) for the year 2015 (Kurokawa & Ohara, 2020) were filtered for these three fetch regions to compare PMF results with the emission inventory. We compare the relative percentage contribution of sources to the total atmospheric pollution burden in the PMF with the relative percentage contribution of sources to the total emissions for the emission inventories. This approach has been routinely used to evaluate emission inventories with the help of PMF results at different sites around the world (Buzcu-Guven and Fraser, 2008, Morino et al. 2011, Sarkar et al., 2017; Li et al., 2019, Qin et al., 2022). For the purpose of emission inventory comparison of anthropogenic sources, natural sources such as biogenic emissions and the photochemistry factor were removed from the PMF output, while the solid fuel-based cooking and residential heating and waste burning emissions were summed up in residential & waste management. In addition, CNG and Petrol 2 & 4-wheeler factors were combined into the consolidated transport sector emissions.”

The conclusions drawn here seem too strong (i.e. lines 536-539). Also, please justify why the PMF results are more correct than model outputs? (i.e. when you state that sources are under-/over-estimated in the models).

The PMF results are based on the primary data acquired at the airshed site. On the other hand, emission inventories rely on activity data and emission factors that is often lags behind in terms of updation by few years and therefore less well constrained. E.g. residential emissions are at times simply scaled with the increase in the population without adjusting for fuel shifts, while transport sector emissions may be scaled with fuel sales without accounting for the shift to lower emission control technologies (e.g. Euro-6=BS-VI). Due to this, routine updates often fail to encompass the technological advancements as well as measures effected by the policy change in a particular region. Hence, we assume that the results arising from the direct ambient measurements are closer to the reality of 2022 than emission inventories that have last been updated in 2015. However, we appreciate the reviewer’s valid comment and now support each of the points we are making with more references to supporting literature with similar findings. In lines 536-539 the text now reads as follows

“Table S6 shows that for residential sector VOCs emissions, the absolute emissions in the EDGARv6.1 inventory are almost twice as large as those in the REASv3.2.1 inventory, even though the percentage contribution of this sector to the VOC emissions in the inventory in Figure 10 appears to be similar for both, because of larger VOC emissions from solvent use and industries in the EDGARv6.1 inventory. Both inventories overestimate the relative importance of residential sector emissions in relation to VOC emissions from other sectors by more than a factor of two when compared to our PMF estimate, most likely because they have not been updated with recent fuel shifts towards LPG in the relatively prosperous Delhi NCR region (Sharma et al., 2022).”

Lines 550-560 containing statements about the agricultural sector emissions in various inventories have also been revised as follows including by addition of new Figure S10 showing coarse mode aerosol from the use of paddy in an industrial burner:

“The EDGARv6.1 inventory significantly underestimates PM_{2.5} & PM₁₀ emissions from agricultural activities, which include, but are not limited to crop residue burning, in comparison to our PMF results, particularly over NW-India (Table S6). Over this fetch region EDGARv6.1 attributes as much PM_{2.5} to all agricultural activities combined for the full year as the FINNv2.5 inventory (Wiedinmyer et al., 2023) attributes just to agricultural residue burning activities taking place between 15th August and 26th November 2021 (a time period comparable to the period in our model run), without including the emissions from rabi crop residue burning in summer (Kumar et al., 2016) and other agricultural activities such as harvest and ploughing. For PM₁₀ the fire count based FINNv2.5 estimate is twice as high as the emission estimate of EDGARv6.1 for this fetch region, and more likely to be correct, because the phytoliths present in rice straw form coarse mode ash during the combustion process (Figure S10). The fact that EDGAR appears to underestimate residue-burning emissions over this fetch region has been flagged earlier (Pallavi et al., 2019; Kumar et al., 2021; Singh et al., 2023). Our PMF analyses also reveals that the relative contribution of agricultural residue burning to the PM burden over

the North-Western IGP (24 % and 27 % of $PM_{2.5}$ and PM_{10} , respectively) and South-Eastern IGP (24 % and 27 % of $PM_{2.5}$ and PM_{10} , respectively) is comparable, despite the much lower fire counts over the South-Eastern IGP (17,810), when compared to the North Western IGP (61,334). This indicates that either fires to the SE are burning closer to the receptor site or the fire detection efficiency in this fetch region is lower. Table S6 reveals that the relative importance of agricultural emissions over the SE fetch region is even more severely underestimated in the FINNv2.5 inventory than in the EDGARv6.1 inventory due to poorer fire detection (close to 100% omission error) for the partial burns prevalent over this region (Lui et al. 2019, 2020, Figure S8) when compared to the complete burns prevalent over the NW IGP (Lui et al. 2019, 2020, Figure S7).”

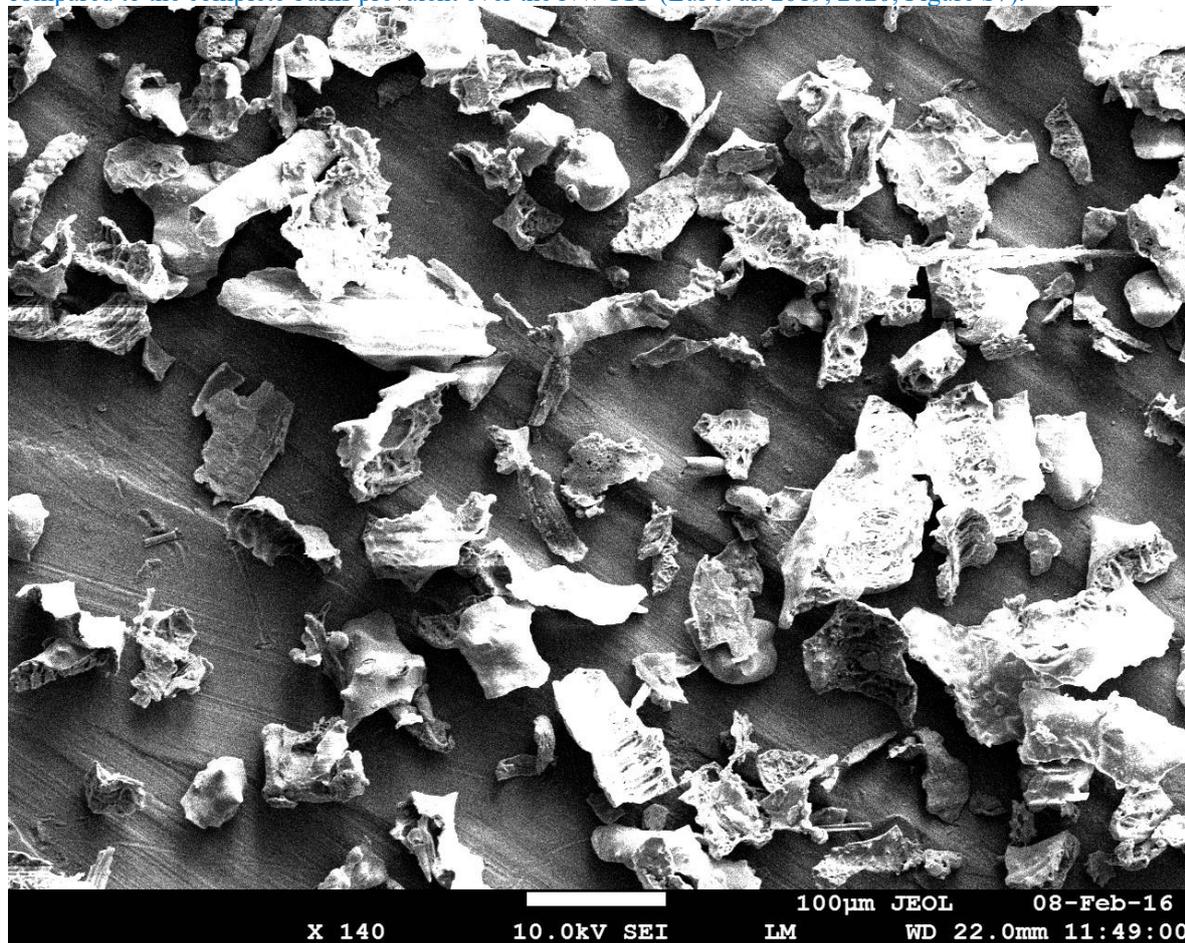


Figure S10 SEM image of rice ash from the electrostatic precipitator of an industrial boiler fired with rice husk and straw illustrating the coarse mode nature of the ash generated during the combustion of phytolith containing biomass.

Be more concise when you present the description of the factors, the fact that all the values and VOC m/z are written in the main text makes it tedious to read. Use only VOC names (or formula if unclear what the compound is, but the m/z are already all listed in Table S1). Also, delete all the concentration and % values in the main text if they are already on the figures, except if it is useful to emphasize the point (example in line 632: “a considerable portion of the PM_{10} (18%) and $PM_{2.5}$ (28%)”). Same for $\log_{10}C_0$, find a clearer way to present them. Another option would be to put the extensive description of factors in SI and a summary and interpretation in the main text.

We appreciate the referee’s suggestion and have made the requested changes to the manuscript. Since the number of changes is large we are not listing each individual one into the response file. However, we chose to retain the factor descriptions in the main text. The $\log_{10}C_0$ values are now presented in a new supplementary figure (Figure S9) that also helps to address the reviewer’s comment regarding SOA formation from these factors.

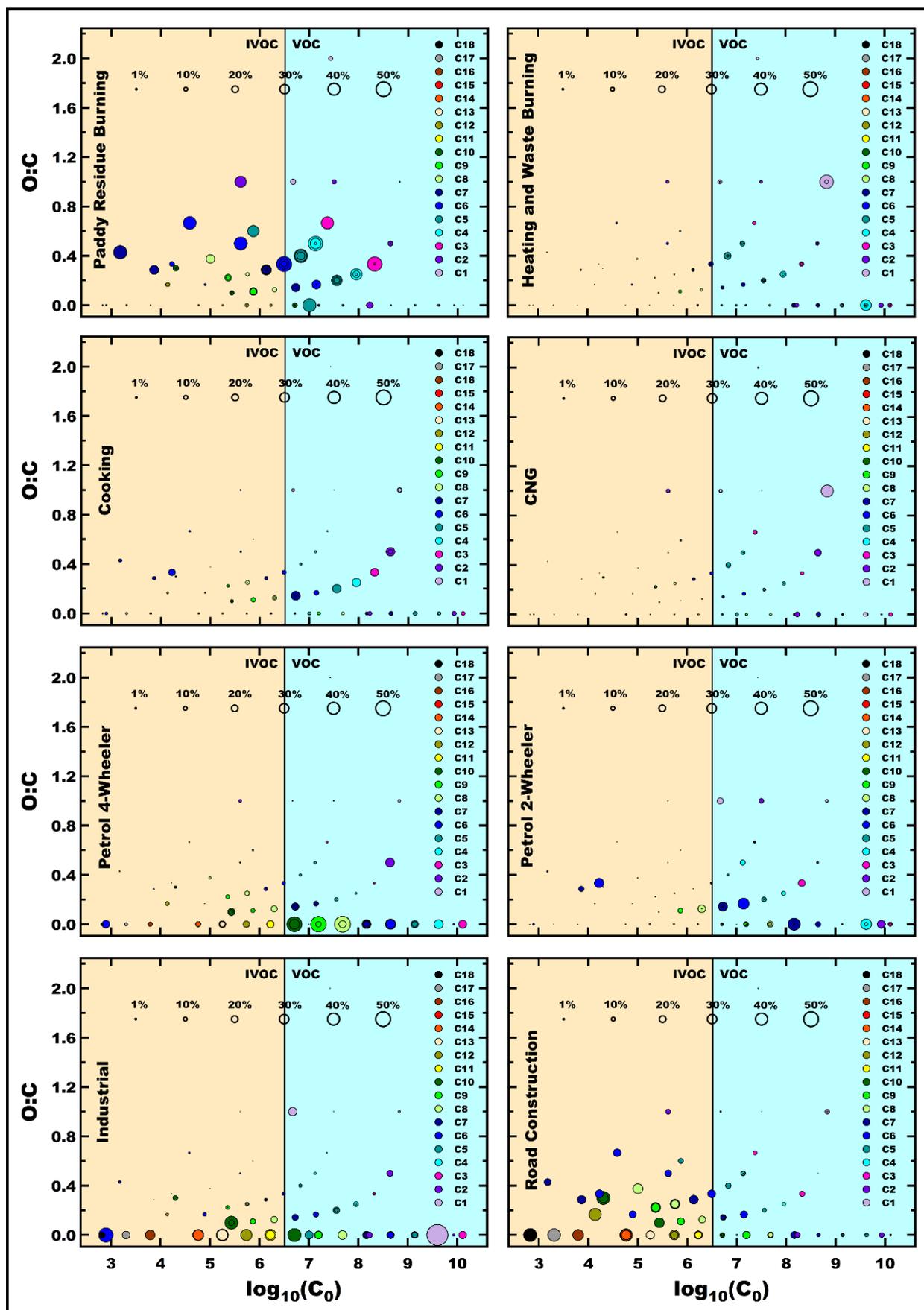


Figure S9: Volatility oxidation state plots for all factors that individually contribute more than 3% to the total SOA formation potential.

Specific comments/questions

Line 87-88: I would suggest adding a map of the receptor site with the surroundings (i.e. roads, industries, agriculture...), and referencing it when needed.

We have added a detailed map of the receptor site with its surroundings to the supplement as Figure S1 and have referenced it in the text as follows:

Ambient air was sampled into the instruments from the roof-top of a tall building (28.5896°N - 77.2210°E) at ~35 m above ground, located within the premises of the Indian Meteorological Department (IMD) at Lodhi Road, New Delhi situated in Central Delhi. The sampling site is a typical urban area surrounded by green spaces, government offices, and residential areas, but not in the direct vicinity of any major industries (Fig. S1)

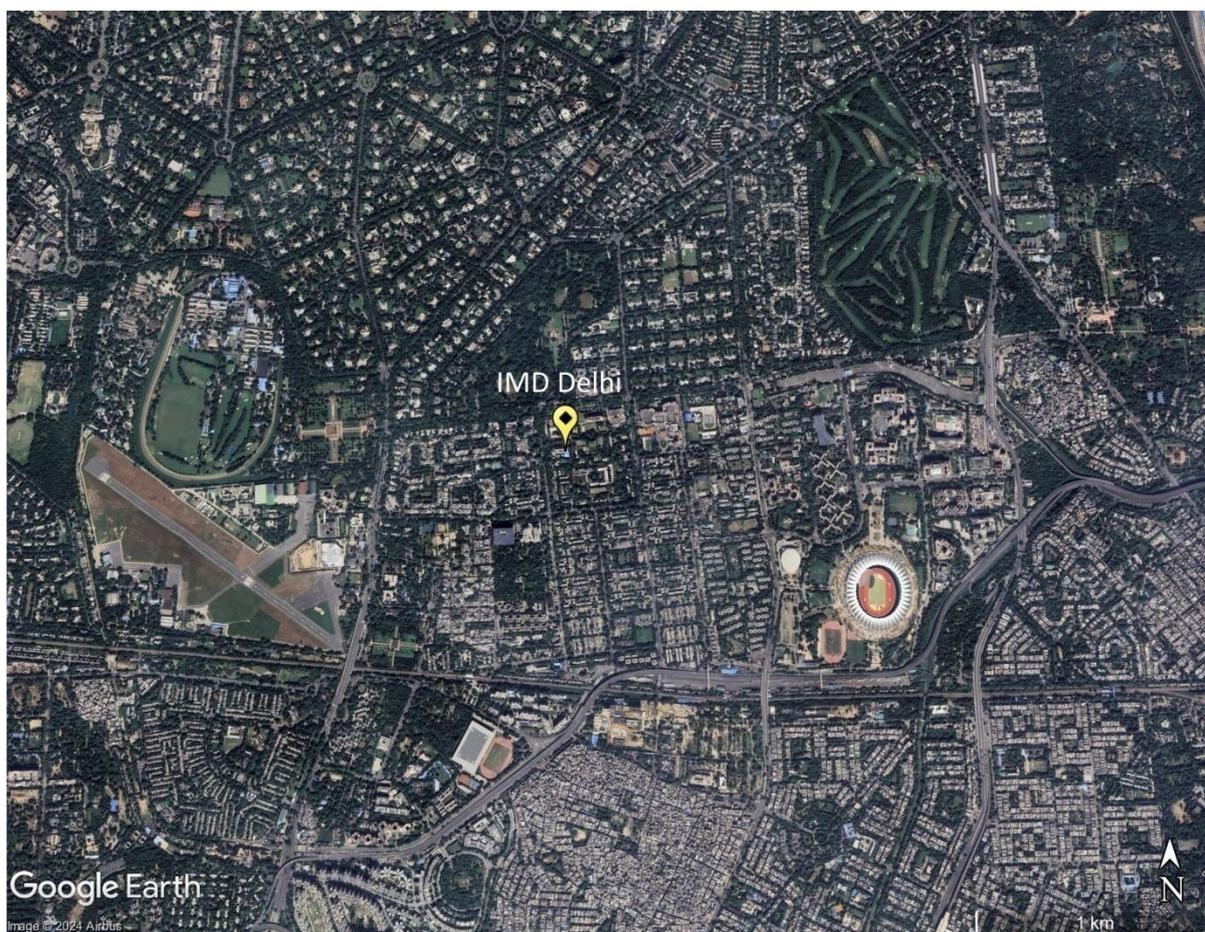


Figure S1: Map of the immediate surroundings of the IMD (28.5896°N - 77.2210°E) sampling site in Central Delhi. (Google Earth Imagery ©Google Earth)

Line 108: I think it would be worth summarizing the main differences between the 3 wind sectors (in terms of typology, specificity, and later on results).

We appreciate the referee's suggestions, hence a map has been added to Figure 1 and we have shifted the description from section 3.3 to section 2.1. The relevant sentences now read as follows:

“Figure 1 shows the location of the site and the 120 h back trajectories of air masses arriving at the site that were grouped according to the dominant synoptic regional scale transport into a) south-westerly (orange and yellow) flows carrying emissions from southern Punjab, Haryana, Uttar Pradesh, Madhya Pradesh, Rajasthan and Gujarat towards the receptor, b) north-westerly (light and dark blue) flows carrying emissions from Pakistan Punjab, Indian Punjab, Haryana, Western Uttar Pradesh, Himachal Pradesh, and Uttarakhand towards the receptor, and c) south-easterly flows (light and dark red) carrying emissions from Haryana, Southern Uttarakhand, Uttar Pradesh, Bihar and Nepal towards the receptor. Figure 1d shows a Google Earth image with a spatial map of the daily fire counts in the region for the post-monsoon season alongside with the maximum 24-h fetch region for each of these synoptic flow situations marked by coloured square. Figure 1e-h shows the e) photosynthetic active radiation, f) daily fire counts in the fetch region (21 - 32°N , 72 - 88°E), g) temperature and relative humidity, and h) the ventilation coefficient and the sum of the daily rainfall during the study period (15th August 2022– 26th November 2022).”

Section 3.1 & Figure 3: I would suggest putting Figure 3 in supplementary and replacing it with only this study's factors profiles in concentration (instead of normalized).

We are now displaying the factor profiles in concentration units on the left-hand axis of Figure 3 in the main text. However, we have chosen to retain the visual comparison with the source profiles. To do so we have shifted towards showing the source profiles on a secondary axis which continues to be normalized, because many individual panels have mixed units (e.g. samples from a traffic junction in the units $\mu\text{g}/\text{m}^3$ and tailpipe exhaust with the units g/kg). It is important to note normalization does not alter the fingerprint of the PMF output and does not affect the R of the cross-correlation analysis between source samples and PMF output either. It just permits us to easily combine things in different units and sources with different absolute emission intensity into one plot. However, we assume that the spirit of the reviewers' suggestion is related to the fact that the figure is a little congested. Hence, we reduced the number of source profiles shown in addition to the PMF fingerprint to at most 3 per panel to reduce the congestion in this figure. The revised Figure 3 looks as follows:

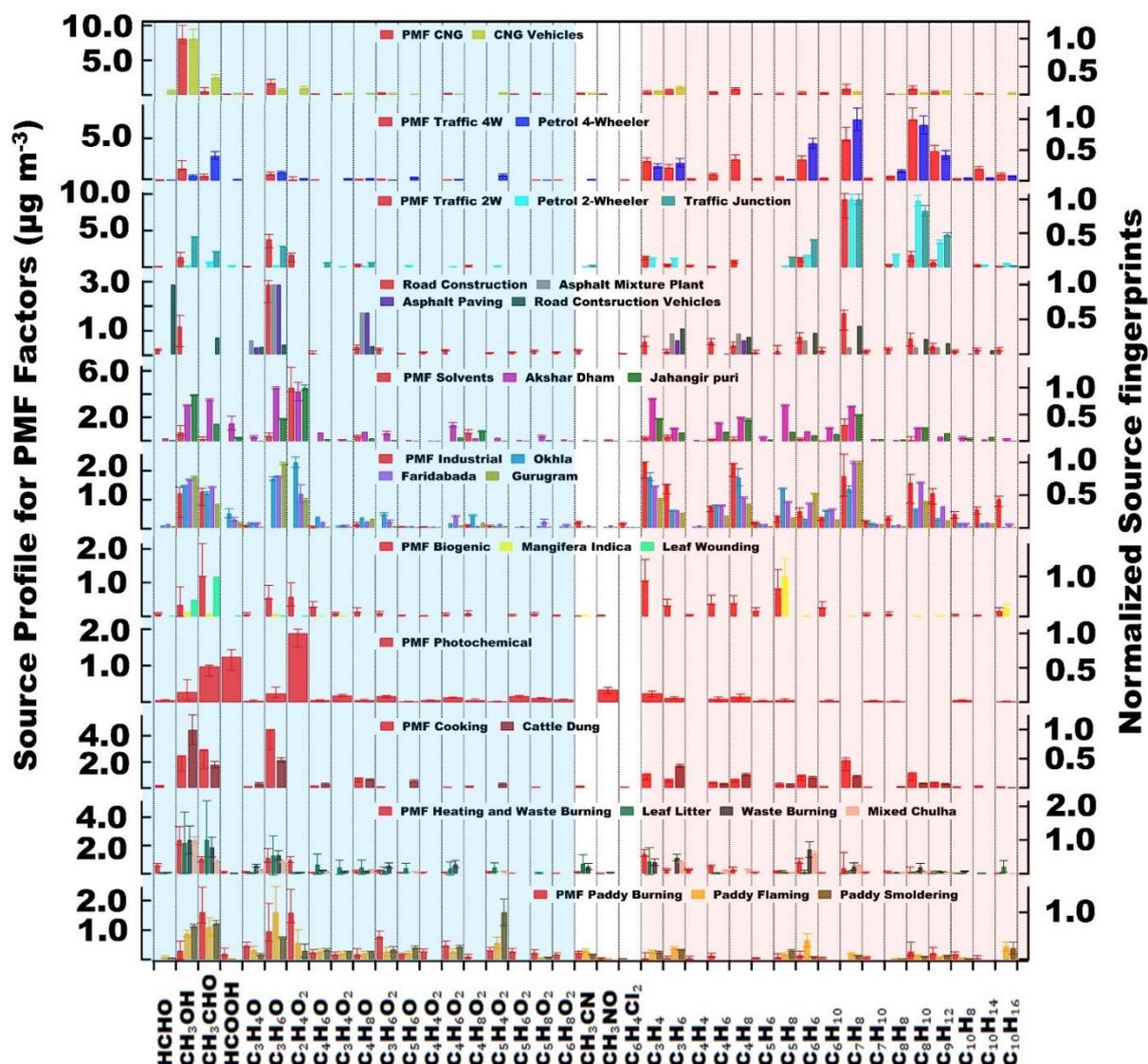


Figure 2: PMF factor profile of the 11 factors identified. The source profile in $\mu\text{g m}^{-3}$ (left in red) and the normalized source fingerprint of grab samples collected at the source (right in various colours). The Error bars indicate the 2σ uncertainty range from the bootstrap runs for PMF factor profiles and the 1σ error of the mean of the emission factors for source samples.

In text 3.1, I would add the R correlation (of profile and/or diurnal cycle) of this study's factors with the mentioned reference factors to justify the factors' interpretation.

We thank the reviewer for this suggestion and have added the R of the correlation of the source profiles of PMF output with the source fingerprints of the source samples in the text in this section. The revised text now reads as follows.

“Figure 3 shows the source profile of the eleven factors that our PMF analyses resolved. Out of the 111 VOCs only those whose normalized source contribution exceeded 0.1 when divided by the most abundant compound in the same source profile in at least one of the sources, were included in the figure. The source identity of the PMF factors was confirmed by matching the PMF factor profiles with the unit $\mu\text{g m}^{-3}$ with normalized source fingerprints of grab samples collected from the potential sources. To facilitate the comparison of emission factors and grab samples from different studies with the PMF output, the source samples were normalized by dividing each species’ mass/emission factor by the mass/emission factor of the most abundant species in a given fingerprint. The PMF factor profile matched best against source samples collected from burning paddy fields ($R=0.6$, Kumar et al., 2020) for the paddy residue burning factor. The cooking factor matched emissions from a cow-dung-fired traditional stove called angithi ($R=0.7$, Fleming et al., 2018). The residential heating & waste burning factor had a source fingerprint matching emission from leaf litter burning, ($R=0.7$, Chaudhary et al., 2022), waste burning ($R=0.7$, Sharma et al., 2022), and cooking on a chulha fired with a mixture of firewood and cow dung ($R=0.9$, Fleming et al., 2018). The factors identified as CNG ($R=1.0$), petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.6$) matched tailpipe emissions of the respective vehicle types and fuels (Hakkim et al., 2021). The petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.7$) also matched traffic junction grab samples from Delhi (Chandra et al., 2018). The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving ($R=0.9$, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles ($R=0.6$, Che et al., 2023). The factors identified as solvent usage and evaporative emissions matched ambient air grab samples collected from an industrial area at Jahangirpuri ($R=0.7$), and Dhobighat at Akshar Dham ($R=0.5$) in this study. The factor identified as industrial emissions showed the greatest similarity to ambient air grab samples from the vicinity of the Okhla waste-to-energy plant ($R=0.8$), Gurugram ($R=0.7$) and Faridabad ($R=0.8$) industrial area. The biogenic factor showed the greatest similarity to leaf wounding compounds released from *Populus tremula* ($R=0.8$, Portillo-Estrada et al., 2015) as well as BVOC fluxes from *Mangifera indica* ($R=0.4$, Datta et al., 2021).”

Sections 3.1 & 3.2: Since you have a dedicated subsection for the comparison of the sources with references, you don’t have to repeat them when describing each factor.

Thank you for the helpful suggestion. We have significantly shortened the text of section 3.1 lines 246-254 and where appropriate the description of individual factors and are now avoiding repetition of text and numbers between section 3.1 and 3.2.:

Figure 4 shows the relative contribution of different sources to the total pollution burden of VOCs, $\text{PM}_{2.5}$ and PM_{10} at the receptor site. In the megacity of Delhi, transport sector sources contributed most ($42\pm 4\%$) to the total VOC burden, while it contributed much less (only 24 %) to the total VOC burden in Mohali a suburban site 250 km North of Delhi during the same season (Singh et al., 2023). On the other hand, the contribution of paddy residue burning ($6\pm 2\%$) and the summed residential sector emissions ($17\pm 3\%$ in Delhi and 18 % in Mohali) to the total VOC burden during post-monsoon season were similar at both sites. The contribution of the different factors to the SOA formation potential (Fig. 4e), stands in stark contrast to their contribution to primary particulate matter emissions. SOA formation potential was dominated by the transport sector (54 %) while direct PM_{10} (52%) and $\text{PM}_{2.5}$ (48%) emissions were dominated by different biomass burning sources (Fig. 4 b & c). CNG-fuelled vehicles also contribute significantly to the PM_{10} ($15\pm 3\%$) and $\text{PM}_{2.5}$ ($11\pm 3\%$) burden. A significant share of the PM_{10} (18 %) and $\text{PM}_{2.5}$ (28 %) burden is associated with the residual and not directly linked to combustion tracers. This share can likely be attributed to windblown dust arriving at the site through long-range transport (Pawar et al., 2015) and to secondary organic, and secondary inorganic aerosols such as ammonium sulphate and ammonium nitrate. Due to the complex relationship of secondary aerosol with gas-phase precursors and emission tracers, VOC tracers are not a suitable tool to source-apportion this aerosol component. Meteorological conditions, homogeneous, heterogeneous, and multiphase chemistry control how fast primary emissions are converted to secondary aerosol. To explain the source of those species, one also needs to invoke the physicochemical and thermodynamical properties of the aerosol. (Acharja et al., 2022).

Figure 3: How were the displayed compounds chosen for this graph? And please use the compounds’ names so that it is clearer.

We display all compounds whose normalized mass is at least 0.1 in at least one of the factor profiles to limit the number of species displayed and keep the figure legible. We have now included this information in the text. We prefer not to name compounds, since particularly at the higher m/z there can be many different chemical compounds with the same monoisotopic mass. Hence, we felt it better to consistently use the chemical formula in the figures. We are discussing names alongside the chemical formula where appropriate in the text of section 3.2,

but a figure x-axis is not the appropriate place accommodate a differentiated discussion of possible names. Hence, we retain the chemical formula instead. The revised text now reads:

“Figure 3 shows the source profile of the eleven factors that our PMF analyses resolved. Out of the 111 VOCs only those whose normalized source contribution exceeded 0.1 when divided by the most abundant compound in the same source profile in at least one of the sources, were included in the figure.”

I would suggest adding Figure S3 in the main text as it is referenced a lot, and that way you don't need to put the % in the main text.

Thank you for the kind suggestion. In accordance with the referee's suggestion, we have removed the percentages from the text, and added Figure S3 to the main text. In response to the editor's comments on our manuscript, we have converted each panel to a separate Figure which is now being referenced as Figure 6 to Figure 9. The revised text segments read as follows:

Section 3.2.1

“Figure 6 shows that this factor explained the largest percentage share of O-heteroarene compounds such as furfural ($C_5H_4O_2$), methyl furfural ($C_6H_6O_2$), hydroxy methyl furfural ($C_6H_6O_3$), furanone ($C_4H_4O_2$), hydroxymethyl furanone ($C_5H_6O_3$), furfuryl alcohol ($C_5H_6O_2$), furan (C_4H_4O), methyl furans (C_5H_6O), C2-substituted furans (C_6H_8O), and C3-substituted furans ($C_7H_{10}O$), which are produced by the pyrolysis of cellulose and hemicellulose, and have previously been detected in biomass burning samples (Coggon et al., 2019; Hatch et al., 2015; 2017; Koss et al., 2018; Stockwell et al., 2015). Figure 6 also shows that this factor explains the largest share of the most abundant oxidation products that result from the nitrate radical-initiated oxidation of toluene as well as from OH-initiated oxidation of aromatic compounds under high NO_x conditions, namely nitrotoluene ($C_7H_7NO_2$) and nitrocresols ($C_7H_7NO_3$) (Ramasamy et al., 2019), which indicates a certain degree of aging of the plumes. These nitroaromatic compounds are significant contributors to SOA and BrC, (Palm et al., 2020, Harrison et al., 2005). It also explains several other nitrogen containing VOCs such as nitroethane ($C_2H_6NO_2$), the biomass burning tracer acetonitrile (CH_3CN) and pentanenitrile (C_5H_9N). The presence of pentanenitrile isomers in biomass burning smoke has previously been confirmed using gas chromatography-based studies (Hatch et al., 2015, Hatch et al., 2017). In addition the factor explains the largest percentage share of acrolein (C_3H_4O), hydroxyacetone ($C_3H_6O_2$), cyclopentadienone (C_5H_4O), cyclopentanone (C_5H_8O), diketone ($C_4H_6O_2$), pentanedione ($C_5H_8O_2$), hydroxybenzaldehyde ($C_7H_6O_2$), guaiacol ($C_7H_8O_2$), and the levoglucosan fragment ($C_6H_8O_4$), many of these compounds are known to form during lignin pyrolysis (Hatch et al., 2015, Koss et al., 2018; Nowakowska et al., 2018), while dimethylbutenedial ($C_6H_8O_2$), trimethylbutenedial ($C_7H_{10}O_2$) are ring opening oxidation products of aromatic compounds (Zaytsev et al., 2019).”

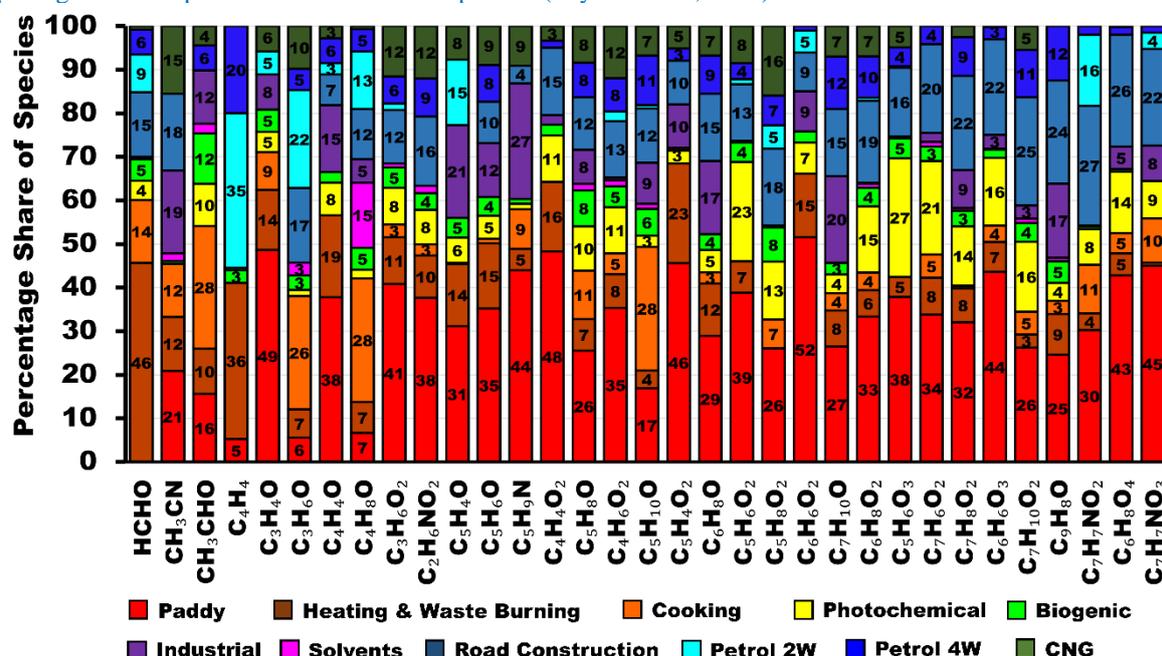


Figure 6: VOC species to which different forms of biomass burning contribute the highest percentage share of the atmospheric burden in Delhi

Section 3.2.2

“Figure 6 shows that this factor explains the largest percentage share of the total mass for formaldehyde (HCHO) and vinylacetylene + 1-buten-3-yne (C_4H_4), and the second largest percentage share of furfural

(C₅H₄O₂), methylfurfural (C₆H₆O₂), furan (C₄H₄O), methyl furan (C₅H₆O), furanone (C₄H₄O₂) and acrolein (C₃H₄O). All these compounds are characteristic of biomass burning smoke (Hatch et al., 2015, Stockwell et al., 2015, Koss et al., 2018).”

Section 3.2.3

“Figure 6 shows that factor explains the largest percentage share of butanone (C₄H₈O), pentanone (C₅H₁₀O), acetaldehyde (CH₃CHO), acetone (C₃H₆O), and benzaldehyde (C₇H₆O). All these compounds are characteristic of biomass burning smoke (Hatch et al., 2015, Stockwell et al., 2015, Koss et al., 2018).”

Section 3.2.4

“Figure 7 shows that the factor explains the largest percentage share of methanol (CH₃OH) and the second largest percentage share of ethanol (C₂H₆O). These compounds are formed by the incomplete combustion of CNG that is catalytically converted to methanol and ethanol (Singh et al., 2016).”

Section 3.2.5

“Figure 7 shows that the factor explains the largest percentage share of most aromatic compounds, namely C₈-aromatics, toluene, C₉-aromatics (C₈H₁₂), C₄-substituted benzene + C₂-substituted xylene, benzene, styrene (C₈H₈), methylstyrenes + indane (C₉H₁₀), and C₂-substituted styrenes (C₁₀H₁₂) and a few oxygenated aromatic hydrocarbons such as methyl phenol isomers (C₇H₈O) and methyl chavicol (C₁₀H₁₂O). The fact that the factor explains the largest percentage share of ethanol and the MTBE fragment (C₄H₈) can likely be attributed to ethanol blending and the use of MTBE in petrol (Achten et al., 2001). This factor also explains the largest percentage share of several other hydrocarbons such as propyne (C₃H₄), propene (C₃H₆), cyclopentadiene (C₅H₆), hexane (C₆H₁₄), C₇H₆, C₇H₁₀, and cycloheptene (C₇H₁₂).”

Section 3.2.6

“Figure 7 shows that this factor explains the largest percentage share of toluene, and a number of oxygenated aromatic compounds such as benzaldehyde (C₇H₆O), tolualdehyde (C₈H₈O), and phenol (C₆H₆O). It also explains the largest percentage share of nitrobenzene (C₆H₅NO₂), cyclohexanone (C₆H₁₁O), and vinyl chloride (C₂H₃Cl). It also explains the second largest percentage share of benzene, vinylacetylene (C₄H₄), acetone + propanal, methoxyamine (CH₃NO) and butanoic acid/ethyl acetate (C₄H₉O₂).”

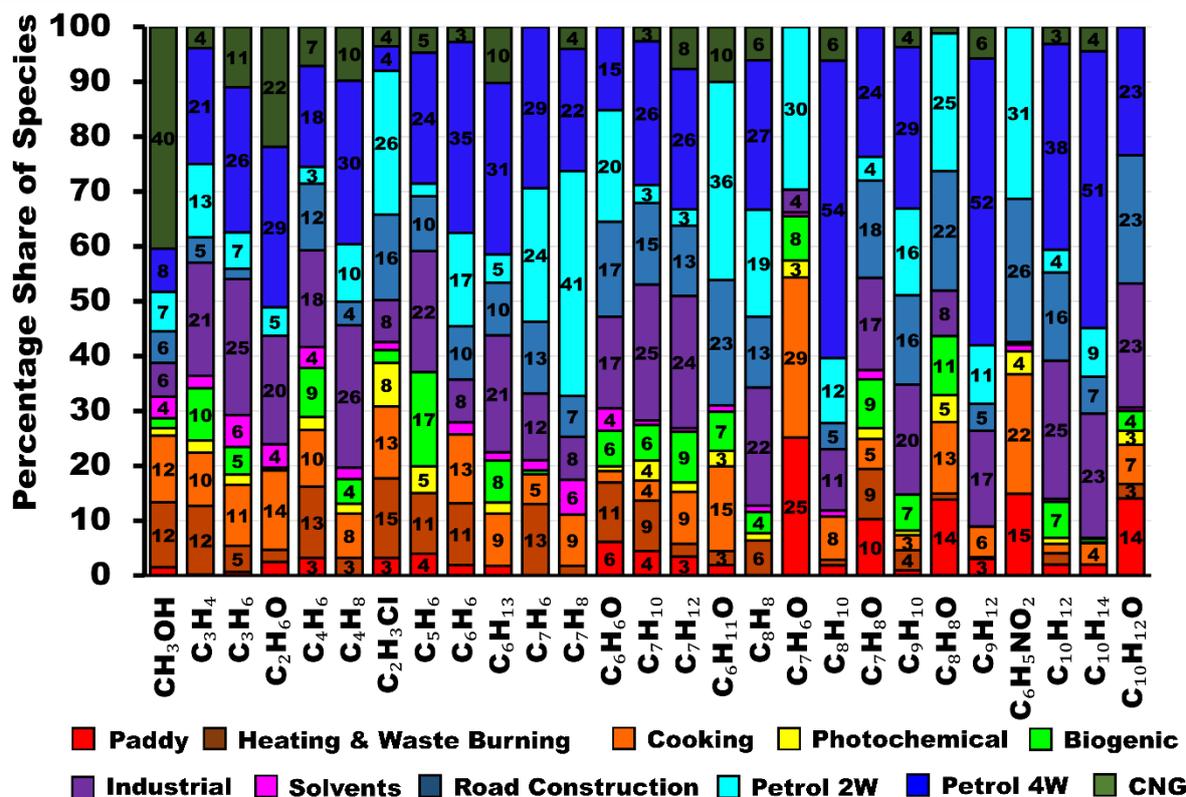


Figure 7: VOC species to which the transport sector contributes the highest percentage share of the atmospheric burden in Delhi

Section 3.2.7

“Figure 8 shows that the factor explains the largest percentage share of methanethiol (CH₃S), a chemical used in the manufacture of the essential amino acid methionine, in the plastic industry and the manufacturing of

pesticides, dichlorobenzenes ($C_6H_4Cl_2$), a chemical used in the synthesis of dyes, pesticides, and other industrial products and methoxyamine (CH_5NO). Analyses of the primary dataset by Mishra et al. (2024) also qualitatively inferred an industrial source for methanethiol and dichlorobenzene. It also explains the largest percentage share of the sum of monoterpenes, camphor/pinene oxide ($C_{10}H_{16}O$), santene (C_9H_{14}) the terpene fragment (C_8H_{12}), C_8H_{14} , C_9H_{16} , cyclohexene (C_6H_{10}) and cyclopentylbenzene ($C_{11}H_{14}$). Terpenes are used in the food and beverages, cosmetics, pharmaceutical, and rubber industry. In addition, this factor also explains the largest percentage share of a large suite of volatile and IVOC aromatic hydrocarbons including naphthalene ($C_{10}H_8$), methyl naphthalene ($C_{11}H_{10}$), $C_{12}H_{16}$, $C_{13}H_{18}$, $C_{13}H_{20}$, $C_{13}H_{22}$, $C_{14}H_{20}$, and $C_{14}H_{22}$. Ambient observations for most of these IVOCs have not been reported in the literature so far. Only, C_9H_{14} , $C_{12}H_{12}$, and $C_{12}H_{16}$ have been reported from aircraft engine emissions (Kılıç et al., 2018) while terpenes, C_9H_{16} , cyclopentylbenzene, naphthalene and methyl naphthalene have been reported from a wide range of combustion sources (Hatch et al., 2015, Bruns et al., 2017). Most other compounds have so far only been reported to degas from heated asphalt (Khare et al., 2020). Due to the high abundance of IVOCs in this factor, it contributes 15 % to the total SOA formation potential.”

Section 3.2.8

“Figure 8 shows that the factor explains the largest share of organic acids namely butanoic acid, acetic acid and isocyanic acid (HNCO) and the second largest share of butanal + butanone + MEK (C_4H_8O). These compounds point towards stack venting of VOCs from chemical-, food-, or pharmaceutical industries or polymer manufacturing as likely sources of these emissions (Hodgson et al., 2000, Villberg et al., 2001, Jankowski et al., 2017, Gao et al., 2019). This assessment is broadly confirmed by the fact that the best source match ($R=0.7$) for this source was collected from a plot situated opposite a polymer manufacturing unit and next to a pet food manufacturer in an industrial area at Jahangirpuri N of the receptor site.”

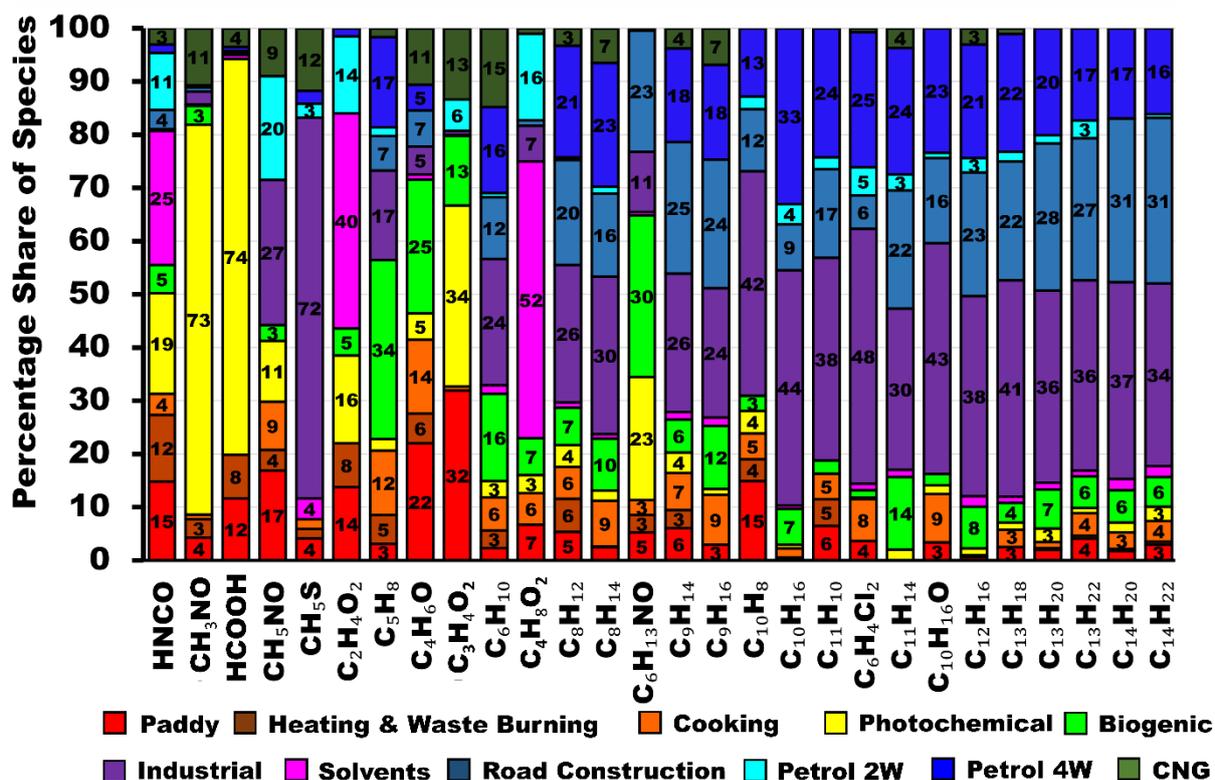


Figure 8: VOC species to which the industries, solvent usage, photochemistry or biogenic sources contribute the largest percentage share of the atmospheric burden in Delhi

Section 3.2.10

“Figure 8 shows that the factor explains the largest percentage share of formic acid, formamide, and methyl glyoxal ($C_3H_5O_2$). It also explains the second largest percentage share of isocyanic acid (HNCO) and hexanamide ($C_6H_{13}NO$), which are formed by the photooxidation of amines (Yao et al., 2016; Wang et al., 2022). Some compounds point towards a significant contribution of photochemically aged biomass burning emissions to this factor for example furfuryl alcohol ($C_5H_6O_2$), hydroxymethyl furanone ($C_5H_6O_3$), and hydroxybenzaldehyde ($C_7H_6O_2$).”

Section 3.2.11

“Figure 8 shows that this factor explains the largest percentage share of two BVOCs namely Isoprene + 2-methyl-3-butene-2-ol fragment, and its oxidation product, methyl vinyl ketone, methacrolein and 2-butenal. It

also explains the largest percentage share of C6 amides (C₆H₁₃NO) which are produced by the photo-oxidation of amines (Yao et al., 2016).”

Section 3.2.9

“As represented by Fig. 9, this factor explains the largest percentage share of a large suite of volatile and IVOC hydrocarbons namely, heptene (C₇H₁₄), C₁₁H₁₂, C₁₂H₁₂, C₁₄H₁₄, C₁₄H₁₈, C₁₆H₂₄, C₁₇H₂₈, and C₁₈H₃₀. In addition, it explains the second largest percentage share of many other IVOC hydrocarbons namely C₉H₁₄, C₉H₁₆, C₁₁H₁₄, C₁₂H₁₆, C₁₃H₁₈, C₁₃H₂₀, C₁₃H₂₂, C₁₄H₂₀, C₁₄H₂₂. Except for the four hydrocarbons C₇H₁₄, C₉H₁₄, C₉H₁₆, and C₁₁H₁₂, all of these IVOCs have been reported to degas at 60°C from asphalt pavement (Khare et al., 2020). So far only C₁₄H₁₈ has been reported as fresh gas phase emissions (transport time <2.5 min) from a farm (Loubet et al., 2022) in ambient air, while C₁₇H₂₈ has been reported in the aerosol phase (Xu et al., 2022). The road construction factor also explains the largest percentage share of a long list of OVOCs namely, C6 diketone isomers (C₆H₁₀O₂), C2-substituted phenol (C₈H₁₀O), C₇H₁₂O₂, C₈H₁₄O₂, C₈H₁₆O₂, phthalic anhydride (C₈H₄O₃), which is a naphthalene oxidation product (Bruns et al., 2017), C₉H₁₀O, C₉H₁₂O₂, C₉H₁₄O₂, C₉H₁₆O₂, C₉H₁₈O₂, C₁₀H₁₂O, C₁₀H₁₈O, C₁₀H₈O₃, C₁₀H₁₆O₃, and C₁₂H₁₈O₂. However, out of these only C₁₀H₁₂O and C₁₀H₁₈O have been detected as direct emissions from heated asphalt pavement (Khare et al., 2020) indicating that most OVOCs in this factor are possibly oxidation products of short-lived IVOCs hydrocarbons emitted by this source. This assessment is supported by the volatility oxidation state plot for the road construction factor (Figure S9) which demonstrates that both precursors and oxidation products are present in this factor and that C6 to C10 hydrocarbons appear to be progressing from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009).”

