Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization

by Chinmoy Sarkar et al., 2017 (ACPD)

REFEREE 1:

Summary: The authors conducted VOC measurements using PTR-TOF in Kathmandu, Nepal. PMF was used to separate various source contributions to ambient VOC as a function of time. The authors then used a PMF "nudging" tool and some a priori knowledge of source profiles to move the PMF solution into a more physically realistic space. The various PMF factors are identified by comparing their VOC composition with known sources, and their diurnal behavior. The contribution of each PMF factor to total VOC mass is compared to comparable sources in several emissions inventories. The PMF derived source contributions are quite different from the emissions inventories, which are also quite different from one another. Contributions of the different sources to VOC mass, O_3 formation potential, and SOA potential are discussed. This is a thorough, detailed, well-written manuscript that provides valuable new information about an important, but underreported, region of the world. I highly recommend publication in ACP, if the following revisions are considered:

We thank the referee for appreciating and highlighting the importance of the work and for highly recommending the manuscript for the publication in ACP subject to revisions. We have found several of Referee 1's comments and suggestions very helpful and these are now reflected in the revised submission (changes are specified in replies and manuscript version with "tracked changes" given at the end of the responses here).

General comments: A general concern is that some important information about the data and PMF implementation are missing. I understand that many of the details of data collection and quality assurance have been published elsewhere (Sarkar et al., 2016). Nonetheless, this paper should be able to stand on its own. Several basic pieces of information should be included. For example: - a small map showing the geographical context of the measurement site - a brief description of the instrument & its measurement capability. - a list of the 37 ions used as PMF input, and the reasoning behind their selection. On a similar note, in many locations values for PMF input or assessment are given, but there is no explanation for why these particular values are chosen. (For example: bootstrapping factor assignment with R>0.6; Line 219- Why this particular length of time?). Could you provide a reason for selecting these values, or a literature citation?

We appreciate the reviewer's comment and suggestions. As per the referee's suggestions, we have modified the Materials and Methods (Section 2), by adding two new subsections and moving Section 2.3 pertaining to collection of grab samples, to precede the existing Section 2.1: Positive Matrix Factorization. The two new sections contain a description of the measurement site (as revised Section 2.1) and a description of the PTR-TOF-MS instrument as well as the list of 37 ions measured (as revised Section 2.2) in the revised manuscript. The existing Section 2.1 is now re-numbered as Section 2.4 and other Sections remain as before:

PMF model (as section 2.5) and the conditional probability analyses (as section 2.6). Lines 175-185 were shifted from the previous section 2.2 to the new section 2.1

Thus, the new Sections are as follows:

2.1 Site Description

NMVOC measurements during this study were performed in the winter season from 19 December 2012 until 30 January 2013 at Bode (27.689^{o} N, 85.395^{o} E, 1345 m a.m.s.l.) in Bhaktapur district, which is a suburban site located in the westerly outflow of the Kathmandu Metropolitan City. The land use in the vicinity of the measurement site consisted of the following cities – Kathmandu Metropolitan City (~ 10 km to the west), Lalitpur Sub-Metropolitan City (~ 12 km south west of the site) and Bhaktapur Municipality (~ 5 km south-east of the site). The site is located in the Madhyapur Thimi Municipality. In addition, the region north of the site had a small forested area (Nilbarahi Jungle; ~ 0.5 km^2 area) and a reserve forest (Gokarna Reserve Forest; ~ 1.8 km^2 area) at approximately 1.5 km and 7 km from the measurement site, respectively. Several brick kilns were located in the south-east of the site within a distance of 1 km. Major industries were located mainly in the Kathmandu and Patan cities whereas Bhaktapur industrial estate was located at around 2 km from the measurement site (in the south-eastern direction). A substantial number of small industries were also located in the south-eastern direction. The Tribhuvan International Airport is located about 4 km to the west of the Bode site. A detailed description of the measurement site and prevalent meteorology is already provided in the companion paper to this special issue (Sarkar et al., 2016). A zoomed view of the land use in the vicinity of the measurement site is provided in Figure 1.

