



Source apportionment of volatile organic compounds in the north-west Indo-Gangetic Plain using positive matrix factorisation model

Pallavi¹, Baerbel Sinha¹, and Vinayak Sinha¹

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.

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

Abstract. In this study we undertook quantitative source apportionment for 32 volatile organic compounds (VOCs) measured at a suburban site in the densely populated North-West Indo-Gangetic Plain using the US EPA PMF 5.0 Model. Six sources were resolved by the PMF model namely “biofuel use and waste disposal”, “wheat-residue burning”, “industrial emissions and solvent use”, “cars”, “two-wheelers” and “mixed daytime sources”. The biofuel and waste disposal, wheat residue burning, industrial emissions and solvent use, combined traffic sources, mixed daytime sources accounted for 23.2 %, 22.4 %, 11.8 %, 25.1 %, and 15.7 % of the total VOC mass concentration respectively; 18.1 %, 32.4 %, 7.3 %, 21.9 %, and 20.3 % of the total O_3 formation potential respectively; and 14.9 %, 13.9 %, 10.1 %, 59.0 %, and 2.2 % of the SOA formation potential, respectively. Further the factors contributed 24.6 %, 8.5 %, 20.1 %, 46.8 %, and 0 %, respectively, to the human class I carcinogen benzene and 18.4 %, 25.4 %, 5.9 %, 13.3 %, and 36.9 %, respectively, to the toxic emerging contaminant isocyanic acid. Evaluation of emission inventories using the in-situ data derived PMF solution revealed that among EDGARv4.2, REASv2.1 and GAINsv5.0, the GAINsv5.0 emission inventory for year 2010, best agreed with the in-situ data derived PMF results for May 2012.

1 Introduction

Volatile organic compounds (VOCs) have diverse natural (760 Tg(C) y^{-1} (Sindelarova et al., 2014)) and anthropogenic sources (127 Tg y^{-1} average value (IPCC, 2013)). Certain VOCs emitted primarily by anthropogenic sources such as benzene and isocyanic acid, have direct adverse impacts on human health even at low ppb concentration exposures (Chandra and Sinha, 2016). In densely populated regions like the Indo Gangetic Plain (IGP), reactive anthropogenic VOCs contribute significantly towards the formation of health relevant secondary pollutants such as ozone and secondary organic aerosol (Chandra and Sinha, 2016; Sarkar et al., 2016). At our study site, a representative suburban site in the NW-IGP, the 8 h average NAAQS (national ambient air quality standard) for ozone limit of $100 \mu\text{g m}^{-3}$ was exceeded on 29 out of 31 days during May 2012 (Sinha et al., 2014), while the 24 h average NAAQS for $PM_{2.5}$ of $60 \mu\text{g m}^{-3}$ was exceeded during 27 out of 31 days in the same period. It has been shown that wheat residue burning results in significant enhancement (by 19 ppb) of the daytime ozone mixing ratios in pre-monsoon season (Kumar et al., 2016) and long range transport in the form of dust storms from the Arabian Peninsula



brings extremely high $PM_{2.5}$ mass loadings (with peak $PM_{2.5}$ mass loadings of $950 \mu\text{g m}^{-3}$ on 17th of May 2012) (Sinha et al., 2014; Pawar et al., 2015) and enhances the $PM_{2.5}$ mass by $\sim 30\%$.

However, ozone mixing ratios exceed the NAAQS even during the non-fire influenced days of the pre-monsoon season and the NAAQS for $PM_{2.5}$ is exceeded 60% of the time for air masses with no history of long range transport. This indicates that
5 local ozone and $PM_{2.5}$ precursor emissions deserve further study.

Previous source receptor modelling studies of VOC emission from India (Srivastava, 2004; Srivastava et al., 2005; Majumdar et al., 2008) produced results that conflicted strongly with the bottom up emission inventories, all of which have residential fuel usage (43% - 68%) as their largest VOC source. All previous studies employed a chemical mass balance (CMB) technique for ambient VOC source attribution and identified the transport sector as the main source in the form of evaporative emissions
10 (40-87%) in Mumbai (Srivastava, 2004)), diesel internal combustion engines (26-58%) in Delhi (Srivastava et al., 2005) and roadway/refuelling exhaust (40%) in Kolkata city (Majumdar et al., 2008). Different bottom up emission inventories have large discrepancies in this understudied region. For our study region (27.4-34.9°N and 72-79.8°E), EDGAR v4.2 estimates that the road transport sector contributes only 10.9% of the total anthropogenic VOC emissions (220 Gg y^{-1}), while REAS
15 v2.1 attributes 35.8% of the total anthropogenic VOC emissions (1227 Gg y^{-1}) to this sector. For industrial emissions and solvent use, EDGAR v4.2 again has the lowest (277 Gg y^{-1}) and REASv2.1 the highest absolute emissions of 736 Gg y^{-1} . Crop residue burning as VOC source is missing in REAS but accounted for a 4.7% (95 Gg y^{-1}) and 7% (163 Gg y^{-1}) share of the annual VOC emissions in EDGAR and GAINS, respectively. Considering the large discrepancies between bottom up inventories and different source receptor modelling studies, more source receptor modelling studies using robust statistical tools and better tracers for different biomass burning sources are necessary.

20 In the present study, we applied the US EPA's PMF 5.0 model in constrained mode for source apportionment of 32 VOCs measured at IISER Mohali Atmospheric Chemistry Facility in May 2012 with the objective of quantifying the most important sources of ozone and SOA precursors, the human class I carcinogen benzene and the emerging contaminant isocyanic acid (Chandra and Sinha, 2016), so that strategies for air pollution mitigation can benefit from quantitative evidence concerning the contribution of major sources.

25 2 Methods

2.1 Receptor site

The measurement facility is situated inside Indian Institute for Science Education and Research Mohali (IISER Mohali) campus (Figure 1a) which is a suburban site (30.667°N, 76.729°E, 310 m above mean sea level) in Mohali near Chandigarh in India (Figure 1b). Collectively the metropolitan of Chandigarh-Mohali-Panchkula forms a tri-city with a total population of
30 1,941,118 (Census, 2011). The main air transport toward the site was from the North West and the period studied was impacted by wheat residue burning, a dust storm and strong photochemistry. Figure 1a shows 72 h HYSPLIT back trajectories arriving at the site. With average wind speeds of 5.6 m s^{-1} during the study period (range 1-20 m s^{-1}) the meteorological conditions



were conducive for capturing the contribution of regional emission sources. The measurement site, the meteorology and the primary dataset acquired during May 2012 have been discussed in detail elsewhere (Sinha et al., 2014).

2.2 VOCs and other Auxiliary measurements

We used hourly data of 32 measured organic ions which were assigned to volatile organic compounds (Supplementary Table S1) based on PTR-TOF-MS studies conducted by our group within the South Asian environment (Sarkar et al., 2016) to initialize the US EPA PMF 5.0 model and employed CO , SO_2 , O_3 and NO_y as independent tracers to validate the results. Since the technical details of the measurements and the QA/QC protocol have already been described in detail (Sinha et al., 2014), we provide only a quick summary here. Ambient air sampling was performed continuously through a Teflon inlet line protected by an in-line Teflon filter. A high sensitivity proton transfer reaction quadrupole mass spectrometer PTR-QMS (HS Model 11-07HS-088, Ionicon Analytik Gesellschaft, Austria) was operated at drift tube pressure of 2.2 mbar, a drift tube temperature of 60 °C and a drift tube voltage of 600 V, which resulted in an operating E/N ratio of ~ 135 . Carbon monoxide (CO), Sulphur dioxide (SO_2), Ozone (O_3) and NO_y (NO, NO_2 and other nitrogen species converted to NO by a molybdenum converter such as nitric acid and PAN) were measured using Thermo Fischer Scientific 48i (IR filter correlation based spectroscopy), 43i (pulsed UV fluorescence), 49i (UV absorption photometry) and 42i trace level air quality analysers (chemiluminescence), respectively.