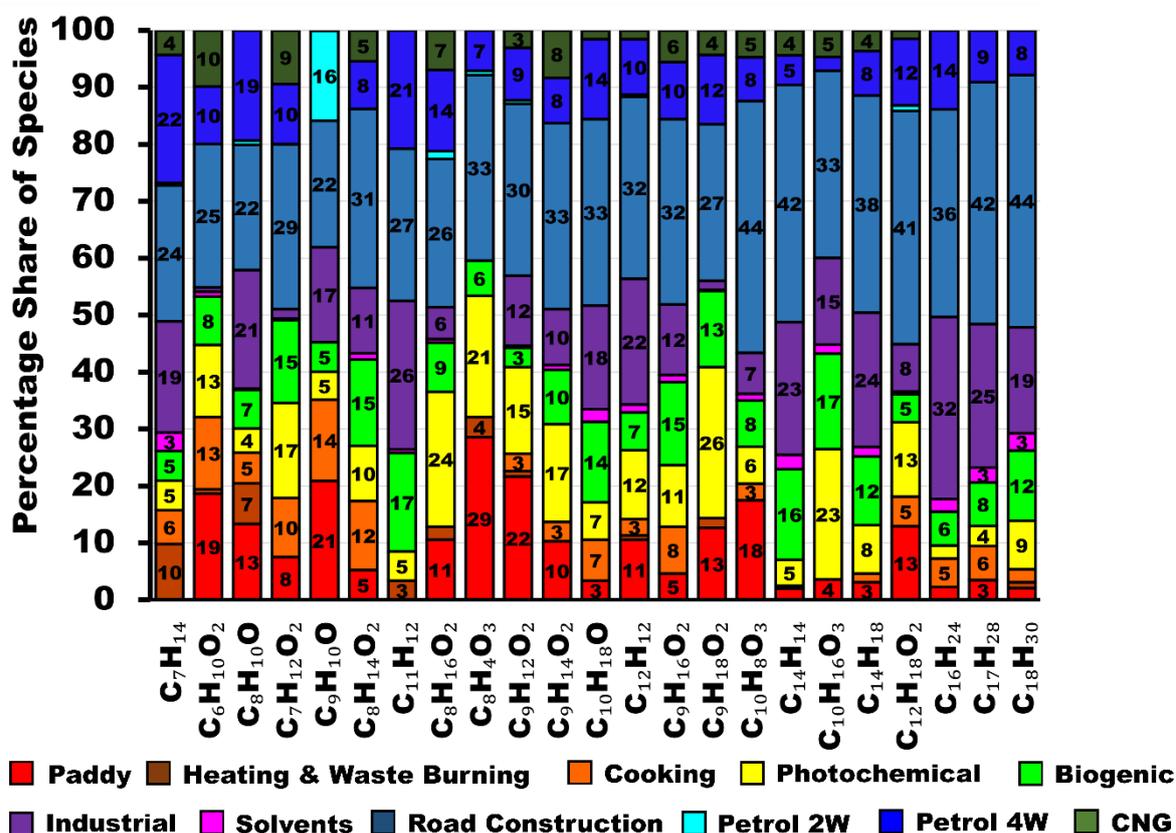


Figure 9: VOC species to which road construction contributes the largest percentage share of the atmospheric burden in Delhi.

Line 224: “The source identity of the PMF factors was confirmed by matching the normalized PMF factor profiles with normalized source fingerprints”. Could you add more detail about this, did you check the R correlations? Or was it just by visually comparing them?

Yes, we have done an R correlation of the source profiles with the PMF factors. We have Now included R values in the manuscript. The revised text reads as follows

“Figure 3 shows the source profile of the eleven factors that our PMF analyses resolved. Out of the 111 VOCs only those whose normalized source contribution exceeded 0.1 when divided by the most abundant compound in the same source profile in at least one of the sources, were included in the figure. The source identity of the PMF factors was confirmed by matching the PMF factor profiles with the unit $\mu\text{g m}^{-3}$ with normalized source

fingerprints of grab samples collected from the potential sources. To facilitate the comparison of emission factors and grab samples from different studies with the PMF output, the source samples were normalized by dividing each species' mass/emission factor by the mass/emission factor of the most abundant species in a given fingerprint. The PMF factor profile matched best against source samples collected from burning paddy fields ($R=0.6$, Kumar et al., 2020) for the paddy residue burning factor. The cooking factor matched emissions from a cow-dung-fired traditional stove called angithi ($R=0.7$, Fleming et al., 2018). The residential heating & waste burning factor had a source fingerprint matching emission from leaf litter burning, ($R=0.7$, Chaudhary et al., 2022), waste burning ($R=0.7$, Sharma et al., 2022), and cooking on a chulha fired with a mixture of firewood and cow dung ($R=0.9$, Fleming et al., 2018). The factors identified as CNG ($R=1.0$), petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.6$) matched tailpipe emissions of the respective vehicle types and fuels (Hakkim et al., 2021). The petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.7$) also matched traffic junction grab samples from Delhi (Chandra et al., 2018). The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving ($R=0.9$, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles ($R=0.6$, Che et al., 2023). The factors identified as solvent usage and evaporative emissions matched ambient air grab samples collected from an industrial area at Jahangirpuri ($R=0.7$), and Dhobighat at Akshar Dham ($R=0.5$) in this study. The factor identified as industrial emissions showed the greatest similarity to ambient air grab samples from the vicinity of the Okhla waste-to-energy plant ($R=0.8$), Gurugram ($R=0.7$) and Faridabad ($R=0.8$) industrial area. The biogenic factor showed the greatest similarity to leaf wounding compounds released from *Populus tremula* ($R=0.8$, Portillo-Estrada et al., 2015) as well as BVOC fluxes from *Mangifera indica* ($R=0.4$, Datta et al., 2021)."

Line 235-236: Did you measure the Munirka furniture market and Dhobighat at Akshar Dham samples? If not, could you add their reference?

Yes, all source-samples not referenced were collected by us. Meanwhile we have collected more samples and found a more relevant match for the solvent factor and on a plot situated opposite a polymer manufacture and next to a pet food manufacturer have updated the figure with better matches. We now also clearly state that samples were collected by us:

The factors identified as solvent usage and evaporative emissions matched ambient air grab samples collected in an industrial area at Jahangirpuri ($R=0.7$), and Dhobighat at Akshar Dham ($R=0.5$) in this study.

Figure5: I would suggest enlarging (by the x axis) the timeseries plot, to make them easier to read. You should keep the same order of the factors as in description (& throughout the paper). What do the lines/shaded areas for the diurnal cycles represent (mean, median...)?

We have modified the plot and figure caption as per the suggestions:

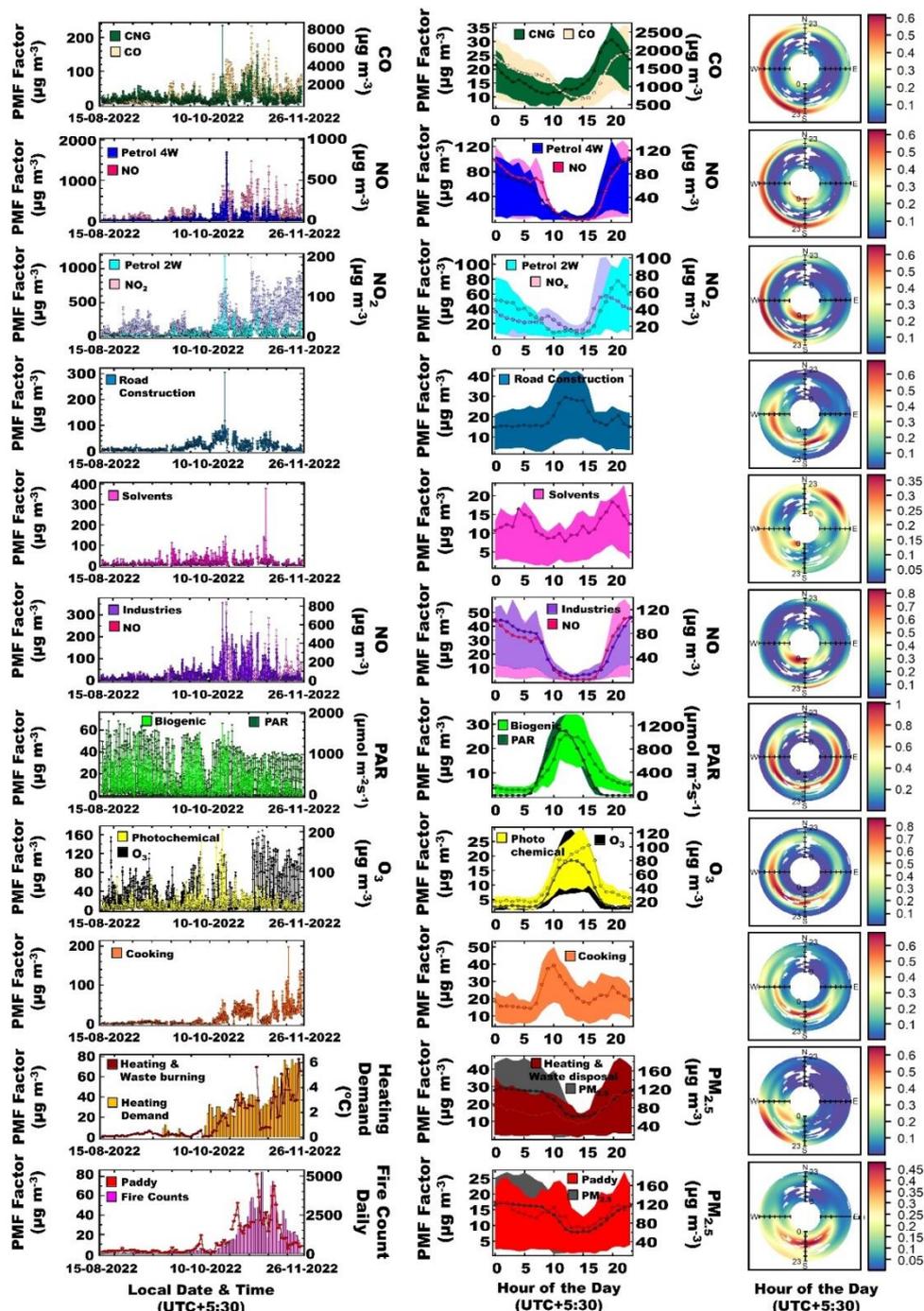


Figure 3: Time series of each factor in $\mu\text{g m}^{-3}$ (left column) with respective normalized diurnal profiles (centre column). The shaded region in the diurnal profiles depicts the area between the 25th and 75th percentile while the median of the dataset is marked as the line. The polar plots (right column) depict the conditional probability of a factor having a mass contribution above the 75th percentile of the dataset during a certain hour of the day between midnight (centre of rose) and 23:00 local time (outside of rose) from a certain wind direction. This probability is determined by dividing the number of observations above the 75th percentile by the total number of measurements in each bin.

3.2.2. There is a mention that this factor may not be always fresh, which I found interesting, you could add a few words at the end of the paragraph about the fresh/aged nature of the factors based on all the information.

We appreciate the suggestions. This assessment was primarily based on the fact that this night time factor shows a lower R with NO than with NO₂. Our comment was primarily meant to contrast this factor with some of the transport sector and industrial emissions that have a much higher R with NO than NO₂ indicating the night time plumes of these factors are so fresh that their atmospheric lifetime is more likely on the scale of minutes rather than in hours, while heating and waste disposal plumes are occasionally fresh but often also aged. We have modified the text in line 330 to clarify as follows:

“The lower correlation with NO ($R=0.4$) (Table S5), indicated that emissions are combustion-related but not always fresh. Occasionally, fresh plumes reach the receptor within minutes, however the majority of plumes have a higher atmospheric age, as NO is a short-lived species and oxidized to NO_2 on the timescale of minutes in the presence of ozone”

3.2.3. Some of these compounds (i.e. aromatics) can also be associated with cooking activities (e.g. Crippa et al (2013), doi.org/10.5194/acp-13-8411-2013).

Thank you for the suggestion. The reference has been added to the manuscript.

These aromatic compounds have been reported to originate from cooking emissions (Crippa et al., 2013).

Crippa, M., Canonaco, F., Slowik, J. G., El Haddad, I., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Temime-Roussel, B., Abidi, E., Poulain, L., Wiedensohler, A., Baltensperger, U., and Prévôt, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, *Atmos. Chem. Phys.*, 13, 8411–8426, <https://doi.org/10.5194/acp-13-8411-2013>, 2013.

3.2.4. You could add one sentence about the interpretation of VOCs (i.e. CH_3OH methanol and ethanol) for this factor.

We appreciate the suggestions. We have modified the text in the manuscript as follows:

“Figure 7 shows that the factor explains the largest percentage share of CH_3OH methanol and the second largest percentage share of $\text{C}_2\text{H}_6\text{O}$ ethanol. These compounds are formed by the incomplete combustion of CNG that is catalytically converted to methanol and ethanol (Singh et al., 2016).”

Singh, S., Mishra, S., Mathai, R., Sehgal, A. K., & Suresh, R.: Comparative study of unregulated emissions on a heavy duty CNG engine using CNG & hydrogen blended CNG as fuels. *SAE Int. J. Engines*, 9(4), 2292-2300, <http://dx.doi.org/10.4271/2016-01-8090>, 2016.

3.2.5. & 3.2.6. Add a sentence (or change existing text) to highlight the differences between 2-wheeler & 4-wheeler factors.

We appreciate the suggestions. We have revised the manuscript to better contrast the two. Firstly petrol 4-wheeler emissions are on average much fresher as central Delhi is a prosperous neighbourhood dominated by private cars. Petrol 2-wheeler plumes are on average more aged. Section 3.2.5 now starts as follows:

“Figure 4 shows petrol 4-wheeler contributed to 20 %, 25 %, and 30 % to the VOC mass loading, OFP, and SOAP, respectively. The source fingerprint of this source matched tailpipe emissions of petrol-fuelled 4-wheelers (Hakkim et al., 2021) and is characterized, in descending rank of contribution, by C8-aromatics, toluene, C9-aromatics (C_9H_{12}), benzene, butene + methyl tert-butyl ether (MTBE) fragment, propyne, propene, methanol and C2-substituted xylenes + C4-substituted benzenes ($\text{C}_{10}\text{H}_{14}$). Figure 5 shows that emissions peak in the evening between 7 pm and midnight with average VOC mass loadings $>70 \mu\text{g m}^{-3}$ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO ($R=0.8$), CO ($R=0.7$), and CO_2 ($R=0.7$) indicating the receptor site is impacted by fresh combustion emissions from this source and the atmospheric age of most plumes is on the timescale of minutes.”

Section 3.2.6 now starts as follows:

“Figure 4 shows petrol 2-wheeler contributed to 14 %, 12 %, and 20 % to the VOC mass loading, OFP, and SOAP respectively. The source fingerprint of this source matched tailpipe emissions of petrol-based 2-wheelers (Hakkim et al., 2021) and are characterized, in descending rank of contribution, by toluene, acetone + propanal, C-8 aromatic compounds, acetic acid ($\text{C}_2\text{H}_4\text{O}_2$), propyne (C_3H_4), methanol (CH_3OH), benzene (C_6H_6), the MTBE fragment and C-9 aromatics (C_9H_{12}). A key difference of the petrol 2-wheeler source profile in comparison to the petrol 4-wheeler source profile is the lower benzene to toluene ratio, which is supported by the GC-FID analysis of tailpipe exhaust (Kumar et al., 2020). Figure 5 shows that emissions peak in the evening between 8 pm and 10 pm with average VOC mass loadings $>50 \mu\text{g m}^{-3}$ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO_x ($R=0.6$), CO ($R=0.6$) and CO_2 ($R=0.7$), but have a lower correlation with NO ($R=0.5$) (Table S5), and a larger contribution of oxygenated compounds to the source profile, indicating that the emissions have been photochemically aged. This suggests that contrary to 4-wheeler plumes which originate from the immediate vicinity of the site in central Delhi (Figure S1), 2-wheeler plumes reach the receptor after prolonged transport from more distant rural and suburban areas on the outskirts of the city. In such areas, people often favour two-wheelers over four-wheelers.”

Line 422: Interesting! Could this last sentence mean that part of $\text{PM}_{2.5}$ for this factor would be SOA?

Yes, likely most of it because $\text{PM}_{10}=\text{PM}_{2.5}$ for this factor. We have added a sentence to this effect along with a volatility oxidation state plot for this factor.

“Figure S9 shows the volatility oxidation state plot for all 111 VOCs in which the marker size represents the

percentage share of each compound explained by the industrial factor and markers are colour coded by the number of carbon atoms. The plot shows evidence of the first- and second-generation oxidation products of C6 to C10 hydrocarbon transitioning from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009). This and the fact that the entire aerosol associated with this factor is PM_{2.5}, indicates that most of the aerosol associated with this factor is likely SOA.”

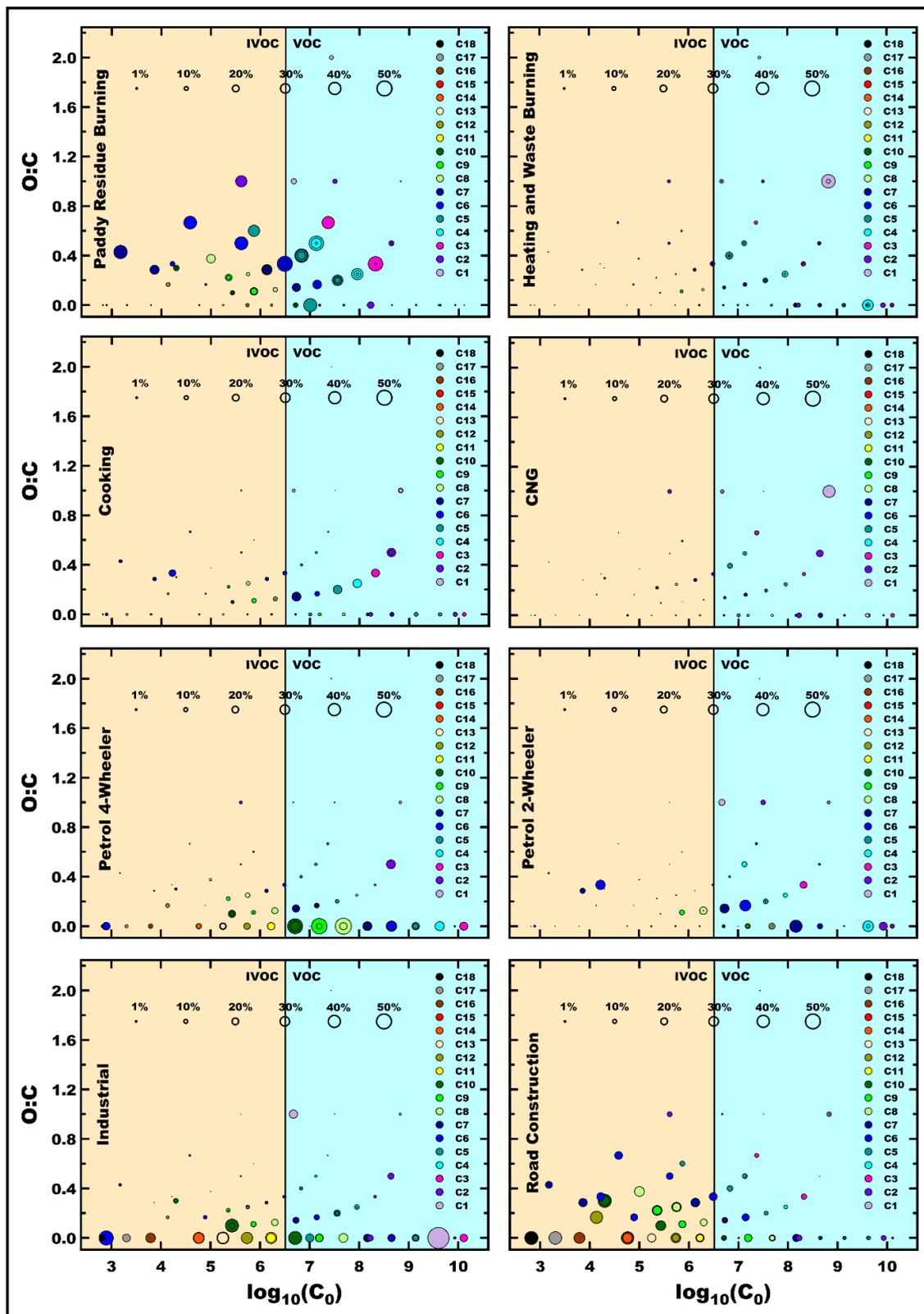


Figure S9: Volatility oxidation state plots for all factors that individually contribute more than 3% to the total SOA formation potential.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, *Science*, 326, 1525-1529, <https://doi.org/10.1126/science.1180353>, 2009.

Line 432-433: Do you have references for this last statement?

Yes, we have added references to this statement.

“Figure 8 shows that the factor explains the largest share of organic acids namely butanoic acid, acetic acid and isocyanic acid (HNCO) and the second largest share of butanal + butanone + MEK (C₄H₈O). These compounds point towards stack venting of VOCs from chemical-, food-, or pharmaceutical industries or polymer manufacturing as likely sources of these emissions (Hodgson et al., 2000, Villberg et al., 2001, Jankowski et al., 2017, Gao et al., 2019). This assessment is broadly confirmed by the fact that the best source match for this source was collected from a plot situated opposite a polymer manufacture and next to a pet food manufacturer in an industrial area at Jahangirpuri (R=0.7) N of the receptor site.”

Gao, Z., Hu, G., Wang, H., Zhu, B.: Characterization and assessment of volatile organic compounds (VOCs) emissions from the typical food manufactures in Jiangsu province, China, *Atmos. Pollut. Res.* **10**(2), 571-579, <https://doi.org/10.1016/j.apr.2018.10.010>, 2019.

Hodgson, S. C., Casey, R. J., Bigger, S. W., & Scheirs, J.: Review of volatile organic compounds derived from polyethylene. *Polym-Plast Technol.* **39**(5), 845-874. <https://doi.org/10.1081/PPT-100101409>, 2000.

Jankowski, M. J., Olsen, R., Thomassen, Y., & Molander, P.: Comparison of air samplers for determination of isocyanic acid and applicability for work environment exposure assessment. *Environm. Sci-Proc. Imp.*, **19**(8), 1075-1085, <https://doi.org/10.1039/C7EM00174F>, 2017.

Villberg, K., & Veijanen, A.: Analysis of a GC/MS thermal desorption system with simultaneous sniffing for determination of off-odor compounds and VOCs in fumes formed during extrusion coating of low-density polyethylene. *Anal. Chem.* **73**(5), 971-977. <https://doi.org/10.1021/ac001114w>, 2001.

3.2.9. Interesting, the last sentence suggests a possible link of the OVOCs with SOA?

Yes. We have added a statement to this effect.

This assessment is supported by the volatility oxidation state plot for the road transport factor (Figure S10) which demonstrates that both precursors and oxidation products are present in this factor and that C6 to C10 hydrocarbons appear to be progressing from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009).

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, *Science*, 326, 1525-1529, <https://doi.org/10.1126/science.1180353>, 2009.

3.3. It's a little tedious to read with all the emission values, please select when it is truly important to have them.

We appreciate the referee's suggestions and added the numbers as Table S5 to the supplement. We have revised the text to reduce the numbers. It now reads as follows:

“Figure 10 shows a comparison of different anthropogenic emission inventories with the PMF output data from this study for three overlapping fetch regions corresponding to the fetch region from which air masses will reach the receptor site within 24 hours for different airflow patterns (Figure 1).

One feature that stands out in this comparison is that all inventories appear to significantly overestimate the relative contribution of residential fuel usage to the VOC and particulate matter emissions for all fetch regions. In

absolute terms, the Regional Emission Inventory in Asia (REAS v3.2.1) for the year 2015 (Kurokawa & Ohara, 2020) and the Emission Database for Global Atmospheric Research (EDGARv6.1) for the year 2018 (Crippa et al., 2022), agree on the residential sector PM_{2.5} emissions for the NW fetch region (Table S6). According to the latest estimates (Pandey et al., 2021), the NW-IGP region has the lowest prevalence of solid fuel usage in the entire IGP and the inventories appear to overestimate the PM_{2.5} emissions from this fetch region only by a factor of 1.5-1.9. For the SW and SE fetch region, respectively, REAS v3.2.1 estimates much larger residential sector PM_{2.5} emissions than EDGARv6.1 and overestimates the PMF estimates by a factor of 3.7 and 4.6. In contrast, EDGARv6.1 only overestimates PMF estimates by a factor of 1.8 and 3.2, for the SW and SE fetch region respectively. Solid fuel-based cooking is more prevalent in both Central and Western India and the Eastern IGP than in the NW-IGP (Pandey et al., 2021). The overestimation in both inventories may be caused by a gradual adoption of cleaner technology. Sharma et al., (2022) calculated a 13 % drop in residential sector PM_{2.5} emissions between 2015 and 2020 due to higher LPG sales and a continuation of that trend to 2022 could explain the overestimation of residential fuel usage in the present emission inventory data. For PM₁₀, the EDGARv6.1 emission estimates for the NW, SW and SE fetch region, are greater than the REASv3.2.1 emission inventory. The EDGARv6.1 and REASv3.2.1 inventory both overestimate our PMF PM₁₀ results by a factor of 1.5 to 3.0. However, while the REASv3.2.1 inventory appears to assume that most of the residential sector aerosol emissions occur in the fine mode, our PMF results (Fig. 10) clearly agree with the EDGARv6.1 inventory on the fact that there are significant coarse aerosol emissions associated with solid-fuel based cooking and heating. Table S6 shows that for residential sector VOCs emissions, the absolute emissions in the EDGARv6.1 inventory are almost twice as large as those in the REASv3.2.1 inventory, even though the percentage contribution of this sector to the VOC emissions in the inventory in Figure 10 appears to be similar for both, because of larger VOC emissions from solvent use and industries in the EDGARv6.1 inventory. Both inventories overestimate the relative importance of residential sector emissions in relation to VOC emissions from other sectors by more than a factor of two when compared to our PMF estimate, most likely because they have not been updated with recent fuel shifts towards LPG in the relatively prosperous Delhi NCR region (Sharma et al., 2022).

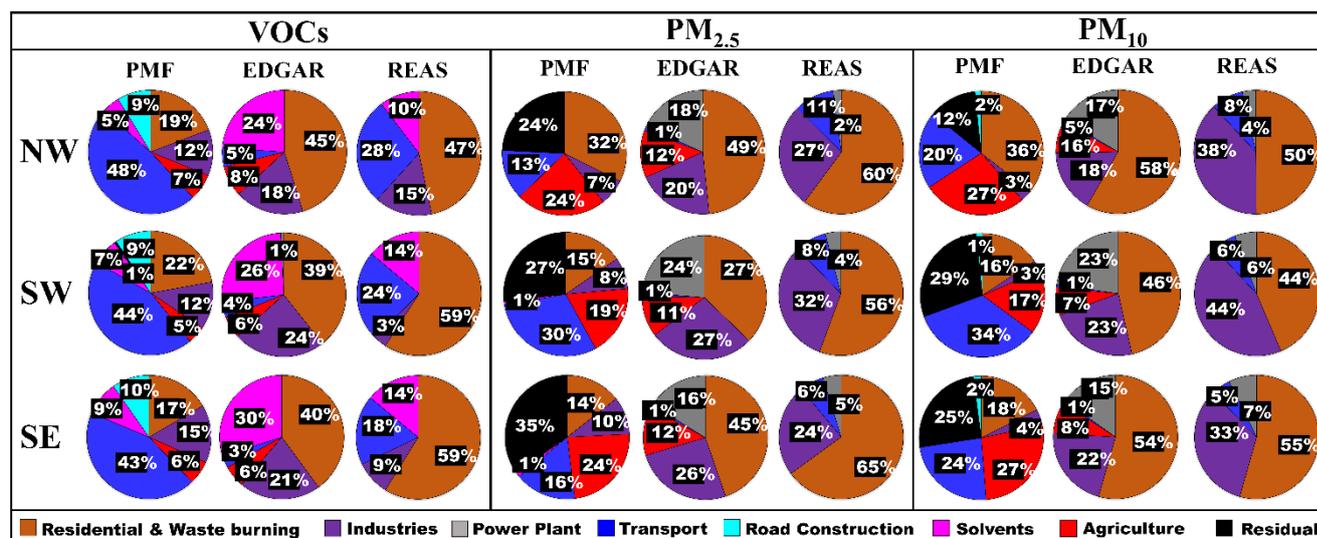


Figure 10: Comparison of different anthropogenic emission inventories with the PMF output from this study for three overlapping fetch regions corresponding to different airflow patterns.

With respect to industrial emissions of VOCs for the NW fetch region, our PMF results indicate that the actual emissions are slightly smaller than those in the REASv3.2.1 inventory, while the EDGARv6.1 inventory overestimates emissions. For the SW and SE fetch region, our PMF estimates fall in between those of the EDGARv6.1 inventory and the REASv3.2.1 inventory. For industrial PM_{2.5} emissions, both EDGARv6.1 & REASv3.2.1 are close and agree on the magnitude of emissions for the NW, SW and SE fetch region, respectively, and both inventories appear to overestimate emissions when compared to our PMF results. Our findings seem to suggest that the pollution boards have been somewhat successful in clamping down on industrial emissions and the technology employed is better than what is currently reflected in emission inventories. Industrial fly ash (PM₁₀) emissions are larger in the REASv3.2.1 inventory for all the fetch regions compared to EDGARv6.1 inventory. Yet both inventories appear to significantly overestimate industrial emissions when compared to our PMF results. These findings also indicate the pollution boards have been somewhat successful in clamping down on large and visible fly ash sources and that the EDGARv6.1 inventory has captured this clean-technology transition better. The REASv3.2.1 inventory completely misses direct VOC and PM emissions from the agricultural sector. The EDGARv6.1 inventory significantly underestimates PM_{2.5} & PM₁₀ emissions from agricultural activities, which

include, but are not limited to crop residue burning, in comparison to our PMF results, particularly over NW-India (Table S6). Over this fetch region EDGARv6.1 attributes as much PM_{2.5} to all agricultural activities combined for the full year as the FINNV2.5 inventory (Wiedinmyer et al., 2023) attributes just to agricultural residue burning activities taking place between 15th August and 26th November 2021 (a time period comparable to the period in our model run), without including the emissions from rabi crop residue burning in summer (Kumar et al., 2016) and other agricultural activities such as harvest and ploughing. For PM₁₀ the fire count based FINNV2.5 estimate is twice as high as the emission estimate of EDGARv6.1 for this fetch region, and more likely to be correct, because the phytoliths present in rice straw form coarse mode ash during the combustion process (Figure S10). The fact that EDGAR appears to underestimate residue-burning emissions over this fetch region has been flagged earlier (Pallavi et al., 2019; Kumar et al., 2021; Singh et al., 2023). Our PMF analyses also reveals that the relative contribution of agricultural residue burning to the PM burden over the North-Western IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) and South-Eastern IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) is comparable, despite the much lower fire counts over the South-Eastern IGP (17,810), when compared to the North Western IGP (61,334). This indicates that either fires to the SE are burning closer to the receptor site or the fire detection efficiency in this fetch region is lower. Table S6 reveals that the relative importance of agricultural emissions over the SE fetch region is even more severely underestimated in the FINNV2.5 inventory than in the EDGARv6.1 inventory due to poorer fire detection (close to 100% omission error) for the partial burns prevalent over this region (Lui et al. 2019, 2020, Figure S8) when compared to the complete burns prevalent over the NW IGP (Lui et al. 2019, 2020, Figure S7).

Transport sector VOC emissions appear to be severely underestimated in the EDGARv6.1 inventory for the NW, SW, and SE fetch region, which has been previously flagged for earlier versions of the same inventory (Sarkar et al., 2017; Pallavi et al., 2019; Singh et al., 2023). The REASv3.2.1 inventory also underestimates our PMF results. This indicates that the contribution of the transport sector to ambient VOC pollution levels in a megacity like Delhi may not be adequately reflected in both the emission inventories. Our PMF suggests that the overall contribution of the transport sector to the total PM_{2.5} and PM₁₀ pollution levels occurs primarily due to non-exhaust emissions from the CNG-fuelled public transport fleet. These non-exhaust emissions are much larger than what is accounted for both in the EDGARv6.1 and REASv3.2.1 inventories for PM_{2.5} & PM₁₀ emissions from the NW, SW and SE fetch region. The transport sector-related findings of this PMF source apportionment study are in agreement with earlier source apportionment studies that often attributed a quarter or more of the total PM emissions to the transport sector. Some prior studies used metals, Pb and/or OC/EC as transport sector activity tracers (Jain et al., 2017, 2020; Sharma et al., 2016, Jaiprakash et al., 2016; Sharma & Mandal, 2017), while others attributed almost the entire HOA component of organic aerosol to transport sector emissions (Reyes-Villega et al., 2021; Cash et al., 2021; Kumar et al., 2022, Shukla et al., 2023) or used a Chemical Mass Balance (CMB) model with source fingerprints from the EPA database (Nagar et al., 2017). Our PMF results differ to emission-inventory-based assessments, which only attribute a minor share of the total PM burden to this activity (Guo et al., 2017). Our findings also add insights to the reasons why the transport sector targeted air quality interventions yielded such poor results (Chandra et al., 2018). Public transport availability was ramped up during the periods when road-rationing schemes restricted the use of private 4-wheelers. Our results suggest that investments into the road infrastructure, that reduce resuspension, modal shifts from buses towards metro-based public transport and electric vehicles with >50 % regenerative braking (Liu et al., 2021) that limit brake wear can yield meaningful reductions in the transport sector-related PM emissions.

Our PMF results indicate that solvent usage results in VOC emissions that are more in line with the REASv3.2.1 inventory while the EDGARv6.1 inventory overestimates emissions by a factor of 4 for all the fetch regions.

Power generation is not considered to be a significant VOC source in both emission inventories (<1 % of the total VOC mass), and fails to show up as a separate sector in our PMF results, as our model runs rely on VOC tracers to track pollution sources. The contribution of energy generation towards the PM burden particularly in the EDGARv6.1 emission inventory, however, is significant. It is, however, striking to note that the PMF features a residual that is of similar magnitude as the PM_{2.5} and PM₁₀ emissions attributed to power generation in the EDGARv6.1 inventory. Power generation is believed to primarily contribute secondary sulfate and nitrate aerosol (Atabakhsh et al., 2023), which is unlikely to be directly associated with a fresh combustion signature. It is hence likely, that much of our PMF residual can be attributed primarily to this source. The amount of emissions attributed to power generation in the REASv3.2.1 inventory is much smaller than those reflected in EDGARv6.1, likely because the inventory misses several coal generation units that were commissioned between 2015-2018.

Our PMF results identify road construction and asphalt pavements as an additional VOC source that is at present not reflected in emission inventories.”

Table S5: Emissions from different sectors for north-western, south-western, and south-eastern fetch regions.

VOC (Gg y ⁻¹)			
	NW	SW	SE

Sector	EDGAR	REAS	FINN	EDGAR	REAS	FINN	EDGAR	REAS	FINN
Residential fuel usage	764	353	-	1421	947	-	1196	862	-
Industrial	302	113	-	867	55	-	635	133	-
Agricultural residue	135	0	760	204	0	801	171	0	207
Transport	84	212	-	154	378	-	96	266	-
Solvents	403	78	-	939	222	-	896	204	-
Power Industry	7	2	-	27	4	-	12	4	
PM_{2.5} (Gg y⁻¹)									
Sector	NW			SW			SE		
Residential fuel usage	382	379	-	713	934	-	597	830	-
Industrial	158	173	-	524	541	-	342	307	-
Agricultural residue	97	0	95	206	0	100	168	0	26
Transport	8	65	-	18	137	-	12	80	-
Solvents	0	0	-	0	0	-	0	0	-
Power Industry	144	14	-	453	68	-	215	61	-
PM₁₀ (Gg y⁻¹)									
Sector	NW			SW			SE		
Residential fuel usage	750	401	-	1391	994	-	1157	882	-
Industrial	211	308	-	684	1015		458	539	-
Agricultural residue	103	0	192	217	0	203	177	0	52
Transport	10	67	-	22	140	-	14	83	-
Solvents	0	0	-	0	0	-	0	0	-
Power Industry	213	28	-	679	130		321	118	

Lines 500-505: A map could be useful here as well.

We appreciate the referee's suggestions and added the map to Figure 1.

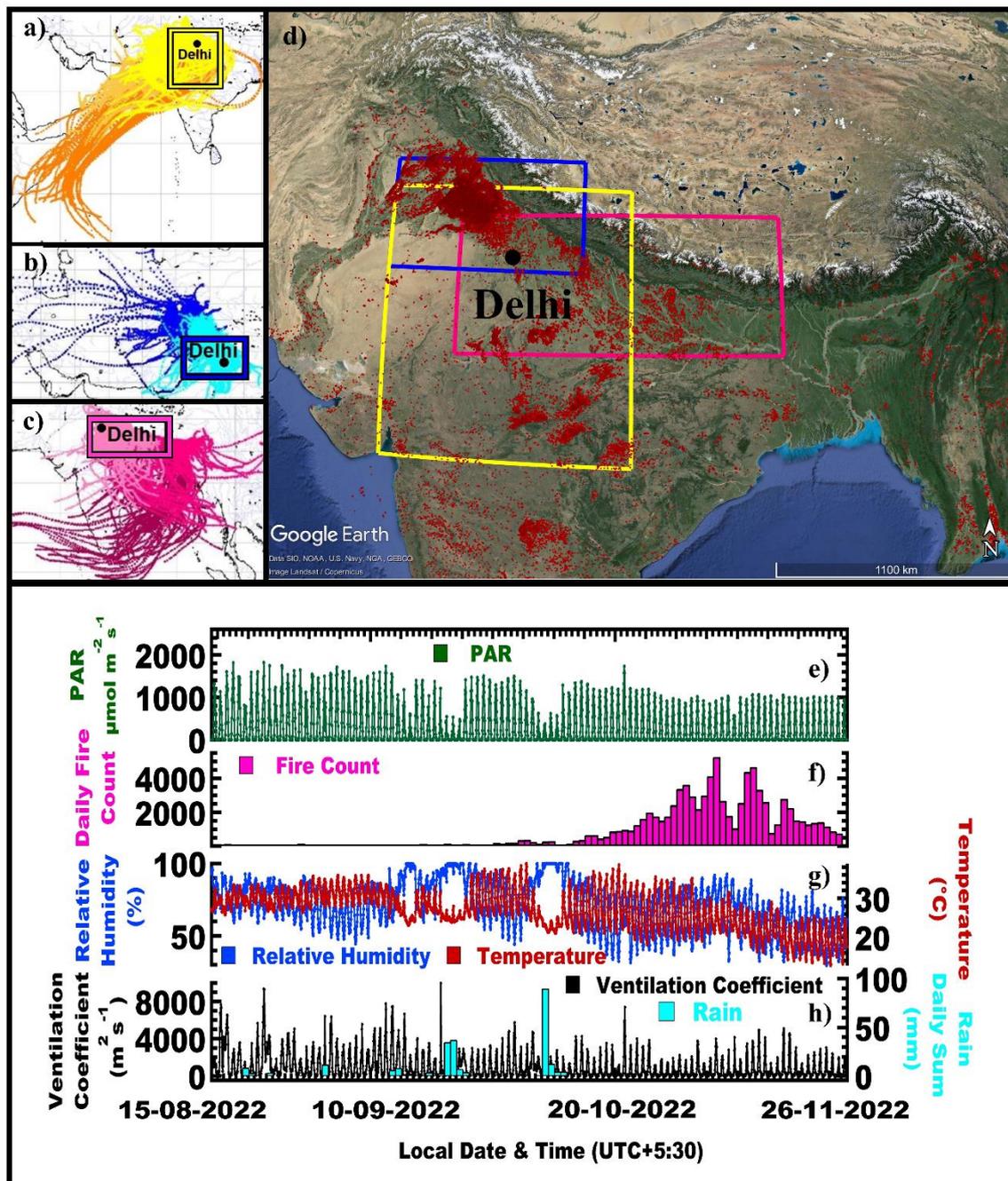


Figure 4: 120 h back trajectory air mass reaching receptor site at Mausam Bhawan building (28.5896°N - 77.2210°E , 50 m above ground level) grouped according to the dominant synoptic scale transport into a) South-Westerly, b) North-Westerly, and c) South-Easterly flow. Square boxes indicate the fetch region from which air masses typically reach the receptor site within 24 hrs for a given flow situation with the d) spatial map of the daily fire counts in the region for the post-monsoon season. The bottom panels show the e) photosynthetically active radiation, f) daily fire counts in the fetch region, g) temperature and relative humidity, and h) the ventilation coefficient and the sum of the daily rainfall for the study period.

Line 536-539: “our PMF results indicate that the actual emissions are slightly smaller than those” “our PMF estimates fall in between those of the EDGARv6.1 inventory and the REASv3.2.1 inventory” I don’t understand how you come to these conclusions, did you calculate emissions out of the PMF concentrations? If yes, please state. If not, I don’t think you can directly compare PMF results and emissions, only in terms of contributions to the total “measured” compounds for each method.

While evaluating the percentage contribution of different sources to the burden of specific pollutants such as PM_{2.5} over a fetch region that is reasonable and related to the atmospheric lifetime of the pollutant in question, the comparison can be considered valid. After all, the lifetime of a given VOC (e.g. benzene) is independent of

their source. Hence the percentage share each source contributes to the measured burden at a site should be proportional to the percentage share the different sources within the fetch region contribute to emissions, provided that the emissions are correctly represented in the emission inventory and the fetch region is chosen suitably small to ensure that emissions from a source within the fetch region can reach the receptor without significant loss. In this study, we are not comparing the absolute concentrations of the PMF and emission inventories, but rather a relative percentage contribution of sources to the total burden. This approach has been routinely used at many other sites of the world (e.g. Buzcu-Guven and Fraser, 2008 <https://doi.org/10.1016/j.atmosenv.2008.02.025>, Morino et al. 2011, <https://doi.org/10.1029/2010JD014762> Li et. al., 2019 <https://doi.org/10.5194/acp-19-5905-2019>; Qin et al., 2022 <https://doi.org/10.1007/s11356-022-19145-7>,) to compare PMF outputs with emission inventories. The reason why absolute concentrations are also brought into the discussion is, that at times the look of pie charts can be deceptive as is the case e.g. for industrial PM_{2.5} emissions. Both the EDGAR and the REAS inventory have almost identical industrial PM_{2.5} emissions in the inventory, yet the pie charts look visibly different, because the larger energy sector emissions and the presence of agricultural burning emissions in the EDGAR inventory visually shrink the size of that “pie slice” compared to how it looks like for the REAS inventory. Looking at the absolute numbers helps to resolve which inventory is more likely to be wrong and for which source. We have reworded this paragraph to make it clearer that we are comparing the relative contribution to the total VOC burden with the relative contribution to the total emissions for the inventory.

Table S6 shows that for residential sector VOCs emissions the absolute emissions in the EDGARv6.1 inventory are almost twice as large as those in the REASv3.2.1 inventory, even though the percentage contribution of this sector to the VOC emissions in the inventory in Figure 10 appears to be similar for both, because of larger VOC emissions from solvent use and industries in the EDGARv6.1 inventory. Both inventories overestimate the relative importance of residential sector emissions in relation to VOC emissions from other sectors by more than a factor of two when compared to our PMF estimate, most likely because they have not been updated with recent fuel shifts to LPG in the relatively prosperous Delhi NCR region.

Line 551: “The EDGARv6.1 inventory significantly underestimates PM_{2.5} & PM₁₀ from agricultural activities” Please backup this statement with a map for example to justify that agricultural emissions should be high.

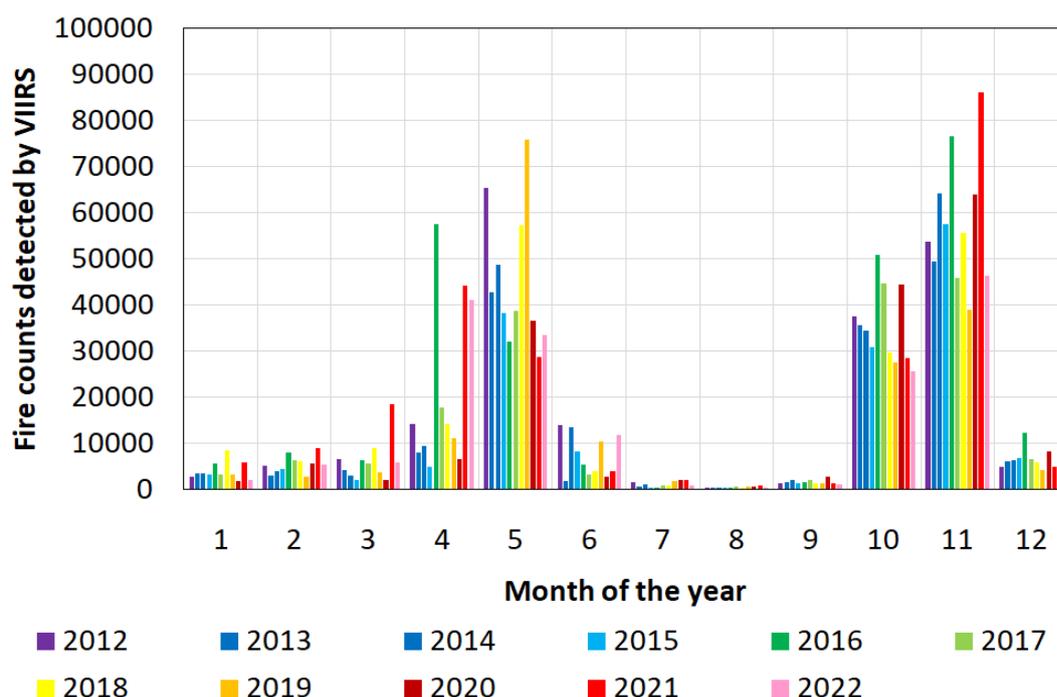
We have already backed up this statement with numbers and a comparison to the FINNv2.5 inventory. Now we also included fire counts in a map in Figure 1, have added images of ash from paddy burning and have simplified the text to make it clearer as follows:

“The REASv3.2.1 inventory completely misses direct VOC and PM emissions from the agricultural sector. The EDGARv6.1 inventory significantly underestimates PM_{2.5} & PM₁₀ emissions from agricultural activities, which include, but are not limited to crop residue burning, in comparison to our PMF results, particularly over NW-India (Table S6). Over this fetch region EDGARv6.1 attributes as much PM_{2.5} to all agricultural activities combined for the full year as the FINNv2.5 inventory (Wiedinmyer et al., 2023) attributes just to agricultural residue burning activities taking place between 15th August and 26th November 2021 (a time period comparable to the period in our model run), without including the emissions from rabi crop residue burning in summer (Kumar et al., 2016) and other agricultural activities such as harvest and ploughing. For PM₁₀ the fire count based FINNv2.5 estimate is twice as high as the emission estimate of EDGARv6.1 for this fetch region, and more likely to be correct, because the phytoliths present in rice straw form coarse mode ash during the combustion process (Figure S10). The fact that EDGAR appears to underestimate residue-burning emissions over this fetch region has been flagged earlier (Pallavi et al., 2019; Kumar et al., 2021; Singh et al., 2023). Our PMF analyses also reveals that the relative contribution of agricultural residue burning to the PM burden over the North-Western IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) and South-Eastern IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) is comparable, despite the much lower fire counts over the South-Eastern IGP (17,810), when compared to the North Western IGP (61,334). This indicates that either fires to the SE are burning closer to the receptor site or the fire detection efficiency in this fetch region is lower. Table S6 reveals that the relative importance of agricultural emissions over the SE fetch region is even more severely underestimated in the FINNv2.5 inventory than in the EDGARv6.1 inventory due to poorer fire detection (close to 100% omission error) for the partial burns prevalent over this region (Lui et al. 2019, 2020, Figure S8) when compared to the complete burns prevalent over the NW IGP (Lui et al. 2019, 2020, Figure S7).”

Line 554-556: There were any more results available from FINNv2.5? “between 15th and August and 26th November 2021 alone” please clarify, was it 15/08-26/11? Then it’s the same length as the current dataset...

The time period matches our observational period just that the data is for the previous year. Unfortunately, 2022 data is not yet available for download, hence we cannot match it with same year data. However, the fact that there are two main crop residue burning seasons of which only one is included in the FINN estimate but both of which should be included in the annual EDGAR number doesn’t change from year to year as can be seen in the figure below.

Fire counts 25-35N 72-82 E



We have now clarified in the text why this period was selected.

Over this fetch region EDGARv6.1 attributes as much $PM_{2.5}$ to all agricultural activities combined for the full year as the FINNv2.5 inventory (Wiedinmyer et al., 2023) attributes just to agricultural residue burning activities taking place between 15th August and 26th November 2021 (a time period comparable to the period in our model run), without including the emissions from rabi crop residue burning in summer (Kumar et al., 2016) and other agricultural activities such as harvest and ploughing.

Table S1: You could add calculated uncertainties and detection limits here. Also, if the “Sr. No” numbers are not used, you can delete them from the table. Are the “Mean” and range values here the detection limits or the averaged concentrations throughout the campaign?

We appreciate the suggestions. We have included the precision error and detection limit used to initiate the model in the supplement Table S1. The mean value is the campaign averaged value and the range represents the minimum and maximum observed throughout the campaign. We now clarify this in the Table caption.

Table S1: 111 NMVOCs species used in the PMF model, the table lists the major compound identifications and the references supporting such assignments from previous works, along with average of the observational period reported in this study (with range min-max), detection limits, precision error.

Table S2 & S3: Same comment about the “Sr. No”.

Deleted

Figure S1: Are these figures referenced in the paper?

Yes, these Figures now numbered as S2-S4 in response to an editors comment are referenced as follows:

Figure 2 shows how the percentage of total VOC, $PM_{2.5}$, and PM_{10} , attributable to various sources changes when the number of factors increases from 3 to 12, while Fig. S2-S4 illustrates the evolution in the factor contribution time series, source profile, and percentage of species explained by different sources when the number of factors in the PMF increases.