2.2 PTR-TOF-MS measurements

NMVOC measurements were performed using a high-sensitivity PTR-TOF-MS (model 8000; Ionicon Analytic GmbH, Innsbruck, Austria) over a mass range of 21-210 amu. The PTR-TOF-

MS instrument works on the basic principle of soft chemical ionization (CI) where reagent hydronium ions (H_3O^+) react with analyte NMVOC molecules having proton affinity (P.A) greater than that of water vapour (165 Kcal/mol) to form protonated molecular ions (with m/z ratio = molecular ion + 1), enabling the identification of NMVOCs (Lindinger et al., 1998). As all the relevant analytical details pertaining to the PTR-TOF-MS instrument, ambient air sampling and the quality assurance of the NMVOC dataset has already been provided in detail in Sarkar et al., 2016, only a brief description of the ambient air sampling and the analytical operating conditions is provided here.

Ambient air sampling was performed continuously through a Teflon inlet line protected from floating dust and debris using an in-line Teflon membrane particle filter. The PTR-TOF-MS was operated at a drift tube pressure of 2.2 mbar, a drift tube temperature of 60^{o}C and a drift tube voltage of 600 V which resulted in an operating E/N ratio of ~ 135 Td (E = electrical field strength in V cm^{-1}; N = buffer gas number density in molecule cm^{-3}; 1 Td = 10^{-17} V cm^{-2}). Identification of several previously unmeasured and rarely measured NMVOCs were achieved due to the high mass resolution ($m/\Delta m > 4000$) and low detection limit (few tens of ppt) of the instrument. For the quality assurance of the measured NMVOC dataset, the instrument was calibrated twice during the measurement period and regular instrumental background checks were performed using zero air at frequent intervals. Detailed description of the sensitivity characterization of the instrument and the quality assurance of the primary dataset is available in Sarkar et al., 2016.

During the measurement period, a total of 37 NMVOC signals (m/z) were observed in the PTR-TOF-MS mass spectra that had an average concentration of > 200 ppt. The cut-off of an average concentration of > 200 ppt was employed keeping in mind the highest instrumental background signals observed during the campaign, so as to have complete confidence that the ions signals were attributable to ambient compounds. For mass identifications at a particular m/z ratio, further quality control was applied. Firstly, only those ion peaks were considered for the mass assignments for which there were no contribution from the major shoulder ion peaks within a mass width bin of 0.005 amu. Next, ion peaks devoid of any variability (that is the time series profile was flat) were not considered for mass assignments at all. Further details including some known interferences that were identified and taken into account are available in Sarkar et al., 2016. Table S1 in the supplementary information lists the identified 37 NMVOCs the corresponding m/z attributions (with references to few previous works which reported the same compound assignment, wherever applicable), and the elemental molecular formula.

Table S1. Most likely identity of VOCs (having average mixing ratios > 0.2 ppb) detected at specific protonated m/z ratios, molecular formula, likely mass assignment, reference of previous mass assignment, sensitivity, limit of detection (LOD), average ambient mixing ratios ($\pm 1\ \sigma$)

Protonated m/z or ion	Formula	Most Likely Identity	References of some previously reported studies	Sensitivity (ncps/ppb)	LOD (ppb)	Average (sdev) mixing ratio (ppb)
28.007	HCN	Hydrogen Cyanide	Stockwell et al., 2015; Karl et al., 2003	18.48	0.241	1.56 (0.24)
31.018	HCHO	Formaldehyde	Inomata et al., 2010; Stockwell et al., 2015	18.88	0.103	1.78 (0.50)
33.034	CH_3OH	Methanol	Seco et al., 2011; de Gouw et al., 2003	19.16	0.090	7.42 (1.28)