2.3 Positive Matrix Factorisation model

In the current study, US EPA PMF 5.0 receptor model (Norris et al., 2014) was applied to the ambient VOC dataset (in $\mu\text{g m}^{-3}$) from May 2012 measured at the IISER-Mohali Atmospheric chemistry facility comprising of data matrix of 721 samples (rows) and 32 species (columns). A detailed description of the model can be found elsewhere (Paatero and Tapper, 1994; Paatero, 1997). All 32 species were assigned a fixed 20% in the uncertainty and 18 were identified as weak based on the signal to noise ratio and the presence of potential isobaric interferences as detailed in table S2. For weak species, the PMF model triples the stated uncertainty to reduce their impact on the models solution. The extra modelling uncertainty was kept to zero and missing values were excluded. For every base model run, we used 20 runs with random seeds. Stable Q-values were obtained for all runs. The model was run with 3 to 7 factors, to identify the appropriate number of factors as discussed in the supplementary text in greater detail. Figure 2 shows the percentage contribution of the identified sources to the VOC burden for these runs. Figure 2 shows that a 7 Factor solution provides little advantage over a 6 Factor solution while a 5 Factor solution does not resolve the wheat residue burning source which is independently verified by MODIS fire counts over the region. The residuals for all species for the 6 Factor solution depicted a normal curve and fall within -3.3 sigma and +3.3 sigma for all species indicating a good model fit. The constraints feature of the 5.0 version of the model was utilised to improve the performance of the model further as described in greater detail in the supplementary text. The constrained model operation of the PMF version 5.0 allows to reduce the rotational ambiguity of the model using external knowledge. For example, if a source is inactive for a particular period (as is photochemistry at night), then the source contribution (g_{ik}) due to that factor during that time period can be pulled to zero in the model to provide more robust output. Similarly, a compound that is known to be present



only in primary emissions can be pulled down in the source composition (f_{kj}) matrix of the photochemistry factor. A detailed discussion of the use of constraints in a receptor model has been provided in previous studies (Paatero et al., 2002, 2014; Paatero and Hopke, 2009; Norris et al., 2014; Sarkar et al., 2016). Bootstrap model runs (Brown et al., 2015) were performed to assess the model uncertainty. Input parameters for the bootstrap runs constituted random seed, 100 number of bootstraps and default values for block size (10) and minimum correlation R-value (0.6) and there were no unmapped factors. Except for the car and two-wheeler factor (R=0.6) for which a certain degree of co-linearity is expected, none of the other factors showed cross correlation with each other (R<0.3) and the g-space plot even of this factor pair is well filled. The constraint mode was unable to force the PMF model to separate the wheat residue burning factor in a 5-factor solution without imposing a split between the car and 2-wheeler factor, indicating that these two indeed represent distinct source profiles.

10 2.4 Conditional Probability Function analysis

We perform a conditional probability function analysis (Leuchner and Rappenglück, 2010) by calculating the probability of observing mass concentrations above the 75th percentile of a given factor contribution for every wind direction. This aids in identifying physical locations of different PMF source factors without using back trajectories (Xie and Berkowitz, 2006).

2.5 Calculation of the ozone formation potential and SOA formation potential

15 Ozone production potential for each of the PMF derived source factors was calculated based on the method used by Sinha and co-workers (Sinha et al., 2012) as described in the supplementary text in greater detail. Secondary organic aerosol (SOA) potential was calculated for the PMF source factors using the literature SOA yields (Derwent et al., 2010) as described in the supplementary text.

3 Results and Discussion

20 3.1 Identification of PMF factors

Six source factors were resolved by the PMF model. These were identified as “biofuel use and waste disposal”, “wheat-residue burning”, “four-wheelers”, “two-wheelers”, “industrial emissions and solvent use” and “mixed daytime sources”, respectively. Factor profiles were cross-correlated with the fingerprints of source samples collected from a number of potential sources including wheat residue fires (Chandra et al., 2017; Kumar et al., 2018), a busy traffic junction (Chandra et al., 2017), tail-pipes of various vehicles (this study), waste burning (Sharma et al., Under Review), leaf litter burning (this study) and domestic biofuel use (Stockwell et al., 2016) to identify the sources. Figure 3 shows the factor profiles obtained from the PMF run (in dark blue), the percentage of each species explained by the respective PMF factor (red squares) and the source profiles of those sources which best matched the factor profile (in various colors as indicated in the legend). The identification of the factors is further supported by independent tracers such as the criteria air pollutants (NO_y , CO , SO_2 , O_3) and MODIS (Moderate Resolution Imaging Spectroradiometer) fire counts as discussed in detail below.



3.2 Factor 1 - Biofuel use & waste disposal

Figure 4 shows that biofuel use & waste disposal contributes 23.2 %, 18.1 % and 14.9 % of the total VOC mass, ozone formation potential and SOA formation potential, respectively. The factor profile correlates most strongly with the measured VOC source speciation profiles of domestic cooking ($R=0.8$), leaf-litter burning ($R=0.7$) and smoldering garbage fires ($R=0.6$).

5 As discussed previously for other South Asian atmospheric environments (Sarkar et al., 2017), the source contributions of domestic biofuel use and domestic waste burning are difficult to segregate due to the high spatio-temporal overlap of the two activities. As can be seen in Figure 5, the factor shows a weak bimodal behaviour with an early morning and late evening peak as both domestic biofuel use and waste disposal fires peak in the early morning and in the evening hours (Nagpure et al., 2015). CO serves as the best independent tracer (Figure 5) indicating that this factor represents a low temperature combustion with a
10 low combustion efficiency. Figure 5 shows that highest conditional probability for this factor is from the N (>0.4), the direction of the Dadu Majra landfill in Chandigarh, followed by the wind direction NW where a large village (Maui Baidwan) can be found within 1 km of the receptor and NE, the direction of Panchkula's garbage dump in Sector 23. This and the fact that the average contribution of this factor remains above $30 \mu\text{g m}^{-3}$ throughout the night indicates that garbage burning contributes significantly to the biofuel use & waste disposal factor.

15 Figure 3 and Figure 6 show that this factor explains 35 %, 35 %, 29 %, 42 %, 37 %, 34 % and 37 % of the total acetonitrile, acrolein, methanol, acetaldehyde, methyl vinyl ketone, propyne and propene mass concentration, respectively, in the PMF model. Most of the NMVOC mass in this factor was contributed by methanol ($\sim 10.6 \mu\text{g m}^{-3}$), formic acid ($1.9 \mu\text{g m}^{-3}$) and acetic acid ($7.4 \mu\text{g m}^{-3}$). High emission of oxygenated VOCs have been reported previously for source profiles of biofuel-stoves (Wang et al., 2009; Paulot et al., 2011; Stockwell et al., 2016) open waste burning (Sharma et al., Under Review)
20 and PMF factors results of residential biofuel use and waste disposal factor in Kathmandu, Nepal (Sarkar et al., 2017). It should be noted, that this factor is responsible for approximately 25 % of the total benzene emissions. Since benzene is an identified Group-1 carcinogen (IARC, 1987) and emissions occur within the household itself (domestic cooking) or within close proximity of the house (waste disposal) this factor deserves special attention in programs targeted at emission reductions. Direct emission of isocyanic acid, a highly toxic emerging contaminant and its photochemical precursors (Alkyl amines and
25 Amides) was observed from this source and explained 18 % of the isocyanic acid mass concentration and 7-15 % of all the alkyl amines and amides in the PMF model, respectively.

3.3 Factor 2 - Wheat Residue burning

Wheat residue burning takes place every year in the NW-IGP in the post-harvest season and generally peaks in the month of May. It has been shown that wheat residue burning has a major impact on both ozone mixing ratios (Kumar et al., 2016) and
30 VOC mixing ratios and hydroxyl radical reactivity (Kumar et al., 2018), resulting in a large suite of unknown ($\sim 40\%$) and poorly quantified reactive gaseous emissions. Figure 4 shows that wheat residue burning, contributes 22.4 % of the total VOC mass concentration, 32.4 % of the total ozone formation potential and 13.9 % of the total SOA formation potential. Figure 3 shows that the factor profile correlates most strongly with flaming wheat residue burning ($R=0.9$) and Figure 5 illustrates that



the best independent tracer for the average contribution of wheat residue burning to the total NMVOC mass are the daily fire counts with a cross correlation of $R=0.4$ and a lag of 2 days. Since wheat residue burning is an area source and emissions are transported to the receptor site from a large fetch region and often with a significant lag time, there is no strong conditional probability for enhancements from any specific wind direction.

5 Figure 3 shows that the four largest contributors to the total NMVOC mass in the wheat residue burning factor are acetic acid ($11.4 \mu\text{g m}^{-3}$), methanol ($3.3 \mu\text{g m}^{-3}$) acetaldehyde ($2.1 \mu\text{g m}^{-3}$) and acetone ($1.1 \mu\text{g m}^{-3}$). Figure 3 and Figure 6 demonstrate that more than 55 % of the hydroxyacetone, 37 % of the acetic acid, 32 % of the total methyl ethyl ketone and 28-39 % of the amides/amines as well as 28 % of the isocyanic acid mass in the model can be explained by this factor. This makes wheat residue burning the largest contributor to the human exposure to isocyanic acid in the month of May both through
10 direct emissions of isocyanic acid and by virtue of being the largest source for its photochemical precursors.