Technical corrections

Throughout the paper, add · in units (ex $\mu g \cdot m^{-3}$) done

Title: There shouldn't be an abbreviation in the title, please use volatile organic compounds instead of VOC. done
the revised title is:

Biomass burning sources control ambient particulate matter but traffic and industrial sources control volatile organic compound emissions and secondary pollutant formation during extreme pollution events in Delhi

Line 16: There is a repetition of the word “using”, please change. Revised to:

Here, we source-apportioned VOCs and PM, using a high-quality recent (2022) dataset of 111 VOCs, PM_{2.5}, and PM₁₀ in a positive matrix factorization (PMF) model.

Line 23: Replace “(<2)” by “at least by a factor of 2”.(done)

Line 36: Please reformulate “continues to add”.Revised to:

Delhi with a population of 31.7 million people (UN World Population Prospects 2022), sees an addition of over six hundred thousand vehicles per year (2022 VAHAN-Ministry of Road Transport and Highways (MoRTH), Government of India).

Line 70: Delete the first “source” in “quantify the source contribution of the different sources”.(done)

Line 80: Delete “:” in the title and check all the titles.(done)

Line 113: “in blue” aren’t there other colours used on the graph too? Yes, the revised figure caption reads as follows:

Figure 5: 120 h back trajectory air mass reaching receptor site at Mausam Bhawan building (28.5896°N-77.2210°E, 50 m above ground level) grouped according to the dominant synoptic scale transport into a) South-Westerly, b) North-Westerly, and c) South-Easterly flow. d) spatial map of the daily fire counts in the region for the post-monsoon season with square boxes indicate the fetch region from which air masses typically reach the receptor site within 24 hrs for a given flow situation with the. The bottom panels show the e) photosynthetically active radiation (PAR), f) daily fire counts in the fetch region, g) temperature and relative humidity, and h) the ventilation coefficient and the sum of the daily rainfall for the study period.

Line 116: Correct to “solar radiation as photosynthetically active radiation (PAR)”.(done)

Line 119: Please add the dates of monsoon and post-monsoon seasons. Revised to:

During the monsoon season (15.08-30.09.2022), the air masses from the south-west direction (western arm of the monsoon) were more prevalent than air masses reaching the site from the south-east (Bay of Bengal arm of the monsoon). During the post-monsoon season (01.10-26.11.2022) air masses remain confined over the NW-IGP for prolonged periods and primarily reach the site from the north-west (Fig. 1b), except during the passage of western disturbances (05.10-10.10.2022 and 04.11-10.11.2022), which result in brief periods with south westerly and south-easterly flow and rain (Fig. 1h).

Line 151-152: The structure of the sentence seems wrong, please correct. The sentence has been split into 2 sentences.

The US EPA PMF 5.0 (Paatero et al., 2002, 2014; Paatero & Hopke, 2009; Noris et al., 2014) was applied to a sample matrix of 2496 hourly observations and 111 VOC species. The species with S/N greater than 2.0 were designated as strong species (94) while others were designated as weak species (17).

Line 180-181: There is a repetition of the word “model”, please change. This was a typo. The sentence now reads as follows:

The model was run in the constrained mode elaborately described in Sarkar et al., (2017) and Singh et al., (2023).

Line 190: “T” to delete at the beginning of the paragraph. (done)

Line 180-181 190-191: There is a repetition of the word “using”, please change. Revised the sentence now reads as follows:

The contribution of VOCs to ozone production was derived with the maximum incremental reactivity (MIR) (Carter, 2010) method using the following equation

Line 194: Change to “The secondary organic aerosol production (SOAP)” in small case.(done)

Line 196 & 197: Replace NO_x with NO_x and check this throughout the paper .(done)

Line 197-199: This sentence is a bit unclear. The sentence was split and now reads as follows:

This equation evaluates each VOC species' ability to make SOA in relation to the amount of SOA the same mass of toluene would make when introduced to the ambient environment. This is represented by the SOAP_i.

Line 219: Replace “while” by starting a new sentence with “In addition,”.(done)

Line 220: Replace “are” with “were” and check that it is the right tense throughout the paper. (done)

Line 247: Delete “,” in “(Fig. 4 a & d) were petrol”. This section was shortened in response to a previous comment and now reads as follows

“Figure 4 shows the relative contribution of different sources to the total pollution burden of VOCs, PM_{2.5} and PM₁₀ at the receptor site. In the megacity of Delhi, transport sector sources contributed most (42±4 %) to the total VOC burden, while it contributed much less (only 24 %) to the total VOC burden in Mohali a suburban site 250 km North of Delhi during the same season (Singh et al., 2023). On the other hand, the contribution of paddy residue burning (6±2 %) and the summed residential sector emissions (17±3 % in Delhi and 18 % in Mohali) to the total VOC burden during post-monsoon season were similar at both sites. The contribution of the different factors to the SOA formation potential (Fig. 4e), stands in stark contrast to their contribution to primary particulate matter emissions. SOA formation potential was dominated by the transport sector (54 %) while direct PM₁₀ (52%) and PM_{2.5} (48%) emissions were dominated by different biomass burning sources (Fig. 4 b & c).”

Figure 4: “Photo”, “P2W” & “P4W” could be written in the full name. (done)

Line 252: Delete “,” between “both” & “paddy”.(done)

Line 293: Put “-3” in superscript.(done)

Line 286 & 1288: Delete “,” in “A recent study in Punjab indicated that” and “increased by 0.027 and 0.047 µg·m⁻³ respectively”.(done)

Line 357: There is a repetition of the word “identified”, please change. Deleted

Line 362: I would suggest deleting the sentence “this is consistent with our results”, as “confirms” in line 358 already suggests this.(done)

Line 368-369: Keep “µg·m⁻³” on the same line.(done)

Line 383: Delete the first “source” in “The source fingerprint of this source”.(done)

Line 397: Correct the start of the sentence to “This factor contributes on average more than 30 µg·m⁻³”(done)

Line 397-398: The second part of the sentence, “due to...”, to reformulate and you could reference the added map of surroundings. Rephrased, the sentence now reads as follows:

“This suggests that contrary to 4-wheeler plumes which originate from the immediate vicinity of the site in central Delhi (Figure S1), 2-wheeler plumes reach the receptor after prolonged transport from more distant rural and suburban areas on the outskirts of the city. In such areas, people often favour two-wheelers over four-wheelers.”

Line 399: Add space in “NO (R=0.7)” and correct “CH₄”.done

Line 402 & 404: Once you have written full MTBE and MT, abbreviation is fine. For monoterpenes, you can also write only full name.done

Line 403: There are 2 “,” after “acetaldehyde (1.2 µg·m⁻³)”. This has been rephrased in response to other comments “The main contributors towards the VOC mass in the industrial factor, are in descending order of contribution propyne (C₃H₄), methyl tert-butyl ether (C₄H₈), toluene (C₇H₈), C-8 aromatic compounds (C₈H₁₀), propene (C₃H₆), acetaldehyde (CH₃CHO), methanol (CH₃OH), C-9 aromatics and the sum of monoterpenes (C₁₀H₁₆).”

Line 415-418: This part is a little difficult to read, cf general comment about writing all the values.

We deleted the values and have instead created a figure for the supplement. The revised text reads as follows

“In addition, this factor also explains the largest percentage share of a large suite of volatile and IVOC aromatic hydrocarbons including naphthalene (C₁₀H₈), methyl naphthalene (C₁₁H₁₀), C₁₂H₁₆, C₁₃H₁₈, C₁₃H₂₀, C₁₃H₂₂, C₁₄H₂₀, and C₁₄H₂₂.”

Line 438: Use “acetone + propanal” as before. Changed

Line 452-460: This part is quite difficult to read and understand, cf general comment about writing all the values. **It has been revised as follows:**

As represented by Fig. 9, this factor explains the largest percentage share of a large suite of volatile and IVOC hydrocarbons namely, heptene (C₇H₁₄), C₁₁H₁₂, C₁₂H₁₂, C₁₄H₁₄, C₁₄H₁₈, C₁₆H₂₄, C₁₇H₂₈, and C₁₈H₃₀. In addition, it explains the second largest percentage share of many other IVOC hydrocarbons namely C₉H₁₄, C₉H₁₆, C₁₁H₁₄, C₁₂H₁₆, C₁₃H₁₈, C₁₃H₂₀, C₁₃H₂₂, C₁₄H₂₀, C₁₄H₂₂. Except for the four hydrocarbons C₇H₁₄, C₉H₁₄, C₉H₁₆, and C₁₁H₁₂, all of these IVOCs have been reported to degas at 60°C from asphalt pavement (Khare et al., 2020). So far only C₁₄H₁₈ has been reported as fresh gas phase emissions (transport time <2.5 min) from a farm (Loubet et al., 2022) in ambient air, while C₁₇H₂₈ has been reported in the aerosol phase (Xu et al., 2022). The road construction factor also explains the largest percentage share of a long list of OVOCs namely, C₆ diketone isomers (C₆H₁₀O₂), C₂-substituted phenol (C₈H₁₀O), C₇H₁₂O₂, C₈H₁₄O₂, C₈H₁₆O₂, phthalic anhydride (C₈H₄O₃), which is a naphthalene oxidation product (Bruns et al., 2017), C₉H₁₀O, C₉H₁₂O₂, C₉H₁₄O₂, C₉H₁₆O₂, C₉H₁₈O₂, C₁₀H₁₂O, C₁₀H₁₈O, C₁₀H₈O₃, C₁₀H₁₆O₃, and C₁₂H₁₈O₂. However, out of these only C₁₀H₁₂O and C₁₀H₁₈O have been detected as direct emissions from heated asphalt pavement (Khare et al., 2020) indicating that most OVOCs in this factor are possibly oxidation products of short-lived IVOCs hydrocarbons emitted by this source. This assessment is supported by the volatility oxidation state plot for the road transport factor (Figure S10) which demonstrates that both precursors and oxidation products are present in this factor and that C₆ to C₁₀ hydrocarbons appear to be progressing from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009).

Line 531-532: Keep “y⁻¹” on the same line. **done**

Line 558: Delete “to” in “Our PMF results reveal that to agricultural”. **done**

Line 608: “two criteria air pollutants” do you mean “critical”?

No. India has a National Ambient Air Quality Standards (NAAQS) for six commonly found air pollutants known as criteria air pollutants. PM₁₀ and PM_{2.5} are two of the six criteria for air pollutants regulated under this law. The text has been revised as follows

“While fresh paddy burning was a negligible source of VOCs (6 %), it was the largest source of PM_{2.5} & PM₁₀ (23 % & 25 %) in the Delhi NCR regions during our study period, likely because combustion of phytolite containing rice straw triggers the formation of coarse mode ash (Figure S10) that contributes significantly to the PM burden. PM_{2.5} & PM₁₀ are the two main criteria air pollutants regulated under the national ambient air quality standard that are thought to be the leading cause of the air pollution emergency in November in Delhi annually (Khan et. al., 2023).

Line 622: What is EDGARv6.1 better than in this sentence?

Revised to:

“The PMF results based on primary in-situ data indicate that the EDGARv6.1 inventory provides a better representation of emissions than the REASv3.2.1 inventory for most sectors, with the exception of transport sector emissions and VOC emissions from solvent use. Agricultural burning emissions over the NW-IGP are best represented in FINNv2.5, while agricultural emissions over the SE-IGP are better captured by EDGARv6.1.”

Line 635: Add “in Delhi”: “Despite including the most comprehensive set of organic species in Delhi to date”

Revised to

“Despite including the most comprehensive set of organic species measured in Delhi to date, our study does not include similar information about these other species.”

Line 644: Add “,” after “that” **done**

Line 651: Replace “till date” by “to date” **done**

Response to reviewer 2

The paper “*Biomass burning sources control ambient particulate matter but traffic and industrial sources control VOCs and secondary pollutant formation during extreme pollution events in Delhi*” discusses the sources responsible for air pollution problems in Delhi. For this, they made stationary ambient gas-phase measurements at a prominent location in urban New Delhi and performed source apportionment analysis on the collected data. The chemical profiles of the factors were compared with previous measurements and tracers to identify sources. The work is quite timely since New Delhi is one of the most polluted cities in the world, and regulatory policies are currently being restricted by our limited understanding of the sources in the region.

We thank the referee for recognizing and highlighting the importance of this research work.

Yet I have significant concerns, which I think should be resolved prior to proceeding with publication. Some of my biggest concerns are with the conclusions drawn and stated quite imposingly in the conclusion section. Hence, I'll discuss those first before moving to the next major ones.

Line 606-607: fresh paddy burning is shown to be a negligible source of VOCs but the largest sources of PM_{2.5} and PM₁₀. This is highly confusing to me. PM_{2.5} would be formed from the secondary oxidation of a lot of gas-phase organic molecules emitted from paddy burning. As such it should be emitting precursors of SOA. Or are the authors suggesting that paddy-burning directly emits particulate matter into the atmosphere but no VOCs?

Thank you for seeking this important clarification. Yes, we are suggesting that paddy straw burning is a source of primary aerosol. It is important to note that paddy straw contains a rigid, microscopic structures made of silica known as phytolite. Upon burning, this structure is converted into a glassy ash. The mass of this type of aerosol emitted during the combustion process appears to be quite high when compared to the mass of VOCs emitted in the same combustion process. The high ash formation is a well known fact in engineering circles. Co-combustion of more than 10% of paddy straw alongside with other fuels in power generation units causes severe equipment fouling, due to the potassium (K) rich glassy ash formed (Lui et al. 2022 <https://doi.org/10.1016/j.energy.2022.123950>, Madhiyanon et al. 2020 <https://doi.org/10.1016/j.joei.2020.04.001>). We have inserted a scanning electron microscopy image of ash collected from the electrostatic precipitator of an industrial boiler fired with rice husk and straw to our supplement to illustrating the coarse mode nature of the ash generated during the combustion of phytolite containing biomass. The reality appears to be that much of the aerosol emitted during paddy residue burning is 1) primary and 2) relatively coarse. The root cause of the discrepancy between the contribution to the VOC mass and the contribution to the PM mass appears to be that the glassy ash particles are bigger and have a higher density than organic aerosol and contribute more to the total aerosol mass, that secondary aerosol particles with smaller size and lower density. Just like dust, this ash cannot be detected by AMS and since the chemical composition is >96% SiO₂ with minor amounts of K, any routine CMB analysis would likely attribute this type of aerosol to the natural dust fraction. This explains why earlier studies may have failed to recognize the importance of ash aerosol. We have added a figure (Figure S10) and revised the text to reflect this more clearly as follows:

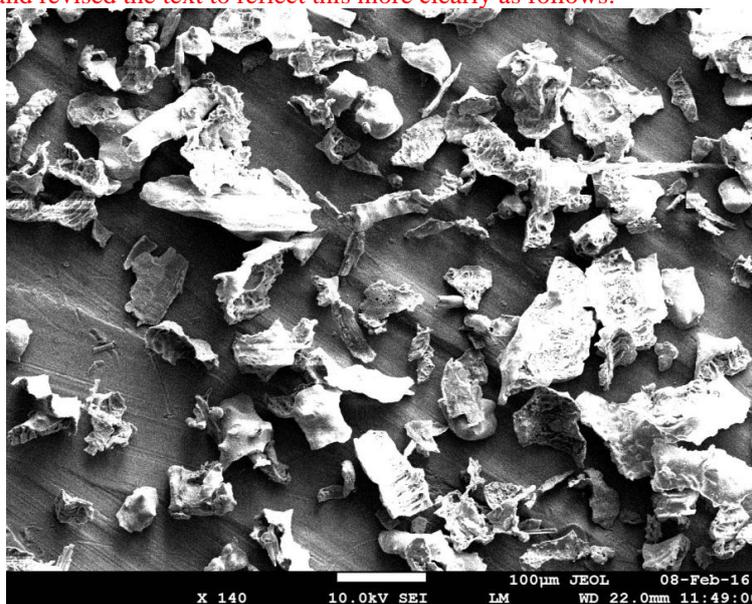


Figure S10 SEM image of rice ash from the electrostatic precipitator of an industrial boiler fired with rice husk

and straw, illustrating the coarse mode nature of the ash generated during the combustion of phytolith containing biomass.

“While fresh paddy burning was a negligible source of VOCs (6 %), it was the largest source of PM_{2.5} & PM₁₀ (23 % & 25 %) in the Delhi NCR regions during our study period, likely because combustion of phytolite containing rice straw triggers the formation of coarse mode ash (Figure S10) that contributes significantly to the PM burden. PM_{2.5} & PM₁₀ are the two main criteria air pollutants regulated under the national ambient air quality standard that are thought to be the leading cause of the air pollution emergency in November in Delhi annually (Khan et al., 2023).”

It is, however, important to note that we do not claim that paddy burning is not a VOC source. In Section 3.2., we clearly state that it is the largest source of a relatively long list of VOCs.

The main point that we are making is that in terms of its relative contribution to the overall pollution levels of certain pollutants when compared to other sectors such as e.g. road transport, this activity is far more important as a PM_{2.5} and PM₁₀ source than it is as a VOC source/SOA precursor source. We have shifted the supplementary figure to the main text and revised the text to make this clearer. The revised text to clarify this aspect reads as follows:

“Figure 6 shows that this factor explained the largest percentage share of O-heteroarene compounds such as furfural (C₅H₄O₂), methyl furfural (C₆H₆O₂), hydroxy methyl furfural (C₆H₆O₃), furanone (C₄H₄O₂), hydroxymethyl furanone (C₅H₆O₃), furfuryl alcohol (C₅H₆O₂), furan (C₄H₄O), methyl furans (C₅H₆O), C2-substituted furans (C₆H₈O), and C3-substituted furans (C₇H₁₀O), which are produced by the pyrolysis of cellulose and hemicellulose, and have previously been detected in biomass burning samples (Coggon et al., 2019; Hatch et al., 2015; 2017; Koss et al., 2018; Stockwell et al., 2015). Figure 6 also shows that this factor explains the largest share of the most abundant oxidation products that result from the nitrate radical-initiated oxidation of toluene as well as from OH-initiated oxidation of aromatic compounds under high NO_x conditions, namely nitrotoluene (C₇H₇NO₂) and nitrocresols (C₇H₇NO₃) (Ramasamy et al., 2019), which indicates a certain degree of aging of the plumes. These nitroaromatic compounds are significant contributors to SOA and BrC, (Palm et al., 2020, Harrison et al., 2005). It also explains several other nitrogen containing VOCs such as nitroethane (C₂H₅NO₂), the biomass burning tracer acetonitrile (CH₃CN) and pentanenitrile (C₅H₉N). The presence of pentanenitrile isomers in biomass burning smoke has previously been confirmed using gas chromatography-based studies (Hatch et al., 2015, Hatch et al., 2017). In addition the factor explains the largest percentage share of acrolein (C₃H₄O), hydroxyacetone (C₃H₆O₂), cyclopentadienone (C₅H₄O), cyclopentanone (C₅H₈O), diketone (C₄H₆O₂), pentanedione (C₅H₈O₂), hydroxybenzaldehyde (C₇H₆O₂), guaiacol (C₇H₈O₂), and the levoglucosan fragment (C₆H₈O₄), many of these compounds are known to form during lignin pyrolysis (Hatch et al., 2015, Koss et al., 2018; Nowakowska et al., 2018), while dimethylbutenedial (C₆H₈O₂), trimethylbutenedial (C₇H₁₀O₂) are ring opening oxidation products of aromatic compounds (Zaytsev et al., 2019).”

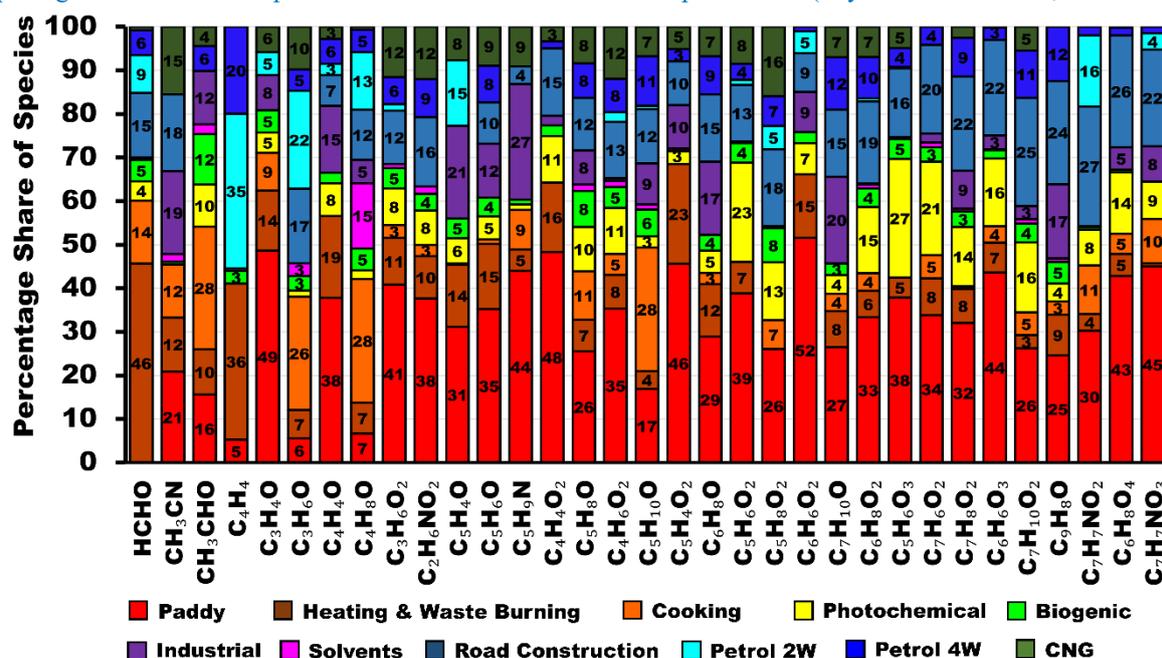


Figure 6: VOC species to which different forms of biomass burning contribute the highest percentage share of the atmospheric burden in Delhi

Is it possible that the PTR-TOF did not measure or fragment a lot of precursor species emitted from paddy burning?

The proton transfer reaction technology is a soft ionization technique and the operating conditions of 120 Td during the deployment further facilitate negligible fragmentation. In addition, the instrument deployed in this work was equipped with extended volatility range technology which has been missing from previous PTR-TOF studies conducted in Delhi. This has been explained in detail in the companion paper (Mishra et al., 2024) and such a system enabled us to detect and measure an unprecedented long list of IVOCs emitted from other sources (industries and asphalt paving), hitherto undetected in ambient gas phase observations without fragmentation. The PMF VOC source signature further matched observational data obtained via source samples collected directly on burning paddy fields. Hence, there is no evidence of loss of VOCs. It is also important to note that we could measure a lot of SOA precursor species and some of their first-generation oxidation products e.g. nitrotoluene ($C_7H_7NO_2$), nitrocresols and ring opening oxidation products of aromatic compounds such as dimethylbutenedial ($C_6H_8O_2$), trimethylbutenedial ($C_7H_{10}O_2$) and could successfully attribute them to paddy burning factor. However, the total VOC mass attributed to this factor $11.6 \mu\text{g m}^{-3}$ is less than the $PM_{2.5}$ ($20.7 \mu\text{g m}^{-3}$) and PM_{10} ($36.5 \mu\text{g m}^{-3}$) mass attributed to this factor. Since the factor has a photochemical age of less than 24 hours, and the SOA yields (in terms for % of mass converted to aerosol) for many compounds on such timescale are <20%, the overall SOA contribution to the $PM_{2.5}$ mass is smaller than the mass contribution of primary ash particles. Thus, it is highly unlikely that the PTR-TOF-MS missed measuring many precursors due to inlet losses or that the compounds fragmented massively in the system used in this work. We have added the following text and supplementary figure S9 and S10 to the end of this section to make this clearer:

“Figure S9 shows the volatility oxidation state plot for all 111 VOCs in which the marker size represents the percentage share of each compound explained by the paddy residue burning factor and markers are colour coded by the number of carbon atoms. The plot shows evidence of the first- and second-generation oxidation products of C5 and C6 hydrocarbon transitioning from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009), while C7 hydrocarbons progress along trajectories expected for both the addition of -OH and =O functionality. This indicates that paddy residue burning contributes significantly to the SOA burden. However, the fact that the PM_{10} mass associated with this factor ($36.5 \mu\text{g m}^{-3}$) is 1.8 times larger than the $PM_{2.5}$ mass ($20.7 \mu\text{g m}^{-3}$) and 3 times larger than the VOC mass ($11.6 \mu\text{g m}^{-3}$) released during the same combustion process, points towards the relatively coarse ash formed from the phytolith skeleton of rice straw (Figure S10) as the dominant aerosol source.”

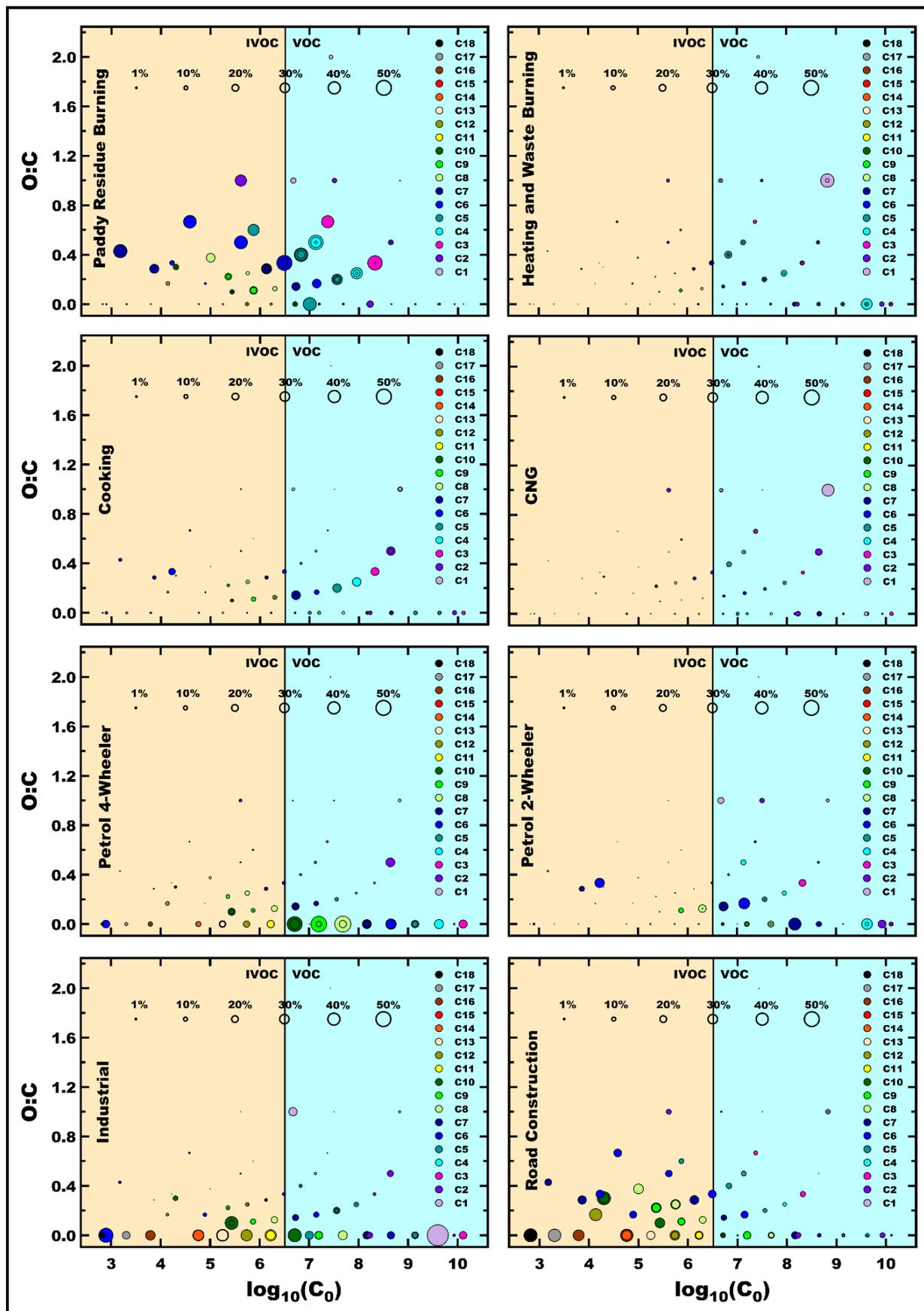


Figure S9: Volatility oxidation state plots for all factors that individually contribute more than 3% to the total SOA formation potential.

Line 620 (also 566-568): “The transport sector’s PM emissions are dominated by the non-exhaust emissions of the CNG-fuelled commercial vehicle fleet.” This sounds somewhat unlikely. Which non-exhaust emissions are the authors referring to emitting from CNG vehicles? I can think of break/tyre-wear as a possible source but that contributes primarily to coarse PM, not so much to fine. Are there evaporative emissions of some kind? I imagine CNG itself would have negligible potential to form ambient PM given its small molecular size.

Yes, tire wear, break wear and dust re-suspension are precisely the sources we are implicating and those sources are well supported by the PMF output, because the PM₁₀ emissions attributed to this source in the PMF (22.5 μgm⁻³) are indeed are twice as large as the PM_{2.5} emissions attributed to this source (10.4 μgm⁻³). According to a recent emission inventory for Delhi (Nagpure et al., 2016 <https://doi.org/10.1016/j.atmosenv.2015.12.026>) their contribution to transport sector PM is one order of magnitude larger than that of the tailpipe emissions and two orders of magnitude larger than the VOC mass. The SOA formation potential of the dominant VOC in the tailpipe exhaust of this vehicle class, methanol, and ethanol is very small hence SOA is not a significant contributor to the PM mass associated with the CNG factor. We have now expressed this more clearly.

“Also, sources that are generally targeted by most clean air action plans such as tailpipe exhaust emissions of private vehicles and industries are responsible for less than one-quarter of the particulate matter mass loading that can be traced with the help of gas-phase organic molecular tracers. Instead, the transport sector’s PM emissions are dominated by the non-exhaust emissions such as road dust suspension, break wear and tire wear of the CNG-fueled commercial vehicle fleet, which according to a recent emission inventory for Delhi are one order of magnitude larger than the transport sector tailpipe exhaust emissions (Nagpure et al., 2016). “

On the other hand, the transport sector in Delhi would have diesel trucks which are known to be large emitters of SOA precursors.

The majority of heavy-duty vehicles in Delhi, have transitioned to CNG fuel. So have internal delivery vehicles and most taxis. There are strict restrictions on the entry of diesel trucks. This shift aligns with Delhi's strict adoption of Euro-6 norms in 2018, and restrictions that completely ban the use of more than 10-year-old diesel vehicles within city limits, which forces owners to sell these into the second hand market of less restrictive states or convert them to CNG with a conversion kit. To incentivize cleaner technologies like CNG kit (Krelling et al., 2022 <https://doi.org/10.1016/j.tranpol.2021.10.019>), the administration heavily subsidizes the price of CNG which was 1.5 times lower per km than that of diesel during the study period. Commercial transport is price sensitive, hence, the number of diesel vehicles on the roads is very low. Diesel-fuelled trucks typically circumvent city borders when they pass by Delhi. This practice is influenced by heavy fines on entry of old vehicles and also stringent time regulations imposed on the entry of diesel trucks. The recent changes in both regulations and their enforcement have resulted in halving the diesel sales in the Delhi NCR over the past 5 years. A random selection of pictures clicked in the timespan of less than 10 minutes while driving around Delhi pasted below supports the fact that diesel trucks are hardly plying across the city and hence not important enough to get their own PMF factor. The CNG cylinders mounted in the place where the diesel tank used to be are easily visible on most trucks. We have now inserted the following supplementary figures and text into Section 3.2.4. to clarify:

“This study attributes a large share of these non-tailpipe emissions to trucks, buses and other commercial vehicles that are typically fuelled by CNG, because commercial diesel vehicles of <10 years age face severe entry restrictions, that limit their use within the Delhi NCR while older diesel vehicles have been completely banned from plying within City limits. Policy interventions in favour of CNG use (Krelling & Badami, 2022) have resulted in a halving of diesel sales, a rapid conversion of Delhi’s HDV fleet to CNG (Figure S12), and a significant reduction in tailpipe exhaust emissions.”



Figure S12: Random selection of photographs clicked while driving around Delhi. One can clearly see the white CNG cylinders mounted in the place where the fuel tank used to be during vehicle conversion. Photo credits: Kriti Annika Sinha

Dust resuspension has been attributed to non-exhaust emissions, but I am not sure if I agree with that classification. Dust is not a vehicular source. Hence, I would like the authors to extensively elaborate on what forms PM from non-exhaust emissions from CNG vehicles.

Dust can be natural and windblown but it can also be anthropogenic. When dust is suspended from the road or more importantly by off-road usage of heavy vehicles e.g. during construction activity it is classified as an anthropogenic transport sector emission (see e.g. the recent review by Harrison et al. 2021, <https://doi.org/10.1016/j.atmosenv.2021.118592>). Particularly when it comes to avoidable road transport, these emissions would not be present if the same transport demand had been satisfied via freight train/passenger train/metro rail/tram. The road transport share of these anthropogenic emission can be reduced by 1) modal shifts of the transport demand towards rail and 2) wet road cleaning to reduce the silt burden on the road in case of road dust suspension and 3) regenerative braking in case of EVs. However, certain types of activity by HDVs that suspend a very large amount of dust e.g. during transport of construction material to construction sites or during the movement of waste are hard to eliminate. A picture of a waste disposal truck (marked with a red arrow) moving on one of the Delhi's "garbage mountains" with a dust plume (encircled with a red square) trailing behind the vehicle is pasted below. We hope that the reviewer agrees such a "dust plume" should not be labelled as a "natural" dust plume, even though in the case of trucks moving on construction sites the suspended dust aerosol will be chemically indistinguishable from natural soil minerals suspended by wind alone. Since the anthropogenic dust is transported in the air alongside with the tailpipe emission plume, our PMF is capable of tracking this type of "dust plumes" to their primary source, namely, HDV movement.



We also modified the text in Section 2.4 to make things clearer:

“CNG-fuelled vehicles are identified as the third largest source of PM₁₀ (15 %) and PM_{2.5} (11 %) and contribute 9 % to the total VOC burden (Fig. 4). The much higher contribution of this source to the coarse mode particulate matter burden (22.5 μgm⁻³ PM₁₀) when compared to the fine mode particular matter burden (10.4 μgm⁻³ PM_{2.5}), confirms earlier emission-inventory-based estimates which flagged that non-tailpipe emissions such as brake and tire wear and road dust resuspension have become the dominant transport sector related particulate matter sources in the Delhi-NCR region (Nagpure et al., 2016). Non-tailpipe emissions such as brake and tire wear and road dust resuspension contribute most to the PM₁₀ burden, although they have also become the largest source of transport sector fine mode aerosol and VOC emissions in some countries that have transitioned to Euro-6 norms (Harrison et al., 2021).”

This also reads somewhat contrary to lines 260-264 where petrol vehicles are shown to be major contributors to SOA.

We did not state that petrol vehicles contribute most to the SOA at the receptor sited. Petrol 4W and 2W vehicles contribute significantly to the SOA formation potential because they are the largest source of several aromatic compounds in the NCR. However their contribution to the PM₁₀ (10.8 μgm⁻³) and PM_{2.5} (4 μgm⁻³) mass in the PMF is much smaller than that of CNG vehicles, primarily because non-tailpipe emissions such as break-wear and dust suspension scale with vehicle weight and these vehicles tend to be lightweight. It is very important to keep in mind that 1) SOA particles are smaller and 2) SOA particles are less dense (~1.4 g/cm³) than break wear (up to 5 g/cm³) and road dust (~2.6 g/cm³) and 3) only a small percentage share of the VOC burden is converted to SOA. This is why the non-tailpipe PM emissions of vehicles that are following emissions norms of EURO 4 or better tend to be larger than the tailpipe aerosol emissions and the mass of the SOA formed. However, a large contribution to the SOA formation potential at the receptor site does not necessarily equate a large contribution to the SOA mass at the receptor site, because the 4-wheeler emissions impacting the site mostly have a photochemical age on the timescale of minutes. While the factor contributes the largest percentage share of many C6 to C10 hydrocarbons, it hardly contributes towards the measured mass of the first- and second-generation oxidation products of those very same compounds at the receptor site. We have added the following text to make this clear:

“Figure S9 shows that this factor contributes significantly to the burden of C6- to C10 hydrocarbons, and hence SOA formation potential. However, due to freshly emitted plumes, it hardly contributes to the burden of the first- and second-generation oxidation products of these hydrocarbons at the receptor site. Instead, this factor is likely to contribute to secondary pollution formation downwind of the Delhi NCR.”

while a distinction has been made between 2-wheeler and 4-wheeler petrol vehicles, no significant discussion exists on the contribution of diesel vehicles. This needs to be explained in more detail.

Diesel vehicles do have a distinctive source fingerprint and would have been identified by the PMF if they had major impact on the air quality in the Delhi NCR. They do not have major impact, because policies have restricted their usage in recent years including 2022. The only noticeable diesel emissions impacting the receptor site are mixed into the road construction factor and reach the receptor simultaneously with the evaporative emissions of freshly laid asphalt. This is now clarified in section 3.1 as follows:

The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving ($R=0.9$, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles ($R=0.6$, Che et al., 2023).

Line 650-651: Authors state that “all” previous studies from the region have attributed PM to BB or fossil-fuel burning, and that we need to look beyond these sources. While I agree that a larger set of sources need to be identified, I think there is already some work done on this front. Kumar et al. 2022 ACP <https://acp.copernicus.org/articles/22/7739/2022/acp-22-7739-2022.pdf>

Kumar et al. 2022 identified “cooking-related OA using EESI-TOF PMF analysis, but sadly did not include this source in the final pie charts of their paper, because their AMS-PMF analysis could not find this particular source, while ESI data was only reported in counts per second and not quantitatively. Instead, the paper reported three primary factor HOA, BBOA-1 and BBOA-2 and one aged biomass burning factor in addition to three SOA factors without naming the activities responsible for the formation of BBOA-1 and BBOA-2. The paper that reported results that most closely relate to the results of our study would be Cash et al. 2021. We have added a discussion of the Cash et al results to these lines

“This new approach of combining VOC tracers with PM measurements provides great potential for improved source apportionment in complex emission environments, at a level of detail that is more meaningful than just attributing emissions to biomass burning or fossil-fuel burning, which has been the case in all previous studies from the region to date. Previously in Delhi-MCR region, Kumar et al. 2022 identified “cooking-related OA using EESI-TOF analysis but due to analytical limitations, the paper only reported quantitative pie charts for three primary factors, namely HOA, BBOA-1 and BBOA-2, without naming the activities responsible for the formation of BBOA-1 and BBOA-2. One of the more comprehensive AMS based studies (Cash et al., 2021) spanning pre-monsoon, monsoon and post monsoon season of the year 2018 identified three different primary biomass burning factors, namely cooking organic aerosol (6% of PM_{10}), solid fuel organic aerosol ($\leq 11\%$ of PM_{10}), and semi-volatility biomass burning organic aerosol ($\leq 13\%$ of PM_{10}), that broadly appear to correspond to our solid fuel-based cooking (4% of PM_{10}), residential heating and waste burning (23% of PM_{10}), and paddy residue burning (25% of PM_{10}) factors. However, the study failed to name and attribute two of these three factors in policy relevant ways, could not identify the significant contribution of coarse mode fly ash to the total aerosol burden, and also was unable to distinguish between different fossil-fuel related sources.”

Figure 5: I notice that road construction and solvent factors show opposing temporal trends. Road construction peaks in the afternoon while solvents are higher during early morning or night hours. The authors state in lines 425-426 that the solvents contribute the most to the VOC burden at night. Given that both these sources are evaporative in nature, how could they show opposing temporal trends? Are there any specific sources of solvents in Delhi that are prominent during nighttime?

The road construction factor primarily involves evaporative emissions released during degassing from the road surface, emissions are greatest when the asphalt has just been paved and hence peak during the hours when construction activity is more prevalent. Conversely, compounds associated with the solvent factor predominantly originate from an industrial point sources that appears to operate 24/7. They reach the receptor in episodic but intense plumes that are not accompanied by combustion tracers. Hence, we attribute this factor to the venting of chemicals from some industrial stacks. This type of activity results in the highest concentrations at night when emissions mix into a shallower nocturnal boundary layer. However, the factor also displays episodes with high concentrations during daytime. This indicates that daytime fluxes can actually be quite high and just mix into a larger volume. We looked for some specific types of industrial units located 1) SW of the receptor with the highest source strength after midnight and 2) NE of the receptor with the highest source strength in the evening before midnight. The best source match we found for this type of source was collected from a plot situated opposite a polymer manufacturing unit and next to a pet food manufacturer in an industrial area at Jahangirpuri ($R=0.7$) N of the receptor. We have updated the text in Section 3.1 and lines 425ff to make the difference clearer:

“The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving ($R=0.9$, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles ($R=0.6$, Che et al., 2023). The factors identified as solvent usage and evaporative

emissions matched ambient air grab samples collected from an industrial area at Jahangirpuri (R=0.7), and Dhobighat at Akshar Dham (R=0.5) in this study.”

“These compounds point towards stack venting of VOCs from chemical-, food-, or pharmaceutical industries or polymer manufacturing as likely sources of these emissions (Hodgson et al., 2000, Villberg et al., 2001, Jankowski et al., 2017, Gao et al., 2019). This assessment is broadly confirmed by the fact that the best match (R=0.7) for this source was collected from a plot situated opposite a polymer manufacturing unit and next to a pet food manufacturer in an industrial area at Jahangirpuri N of the receptor site.”

One can also check the temporal trends in PCBTF, Texanol and p-dichlorobenzene, D4- and D5-siloxane that are known tracers of VCP sources. Some of these can be measured with PTR-ToF.

Thank you for this comment. We appreciate that Volatile Chemical Products (VCP) have emerged as an important source in recent studies conducted in western countries, in which emissions from industrial and other sources have been regulated to a greater extent and VCP emissions from cosmetic and perfume usage and that of sanitation products have transitioned to becoming major sources of some VOCs in the urban environment. In India, however, these products are expensive and can only be afforded by a limited subset of the urban population. The vast majority of the population, even in a city like Delhi, struggle to meet their daily food needs and expenses on education and healthcare, which likely has kept VCPS from being a major source so far. Hence it is not surprising that this factor did not appear in the 11-factor PMF solution. Except for dichlorobenzene none of the compounds passed our quality control filter, which means either the signal at the m/z was not above the detection limit or the compound presence could not be confirmed via isotopic peaks of the correct height for the isotopes of the compound. At our site, the dominant dichlorobenzene sources appear to be industrial in nature.

The authors should more clearly discuss how they calculated the total VOC mass in the paper. This is important because the fractions of other measured species are drawn from the total, and this can introduce significant bias in the conclusions regarding source contributions if the total VOC mass is not comprehensive enough.

We have added a few sentences to clarify this point:

“The total VOC mass was included as a weak species and was calculated as the sum of the mass of the individual 111 VOC species included in the PMF. Overall, the 111 VOC species included in our analysis and their isotopic peaks explained 86% of the VOC mass detected during our study period. The remaining 119 m/z that accounted for 14% of the detected VOC mass could not be included in our PMF analysis mostly because signals were below the detection limit for close to 50% of the observation period, or because compound identity could not be confirmed via isotopic peaks.”

The chemical profiles shown in Figure 3 run up to C₁₀H₁₆ and there is some additional discussion in the paper about IVOCs. However, sources such as road construction emit minimally in the VOC space, and more in the IVOC and SVOC space. The authors should discuss how they prevented biases from creeping into their conclusions.

The main rationale behind choosing these compounds shown in Figure 3 is, that the normalized height of the bar displaying that compound should be at least 0.1x the height of the tallest peak in at least one of the factor profiles. Most of the IVOCs did not meet the inclusion criteria for Figure 3. The figure serves to depict the chemical fingerprints of all the factors and very low bar heights are invisible on the y-axis, while the compound formula will clutter the x-axis of the figure and will make it hard to read. We have now clarified this as follows:

“Figure 3 shows the source profile of the eleven factors that our PMF analyses resolved. Out of the 111 VOCs only those whose normalized source contribution exceeded 0.1 when divided by the most abundant compound in the same source profile in at least one of the sources, were included in the figure.”

Fresh asphalt does have a very characteristic VOC signature in the OVOCs space in the form of a very distinct double peak at C₃H₆O and C₄H₈O that is not accompanied by methanol peaks as is usually the case in a solvent factor, and neither accompanied by the furanes, aldehydes and organic acids that are usually seen in biomass burning source fingerprints. This makes the identification so easy and clear. The factor identity has been confirmed via cross correlation analysis as follows:

“The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving (R=0.9, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles (R=0.6, Che et al., 2023).”

Also, there should be at least some discussion in the paper about the inlet system used upstream of the PTR-ToF as this can prove crucial in the detection of many species (lines 132-133).

A detailed discussion about this as well as the inlet system is given in the companion paper (Mishra et al., 2024 10.5194/egusphere-2024-500) and now also mentioned and included in the revised MS as follows:

“It is worth mentioning again that as a significant improvement over other previous PTR-TOF-MS deployments in Delhi, the inlet system of the instrument used in this work was designed for sampling and detection of low-volatility compounds with the extended volatility range technology (Piel et al., 2021). The inlet system of the instrument as well as the ionization chamber is fully built into a heated chamber and the inlet capillary is further fed through a heated hose to ensure there are no “cold” spots for condensation. The entire inlet system is made of inert material (e.g. PEEK or siliconert treated steel capillaries to keep surface effects minimal. Further the overall inlet residence time was less than 3 seconds, throughout the campaign.”

Piel, F., Müller, M., Winkler, K., Skytte af Sætra, J., and Wisthaler, A.: Introducing the extended volatility range proton-transfer-reaction mass spectrometer (EVR PTR-MS), *Atmos. Meas. Tech.*, 14, 1355–1363, <https://doi.org/10.5194/amt-14-1355-2021>, 2021.

Furthermore:

Lines 182-184: The “pulling up” and “pulling down” should be briefly explained. It sounds vague in its current form.

We regret the confusion, as "pulling up" and "pulling down" are normally common terms used in PMF discussions. We have added the following text to clarify:

“The rotational ambiguity can be reduced using this option with the aid of prior knowledge by encouraging the model to minimize (pull down) or maximize (pull up) the total mass assigned to specific hourly observations or compounds in source profiles as much as possible within a pre-defined permissible penalty on Q.”

Lines 187-188: It is quite amazing that the bootstrap found all 100% of the runs stable and well-mapped to the base solution. In principle, this may suggest that your dataset yields only one solution which is super robust. Is this what you are saying?

The reviewer has interpreted this correctly. 100% stable bootstrap solutions for the constraint run indeed indicate that with the help of the constraints a robust solution has been achieved for this particular dataset.

I acknowledge citations, but in lines 180-187, I recommend briefly describing the rationale behind application of different constraints to help the reader assess.

Thank you for highlighting this important aspect. As can be seen from our list of constraints this has been primarily accomplished by pulling down some night-time plumes from combustion sources that the model’s base model run had left behind in the BVOC and photochemistry factors. Stopping the model from mixing some of the combustion emissions into non-combustion sources turns out to be an extremely efficient way to force the model to resolve the combustion sources properly, and prevent factor swapping. The solution can be further refined by identifying one major characteristic plume per source and pulling it up. The selected plumes should originate from one dominant source only, should have plume enhancement ratios that matches source samples as closely as possible, and should be among the strongest plumes attributed to the source in the base run. Pulling up as little as one such plume for each of the anthropogenic sources, minimizes the factor swapping between similar sources

We have added the following details:

“The rotational ambiguity can be reduced using this option with the aid of prior knowledge by encouraging the model to minimize (pull down) or maximize (pull up) the total mass assigned to specific hourly observations or compounds in source profiles as much as possible within a pre-defined permissible penalty on Q. The primary problem of the base run solutions is that night-time biomass burning plumes contaminate both the biogenic and the photochemical factor. To minimize this in our constrained run, we have pulled down primary emissions (acetonitrile, toluene, C8 aromatics, and C9 aromatics) in the biogenic and photochemical factors. We also pulled down the top-7 strongest nighttime plumes contaminating the biogenic and photochemical factors. In addition, we pulled up the highest plume event for all the anthropogenic emission-related factors as detailed in Table S2. The overall penalty to Q (the object function) was 4.9 %, which is within the recommended limit of 5 % (Norris et al., 2014; Rizzo & Scheff, 2007).”

Lines 229-234: The comparisons stated here are very on point, which is great. But it is not clear how contributions from heavy vehicles, e.g. road construction vehicles, were separated from other diesel-based sources, such as transport trucks.

We appreciate the referee’s comment. As also mentioned in the above replies, the majority of heavy-duty vehicles in Delhi have been switched to CNG. This includes all commercial diesel vehicles such as trucks, buses and taxis. Even private diesel vehicles that are older than 10 years are not permitted to ply in the city. This is why the only diesel emissions we see in the PMF are emissions that reach the receptor site during daytime simultaneously with compounds that are diagnostic of fresh asphalt paving. It appears that construction machinery is the dominant HDV class that has not yet completely been converted. The revised text now reads as follows:

“The factors identified as CNG (R=1.0), petrol 4-wheelers (R=0.9), and petrol 2-wheelers (R=0.6) matched tailpipe emissions of the respective vehicle types and fuels (Hakkim et al., 2021). The petrol 4-wheelers (R=0.9), and petrol 2-wheelers (R=0.7) also matched traffic junction grab samples from Delhi (Chandra et al., 2018). The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving (R=0.9, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles (R=0.6, Che et al., 2023).”

I recommend to put some correlation plots in the SI that compare the chemical profiles of the source factors obtained in this study with the sources from literature that are discussed here.