41.039	C_3H_4	Propyne	Akagi et al., 2011; Stockwell et al., 2015	7.167	0.080	7.67 (1.80)
42.034	CH_3CN	Acetonitrile*	Seco et al., 2011; de Gouw et al., 2003	20.91	0.043	1.08 (0.38)
43.055	C_3H_6	Propene	Stockwell et al., 2015; Park et al., 2013	7.45	0.048	3.98 (1.21)
44.014	NHCO	Isocyanic acid	Warneke et al., 2011	20.64	0.067	0.90 (0.08)
45.033	C_2H_4O	Acetaldehyde*	De Gouw et al., 2003; Seco et al., 2011	20.04	0.262	8.81 (4.58)
45.990	NO_2^+	Nitronium ion from fragmentation of C1-C5 alkyl nitrates	Aoki et al., 2007	20.91	0.094	1.08 (0.24)
46.029	CH_3NO	Methanamide		20.91	0.093	0.76 (0.16)
47.013	CH_2O_2	Formic acid	Jordan et al., 2009; Williams et al., 2001	21.04	0.041	4.96 (1.02)
47.049	C_2H_6O	Ethanol	Park et al., 2013; Seco et al., 2011	21.05	0.361	1.59 (0.85)
51.044	C_4H_2	1,3-Butadiyne$^{\$}$	Yokelson et al., 2013	8.56	0.013	0.67 (0.14)
56.060	C_3H_5N	Propanenitrile$^{\$}$	Yokelson et al., 2013	22.27	0.022	0.21 (0.05)
57.034	C_3H_4O	Acrolein* + Methylketene	Stockwell et al., 2015; Jordan et al., 2009	22.26	0.034	0.80 (0.26)
59.049	C_3H_6O	Acetone* + Propanal	de Gouw et al., 2003; Seco et al., 2011	23.47	0.074	4.21 (0.65)
60.051	C_2H_5NO	Acetamide		22.80	0.069	0.39 (0.05)
61.027	$C_2H_4O_2$	Acetic acid	de Gouw et al., 2007; Stockwell et al., 2015; Seco et al., 2011	22.94	0.440	4.24 (1.21)
62.026	CH_3NO_2	Nitromethane$^{@}$	Inomata et al., 2014; Akagi et al., 2013	23.07	0.020	0.24 (0.08)
63.026	C_2H_6S	Dimethyl Sulfide	Akagi et al., 2011; Park et al., 2013	23.21	0.049	0.26 (0.03)
67.054	C_5H_6	1,3-Cyclopentadiene	Stockwell et al., 2015	10.78	0.008	0.23 (0.06)
69.033	C_4H_4O	Furan	Stockwell et al., 2015; Jordan et al., 2009	24.02	0.009	0.46 (0.17)
69.070	C_5H_8	Isoprene*	Stockwell et al., 2015; de Gouw et al., 2003; Seco et al., 2011	10.02	0.013	1.11 (0.24)
71.049	C_4H_6O	Methyl vinyl ketone; Methacrolein; Crotonaldehyde*	Seco et al., 2011; Stockwell et al., 2015; de Gouw et al., 2007	27.17	0.017	0.35 (0.10)
73.027	$C_3H_4O_2$	Methylglyoxal	Stockwell et al., 2015; Muller et al., 2012	24.56	0.021	0.31 (0.10)
73.063	C_4H_8O	Methyl ethyl ketone*	de Gouw et al., 2003; Stockwell et al., 2015; Park et al., 2013	21.91	0.036	0.69 (0.12)
75.042	$C_3H_6O_2$	Hydroxyacetone	Christian et al., 2003; Heigenmoser et al., 2013; Stockwell et al., 2015	24.83	0.066	0.63 (0.18)
79.054	C_6H_6	Benzene*	Jordan et al., 2009; de Gouw et al., 2003	13.43	0.013	2.71 (1.17)
83.085	C_6H_{10}	Assorted Hydrocarbons	Stockwell et al., 2015	13.01	0.008	0.45 (0.09)
87.042	$C_4H_6O_2$	2,3-Butanedione	Stockwell et al., 2015; Karl et al., 2007	26.45	0.028	0.35 (0.08)
93.070	C_7H_8	Toluene*	Seco et al., 2011; Jordan et al., 2009	15.78	0.006	1.53 (0.38)
97.031	$C_5H_4O_2$	2-Furaldehyde (furfural)	Ruuskanen et al., 2011; Liu et al., 2012; Li et al., 2013	27.80	0.010	0.26 (0.07)
97.102	C_7H_{12}	Assorted Hydrocarbons	Stockwell et al., 2015	14.96	0.006	0.23 (0.05)
105.070	C_8H_8	Styrene	Jordan et al., 2009; Stockwell et al., 2015	16.07	0.004	0.21 (0.08)
107.086	C_8H_{10}	Xylenes*	Jordan et al., 2009; Stockwell et al., 2015	15.36	0.004	0.97 (0.27)
121.101	C_9H_{12}	Trimethylbenzenes	Muller et al., 2012; Jordan et al., 2009	18.30	0.004	0.38 (0.10)
129.070	$C_{10}H_8$	Naphthalene	Jordan et al., 2009; Stockwell et al., 2015	19.40	0.009	0.33 (0.09)

* VOC sensitivities determined using VOC gas standards in calibration experiments
$ Observed mass accuracy for 1,3-Butadiyene and Propanenitrile were 21 mDa and 10 mDa, respectively
@ Corrected for the ^{13}C isotopologues of acetic acid

Lindinger, W., Hansel, A., and Jordan, A.: On-line monitoring of volatile organic compounds at pptv levels by means of proton transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research, Int. J. Mass Spectrom., 173, 191–241, doi:10.1016/s0168-1176(97)00281-4, 1998.