3.4 Factor 3 - Industrial emissions and solvent use

Figure 4 shows that the industrial emissions and solvent use jointly contribute 11.8 % of the total VOC mass concentration, 7.3 % of the total ozone formation potential and 10.1 % of the total SOA formation potential. Methanol ($7.3 \mu\text{g m}^{-3}$), acetic acid ($3.9 \mu\text{g m}^{-3}$) and acetone ($2.9 \mu\text{g m}^{-3}$) are the largest contributors to this factor profile. This points towards solvent use
15 (Gaimoz et al., 2011) and/or polymer manufacturing (Sarkar et al., 2017) contributing to the industrial emission and solvent use factor. In addition, Figure 3 and Figure 6 show that this factor explains a significant fraction of the benzene (20 %) and acetonitrile (17 %) mass in the PMF model. While both are known for their use as solvents (Brown et al., 2007) they can also be emitted from the combustion. Figure 5 shows that the factor contribution of the industrial emissions and solvent use factor correlated with the SO_2 time series ($R=0.6$), indicates that the emissions of coal or biofuel burning in industrial units and/or
20 coal fired power plants may also be contributing to this factor profile. Figure 5 shows that the highest conditional probability of this factor is to the South East direction (120° - 150° wind sector). The receptor site is downwind of a 600 MW coal fired power plant located in Jagadhri (80 km SE) as well as downwind of several industrial areas and brick kiln clusters located around Dera Bassi (15 km), Lalru (20 km) and Jagadhari (80 km) when the wind blows from this direction. In the Kathmandu valley, biofuel co-fired brick kilns explained a significant fraction of the benzene and acetonitrile mass (Sarkar
25 et al., 2017) and hence a combustion contribution from brick kilns to the factor profile cannot be ruled out. The diel profile broadly reflects boundary layer dynamics with factor contributions increasing continuously throughout the night indicating a buildup of constant emissions in the nocturnal boundary layer. Factor contributions peak in the early morning $17\text{-}26 \mu\text{g m}^{-3}$ between 5-9 am local time and the factor contribution of this factor decreases from 9 am onwards after the breakup of the nocturnal boundary layer. This factor has higher average than the median factor contributions at night due to strong plumes
30 (max $\sim 200 \mu\text{g m}^{-3}$) reaching the receptor when it is downwind of the industrial sector but not during other nights when the wind direction is from rural Punjab (NW) or the urban sector (NE).



3.5 Factor 4 and 5 - cars and two-wheelers

Figure 4 shows that cars and two-wheelers contribute 16.2 %, 8.9 % of the total VOC mass concentration, 16.5 %, 5.4 % of the total ozone formation potential and 36.9 %, 22.1 % of the total SOA formation potential, respectively, at the receptor site.

As can be seen in Figure 3, factor 4 was identified as a factor dominated by car exhaust because it correlated best with the tailpipe exhaust of petrol-fueled cars ($R=0.5$), urban traffic junction grab samples ($R=0.8$) and the independent tracer NO_y ($R=0.7$) which is considered to be a vehicular exhaust marker (Ramanathan et al., 1985). The factor profile is characterized by elevated concentration levels of benzene ($1.4 \mu\text{g m}^{-3}$), toluene ($2.3 \mu\text{g m}^{-3}$), sum of C-8 aromatics ($3.5 \mu\text{g m}^{-3}$) and sum of C-9 aromatics ($2.7 \mu\text{g m}^{-3}$) and explained 35 %, 30 %, 53 % and 58 % of the total benzene, toluene, C-8 aromatics and C-9 aromatics mass in the PMF model, respectively. Features of car's factor profile also resemble gasoline evaporation headspace for diesel ($R=0.5$) collected at a petrol pump. This indicates that the factor profile consists of multiple components contributed by fuel exhaust and fuel evaporative running losses from vehicles. Similar profiles have been observed during field measurements in Beirut, Lebanon (Salameh et al., 2014, 2016) and Hong Kong (Ho et al., 2004). The toluene to benzene ratio of this profile (1.7) is typical for traffic emissions (1.7-2.3) (Som et al., 2007; Hoque et al., 2008; Chandra et al., 2018) and the highest conditional probability is observed for the Chandigarh wind sector ($0-90^\circ$). As reported previously from Mexico City during the Milagro campaign (Bon et al., 2011), a significant mass of methanol ($4.3 \mu\text{g m}^{-3}$) and other oxygenated VOCs were present in the traffic emissions factor. The fact that this factor explains 28 % of the total m/z 57 is consistent with the gasoline additive MTBE landing at this m/z ratio as an interference to acrolein/methylketone (Karl et al., 2003; Warneke et al., 2003, 2005; Rogers et al., 2006). Signals at m/z 31, 47, 59, 61, 73, 87 in aged traffic plumes can be attributed to formaldehyde, glyoxal, formic acid, acetic acid, methylglyoxal and 2-butanedione which are products of the gas phase oxidation of toluene, C-8 and C-9 aromatic compounds (Bethel et al., 2000; Ervens et al., 2004). In addition, car exhaust also explained 34 % of the propyne mass in the model. Factor 5 was identified as 2-wheeler exhaust, as the factor profile showed the highest correlation with the tailpipe exhaust of petrol-fuelled 4-stroke two-wheelers ($R=0.6$) and the independent tracer NO_y ($R=0.6$). Toluene ($3.9 \mu\text{g m}^{-3}$), acetic acid ($4.2 \mu\text{g m}^{-3}$) and methanol ($2.4 \mu\text{g m}^{-3}$) feature as the most abundant compounds in this factor profile, which explains 50 % of the total toluene mass as well as 17 %, 12 % and 9 %, of the total C-8 aromatics, benzene and C-9 aromatics in the PMF model, respectively. While part of the signal at m/z 61 (acetic acid) may be due to fragmentation of octane or ethyl acetate (Warneke et al., 2003; Rogers et al., 2006) which could be present in fuel, the mass has also been attributed to acetic acid in a previous study of diesel tailpipe emissions (Jobson et al., 2005). Nevertheless, it still seems that the 2-wheeler factor profile has a higher contribution from oxidised compounds compared to the car factor profile indicating that the plumes are typically more aged. Figure 7 shows that this factor displays higher conditional probability than the car factor towards the towns Kharar (8 km N), Dera Bassi (15 km SE) and Lalru (20 km SE), and a lower conditional probability than the car factor towards Chandigarh (NE) indicating 2-wheelers are more abundant in small towns, while cars dominate the traffic emissions in urban Chandigarh. Figure 7 illustrates that both the traffic factors show bimodal peaks in morning ($10.3 \mu\text{g m}^{-3}$ at 5-9 am local time) and evening ($20 \mu\text{g m}^{-3}$ at 7-9 pm local time) during peak traffic hours. When the wind blows from the urban sector ($0-90^\circ$) during peak traffic hour (7-9 pm) peak factor contributions of $>140 \mu\text{g m}^{-3}$ for cars and $>80 \mu\text{g m}^{-3}$ for



2-wheelers are observed. As can be seen from Figure 6, the two traffic factors jointly explain 47 %, 80 %, 70 % and 67 % of the total benzene, toluene, C-8 and C-9 aromatic compounds in the model consistent with findings from the Kathmandu valley that traffic, not residential biofuel use and waste disposal is the more important source of aromatic compounds in South Asia. It is also clear that despite stringent regulations, the transport sector in the region is still the largest contributor to human benzene exposure.

3.6 Factor 6 - mixed daytime sources

Figure 4 shows mixed daytime sources comprising of biogenic emissions and photochemically formed compounds contribute 15.7 % of the total VOC mass, 20.3 % of the total ozone formation potential and 2.2 % of the total SOA formation potential. Figure 7 illustrates that the mixed daytime factor correlates most strongly with the independent tracer O_3 ($R=0.8$). It can be seen from Figure 6 that biogenic daytime emissions explained 22 % of the monoterpenes and 25 % of the measured isoprene, respectively. Isoprene has a short chemical lifetime of 1.5 hours during the day and 16 % and 11 % of its first generation oxidation products MVK and MEK were also attributed to this factor (Kesselmeier and Staudt, 1999). In addition, the mixed daytime factor explains 41 %, 44 %, 24 % and 22 % of the total formaldehyde, formic acid/ethanol, methanol and acetone mass, respectively. Photochemically formed isocyanic acid, formamide, acetamide and propanamide explain a slightly lower fraction (27-37 %) of the total mass concentration of these compounds compared to what has been reported from wintertime Kathmandu valley (36-41 %). Methanol ($8.9 \mu\text{g m}^{-3}$), formic acid ($4.4 \mu\text{g m}^{-3}$), acetic acid ($2.5 \mu\text{g m}^{-3}$) were the most important contributors to the factor fingerprint. Figure 7 illustrates that the mixed daytime factor peaks between 9 am and 4 pm and shows a slightly enhanced conditional probability for the 180° - 330° rural wind sector (0.2-0.3) due to agroforestry plantations of poplar in the rural landscape.

3.7 Split up of VOC Emission Sources in Mohali and their contribution to Ozone and SOA Formation Potential

Figure 4 (a) shows the contribution of the different sectors to ambient VOC mass concentration loadings during May 2012 in Mohali. The two traffic factors combined together were found to be the strongest contributors to the total VOC mass concentration (25.1 %) followed by biofuel use and waste disposal factor (23.2 %), wheat-residue burning (22.4 %), the mixed daytime factor (15.7 %) and industrial emissions (11.8 %), with the residual not apportioned VOC mass only amounting to 1.7 % of the total. Early source receptor modelling studies from India attributed a slightly larger share 26-58 % of the total VOC mass to traffic related emissions (Srivastava, 2004; Srivastava et al., 2005), suggesting that the progression to the emission norms Bharat stage III & IV (which are equivalent to Euro 3 and Euro 4 norms, <http://cpcb.nic.in/vehicular-exhaust/>) may have brought down VOC emissions from the traffic sector.