We thank the referee for the suggestion. We have now modified the text (see below) and added the R values to the main text. but don't think correlation plots are necessary since the source profiles as presented in the revised version can already be visually compared with the help of Figure 3.,

“The PMF factor profile matched best against source samples collected from burning paddy fields (R=0.6, Kumar et al., 2020) for the paddy residue burning factor. The cooking factor matched emissions from a cow-dung-fired traditional stove called angithi (R=0.7, Fleming et al., 2018). The residential heating & waste burning factor had a source fingerprint matching emission from leaf litter burning, (R=0.7, Chaudhary et al., 2022), waste burning (R=0.7, Sharma et al., 2022), and cooking on a chulha fired with a mixture of firewood and cow dung (R=0.9, Fleming et al., 2018). The factors identified as CNG (R=1.0), petrol 4-wheelers (R=0.9), and petrol 2-wheelers (R=0.6) matched tailpipe emissions of the respective vehicle types and fuels (Hakkim et al., 2021). The petrol 4-wheelers (R=0.9), and petrol 2-wheelers (R=0.7) also matched traffic junction grab samples from Delhi (Chandra et al., 2018). The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving (R=0.9, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles (R=0.6, Che et al., 2023). The factors identified as solvent usage and evaporative emissions matched ambient air grab samples collected from an industrial area at Jahangirpuri (R=0.7), and Dhobighat at Akshar Dham (R=0.5) in this study. The factor identified as industrial emissions showed the greatest similarity to ambient air grab samples from the vicinity of the Okhla waste-to-energy plant (R=0.8), Gurugram (R=0.7) and Faridabad (R=0.8) industrial area. The biogenic factor showed the greatest similarity to leaf wounding compounds released from *Populus tremula* (R=0.8, Portillo-Estrada et al., 2015) as well as BVOC fluxes from *Mangifera indica* (R=0.4, Datta et al., 2021).”

Lines 252-253: As a reader, I was surprised to see a comparison with NW-IGP and Mohali. It was quite sudden and not consistent throughout the paper. This should be rephrased in a way that gives a reader some context on which regions are being compared and why.

We have rephrased this as follows:

“Figure 4 shows the relative contribution of different sources to the total pollution burden of VOCs, PM_{2.5} and PM₁₀ at the receptor site. In the megacity of Delhi, transport sector sources contributed most (42±4 %) to the total VOC burden, while it contributed much less (only 24 %) to the total VOC burden in Mohali a suburban site 250 km North of Delhi during the same season (Singh et al., 2023). On the other hand, the contribution of paddy residue burning (6±2 %) and the summed residential sector emissions (17±3 % in Delhi and 18 % in Mohali) to the total VOC burden during post-monsoon season were similar at both sites.”

Lines 262-270: Add error values to the average percentages to account for the variability in these fractions during the study period.

This text has been substantially simplified. Uncertainties have been added to the segments that were retained. However, uncertainties reflect the uncertainty of the PMF model imposed by the stability of the bootstrap runs, not the ambient variability. No uncertainty was added to the SOA formation potential because the uncertainty of the widely used SOAP factors has not been quantified so far. This prevents meaningful error propagation from VOC mass to SOAP.

“Figure 4 shows the relative contribution of different sources to the total pollution burden of VOCs, PM_{2.5} and PM₁₀ at the receptor site. In the megacity of Delhi, transport sector sources contributed most (42±4 %) to the total VOC burden, while it contributed much less (only 24 %) to the total VOC burden in Mohali a suburban site 250 km north of Delhi during the same season (Singh et al., 2023). On the other hand, the contribution of paddy residue burning (6±2 %) and the summed residential sector emissions (17±3 % in Delhi and 18 % in Mohali) to the total VOC burden during post-monsoon season were similar at both sites. The contribution of the different factors to the SOA formation potential (Fig. 4e), stands in stark contrast to their contribution to primary particulate matter emissions. SOA formation potential was dominated by the transport sector (54 %) while direct PM₁₀ (52±8%) and PM_{2.5} (48±12%) emissions were dominated by different biomass burning sources (Fig. 4 b & c). CNG-fuelled vehicles also contribute significantly to the PM₁₀ (15±3 %) and PM_{2.5} (11±3 %) burden.”

Line 284: I am not sure whether a correlation R of 0.5 could be considered significant.

We have deleted this.

Line 288: 0.027 and 0.047 are quite small values. What is your error bound on these numbers?

We have added the uncertainties of the slope. Numbers appear to be small because fire counts are very high (several thousands) but the slopes and resulting PM enhancements are very significant:

“Figure S6 shows that the PM_{2.5} and PM₁₀ mass loadings at the receptor site increased by 0.027±0.006 and 0.047±0.01 μg m⁻³ respectively for each additional fire count within the 24-hour fetch region whenever the trajectories are arriving through north-west and south-west region. It is very interesting to note that the incremental increase in PM_{2.5} and PM₁₀ mass loadings for each additional fire count were almost four times higher than the former regions when the trajectory fetch region was south-east with 0.11±0.01 and 0.19±0.02 μg m⁻³, respectively, likely because the complete burns of entire fields (Figure S7) that are prominent in Punjab can be more easily identified as a fire activity with satellite-based detection (Liu et al., 2019, 2020), while the partial burns (Figure S8) that are more prevalent in the eastern IGP and in Haryana have larger omission errors (Liu et al., 2019, 2020).”

Figure 5: The increase in NO_x in petrol 2W panel during morning commute hours is not reflected in 2W or 4W factors. Does this make sense? Also why are the 2-wheeler petrol vehicle factor contributions high throughout the night and drop near the morning commute hours? I would imagine the 2W vehicles on the road to decrease substantially during the night.

Our PMF has a 4W dominated petrol vehicle factor, containing emissions from the immediate vicinity of the receptor (i.e. within Central Delhi), that have a plume age on the timescale of minutes based on the highest correlation of the factor being with NO emissions. The 2W dominated petrol vehicle factor is aged shows a much higher correlation with NO₂ than with NO and emissions appear to occur in the rural hinterland and outskirts. Their transport time appears to be on the scale of hours. The 2-wheeler factor has been identified as such primarily based on the benzene to toluene ratio which differs between 2-wheelers and 4-wheelers and is better preserved longer during photochemical aging when compared to the emission ratios of C8 and C9 aromatics. However, it is clear from the presence of OVOCs in the source profile and the low correlation of the factor time series with NO that most plumes in the 2W factor are aged, and hence are expected to reach the receptor several hours after the peak evening traffic. The morning peak in NO_x coincides primarily with the peak in cooking emissions and is not triggered by either of these factors. We have revised the text of both sections to make this distinction clearer:

“Figure 4 shows petrol 4-wheeler contributed 20 %, 25 %, and 30 % to the VOC mass loading, OFP, and SOAP, respectively. The source fingerprint of this source matched tailpipe emissions of petrol-fuelled 4-wheelers (Hakkim et al., 2021) and is characterized, in descending rank of contribution, by C8-aromatics, toluene, C9-aromatics (C₉H₁₂), benzene, butene + methyl tert-butyl ether (MTBE) fragment, propyne, propene, methanol and C2-substituted xylenes + C4-substituted benzenes (C₁₀H₁₄). Figure 5 shows that emissions peak in the evening between 7 pm and midnight with average VOC mass loadings >70 μg m⁻³ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO (R=0.8), CO (R=0.7), and CO₂ (R=0.7) indicating the receptor site is impacted by fresh combustion emissions from this source and the atmospheric age of most plumes is on the timescale of minutes.”

“Figure 4 shows petrol 2-wheeler contributed 14 %, 12 %, and 20 % to the VOC mass loading, OFP, and SOAP respectively. The source fingerprint of this source matched tailpipe emissions of petrol-based 2-wheelers (Hakkim et al., 2021) and are characterized, in descending rank of contribution, by toluene, acetone + propanal, C-8 aromatic compounds, acetic acid (C₂H₄O₂), propyne (C₃H₄), methanol (CH₃OH), benzene (C₆H₆), the MTBE fragment and C-9 aromatics (C₉H₁₂). A key difference of the petrol 2-wheeler source profile in comparison to the petrol 4-wheeler source profile is the lower benzene to toluene ratio, which is supported by the GC-FID analysis of tailpipe exhaust (Kumar et al., 2020). Figure 5 shows that emissions peak in the evening between 8 pm and 10 pm with average VOC mass loadings >50 μg m⁻³ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO_x (R=0.6), CO (R=0.6) and CO₂ (R=0.7), but have a lower correlation with NO (R=0.5) (Table S5), and a larger contribution of oxygenated compounds to the source profile, indicating that the emissions have been photochemically aged. This suggests that contrary to 4-wheeler plumes which originate from the immediate vicinity of the site in central Delhi (Figure S1), 2-wheeler plumes reach the receptor after prolonged transport from more distant rural and suburban areas on the outskirts of the city. In such areas, people often favour two-wheelers over four-wheelers.”

Line 326: 3.2.2 Title: By waste disposal, do the authors mean waste burning? These can be very different things with different mechanisms of emissions if combustion is not involved in one versus the other.

We appreciate the referee's helpful comments and have changed the names to heating and waste-burning

Line 354: BB emissions are attributed to solid fuel-based cooking and a cow dung-fired traditional stove is discussed. These measurements were made at IMD Lodhi Road, which appears to be a highly urbanized area. How do the authors justify BB-based cooking activities near such location? Is regional transport important for fresh emissions?

In a mega city like Delhi, there is a socio-economic spectrum of society. Among its residents are many who continue to rely on traditional solid fuels for their cooking and heating needs due to financial constraints. It is also important to remember that the air shed is much larger than just Central Delhi and this is not a PMF factor containing fresh emissions as can be seen by the low $R=0.1$ with NO. Solid fuel usage is very much prevalent in the villages of Haryana and Uttar Pradesh which are located within a radius of less than 60 km from the receptor. We have modified the text as follows:

The activity peaks from 8 am to noon time, with a secondary peak in the early evening hours and persists throughout monsoon and post-monsoon season. Emissions reaching the receptor site show no correlation with NO ($R=0.1$) indicating plumes are not fresh.

Furthermore, cooking's contribution to PM₁₀ is discussed, which is understandably low. However, what about PM_{2.5} that can be formed from the oxidation of gas-phase cooking emissions?

Yes cooking appears to be more of a VOC than a PM source. However, for this factor the percentage contribution to the SOA formation potential is lower than its percentage contribution to the VOC burden because most VOCs emitted are small OVOCs with limited SOA formation potential. The volatility oxidation space plot also shows very little evidence of first- and second- generation oxidation products progressing from the VOC into the IVOC region. The text has been modified to include this information:

The cooking factor is a daytime factor and explains 10 % of the total VOC mass loading, 10% and 8 % to the ozone and SOA formation potential (Fig. 4) but only a negligible share of the total PM₁₀ (≤ 4 %) burden. The volatility oxidation space plot (Figure S9) also shows very little evidence of IVOC oxidation products that could partition into the aerosol phase.

Minor points:

Line 86: "at" Lodhi Road. **done**

Line 190: extra "T" at the start. **done**

Line 264: "Direct", do you mean "Primary"? **yes, however these details were deleted in response to a comment of reviewer 1**

Line 642: "at this time of the year..." Which time of the year? This is written casually. **Revised to:**

While several recent efforts in some sectors (e.g. residential biofuel and cooking) appear to have yielded emission reduction benefits, the narrative to blame the post-monsoon pollution exclusively on the more visible sources (e.g. paddy residue burning), needs to be corrected so other sources are also mitigated.

Figure 3: Remove the word "PMF" from all figure legends. **We prefer to retain the legend in all panels. The figure is easier to comprehend when legends of all individual panels are complete.**

Figure 5: Add y-axis labels to the wind rose plots. **These are conditional probability roses showing a probability between 0 and 1. We now explain this more clearly in the figure legend**

The polar plots (right column) depict the conditional probability of a factor having a mass contribution above the 75th percentile of the dataset during a certain hour of the day between midnight (centre of rose) and 23:00 local time (outside of rose) from a certain wind direction. This probability is determined by dividing the number of observations above the 75th percentile by the total number of measurements in each bin.

Track changes version of the manuscript

Biomass burning sources control ambient particulate matter but traffic and industrial sources control volatile organic compound emissions and secondary pollutant formation during extreme pollution events in Delhi

Arpit Awasthi¹, Baerbel Sinha¹, Haseeb Hakkim¹, Sachin Mishra¹, Varkrishna Mummidivarapu¹, Gurmanjot Singh¹, Sachin D. Ghude², Vijay Kumar Soni³, Narendra Nigam³, Vinayak Sinha¹, Madhavan N. Rajeevan⁴

¹Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research Mohali, Sector 81,

S.A.S Nagar, Manauli PO, Punjab, 140306, India

²Indian Institute of Tropical Meteorology, Pashan, Pune 411 008, Ministry of Earth Sciences, India

³India Meteorological Department, New Delhi 110 003, Ministry of Earth Sciences, India

⁴Ministry of Earth Sciences, Government of India, New Delhi 110 003, India

Correspondence to: Baerbel Sinha (bsinha@iisermohali.ac.in)

Abstract. Volatile organic compounds (VOCs) and particulate matter (PM) are major constituents of smog. Delhi experiences severe smog during post-monsoon season, but a quantitative understanding of VOCs and PM sources is still lacking. Here, we source-apportioned VOCs and PM, using a high-quality recent (2022) dataset of 111 VOCs, PM_{2.5}, and PM₁₀ in a positive matrix factorization (PMF) model. Contrasts between clean-monsoon and polluted-post-monsoon air, VOC source fingerprints, and molecular-tracers, enabled differentiating paddy-residue burning from other biomass-burning sources, which has hitherto been impossible. Fresh paddy-residue burning and residential heating & waste-burning contributed the highest to observed PM₁₀ (25% & 23%), PM_{2.5} (23% & 24%), followed by heavy-duty CNG-vehicles 15% PM₁₀ and 11% PM_{2.5}. For ambient VOCs, ozone, and SOA formation potentials, top sources were petrol-4-wheelers (20%, 25%, 30%), petrol-2-wheelers (14%, 12%, 20%), industrial emissions (12%, 14%, 15%), solid fuel-based cooking (10%, 10%, 8%) and road construction (8%, 6%, 9%). Emission inventories tended to overestimate residential-biofuel emissions at least by a factor of 2, relative to the PMF output. The major source of PM pollution was regional biomass burning, whereas traffic and industries governed VOC emissions and secondary pollutant formation. Our novel source-apportionment method quantitatively resolved even similar biomass and fossil-fuel sources, offering insights into both VOC and PM sources affecting extreme-pollution events. It represents a notable advancement over current source apportionment approaches, and would be of great relevance for future studies in other polluted cities/regions of the world with complex source mixtures.

1 Introduction

The Delhi National Capital Region (NCR) is located in the Indo-Gangetic plains and experiences some of the highest air pollution events worldwide, exposing its inhabitants to hazardous air quality. New Delhi had the world's highest population-weighted annual average PM_{2.5} exposures of 217.6 $\mu\text{g m}^{-3}$ and the sixth-highest PM_{2.5}-attributable death (85 deaths per lakh) (Pandey et al., 2021). India is currently among the world's foremost developing countries and Delhi being its capital has witnessed rapid population growth and urbanization in the

past decade, but a significant fraction of the population still lacks access to cleaner technologies for cooking and heating (Thakur M. 2023; Fadly et al., 2023). Delhi with a population of 31.7 million people (UN World Population Prospects 2022), sees an addition of over six hundred thousand vehicles per year (2022 VAHAN-Ministry of Road Transport and Highways (MoRTH), Government of India). The sources of air pollutants over the region have received much attention recently and a number of source apportionment methods have been applied. Several studies have relied on chemical mass balance models (CMB) that are unable to sniff out unknown fugitive sources since their application rests on prior knowledge of all relevant sources and their source profiles (Prakash et al., 2021; Srivastava et al., 2008). Clearly, in a dynamic developing world megacity like Delhi, where wide disparities exist in terms of access to clean energy and waste burning, and many other activities continue to be carried out by the informal sector, the CMB approach may misattribute emissions only to known sources, with no possibility of identifying other major sources that may be active. While much information has come to light through previous aerosol mass spectrometry-based source apportionment studies, a key limitation of the previous studies has been an inability to distinguish between different similar types of fossil fuel and biomass-burning sources (Kumar et al., 2022; Mishra et al., 2023). The VOCs source-fingerprints of many combustion sources are well constraint and understood, and have recently been used in PMF-based studies to source apportion co-emitted greenhouse gasses such as methane, CO₂ and N₂O (Guha, et al. 2015; Assan et al. 2018; Schulze et al. 2023). We now extend the use of this promising new technique towards source-apportionment of co-emitted PM_{2.5} and PM₁₀. This helps us overcome another major limitation of existing studies which has been the piece-meal approach where either VOCs (Jain et al., 2022) or PM or a subset thereof have been investigated, that too only on datasets that were acquired in 2019 or earlier, i.e. pre-COVID19 period after which significant changes have been implemented. For example, the Bharat Stage VI which complies with the Euro VI norms was implemented in 2018 in Delhi and 2019 for Delhi NCR (Gajbhiye et al., 2023). This significant decision was prompted by the severe air pollution challenges faced by Delhi, particularly worsening around 2019 (Gajbhiye et al., 2023). Still air pollution continues to pose major health risks. Overall, a continued lack of strategic knowledge and inability to pinpoint the exact sources and their contribution, hampers efforts to propose evidence-based strategies for mitigation of major sources. In our previous studies from another site in the Indo-Gangetic Plain (Pallavi et al., 2019; Singh et al., 2023), we demonstrated that source apportionment carried out by PMF when combined with measured VOC chemical fingerprints of sources, can distinguish and quantify the contribution of even similar types of sources (e.g. within traffic source: to distinguish 4-wheelers from 2-wheelers and diesel vehicles; and within biomass burning sources to distinguish paddy stubble burning from residential biofuel combustion). We improve upon those studies that were carried out on datasets acquired using a unity mass resolution VOC proton transfer reaction mass spectrometer by recent new data acquired using the latest state-of-the-art enhanced volatile range high mass resolution and high sensitivity PTR-TOF-10 K technology over Delhi (Mishra et al., 2024). The dataset used for source apportionment in this study using the positive matrix factorization modelling includes the high sensitivity (few ppt), high mass resolution (>10000) real-time acquisition of 111 speciated volatile organic compounds measured (15th August 2022–26th November 2022) using a Proton Transfer Reaction Time of Flight Mass Spectrometer 10 K (PTR-TOF10K-MS) instrument in Delhi, along with hourly averaged PM_{2.5} and PM₁₀ measurements. This dataset is novel in that it contains all major known gas phase molecular tracers for varied sources and VOC profiles of major agricultural and urban sources extant over Indo-Gangetic Plain. The dataset covered the relatively cleaner monsoon season which provides a baseline air pollution over the city and

the post-monsoon season when post-harvest agricultural paddy residue burning in the Indo-Gangetic Plain perturbs the atmospheric chemical composition by providing an additional source of VOC and PM emissions. This comprehensive approach ensured that the positive matrix factorization model, which provides the advantage of determining air pollution sources without any prior knowledge of the source fingerprints, was able to quantify the contribution of different sources to the ambient VOC, PM_{2.5}, and PM₁₀ mass concentrations reliably as its solutions are sensitive to contrasts in ambient time series data. The statistical solution obtained using the model were verified against real-world measured source profiles from the region and thus presents a significant advancement over previous PMF source apportionment studies reported from the Delhi-NCR region. Furthermore, by combining this molecular tracer-based methodology and analyses with additional air mass back trajectory and statistical analyses, we also constrain the location of the major pollution sources and regions and compare the results of our source apportionment study with two widely used gridded emission inventories in chemical transport models, namely the Emission Database for Global Atmospheric Research (EDGARv6.1) (Crippa et al., 2022), and the Regional Emission inventory in Asia (REAS v3.2.1 (Kurokawa & Ohara, 2020)).

2 Methodology

2.1 Measurement site and meteorological conditions

The new PTR-TOF-MS 10 K enhanced volatility range mass spectrometer, as well as the primary VOC dataset and site, have already been described and analyzed in detail in the companion paper (Mishra et al., 2024). Hence only a brief description of these aspects and complementary aspects such as the air mass flow trajectories at the site during the study period from August 2022 to November 2022 are provided below.

Ambient air was sampled into the instruments from the roof-top of a tall building (28.5896°N-77.2210°E) at ~35 m above ground, located within the premises of the Indian Meteorological Department (IMD) at Lodhi Road, New Delhi situated in Central Delhi. The sampling site is a typical urban area surrounded by green spaces, government offices, and residential areas, but not in the direct vicinity of any major industries (Fig. S1) and representative of the airflow patterns observed in Delhi seasonally. Figure 1 shows the location of the site and the 120 h back trajectories of air masses arriving at the site that were grouped according to the dominant synoptic regional scale transport into a) south-westerly (orange and yellow) flows carrying emissions from southern Punjab, Haryana, Uttar Pradesh, Madhya Pradesh, Rajasthan and Gujarat towards the receptor, b) north-westerly (light and dark blue) flows carrying emissions from Pakistan Punjab, Indian Punjab, Haryana, Western Uttar Pradesh, Himachal Pradesh, and Uttarakhand towards the receptor, and c) south-easterly flows (light and dark red) carrying emissions from Haryana, Southern Uttarakhand, Uttar Pradesh, Bihar and Nepal towards the receptor. Figure 1d shows a Google Earth image with a spatial map of the daily fire counts in the region for the post-monsoon season alongside with the maximum 24-h fetch region for each of these synoptic flow situations marked by coloured square. Figure 1e-h shows the e) photosynthetic active radiation, f) daily fire counts in the fetch region (21-32°N, 72-88°E), g) temperature and relative humidity, and h) the ventilation coefficient and the sum of the daily rainfall during the study period (15th August 2022– 26th November 2022). Wind speed, wind direction, ambient temperature, relative humidity, and photosynthetic active radiation were measured using meteorological sensors (Campbell Scientific portable sensors equipped with CS215 RH and temperature sensor, PQS1 PAR sensor, TE525-L40 v rain gauge, Campbell Scientific Inc.). Boundary layer height was taken from the

ERA5 dataset (Hersbach et al., 2023) and the ventilation coefficient was calculated as the product of the measured wind speed and boundary layer height. Fire counts were obtained using the Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m thermal anomalies / active fire product data from the VIIRS sensor aboard the joint NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites, for high and normal confidence intervals only. The back trajectories in Fig. 1 showing the 5-day runs were obtained using Hysplit Desktop, version 5.2.1 (Stein et al., 2015; Rolph et al., 2017) with GFSv1 0.25° resolution meteorological fields as input data. The model was initialized every 3 hours (0, 3, 6, 9, 12, 15, 18, and 21 UTC) at 50 m above ground level for the year 2022 and trajectories were subjected to back trajectory cluster analysis via k-means clustering (Bow, 1984) with Euclidean distance metrics using the open-air package (v2.11, Carslaw & Ropkins, 2012). Three basic air transport situations occur at this site, namely from the South West (Fig. 1a), North-West (Fig. 1b), and South-East (Fig. 1c). These regional transport situations in the shared air-shed have been described for another receptor site located 300 km north of Delhi previously in great detail (Pawar et al, 2015). At Delhi, each of these large-scale flow patterns can occur with three different transport speeds; fast (darkest colour), medium (intermediate colour) and slow (lighter colour), resulting in 9 clusters.

During the monsoon season (15.08-30.09.2022), the air masses from the south-west direction (western arm of the monsoon) were more prevalent than air masses reaching the site from the south-east (Bay of Bengal arm of the monsoon). During the post-monsoon season (01.10-26.11.2022) air masses remain confined over the NW-IGP for prolonged periods and primarily reach the site from the north-west (Fig. 1b), except during the passage of western disturbances (05.10-10.10.2022 and 04.11-10.11.2022), which result in brief periods with south westerly and south-easterly flow and rain (Fig. 1h). Figure 1f shows that paddy residue burning of short-duration varieties commences even before the monsoon withdrawal on 29th September 2022, however, the burning peaks during the harvest of late varieties in late October and early November. During this period a drop in temperature (Fig. 1g) and increased fire activity (Fig. 1f) results in the build-up of a persistent haze layer leading to suppressed photosynthetically active radiation (Fig. 1e). This is associated with prolonged periods of poor ventilation (Fig. 1h).

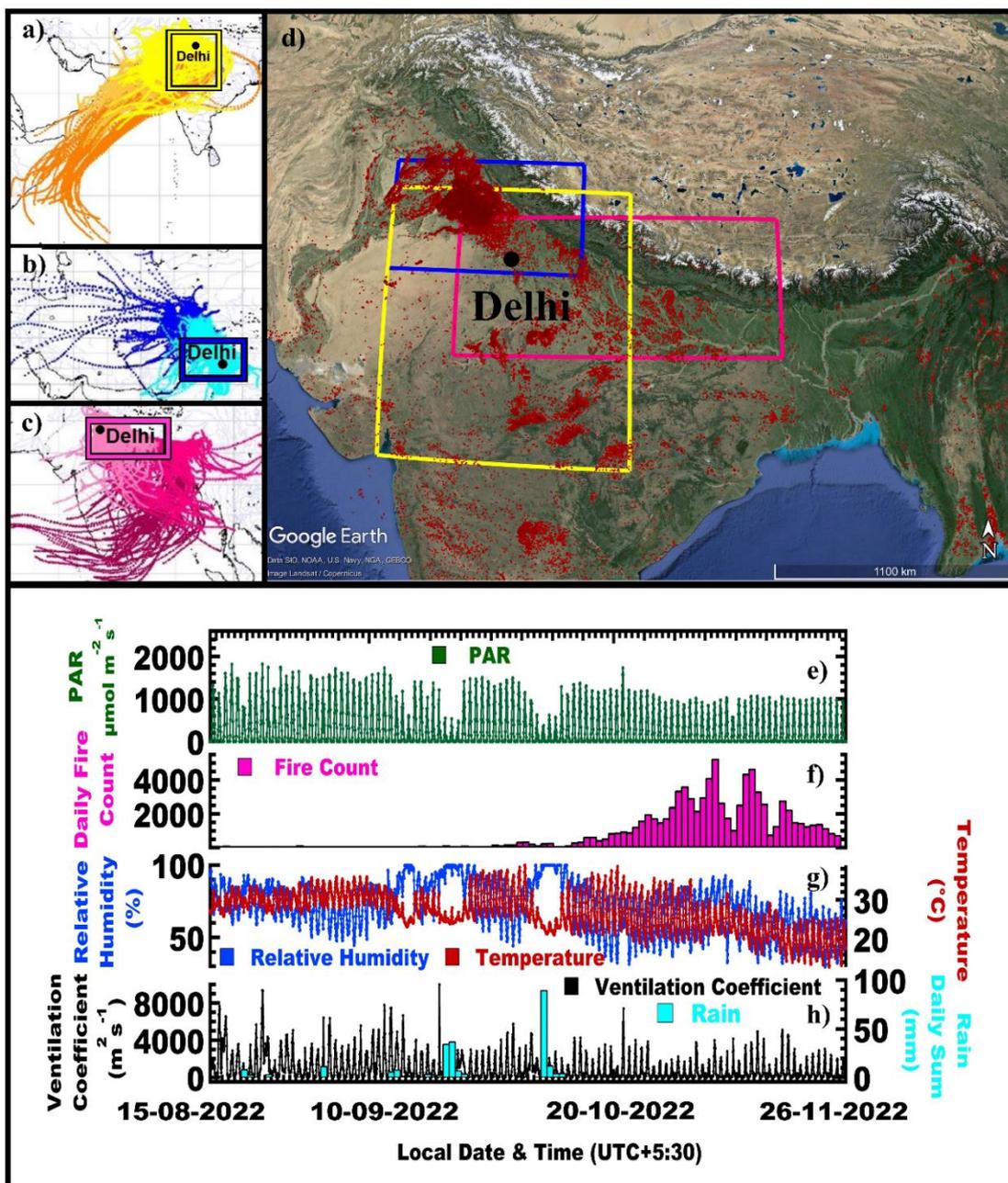


Figure 6: 120 h back trajectory air mass reaching receptor site at Mausam Bhawan building (28.5896°N - 77.2210°E , 50 m above ground level) grouped according to the dominant synoptic scale transport into a) South-Westerly, b) North-Westerly, and c) South-Easterly flow. d) spatial map of the daily fire counts in the region for the post-monsoon season with square boxes indicating the fetch region from which air masses typically reach the receptor site within 24 hrs for a given flow situation. The bottom panels show the e) photosynthetically active radiation (PAR), f) daily fire counts in the fetch region, g) temperature and relative humidity, and h) the ventilation coefficient and the sum of the daily rainfall for the study period. .

2.2 Measurement of Volatile Organic Compounds, trace gases, and $\text{PM}_{2.5}$ and PM_{10} mass concentrations

Measurements of volatile organic compounds were performed using a high mass resolution and high sensitivity proton transfer reaction time of flight mass spectrometer (PTR-TOF10k; model PT10-004 manufactured by Ionicon Analytik GmbH, Austria). Details pertaining to the characterization, calibration, and QA/QC of the acquired dataset have been provided in Mishra et al., 2024. It is worth mentioning again that as a significant improvement over other previous PTR-TOF-MS deployments in Delhi, the inlet system of the instrument used in

this work was designed for sampling and detection of low-volatility compounds with the extended volatility range technology (Piel et al., 2021). The inlet system of the instrument as well as the ionization chamber is fully built into a heated chamber and the inlet capillary is further fed through a heated hose to ensure there are no “cold” spots for condensation. The entire inlet system is made of inert material (e.g. PEEK or siliconert treated steel capillaries) to keep surface effects minimal. Further, the overall inlet residence time was less than 3 seconds, throughout the campaign. Compared to previous PTR-TOF-MS instruments deployed in Delhi, this instrument also had unprecedented higher mass resolution (greater than 10000 $m/\Delta m$ (FWHM) for $m/z \geq 79$ Th even reaching as high as 15000 at m/z 330) coupled with high detection sensitivity (~ 1 ppt or better for 60 s averaged data), providing unprecedented ability for identification and quantification of new ambient compounds. Mass spectra were acquired over the m/z 15 to 450 amu range at a frequency of 1 Hz. Table S1 lists information pertaining to m/z , compound names, and sources supported by references to previous studies where available, averaged ambient mass concentrations and classification of the species as weak or strong for the PMF model runs. The accuracy error was minimized by conducting a total of 8 span calibrations throughout the study period. The details of these calibrations can be found in Mishra et al., 2024. The precision error for each m/z listed in table S1, which needs to be included into the PMF model runs, was calculated from the average observed count rate in counts per second (cps) of each m/z with the help of Poisson statistics. The detection limit was determined as 2σ of the noise observed in clean zero air.

ThermoFisher Scientific 48i (IR filter correlation-based spectroscopy), 43i (pulsed UV fluorescence), 49i (UV absorption photometry), and 42i trace level air quality analyzers (chemiluminescence) were used to quantify carbon monoxide (CO), ozone (O₃), and NO and NO₂, respectively. The overall uncertainty of the measurements was less than 6 %. Measurements of PM_{2.5} and PM₁₀ were made using ThermoFisher Scientific Model 5014i series which is based on the beta-attenuation technique. Technical details pertaining to QA/QC of these instruments have been comprehensively described in our previous works (Chandra and Sinha, 2016; Kumar et al., 2016; Sinha et al., 2014). Carbon dioxide and methane were measured using a cavity ring down spectrometer (Model G2508, Picarro, Santa Clara, USA). The overall uncertainty of these measurements was below 4 % and technical details pertaining to the instrument are available in Chandra et al., 2018.

2.3 Positive matrix factorization (PMF) model analysis

The US EPA PMF 5.0 (Paatero et al., 2002, 2014; Paatero & Hopke, 2009; Noris et al., 2014) was applied to a sample matrix of 2496 hourly observations and 111 VOC species. The species with S/N greater than 2.0 were designated as strong species (94) while others were designated as weak species (17). The total VOC mass was included as a weak species and was calculated as the sum of the mass of the individual 111 VOC species included in the PMF. Overall, the 111 VOC species included in our analysis and their isotopic peaks explained 86% of the VOC mass detected during our study period. The remaining 119 m/z that accounted for 14% of the detected VOC mass could not be included in our PMF analysis mostly because signals were below the detection limit for close to 50% of the observation period, or because compound identity could not be confirmed via isotopic peaks. PM_{2.5} and PM₁₀ were included as additional weak species in the model. The specified uncertainty for weak species is tripled by the PMF model, to limit the influence of such species on the PMF solution. Several authors have recently pioneered the use of VOC tracers in a PMF to source apportion co-emitted greenhouse gasses such as methane, CO₂ and N₂O (Guha, et al. 2015; Assan et al. 2018; Schulze et al. 2023). Since the VOCs source-fingerprints of

many combustion sources are well constrained and understood, we now extend the use of this promising new technique towards source-apportionment of co-emitted PM_{2.5} and PM₁₀. The PMF is a matrix decomposition factor analysis model that decomposes a time series of measured species into a set of factors with fixed source fingerprints whose contributions to the input data set varies with time. This makes the model well suited to accommodate all chemical species co-emitted from the same source.

The EPA PMF 5.0 is a multivariate factor analysis tool and a receptor model that divides the data matrix X_{ij} (time series of measured concentrations of VOCs with i distinct observations and j measured species) into two matrices, F_{kj} (source fingerprint) and G_{ik} (source contribution), along with a residual matrix, E_{ij}, using the simultaneous application of the linear least square method in multiple dimensions.

$$\mathbf{X}_{ij} = \sum_{k=1}^p \mathbf{G}_{ik} \times \mathbf{F}_{kj} + \mathbf{E}_{ij} \quad (1)$$

The user must provide the number of variables or sources (k). To determine the number of VOC sources the model can resolve in this atmospheric environment, the model was run with 3 to 12 factors. The model was initiated for 20 base runs with the recommended block size of 379, and the run with the lowest Q_{robust} and Q_{true} was chosen for further analysis and display in Figure 2. Figure 2 shows how the percentage of total VOC, PM_{2.5}, and PM₁₀, attributable to various sources changes when the number of factors increases from 3 to 12, while Fig. S2-S4 illustrates the evolution in the factor contribution time series, source profile, and percentage of species explained by different sources when the number of factors in the PMF increases. Figure S5 shows how the Q_{true}/Q_{theoretical} ratio and Q_{robust}/Q_{theoretical}, and scaled residuals beyond 3 standard deviations drop exponentially when the number of factors increases. It can be seen that initially the Q_{true}/Q_{theoretical} ratio drops faster than Q_{robust}/Q_{theoretical} ratio on account of additional major plumes being better explained with each additional factor. However, with the increase from 11 to 12 factors both drop in a parallel fashion indicating that the point of diminishing returns has been reached.

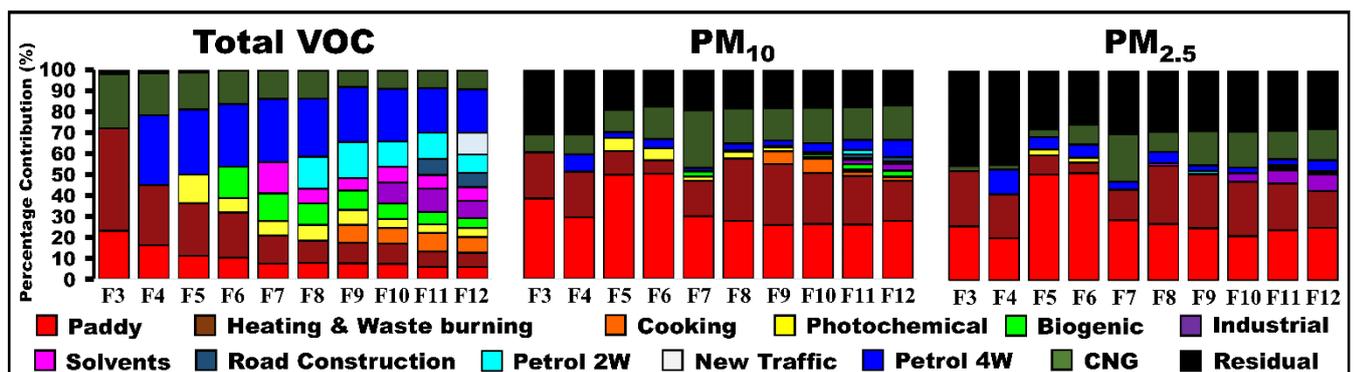


Figure 7: Percentage of the total VOC, PM₁₀ and PM_{2.5} mass explained by each factor in the PMF model output results when the number of PMF factors in the model is increased from 3 to 12. The balance to 100 % shown in black indicates the percentage share of the total mass in the PMF residuals.

While the three major traffic factors namely; CNG, petrol 4-wheeler, and petrol 2-wheeler are completely resolved with the 8 factors solution, three major biomass-burning related sources namely paddy residue burning, heating, and waste burning, and solid fuel-based cooking are separated with a 9-factors solution. Until the PMF opens distinct factors for the industrial OVOC emissions in the 7-factor solution, the partitioning between paddy residue burning and heating and waste burning PM_{2.5} and PM₁₀ emissions in the model remains unstable, because these sources with their strong OVOC emissions are most agreeable to accommodating additional OVOC sources in

their fingerprint at the expense of explaining the PM_{2.5} and PM₁₀ emissions. Once the industrial OVOC emissions have their own factor, this split becomes stable. The amount of PM attributed to residential heating and waste burning stabilizes after a separate factor for cooking emissions opens up in the 9-factor solution. Industrial emissions are separated from solvent usage and other evaporative emissions with a 10-factor solution, and road construction activity emerges as a separate source with an 11-factor solution. While attempting to resolve 12-factors, the model splits transport sector emissions into four separate factors. However, this new transport sector factor shows a time series correlation (R=0.8) with the petrol 4-wheeler factor, and the 12-factor solution was found to be rotationally unstable during bootstrap runs, indicating that the model cannot resolve more than 11 factors with the available VOC tracers. The 12-factor solution also hardly improves the $Q_{\text{robust}}/Q_{\text{theoretical}}$ and $Q_{\text{true}}/Q_{\text{theoretical}}$ ratio (Fig. S5). Therefore, the 11-factor solution was analyzed further. The model was run in the constrained mode elaborately described in Sarkar et al., (2017) and Singh et al., (2023). The rotational ambiguity can be reduced using this option with the aid of prior knowledge by encouraging the model to minimize (pull down) or maximize (pull up) the total mass assigned to specific hourly observations or compounds in source profiles as much as possible within a pre-defined permissible penalty on Q. The primary problem of the base run solutions is that night-time biomass burning plumes contaminate both the biogenic and the photochemical factor. To minimize this in our constrained run, we have pulled down primary emissions (acetonitrile, toluene, C8 aromatics, and C9 aromatics) in the biogenic and photochemical factors. We also pulled down the top-7 strongest nighttime plumes contaminating the biogenic and photochemical factors. In addition, we pulled up the highest plume event for all the anthropogenic emission-related factors as detailed in Table S2. The overall penalty to Q (the object function) was 4.9 %, which is within the recommended limit of 5 % (Norris et al., 2014; Rizzo & Scheff, 2007). The model uncertainty was assessed using bootstrap runs. The constrained model was found to be rotationally stable and robust with 100 % of all bootstrap runs for each individual factor mapped onto the base factor with R>0.6 and no unmapped bootstraps.

2.4 Calculation of the ozone formation potential, secondary organic aerosol formation and volatility

The contribution of VOCs to ozone production was derived with the maximum incremental reactivity (MIR) (Carter, 2010) method using the following equation

$$OFP = \sum(c_iMIR_i) \quad (2)$$

where c_i is the measured concentration of VOC species i and MIR_i is the maximum incremental reactivity of VOC species i .

The secondary organic aerosol production (SOAP) was determined using the following equation

$$SOAP = \sum(c_iSOAP_i) \quad (3)$$

SOAP_{*i*} values were calculated with the SOA yields for high NO_x emission environments reported in Table S3 according to the equation of Derwent et al., (1998; 2010), as Delhi being a megacity is a high NO_x emission environment. This equation evaluates each VOC species' ability to make SOA in relation to the amount of SOA the same mass of toluene would make when introduced to the ambient environment. This is represented by the SOAP_{*i*}.

The saturation vapour pressure of VOCs was calculated using EPA EPI Suite v4.1 (MPBPWINv.1.43; KOAWIN v.1.00) provided by the US Environmental Protection Agency (US EPA, 2015) according to the method described in Li et al., (2016). The vapour pressure of liquids and gases is estimated using the average of the Antoine method (Lyman et al., 1990) and the modified Grain method (Lyman 1985). The vapour pressure is then converted to saturation mass concentration C_0 in $\mu\text{g m}^{-3}$ using the following equation:

$$C_0 = \frac{M 10^6 p_0}{760 R T} \quad (4)$$

wherein M is the molar mass [g mol^{-1}], R is the ideal gas constant [$8.205 \times 10^{-5} \text{ atm K}^{-1} \text{ mol}^{-1} \text{ m}^3$], p_0 is the saturation vapor pressure [mm Hg], and T is the temperature (K). Organic compounds with $C_0 > 3 \times 10^6 \mu\text{g m}^{-3}$ are classified as VOCs while compounds with $300 < C_0 < 3 \times 10^6 \mu\text{g m}^{-3}$ as Intermediate VOCs (IVOCs).

2.5 Comparison of existing emission inventories with PMF derived output

The observational data was grouped according to the predominant airflow into a south-westerly, north-westerly, and south-easterly group, and the fetch region from which air masses would reach the receptor site within 24 h was determined for each group separately spanning latitude 21–31 N and longitude 72–82 E, latitude 28–32 N and longitude 72–80 E and latitude 25–30 N and longitude 75–88 E, respectively, for the three flow regimes. Two gridded emission inventories namely the Emission Database for Global Atmospheric Research (EDGARv6.1) for the year 2018 (Crippa et al., 2022), and the Regional Emission inventory in Asia (REAS v3.2.1) for the year 2015 (Kurokawa & Ohara, 2020) were filtered for these three fetch regions to compare PMF results with the emission inventory. We compare the relative percentage contribution of sources to the total atmospheric pollution burden in the PMF with the relative percentage contribution of sources to the total emissions for the emission inventories. This approach has been routinely used to evaluate emission inventories with the help of PMF results at different sites around the world (Buzcu-Guven and Fraser, 2008; Morino et al. 2011; Sarkar et al., 2017; Li et al., 2019; Qin et al., 2022). For the purpose of emission inventory comparison of anthropogenic sources, natural sources such as biogenic emissions and the photochemistry factor were removed from the PMF output, while the solid fuel-based cooking and residential heating and waste burning emissions were summed up in residential & waste management. In addition, CNG and Petrol 2 & 4-wheeler factors were combined into the consolidated transport sector emissions.

3 Results and Discussions

3.1 Validation of the PMF output with source fingerprints

Figure 3 shows the source profile of the eleven factors that our PMF analyses resolved. Out of the 111 VOCs only those whose normalized source contribution exceeded 0.1 when divided by the most abundant compound in the same source profile in at least one of the sources, were included in the figure. The source identity of the PMF factors was confirmed by matching the PMF factor profiles with the unit $\mu\text{g m}^{-3}$ with normalized source fingerprints of grab samples collected from the potential sources. To facilitate the comparison of emission factors and grab samples from different studies with the PMF output, the source samples were normalized by dividing each species' mass/emission factor by the mass/emission factor of the most abundant species in a given fingerprint.

The PMF factor profile matched best against source samples collected from burning paddy fields ($R=0.6$, Kumar et al., 2020) for the paddy residue burning factor. The cooking factor matched emissions from a cow-dung-fired traditional stove called angithi ($R=0.7$, Fleming et al., 2018). The residential heating & waste burning factor had a source fingerprint matching emission from leaf litter burning, ($R=0.7$, Chaudhary et al., 2022), waste burning ($R=0.7$, Sharma et al., 2022), and cooking on a chulha fired with a mixture of firewood and cow dung ($R=0.9$, Fleming et al., 2018). The factors identified as CNG ($R=1.0$), petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.6$) matched tailpipe emissions of the respective vehicle types and fuels (Hakkim et al., 2021). The petrol 4-wheelers ($R=0.9$), and petrol 2-wheelers ($R=0.7$) also matched traffic junction grab samples from Delhi (Chandra et al., 2018). The OVOC source fingerprint of the road construction factor matched the source fingerprint of asphalt mixture plants and asphalt paving ($R=0.9$, Li et al., 2020), while the hydrocarbon source fingerprint matched diesel-fuelled road construction vehicles ($R=0.6$, Che et al., 2023). The factors identified as solvent usage and evaporative emissions matched ambient air grab samples collected from an industrial area at Jahangirpuri ($R=0.7$), and Dhobighat at Akshar Dham ($R=0.5$) in this study. The factor identified as industrial emissions showed the greatest similarity to ambient air grab samples from the vicinity of the Okhla waste-to-energy plant ($R=0.8$), Gurugram ($R=0.7$) and Faridabad ($R=0.8$) industrial area. The biogenic factor showed the greatest similarity to leaf wounding compounds released from *Populus tremula* ($R=0.8$, Portillo-Estrada et al., 2015) as well as BVOC fluxes from *Mangifera indica* ($R=0.4$, Datta et al., 2021).

emissions. SOA formation potential was dominated by the transport sector (54 %) while direct PM₁₀ (52±8%) and PM_{2.5} (48±12%) emissions were dominated by different biomass burning sources (Fig. 4 b & c). CNG-fuelled vehicles also contribute significantly to the PM₁₀ (15±3 %) and PM_{2.5} (11±3 %) burden. A significant share of the PM₁₀ (18 %) and PM_{2.5} (28 %) burden is associated with the residual and not directly linked to combustion tracers. This share can likely be attributed to windblown dust arriving at the site through long-range transport (Pawar et al., 2015) and to secondary organic, and secondary inorganic aerosols such as ammonium sulphate and ammonium nitrate. Due to the complex relationship of secondary aerosol with gas-phase precursors and emission tracers, VOC tracers are not a suitable tool to source-apportion this aerosol component. Meteorological conditions, homogeneous, heterogeneous, and multiphase chemistry control how fast primary emissions are converted to secondary aerosol. To explain the source of those species, one also needs to invoke the physicochemical and thermodynamical properties of the aerosol. (Acharja et al., 2022).

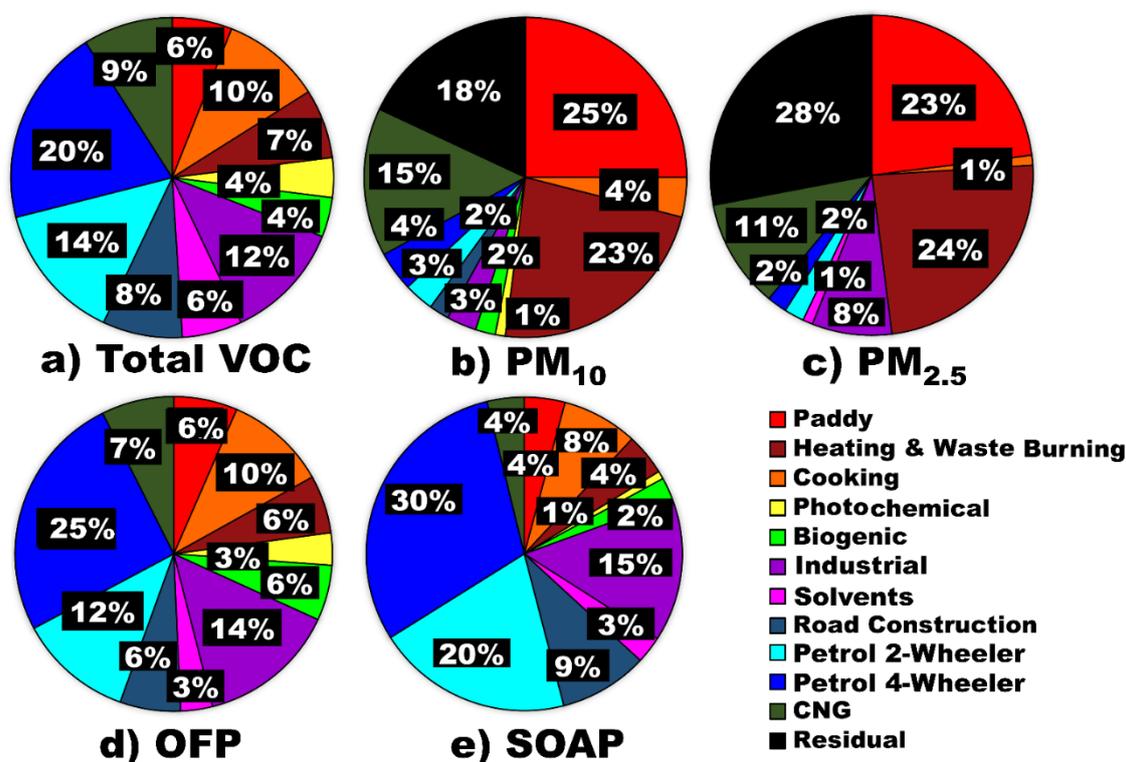


Figure 9: Source contribution of the 11 sources to the (a) total ambient VOC mass loading, (b) PM₁₀ mass loading and (c) PM_{2.5} mass loading (d) ozone formation potential and (e) SOA formation potential. .