A selection of the Pearson's coefficient R > 0.6 has been recommended by Norris et al. (2008) in the EPA PMF v3.0 user manual and more recently the same suggestion has been repeated in the user manual of the v5.0 (Norris et al. 2014). The recommendation has been generally adhered to by other authors using the same software e.g. Baudic et al., 2016.

We have now included the citation to Norris et al., 2008 and 2014 in the revised manuscript.

Norris, G., Vedantham, R.,Wade, K. S., Brown, S. G., Prouty, J. D., and Foley, C.: EPA positive matrix factorization (PMF) 3.0 fundamentals and user guide. Prepared for the US. Environmental Protection Agency, Washington, DC, by the National Exposure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma, CA; and Lockheed Martin Systems Engineering Center, Arlington, VA, EP-D-05-004; STI-907045.05- 3347-UG, October, 2008.

Norris, G., Duvall, R., Brown, S., and Bai, S.: EPA Positive Matrix Factorization (PMF) 5.0: Fundamentals & User Guide, Prepared for the US, Environmental Protection Agency (EPA), Washington DC, by the National Exposure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma, 2014.

Previously the sentence was (P5, L132-135):

"The model output of each bootstrap run is mapped onto the original solution using a cross correlation matrix of the factor contributions g_{ik} of a given bootstrap run with the factor contributions g_{ik} of the same time segment of the original solution using a threshold of the Pearson's correlation coefficient (R) > 0.6."

The revised sentence now reads as:

"The model output of each bootstrap run is mapped onto the original solution using a cross correlation matrix of the factor contributions g_{ik} of a given bootstrap run with the factor contributions g_{ik} of the same time segment of the original solution using a threshold of the Pearson's correlation coefficient (R) > 0.6, as suggested by Norris et al., 2008 and 2014."

Previously, P7, L219 was:

"To identify the uncertainty associated with the PMF solution, bootstrap runs were performed 100 times taking 96 hours as the segment length."

The revised sentence now reads as :

"To identify the uncertainty associated with the PMF solution, bootstrap runs were performed 100 times taking 96 hours as the segment length. This is slightly shorter than the recommended length based on the equation of Politis and White (2004), of 108 hours but represents a multiple of 24 hours and hence ensures each bootstrap run contains four full days' worth of data."

Politis, D. N. and White, H.: Automatic block-length selection for the dependent bootstrap, Econometrics Reviews, 23, 53–70, doi:10.1081/ETC-120028836, 2004.

Specific comments: Section 2.1 This section needs some minor reorganization. A brief description of the PTR-ToF measurement should come first, then the description of the grab-sampling, then the PMF implementation. I suggest this because the PTR-ToF measurements, and the grab-samples, are referred to several times in the discussion of PMF; however, they had not yet been introduced.

Done.

We have now reorganized the Materials and Methods section (section 2) as per the suggestion of the referee in the earlier comment, put new captions for the description of the PTR-ToF measurements and the grab-sampling and shifted the text to the relevant sections

Line 127 I found this sentence very hard to parse. Perhaps you can break this down to provide a clearer explanation of the information provided by bootstrapping.

Done.

Previously the sentence was (P5, L127):

"To ascertain the magnitude of random errors that can be caused due to the use of random seeds followed by the selection of the run with the lowest Q due to the existence of infinite solutions with different g_{ik}, f_{kj} and e_{ij} matrices but identical $Q = \Sigma_{i=1}^{n}\Sigma_{j=1}^{m}({}^{e_{ij}}/_{\sigma_{ij}})^2$, bootstrap runs were performed."

The revised sentence now reads:

"Bootstrap runs were performed to ascertain the magnitude of random errors of the dataset (Norris et al. 2014, Paatero et al. 2014). Random errors can be caused due to the existence of infinite solutions with different g_{ik}, f_{kj} and e_{ij} matrices but identical $Q = \Sigma_{i=1}^{n}\Sigma_{j=1}^{m}({}^{e_{ij}}/_{\sigma_{ij}})^2$."

Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, Atmos. Meas. Tech., 7, 781–797, doi:10.5194/amt-7-781-2014, 2014.

Lines 127-139 This section could benefit from literature citations describing the use of bootstrapping. I'd also like to see a citation supporting your assertion that a fraction <20% of unmapped factors indicates low random error.