Figure 4 (b) shows the contribution of the different sectors to the ozone formation potential during May 2012 in Mohali. Wheat residue burning factor was found to be the largest contributor to the ozone formation potential (32.4 %) and has been shown to enhance ambient tropospheric ozone mixing ratios by 19 ppb (Kumar et al., 2016). Both traffic sources combined, the mixed daytime sources, biofuel use & waste disposal, and industrial emissions and solvent use contributed 21.9 %, 20.3 %, 18.1 % and 7.3 %, respectively, to the ozone formation potential. It is clear that in order to bring ozone levels into compliance



with the NAAQS, the wheat residue burning source of ozone precursors deserves the largest attention at this point, but the transport sector and biofuel use and waste disposal should not be neglected, either.

Figure 4 (c) shows the contribution of the different sectors to the SOA formation potential. Traffic is the single largest contributor and is responsible for contributing 59.0% of the SOA formation potential followed by biofuel use and waste disposal (14.9%), wheat residue burning (13.9%), industrial emissions and solvent use (10.1%) and the mixed daytime factor (2.2%). Total SOA formation potential amounting to $\sim 17 \mu\text{g m}^{-3}$, a resultant from all VOC source sectors indicates that at least 16% of the average PM_{2.5} mass loading ($104 \mu\text{g m}^{-3}$) for May2012 at IISER-Mohali could be secondary organic aerosols and that transport sector VOC emissions need to be targeted to reduce SOA formation.

3.8 Comparison of PMF source factors with existing Emission Inventories

Global Emission Database for Global Atmospheric Research (EDGARv4.2) inventory for the year 2008 (EDGARv4.2, 2011) and two regional emission inventories: Regional Emission inventory in Asia (REAS v2.1) for the year 2008 (Kurokawa et al., 2013) and the Greenhouse Gas and Air Pollution Interactions and Synergies model (GAINS) (Amann et al., 2011) for the year 2010 (Stohl et al., 2015) were compared with our PMF output. The gridded inventory was filtered for Latitude: 27.4-34.9°N and Longitude: 72-79.8°E, i.e. the fetch region from which the air mass trajectories reach the receptor site within one day. Annual emissions were available for EDGAR (2008) and GAINS (2010), whereas, REAS provided monthly data (May 2008). However, Figure S5 shows that despite providing monthly data, the REAS emission inventory has very little seasonality for any of the sources.

Figure 8 shows pie charts depicting the contribution of different sectors to the total VOC mass burden for the emission inventories and our PMF output. Biofuel use and waste disposal were responsible for 28.1% of the mass in our PMF but 67.9%, 44.2% and 41.7% of the mass in EDGAR, GAINS and REAS respectively. The contribution of crop residue burning (27.1%) to the VOC mass in the month of May would be highly underestimated by both GAINS (7%) and EDGAR (4.7%) if the annual emissions are attributed equally to all months of the year. However, if both emission inventories would attribute their annual crop residue burning emissions over the region only to the 2.5 months when crop residue burning actually occurs (middle of October to end of November and May), these emission inventory could be reconciled with the PMF solution, as emissions in May would amount to 26.5% and 19.2% of the monthly VOC emissions for the month of May for GAINS and EDGAR, respectively as shown in Figure 8. At the same time the percentage share of domestic fuel use and waste disposal would drop to 54% and 35% in EDGAR and GAINS, respectively and the contribution of industrial emissions and solvent drops to 18% in GAINS and 11% in EDGAR, respectively. Our PMF (14.3%) lies in between the estimate of these two emission inventories for industrial emissions and solvent use but closer to GAINS for wheat residue burning and domestic biofuel use and waste disposal. REAS overestimates the contribution of industrial activity and solvent use in the month of May (22%). Our PMF solution for road transport sector emissions (30.5%) lies in between the estimates of GAINS (558 Gg y⁻¹, 24%) and REAS (1230 Gg y⁻¹, 36.2%), possibly, because not all pre-2000 super-emitters for which the 20-year vehicle lifetime has been exceeded have been retired as planned.



Overall it appears that GAINS, the emission inventory with the lowest absolute emissions from residential and commercial biofuel use shows the best agreement with our PMF solution. Our PMF solution suggests that transport sector emissions are underestimated by approximately a factor of 1.5 in GAINS, while the combined effect of residential biofuel use and waste disposal emissions as well as the VOC burden associated with solvent use may be overestimated by a factor of 1.3 in the same emission inventory. Similar results have been reported previously. Sarkar and co-workers (Sarkar et al., 2017) reported an underestimation of transport sector emissions for the REAS and EDGAR emission inventory for the Kathmandu valley in Nepal and an overestimation of the residential biofuel use and waste disposal source in all emission inventories, while Gaimoz and co-workers (Gaimoz et al., 2011) reported an overestimation of the VOC emissions from solvent use in Paris.

REAS and EDGAR overestimated residential bio fuel usage emissions even more than GAINS. EDGAR underestimated transport sector emissions and industrial emissions and solvent usage while REAS overestimates the importance of the same two sources. REAS also fails to include agricultural residue burning as a source.

Our results highlight that for accurate air quality forecasting and modelling it is essential that emissions are attributed only to the months in which the activity actually occurs. This is important both for emissions from crop residue burning (which occur in May and from Mid-October to the end of November) and emissions from wildfires (which are restricted to the dry season and peak in April and May). Annually averaged emissions are unlikely to yield accurate air quality forecast in regions affected by such seasonal events. At present, more specialized fire emission inventories such as FINN (Wiedinmyer et al., 2011) must be used to account for the full seasonality and day to day variations of open burning emissions. We also demonstrate, that the source profiles obtained as PMF output can be validated and matched against samples collected at the potential sources to validate the factor identification.

We find that the GAINsv5.0 emission inventory for the year 2010 agreed best with the in-situ data derived PMF solution for May 2012.

4 Conclusions

Six VOC emission sources were extracted via PMF simulations from the dataset comprising of 32 VOC species measured online at primary temporal resolution of 1 minute at a sub-urban site in Mohali in the summer of 2012. US EPA PMF 5.0 Model was used for source apportionment of VOCs and PMF-resolved factors included traffic exhaust, biofuel use and waste disposal, wheat-residue burning and mixed daytime sources (comprising of biogenic emissions and photochemical formation), industrial emissions and solvent use, which along with the residuals, accounted for 25.1 %, 23.2 %, 22.4 %, 15.7 %, 11.8 % and 1.7 %, respectively, of the total VOC mass concentration. For the human class I carcinogen benzene, the traffic factor alone contributed to 47 % of the total benzene mass at this receptor site followed by residential biofuel use and waste disposal (25 %) and industrial emissions and solvent use (20 %). Since the annual NAAQS for benzene is exceeded at this receptor site (Chandra and Sinha, 2016), all three sectors must be targeted for emission reductions. For the emerging contaminant isocyanic acid, photochemical formation from precursors (37 %), wheat residue burning (25 %) and biofuel usage and waste disposal (18 %) were the largest contributors to human exposure. The monthly average isocyanic mixing ratio of 1.4 ppb



exceeds concentrations that can, after dissociation at blood pH, result in blood cyanate ion concentrations (Roberts et al., 2011) high enough to produce significant health effects in humans (Wang et al., 2007) such as atherosclerosis, cataracts and rheumatoid arthritis due to protein damage. Peak mixing ratios of this compound exceed 3 ppb in some night time wheat residue burning plumes. Wheat residue burning was also the single largest source of the photochemical precursors of isocyanic acid, namely, formamide, acetamide and propanamide, indicating that this source must be most urgently targeted to reduce human concentration exposure to isocyanic acid. Our results highlight that for accurate air quality forecasting and modelling it is essential that emissions that are both large in terms of their absolute contribution and display a significant seasonality in their occurrence are attributed only to the months in which the activity actually occurs. This is important both for emissions from crop residue burning (which occur in May and from Mid-October to the end of November) and emissions from wildfires (which are restricted to the dry season and peak in April and May). Annually averaged emissions are unlikely to yield accurate air quality forecast in regions affected by such seasonal events. We find that the GAINSv5.0 emission inventory for the year 2010 was best agreed with the in-situ data derived PMF solution for May 2012, as long as crop residue burning emissions were attributed to 2.5 months of the year only, and emissions from domestic biofuel use and solvent use were scaled down by a factor of 1.3 and transport sector emissions were scaled up by a factor of 1.5. The quantitative source apportionment results reported in this study for benzene, isocyanic acid and ozone and SOA precursors will provide much needed information for targeted mitigation efforts to improve the regional air quality.

Data availability. Data is available from the corresponding author upon request.

Author contributions. Pallavi performed the analysis and wrote the first draft of the paper. Dr. Baerbel Sinha conceived the analysis and revised the paper draft. Dr. Vinayak Sinha collected the data and commented on the paper draft.

20 *Competing interests.* The authors have no competing interests to declare.

Acknowledgements. We acknowledge the IISER Mohali Atmospheric Chemistry facility for data and the Ministry of Human Resource Development (MHRD), India for funding the facility. Pallavi acknowledges IISER Mohali for Institute PhD fellowship. This work was also partially supported through grant (SPLICE) DST/CCP/MRDP/100/2017(G) under the National Mission on Strategic knowledge for Climate Change (NMSKCC) MRDP Program of the Department of Science and Technology, India.