3.2 Detailed discussion of individual emission sources

3.2.1 Factor 1: Paddy residue burning

Paddy residue burning was one of the largest contributors to the total observed PM₁₀ (25 %) and PM_{2.5} (23 %) (Fig.4b,4c) mass concentrations in Delhi. An earlier WRF-Chem-based study with the FINNv1.5 inventory had attributed 20 % of the PM_{2.5} burden to this source for the year 2018 (Kulkarni et al., 2020). Its importance as a PM source stands in stark contrast to its minor contribution to the overall VOC mass loading in Delhi (6 %). In Mohali, Punjab, this source was also found to only contribute 6 % to the VOC burden in October and November (Singh et al., 2023). In descending rank of mass contribution, acetaldehyde (CH₃CHO), acetic acid (C₂H₄O₂),

acetone + propanal (C_3H_6O), hydroxyacetone ($C_3H_6O_2$), acrolein (C_3H_4O), diketone ($C_4H_6O_2$) and furfural ($C_5H_4O_2$) contributed most to the total VOC mass of this factor. Figure 5 shows that the 24-h averaged factor contribution time series has the highest cross correlation with same day fire counts ($R=0.8$), while hourly average source contributions correlate most with $PM_{2.5}$ (0.7) and PM_{10} (0.7) (Table S4). The high correlation with same-day fire counts points towards nearby fire activity as the dominant source of paddy burning-related pollution in the Delhi NCR. A recent study from Punjab indicated that the largest PM enhancements at a receptor are caused by fire occurring within 50 km radius around the receptor site (Pawar & Sinha, 2022). Figure S6 shows that the $PM_{2.5}$ and PM_{10} mass loadings at the receptor site increased by 0.027 ± 0.006 and 0.047 ± 0.01 $\mu g\ m^{-3}$ respectively for each additional fire count within the 24-hour fetch region whenever the trajectories are arriving through north-west and south-west region. It is very interesting to note that the incremental increase in $PM_{2.5}$ and PM_{10} mass loadings for each additional fire count were almost four times higher than the former regions when the trajectory fetch region was south-east with 0.11 ± 0.01 and 0.19 ± 0.02 $\mu g\ m^{-3}$, respectively, likely because the complete burns of entire fields (Figure S7) that are prominent in Punjab can be more easily identified as a fire activity with satellite-based detection (Liu et al., 2019; 2020), while the partial burns (Figure S8) that are more prevalent in the eastern IGP and in Haryana have larger omission errors (Liu et al., 2019; 2020). Regional gradients in fire detection efficiency can complicate attempts to model air quality with the help of fire-count-based emission inventories (Kulkarni et al., 2020).

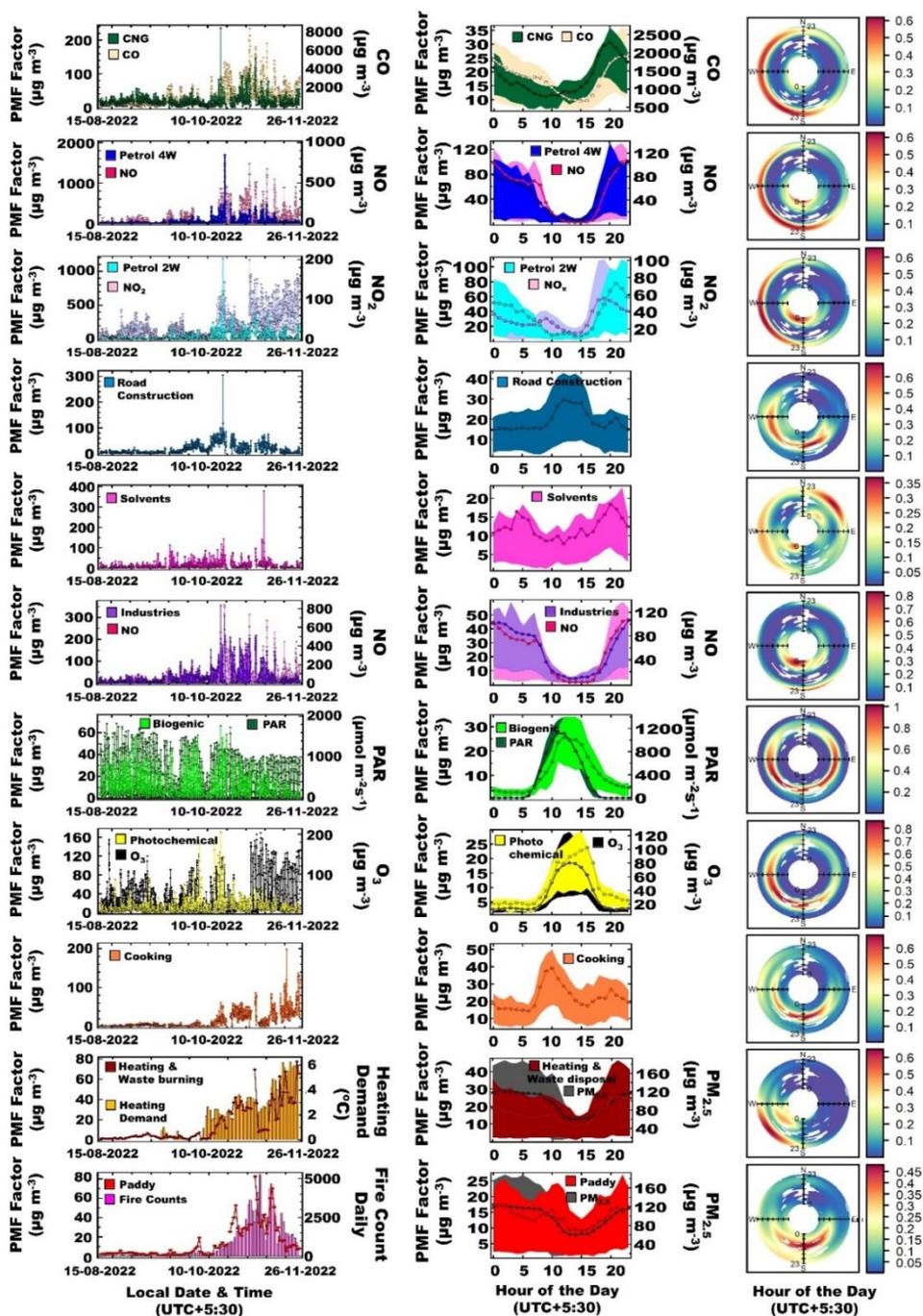


Figure 10: Time series of each factor in $\mu\text{g m}^{-3}$ (left column) with respective normalized diurnal profiles (centre column). The shaded region in the diurnal profiles depicts the area between the 25th and 75th percentile while the median of the dataset is marked as the line. The polar plots (right column) depict the conditional probability of a factor having a mass contribution above the 75th percentile of the dataset during a certain hour of the day between midnight (centre of rose) and 23:00 local time (outside of rose) from a certain wind direction. This probability is determined by dividing the number of observations above the 75th percentile by the total number of measurements in each bin.

Figure 6 shows that this factor explained the largest percentage share of O-heteroarene compounds such as furfural ($\text{C}_5\text{H}_4\text{O}_2$), methyl furfural ($\text{C}_6\text{H}_6\text{O}_2$), hydroxy methyl furfural ($\text{C}_6\text{H}_6\text{O}_3$), furanone ($\text{C}_4\text{H}_4\text{O}_2$), hydroxymethyl furanone ($\text{C}_5\text{H}_6\text{O}_3$), furfuryl alcohol ($\text{C}_5\text{H}_6\text{O}_2$), furan ($\text{C}_4\text{H}_4\text{O}$), methyl furans ($\text{C}_5\text{H}_6\text{O}$), C2-substituted furans ($\text{C}_6\text{H}_8\text{O}$), and C3-substituted furans ($\text{C}_7\text{H}_{10}\text{O}$), which are produced by the pyrolysis of cellulose and hemicellulose, and have previously been detected in biomass burning samples (Coggon et al., 2019; Hatch et al., 2015; 2017; Koss et al., 2018; Stockwell et al., 2015). Figure 6 also shows that this factor explains the largest share of the most

abundant oxidation products that result from the nitrate radical-initiated oxidation of toluene as well as from OH-initiated oxidation of aromatic compounds under high NO_x conditions, namely nitrotoluene ($\text{C}_7\text{H}_7\text{NO}_2$) and nitrocresols ($\text{C}_7\text{H}_7\text{NO}_3$) (Ramasamy et al., 2019), which indicates a certain degree of aging of the plumes. These nitroaromatic compounds are significant contributors to SOA and BrC, (Palm et al., 2020, Harrison et al., 2005). It also explains several other nitrogen containing VOCs such as nitroethane ($\text{C}_2\text{H}_6\text{NO}_2$), the biomass burning tracer acetonitrile (CH_3CN) and pentanenitrile ($\text{C}_5\text{H}_9\text{N}$). The presence of pentanenitrile isomers in biomass burning smoke has previously been confirmed using gas chromatography-based studies (Hatch et al., 2015, Hatch et al., 2017). In addition the factor explains the largest percentage share of acrolein ($\text{C}_3\text{H}_4\text{O}$), hydroxyacetone ($\text{C}_3\text{H}_6\text{O}_2$), cyclopentadienone ($\text{C}_5\text{H}_4\text{O}$), cyclopentanone ($\text{C}_5\text{H}_8\text{O}$), diketone ($\text{C}_4\text{H}_6\text{O}_2$), pentanedione ($\text{C}_5\text{H}_8\text{O}_2$), hydroxybenzaldehyde ($\text{C}_7\text{H}_6\text{O}_2$), guaiacol ($\text{C}_7\text{H}_8\text{O}_2$), and the levoglucosan fragment ($\text{C}_6\text{H}_8\text{O}_4$), many of these compounds are known to form during lignin pyrolysis (Hatch et al., 2015, Koss et al., 2018; Nowakowska et al., 2018), while dimethylbutenedial ($\text{C}_6\text{H}_8\text{O}_2$), trimethylbutenedial ($\text{C}_7\text{H}_{10}\text{O}_2$) are ring opening oxidation products of aromatic compounds (Zaytsev et al., 2019). Figure S9 shows the volatility oxidation state plot for all 111 VOCs in which the marker size represents the percentage share of each compound explained by the paddy residue burning factor and markers are colour coded by the number of carbon atoms. The plot shows evidence of the first- and second-generation oxidation products of C5 and C6 hydrocarbon transitioning from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009), while C7 hydrocarbons progress along trajectories expected for both the addition of -OH and =O functionality. This indicates that paddy residue burning contributes significantly to the SOA burden. However, the fact that the PM_{10} mass associated with this factor ($36.5 \mu\text{g m}^{-3}$) is 1.8 times larger than the $\text{PM}_{2.5}$ mass ($20.7 \mu\text{g m}^{-3}$) and 3 times larger than the VOC mass ($11.6 \mu\text{g m}^{-3}$) released during the same combustion process, points towards the relatively coarse ash formed from the phytolith skeleton of rice straw (Figure S10) as the dominant aerosol source.

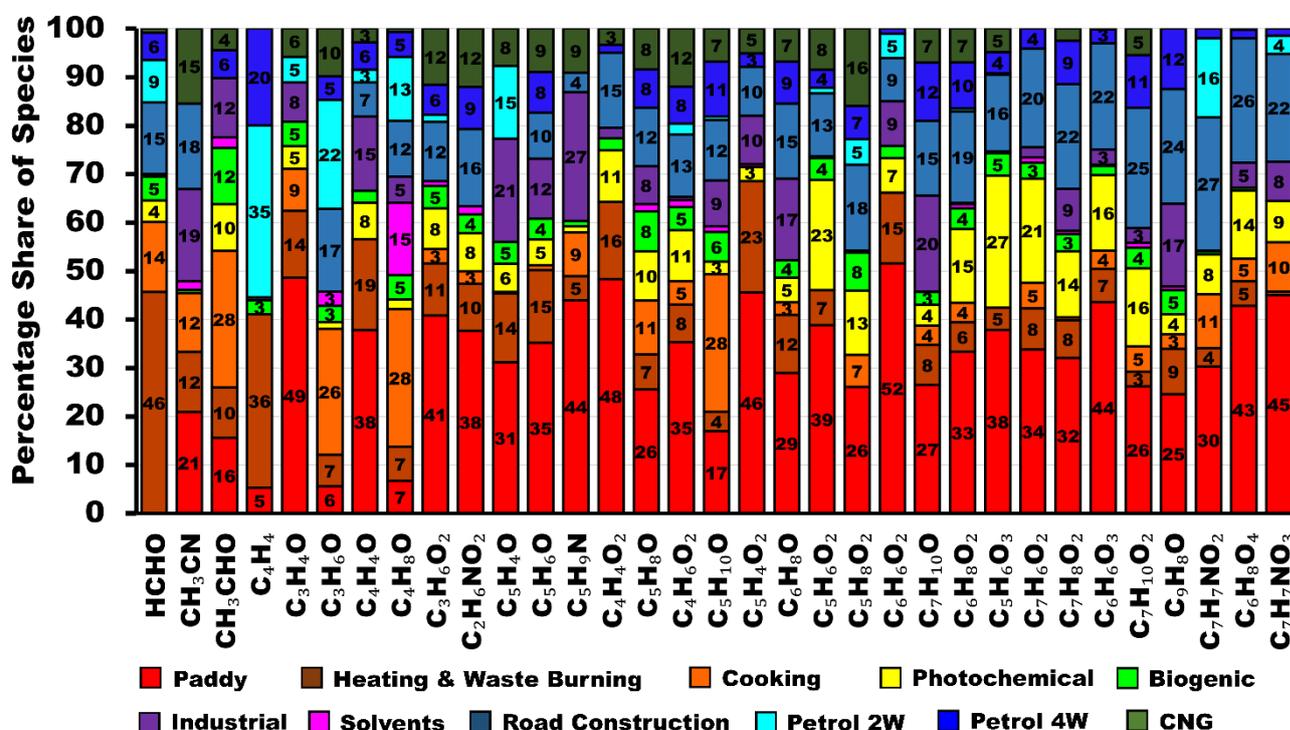


Figure 6: VOC species to which different forms of biomass burning contribute the highest percentage share of the atmospheric burden in Delhi

3.2.2 Factor 2: Residential heating and waste burning

The residential heating and waste burning factor is the second largest particulate matter source at the receptor site and contributes 23 % and 24 % to the total PM₁₀ and PM_{2.5} mass loadings, respectively (Fig. 4), while it contributed only 7% to the total VOC mass loading, 6% and 4% to the ozone and SOA formation potential respectively (Fig.4). Emissions peak at nighttime (Fig. 5) and the factor contribution time series displays the largest cross-correlation with the 24 h averaged heating demand (R=0.8) (Fig. S6), PM₁₀ (R=0.7), PM_{2.5} (R=0.6) and NO₂ (R=0.7) and CO (R=0.5) (Table S4). The lower correlation with NO (R=0.4) (Table S4), indicated that emissions are combustion-related but not always fresh. Occasionally, fresh plumes reach the receptor within minutes, however the majority of plumes have a higher atmospheric age, as NO is a short-lived species and oxidized to NO₂ on the timescale of minutes in the presence of ozone (Sinha et al., 2014). The factor contribution time series is anti-correlated with temperature (R=-0.6) and has its strong correlation with the 24 h averaged heating demand (R=0.8) indicating that this combustion activity is primarily triggered by the need to keep warm. Figure S11 shows that the PM_{2.5} and PM₁₀ mass loadings at the receptor site increase by 13.9 µg m⁻³ and 22.3 µg m⁻³, respectively for each degree increase in the 24-h average heating demand. Earlier studies have documented the strong seasonality of open waste burning emissions over Delhi, as well as the diversity of fuel used in wintertime heating-related fires (Nagpure et al., 2015). This factor explains 7 % of the total VOC mass loading. The top contributors to the VOC mass of this factor are in descending rank of contribution: methanol (CH₃OH), propyne (C₃H₄), acetone + propanal (C₃H₆O), acetaldehyde (CH₃CHO), acetic acid (C₂H₄O₂) and benzene (C₆H₆). Figure 6 shows that this factor explains the largest percentage share of the total mass for formaldehyde (HCHO) and vinylacetylene + 1-buten-3-yne (C₄H₄), and the second largest percentage share of furfural (C₅H₄O₂), methylfurfural (C₆H₆O₂), furan (C₄H₄O), methyl furan (C₅H₆O), furanone (C₄H₄O₂) and acrolein (C₃H₄O). All these compounds are characteristic of biomass burning smoke (Hatch et al., 2015; Stockwell et al., 2015; Koss et al., 2018).

3.2.3 Factor 3: Solid fuel-based cooking

The cooking factor is a daytime factor and explains 10 % of the total VOC mass loading, 10% and 8 % of the ozone and SOA formation potential (Fig. 4), but only a negligible share of the total PM₁₀ (≤4 %) burden. The volatility oxidation space plot (Figure S9) also shows very little evidence of IVOC oxidation products that could partition into the aerosol phase. The activity peaks from 8 am to noon time, with a secondary peak in the early evening hours and persists throughout monsoon and post-monsoon season. Emissions reaching the receptor site show no correlation with NO (R=0.1) indicating plumes are not fresh. In descending rank of mass contribution (C₃H₆O), acetaldehyde (CH₃CHO), methanol (CH₃OH), toluene (C₇H₈), the sum of C₈ aromatics (C₈H₁₀), propyne (C₃H₄) and benzene (C₆H₆) contribute most to this factor. These aromatic compounds have been reported to originate from cooking emissions (Crippa et al., 2013). Figure 6 shows that factor explains the largest percentage share of butanone (C₄H₈O), pentanone (C₅H₁₀O), acetaldehyde (CH₃CHO), acetone (C₃H₆O), and benzaldehyde (C₇H₆O). All these compounds are characteristic of biomass burning smoke (Hatch et al., 2015; Stockwell et al., 2015; Koss et al., 2018).

3.2.4 Factor 4: CNG

CNG-fuelled vehicles are identified as the third largest source of PM₁₀ (15 %) and PM_{2.5} (11 %) and contribute 9 % to the total VOC burden (Fig. 4). The much higher contribution of this source to the coarse mode particulate

matter burden ($22.5 \mu\text{g m}^{-3} \text{PM}_{10}$) when compared to the fine mode particular matter burden ($10.4 \mu\text{g m}^{-3} \text{PM}_{2.5}$), confirms earlier emission-inventory-based estimates which flagged that non-tailpipe emissions such as brake and tire wear and road dust resuspension have become the dominant transport sector related particulate matter sources in the Delhi-NCR region (Nagpure et al., 2016). Non-tailpipe emissions such as brake and tire wear and road dust resuspension contribute most to the PM_{10} burden, although they have also become the largest source of transport sector fine mode aerosol and VOC emissions in some countries that have transitioned to Euro-6 norms (Harrison et al., 2021). This study attributes a large share of these non-tailpipe emissions to trucks, buses and other commercial vehicles that are typically fuelled by CNG, because commercial diesel vehicles of <10 years age face severe entry restrictions, that limit their use within the Delhi NCR while older diesel vehicles have been completely banned from plying within City limits. Policy interventions in favour of CNG use (Krelling & Badami, 2022) have resulted in a halving of diesel sales, a rapid conversion of Delhi's HDV fleet to CNG (Figure S12), and a significant reduction in tailpipe exhaust emissions. In descending order methanol (CH_3OH), acetone + propanal ($\text{C}_3\text{H}_6\text{O}$), toluene (C_7H_8), C8 aromatic compounds (C_8H_{10}), butane (C_4H_8), propene (C_3H_6), and acetaldehyde (CH_3CHO) contribute most to the VOC mass in this source. Figure 7 shows that the factor explains the largest percentage share of methanol (CH_3OH) and the second largest percentage share of ethanol ($\text{C}_2\text{H}_6\text{O}$). These compounds are formed by the incomplete combustion of CNG that is catalytically converted to methanol and ethanol (Singh et al., 2016).

3.2.5 Factor 5: Petrol 4-wheeler factor

Figure 4 shows petrol 4-wheeler contributed 20 %, 25 %, and 30 % to the VOC mass loading, OFP, and SOAP, respectively. The source fingerprint of this source matched tailpipe emissions of petrol-fuelled 4-wheelers (Hakkim et al., 2021) and is characterized, in descending rank of contribution, by C8-aromatics, toluene, C9-aromatics (C_9H_{12}), benzene, butene + methyl tert-butyl ether (MTBE) fragment, propyne, propene, methanol and C2-substituted xylenes + C4-substituted benzenes ($\text{C}_{10}\text{H}_{14}$). Figure 5 shows that emissions peak in the evening between 7 pm and midnight with average VOC mass loadings $>70 \mu\text{g m}^{-3}$ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO ($R=0.8$), CO ($R=0.7$), and CO_2 ($R=0.7$) indicating the receptor site is impacted by fresh combustion emissions from this source and the atmospheric age of most plumes is on the timescale of minutes. Figure 7 shows that the factor explains the largest percentage share of most aromatic compounds, namely C8-aromatics, toluene, C9-aromatics (C_8H_{12}), C4-substituted benzene + C2-substituted xylene, benzene, styrene (C_8H_8), methylstyrenes + indane (C_9H_{10}), and C2-substituted styrenes ($\text{C}_{10}\text{H}_{12}$) and a few oxygenated aromatic hydrocarbons such as methyl phenol isomers ($\text{C}_7\text{H}_8\text{O}$) and methyl chavicol ($\text{C}_{10}\text{H}_{12}\text{O}$). The fact that the factor explains the largest percentage share of ethanol and the MTBE fragment (C_4H_8) can likely be attributed to ethanol blending and the use of MTBE in petrol (Achten et al., 2001). This factor also explains the largest percentage share of several other hydrocarbons such as propyne (C_3H_4), propene (C_3H_6), cyclopentadiene (C_5H_6), hexane (C_6H_{14}), C_7H_6 , C_7H_{10} , and cycloheptene (C_7H_{12}).

Figure S9 shows that this factor contributes significantly to the burden of C6- to C10 hydrocarbons, and hence SOA formation potential. However, due to freshly emitted plumes, it hardly contributes to the burden of the first- and second-generation oxidation products of these hydrocarbons at the receptor site. Instead, this factor is likely to contribute to secondary pollution formation downwind of the Delhi NCR.

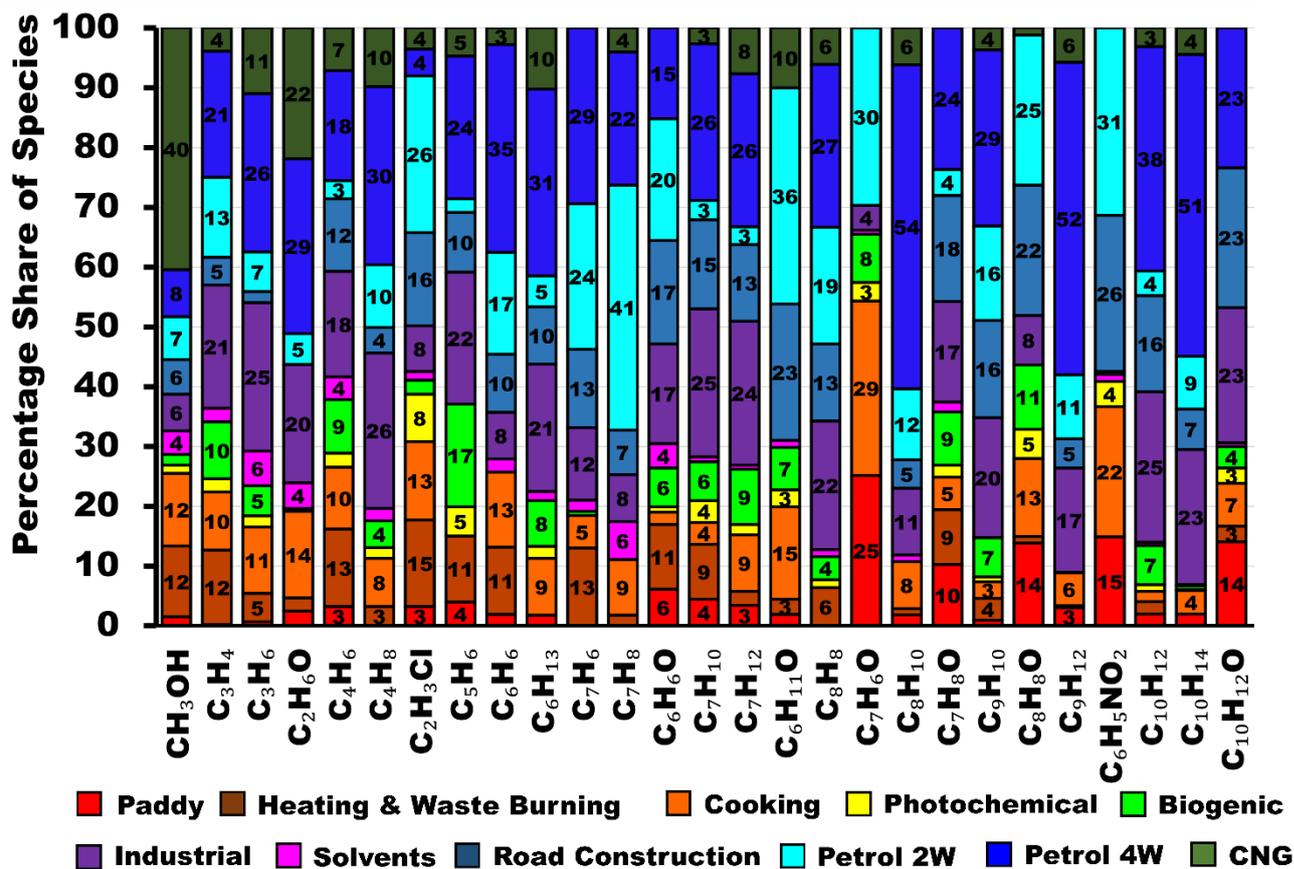


Figure 7: VOC species to which the transport sector contributes the highest percentage share of the atmospheric burden in Delhi.

3.2.6 Factor 6: Petrol 2-wheeler factor

Figure 4 shows petrol 2-wheeler contributed 14 %, 12 %, and 20 % to the VOC mass loading, OFP, and SOAP respectively. The source fingerprint of this source matched tailpipe emissions of petrol-based 2-wheelers (Hakkim et al., 2021) and are characterized, in descending rank of contribution, by toluene, acetone + propanal, C-8 aromatic compounds, acetic acid (C₂H₄O₂), propyne (C₃H₄), methanol (CH₃OH), benzene (C₆H₆), the MTBE fragment and C-9 aromatics (C₉H₁₂). A key difference of the petrol 2-wheeler source profile in comparison to the petrol 4-wheeler source profile is the lower benzene to toluene ratio, which is supported by the GC-FID analysis of tailpipe exhaust (Kumar et al., 2020). Figure 5 shows that emissions peak in the evening between 8 pm and 10 pm with average VOC mass loadings >50 µg m⁻³ and reach the receptor site from most wind directions. Emissions are strongly correlated with NO_x (R=0.6), CO (R=0.6) and CO₂ (R=0.7), but have a lower correlation with NO (R=0.5) (Table S4), and a larger contribution of oxygenated compounds to the source profile, indicating that the emissions have been photochemically aged. This suggests that contrary to 4-wheeler plumes which originate from the immediate vicinity of the receptor site in central Delhi (Figure S1), 2-wheeler plumes reach the receptor after prolonged transport from more distant rural and suburban areas on the outskirts of the city. In such areas, people often favour two-wheelers over four-wheelers. Figure 7 shows that this factor explains the largest percentage share of toluene, and a number of oxygenated aromatic compounds such as benzaldehyde (C₇H₆O), tolualdehyde (C₈H₈O), and phenol (C₆H₆O). It also explains the largest percentage share of nitrobenzene (C₆H₅NO₂),

cyclohexanone (C₆H₁₀O), and vinyl chloride (C₂H₃Cl). It also explains the second largest percentage share of benzene, vinylacetylene (C₄H₄), , acetone + propanal, methoxyamine (CH₅NO) and butanoic acid/ethyl acetate (C₄H₉O₂).

3.2.7 Factor 7: Industrial

This factor contributes 12 %, 14 %, 15%, 8% and 3% to the VOC mass loading, OFP, SOAP, PM_{2.5} and PM₁₀ mass loading. On average more than 30 μg m⁻³ to the VOC burden throughout the night from 9 pm to 7 am (Fig. 5) is from this factor. This factor is identified as an industrial point sources located in the wind sector S to SW of the receptor site. Emissions are most strongly correlated with CO (R=0.7), NO (R=0.7), CH₄ (R=0.8), and CO₂ (R=0.8) indicating that the emissions are fresh and originate from combustion processes. The main contributors towards the VOC mass in the industrial factor, are in descending order of contribution propyne (C₃H₄), butene + MTBE fragment (C₄H₈), toluene (C₇H₈), C-8 aromatic compounds (C₈H₁₀), propene (C₃H₆), acetaldehyde (CH₃CHO), methanol (CH₃OH), C-9 aromatics and the sum of monoterpenes (C₁₀H₁₆). The source fingerprint is most similar to ambient air grab samples collected near the Okhla waste to energy plant and industrial area in Faridabad.

Figure 8 shows that the factor explains the largest percentage share of methanethiol (CH₃S), a chemical used in the manufacture of the essential amino acid methionine, in the plastic industry and the manufacturing of pesticides, dichlorobenzenes (C₆H₄Cl₂), a chemical used in the synthesis of dyes, pesticides, and other industrial products and methoxyamine (CH₅NO). Analyses of the primary dataset by Mishra et al. (2024) also qualitatively inferred an industrial source for methanethiol and dichlorobenzene. It also explains the largest percentage share of the sum of monoterpenes, camphor/pinene oxide (C₁₀H₁₆O), santene (C₉H₁₄) the terpene fragment (C₈H₁₂), C₈H₁₄, C₉H₁₆, cyclohexene (C₆H₁₀) and cyclopentylbenzene (C₁₁H₁₄). Terpenes are used in the food and beverages, cosmetics, pharmaceutical, and rubber industry. In addition, this factor also explains the largest percentage share of a large suite of volatile and IVOC aromatic hydrocarbons including naphthalene (C₁₀H₈), methyl naphthalene (C₁₁H₁₀), C₁₂H₁₆, C₁₃H₁₈, C₁₃H₂₀, C₁₃H₂₂, C₁₄H₂₀, and C₁₄H₂₂. Ambient observations for most of these IVOCs have not been reported in the literature so far. Only, C₉H₁₄, C₁₂H₁₂, and C₁₂H₁₆ have been reported from aircraft engine emissions (Kılıç et al., 2018) while terpenes, C₉H₁₆, cyclopentylbenzene, naphthalene and methyl naphthalene have been reported from a wide range of combustion sources (Hatch et al., 2015, Bruns et al., 2017). Most other compounds have so far only been reported to degas from heated asphalt (Khare et al., 2020). Due to the high abundance of IVOCs in this factor, it contributes 15 % to the total SOA formation potential. Figure S9 shows the volatility oxidation state plot for all 111 VOCs in which the marker size represents the percentage share of each compound explained by the industrial factor and markers are colour coded by the number of carbon atoms. The plot shows evidence of the first- and second-generation oxidation products of C₆ to C₁₀ hydrocarbon transitioning from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009). This and the fact that the entire aerosol associated with this factor is PM_{2.5}, indicates that most of the aerosol associated with this factor is likely SOA.

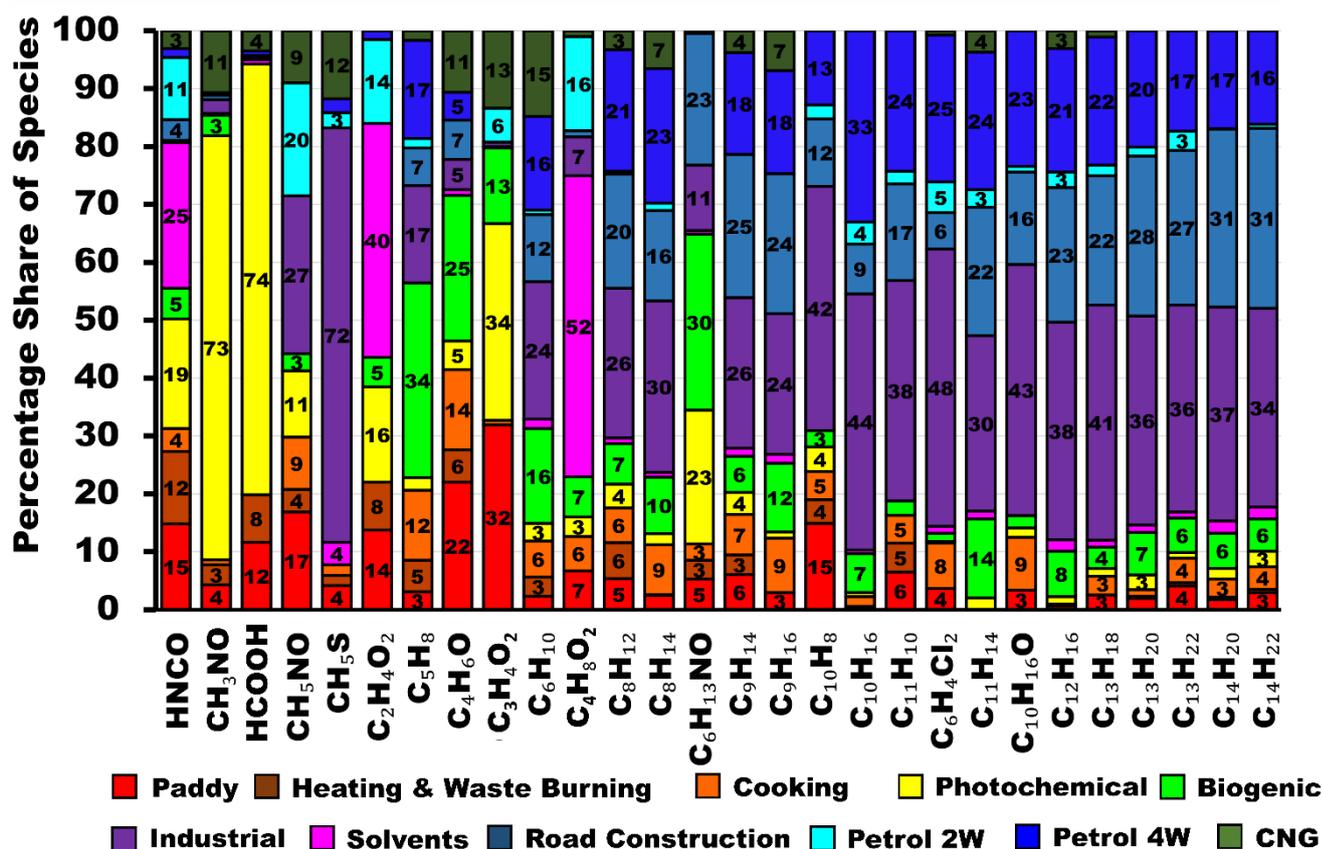


Figure 8: VOC species to which the industries, solvent usage, photochemistry or biogenic sources contribute the largest percentage share of the atmospheric burden in Delhi.

3.2.8 Factor 8: Solvents and Evaporative Emissions

Solvent usage and evaporative emissions reach the site from several point sources and wind directions often in the form of short and intense plumes that show no correlation with combustion tracers. This source contributes most to the VOC burden at night and explains 6 % of the total VOC but ≤ 1 % of the total PM_{2.5} and PM₁₀ mass (Fig. 4). The source fingerprint of the solvents factor (Fig. 3) is characterized in descending rank of mass contribution by acetic acid + glycolaldehyde (C₂H₄O₂), toluene (C₇H₈), methanol (CH₃OH), butanoic acid/ethyl acetate (C₄H₈O₂), acetone + propanal (C₃H₆O) and butanal + butanone + MEK (C₄H₈O). Figure 8 shows that the factor explains the largest share of organic acids namely butanoic acid, acetic acid and isocyanic acid (HNCO) and the second largest share of butanal + butanone + MEK (C₄H₈O). These compounds point towards stack venting of VOCs from chemical-, food-, or pharmaceutical industries or polymer manufacturing as likely sources of these emissions (Hodgson et al., 2000, Villberg et al., 2001, Jankowski et al., 2017, Gao et al., 2019). This assessment is broadly confirmed by the fact that the best match (R=0.7) for this source was collected from a plot situated opposite a polymer manufacturing unit and next to a pet food manufacturer in an industrial area at Jahangirpuri N of the receptor site.

3.2.9 Factor 9: Road construction

The road construction factor contributed 8% to the total VOC mass loading and 2% to the total PM₁₀. This factor is almost absent during monsoon season, as road repair work is mostly avoided during this period due to water

logging risks, and emissions from this source generally peak during the day as degassing of compounds from asphalt is temperature-driven and continues for days after the initial paving (Khare et al., 2020). The source fingerprint of the road construction factor is characterized in descending order of the mass concentrations by acetone + propionaldehyde, toluene, methanol, benzene and C8-aromatics. Acetone and propionaldehyde were found to be the most abundant oxygenated volatile compounds emitted during asphalt paving (Li et al., 2020). The source profile had the greatest similarity with the mix of emissions that would originate from asphalt paving (Li et al., 2020) and the tailpipe of road construction vehicles (Che et al., 2023). As represented by Fig. 9, this factor explains the largest percentage share of a large suite of volatile and IVOC hydrocarbons namely, heptene (C_7H_{14}), $C_{11}H_{12}$, $C_{12}H_{12}$, $C_{14}H_{14}$, $C_{14}H_{18}$, $C_{16}H_{24}$, $C_{17}H_{28}$, and $C_{18}H_{30}$. In addition, it explains the second largest percentage share of many other IVOC hydrocarbons namely C_9H_{14} , C_9H_{16} , $C_{11}H_{14}$, $C_{12}H_{16}$, $C_{13}H_{18}$, $C_{13}H_{20}$, $C_{13}H_{22}$, $C_{14}H_{20}$, $C_{14}H_{22}$. Except for the four hydrocarbons C_7H_{14} , C_9H_{14} , C_9H_{16} , and $C_{11}H_{12}$, all of these IVOCs have been reported to degas at 60°C from asphalt pavement (Khare et al., 2020). So far only $C_{14}H_{18}$ has been reported as fresh gas phase emissions (transport time <2.5 min) from a farm (Loubet et al., 2022) in ambient air, while $C_{17}H_{28}$ has been reported in the aerosol phase (Xu et al., 2022). The road construction factor also explains the largest percentage share of a long list of OVOCs namely, C6 diketone isomers ($C_6H_{10}O_2$), C2-substituted phenol ($C_8H_{10}O$), $C_7H_{12}O_2$, $C_8H_{14}O_2$, $C_8H_{16}O_2$, phthalic anhydride ($C_8H_4O_3$), which is a naphthalene oxidation product (Bruns et al., 2017), $C_9H_{10}O$, $C_9H_{12}O_2$, $C_9H_{14}O_2$, $C_9H_{16}O_2$, $C_9H_{18}O_2$, $C_{10}H_{12}O$, $C_{10}H_{18}O$, $C_{10}H_8O_3$, $C_{10}H_{16}O_3$, and $C_{12}H_{18}O_2$. However, out of these only $C_{10}H_{12}O$ and $C_{10}H_{18}O$ have been detected as direct emissions from heated asphalt pavement (Khare et al., 2020) indicating that most OVOCs in this factor are possibly oxidation products of short-lived IVOCs hydrocarbons emitted by this source. This assessment is supported by the volatility oxidation state plot for the road construction factor (Figure S9) which demonstrates that both precursors and oxidation products are present in this factor and that C6 to C10 hydrocarbons appear to be progressing from the VOC to the IVOC range along trajectories expected for the addition of =O functionality to the molecule (Jimenez, et al. 2009).

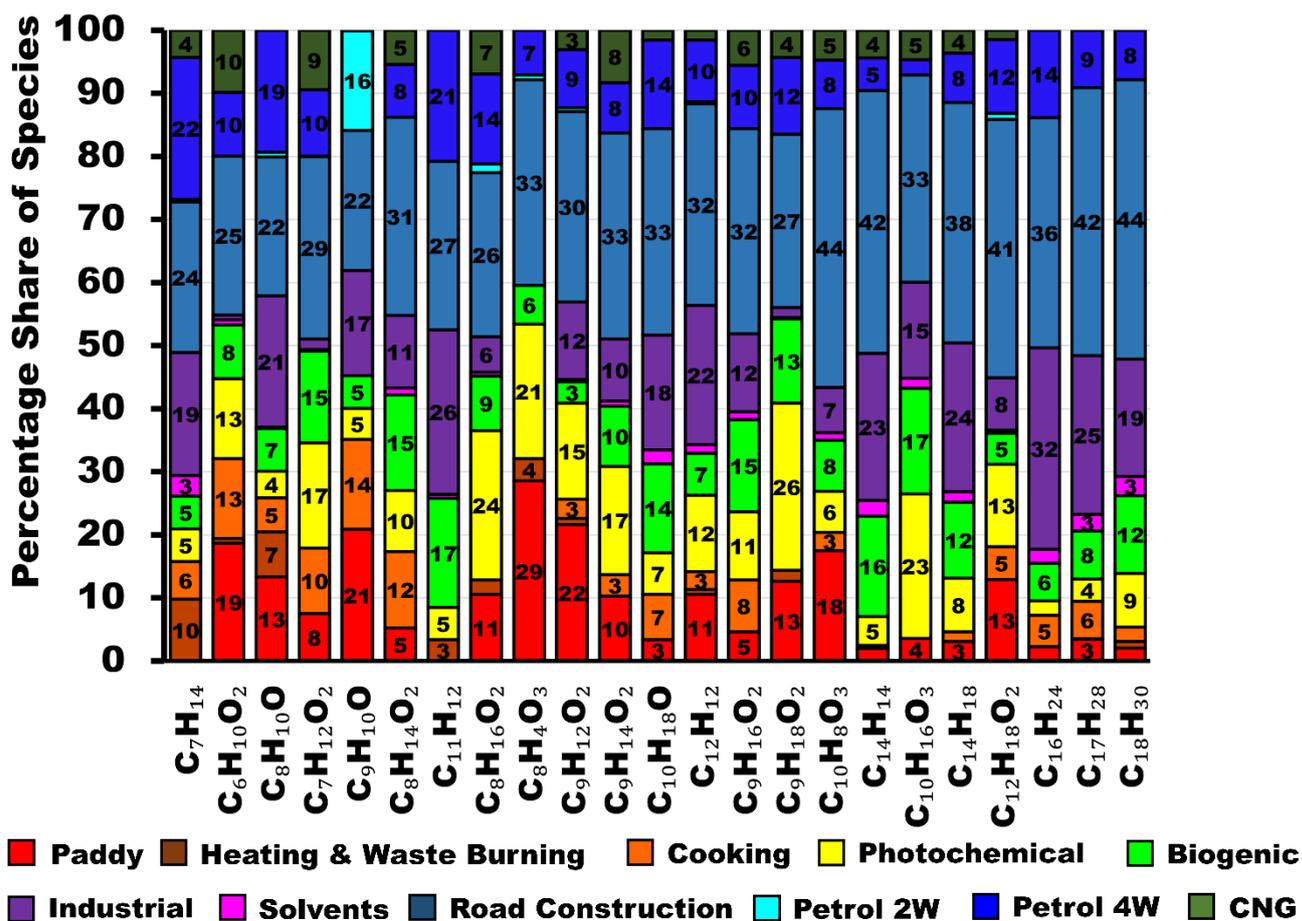


Figure 9: VOC species to which road construction contributes the largest percentage share of the atmospheric burden in Delhi.

3.2.10 Factor 10: Photochemistry

The photochemical factor has a diurnal profile that follows the diurnal profile of ozone (R=0.4). The factor profile is dominated by OVOCs such as acetic acid (C₂H₄O₂), formic acid (CH₃O₂), acetaldehyde (CH₃CHO), formamide (CH₄NO), and methanol (CH₃OH). Figure 8 shows that the factor explains the largest percentage share of formic acid, formamide, and methyl glyoxal (C₃H₅O₂). It also explains the second largest percentage share of isocyanic acid (HNCO) and hexanamide (C₆H₁₃NO), which are formed by the photooxidation of amines (Yao et al., 2016; Wang et al., 2022). Some compounds point towards a significant contribution of photochemically aged biomass burning emissions to this factor for example furfuryl alcohol (C₅H₆O₂), hydroxymethyl furanone (C₅H₆O₃), and hydroxybenzaldehyde (C₇H₆O₂). While this factor explained ≤4 % of the total VOC share and negligible share of PM_{2.5} and PM₁₀ mass in Delhi, photochemically aged biomass burning emissions were a significant source of VOCs at a suburban site in Punjab during the post-monsoon season of 2017 (Singh et al., 2023). The difference is likely due to the fact that great smog episode of 2017 was primarily driven by low wind speeds a shallow boundary layer and regional-scale build-up of emissions over a prolonged period (Dekker et al., 2019, Roozitalab, et al., 2021), while the post-monsoon season of 2022 experienced western disturbances and higher ventilation coefficients. The factor also explains the largest percentage share of the total mass for organic acids such as nonanoic acid (C₉H₁₈O₂), n-octanoic acid (C₈H₁₆O₂) which have been detected in biomass-burning impacted environments in China (Mochizuki et al., 2019), C₁₂H₁₈O₂ which has been found in aged wildfire plumes in the

US (Haeri, 2023), and the terpene ozonolysis products norpinonaldehyde ($C_9H_{14}O_2$) and cis-Pinonic acid ($C_{10}H_{16}O_3$) (Camredon et al., 2010) and $C_7H_{12}O_2$. Pinonic acid was found to be an important aerosol phase tracer of biogenic SOA formation in India (Mahilang et al., 2021) and $C_7H_{12}O_2$ has been reported as a pinonic acid aqueous-phase photolysis product (Lignell et al., 2013) Fig. 8.

3.2.11 Factor 11: Biogenic

Biogenic VOC emissions at the receptor site show the highest cross-correlation with photosynthetic active radiation (PAR, $R=0.7$) and temperature ($R=0.7$) (Table S4) and explain 4 % of the total VOC burden and 2 % of the PM_{10} burden in the PMF. The BVOC emission in this factor is relatively fresh as the ratio of isoprene to its first-generation oxidation products MEK (C_4H_8O) and MVK+MACR (C_4H_6O) is 5.9 and 3.0 respectively. At the site, the top of the tree canopy of roadside trees is located approximately 20 m below the inlet height. Figure 3 shows that in descending rank of mass contribution, acetaldehyde (CH_3CHO), C_3H_4 , isoprene (C_5H_8), acetic acid + glycolaldehyde ($C_2H_4O_2$), and acetone + propanal (C_3H_6O) are the major contributors for biogenic factor indicating that leaf wounding compounds contribute significantly to the BVOC burden in Delhi (Portillo-Estrada et al., 2015). The signal at $m/z41.035$ can potentially be attributed to a C_3H_4 the 2-methyl-3-butene-2-ol fragment (Kim et al., 2010; Park et al., 2013). Figure 8 shows that this factor explains the largest percentage share of two BVOCs namely Isoprene + 2-methyl-3-butene-2-ol fragment, and its oxidation product, methyl vinyl ketone, methacrolein and 2-butenal. It also explains the largest percentage share of C6 amides ($C_6H_{13}NO$) which are produced by the photo-oxidation of amines (Yao et al., 2016). The potential precursor, C6-amines has previously been detected in forested environments (You et al., 2014). However, it is also possible that C-6 amides are only attributed to the biogenic factor because their diurnal concentration profile matches that of first-generation oxidation products, and the source strength is high during both monsoon and post-monsoon season. This type of time series would also be expected if the precursors of this oxidation product are emitted from agricultural activities.

3.3 Comparison with emission inventories

Figure 10 shows a comparison of different anthropogenic emission inventories with the PMF output data from this study for three overlapping fetch regions corresponding to the fetch region from which air masses will reach the receptor site within 24 hours for different airflow patterns (Figure 1).

One feature that stands out in this comparison is that all inventories appear to significantly overestimate the relative contribution of residential fuel usage to the VOC and particulate matter emissions for all fetch regions. In absolute terms, the Regional Emission Inventory in Asia (REAS v3.2.1) for the year 2015 (Kurokawa & Ohara, 2020) and the Emission Database for Global Atmospheric Research (EDGARv6.1) for the year 2018 (Crippa et al., 2022), agree on the residential sector $PM_{2.5}$ emissions for the NW fetch region (Table S5). According to the latest estimates (Pandey et al., 2021), the NW-IGP region has the lowest prevalence of solid fuel usage in the entire IGP and the inventories appear to overestimate the $PM_{2.5}$ emissions from this fetch region only by a factor of 1.5-1.9. For the SW and SE fetch region, respectively, REAS v3.2.1 estimates much larger residential sector $PM_{2.5}$ emissions than EDGARv6.1 and overestimates the PMF estimates by a factor of 3.7 and 4.6. In contrast, EDGARv6.1 only overestimates PMF estimates by a factor of 1.8 and 3.2, for the SW and SE fetch region respectively. Solid fuel-based cooking is more prevalent in both Central and Western India and the Eastern IGP

than in the NW-IGP (Pandey et al., 2021). The overestimation in both inventories may be caused by a gradual adoption of cleaner technology. Sharma et al., (2022) calculated a 13 % drop in residential sector PM_{2.5} emissions between 2015 and 2020 due to higher LPG sales and a continuation of that trend to 2022 could explain the overestimation of residential fuel usage in the present emission inventory data. For PM₁₀, the EDGARv6.1 emission estimates for the NW, SW and SE fetch region, are greater than the REASv3.2.1 emission inventory. The EDGARv6.1 and REASv3.2.1 inventory both overestimate our PMF PM₁₀ results by a factor of 1.5 to 3.0. However, while the REASv3.2.1 inventory appears to assume that most of the residential sector aerosol emissions occur in the fine mode, our PMF results (Fig. 10) clearly agree with the EDGARv6.1 inventory on the fact that there are significant coarse aerosol emissions associated with solid-fuel based cooking and heating. Table S5 shows that for residential sector VOCs emissions, the absolute emissions in the EDGARv6.1 inventory are almost twice as large as those in the REASv3.2.1 inventory, even though the percentage contribution of this sector to the VOC emissions in the inventory in Figure 10 appears to be similar for both, because of larger VOC emissions from solvent use and industries in the EDGARv6.1 inventory. Both inventories overestimate the relative importance of residential sector emissions in relation to VOC emissions from other sectors by more than a factor of two when compared to our PMF estimate, most likely because they have not been updated with recent fuel shifts towards LPG in the relatively prosperous Delhi NCR region (Sharma et al., 2022).