The statement that a solution in which >80% of the bootstrap runs are successfully mapped onto the same factor can still be considered a relatively stable solution can be found in Norris et al. (2014) in several places (e.g. discussion of table 9). The manual recommends that when the number of unmapped factors (= bootstrap factors that could not be mapped onto any one of the source profiles) is high, the users should carefully investigate which observations/outliers have a disproportionate influence on the factor profiles but could also explore lowering the threshold below 0.6. However, this is only of academic interest, since our model runs have no unmapped factors, while retaining the recommended threshold of 0.6. We have included a citation and elaborated further. Line 138ff now reads:

"The presence of a high fraction unmapped factors (> 20%) is a clear indication of large random errors (introduced by a few critical observations that drastically impact factor profiles) and should be investigated carefully (Norris et al. 2014). In our analysis, no unmapped factors were present."

Lines 155-160 This explanation of rotational ambiguity is a little convoluted. Can you rephrase to make this easier to follow? See Ulbrich et al. (2009) as an example.

Done.

Previously it was (P5, L155-160):

"In addition to the random error, the PMF model also has rotational ambiguity. There can be multiple solutions with a different factor profile for all factors for which the model will find a different local minimum of the residual matrix while determining the factor contribution matrix. This fact that different solution for $g_{ik}f_{kj}$ with the same sum of the scaled residuals $Q = \Sigma_{i=1}^{n}\Sigma_{j=1}^{m}({}^{e_{ij}}/_{\sigma_{ij}})^2$ exist, is called the rotational ambiguity of the model."

In the revised manuscript, we have included the reference of Ulbrich et al., 2009 and modified the explanation of rotational ambiguity as follows:

"In addition to the random error, the PMF model also has rotational ambiguity (Ulbrich et al., 2009, Paatero et al. 2014). This rotational ambiguity is caused due to the existence of multiple solutions which have a Q similar to the solution produced by the PMF model but different factor profiles and factor contributions. Thus, the model will find different local minima of the residual matrix, while determining the factor contribution matrix ($g_{ik}f_{kj}$). The coexistence of different solutions for the factor contribution matrix ($g_{ik}f_{kj}$) with the same sum of the scaled residuals ($Q = \Sigma_{i=1}^{n}\Sigma_{j=1}^{m}({}^{e_{ij}}/_{\sigma_{ij}})^2$) is called the rotational ambiguity of the model."

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891-2918, 10.5194/acp-9-2891-2009, 2009.

Line 192 What does it mean to be classified as a "weak" vs. a "strong" species?

Weak species are considered to have a larger uncertainty. The specified uncertainty in the uncertainty file is tripled for weak species while for strong species the uncertainty is retained as is. This reduces the influence of the weak species on the factor profiles as it reduces the magnitude of the uncertainty scaled residual and hence the contribution to Q.

We have clarified this P6, L191-197 of the ACPD version of the manuscript (changes bold):

"Due to its erratic timeseries profile, HCN (m/z = 28.007) was classified as a weak species in the PMF input while all other ions were classified as strong species. **For weak species, the stated uncertainty is tripled, to reduce their impact on the scaled residual and hence Q.**"

Line 193 The conversion from mixing ratio to mass concentration introduces additional variability due to meteorology. Additionally, higher molecular weight species now have more "pull" in the PMF solution, because their signals are higher. Do you see evidence for meteorological influence in the PMF solution? Any evidence of a bias towards explaining variability of heavy species, at the expense of light VOCs? Why did you choose to run PMF on the mass concentration data rather than mixing ratio data?

We agree that the conversion of mixing ratio to mass concentration introduces additional uncertainty, however, the conversion was done using the relevant temperature, pressure and molecular weight and the uncertainty thus introduced is already accounted for by running the model with 5% measurement uncertainty. It is desirable and recommended to run the PMF after converting to mass concentrations (Norris et al. 2014). Only this conversion allows mass closure and hence the preparation of pie charts quantifying the source contributions (Figure 13; Section 3.3), which can be compared to emission inventories (Figure 14; Section 3.3).

It should be noted that for each observation the effect of each species residual on Q is scaled by the uncertainty of that very same species and observation. Therefore, higher mass does not at all result in a higher pull. A residual that is high compared to the measurement uncertainty of that specific observation exerts a pull. A residual that is negligible when compared to the measurement uncertainty of that specific observation has no influence on the output.

As described in Paatero et al. (2014), the highest "pull" on the PMF solution is due to species with a strong presence in some observations (in our case factor >4 enhancement) above the median mass loading, which is contrasted by low values (in our case less than 1/3 of the median) at other times. A strong presence-absence contrast most strongly defines source profiles.