References

- Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Hoeglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schoepp, W., Wagner, F., and Winiwarter, W.: Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, *Environ. Model Softw.*, 26, 1489–1501., <https://doi.org/10.1016/j.envsoft.2011.07.012>, (<http://www.iiasa.ac.at/web/home/research/researchPrograms/air/Asia.html>), 2011.
- Bethel, H. L., Atkinson, R., and Arey, J.: Products of the gas-phase reactions of OH radicals with p-xylene and 1, 2, 3- and 1, 2, 4-trimethylbenzene: effect of NO₂ concentration, *J. Phys. Chem. A*, 104, 8922–8929, <https://doi.org/10.1021/jp001161s>, 2000.
- Bon, D. M., Ulbrich, I. M., de Gouw, J. A., Warneke, C., Kuster, W. C., Alexander, M. L., Baker, A., Beyersdorf, A. J., Blake, D., Fall, R., Jimenez, J. L., Herndon, S. C., Huey, L. G., Knighton, W. B., Ortega, J., Springston, S., and Vargas, O.: Measurements of volatile organic compounds at a suburban ground site (T1) in Mexico City during the MILAGRO 2006 campaign: measurement comparison, emission ratios, and source attribution, *Atmos. Chem. Phys.*, 11, 2399–2421, <https://doi.org/10.5194/acp-11-2399-2011>, 2011.
- Brown, S. G., Frankel, A., and Hafner, H. R.: Source apportionment of VOCs in the Los Angeles area using positive matrix factorization, *Atmos. Environ.*, 41, 227–237, <https://doi.org/10.1016/j.atmosenv.2006.08.021>, 2007.
- Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, *Sci. Total Environ.*, 518, 626–635, <https://doi.org/10.1016/j.scitotenv.2015.01.022>, 2015.
- Census: Government of India, Ministry of Home Affairs, Office of the Registrar General & Census Commissioner, India. <http://www.censusindia.gov.in/pca/Searchdata.aspx>- accessed on 25july2018, 2011.
- Chandra, B., 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., Sinha, V., Hakkim, H., Kumar, A., Pawar, H., Mishra, A., Sharma, G., Garg, S., Ghude, S. D., and Chate, D.: 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.*, 114, <https://doi.org/10.18520/cs/v114/i06/1318-1325>, 2018.
- Chandra, B. P. and Sinha, V.: Contribution of post-harvest agricultural paddy residue fires in the NW 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.
- 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, 3374–3381, <https://doi.org/10.1016/j.scitotenv.2010.04.013>, 2010.
- EDGARv4.2: European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). Emission Database for Global Atmospheric Research (EDGAR), release version 4.2., 2011.
- Ervens, B., Feingold, G., Frost, G. J., and Kreidenweis, S. M.: A modeling study of aqueous production of dicarboxylic acids: 1. Chemical pathways and speciated organic mass production, *J. Geophys. Res. Atmos.*, 109, D15 205, <https://doi.org/10.1029/2003JD004387>, 2004.
- Gaimoz, C., Sauvage, S., Gros, V., Herrmann, F., Williams, J., Locoge, N., Perrussel, O., Bonsang, B., d'Argouges, O., and Sarda-Estève, R.: Volatile organic compounds sources in Paris in spring 2007. Part II: source apportionment using positive matrix factorisation, *Environ. Chem.*, 8, 91–103, <https://doi.org/10.1071/EN10067>, 2011.



- Ho, K., Lee, S., Guo, H., and Tsai, W.: Seasonal and diurnal variations of volatile organic compounds (VOCs) in the atmosphere of Hong Kong, *Sci. Total Environ.*, 322, 155–166, <https://doi.org/10.1016/j.scitotenv.2003.10.004>, 2004.
- Hoque, R. R., Khillare, P., Agarwal, T., Shridhar, V., and Balachandran, S.: Spatial and temporal variation of BTEX in the urban atmosphere of Delhi, India, *Sci. Total Environ.*, 392, 30–40, <https://doi.org/10.1016/j.scitotenv.2007.08.036>, 2008.
- 5 IARC: International Agency for Research on Cancer. Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42, Supplement 7, 1987.
- IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp,
10 10.1017/CBO9781107415324., <https://doi.org/10.1017/CBO9781107415324>, 2013.
- Jobson, B., Alexander, M. L., Maupin, G. D., and Muntean, G. G.: On-line analysis of organic compounds in diesel exhaust using a proton transfer reaction mass spectrometer (PTR-MS), *Int. J. Mass Spectrom.*, 245, 78–89, <https://doi.org/10.1016/j.ijms.2005.05.009>, 2005.
- Karl, T., Jobson, T., Kuster, W. C., Williams, E., Stutz, J., Shetter, R., Hall, S. R., Goldan, P., Fehsenfeld, F., and Lindinger, W.: Use of
15 protontransferreaction mass spectrometry to characterize volatile organic compound sources at the La Porte super site during the Texas Air Quality Study 2000, *J. Geophys. Res. Atmos.*, 108, D16, <https://doi.org/10.1029/2002JD003333>, 2003.
- Kesselmeier, J. and Staudt, M.: Biogenic volatile organic compounds (VOC): an overview on emission, physiology and ecology, *J. Atmos. Chem.*, 33, 23–88, <https://doi.org/10.1023/A:1006127516791>, 1999.
- Kumar, V., Sarkar, C., and Sinha, V.: Influence of postharvest crop residue fires on surface ozone mixing ratios in the NW IGP analyzed using
20 2 years of continuous in situ trace gas measurements, *J. Geophys. Res. Atmos.*, 121, 3619–3633, <https://doi.org/10.1002/2015JD024308>, 2016.
- Kumar, V., Chandra, B., and Sinha, V.: Large unexplained suite of chemically reactive compounds present in ambient air due to biomass fires, *Sci. Rep.*, 8, 626, <https://doi.org/10.1038/s41598-017-19139-3>, 2018.
- Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima, K., and Akimoto, H.: Emissions of air
25 pollutants and greenhouse gases over Asian regions during 2000–2008: Regional Emission inventory in ASia (REAS) version 2, *Atmos. Chem. Phys.*, 13, 11 019–11 058. (<https://www.nies.go.jp/REAS/>), <https://doi.org/10.5194/acp-13-11019-2013>, 2013.
- Leuchner, M. and Rappenglück, B.: VOC source–receptor relationships in Houston during TexAQS-II, *Atmos. Environ.*, 44, 4056–4067, <https://doi.org/10.1016/j.atmosenv.2009.02.029>, 2010.
- Majumdar, D., Dutta, C., Mukherjee, A., and Sen, S.: Source apportionment of VOCs at the petrol pumps in Kolkata, India; exposure of
30 workers and assessment of associated health risk, *Transp. Res. D Transp. Environ.*, 13, 524–530, <https://doi.org/10.1016/j.trd.2008.09.011>, 2008.
- 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, 12 904–12 912, <https://doi.org/10.1021/acs.est.5b03243>, 2015.
- Norris, G., Duvall, R., Brown, S., and Bai, S.: EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide, EPA/600/R-14/108, 124, https://doi.org/https://www.epa.gov/sites/production/files/2015-03/epa_pmf_5.0_setup.exe,
35 https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf, 2014.
- Paatero, P.: Least squares formulation of robust non-negative factor analysis, *Chemom. Intell. Lab. Syst.*, 37, 23–35, [https://doi.org/10.1016/S0169-7439\(96\)00044-5](https://doi.org/10.1016/S0169-7439(96)00044-5), 1997.
- Paatero, P. and Hopke, P. K.: Rotational tools for factor analytic models, *J. Chemometr.*, 23, 91–100, <https://doi.org/10.1002/cem.1197>, 2009.