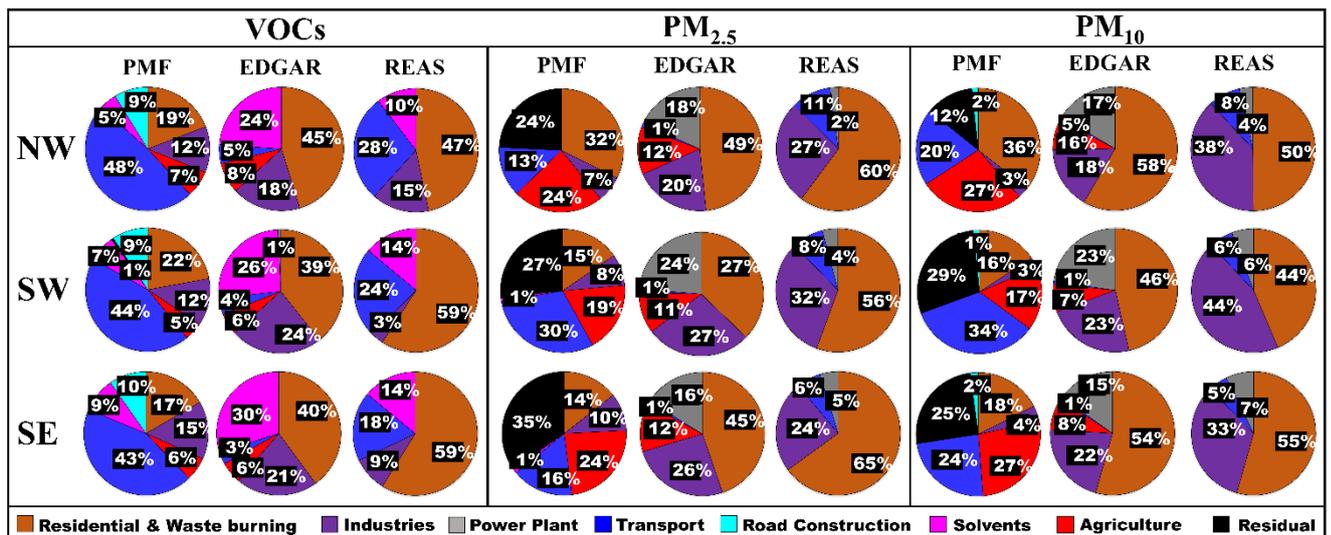


Figure 10: Comparison of different anthropogenic emission inventories with the PMF output from this study for three overlapping fetch regions corresponding to different airflow patterns.

With respect to industrial emissions of VOCs for the NW fetch region, our PMF results indicate that the actual emissions are slightly smaller than those in the REASv3.2.1 inventory, while the EDGARv6.1 inventory overestimates emissions. For the SW and SE fetch region, our PMF estimates fall in between those of the EDGARv6.1 inventory and the REASv3.2.1 inventory. For industrial PM_{2.5} emissions, both EDGARv6.1 & REASv3.2.1 are close and agree on the magnitude of emissions for the NW, SW and SE fetch region, respectively, and both inventories appear to overestimate emissions when compared to our PMF results. Our findings seem to suggest that the pollution boards have been somewhat successful in clamping down on industrial emissions and the technology employed is better than what is currently reflected in emission inventories. Industrial fly ash (PM₁₀) emissions are larger in the REASv3.2.1 inventory for all the fetch regions compared to EDGARv6.1 inventory. Yet both inventories appear to significantly overestimate industrial emissions when compared to our PMF results.

These findings also indicate the pollution boards have been somewhat successful in clamping down on large and visible fly ash sources and that the EDGARv6.1 inventory has captured this clean-technology transition better.

The REASv3.2.1 inventory completely misses direct VOC and PM emissions from the agricultural sector. The EDGARv6.1 inventory significantly underestimates PM_{2.5} & PM₁₀ emissions from agricultural activities, which include, but are not limited to crop residue burning, in comparison to our PMF results, particularly over NW-India (Table S5). Over this fetch region EDGARv6.1 attributes as much PM_{2.5} to all agricultural activities combined for the full year as the FINNv2.5 inventory (Wiedinmyer et al., 2023) attributes just to agricultural residue burning activities taking place between 15th August and 26th November 2021 (a time period comparable to the period in our model run), without including the emissions from rabi crop residue burning in summer (Kumar et al., 2016) and other agricultural activities such as harvest and ploughing. For PM₁₀ the fire count based FINNv2.5 estimate is twice as high as the emission estimate of EDGARv6.1 for this fetch region, and more likely to be correct, because the phytoliths present in rice straw form coarse mode ash during the combustion process (Figure S10). The fact that EDGAR appears to underestimate residue-burning emissions over this fetch region has been flagged earlier (Pallavi et al., 2019; Kumar et al., 2021; Singh et al., 2023). Our PMF analyses also reveals that the relative contribution of agricultural residue burning to the PM burden over the North-Western IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) and South-Eastern IGP (24 % and 27 % of PM_{2.5} and PM₁₀, respectively) is comparable, despite the much lower fire counts over the South-Eastern IGP (17,810), when compared to the North Western IGP (61,334). This indicates that either fires to the SE are burning closer to the receptor site or the fire detection efficiency in this fetch region is lower. Table S5 reveals that the relative importance of agricultural emissions over the SE fetch region is even more severely underestimated in the FINNv2.5 inventory than in the EDGARv6.1 inventory due to poorer fire detection (close to 100% omission error) for the partial burns prevalent over this region (Lui et al. 2019; 2020, Figure S8) when compared to the complete burns prevalent over the NW IGP (Lui et al. 2019; 2020, Figure S7).

Transport sector VOC emissions appear to be severely underestimated in the EDGARv6.1 inventory for the NW, SW, and SE fetch region, which has been previously flagged for earlier versions of the same inventory (Sarkar et al., 2017; Pallavi et al., 2019; Singh et al., 2023). The REASv3.2.1 inventory also underestimates our PMF results. This indicates that the contribution of the transport sector to ambient VOC pollution levels in a megacity like Delhi may not be adequately reflected in both the emission inventories. Our PMF suggests that the overall contribution of the transport sector to the total PM_{2.5} and PM₁₀ pollution levels occurs primarily due to non-exhaust emissions from the CNG-fuelled public transport fleet. These non-exhaust emissions are much larger than what is accounted for both in the EDGARv6.1 and REASv3.2.1 inventories for PM_{2.5} & PM₁₀ emissions from the NW, SW and SE fetch region. The transport sector-related findings of this PMF source apportionment study are in agreement with earlier source apportionment studies that often attributed a quarter or more of the total PM emissions to the transport sector. Some prior studies used metals, Pb and/or OC/EC as transport sector activity tracers (Jain et al., 2017, 2020; Sharma et al., 2016, Jaiprakash et al., 2016; Sharma & Mandal, 2017), while others attributed almost the entire HOA component of organic aerosol to transport sector emissions (Reyes-Villega et al., 2021; Cash et al., 2021; Kumar et al., 2022, Shukla et al., 2023) or used a Chemical Mass Balance (CMB) model with source fingerprints from the EPA database (Nagar et al., 2017). Our PMF results differ to emission-inventory-based assessments, which only attribute a minor share of the total PM burden to this activity (Guo et al., 2017). Our findings also add insights to the reasons why the transport sector targeted air quality interventions

yielded such poor results (Chandra et al., 2018). Public transport availability was ramped up during the periods when road-rationing schemes restricted the use of private 4-wheelers. Our results suggest that moving forward only investments into the road infrastructure, that reduce resuspension, modal shifts from buses towards metro-based public transport and electric vehicles with >50 % regenerative braking (Liu et al., 2021) that limit brake wear can yield meaningful reductions in the transport sector-related PM emissions.

Our PMF results indicate that solvent usage results in VOC emissions that are more in line with the REASv3.2.1 inventory while the EDGARv6.1 inventory overestimates emissions by a factor of 4 for all the fetch regions.

Power generation is not considered to be a significant VOC source in both emission inventories (<1 % of the total VOC mass), and fails to show up as a separate sector in our PMF results, as our model runs rely on VOC tracers to track pollution sources. The contribution of energy generation towards the PM burden particularly in the EDGARv6.1 emission inventory, however, is significant. It is, however, striking to note that the PMF features a residual that is of similar magnitude as the $PM_{2.5}$ and PM_{10} emissions attributed to power generation in the EDGARv6.1 inventory. Power generation is believed to primarily contribute secondary sulfate and nitrate aerosol (Atabakhsh et al., 2023), which is unlikely to be directly associated with a fresh combustion signature. It is hence likely, that much of our PMF residual can be attributed primarily to this source. The amount of emissions attributed to power generation in the REASv3.2.1 inventory is much smaller than those reflected in EDGARv6.1, likely because the inventory misses several coal generation units that were commissioned between 2015-2018.

Our PMF results identify road construction and asphalt pavements as an additional VOC source that is at present not reflected in emission inventories.

4 Conclusions

This study presents source-apportionment results derived from application of the positive matrix factorization model to a recently acquired high-quality dataset of $PM_{2.5}$ & PM_{10} , and 111 VOCs measured using the new PTR-TOF-MS10K enhanced volatility instrument, during monsoon and post-monsoon seasons of 2022, from one world's most polluted megacities: Delhi. We found that the top ranked major emission source of gas phase and aerosol phase differed from each other, highlighting the complexity of air pollution sources in such atmospheric environments. While fresh paddy burning was a negligible source of VOCs (6 %), it was the largest source of $PM_{2.5}$ & PM_{10} (23 % & 25 %) in the Delhi NCR regions during our study period, likely because combustion of phytolith containing rice straw triggers the formation of coarse mode ash (Figure S10) that contributes significantly to the PM burden. $PM_{2.5}$ & PM_{10} are the two main criteria air pollutants regulated under the national ambient air quality standard that are thought to be the leading cause of the air pollution emergency in November in Delhi annually (Khan et al., 2023). The strong correlation of $PM_{2.5}$ & PM_{10} with same-day fire counts, and VOC emission signatures of fresh paddy burning plumes showed that fires burning in and within the vicinity of Delhi-NCR and plumes that reached the receptor on the same day were the stronger contributory source of the high pollution levels, compared to plumes from more distant states such as Punjab and Pakistan Punjab. Both are located north-west of Delhi-NCR and were thought to be the stronger contributors to the pollution levels because the detected fire activity is more prevalent there. Furthermore, $PM_{2.5}$ & PM_{10} emissions from residential heating and waste burning (24 % & 23 %) rival those from crop residue burning and unlike paddy residue burning emissions, which are episodic, this activity persists into winter. While popular perception generally blames

burning in Punjab for the high particulate matter burden due to paddy stubble burning, our PMF reveals that despite the much lower fire counts over the Eastern IGP (17,810) when compared to the North Western IGP (61,334) both are a significant source of paddy stubble burning PM in the NCR region. Also, sources that are generally targeted by most clean air action plans such as tailpipe exhaust emissions of private vehicles and industries are responsible for less than one-quarter of the particulate matter mass loading that can be traced with the help of gas-phase organic molecular tracers. Instead, the transport sector's PM emissions are dominated by the non-exhaust emissions such as road dust suspension, break wear and tire wear of the CNG-fueled commercial vehicle fleet, which according to a recent emission inventory for Delhi are one order of magnitude larger than the transport sector tailpipe exhaust emissions (Nagpure et al., 2016).

The PMF results based on primary in-situ data indicate that the EDGARv6.1 inventory provides a better representation of emissions than the REASv3.2.1 inventory for most sectors, with the exception of transport sector emissions and VOC emissions from solvent use. Agricultural burning emissions over the NW-IGP are best represented in FINNv2.5 (Wiedinmyer et al., 2023) while agricultural emissions over the SE-IGP are better captured by EDGARv6.1. At present none of the residential sector inventories appears to have incorporated the change in the magnitude and spatial patterns due to the recent adoption of cleaner cooking technology interventions since 2018. Transport sector non-exhaust emissions are still absent (REASv3.2) or underestimated (EDGARv6.1) in all inventories. For VOC emissions from solvent usage, REASv3.2 provides better emissions than EDGARv6.1. There is also a road construction sector in our PMF results which has a significant (9-10 %) contribution to the VOC burden but hasn't been addressed in any of the emission inventories so far, and our study by including measurements of specific molecular markers of this activity has been able to shed new strategic insights concerning this missing source.

A considerable portion of the PM₁₀ (18 %) and PM_{2.5} (28 %) load is connected to residual sources, not directly related to combustion tracers. This contribution is likely due to windblown dust transported over long distances (Pawar et al., 2015) as well as secondary inorganic aerosols like ammonium sulfate and ammonium nitrate whose precursors are primarily emitted from power plants. Despite including the most comprehensive set of organic species measured in Delhi to date, our study does not include similar information about these other species.

Residential heating and waste burning was identified as one of the largest contributors to PM pollution, and this source is active year-round with strengths varying depending on seasonality. The total contribution of residential sector solid fuel usage and waste burning (17 % in Delhi and 18 % in Mohali) to the VOC burden during post-monsoon season was similar at both sites. So, targeting these through improved access to cleaner energy sources for heating and cooking would likely improve air quality significantly in other seasons. Future similarly designed quantitative studies would be needed to confirm this hypothesis.

The findings and insights from this study emphasize the necessity for a comprehensive, multi-sectoral approach to reduce primary emissions. While several recent efforts in some sectors (e.g. residential biofuel and cooking) appear to have yielded emission reduction benefits, the narrative to blame the post-monsoon pollution exclusively on the more visible sources (e.g. paddy residue burning), needs to be corrected so other sources are also mitigated. Our findings support the assertions of (Ganguly et al., 2020), who have pointed out previously that, rather than solely focusing on specific sources like agricultural residue burning or transport emissions, it's crucial to address the disparity between the primary targets of clean air action plans and the actual dominant sources of particulate matter. Future action plans need to account for more targeted and impactful pollution control measures and also

a more comprehensive approach to address the diverse urban mixed sources highlighted in this study, such as industries and residential solid fuel/waste burning, non-exhaust road emissions, and emissions from road construction.

This new approach of combining VOC tracers with PM measurements provides great potential for improved source apportionment in complex emission environments, at a level of detail that is more meaningful than just attributing emissions to biomass burning or fossil-fuel burning, which has been the case in all previous studies from the region to date. Previously in Delhi-MCR region, Kumar et al. 2022 identified “cooking-related OA using EESI-TOF analysis but due to analytical limitations, the paper only reported quantitative data for three primary factors, namely HOA, BBOA-1 and BBOA-2, without naming the activities responsible for the formation of BBOA-1 and BBOA-2. One of the more comprehensive AMS based studies (Cash et al., 2021) spanning pre-monsoon, monsoon and post monsoon season of the year 2018 only identified three different primary biomass burning factors, namely cooking organic aerosol (6% of PM_{10}), solid fuel organic aerosol ($\leq 11\%$ of PM_{10}), and semi-volatility biomass burning organic aerosol ($\leq 13\%$ of PM_{10}), that broadly appear to correspond to our solid fuel-based cooking (4% of PM_{10}), residential heating and waste burning (23% of PM_{10}), and paddy residue burning (25% of PM_{10}) factors. However, the study failed to name and attribute two of these three factors in policy relevant ways, could not identify the significant contribution of coarse mode fly ash to the total aerosol burden, and also was unable to distinguish between different fossil-fuel related sources. Our study design which captured contrasts between clean-monsoon and polluted-post-monsoon air, and included measured VOC source fingerprints and molecular tracers enabled us to distinguish paddy-residue burning from other biomass burning sources, and resolve similar traffic emission sources (e.g. 2-wheelers from 4-wheelers and CNG vehicles). This provides a significant advance over existing source-apportionment studies and its application would be of great relevance in other complex emission environments suffering from high air pollution where quantitative knowledge of sources can lead to evidence-based emission reduction prioritization efforts and a better understanding of the atmospheric chemistry of polluted environments around the world.

Data availability

PMF model simulations and input data can be obtained by contacting Baerbel Sinha.

Author Contribution

Arpit Awasthi: Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft preparation. Baerbel Sinha: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Resources, Software, Supervision, Validation, review & editing, Writing – review & editing. Haseeb Hakkim: Data curation, Formal analysis, Investigation, Writing – review & editing. Sachin Mishra: Data curation, Formal analysis, Investigation. Varkrishna M.: Investigation. Gurmanjot Singh: Investigation. Sachin D. Ghude: Resources. Vijay Kumar Soni: Resources. N. Nigam: Resources. Vinayak Sinha: Conceptualization, Data curation, Project administration, Methodology, Supervision, Writing – review & editing. M. Rajeevan: Resources.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgements:

We acknowledge the financial support given by the Ministry of Earth Sciences (MOEs), the government of India, to support the RASAGAM (Realtime Ambient Source Apportionment of Gases and Aerosol for Mitigation) project at IISER Mohali vide grant MOES/16/06/2018-RDEAS Dt. 22.6.2021. S.M acknowledges IISER Mohali for Institute PhD fellowship. AA acknowledges MoE for PMRF PhD fellowship. We thank Dr. R. Mahesh, Dr. Gopal Iyengar, Dr. R. Krishnan (Director, IITM Pune), Prof. Gowrishankar (Director, IISER Mohali), Dr. M. Mohapatra (DG, IMD), Dr. M. Ravichandran (Secretary Ministry of Earth Science) for their encouragement and support. We thank student members of the Atmospheric Chemistry and Emissions (ACE) research group and Aerosol Research Group (ARG) of IISER Mohali and IITM Pune in particular AkashVispute and PrassanaLonkar and local scientists of IMD for their local logistics support. The authors gratefully acknowledge the NASA/ NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites VIIRS fire count data used in this publication. The authors gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model used in this publication.

References

- Achten, C., Kolb, A., and Püttmann, W.: Methyl tert-butyl ether (MTBE) in urban and rural precipitation in Germany. *Atmos. Environ.*, 35(36), 6337-6345, [https://doi.org/10.1016/S1352-2310\(01\)00423-X](https://doi.org/10.1016/S1352-2310(01)00423-X), 2001.
- Acharja, P., Ali, K., Ghude, S. D., Sinha, V., Sinha, B., Kulkarni, R., Gultepe, I., Rajeevan, M. N.: Enhanced secondary aerosol formation driven by excess ammonia during fog episodes in Delhi, India, *Chemosphere* 289, 133155, <https://doi.org/10.1016/j.chemosphere.2021.133155>, 2022.
- Atabakhsh, S., Poulain, L., Chen, G., Canonaco, F., Prévôt, A. S. H., Pöhlker, M., Wiedensohler, A., and Herrmann, H.: A 1-year aerosol chemical speciation monitor (ACSM) source analysis of organic aerosol particle contributions from anthropogenic sources after long-range transport at the TROPOS research station Melpitz, *Atmos. Chem. Phys.*, 23, 6963–6988, <https://doi.org/10.5194/acp-23-6963-2023>, 2023.
- Bow, S.-T.: *Pattern Recognition, Application to Large Data Set Problems*, Marcel Dekker, Inc., 1984. 12.
- Therrien, C. W., *Decision, Estimation and Classification, An Introduction to Pattern Recognition and Related Topics*, John Wiley and Sons, <https://ui.adsabs.harvard.edu/abs/1984prat.book....B>, 1984.
- Bruns, E. A., Slowik, J. G., El Haddad, I., Kilic, D., Klein, F., Dommen, J., Temime-Roussel, B., Marchand, N., Baltensperger, U., and Prévôt, A. S. H.: Characterization of gas-phase organics using proton transfer reaction time-of-flight mass spectrometry: Fresh and aged residential wood combustion emissions, *Atmos. Chem. Phys.*, 17(1), 705–720, <https://doi.org/10.5194/acp-17-705-2017>, 2017.
- Buzcu-Guven, B., and Fraser, M. P.: Comparison of VOC emissions inventory data with source apportionment results for Houston, TX, *Atmos. Environ.* 42(20), 5032-5043, <https://doi.org/10.1016/j.atmosenv.2008.02.025>, 2008. ^b

Camredon, M., Hamilton, J. F., Alam, M. S., Wyche, K. P., Carr, T., White, I. R., Monks, P. S., Rickard, A. R., and Bloss, W. J.: Distribution of gaseous and particulate organic composition during dark α -pinene ozonolysis, *Atmos. Chem. Phys.*, 10, 2893–2917, <https://doi.org/10.5194/acp-10-2893-2010>, 2010.

Carslaw, D. C. and Ropkins, K.: openair – An R package for air quality data analysis, *Environ. Modell. Softw.*, 27–28, 52–61, <https://doi.org/10.1016/j.envsoft.2011.09.008>, 2012.

Carter, W. P. L.: Updated maximum incremental reactivity scale and hydrocarbon bin reactivities for regulatory applications, prepared for California Air Resources Board Contract 07-339, available at: <http://cmscert.engr.ucr.edu/~carter/SAPRC/MIR10.pdf> (last access: July 2018), 2000.

Cash, J. M., Langford, B., Di Marco, C., Mullinger, N. J., Allan, J., Reyes-Villegas, E., Joshi, R., Heal, M. R., Acton, W. J. F., Hewitt, C. N., Misztal, P. K., Drysdale, W., Mandal, T. K., Shivani, Gadi, R., Gurjar, B. R., and Nemitz, E.: Seasonal analysis of submicron aerosol in Old Delhi using high-resolution aerosol mass spectrometry: chemical characterisation, source apportionment and new marker identification, *Atmos. Chem. Phys.*, 21, 10133–10158, <https://doi.org/10.5194/acp-21-10133-2021>, 2021.

Chandra, B. P., and Sinha, V.: Contribution of post-harvest agricultural paddy residue fires in the N.W. Indo-Gangetic Plain to ambient carcinogenic benzenoids, toxic isocyanic acid and carbon monoxide. *Environ. Int.*, 88, 187–197, <https://doi.org/10.1016/J.ENVINT.2015.12.025>, 2016.

Chandra, B. P., Sinha, V., Hakkim, H., and Sinha, B.: Storage stability studies and field application of low-cost glass flasks for analyses of thirteen ambient VOCs using proton transfer reaction mass spectrometry. *Int. J. Mass. Spectrom.*, 419, 11–19, <https://doi.org/10.1016/j.ijms.2017.05.008>, 2017.

Chandra, B. P., Sinha, V., Hakkim, H., Kumar, A., Pawar, H., Mishra, A. K., Sharma, G., Pallavi, Garg, S., Ghude, S. D., Chate, D. M., Pithani, P., Kulkarni, R., Jenamani, R. K. and Rajeevan M.: Odd–even traffic rule implementation during winter 2016 in Delhi did not reduce traffic emissions of VOCs, carbon dioxide, methane and carbon monoxide. *Curr. Sci. India.*, Vol. 114, 1318–1325, <https://www.jstor.org/stable/26797338>, 2018.

Chaudhary, P., Garg, S., George, T., Shabin, M., Saha, S., Subodh, S., Sinha, B.: Underreporting and open burning—the two largest challenges for sustainable waste management in India. *Resour. Conserv. Recycl.* 175, 105865, <https://doi.org/10.1016/j.resconrec.2021.105865>, 2021.

Chaudhary, P., Singh, R., Shabin, M., Sharma, A., Bhatt, S., Sinha, V., and Sinha, B.: Replacing the greater evil: Can legalizing decentralized waste burning in improved devices reduce waste burning emissions for improved air quality? *Environ. Pollut.*, 311, 119897, <https://doi.org/10.1016/J.ENVPOL.2022.119897>, 2022.

Che, H., Shen, X., Yao, Z., Wua, B., Gou, R., Hao, X., Cao, X., Li, X., Zhang, Z., Wang, S. and Chen, Z.: Real-world emission characteristics and inventory of volatile organic compounds originating from construction and agricultural machinery. *Sci. Total. Environ.* 894, 164993, <https://doi.org/10.1016/j.scitotenv.2023.164993>, 2023.

Crippa, M., Canonaco, F., Slowik, J. G., El Haddad, I., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Temime-Roussel, B., Abidi, E., Poulain, L., Wiedensohler, A., Baltensperger, U., and Prévôt, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, *Atmos. Chem. Phys.*, 13, 8411–8426, <https://doi.org/10.5194/acp-13-8411-2013>, 2013.

Crippa, M., Guizzardi, D., Muntean, M., and Schaaf, E.: EDGAR v5.0 Global Air Pollutant Emissions, European Commission, Joint Research Centre (JRC) [data set], <http://data.europa.eu/89h/377801af-b094-4943-8fdc-f79a7c0c2d19> (last access: 2 April 2020), 2019a.

Coggon, M. M., Lim, C. Y., Koss, A. R., Sekimoto, K., Yuan, B., Gilman, J. B., Hagan, D. H., Selimovic, V., Zarzana, K. J., Brown, S. S., M Roberts, J., Müller, M., Yokelson, R., Wisthaler, A., Krechmer, J. E., Jimenez, J. L., Cappa, C., Kroll, J. H., De Gouw, J., and Warneke, C.: OH chemistry of non-methane organic gases (NMOGs) emitted from laboratory and ambient biomass burning smoke: Evaluating the influence of furans and oxygenated aromatics on ozone and secondary NMOG formation. *Atmos. Chem. Phys.*, 19(23), 14875–14899, <https://doi.org/10.5194/ACP-19-14875-2019>, 2019.

Datta, S., Sharma, A., Parkar, V., Hakkim, H., Kumar, A., Chauhan, A., Tomar S. S., and Sinha, B.: A new index to assess the air quality impact of urban tree plantation. *Urban Climate* 40 (2021) 100995, <https://doi.org/10.1016/j.uclim.2021.100995>, 2021.

Dekker, I.N., Houweling, S., Pandey, S., Krol, M., Röckmann, T., Borsdorff, T., Landgraf, J., and Aben, I.: What caused the extreme CO concentrations during the 2017 high pollution episode in India? *Atmos. Chem. Phys.* 19, 3433–3445. <https://doi.org/10.5194/acp-19-3433-2019>, 2019.

Derwent, R. G., Jenkin, M. E., Saunders, S. M., and Pilling, M. J.: Photochemical ozone creation potentials for organic compounds in northwest Europe calculated with a master chemical mechanism, *Atmos. Environ.*, 32(14–15), 2429–2441, [https://doi.org/10.1016/S1352-2310\(98\)00053-3](https://doi.org/10.1016/S1352-2310(98)00053-3), 1998.

Derwent, R. G., Jenkin, M. E., Utembe, S. R., Shallcross, D. E., Murrells, T. P., and Passant, N. R.: Secondary organic aerosol formation from a large number of reactive man-made organic compounds. *Sci. Total. Environ.*, 408(16), 3374–3381, <https://doi.org/10.1016/J.SCITOTENV.2010.04.013>, 2010.

Fadly, D., Fontes, F., and Maertens, M.: Fuel for food: Access to clean cooking fuel and Food Secur in India. *Food Secur*, 15(2), 301–321, <https://doi.org/10.1007/s12571-023-01350-y>, 2023.

Fleming, L. T., Weltman, R., Yadav, A., Edwards, R. D., Arora, N. K., Pillarisetti, A., Meinardi, S., Smith, K. R., Blake, D. R., and Nizkorodov, S. A.: Emissions from village cookstoves in Haryana, India, and their potential impacts on air quality, *Atmos. Chem. Phys.*, 18, 15169–15182, <https://doi.org/10.5194/acp-18-15169-2018>, 2018.

Gajbhiye, M. D., Lakshmanan, S., Aggarwal, R., Kumar, N., and Bhattacharya, S.: Evolution and mitigation of vehicular emissions due to India's Bharat Stage Emission Standards—A case study from Delhi. *Environmental Development*, 45, 100803, <https://doi.org/10.1016/j.envdev.2023.100803>, 2023.

Ganguly, T., Selvaraj, K. L. and Guttikunda, S. K. National Clean Air Programme (NCAP) for Indian cities: Review and outlook of clean air action plans. *Atmos. Environ.*: X, 8, 100096, <https://doi.org/10.1016/j.aeaoa.2020.100096>, 2020.

Guo, H., Kota, S. H., Sahu, S. K., Hu, J., Ying, Q., Gao, A., and Zhang, H.: Source apportionment of PM_{2.5} in North India using source-oriented air quality models, *Atmos. Chem. Phys.*, 17(1), 426–436, <https://doi.org/10.1016/j.envpol.2017.08.016>, 2017.

Gao, Z., Hu, G., Wang, H., Zhu, B.: Characterization and assessment of volatile organic compounds (VOCs) emissions from the typical food manufactures in Jiangsu province, China, *Atmos. Pollut. Res.* 10(2), 571–579, <https://doi.org/10.1016/j.apr.2018.10.010>, 2019.

Haeri, F.: Molecular Speciation of Organic Nitrogen Compounds Separated in Smoke Particles Emitted from Burning Western U.S. Wildland Fuels, PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, <https://doi.org/10.1184/R1/22670578.v1>, 2023.

Hakkim, H., Kumar, A., Annadate, S., Sinha, B., and Sinha, V.: Air pollution scenario analyses of fleet replacement strategies to accomplish reductions in criteria air pollutants and 74 VOCs over India. *Atmos. Environ.*: X, 13, 100150, <https://doi.org/10.1016/j.aeaoa.2022.100150>, 2021.

Harrison, M. A. J., Heal, M. R., and Cape, J. N.: Evaluation of the pathways of tropospheric nitrophenol formation from benzene and phenol using a multiphase model. *Atmos. Chem. Phys.*, 5, <https://doi.org/10.5194/acp-5-1679-2005>, 2022.

Harrison, M. A., Barra, S., Borghesi, D., Vione, D., Arsene, C., and Olariu, R. I.; Nitrated phenols in the atmosphere: a review. *Atmos. Environ.*, 39(2), 231-248, <https://doi.org/10.1016/j.atmosenv.2004.09.044>, 2005.

Hatch, L. E., Luo, W., Pankow, J. F., Yokelson, R. J., Stockwell, C. E., and Barsant K. C.: Identification and quantification of gaseous organic compounds emitted from biomass burning using two-dimensional gas chromatography–time-of-flight mass spectrometry. *Atmos. Chem. Phys.*, 15, 1865–1899, <https://doi.org/10.5194/acp-15-1865-2015>, 2015.

Harrison, R. M., Allan, J., Carruthers, D., Heal, M. R., Alastair, L. C., Marnier, B., Murrells, T., and Williams, A.: Non-exhaust vehicle emissions of particulate matter and VOC from road traffic: A review. *Atmos. Environ.* 262, 118592, <https://doi.org/10.1016/j.atmosenv.2021.118592>, 2021.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J., and Barsanti K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors. *Atmos. Chem. Phys.*, 17, 1471–1489, <https://doi.org/10.5194/acp-17-1471-2017>, 2017.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N.: ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), <https://doi.org/10.24381/cds.adbb2d47>, 2023.

Hodgson, S. C., Casey, R. J., Bigger, S. W., & Scheirs, J.: Review of volatile organic compounds derived from polyethylene. *Polym-Plast Technol*, 39(5), 845-874. <https://doi.org/10.1081/PPT-100101409>, 2000.

Jain, S., Sharma, S. K., Choudhary, N., Masiwal, R., Saxena, M., Sharma, A., Mandal, T. K., Gupta, A., Gupta, N. C., and Sharma, C.: Chemical characteristics and source apportionment of PM_{2.5} using PCA/APCS, UNMIX, and PMF at an urban site of Delhi, India. *Environ. Sci. Pollut. R.*, 24(17), 14637–14656, <https://doi.org/10.1007/s11356-017-8925-5>, 2017.

Jain, V., Tripathi, S.N., Tripathi, N., Sahu, L.K., Gaddamidi, S., Shukla, A.K., Bhattu, D., Ganguly, D.: Seasonal variability and source apportionment of non-methane VOCs using PTR-TOF-MS measurements in Delhi, India, *Atmos. Environ.* 283, 119163, <https://doi.org/10.1016/j.atmosenv.2022.119163>, 2022.

Jain, S., Sharma, S. K., Vijayan, N., and Mandal, T. K.: Seasonal characteristics of aerosols (PM_{2.5} and PM₁₀) and their source apportionment using PMF: A four-year study over Delhi, India. *Environ. Pollut.*, 262, <https://doi.org/10.1016/j.envpol.2020.114337>, 2020.

Jaiprakash, Singhai, A., Habib, G., Sunder Raman, R. and Gupta, T.: Chemical characterization of PM_{1.0} aerosol in Delhi and source apportionment using positive matrix factorization. *Environ. Sci. Pollut. R.*, 24(45–462), <https://doi.org/10.1007/s11356-016-7708-8>, 2016.

Jankowski, M. J., Olsen, R., Thomassen, Y., & Molander, P.: Comparison of air samplers for determination of isocyanic acid and applicability for work environment exposure assessment. *Environm. Sci-Proc. Imp.*, 19(8), 1075-1085, <https://doi.org/10.1039/C7EM00174F>, 2017.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, *Science*, 326, 1525-1529, <https://doi.org/10.1126/science.1180353>, 2009.

Khan, A. A., Garsa, K., Jindal, P., and Devara, P. C. S.: Effects of stubble burning and firecrackers on the air quality of Delhi. *Environ. Monit Assess.*, 195(10), 1170, <https://doi.org/10.1007/s10661-023-11635-6>, 2023.

Khare, P., Machesky, J., Soto, R. He, M., Presto, A. A. and Gentner, D. R.: Asphalt-related emissions are a major missing nontraditional source of secondary organic aerosol precursors. *Science Advances*, 6, eabb9785, <https://doi.org/10.1126/sciadv.abb9785>, 2020.

Kılıç, D., El Haddad, I., Brem, B. T., Bruns, E., Bozetti, C., Corbin, J., Durdina, L., Huang, R.-J., Jiang, J., Klein, F., Lavi, A., Pieber, S. M., Rindlisbacher, T., Rudich, Y., Slowik, J. G., Wang, J., Baltensperger, U., and Prévôt, A. S. H.: Identification of secondary aerosol precursors emitted by an aircraft turbofan. *Atmos. Chem. Phys.*, 18, 7379–7391, <https://doi.org/10.5194/acp-18-7379-2018>, 2018.

Kim, S., Karl, T., Guenther, A., Tyndall, G., Orlando, J., Harley, P., Rasmussen, R., and Apel, E.: Emissions and ambient distributions of Biogenic Volatile Organic Compounds (BVOC) in a ponderosa pine ecosystem: interpretation of PTR-MS mass spectra. *Atmos. Chem. Phys.* 10, 1759–1771, <https://doi.org/10.5194/acp-10-1759-2010>, 2010.

Koss, A. R., Sekimoto, K., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., Jimenez, J. L., Krechmer, J., Roberts, J. M., Warneke, C., Yokelson, R. J., and De Gouw, J.: Non-methane organic gas emissions from biomass burning: Identification, quantification, and emission factors from PTR-ToF during the FIREX 2016 laboratory experiment. *Atmos. Chem. Phys.* 18(5), 3299–3319, <https://doi.org/10.5194/ACP-18-3299-2018>, 2018.

Kulkarni, S. H., Ghude, S. D., Jena, C., Karumuri, R. K., Sinha, B., Sinha, V., Kumar, R., Soni, V. K., and Khare M.: How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi? *Environ. Sci. Technol.* 2020, 54, 4790–4799. <https://doi.org/10.1021/acs.est.0c00329>, 2020.

Krelling, C., and Badami, M.: Cost-effectiveness analysis of compressed natural gas implementation in the public bus transit fleet in Delhi, India. *Transport Policy* 115, 49-61, <https://doi.org/10.1016/j.tranpol.2021.10.019>, 2022.

Kumar, V., Sarkar, C., and Sinha, V.: Influence of post-harvest crop residue fires on surface ozone mixing ratios in the NW IGP analyzed using 2 years of continuous in situ trace gas measurements, *J. Geophys. Res.-Atmos.*, 121, 3619–3633, <https://doi.org/10.1002/2015JD024308>, 2016.

Kumar, A., Sinha, V., Shabin, M., Hakkim, H., Bonsang, B., and Gros, V.: Non-methane hydrocarbon (NMHC) fingerprints of major urban and agricultural emission sources for use in source apportionment studies. *Atmos. Chem. Phys.* 20(20), 12133–12152, <https://doi.org/10.5194/ACP-20-12133-2020>, 2020.

Kumar, A., Hakkim, H., Sinha, B., and Sinha, V.: Gridded 1 km × 1 km emission inventory for paddy stubble burning emissions over north-west India constrained by measured emission factors of 77 VOCs and district-wise crop yield data. *Sci. Total. Environ.*, 789, 148064, <https://doi.org/10.1016/J.SCITOTENV.2021.148064>, 2021.

Kumar, V., Giannoukos, S., Haslett, S. L., Tong, Y., Singh, A., Bertrand, A., Lee, C. P., Wang, D. S., Bhattu, D., Stefanelli, G., Dave, J. S., Puthussery, J. V., Qi, L., Vats, P., Rai, P., Casotto, R., Satish, R., Mishra, S., Pospisilova, V., ... Slowik, J. G.: Highly time-resolved chemical speciation and source apportionment of organic aerosol components in Delhi, India, using extractive electrospray ionization mass spectrometry *Atmos. Chem. Phys.* 22(11), 7739–7761, <https://doi.org/10.5194/acp-22-7739-2022>, 2022.

Kurokawa, J., and Ohara, T.: Long-term historical trends in air pollutant emissions in Asia: Regional Emission inventory in ASia (REAS) version 3. *Atmos. Chem. Phys.* 20(21), 12761–12793, <https://doi.org/10.5194/ACP-20-12761-2020>, 2020.

Li, J., Hao, Y., Simayi, M., Shi, Y., Xi, Z., and Xie, S.: Verification of anthropogenic VOC emission inventory through ambient measurements and satellite retrievals, *Atmos. Chem. Phys.*, 19, 5905–5921, <https://doi.org/10.5194/acp-19-5905-2019>, 2019.

Li, N., Jiang, Q., Wang, F., Xie, J., Li, Y., Li, J., and Wu, S.: Emission behaviour, environmental impact and priority-controlled pollutants assessment of volatile organic compounds (VOCs) during asphalt pavement construction based on laboratory experiment. *J. Hazard. Mater.* 398, 122904, <https://doi.org/10.1016/j.jhazmat.2020.122904>, 2020.

Li, Y., Pöschl, U., and Shiraiwa, M.: Molecular corridors and parameterizations of volatility in the chemical evolution of organic aerosols. *Atmos. Chem. Phys.* 16, 3327–3344, <https://doi.org/10.5194/acp-16-3327-2016>, 2016.

Lignell, H., Epstein, S. A., Marvin, M. R., Shemesh, D., Gerber, B., and Nizkorodov, S.: Experimental and Theoretical Study of Aqueous cis-Pinonic Acid Photolysis. *J. Phys. Chem. A* 2013, 117, 12930–12945, <https://doi.org/10.1021/jp4093018>, 2013.

Liu, T., Marlier, M. E., Karambelas, A., Jain, M., Singh, S., Singh, M. K., Gautam, R., and DeFries, R. S.: Missing emissions from post-monsoon agricultural fires in northwestern India: regional limitations of MODIS burned area and active fire products. *Environ. Res. Communications* 1, 011007, <https://doi.org/10.1088/2515-7620/ab056c>, 2019.

Liu, T., Mickley, L. J., Singh, S., Jain, M., DeFries, R. S., & Marlier, M. E.: Crop residue burning practices across north India inferred from household survey data: Bridging gaps in satellite observations. *Atmos. Environ.*: X, 8, 100091, <https://doi.org/10.1016/j.aeaoa.2020.100091>, 2020

Liu, Y., Chen, H., Gao, J., Li, Y., Dave, K., Chen, J., Federici, M., and Perricone, G.: Comparative analysis of non-exhaust airborne particles from electric and internal combustion engine vehicles. *J. Hazard. Mater.* 420, 126626, <https://doi.org/10.1016/j.jhazmat.2021.126626>, 2021.

Loubet, B., Buysse, P., Gonzaga-Gomez, L., Lafouge, F., Ciuraru, R., Decuq, C., Kammer, J., Bsaibes, S., Boissard, C., Durand, B., Gueudet, J.-C., Fanucci, O., Zurfluh, O., Abis, L., Zannoni, N., Truong, F., Baisnée, D., Sarda-Estève, R., Staudt, M., and Gros, V.: Volatile organic compound fluxes over a winter wheat field by PTR-

Qi-TOF-MS and eddy covariance, *Atmos. Chem. Phys.* 22, 2817–2842, <https://doi.org/10.5194/acp-22-2817-2022>, 2022.

Lyman WJ.: Estimation of physical properties. In: Neely WB, Blau GE, editors. *Environmental Exposure From Chemicals*. Boca Raton, FL: CRC Press. p. 13–47, <https://doi.org/10.1201/9781351071789>, 1985.

Lyman, W. J., Reehl, W. F., and Rosenblatt, D. H.: *Handbook of chemical property estimation methods*, United States. Available at: <https://www.osti.gov/biblio/6942967>, 1990.

Mahilang, M., Deb, M. K., Pervez, S., Tiwari, S., and Jain, V. K.: Biogenic secondary organic aerosol formation in an urban area of eastern central India: Seasonal variation, size distribution and source characterization, *Environ. Res.* 195, 110802, <https://doi.org/10.1016/j.envres.2021.110802>, 2021.

Mishra, S., Tripathi, S. N., Kanawade, V. P., Haslett, S. L., Dada, L., Ciarelli, G., Kumar, V., Singh, A., Bhattu, D., Rastogi, N., Daellenbach, K. R., Ganguly, D., Gargava, P., Slowik, J. G., Kulmala, M., Mohr, C., El-Haddad, I., and Prevot, A. S. H.: Rapid night-time nanoparticle growth in Delhi driven by biomass-burning emissions. *Nat. Geosci.* 16(3), 224–230, <https://doi.org/10.1038/s41561-023-01138-x>, 2023.

Mishra, S., Sinha, V., Hakkim, H., Awasthi, A., Ghude, S. D., Soni, V K., Nigam, N., Sinha, B., and Rajeevan M.: Biomass burning sources control ambient particulate matter but traffic and industrial sources control VOCs and secondary pollutant formation during extreme pollution events in Delhi, *Atmos. Chem. Phys. Discuss.*, (submitted), 2024.

Mochizuki, T., Kawamura, K., Miyazaki, Y., Kunwar, B., and Boreddy, S. K. R.: Distributions and sources of low-molecular-weight monocarboxylic acids in gas and particles from a deciduous broadleaf forest in northern Japan. *Atmos. Chem. Phys.*, 19(4), 2421–2432, <https://doi.org/10.5194/acp-19-2421-2019>, 2019.

Morino, Y., Ohara, T., Yokouchi, Y., and Ooki, A.: Comprehensive source apportionment of volatile organic compounds using observational data, two receptor models, and an emission inventory in Tokyo metropolitan area, *J. Geophys. Res-Atmos.* 116(D02311), <https://doi.org/10.1029/2010JD014762>, 2011.

Nagar, P.K., Singh, D., Sharma, M., Kumar, A., Aneja, V. P., George, M. P., Agarwal, N., Shukla, S. P.: Characterization of PM_{2.5} in Delhi: role and impact of secondary aerosol, burning of biomass, and municipal solid waste and crustal matter. *Environ. Sci. Pollut. Res.* 24:25179–25189, <https://doi.org/10.1007/s11356-017-0171-3>, 2017.

Nagpure, A. S., Ramaswami, A., and Russell, A.: Characterizing the spatial and temporal patterns of open burning of municipal solid waste (MSW) in Indian cities. *Environ. Sci. Technol.* 49, 12904–12912, <https://doi.org/10.1021/acs.est.5b03243>, 2015.

Nagpure, A. S., Gurjar, B. R., Kumar, V., Kumar, P.: Estimation of exhaust and non-exhaust gaseous, particulate matter and air toxics emissions from on-road vehicles in Delhi. *Atmos. Environ.* 127, 118–124, <https://doi.org/10.1016/j.atmosenv.2015.12.026>, 2016.

Norris, G., Duvall, R., Brown, S., and Bai, S.: EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide, available at: https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf (last access: 31 October 2019), 2014.

Nowakowska, M., Herbinet, O., Dufour, A., and Glaude, P. A.: Kinetic Study of the Pyrolysis and Oxidation of Guaiacol. *J. Phys. Chem. A*, 122(39), 7894–7909, <https://doi.org/10.1021/acs.jpca.8b06301>, 2018.

Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions. *Atmos. Meas. Tech.* 7(3), 781–797, <https://doi.org/10.5194/AMT-7-781-2014>, 2014.

Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models. *Chemometr. Intell. Lab.* 60(1–2), 253–264, [https://doi.org/10.1016/S0169-7439\(01\)00200-3](https://doi.org/10.1016/S0169-7439(01)00200-3), 2002.

Paatero, P., and Hopke, P. K.: Rotational tools for factor analytic models. *J. Chemometrics*, 23(2), 91–100, <https://doi.org/10.1002/CEM.1197>, 2009.

Pagonis, D., Sekimoto, K., and de Gouw, J.: A Library of Proton-Transfer Reactions of H₃O⁺ Ions Used for Trace Gas Detection. *J. Am. Soc. Mass Spectrom.* 30(7), 1330–1335, https://doi.org/10.1007/S13361-019-02209-3/SUPPL_FILE/JS8B06050_SI_001.XLSX.

Pallavi, Sinha, B., and Sinha, V.: Source apportionment of volatile organic compounds in the northwest Indo-Gangetic Plain using a positive matrix factorization model. *Atmos. Chem. Phys.*, 19(24), 15467–15482, <https://doi.org/10.5194/ACP-19-15467-2019>, 2019.

Palm, B. B., Peng, Q. Y., Fredrickson, C. D., Lee, B., Garofalo, L. A., Pothier, M. A., Kreidenweis, S. M., Farmer, D. K., Pokhrel, R. P., Shen, Y. J., Murphy, S. M., Permar, W., Hu, L., Campos, T. L., Hall, S. R., Ullmann, K., Zhang, X., Flocke, F., Fischer, E. V., and Thornton, J. A.: Quantification of organic aerosol and brown carbon evolution in fresh wildfire plumes, *P. Natl. Acad. Sci. USA*, 117, 29469–29477, <https://doi.org/10.1073/pnas.2012218117>, 2020.

Pandey, A., Brauer, M., Cropper, M. L., Balakrishnan, K., Mathur, P., Dey, S., Turkoglu, B., Kumar, G. A., Khare, M., Beig, G., Gupta, T., Krishnankutty, R. P., Causey, K., Cohen, A. J., Bhargava, S., Aggarwal, A. N., Agrawal, A., Awasthi, S., Bennitt, F., ... Dandona, L.: Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019. *The Lancet Planetary Health*, 5(1), e25–e38, [https://doi.org/10.1016/S2542-5196\(20\)30298-9](https://doi.org/10.1016/S2542-5196(20)30298-9), 2021.

Park, J.-H., Goldstein, A. H., Timkovsky, J., Fares, S., Weber, R., Karlik, J., and Holzinger, R.: Eddy covariance emission and deposition flux measurements using proton transfer reaction – time of flight – mass spectrometry (PTR-TOF-MS): comparison with PTR-MS measured vertical gradients and fluxes. *Atmos. Chem. Phys.* 13, 1439–1456, <https://doi.org/10.5194/acp-13-1439-2013>, 2013.

Pawar, H., Garg, S., Kumar, V., Sachan, H., Arya, R., Sarkar, C., Chandra, B. P. and Sinha, B.: Quantifying the contribution of long-range transport to Particulate Matter (PM) mass loadings at a suburban site in the North-Western Indo Gangetic Plain (IGP). *Atmos. Chem. Phys.* 15, 9501–9520, 2015, <https://doi.org/10.5194/ACP-15-9501-2015>, 2015.

Pawar, H., and Sinha, B.: Residential heating emissions (can) exceed paddy-residue burning emissions in rural northwest India. *Atmos. Environ.* 269, 118846, <https://doi.org/10.1016/J.ATMOSENV.2021.118846>, 2022.

Portillo-Estrada, M., Kazantsev, T., Talts, E., Tosens, T., and Niinemets, Ü.: Emission Timetable and Quantitative Patterns of Wound-Induced Volatiles Across Different Leaf Damage Treatments in Aspen (*Populus Tremula*). *J Chem Ecol* 41:1105–1117, <https://doi.org/10.1007/s10886-015-0646-y>, 2015.

Piel, F., Müller, M., Winkler, K., Skytte af Sætra, J., and Wisthaler, A.: Introducing the extended volatility range proton-transfer-reaction mass spectrometer (EVR PTR-MS), *Atmos. Meas. Tech.*, 14, 1355–1363, <https://doi.org/10.5194/amt-14-1355-2021>, 2021.

Prakash, J., Choudhary, S., Raliya, R., Chadha, T. S., Fang, J., and Biswas, P.: Real-time source apportionment of fine particle inorganic and organic constituents at an urban site in Delhi city: An IoT-based approach. *Atmos. Pollut. Res.* 12(11), <https://doi.org/10.1016/j.apr.2021.101206>, 2021.

Qin, G., Gao, S., Fu, Q., Fu, S., Jia, H., Zeng, Q., Fan, L., Ren, H., and Cheng, J.: Investigation of VOC characteristics, source analysis, and chemical conversions in a typical petrochemical area through 1-year monitoring and emission inventory. *Environ Sci Pollut Res* 29, 51635–51650, <https://doi.org/10.1007/s11356-022-19145-7>, 2022.

Ramasamy, S., Nakayama, T., Imamura, T., Morino, Y., Kajii, Y., and Sato, K.: Investigation of dark condition nitrate radical- and ozone-initiated aging of toluene secondary organic aerosol: Importance of nitrate radical reactions with phenolic products. *Atmos. Environ.* 219, 117049, <https://doi.org/10.1016/j.atmosenv.2019.117049>, 2019.