- Paatero, P. and Tapper, U.: Positive matrix factorization: A nonnegative factor model with optimal utilization of error estimates of data values, *Environmetrics.*, 5, 111–126, <https://doi.org/10.1002/env.3170050203>, 1994.
- Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, *Chemometr. Intell. Lab.*, 60, 253–264, [https://doi.org/10.1016/S0169-7439\(01\)00200-3](https://doi.org/10.1016/S0169-7439(01)00200-3), 2002.
- 5 Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, *Atmos. Meas. Tech.*, 7, 23–35, <https://doi.org/10.5194/amt-7-781-2014>, 2014.
- Paulot, F., Wunch, D., Crounse, J. D., Toon, G., Millet, D. B., DeCarlo, P. F., Vigouroux, C., Deutscher, N. M., González Abad, G., and Notholt, J.: Importance of secondary sources in the atmospheric budgets of formic and acetic acids, *Atmos. Chem. Phys.*, 11, 1989–2013, <https://doi.org/10.5194/acp-11-1989-2011>, 2011.
- 10 Pawar, H., Garg, S., Kumar, V., Sachan, H., Arya, R., Sarkar, C., Chandra, B., 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 (NW-IGP), *Atmos. Chem. Phys.*, 15, 9501–9520, <https://doi.org/10.5194/acp-15-9501-2015>, 2015.
- Ramanathan, V., Cicerone, R. J., Singh, H. B., and Kiehl, J. T.: Trace gas trends and their potential role in climate change, *J. Geophys. Res. Atmos.*, 90, 5547–5566, <https://doi.org/10.1029/JD090iD03p05547>, 1985.
- 15 Roberts, J. M., Veres, P. R., Cochran, A. K., Warneke, C., Burling, I. R., Yokelson, R. J., Lerner, B., Gilman, J. B., Kuster, W. C., Fall, R., and de G. J.: OIsocyanic acid in the atmosphere and its possible link to smokerelated health effects, *Proc. Natl. Acad. Sci.*, 108, 8966–8971, 2011.
- Rogers, T., Grimsrud, E., Herndon, S., Jayne, J., Kolb, C. E., Allwine, E., Westberg, H., Lamb, B., Zavala, M., and Molina, L.: On-road measurements of volatile organic compounds in the Mexico City metropolitan area using proton transfer reaction mass spectrometry, *Int. J. Mass Spectrom.*, 252, 26–37, <https://doi.org/10.1016/j.ijms.2006.01.027>, 2006.
- 20 Salameh, T., Afif, C., Sauvage, S., Borbon, A., and Locoge, N.: Speciation of non-methane hydrocarbons (NMHCs) from anthropogenic sources in Beirut, Lebanon, *Environ. Sci. Pollut. Res.*, 21, 10 867–10 877, <https://doi.org/10.1007/s11356-014-2978-5>, 2014.
- Salameh, T., Sauvage, S., Afif, C., Borbon, A., and Locoge, N.: Source apportionment vs. emission inventories of non-methane hydrocarbons (NMHC) in an urban area of the Middle East: local and global perspectives, *Atmos. Chem. Phys.*, 16, 3595–3607, <https://doi.org/10.5194/acp-16-3595-2016>, 2016.
- 25 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-TOF-MS 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.
- 30 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, 8129–8156, <https://doi.org/10.5194/acp-17-8129-2017>, 2017.
- Sharma, G., Sinha, B., Jangra, P., Hakkim, H., Chandra, B. P., Kumar, A., and Sinha, V.: Gridded emissions of CO, NO_x, SO₂, CO₂, NH₃, HCl, CH₄, PM_{2.5}, PM₁₀, BC and NMVOC from open municipal waste burning in India, *Environ. Sci. Technol.*, Under Review.
- 35 Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, *Atmos. Chem. Phys.*, 14, 9317–9341, <https://doi.org/10.5194/acp-14-9317-2014>, 2014.



- Sinha, V., Williams, J., Diesch, J., Drewnick, F., Martinez, M., Harder, H., Regelin, E., Kubistin, D., Bozem, H., and Hosaynali-Beygi, Z.: Constraints on instantaneous ozone production rates and regimes during DOMINO derived using in-situ OH reactivity measurements, *Atmos. Chem. Phys.*, 12, 7269–7283, <https://doi.org/10.5194/acp-12-7269-2012>, 2012.
- 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, 5921–5941, <https://doi.org/10.5194/acp-14-5921-2014>, <https://www.atmos-chem-phys.net/14/5921/2014/>, 2014.
- Som, D., Dutta, C., Chatterjee, A., Mallick, D., Jana, T., and Sen, S.: Studies on commuters' exposure to BTEX in passenger cars in Kolkata, India, *Sci. Total Environ.*, 372, 426–432, <https://doi.org/10.1016/j.scitotenv.2006.09.025>, 2007.
- Srivastava, A.: Source apportionment of ambient VOCS in Mumbai city, *Atmos. Environ.*, 38, 6829–6843, <https://doi.org/10.1016/j.atmosenv.2004.09.009>, 2004.
- Srivastava, A., Sengupta, B., and Dutta, S.: Source apportionment of ambient VOCs in Delhi City, *Sci. Total Environ.*, 343, 207–220, <https://doi.org/10.1016/j.scitotenv.2004.10.008>, 2005.
- Stockwell, C. E., Christian, T. J., Goetz, J. D., Jayarathne, T., Bhave, P. V., Praveen, P. S., Adhikari, S., Maharjan, R., DeCarlo, P. F., and Stone, E. A.: Nepal Ambient Monitoring and Source Testing Experiment (NAMaSTE): emissions of trace gases and light-absorbing carbon from wood and dung cooking fires, garbage and crop residue burning, brick kilns, and other sources, *Atmos. Chem. Phys.*, 16, 11 043–11 081, <https://doi.org/10.5194/acp-16-11043-2016>, 2016.
- Stohl, A., Aamaas, B., Amann, M., Baker, L. H., Bellouin, N., Berntsen, T. K., Boucher, O., Cherian, R., Collins, W., and Daskalakis, N.: Evaluating the climate and air quality impacts of short-lived pollutants, *Atmos. Chem. Phys.*, 15, 10 529–10 566, <https://doi.org/10.5194/acp-15-10529-2015>, 2015.
- Wang, S., Wei, W., Du, L., Li, G., and Hao, J.: Characteristics of gaseous pollutants from biofuel-stoves in rural China, *Atmos. Environ.*, 43, 4148–4154, <https://doi.org/10.1016/j.atmosenv.2009.05.040>, 2009.
- Wang, Z., Nicholls, S. J., Rodriguez, E. R., Kummu, O., Hörkkö, S., Barnard, J., Reynolds, W. F., Topol, E. J., DiDonato, J. A., and L, H. S.: Protein carbamylation links inflammation, smoking, uremia and atherogenesis, *Nat. Med.*, 13, 1176–1184, 2007.
- Warneke, C., De Gouw, J. A., Kuster, W. C., Goldan, P. D., and Fall, R.: Validation of atmospheric VOC measurements by proton-transfer-reaction mass spectrometry using a gas-chromatographic pre-separation method, *Environ. Sci. Technol.*, 37, 2494–2501, <https://doi.org/10.1021/es026266i>, 2003.
- Warneke, C., Kato, S., de Gouw, J. A., Goldan, P. D., Kuster, W. C., Shao, M., Lovejoy, E. R., Fall, R., and Fehsenfeld, F. C.: Online volatile organic compound measurements using a newly developed proton-transfer ion-trap mass spectrometry instrument during New England Air Quality Study Intercontinental Transport and Chemical Transformation 2004: Performance, intercomparison, and compound identification, *Environ. Sci. Technol.*, 39, 5390–5397, <https://doi.org/10.1021/es050602o>, 2005.
- Wiedinmyer, C., Akagi, S., Yokelson, R. J., Emmons, L., Al-Saadi, J., Orlando, J., and Soja, A.: The Fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning, *Geosci. Model Dev.*, 4, 625, <https://doi.org/10.5194/gmd-4-625-2011>, 2011.
- Xie, Y. and Berkowitz, C. M.: The use of positive matrix factorization with conditional probability functions in air quality studies: an application to hydrocarbon emissions in Houston, Texas, *Atmos. Environ.*, 40, 3070–3091, <https://doi.org/10.1016/j.atmosenv.2005.12.065>, 2006.

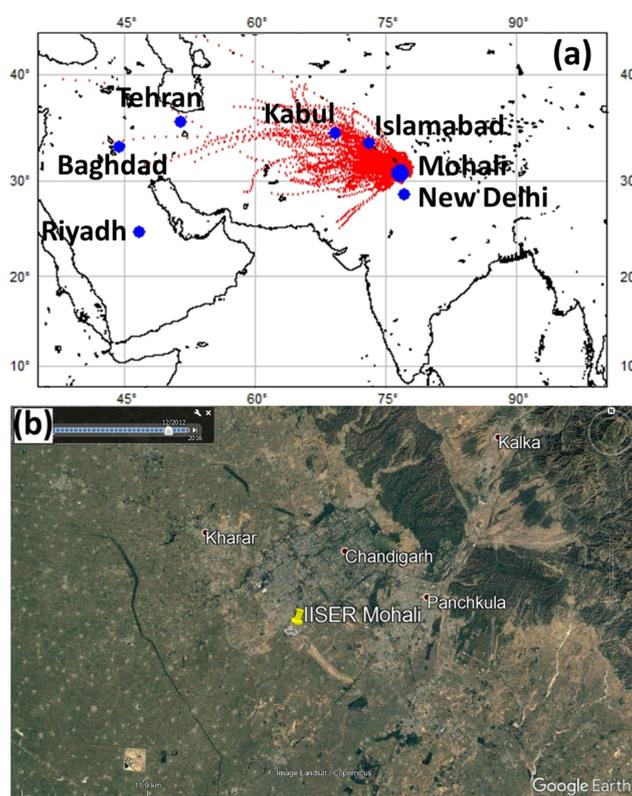


Figure 1. (a) Mohali located on Indian Subcontinent with the overlaid 72 h air mass back trajectories for May 2012 at 09:00 LT and 23:00 LT (UTC+5:30) (b) Precise location of IISER-Mohali Atmospheric chemistry facility with nearby cities on google earth imagery. The campus of IISER Mohali is outlined in white.

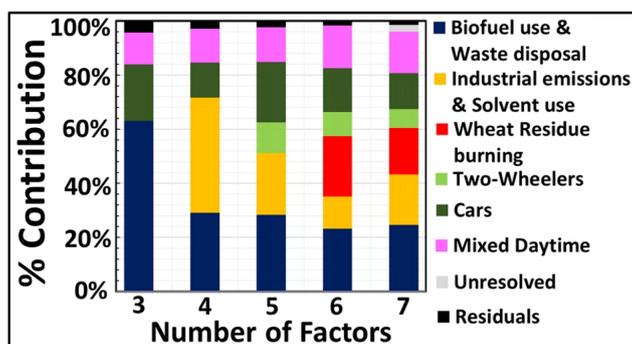


Figure 2. Percentage contribution assignment for various PMF factor number solutions (3-7) to the corresponding VOC emission sources.

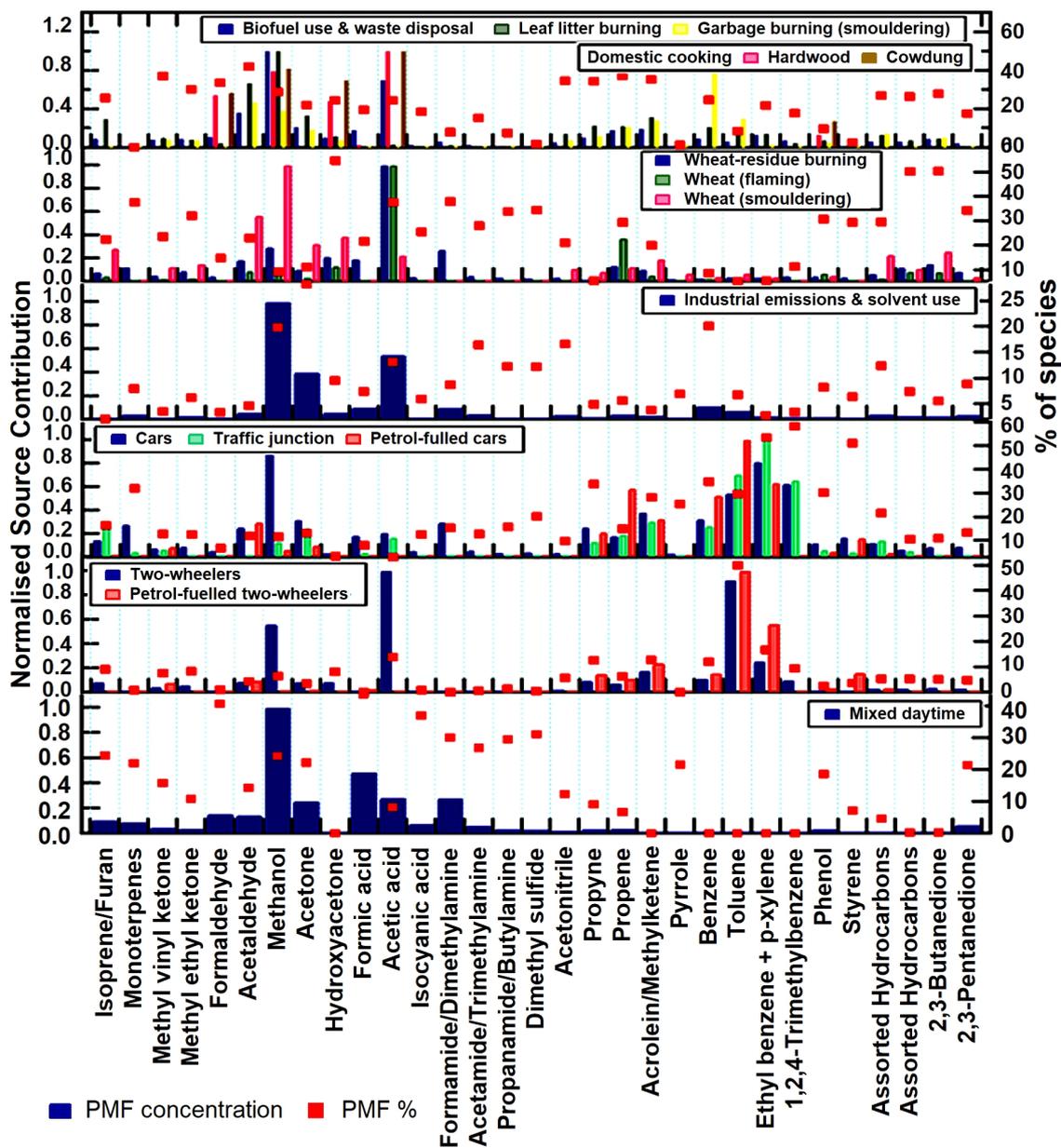


Figure 3. Factor profile composition for (6) PMF resolved factors at IISER-Mohali. It displays the normalized source fingerprints of the PMF factors (dark blue) and samples collected at source (in various colours) in bar-chart form. The value of the normalized species contribution is depicted on the left hand axis. The percentage of each species explained by each of the PMF factors is displayed in the form of a red square to be read from the right hand axis.

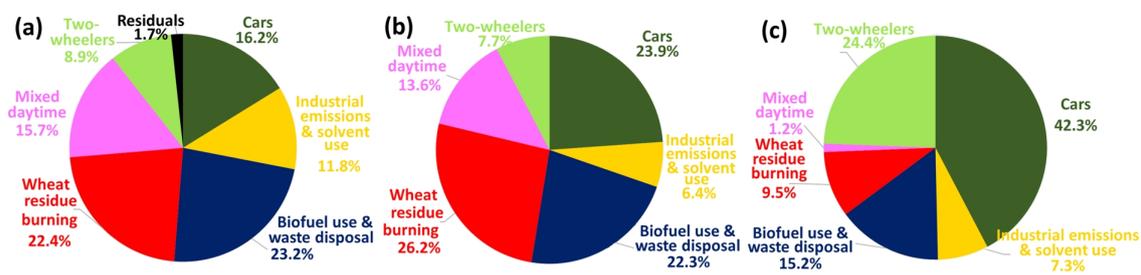


Figure 4. (a) Source contribution to the ambient VOC loading at the receptor site. (b) Ozone formation potential for PMF derived sources (c) SOA potential for PMF factors.

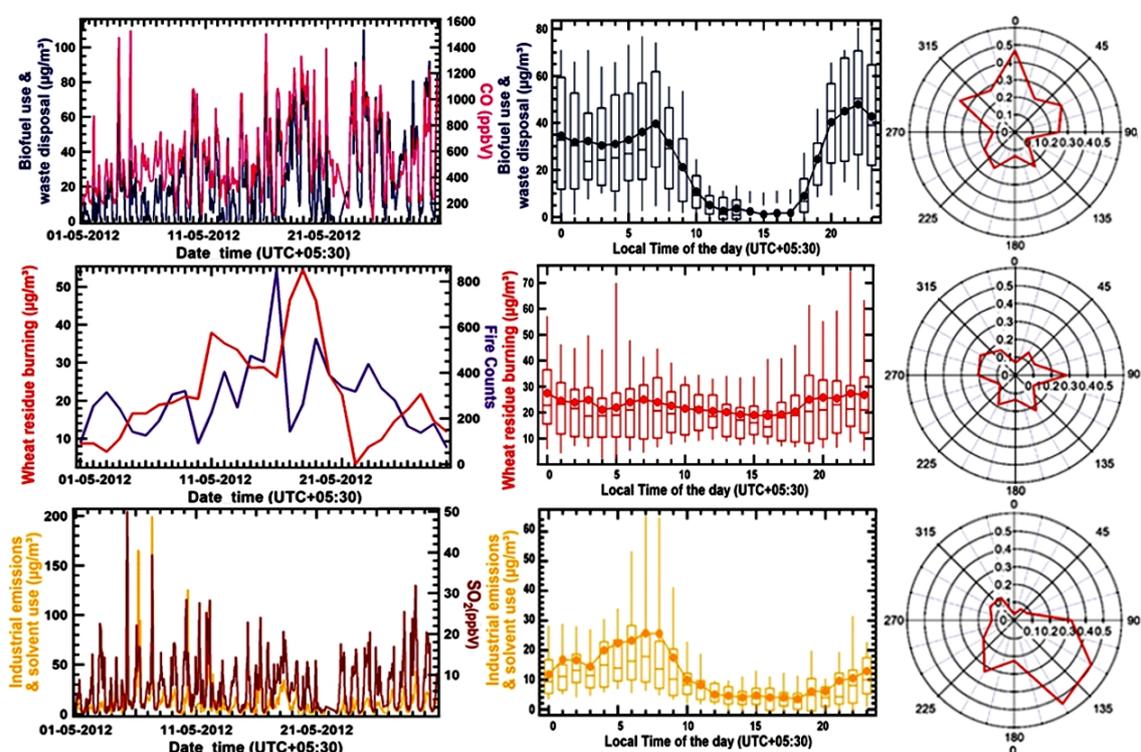


Figure 5. Factor contribution time series, factor diel variability and CPF plot for PMF Factor 1 (Biofuel use and waste disposal), PMF Factor2 (Wheat-residue burning) and PMF Factor3 (Industrial emissions and solvent use) for May2012. The time series of PMF factor's hourly mass in $\mu\text{g m}^{-3}$ is plotted against independent tracer species CO (in ppbv) for the biofuel use and waste disposal factor, daily fire counts for the wheat residue burning factor and SO_2 (in ppbv) for the industrial emission and solvent use factor. The Diel box and whisker plot shows the statistical parameters of factor's hourly mass contribution in $\mu\text{g m}^{-3}$ for every hour of the day plotted against the start time of the hour. The width of the box gives 25th and 75th percentiles, 50th percentile partitions the box; whiskers represent 10th and 90th percentiles of the dataset and average values are given by solid circles.