Reyes-Villegas, E., Panda, U., Darbyshire, E., Cash, J. M., Joshi, R., Langford, B., Di Marco, C. F., Mullinger, N. J., Alam, M. S., Crilley, L. R., Rooney, D. J., Acton, W. J. F., Drysdale, W., Nemitz, E., Flynn, M., Voliotis, A., McFiggans, G., Coe, H., Lee, J., Hewitt, C. N., Heal, M. R., Gunthe, S. S., Mandal, T. K., Gurjar, B. R., Shivani, Gadi, R., Singh, S., Soni, V., and Allan, J. D.: PM1 composition and source apportionment at two sites in Delhi, India, across multiple seasons. *Atmos. Chem. Phys.* 21, 11655–11667, <https://doi.org/10.5194/acp-21-11655-2021>, 2021.

Rizzo, M. J., and Scheff, P. A.: Utilizing the Chemical Mass Balance and Positive Matrix Factorization models to determine influential species and examine possible rotations in receptor modeling results. *Atmos. Environ.* 41(33), 6986–6998, <https://doi.org/10.1016/J.ATMOSENV.2007.05.008>, 2007.

Rolph, G., Stein, A., and Stunder, B.: Real-time Environmental Applications and Display sYstem: READY. *Environmental Modelling and Software* 95, 210-228, <https://doi.org/10.1016/j.envsoft.2017.06.025>, 2017.

Roozitalab, B., Carmichael, G.R., and Guttikunda, S.K.: Improving regional air quality predictions in the Indo-Gangetic Plain – case study of an intensive pollution episode in November 2017. *Atmos. Chem. Phys.* 21, 2837–2860. <https://doi.org/10.5194/acp-21-2837-2021>, 2021.

Sarkar, C., Sinha, V., Kumar, V., Rupakheti, M., Panday, A., Mahata, K. S., Rupakheti, D., Kathayat, B., and Lawrence, M. G.: Overview of VOC emissions and chemistry from PTR-TOFMS measurements during the SusKat-ABC campaign: high acetaldehyde, isoprene and isocyanic acid in wintertime air of the Kathmandu Valley. *Atmos. Chem. Phys.* 16, 3979–4003, <https://doi.org/10.5194/acp-16-3979-2016>, 2016.

Sarkar, C., Sinha, V., Sinha, B., Panday, A. K., Rupakheti, M., and Lawrence, M. G.: Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization. *Atmos. Chem. Phys.* 17(13), 8129–8156, <https://doi.org/10.5194/ACP-17-8129-2017>, 2017.

Sharma, S. K., Sharma, A., Saxena, M., Choudhary, N., Masiwal, R., Mandal, T. K., and Sharma, C.: Chemical characterization and source apportionment of aerosol at an urban area of Central Delhi, India. *Atmos. Pollut. Res.* 7(1), 110–121, <https://doi.org/10.1016/j.apr.2015.08.002>, 2016.

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

Sharma, G., Annadate, S., and Sinha, B.: Will open waste burning become India's largest air pollution source? *Environ. Pollut.* 292, 118310, <https://doi.org/10.1016/j.envpol.2021.118310>, 2022.

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., and Rastogi, N.: 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>, 2023.

[Singh, R., Sinha, B., Hakkim, H., and Sinha, V.: Source apportionment of volatile organic compounds during paddy-residue burning season in north-west India reveals a large pool of photochemically formed air toxics. *Environ. Pollut.* 338, 122656, <https://doi.org/10.1016/j.envpol.2023.122656>, 2023.](#)

Singh, S., Mishra, S., Mathai, R., Sehgal, A. K., & Suresh, R.: Comparative study of unregulated emissions on a heavy duty CNG engine using CNG & hydrogen blended CNG as fuels. *SAE Int. J. Engines*, 9(4), 2292-2300, <http://dx.doi.org/10.4271/2016-01-8090>, 2016.

Sinha, V., Kumar, V., and Sarkar, C.: 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. *Atmos. Chem. Phys.* 14(12), 5921–5941, <https://doi.org/10.5194/ACP-14-5921-2014>, 2014.

Srivastava, A., Gupta, S., and Jain, V. K.: Source Apportionment of Total Suspended Particulate Matter in Coarse and Fine Size Ranges Over Delhi. *Aerosol Air Qual. Res.* 8, 188-200, <https://doi.org/10.4209/aaqr.2007.09.0040>, 2008.

Stein, A.F., Draxler, R.R, Rolph, G.D., Stunder, B.J.B., Cohen, M.D., and Ngan, F.: NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Amer. Meteor. Soc.*, 96, 2059-2077, <https://doi.org/10.1175/BAMS-D-14-00110.1>, 2015

Stockwell, C. E., Veres, P.R., Williams, J. and Yokelson, R.J.: Characterization of biomass burning emissions from cooking fires, peat, crop residue, and other fuels with high-resolution proton-transfer-reaction time-of-flight mass spectrometry. *Atmos. Chem. Phys.* 15, 845–865, <https://doi.org/10.5194/acp-15-845-2015>, 2015.

Thakur, M.: Low-smoke chulha and health consequences in Indian slums. [Doctoral Thesis, Maastricht University]. Maastricht University, <https://doi.org/10.26481/dis.20230403mt>, 2023.

Villberg, K., & Veijanen, A.: Analysis of a GC/MS thermal desorption system with simultaneous sniffing for determination of off-odor compounds and VOCs in fumes formed during extrusion coating of low-density polyethylene. *Anal. Chem.* 73(5), 971-977. <https://doi.org/10.1021/ac001114w>, 2001.

Wang, M., Wang, Q., Ho, S. S. H., Li, H., Zhang, R., Ran, W., Que, L.: Chemical characteristics and sources of nitrogen-containing organic compounds at a regional site in the North China Plain during the transition period of autumn and winter. *Sci. Total. Environ.* 812, 151451, <https://doi.org/10.1016/j.scitotenv.2021.151451>, 2022.

Wiedinmyer, C., Kimura, Y., McDonald-Buller, E. C., Emmons, L. K., Buchholz, R. R., Tang, W., Seto, K., Joseph, M. B., Barsanti, K. C., Carlton, A. G., and Yokelson, R.: The Fire Inventory from NCAR version 2.5: an updated global fire emissions model for climate and chemistry applications, *Geosci. Model Dev.*, 16, 3873–3891, <https://doi.org/10.5194/gmd-16-3873-2023>, 2023.

Xu, C., Gao, L., Lyu, C., Qiao, L. Huang, D., Liu, Y., Li, D. and Zheng, M.: Molecular characteristics, sources and environmental risk of aromatic compounds in particulate matter during COVID-2019: Nontarget screening by ultra-high resolution mass spectrometry and comprehensive two-dimensional gas chromatography. *Environ. Int.* 167 107421, <https://doi.org/10.1016/j.envint.2022.107421>, 2022.

Yáñez-Serrano, A. M., Filella, I., LLusà, J., Gargallo-Garriga, A., Granda, V., Bourtsoukidis, E., Williams, J., Seco, R., Cappellin, L., Werner, C., de Gouw, J., and Peñuelas, J.: GLOVOCS - Master compound assignment guide for proton transfer reaction mass spectrometry users. *Atmos. Environ.* 244, 117929, <https://doi.org/10.1016/J.ATMOSENV.2020.117929>, 2021.

Yao, L., Wang, M.-Y., Wang, X. -K. Liu, Y.-J., Chen, H.-F., Zheng, J., Nie, W., Ding, A.-J., Geng, F.-H., Wang, D.-F., Chen, J.-M., Worsnop, D. R. and Wang, L.: Detection of atmospheric gaseous amines and amides by a high-resolution time-of-flight chemical ionization mass spectrometer with protonated ethanol reagent ions. *Atmos. Chem. Phys.*, 16, 14527–14543. <https://doi.org/10.5194/acp-16-14527-2016>, 2016.

You, Y., Kanawade, V. P., de Gouw, J. A., Guenther, A. B., Madronich, S., Sierra-Hernández, M. R., Lawler, M., Smith, J. N., Takahama, S., Ruggeri, G., Koss, A., Olson, K., Baumann, K., Weber, R. J., Nenes, A., Guo, H., Edgerton, E. S., Porcelli, L., Brune, W. H., Goldstein, A. H., and Lee, S.-H.: Atmospheric amines and ammonia measured with a chemical ionization mass spectrometer (CIMS). *Atmos. Chem. Phys.* 14, 12181–12194, <https://doi.org/10.5194/acp-14-12181-2014>, 2014.

Yuan, B., Shao, M., De Gouw, J., Parrish, D. D., Lu, S., Wang, M., Zeng, L., Zhang, Q., Song, Y., Zhang, J., and Hu, M.: Volatile organic compounds (VOCs) in urban air: How chemistry affects the interpretation of positive matrix factorization (PMF) analysis. *J. Geophys. Res-Atmos.* 117(D24), 24302, <https://doi.org/10.1029/2012JD018236>, 2012.

Yuan, B., Koss, A.R., Warneke, C., Coggon, M., Sekimoto, K., and de Gouw, J.: Proton-Transfer-Reaction Mass Spectrometry: Applications in Atmospheric Sciences. *Chem. Rev.* 117 (21), 13187-13229, <https://doi.org/10.1021/acs.chemrev.7b00325>, 2017.

Zaytsev, A., Koss, A. R., Breitenlechner, M., Krechmer, J. M., Nihill, K. j. Lim, C. Y., Rowe, J. C., Cox, J. L., Moss, J., Roscioli, J. R., Canagaratna, M. R., Worsnop, D. R., Kroll, J. H., and Keutsch1, F. N.: Mechanistic Study of Formation of Ring-retaining and Ring-opening Products from Oxidation of Aromatic Compounds under Urban Atmospheric Conditions. *Atmos. Chem. Phys.* 19, 15117–15129, <https://doi.org/10.5194/acp-19-15117-2019>, 2019.

Track changes version of the supplement

Supporting Information for

Biomass burning sources control ambient particulate matter but traffic and industrial sources control volatile organic compound emissions VOCs and secondary pollutant formation during extreme pollution events in Delhi

Arpit Awasthi¹, Baerbel Sinha¹, Haseeb Hakkim¹, Sachin Mishra¹, Varkrishna Mummidivarapu¹, Gurmanjot Singh¹, Sachin D. Ghude², Vijay Kumar Soni³, Narendra Nigam³, Vinayak Sinha¹, Madhavan N. Rajeevan⁴

¹Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research Mohali, Sector 81,

S.A.S Nagar, Manauli PO, Punjab, 140306, India

²Indian Institute of Tropical Meteorology, Pashan, Pune 411 008, Ministry of Earth Sciences, India

³India Meteorological Department, New Delhi 110 003, Ministry of Earth Sciences, India

⁴Ministry of Earth Sciences, Government of India, New Delhi 110 003, India

Correspondence to: Baerbel Sinha (bsinha@iisermohali.ac.in)

Table S1: 111 NMVOCs species used in the PMF model, the table lists the major compound identifications and the references supporting such assignments from previous works, along with mean average of the observational period reported in this study (with range min-max), detection limits, precision error.

Protonated m/z (H⁺)	Potential contributor	Strong /Weak	Mean-average (with range) $\mu\text{g}/\text{m}^3$	Detection limit $\mu\text{g}/\text{m}^3$	Precision error (%)	Sources	References
31.014	Formaldehyde <u>HCH₂O</u>	Strong	1.063 (0.066-9.464)	<u>0.024</u>	<u>3.7</u>	Photochemical production, traffic, biomass burning	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
33.030	Methanol <u>CH₃+OH</u>	Strong	17.285 (3.522-127.853)	<u>0.038</u>	<u>2.1</u>	Photochemical production, biomass burning, biogenic	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
41.035	Propyne <u>C₃H₄</u>	Strong	9.149 (1.001-96.661)	<u>0.041</u>	<u>2.1</u>	Traffic, biomass burning	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Hakkim et al. 2021
42.030	Acetonitrile <u>C₂H₃N</u>	Strong	0.888 (0.208-5.208)	<u>0.003</u>	<u>4.3</u>	Biomass burning, industries	Hatch et al. 2015, 2017
43.051	Propene <u>C₃H₆</u>	Strong	5.117 (0.766-59.931)	<u>0.027</u>	<u>2.8</u>	Biomass burning, traffic	Hakkim et al. 2021
44.018	Isocyanic acid <u>HNCO</u>	Strong	0.145 (0.007-1.231)	<u>0.007</u>	<u>10.5</u>	Oxidation of amines (secondary	Chandra & Sinha, 2016; Wang et al., 2020

						formation), biomass burning	
45.030	Acetaldehyde C_2H_3CHO	Strong	8.764 (1.432-55.175)	<u>0.059</u>	<u>1.5</u>	Biogenic, biomass burning, photochemical production	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Kumar et al. 2021
46.025	Formamide CH_2NO	Strong	0.447 (0.004-8.3)	<u>0.012</u>	<u>6</u>	Oxidation of amines (secondary formation)	Yao et al. 2016; Wang et al. 2022
47.009	Formic acid $HCHO_2$	Strong	1.717 (0.004-31.063)	<u>0.055</u>	<u>4.2</u>	Oxidation of amines (secondary formation)	Yao et al. 2016; Wang et al. 2022
47.0457	Ethanol C_2H_6O	Strong	0.625 (0.052-7.845)	<u>0.002</u>	<u>5.8</u>	Industrial, Traffic	Bruns et al. 2017; Koss et al. 2018
48.048	methoxyamine <u>Methoxyamine</u> CH_3NO	Weak	0.011 (0-0.148)	<u>0.001</u>	<u>39.6</u>	Oxidation of amines (secondary formation)	Yáñez-Serrano et al. 2021
49.007	Methanethiol CH_4S	Strong	0.15 (0.002-3.295)	<u>0.002</u>	<u>15.4</u>	Industrial	Toda et al. 2010
53.035	Vinylacetylene, 1-Buten- 3-yne C_4H_4	Strong	0.662 (0.003-10.879)	<u>0.007</u>	<u>6.3</u>	Biomass burning	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
55.051	1,2-Butadiene, 1-Butyne, 2-Butyne, 1,3 Butadiene C_4H_6	Strong	3.701 (0.418-30.204)	<u>0.029</u>	<u>3.5</u>	Biomass burning	Koss et al., 2018, Sarkar et al., 2017
57.030	Acrolein C_3H_4O	Strong	0.802 (0-8.038)	<u>0.012</u>	<u>5.7</u>	Biomass burning, waste burning	Kumar et al. 2021
57.067	Methyl tert-butyl ether (MTBE) fragment /Butene C_4H_8	Strong	7.293 (0.983- 100.664)	<u>0.024</u>	<u>2.5</u>	Biomass burning, waste burning, traffic	Hakkim et al. 2021
59.046	Acetone + Propanal C_3H_6O	Strong	14.603 (2.341- 145.458)	<u>0.016</u>	<u>2.4</u>	Biomass burning, industries, photochemical production	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Kumar et al. 2021
61.025	Acetic acid+ Glycolaldehyde $C_2H_4O_2$	Strong	16.086 (0.695- 170.723)	<u>0.033</u>	<u>1.9</u>	Photochemical production, biomass burning, industries	Kumar et al. 2021
62.997	Vinyl chloride C_2H_3Cl	Weak	0.02 (0.001-0.235)	<u>0.001</u>	<u>37.1</u>	PVC burning	Hsu et al. 2022, Fukusaki et al. 2021
67.051	Cyclopentadiene, monoterpene fragment, butanol fragment C_5H_6	Strong	0.781 (0.117-10.93)	<u>0.01</u>	<u>8.2</u>	Biomass burning	Hatch et al. 2015; Stockwell et al., 2015
69.031	Furan C_4H_4O	Strong	0.231 (0.011-3.247)	<u>0.006</u>	<u>11.5</u>	Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
69.067	Isoprene + 2-methyl-3- butene-2-ol fragment C_5H_8	Strong	2.234 (0.294-17.733)	<u>0.014</u>	<u>4.9</u>	Biogenic sources, biomass burning	Stockwell et al., 2015; Jordan et al., 2009
71.047	Methyl Vinyl Ketone, Methacrolein, Butenal C_4H_6O	Strong	1.015 (0.099-4.729)	<u>0.013</u>	<u>5.6</u>	Biomass burning, photochemical production	Stockwell et al., 2015; de Gouw et al., 2007
73.026	Methyl-glyoxal $C_3H_4O_2$	Strong	0.56 (0.004-8.797)	<u>0.024</u>	<u>8.6</u>	Oxidation of amines (secondary formation)	Yao et al. 2016; Wang et al. 2022

73.062	Butanal, b Butanone, MEK C ₄ H ₈ O	Strong	2.534 (0.315-43.128)	<u>0.006</u>	<u>3.6</u>	Biomass burning, biogenic, photochemical production	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
75.042	Hydroxyacetone, p Propanoic acid C ₃ H ₆ O ₂	Strong	1.664 (0.116-14.045)	<u>0.005</u>	<u>8.5</u>	Biomass burning, biogenic, traffic	Kumar et al. 2021
76.037	Nitroethane C ₂ H ₅ NO ₂	Strong	0.036 (0.002-0.308)	<u>0.013</u>	<u>5.4</u>	Biomass burning	Palm et al. 2020, Harrison et al. 2005
79.052	Benzene C ₆ H ₆	Strong	6.07 (0.305-68.694)	<u>0.007</u>	<u>1.7</u>	Biomass burning, traffic	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Hakkim et al. 2021
81.031	Cyclopentadienone C ₅ H ₄ O	Strong	0.102 (0.001-0.896)	<u>0.004</u>	<u>17.3</u>	Biomass burning	Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
83.047	Methyl furan C ₅ H ₆ O	Strong	0.363 (0.032-3.68)	<u>0.018</u>	<u>10.2</u>	Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
83.084	Cyclohexene, h Hexyne isomers C ₆ H ₁₀	Strong	1.406 (0.345-12.044)	<u>0.007</u>	<u>6.6</u>	Biomass burning, biogenic, traffic	Stockwell et al., 2015, Koss et al., 2018; Kumar et al., 2021;
84.080	Pentanenitrile, m Methylbutanenitrile isomers, C5 -amines C ₅ H ₉ N	Strong	0.037 (0-0.495)	<u>0.011</u>	<u>28.7</u>	Biogenic, biomass burning	Hatch et al. 2015, 2017
85.027	Furanone, butenedial C ₄ H ₄ O ₂	Strong	0.466 (0.015-6.719)	<u>0.002</u>	<u>6.4</u>	Biomass burning, photochemical production	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
85.063	Cyclopentanone C ₅ H ₈ O	Strong	0.337 (0.042-1.971)	<u>0.025</u>	<u>10.4</u>	Biomass burning, traffic	Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
85.094	Cyclohexane, Hexene C ₆ H ₁₂	Strong	0.326 (0.051-4.057)	<u>0.009</u>	<u>12.7</u>	Biomass burning, Traffic	Fleming et al. 2018, Hakkim et al. 2021
87.043	Isomers of C4 carboxylic acid/ester/diketone C ₄ H ₆ O ₂	Strong	1.118 (0.06-8.545)	<u>0.006</u>	<u>6.7</u>	Biomass burning, photochemical production	Kumar et al. 2021, Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
87.079	Pentanone, methyl-buteneol, p Pentanal C ₅ H ₁₀ O	Strong	0.295 (0.05-1.841)	<u>0.023</u>	<u>11.6</u>	Biomass burning, biogenic	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
89.058	Isomers of C4-carboxylic acid/ester C ₄ H ₈ O ₂	Strong	1.306 (0.087-17.301)	<u>0.003</u>	<u>6.1</u>	Industrial solvent	Kamarulzaman et al. 2019
91.053	Monoterpene f Fragment C ₇ H ₆	Strong	1.24 (0.058-20.618)	<u>0.008</u>	<u>5.9</u>	Industrial, Traffic	Kamarulzaman et al. 2019
93.069	Toluene C ₇ H ₈	Strong	18.285 (0.43-321.651)	<u>0.004</u>	<u>1.7</u>	Biomass burning traffic, chemical production, biogenic	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Hakkim et al. 2021
95.048	Phenol C ₆ H ₆ O	Strong	0.835 (0.069-9.968)	<u>0.023</u>	<u>6.1</u>	Biomass burning, photochemical production	Hakkim et al., 2021; Koss et al., 2018; Kumar et al., 2021

95.084	Monoterpene fragment C ₇ H ₁₀	Strong	0.826 (0.111-13.348)	<u>0.008</u>	<u>9</u>	Industrial, traffic	Kamarulzaman et al. 2019
97.027	Furfural isomers of diketone/ carboxylic acid / ester C ₅ H ₄ O ₂	Strong	0.649 (0.026-17.29)	<u>0.013</u>	<u>9.5</u>	Biomass burning	Kumar et al. 2021, Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
97.063	C2 substituted furan, methyl Cyclopentenone C ₆ H ₈ O	Strong	0.277 (0.024-2.555)	<u>0.008</u>	<u>12.2</u>	Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
97.100	Cycloheptene, Alkyl fragment C ₇ H ₁₂	Strong	0.634 (0.104-7.483)	<u>0.011</u>	<u>10.1</u>	Biomass burning	Stockwell et al., 2015; Koss et al., 2018; Kumar et al., 2021
99.043	Furfuryl alcohol, Methyl- furanone C ₅ H ₆ O ₂	Strong	0.599 (0.025-4.61)	<u>0.009</u>	<u>9.5</u>	Photochemical production, Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
99.079	Cyclohexanone isomers of C6-aldehyde/ketone C ₆ H ₁₀ O	Strong	1.033 (0.088-23.49)	<u>0.022</u>	<u>8.3</u>	Industrial	Gupta et al. 1979
99.116	Methylcyclohexane, Heptene & other hydrocarbons C ₇ H ₁₄	Weak	0.034 (0.002-0.472)	<u>0.012</u>	<u>42.6</u>	asphalt degassing	Khare et al., 2020
101.059	Pentanedione isomers of C5-diketone/ carboxylic acid/ ester/ aldehyde C ₅ H ₈ O ₂	Strong	0.777 (0.048-5.073)	<u>0.002</u>	<u>8.5</u>	Biomass burning	Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
105.069	Styrene C ₈ H ₈	Strong	1.556 (0.13-35.843)	<u>0.021</u>	<u>7</u>	Biomass burning traffic, chemical production	Jordan et al., 2009; Stockwell et al., 2015
107.050	Benzaldehyde isomers of C7-aldehyde/ ketone C ₇ H ₆ O	Strong	0.484 (0.003-4.687)	<u>0.004</u>	<u>9.7</u>	Biomass burning, photochemical production	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
107.085	Sum of C8-Aromatics C ₈ H ₁₀	Strong	11.214 (0.4-193.065)	<u>0.004</u>	<u>2.6</u>	Biomass burning, traffic	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018, Hakkim et al. 2021
109.064	Methylphenol isomers, Anisole C ₇ H ₈ O	Strong	0.24 (0.021-2.61)	<u>0.005</u>	<u>14.1</u>	Biomass Burning	Hatch et al. 2015
109.100	Terpene fragment/Cyclooctadiene C ₈ H ₁₂	Strong	0.511 (0.089-5.681)	<u>0.006</u>	<u>11.8</u>	Industrial, traffic	Kamarulzaman et al. 2019
111.042	Methylfurfural, Hydroxyphenol C ₆ H ₆ O ₂	Strong	0.208 (0.002-4.11)	<u>0.009</u>	<u>16</u>	Biomass burning, photochemical production	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
111.080	C3-substituted furans, C2-substituted cyclopentene, methyl cyclohexene C ₇ H ₁₀ O	Strong	0.195 (0.024-1.519)	<u>0.016</u>	<u>15.5</u>	Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
111.116	Ethenyl cyclohexane C ₈ H ₁₄	Strong	0.434 (0.078-5.193)	<u>0.005</u>	<u>12.6</u>	Ring-opening products of cyclic alkanes	Wang et al. 2015

113.059	Dimethylbutenedial/ C4-substituted aldehyde C ₆ H ₈ O ₂	Strong	0.403 (0.023-2.736)	<u>0.009</u>	<u>11.8</u>	Oxidation of aromatic compounds, biomass burning	Zaytsev et al., 2019
115.039	Hydroxymethyl furanone/ methylepoxybutanedial C ₅ H ₆ O ₃	Strong	0.119 (0.004-1.797)	<u>0.012</u>	<u>18.8</u>	Photochemical production, Biomass burning	Coggon et al., 2019; Hatch et al. 2015; 2017; Koss et al., 2018; Stockwell et al., 2015
115.075	Isomers of C6-diketones/ aldehyde/ carboxylic acid/ester C ₆ H ₁₀ O ₂	Strong	0.357 (0.024-2.187)	<u>0.013</u>	<u>12.7</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017
116.108	C6-amides C ₆ H ₁₃ NO	Weak	0.035 (0-0.255)	<u>0.012</u>	<u>34.2</u>	Photooxidation of amines	Yao et al. 2016
119.085	Terpene fragment C ₉ H ₁₀	Strong	0.608 (0.062-9.105)	<u>0.001</u>	<u>9.3</u>	Traffic	Erickson et al. 2013
121.064	Tolualdehyde/ isomers of C8-aldehyde/ ketone C ₈ H ₈ O	Strong	0.578 (0.048-6.186)	<u>0.004</u>	<u>9.4</u>	Biomass burning	Hatch et al. 2015, Koss et al. 2018
121.101	Sum of C-9 aromatics C ₉ H ₁₂	Strong	5.67 (0.174- 125.472)	<u>0.004</u>	<u>3.5</u>	Traffic, Biomass burning	Hatch et al. 2015, Stockwell et al. 2015, Koss et al. 2018
123.044	Hydroxybenzaldehyde/is omers of C7- carboxylic acid/ ester C ₇ H ₆ O ₂	Strong	0.307 (0.009-2.77)	<u>0.004</u>	<u>13.6</u>	Photochemical production, Biomass burning	Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
123.080	C2-substituted phenol, methyl anisole C ₈ H ₁₀ O	Strong	0.146 (0.013-1.255)	<u>0.006</u>	<u>18.6</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017
123.12	Santene, Cyclopentadiene & other hydrocarbons C ₉ H ₁₄	Strong	0.349 (0.063-3.73)	<u>0.009</u>	<u>14.6</u>	asphalt degassing	Khare et al., 2020, Kılıç et al., 2018
124.039	Nitrobenzene C ₆ H ₅ NO ₂	Strong	0.054 (0.004-0.871)	<u>0.002</u>	<u>25.3</u>	Traffic	Palm et al. 2020, Harrison et al. 2005
125.060	Guaiacol/ isomers of C7- carboxylic acid/ester C ₇ H ₈ O ₂	Strong	0.162 (0.01-2.15)	<u>0.009</u>	<u>19.2</u>	Biomass burning	Hatch et al. 2015, Koss et al. 2018; Nowakowska et al. 2018
125.133	Nonyne, nondiene C ₉ H ₁₆	Strong	0.157 (0.033-1.702)	<u>0.005</u>	<u>21.4</u>	asphalt degassing	Khare et al., 2020, Kılıç et al., 2018
127.039	Hydroxymethyl furfural C ₆ H ₆ O ₃	Strong	0.116 (0.004-2.045)	<u>0.01</u>	<u>18.9</u>	biomass burning	Koss et al., 2018
127.075	Isomers of C7-carboxylic acid/ ester/ aldehyde/ ketone C ₇ H ₁₀ O ₂	Strong	0.226 (0.014-1.5)	<u>0.013</u>	<u>16.3</u>	Oxidation of aromatic compounds, biomass burning	Zaytesv et al., 2019
129.070	Naphthalene C ₁₀ H ₈	Strong	1.043 (0.099-12.618)	<u>0.012</u>	<u>8.7</u>	Traffic, biomass burning	Hakkim et al., 2021; Koss et al., 2018; Kumar et al., 2021
129.092	Isomers of C9- acetaldehyde/ ketone/ carboxylic acid/ ester C ₇ H ₁₂ O ₂	Strong	0.176 (0.01-1.274)	<u>0.006</u>	<u>18.2</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017; Lignell et al. 2013
133.065	Methyl benzofuran C ₉ H ₈ O	Strong	0.078 (0.007-0.617)	<u>0.003</u>	<u>26.7</u>	Biomass Burning	Hatch et al. 2015
133.102	Ethyl styrene, tetrahydronaphthalene C ₁₀ H ₁₂	Strong	0.457 (0.044-7.886)	<u>0.003</u>	<u>13.4</u>	Traffic	Yáñez-Serrano et al. 2021

135.080	Isomers of C9-acetaldehyde/ ketone C ₉ H ₁₀ O	Strong	0.172 (0.007-1.367)	<u>0.004</u>	<u>17.9</u>	Traffic	Knighon et al. 2007
135.118	p-cymene, C4-substituted benzene, C2-substituted xylene C ₁₀ H ₁₄	Strong	2.509 (0.091-71.697)	<u>0.005</u>	<u>5.7</u>	Traffic	Hakkim et al. 2021
137.133	Sum of Monoterpenes (MT) C ₁₀ H ₁₆	Strong	2.66 (0.167-127.676)	<u>0.006</u>	<u>10.3</u>	Industrial, biogenic, biomass burning, traffic	Guenther et al., 2006; Kamarulzaman et al. 2019; Koss et al., 2018; Kumar et al., 2021
138.056	Nitrotoluene/ salicylamide C ₇ H ₇ NO ₂	Weak	0.036 (0.001-0.707)	<u>0.003</u>	<u>32.9</u>	Oxidation of toluene	Ramasamy et al., 2019
143.086	Methyl naphthalene C ₁₁ H ₁₀	Strong	0.19 (0.015-2.848)	<u>0.004</u>	<u>21.2</u>	Biomass Burning, Traffic	Hatch et al. 2015; Yáñez-Serrano et al. 2021
143.108	Isomers of C8-aldehyde/ketone/carboxylic acid /ester C ₈ H ₁₄ O ₂	Strong	0.153 (0.022-0.942)	<u>0.011</u>	<u>20.7</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017
145.051	Organic acids/ levoglucosan fragment C ₆ H ₈ O ₄	Strong	0.068 (0.002-1.626)	<u>0.009</u>	<u>27.2</u>	Biomass Burning	Hatch et al. 2015; Koss et al. 2018; Nowakowska et al. 2018
145.102	C2-substituted indene C ₁₁ H ₁₂	Weak	0.062 (0.007-0.814)	<u>0.004</u>	<u>36.7</u>	Asphalt degassing	Khare et al., 2020
145.123	Isomer of C8-carboxylic acid/ C8-ester C ₈ H ₁₆ O ₂	Strong	0.072 (0.003-0.96)	<u>0.004</u>	<u>29.3</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017; Mochizuki et al. 2019
146.977	Isomers of dichlorobenzene C ₆ H ₄ Cl ₂	Strong	0.352 (0.009-5.79)	<u>0.001</u>	<u>13.7</u>	Industrial pesticides	Graus et al. 2010; Yáñez-Serrano et al. 2021
147.118	Cyclopentylbenzene & other hydrocarbons C ₁₁ H ₁₄	Strong	0.318 (0.038-4.599)	<u>0.002</u>	<u>16.7</u>	asphalt degassing	Khare et al., 2020
149.024	Phthalic anhydride, Benzofurandione benzofurandione C ₈ H ₄ O ₃	Strong	0.195 (0.003-4.508)	<u>0.007</u>	<u>15.3</u>	Oxidation of polyaromatic hydrocarbons	Bruns et al. 2017
149.096	Isomers of C10-aldehyde/ ketone C ₁₀ H ₁₂ O	Strong	0.094 (0.008-1.25)	<u>0.003</u>	<u>26.4</u>	asphalt degassing	Khare et al., 2020
153.092	Isomers of C9-carboxylic acid/ester C ₉ H ₁₂ O ₂	Strong	0.102 (0.01-1.413)	<u>0.01</u>	<u>25.3</u>	Oxidation of monoterpenes	Gkatzelis et al. 2018
153.128	Isomers of C10-aldehyde/ ketone C ₁₀ H ₁₆ O	Strong	0.31 (0.021-7.93)	<u>0.005</u>	<u>15.2</u>	Oxidation of monoterpenes	Camredon et al.: 2010
154.052	Nitrobenzyl alcohol/ Nitrocresols, methyl-nitrophenol C ₇ H ₇ NO ₃	Strong	0.077 (0.003-0.577)	<u>0.003</u>	<u>27.5</u>	Oxidation of toluene	Ramasamy et al., 2019
155.108	Isomers of C9-ketone/ C9-carboxylic acid/ C9-ester C ₉ H ₁₄ O ₂	Strong	0.104 (0.008-0.759)	<u>0.009</u>	<u>25.5</u>	Oxidation of monoterpenes	Camredon et al.: 2010; Gkatzelis et al. 2018

155.144	Isomers of C10-aldehyde/ ketone C ₁₀ H ₁₈ O	Strong	0.092 (0.008-1.238)	<u>0.005</u>	<u>26.7</u>	asphalt degassing	Khare et al., 2020
157.099	C2-substituted naphthalene C ₁₂ H ₁₂	Strong	0.144 (0.017-1.209)	<u>0.005</u>	<u>23.9</u>	asphalt degassing	Khare et al., 2020; Kılıç et al., 2018
157.122	C9-ester/ C9-organic acid C ₉ H ₁₆ O ₂	Strong	0.109 (0.012-1.035)	<u>0.02</u>	<u>23.6</u>	Oxidation of monoterpenes	Lignell et al. 2013; Camredon et al. 2010
159.140	C9-organic acid C ₉ H ₁₈ O ₂	Weak	0.056 (0.002-0.559)	<u>0.005</u>	<u>32.8</u>	Oxidation of monoterpenes	Mochizuki et al. 2019
161.134	Cyclohexylbenzene, butyl styrene, cyclopentylmethylbenzene C ₁₂ H ₁₆	Strong	0.161 (0.016-2.527)	<u>0.002</u>	<u>24.1</u>	asphalt degassing	Khare et al., 2020; Kılıç et al., 2018
175.150	Trimethyltetralin/ ionene C ₁₃ H ₁₈	Strong	0.086 (0.007-1.677)	<u>0.001</u>	<u>33.6</u>	asphalt degassing	Khare et al., 2020
177.056	Formylcinnamic acid / hydroxy-methyl-coumarin C ₁₀ H ₈ O ₃	Strong	0.105 (0.006-0.852)	<u>0.004</u>	<u>25.9</u>	asphalt degassing	Xing et al. 2023
177.165	C7-substituted benzene, C ₁₃ H ₂₀	Strong	0.13 (0.013-2.594)	<u>0.003</u>	<u>27.8</u>	asphalt degassing	Khare et al., 2020
179.181	C3-substituted adamantane C ₁₃ H ₂₂	Weak	0.06 (0.006-0.878)	<u>0.003</u>	<u>38.5</u>	asphalt degassing	Khare et al., 2020
183.121	Bibenzyl C ₁₄ H ₁₄	Weak	0.029 (0.004-0.2)	<u>0.004</u>	<u>58.5</u>	asphalt degassing	Khare et al., 2020
185.121	cis-Pinonic acid / C10-ester C ₁₀ H ₁₆ O ₃	Weak	0.054 (0.007-1.571)	<u>0.006</u>	<u>35.2</u>	Oxidation of monoterpenes	Camredon et al. 2010; Khare et al. 2020
187.148	C4-substituted dihydroazulene, benzyl cycloheptene C ₁₄ H ₁₈	Weak	0.039 (0.005-0.306)	<u>0.006</u>	<u>47.7</u>	asphalt degassing	Khare et al. 2020, Loubet et al. 2022
189.165	C4-substituted dihydronaphthalene cyclopentylpropyl benzene C ₁₄ H ₂₀	Weak	0.053 (0.006-0.818)	<u>0.002</u>	<u>43.5</u>	asphalt degassing	Khare et al., 2020
191.181	C8-substituted benzene C ₁₄ H ₂₂	Weak	0.068 (0.007-0.892)	<u>0.002</u>	<u>38.1</u>	asphalt degassing	Khare et al., 2020
195.138	Myrtenyl acetate/C12-organic acid/C12-ester C ₁₂ H ₁₈ O ₂	Weak	0.029 (0.003-0.766)	<u>0.005</u>	<u>48.4</u>	Oxidation of biomass burning	Haeri, 2023
217.195	C6-substituted dihydronaphthalene C ₁₆ H ₂₄	Weak	0.026 (0.003-0.307)	<u>0.002</u>	<u>63.1</u>	asphalt degassing	Khare et al., 2020
233.228	C11-substituted benzene C ₁₇ H ₂₈	Weak	0.023 (0.002-0.215)	<u>0.002</u>	<u>67.1</u>	asphalt degassing	Khare et al., 2020, Xu et al. 2022
247.243	C12-substituted benzene C ₁₈ H ₃₀	Weak	0.022 (0.002-0.158)	<u>0.004</u>	<u>66</u>	asphalt degassing	Khare et al., 2020

Table S2: List of Constraints incorporated in the PMF model

Factor	Element	Type
Biogenic	m/z 42.030	Pull Down Maximally
Biogenic	m/z 93.069	Pull Down Maximally
Biogenic	m/z 107.085	Pull Down Maximally
Biogenic	m/z 121.101	Pull Down Maximally
Biogenic	10/16/2022 11:00:00 PM	Pull Down Maximally
Biogenic	10/17/2022 1:00:00 AM	Pull Down Maximally
Biogenic	10/17/2022 12:00:00 AM	Pull Down Maximally
Biogenic	10/17/2022 2:00:00 AM	Pull Down Maximally
Biogenic	10/16/2022 10:00:00 PM	Pull Down Maximally
Biogenic	10/16/2022 9:00:00 PM	Pull Down Maximally
Biogenic	10/21/2022 10:00:00 PM	Pull Down Maximally
Biogenic	10/21/2022 11:00:00 PM	Pull Down Maximally
Biogenic	10/22/2022 12:00:00 AM	Pull Down Maximally
Biogenic	10/21/2022 9:00:00 PM	Pull Down Maximally
Biogenic	10/21/2022 8:00:00 PM	Pull Down Maximally
Biogenic	10/22/2022 6:00:00 PM	Pull Down Maximally
Biogenic	10/23/2022 7:00:00 PM	Pull Down Maximally
Biogenic	10/26/2022 11:00:00 PM	Pull Down Maximally
Photochemical	m/z 42.030	Pull Down Maximally
Photochemical	m/z 93.069	Pull Down Maximally
Photochemical	m/z 107.085	Pull Down Maximally
Photochemical	m/z 121.101	Pull Down Maximally
Photochemical	9/20/2022 2:00:00 AM	Pull Down Maximally
Photochemical	9/20/2022 11:00:00 PM	Pull Down Maximally
Photochemical	9/22/2022 1:00:00 AM	Pull Down Maximally
Photochemical	9/23/2022 10:00:00 PM	Pull Down Maximally
Photochemical	9/24/2022 1:00:00 AM	Pull Down Maximally
Photochemical	10/4/2022 2:00:00 AM	Pull Down Maximally
Photochemical	10/4/2022 4:00:00 AM	Pull Down Maximally
Cooking	10/17/2022 8:00:00 PM	Pull Up Maximally
Solvents	11/8/2022 4:00:00 AM	Pull Up Maximally
Road Construction	10/17/2022 9:00:00 PM	Pull Up Maximally
Petrol 4-Wheeler	10/18/2022 11:00:00 PM	Pull Up Maximally
Paddy	11/12/2022 1:00:00 AM	Pull Up Maximally
Paddy	11/12/2022 12:00:00 AM	Pull Up Maximally
Heating & Waste Disposal	11/19/2022 11:00:00 PM	Pull Up Maximally
Petrol 2-Wheeler	10/17/2022 9:00:00 PM	Pull Up Maximally
Industrial	10/16/2022 11:00:00 PM	Pull Up Maximally
CNG	10/16/2022 7:00:00 PM	Pull Up Maximally

Table S3: Contains the SOA yields for 53 compounds to calculate the SOAP

m/z	SOAP	References
HCHO	0.7	Derwent et al. 2010
CH₃OH	0.3	Derwent et al., 2010
HCOOH	0.1	Derwent et al., 2010
C₂H₆O	0.6	Derwent et al. 2010
C₄H₆	30.6	Xiong et al. 2020
C₃H₄O	11.5	Hakkim et al. 2021
C₃H₆O	0.3	Derwent et al. 2010
C₂H₄O₂	0.1	Derwent et al., 2010
C₄H₄O	29.4	Hakkim et al. 2021
C₅H₈	42.3	Hakkim et al. 2021, Xiong et- al. 2020
C₄H₆O	82.2	Hakkim et al. 2021
C₄H₈O	0.6	Derwent et- al. 2010
C₆H₆	145.6	Chan et- al. 2009, Kılıç et- al. 2018, Hakkim et al. 2021, Xiong et- al. 2020
C₅H₆O	41.2	Hakkim et al. 2021
C₆H₁₃	1	Xiong et- al. 2020
C₄H₉O₂	0.1	Derwent et- al. 2010
C₇H₈	100	Chan et- al. 2009, Kılıç et- al. 2018, Hakkim et al. 2021, Xiong et- al. 2020
C₆H₆O	278.4	Kılıç et- al. 2018, Hakkim et al. 2021, Khare et- al. 2020
C₅H₄O₂	188.2	Hakkim et al. 2021
C₆H₈O	188.2	Kılıç et- al. 2018
C₈H₈	83.5	Kılıç et- al. 2018, Hakkim et al. 2021, Xiong et- al. 2020
C₇H₆O	188.2	Kılıç et- al. 2018
C₈H₁₀	69.4	Chan et- al. 2009, Kılıç et- al. 2018, Hakkim et al. 2021, Xiong et- al. 2020
C₆H₆O₂	229.4	Kılıç et- al. 2018
C₉H₁₀	188.2	Kılıç et- al. 2018
C₉H₁₂	75.3	Chan et- al. 2009, Kılıç et- al. 2018, Hakkim et al. 2021, Xiong et- al. 2020
C₉H₁₄	117.6	Kılıç et- al. 2018
C₇H₈O₂	200	Yee et- al. 2013
C₁₀H₈	185	Chan et- al. 2009, Kılıç et- al. 2018, Hakkim et al. 2021, Khare et- al. 2020
C₁₀H₁₂	129.4	Kılıç et al. 2018, Khare et al. 2020
C₁₀H₁₄	86.3	Kılıç et- al. 2018, Hakkim et al. 2021, Khare et- al. 2020
C₁₀H₁₆	176.5	Xiong et- al. 2020
C₁₁H₁₀	221.6	Chan et- al. 2009, Kılıç et- al. 2018, Khare et- al. 2020
C₁₁H₁₂	88.2	Khare et- al. 2020
C₆H₄Cl₂	9.4	Xiong et- al. 2020
C₁₁H₁₄	102.9	Kılıç et- al. 2018, Khare et- al. 2020
C₁₀H₁₂O	70.6	Khare et- al. 2020
C₁₀H₁₆O	2.5	Khare et- al. 2020
C₁₀H₁₈O	0.9	Khare et al. 2020
C₁₂H₁₂	176.5	Chan et- al. 2009, Khare et- al. 2020
C₁₂H₁₆	114.7	Kılıç et- al. 2018, Khare et- al. 2020
C₁₃H₁₈	147.1	Khare et- al. 2020
C₁₀H₈O₃	170.6	Khare et- al. 2020
C₁₃H₂₀	147.1	Khare et- al. 2020
C₁₃H₂₂	4.6	Khare et- al. 2020
C₁₄H₁₄	288.2	Khare et- al. 2020
C₁₀H₁₆O₃	35.3	Witkowski et- al. 2017
C₁₄H₁₈	194.1	Khare et- al. 2020
C₁₄H₂₀	194.1	Khare et- al. 2020
C₁₄H₂₂	194.1	Khare et- al. 2020
C₁₆H₂₄	252.9	Khare et- al. 2020
C₁₇H₂₈	276.5	Khare et- al. 2020
C₁₈H₃₀	305.9	Khare et- al. 2020

Table S4: Correlation table (R) for the 11 factors with independent tracer species.

Tracer Species	Cooking	Solvents	Road Construction	Biogenic	Petrol 4-Wheeler	Paddy	Heating & Waste Burning	Petrol 2-Wheeler	Industrial	Photochemical	CNG
WS	-0.4	-0.2	-0.1	0.1	-0.3	-0.2	-0.2	-0.2	-0.3	0.0	-0.2
WD	0.1	-0.1	0.2	0.1	0.0	0.1	0.2	0.1	-0.1	0.1	0.1
CO ₂	0.2	0.2	0.1	-0.3	0.7	0.4	0.4	0.7	0.8	-0.1	0.3
N ₂ O	0.0	0.0	-0.1	-0.1	0.2	0.0	0.1	0.2	0.4	-0.1	0.2
CH ₄	0.2	0.3	0.0	-0.3	0.6	0.4	0.3	0.6	0.8	-0.1	0.3
PM ₁₀	0.3	0.2	0.1	-0.2	0.4	0.7	0.7	0.5	0.5	0.0	0.3
PM _{2.5}	0.3	0.2	0.1	-0.3	0.3	0.7	0.6	0.4	0.5	0.0	0.2
CO	0.3	0.3	0.1	-0.2	0.7	0.5	0.5	0.6	0.7	-0.1	0.4
NO	0.1	0.2	0.0	-0.2	0.8	0.3	0.4	0.5	0.7	-0.1	0.3
NO ₂	0.1	0.1	0.1	-0.2	0.3	0.3	0.7	0.4	0.2	0.0	0.2
NO _x	0.1	0.2	0.0	-0.2	0.8	0.3	0.5	0.6	0.7	-0.1	0.3
O ₃	0.1	-0.1	0.3	0.5	-0.2	0.0	-0.1	-0.2	-0.3	0.4	-0.2
SO ₂	0.0	-0.1	-0.2	-0.1	0.2	0.3	0.4	0.2	0.2	0.0	0.2
AT	0.2	-0.1	0.2	0.7	-0.2	-0.3	-0.6	-0.3	-0.4	0.3	0.1
RH	-0.1	0.1	-0.2	-0.5	0.0	-0.2	-0.4	-0.1	0.2	-0.4	-0.1
PAR	0.2	-0.2	0.2	0.7	-0.2	-0.1	-0.2	-0.3	-0.3	0.3	-0.2
VC	-0.2	-0.2	0.1	0.6	-0.3	-0.2	-0.3	-0.3	-0.4	0.2	-0.2
FC	0.4	0.3	0.2	-0.3	0.3	0.8	0.6	0.5	0.4	0.3	0.3
<u>HD</u>	<u>-0.2</u>	<u>-0.3</u>	<u>-0.3</u>	<u>-0.4</u>	<u>-0.1</u>	<u>0.3</u>	<u>0.8</u>	<u>0.1</u>	<u>-0.1</u>	<u>0.0</u>	<u>0.1</u>

AT: Ambient Temperature, RH: Relative Humidity, PAR: Photosynthetic active radiation, VC: Ventilation Coefficient,

FC: Daily Fire Count, HD: Heating Demand

Table S5: Emissions from different sectors for north-western, south-western, and south-eastern fetch regions.