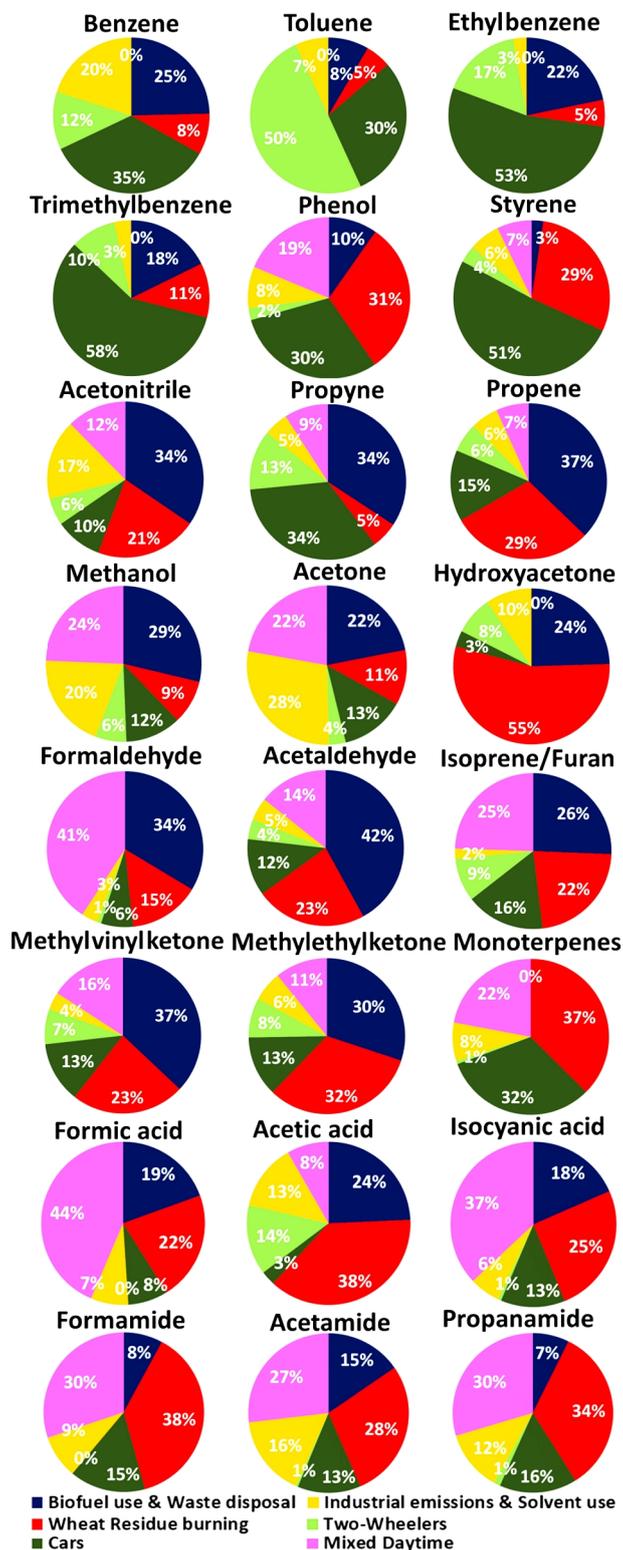


Figure 6. Contribution of individual PMF derived source factors to the total mass of different VOCs.

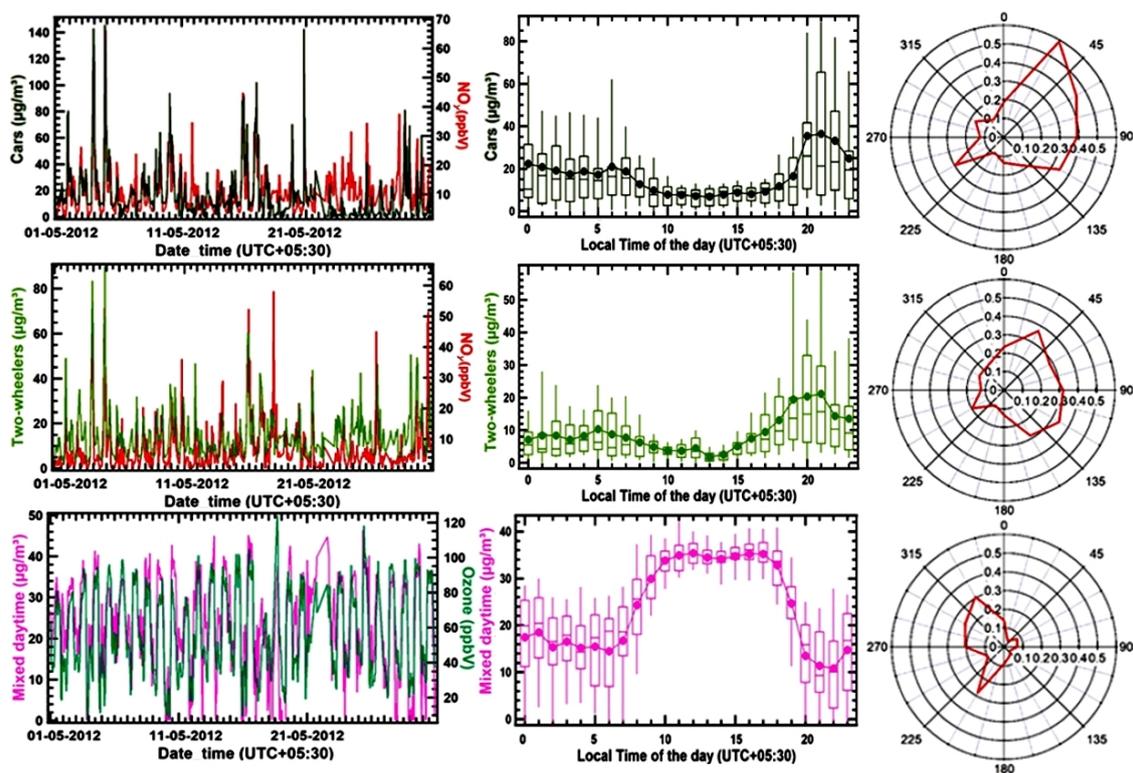


Figure 7. Factor contribution time series, factor diel variability and CPF plot for PMF Factor 4 and Factor 5 (Cars and two-wheelers) and PMF Factor 6 (Mixed daytime) for May2012. The time series of PMF factor’s hourly mass in $\mu\text{g m}^{-3}$ is plotted against independent tracer species NO_y (in ppbv) for the car and two-wheeler factor and O_3 (in ppbv) for the mixed daytime factor. The Diel box and whisker plot shows the statistical parameters of factor’s hourly mass contribution in $\mu\text{g m}^{-3}$ for every hour of the day plotted against the start time of the hour. The width of the box gives 25th and 75th percentiles, 50th percentile partitions the box; whiskers represent 10th and 90th percentiles of the dataset and average values are given by solid circles.

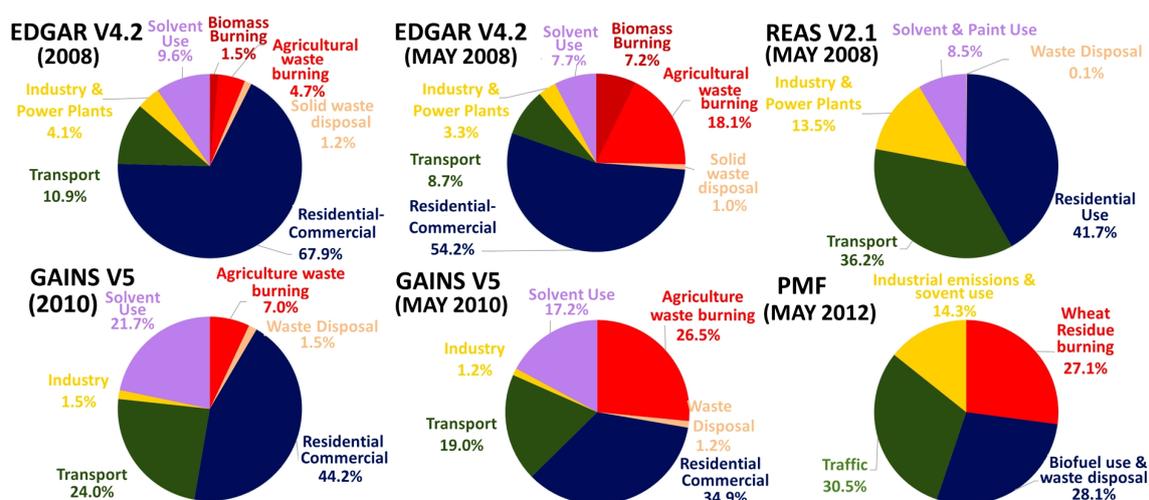


Figure 8. Comparison of PMF derived VOC source contribution to the EDGAR, REAS and GAINS Emission Inventory Database.