<u>VOC (Gg y⁻¹)</u>									
	<u>NW</u>			<u>SW</u>			<u>SE</u>		
<u>Sector</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>
<u>Residential fuel usage</u>	<u>764</u>	<u>353</u>	<u>=</u>	<u>1421</u>	<u>947</u>	<u>=</u>	<u>1196</u>	<u>862</u>	<u>=</u>
<u>Industrial</u>	<u>302</u>	<u>113</u>	<u>=</u>	<u>867</u>	<u>55</u>	<u>=</u>	<u>635</u>	<u>133</u>	<u>=</u>
<u>Agricultural Residue</u>	<u>135</u>	<u>0</u>	<u>760</u>	<u>204</u>	<u>0</u>	<u>801</u>	<u>171</u>	<u>0</u>	<u>207</u>
<u>Transport</u>	<u>84</u>	<u>212</u>	<u>=</u>	<u>154</u>	<u>378</u>	<u>=</u>	<u>96</u>	<u>266</u>	<u>=</u>
<u>Solvents</u>	<u>403</u>	<u>78</u>	<u>=</u>	<u>939</u>	<u>222</u>	<u>=</u>	<u>896</u>	<u>204</u>	<u>=</u>
<u>Power Industry</u>	<u>7</u>	<u>2</u>	<u>=</u>	<u>27</u>	<u>4</u>	<u>=</u>	<u>12</u>	<u>4</u>	
<u>PM_{2.5} (Gg y⁻¹)</u>									
	<u>NW</u>			<u>SW</u>			<u>SE</u>		
<u>Sector</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>
<u>Residential fuel usage</u>	<u>382</u>	<u>379</u>	<u>=</u>	<u>713</u>	<u>934</u>	<u>=</u>	<u>597</u>	<u>830</u>	<u>=</u>
<u>Industrial</u>	<u>158</u>	<u>173</u>	<u>=</u>	<u>524</u>	<u>541</u>	<u>=</u>	<u>342</u>	<u>307</u>	<u>=</u>
<u>Agricultural Residue</u>	<u>97</u>	<u>0</u>	<u>95</u>	<u>206</u>	<u>0</u>	<u>100</u>	<u>168</u>	<u>0</u>	<u>26</u>
<u>Transport</u>	<u>8</u>	<u>65</u>	<u>=</u>	<u>18</u>	<u>137</u>	<u>=</u>	<u>12</u>	<u>80</u>	<u>=</u>
<u>Solvents</u>	<u>0</u>	<u>0</u>	<u>=</u>	<u>0</u>	<u>0</u>	<u>=</u>	<u>0</u>	<u>0</u>	<u>=</u>
<u>Power Industry</u>	<u>144</u>	<u>14</u>	<u>=</u>	<u>453</u>	<u>68</u>	<u>=</u>	<u>215</u>	<u>61</u>	<u>=</u>
<u>PM₁₀ (Gg y⁻¹)</u>									
	<u>NW</u>			<u>SW</u>			<u>SE</u>		
<u>Sector</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>	<u>EDGAR</u>	<u>REAS</u>	<u>FINN</u>
<u>Residential fuel usage</u>	<u>750</u>	<u>401</u>	<u>=</u>	<u>1391</u>	<u>994</u>	<u>=</u>	<u>1157</u>	<u>882</u>	<u>=</u>
<u>Industrial</u>	<u>211</u>	<u>308</u>	<u>=</u>	<u>684</u>	<u>1015</u>		<u>458</u>	<u>539</u>	<u>=</u>
<u>Agricultural Residue</u>	<u>103</u>	<u>0</u>	<u>192</u>	<u>217</u>	<u>0</u>	<u>203</u>	<u>177</u>	<u>0</u>	<u>52</u>
<u>Transport</u>	<u>10</u>	<u>67</u>	<u>=</u>	<u>22</u>	<u>140</u>	<u>=</u>	<u>14</u>	<u>83</u>	<u>=</u>
<u>Solvents</u>	<u>0</u>	<u>0</u>	<u>=</u>	<u>0</u>	<u>0</u>	<u>=</u>	<u>0</u>	<u>0</u>	<u>=</u>
<u>Power Industry</u>	<u>213</u>	<u>28</u>	<u>=</u>	<u>679</u>	<u>130</u>		<u>321</u>	<u>118</u>	

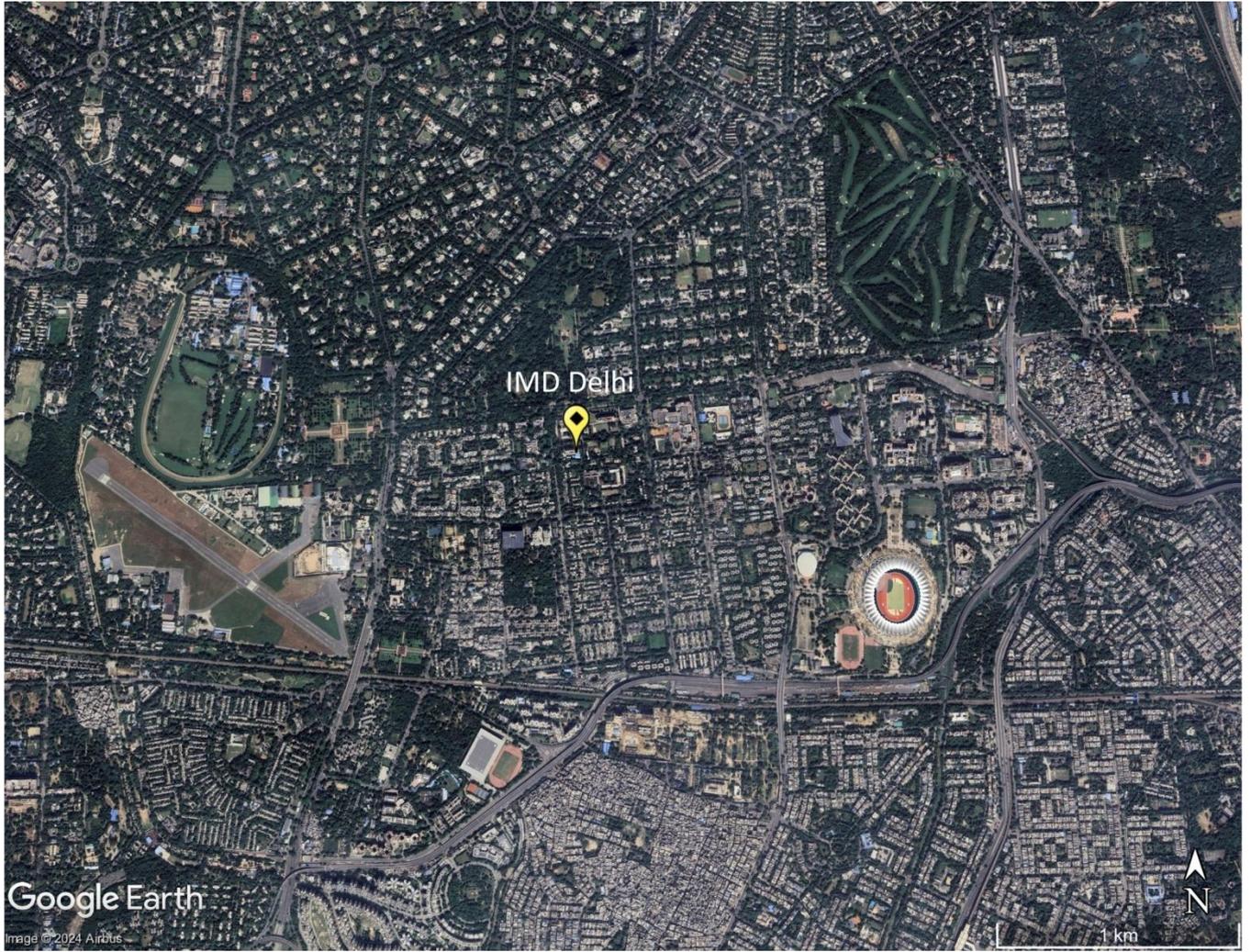


Figure S1: Map of the immediate surroundings of the IMD (28.5896°N-77.2210°E) sampling site in Central Delhi. (Google Earth Imagery ©Google Earth)

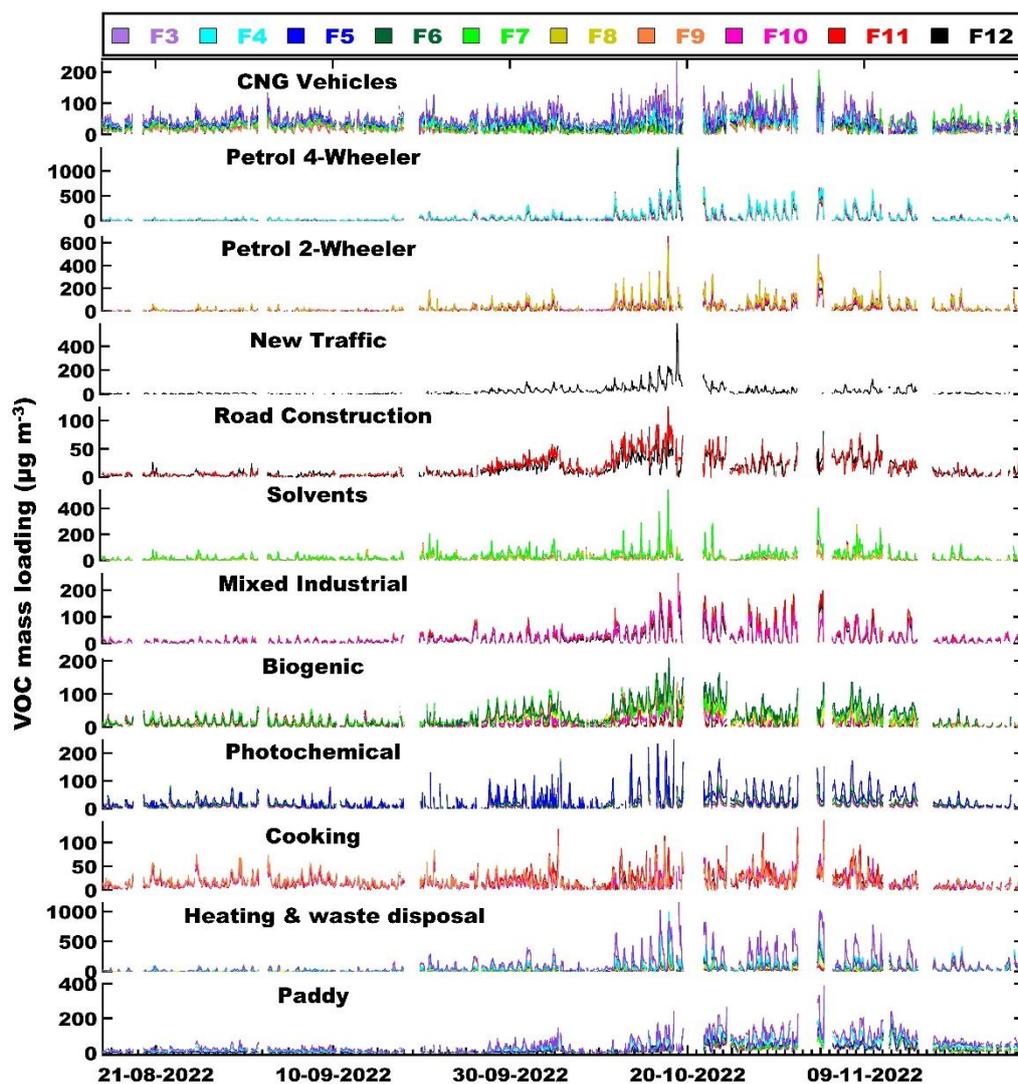


Figure S2(a): Evolution of the factor contribution time series when the number of factors is increased from 3 to 12.

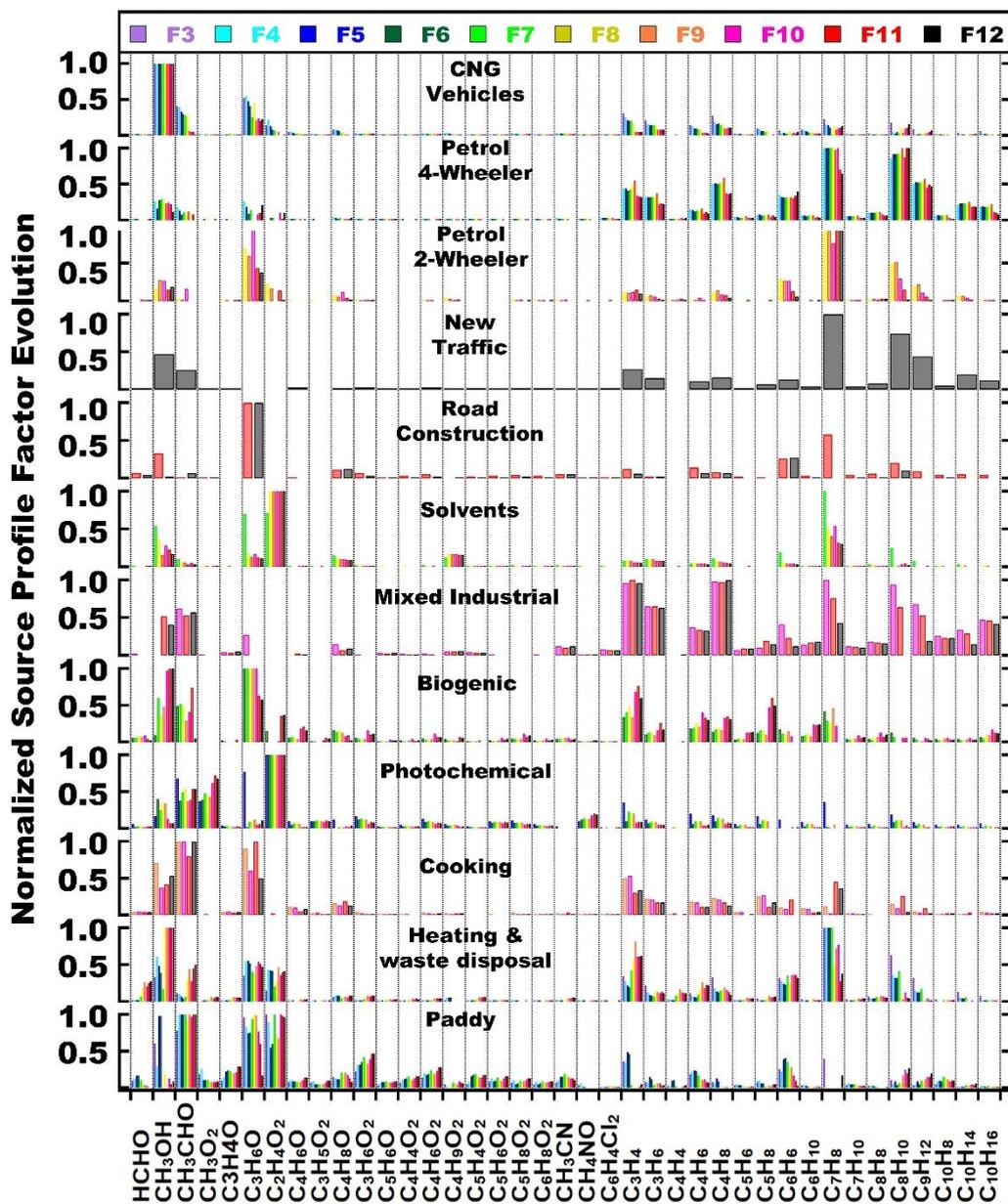


Figure S31(4): Evolution of the normalized PMF factor profile when the number of factors is increased from 3 to 12.

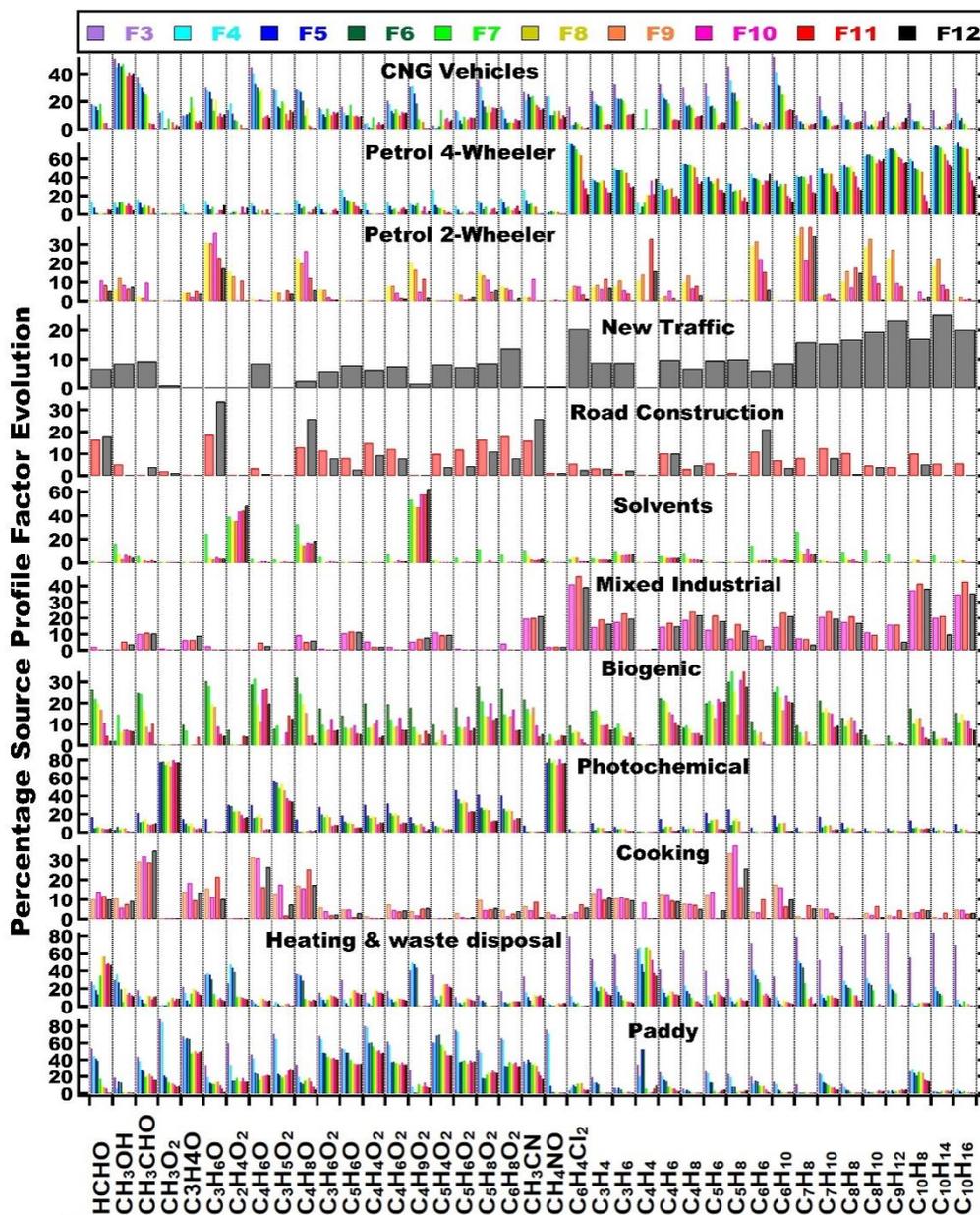


Figure S41(e): Evolution of the percentage of the mass explained by different sources when the number of factors is increased from 3 to 12.

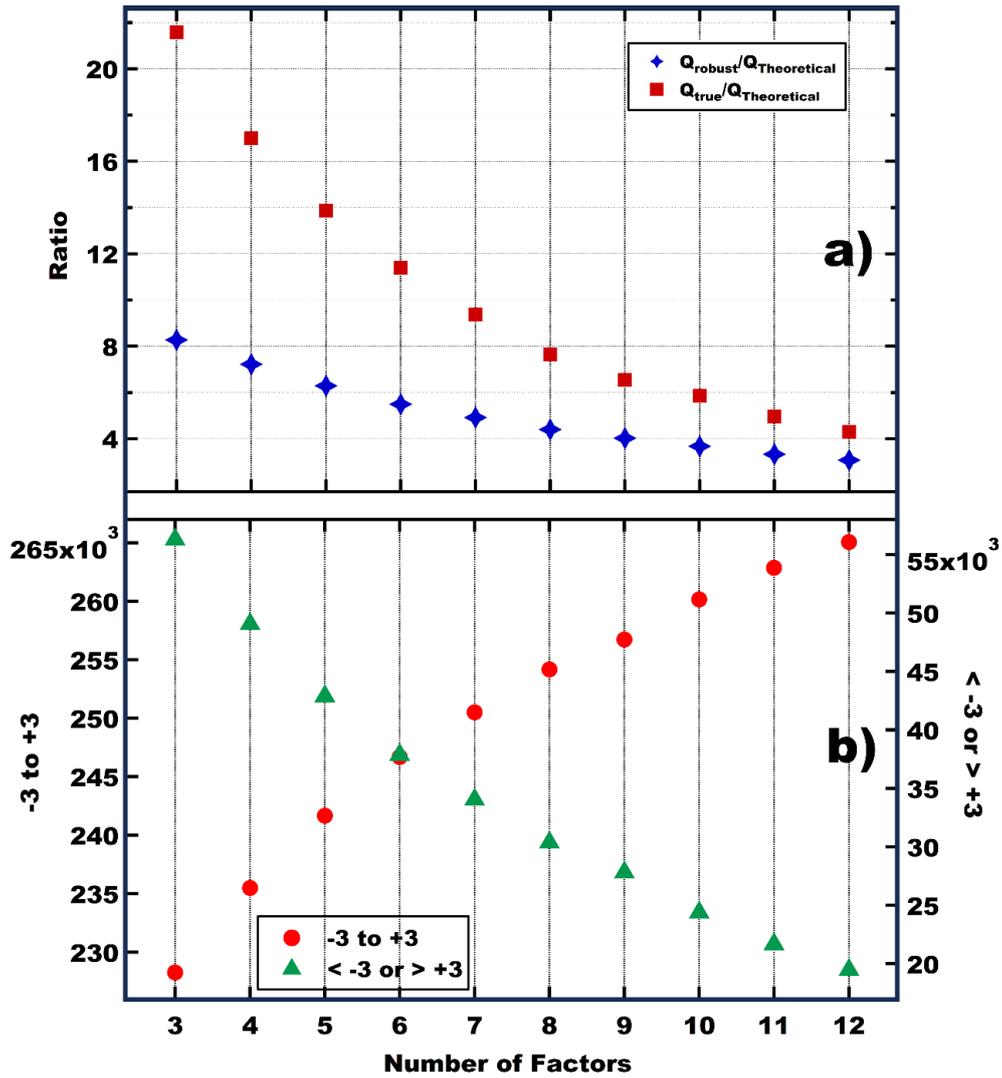


Figure S2S5: represents change in the a) $Q_{\text{true}}/Q_{\text{theoretical}}$ ratio and $Q_{\text{robust}}/Q_{\text{theoretical}}$, b) scaled residuals beyond 3 standard deviations and under 3 standard deviations when the number of factors is increased from 3 to 12.

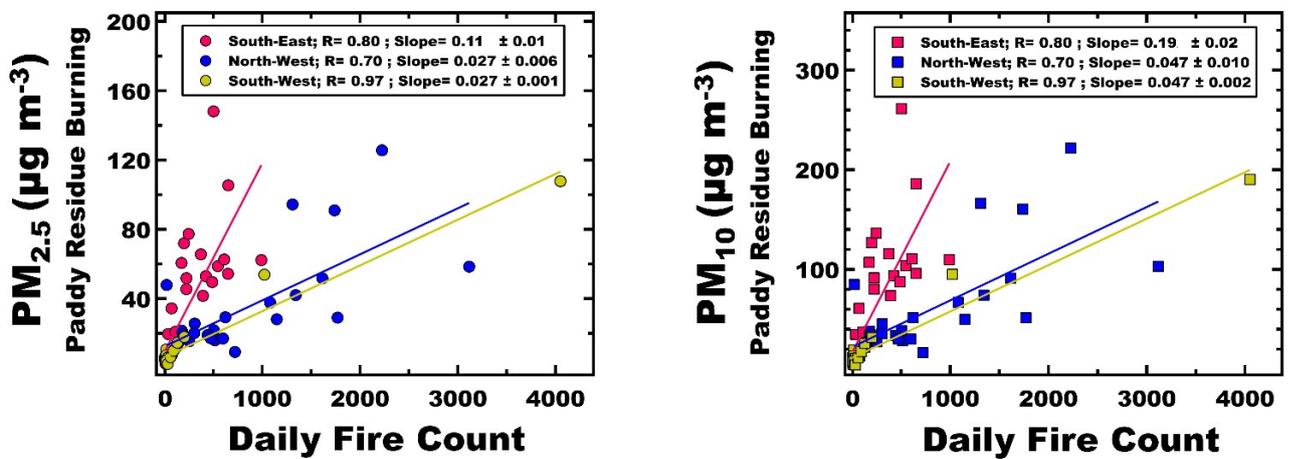


Figure S4S6: ~~represents the~~ Cross correlation analysis of PM_{2.5} and PM₁₀ mass loadings from the paddy residue burning factor at the receptor site influenced bywith the 24-hour averaged fire count in three fetch regions (SE,NW,SW).



Figure S7 Photographs of complete burning practises that prevail over Punjab. Fires of complete burns tend to be larger than fires from partial burns (Figure S8) and suffer lower omission error during satellite detection. Photo credits: Pooja Chaudhary and Vinayak Sinha



Figure S8 Photographs of partial burning practises that prevail in peri-urban areas and in the eastern IGP. The left hand side shows a line burn of residue left behind by the combined harvester while the right hand side shows a heap burn of residue next to the threshing machine after manual harvest. Fires of partial burns tend to be very small in comparison to the 375 x 375 m satellite footprint and suffer from close to 100% omission error during satellite detection. Photo credits: Pooja Chaudhary

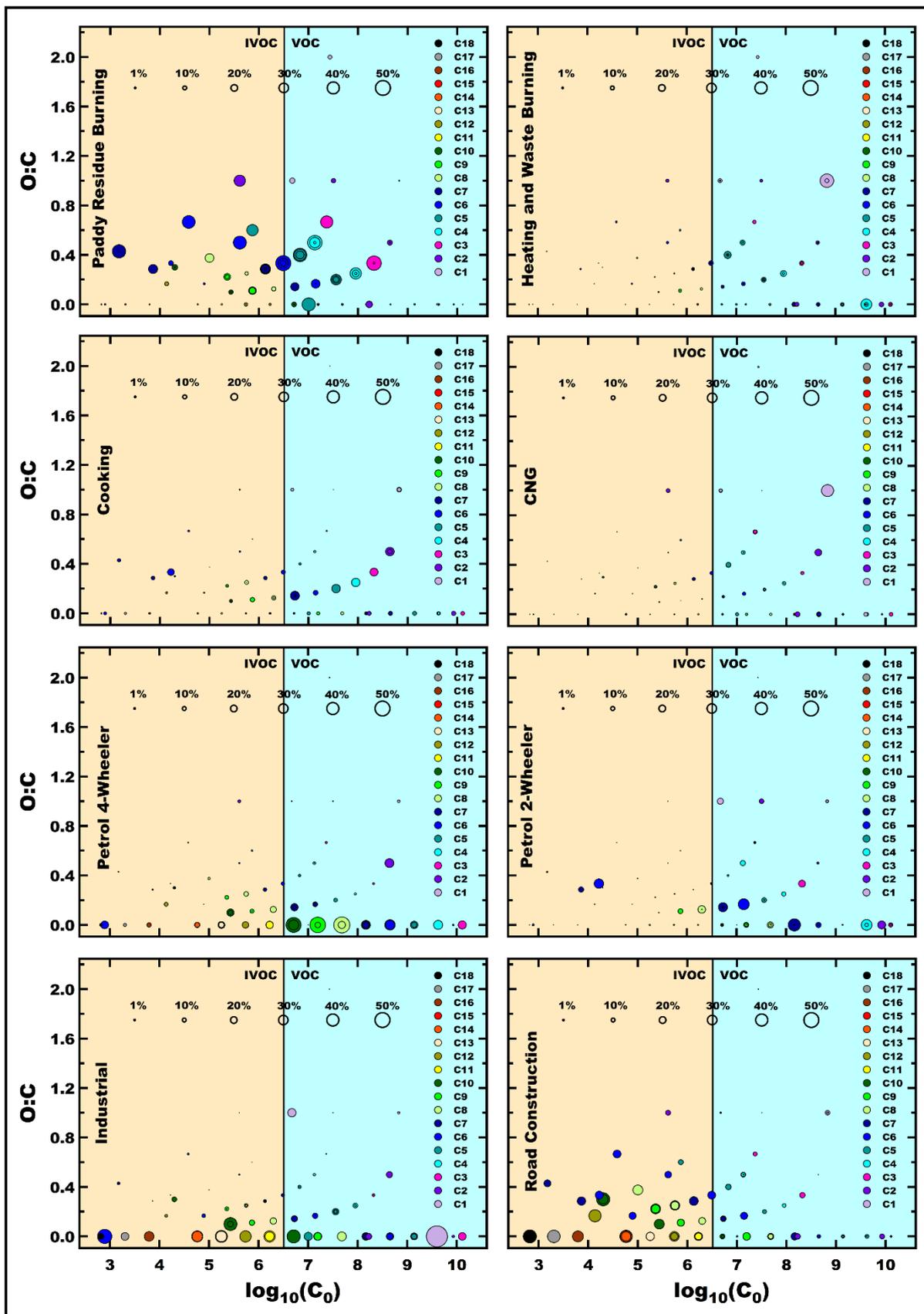


Figure S9: Volatility oxidation state plots for all factors that individually contribute more than 3% to the total SOA formation potential.

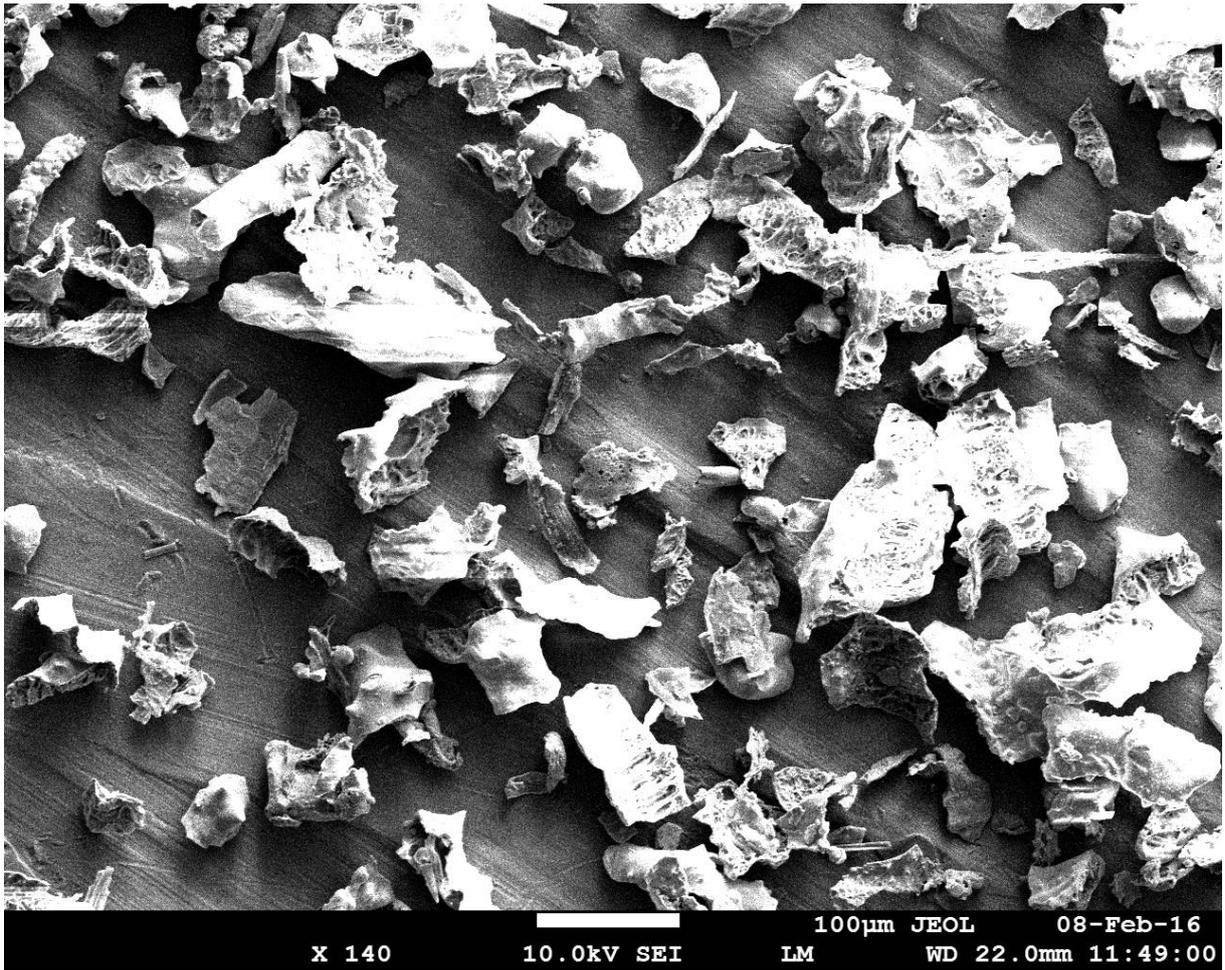


Figure S10 SEM image of rice ash from the electrostatic precipitator of an industrial boiler fired with rice husk and straw illustrating the coarse mode nature of the ash generated during the combustion of phytolith containing biomass.

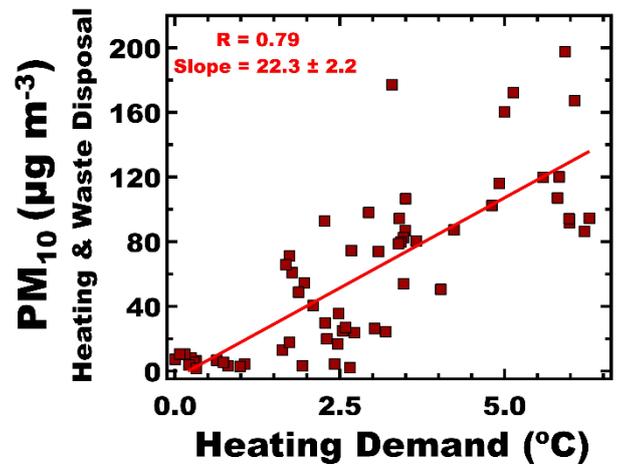
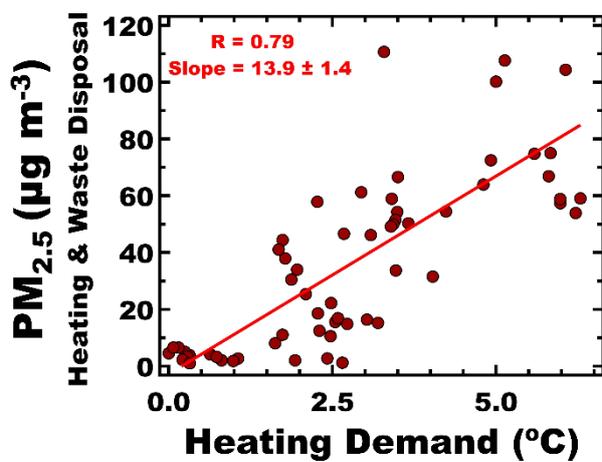


Figure S5S11: represents the PM_{2.5} Cross correlation analysis of the PM_{2.5} and PM₁₀ mass loadings attributed to residential heating and waste burning at the receptor site with the 24-hour averaged heating demand



Figure S12: Random selection of photographs clicked while driving around Delhi. One can clearly see the white CNG cylinders mounted in the place where the fuel tank used to be during vehicle conversion. Photo credits: Kriti Annika Sinha

References:

Bruns, E. A., Slowik, J. G., El Haddad, I., Kilic, D., Klein, F., Dommen, J., Temime-Roussel, B., Marchand, N., Baltensperger, U., and Prévôt, A. S. H.: Characterization of gas-phase organics using proton transfer reaction time-of-flight mass spectrometry: Fresh and aged residential wood combustion emissions. *Atmos. Chem. Phys.*, 17(1), 705–720, <https://doi.org/10.5194/acp-17-705-2017>, 2017.

Camredon, M., Hamilton, J. F., Alam, M. S., Wyche, K. P., Carr, T., White, I. R., Monks, P. S., Rickard, A. R., and Bloss, W. J.: Distribution of gaseous and particulate organic composition during dark α -pinene ozonolysis. *Atmos. Chem. Phys.*, 10, 2893–2917, <https://doi.org/10.5194/acp-10-2893-2010>, 2010.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crouse, J. D., Kürten, A., Wennberg, P. O., Flagan, R. C., and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: implications for oxidation of intermediate volatility organic compounds (IVOCs). *Atmos. Chem. Phys.*, 9, 3049–3060, <https://doi.org/10.5194/acp-9-3049-2009>, 2009.

Chandra, B. P., and Sinha, V.: Contribution of post-harvest agricultural paddy residue fires in the N.W. Indo-Gangetic Plain to ambient carcinogenic benzenoids, toxic isocyanic acid and carbon monoxide. *Environ. Int.*, 88, 187–197, <https://doi.org/10.1016/J.ENVINT.2015.12.025>, 2016.

Coggon, M. M., Lim, C. Y., Koss, A. R., Sekimoto, K., Yuan, B., Gilman, J. B., Hagan, D. H., Selimovic, V., Zarzana, K. J., Brown, S. S., Roberts, J. M., Müller, M., Yokelson, R., Wisthaler, A., Krechmer, J. E., Jimenez, J. L., Cappa, C., Kroll, J. H., De Gouw, J., and Warneke, C.: OH chemistry of non-methane organic gases (NMOGs) emitted from laboratory and ambient biomass burning smoke: evaluating the influence of furans and oxygenated

aromatics on ozone and secondary NMOG formation. *Atmos. Chem. Phys.*, 19, 14875–14899, <https://doi.org/10.5194/acp-19-14875-2019>, 2019.

Derwent, R. G., Jenkin, M. E., Saunders, S. M., and Pilling, M. J.: Photochemical ozone creation potentials for organic compounds in northwest Europe calculated with a master chemical mechanism. *Atmos. Environ.*, 32(14–15), 2429–2441, [https://doi.org/10.1016/S1352-2310\(98\)00053-3](https://doi.org/10.1016/S1352-2310(98)00053-3), 1998.

Derwent, R. G., Jenkin, M. E., Utembe, S. R., Shallcross, D. E., Murrells, T. P., and Passant, N. R.: Secondary organic aerosol formation from a large number of reactive man-made organic compounds. *Sci. Total. Environ.*, 408(16), 3374–3381, <https://doi.org/10.1016/J.SCITOTENV.2010.04.013>, 2010.

Fleming, L. T., Weltman, R., Yadav, A., Edwards, R. D., Arora, N. K., Pillarisetti, A., Meinardi, S., Smith, K. R., Blake, D. R., and Nizkorodov, S. A.: Emissions from village cookstoves in Haryana, India, and their potential impacts on air quality. *Atmos. Chem. Phys.*, 18, 15169–15182, <https://doi.org/10.5194/acp-18-15169-2018>, 2018.

Fukusaki, Y., Kousa, Y., Umehara, M., Ishida, M., Sato, R., Otagiri, K., Hoshi, J., Nudjima, C., Takahashi, K., and Nakai, S.: Source region identification and source apportionment of volatile organic compounds in the Tokyo Bay coastal area, Japan. *Atmos. Environ.: X*, 9, <https://doi.org/10.1016/j.aeaoa.2021.100103>, 2021.

Gkatzelis, G. I., Hohaus, T., Tillmann, R., Gensch, I., Müller, M., Eichler, P., Xu, K.-M., Schlag, P., Schmitt, S. H., Yu, Z., Wegener, R., Kaminski, M., Holzinger, R., Wisthaler, A., and Kiendler-Scharr, A.: Gas-to-particle partitioning of major biogenic oxidation products: a study on freshly formed and aged biogenic SOA. *Atmos. Chem. Phys.*, 18, 12969–12989, <https://doi.org/10.5194/acp-18-12969-2018>, 2018.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys.*, 6(11), 3181–3210, <https://doi.org/10.5194/ACP-6-3181-2006>, 2006.

Haeri, F.: Molecular Speciation of Organic Nitrogen Compounds Separated in Smoke Particles Emitted from Burning Western U.S. Wildland Fuels, PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, <https://doi.org/10.1184/R1/22670578.v1>, 2023.

Hakkim, H., Kumar, A., Annadate, S., Sinha, B., and Sinha, V.: RTEII: A new high-resolution ($0.1^\circ \times 0.1^\circ$) road transport emission inventory for India of 74 speciated NMVOCs, CO, NO_x, NH₃, CH₄, CO₂, PM_{2.5} reveals massive overestimation of NO_x and CO and missing nitromethane emissions by existing inventories. *Atmos. Environ.: X*, 11, 100118, <https://doi.org/10.1016/J.AEAOA.2021.100118>, 2021.

Harrison, M. A., Barra, S., Borghesi, D., Vione, D., Arsene, C., and Olariu, R. I.: Nitrated phenols in the atmosphere: a review. *Atmos. Environ.*, 39(2), 231–248, <https://doi.org/10.1016/j.atmosenv.2004.09.044>, 2005.

Hatch, L. E., Luo, W., Pankow, J. F., Yokelson, R. J., Stockwell, C. E., and Barsant K. C.: Identification and quantification of gaseous organic compounds emitted from biomass burning using two-dimensional gas chromatography–time-of-flight mass spectrometry. *Atmos. Chem. Phys.*, 15, 1865–1899, <https://doi.org/10.5194/acp-15-1865-2015>, 2015.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J., and Barsanti K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors. *Atmos. Chem. Phys.*, 17, 1471–1489, <https://doi.org/10.5194/acp-17-1471-2017>, 2017.

Hsu, C. Y., Wu, P. Y., Chen, Y. C., Chen, P. C., Guo, Y. L., Lin, Y. J., and Lin, P.: An integrated strategy by using long-term monitoring data to identify volatile organic compounds of high concern near petrochemical industrial parks. *Sci. Total. Environ.*, 821, <https://doi.org/10.1016/j.scitotenv.2022.153345>, 2022.

Jordan, C., Fitz, E., Hagan, T., Sive, B., Frinak, E., Haase, K., Cottrell, L., Buckley, S., and Talbot, R.: Long-term study of VOCs measured with PTR-MS at a rural site in New Hampshire with urban influences. *Atmos. Chem. Phys.*, 9(14), 4677–4697, <https://doi.org/10.5194/ACP-9-4677-2009>, 2009.

Kamarulzaman, N. H., Le-Minh, N., and Stuetz, R. M.: Identification of VOCs from natural rubber by different headspace techniques coupled using GC-MS. *Talanta*, 191, 535–544, <https://doi.org/10.1016/j.talanta.2018.09.019>, 2019.

Khare, P., Machesky, J., Soto, R. He, M., Presto, A. A. and Gentner, D. R.: Asphalt-related emissions are a major missing nontraditional source of secondary organic aerosol precursors. *Science Advances* 6, eabb9785, <https://doi.org/10.1126/sciadv.abb9785>, 2020.

Kılıç, D., El Haddad, I., Brem, B. T., Bruns, E., Bozetti, C., Corbin, J., Durdina, L., Huang, R.-J., Jiang, J., Klein, F., Lavi, A., Pieber, S. M., Rindlisbacher, T., Rudich, Y., Slowik, J. G., Wang, J., Baltensperger, U., and Prévôt, A. S. H.: Identification of secondary aerosol precursors emitted by an aircraft turbofan, *Atmos. Chem. Phys.*, 18, 7379–7391, <https://doi.org/10.5194/acp-18-7379-2018>, 2018.

Koss, A. R., Sekimoto, K., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., Jimenez, J. L., Krechmer, J., Roberts, J. M., Warneke, C., Yokelson, R. J., and De Gouw, J.: Non-methane organic gas emissions from biomass burning: Identification, quantification, and emission factors from PTR-ToF during the FIREX 2016 laboratory experiment. *Atmos. Chem. Phys.*, 18(5), 3299–3319, <https://doi.org/10.5194/ACP-18-3299-2018>, 2018.

Kumar, A., Hakkim, H., Sinha, B., and Sinha, V.: Gridded 1 km × 1 km emission inventory for paddy stubble burning emissions over north-west India constrained by measured emission factors of 77 VOCs and district-wise crop yield data. *Sci. Total. Environ.*, 789, 148064, <https://doi.org/10.1016/J.SCITOTENV.2021.148064>, 2021.

Kumar, A., Sinha, V., Shabin, M., Hakkim, H., Bonsang, B., and Gros, V.: Non-methane hydrocarbon (NMHC) fingerprints of major urban and agricultural emission sources for use in source apportionment studies. *Atmos. Chem. Phys.*, 20(20), 12133–12152, <https://doi.org/10.5194/ACP-20-12133-2020>, 2020.

Lignell, H., Epstein, S. A., Marvin, M. R., Shemesh, D., Gerber, B., and Nizkorodov, S.: Experimental and Theoretical Study of Aqueous cis-Pinonic Acid Photolysis. *J. Phys. Chem. A* 2013, 117, 12930–12945, <https://doi.org/10.1021/jp4093018>, 2013.

Loubet, B., Buysse, P., Gonzaga-Gomez, L., Lafouge, F., Ciuraru, R., Decuq, C., Kammer, J., Bsaibes, S., Boissard, C., Durand, B., Gueudet, J.-C., Fanucci, O., Zurfluh, O., Abis, L., Zannoni, N., Truong, F., Baisnée, D., Sarda-Estève, R., Staudt, M., and Gros, V.: Volatile organic compound fluxes over a winter wheat field by PTR-Qi-TOF-MS and eddy covariance, *Atmos. Chem. Phys.* 22, 2817–2842, <https://doi.org/10.5194/acp-22-2817-2022>, 2022.

Mochizuki, T., Kawamura, K., Miyazaki, Y., Kunwar, B., and Boreddy, S. K. R.: Distributions and sources of low-molecular-weight monocarboxylic acids in gas and particles from a deciduous broadleaf forest in northern Japan. *Atmos. Chem. Phys.*, 19(4), 2421–2432, <https://doi.org/10.5194/acp-19-2421-2019>, 2019.

Nowakowska, M., Herbinet, O., Dufour, A., and Glaude, P. A.: Kinetic Study of the Pyrolysis and Oxidation of Guaiacol. *J. Phys. Chem. A*, 122(39), 7894–7909, <https://doi.org/10.1021/acs.jpca.8b06301>, 2018.

Palm, B., Peng, Q., Fredrickson, C. D., Lee B. H., Garofalo, L. A., Pothier, M. A., Kreidenweis S. M., Farmer D. K., Pokhrel, R. P., Shen, Y., Murphy S. M., Permar W., Hu L. Campos T.L., Hall S. R., Ullmann K., Zhang, X., Flocke, F., Fischer, E. V., and Thornton, J. A.: Quantification of organic aerosol and brown carbon evolution in fresh wildfire plumes. *P. Natl. Acad. Sci. USA*, 117 (47) 29469–29477, <https://doi.org/10.1073/pnas.2012218117>, 2020.

Ramasamy, S., Nakayama, T., Imamura, T., Morino, Y., Kajii, Y. and Sato, K.: Investigation of dark condition nitrate radical- and ozone-initiated aging of toluene secondary organic aerosol: Importance of nitrate radical reactions with phenolic products. *Atmos. Environ.* 219, 117049, <https://doi.org/10.1016/j.atmosenv.2019.117049>, 2019.

Sarkar, C., Sinha, V., Kumar, V., Rupakheti, M., Panday, A., Mahata, K. S., Rupakheti, D., Kathayat, B., and Lawrence, M. G.: Overview of VOC emissions and chemistry from PTR-TOFMS measurements during the SusKat-ABC campaign: high acetaldehyde, isoprene and isocyanic acid in wintertime air of the Kathmandu Valley. *Atmos. Chem. Phys.*, 16, 3979–4003, <https://doi.org/10.5194/acp-16-3979-2016>, 2016.

Sarkar, C., Sinha, V., Sinha, B., Panday, A. K., Rupakheti, M., and Lawrence, M. G.: Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization. *Atmos. Chem. Phys.*, 17(13), 8129–8156, <https://doi.org/10.5194/ACP-17-8129-2017>, 2017.

Stockwell, C. E., Veres, P. R., Williams, J., and Yokelson, R. J.: Characterization of biomass burning emissions from cooking fires, peat, crop residue, and other fuels with high-resolution proton-transfer-reaction time-of-flight mass spectrometry. *Atmos. Chem. Phys.*, 15, 845–865, <https://doi.org/10.5194/acp-15-845-2015>, 2015.

Toda, K., Obata, T., Obokin, V. A., Potemkin, V. L., Hirota, K., Takeuchi, M., Arita, S., Khodzher, T. V., and Grachev, M. A.: Atmospheric methanethiol emitted from a pulp and paper plant on the shore of Lake Baikal, *Atmos. Environ.*, 44, 2427–2433, <https://doi.org/10.1016/j.atmosenv.2010.03.037>, 2010.

Wang, Z., Yuan, B., Ye, C., Roberts, J., Wisthaler, A., Lin, Y., Li, T., Wu, C., Peng, Y., Wang, C., Wang, S., Yang, S., Wang, B., Qi, J., Wang, C., Song, W., Hu, W., Wang, X., Xu, W., ... Shao, M.: High Concentrations of Atmospheric Isocyanic Acid (HNCO) Produced from Secondary Sources in China. *Environ. Sci. Technol.*, 54(19), 11818–11826, <https://doi.org/10.1021/acs.est.0c02843>, 2020.

Wang, M., Wang, Q., Ho, S. S. H., Li, H., Zhang, R., Ran, W., Que, L.: Chemical characteristics and sources of nitrogen-containing organic compounds at a regional site in the North China Plain during the transition period of autumn and winter. *Sci. Total. Environ.*, 812, 151451, <https://doi.org/10.1016/j.scitotenv.2021.151451>, 2022.

Witkowski, B., and Gierczak, T.: cis-Pinonic acid oxidation by hydroxyl radicals in the aqueous phase under acidic and basic conditions: kinetics and mechanism. *Environ. Sci. Technol.*, 51(17), 9765–9773, <https://doi.org/10.1021/acs.est.7b02427>, 2017.

Xing, C., Wan, Y., Wang, Q., Kong, S., Huang, X., Ge, X., [Xie, M. and Yu, H.etal](#): Molecular characterization of brown carbon chromophores in atmospherically relevant samples and their gas-particle distribution and diurnal variation in the atmosphere. *J. Geophys Res-Atmos.*, 128, e2022JD038142, <https://doi.org/10.1029/2022JD038142>, 2023.

Xiong, Y., Zhou, J., Xing, Z., and Du, K.: Optimization of a volatile organic compound control strategy in an oil industry center in Canada by evaluating ozone and secondary organic aerosol formation potential. *Environ. Res.*, 191, 110217, <https://doi.org/10.1016/j.envres.2020.110217>, 2020.

Xu, C., Gao, L., Lyu, C., Qiao, L., Huang, D., Liu, Y., Li, D. and Zheng, M.: Molecular characteristics, sources and environmental risk of aromatic compounds in particulate matter during COVID-2019: Nontarget screening by ultra-high resolution mass spectrometry and comprehensive two-dimensional gas chromatography. *Environ. Int.*, 167 107421, <https://doi.org/10.1016/j.envint.2022.107421>, 2022.

Yáñez-Serrano, A. M., Filella, I., Llusà, J., Gargallo-Garriga, A., Granda, V., Bourtsoukidis, E., Williams, J., Seco, R., Cappellin, L., Werner, C., de Gouw, J., and Peñuelas, J.: GLOVOCS - Master compound assignment guide for proton transfer reaction mass spectrometry users. *Atmos. Environ.*, 244, 117929, <https://doi.org/10.1016/J.ATMOSENV.2020.117929>, 2021.

Yao, L., Wang, M.-Y., Wang, X., Liu, Y.-J., Chen, H.-F., Zheng, J., Nie, W., Ding, A.-J., Geng, F.-H., Wang, D.-F., Chen, J.-M., Worsnop, D. R. and Wang, L.: Detection of atmospheric gaseous amines and amides by a high-resolution time-of-flight chemical ionization mass spectrometer with protonated-~~et~~ethanol reagent ions. *Atmos. Chem. Phys.*, 16, 14527–14543. <https://doi.org/10.5194/acp-16-14527-2016>, 2016.

Zaytsev, A., Koss, A. R., Breitenlechner, M., Krechmer, J. M., Nihill, K. j. Lim, C. Y., Rowe, J. C., Cox, J. L., Moss, J., Roscioli, J. R., Canagaratna, M. R., Worsnop, D. R., Kroll, J. H., and Keutsch, F. N.: Mechanistic Study of Formation of Ring-retaining and Ring-opening Products from Oxidation of Aromatic Compounds under Urban Atmospheric Conditions. *Atmos. Chem. Phys.*, 19, 15117–15129, <https://doi.org/10.5194/acp-19-15117-2019>, 2019.