It can be seen in the factor profiles and the sequence in which factors appear in the PMF output, that compounds with a low molecular mass are equally or more important in defining for the PMF solution space compared to higher mass species. Table S3 shows the percentage contribution of PMF derived factors obtained from constrained runs with 5-, 6-, 7-, 8- and 9-Factors. The biogenic factor (which is primarily defined by biogenic isoprene and acetaldehyde emissions) splits from the mixed daytime factor in the 5-factor solution, and does so, before the traffic source (defined primarily by benzene, toluene, xylenes and trimethylbenzens) splits from mixed combustion sources in the six factor solution.

We have rephrased L193ff" "All the input data was converted from mixing ratios ppb to mass concentrations (µg m^{-3}) using the relevant temperature, pressure and molecular weight. The total measured NMVOC concentration was calculated by adding the mass concentrations of all measured NMVOCs and was classified as a weak species in the PMF input."B

The text now reads: "All the input data was converted from mixing ratios of ppb to mass concentrations (µgm−3) using the relevant temperature, pressure and molecular weight and the total measured NMVOC concentration was calculated by adding the mass concentrations of all measured NMVOCs. This conversion allows calculating the explained variability (Gaimoz et al. 2011) for the total VOC mass and comparing the results with emission inventories. The conversion does not introduce significant additional uncertainty and the variability induced by the temperature (average range observed was: 5-20 ºC), has largely been taken into account by running the model with a 5% extra modelling uncertainty. The total VOC mass was classified as a weak species in the PMF input (Norris et al. 2014)."

Lines 200-220 On my first read-through, the presentation of the eight factors was very sudden and it wasn't at all clear to me how their identifications were derived, or that more information would be provided later in the text.

This section mainly focuses on the description of the implementation of the PMF model to the VOC dataset and therefore in P7, L215-218 (of the ACPD version of the manuscript), only the names of the factors are mentioned which were derived from the diagnostics of the 8-factor solution of the PMF run. A detailed description of how the identification and attribution of these PMF derived 8-factors were performed is described in detail in section 3.1 of the Results and Discussion (section 3). However, we have now added a sentence after L218 (of the ACPD version):

"A detailed description for the identification and the attribution of the 8-factor solutions is provided later in section 3.1."

It would be very helpful to see a plot of Q/Qexpected as a function of number of factors, and an additional plot showing what happens to each of your identified factors as the number of factors is changed (perhaps in the supplemental information).

We have added a figure showing the % change in $Q/Q_{expected}$ when the number of factors is increased for all solutions starting from a 3 factor solution (in absolute terms $Q/Q_{expected}$.<1 even for a 3 factor solution). This dataset is highly unusual, due to the fact that wind speed and wind direction have a strong diel cycle throughout the campaign due to mountain meteorology. In this situation all traditional indicators for the quality of a PMF solution fail. We have also added a time series of the modelled mass and the measured mass, which clearly shows perfect mass closure even with a 3 factor solution already. A 3 factor solution distributes compounds into the following factors: Factor 1: "higher during the first part of the campaign no diel profile", Factor 2: "higher during the second part of the campaign and higher at night and Factor 3 "higher during the second part of the campaign and higher during the day". Mass closure, however, does not mean this 3 factor solution is plausible and corresponds to real world source. It merely means that all mathematical quality indicators typically used fail for this dataset. Instead the plausibility of each possible solution must be carefully assessed keeping in mind

the auxiliary information. However, the last unusually high drop in $Q/Q_{expected}$ is seen when the number of factors is increased to 8. Beyond that the relative change stays constant.

We have modified the statement in line 202 of the ACPD manuscript to make this clearer, instead of previously:
"Based on the Q/Qtheoretical ratio, the physical plausibility of the factors and the rotational ambiguity of the solution, an 8-factor solution was deemed to be the best for this dataset."
The text now reads"
"Based on the $Q/Q_{theoretical}$ ratio, the physical plausibility of the factors and constraints imposed by rotational ambiguity of the solution, an 8-factor solution was deemed to be the best for this dataset. For the data presented in this study, the $Q/Q_{theoretical}$ ratio is <1 even for a 3 factor solution with no physical plausibility and hence the absolute number does not help to decide the optimum number of factors. Supplementary Figure S2 shows clearly, that the number of factors has almost no impact on how well the total mass is reproduced by the model but the last distinct drop in the $Q/Q_{theoretical}$ ratio is seen when the number of factors is increased to 8."

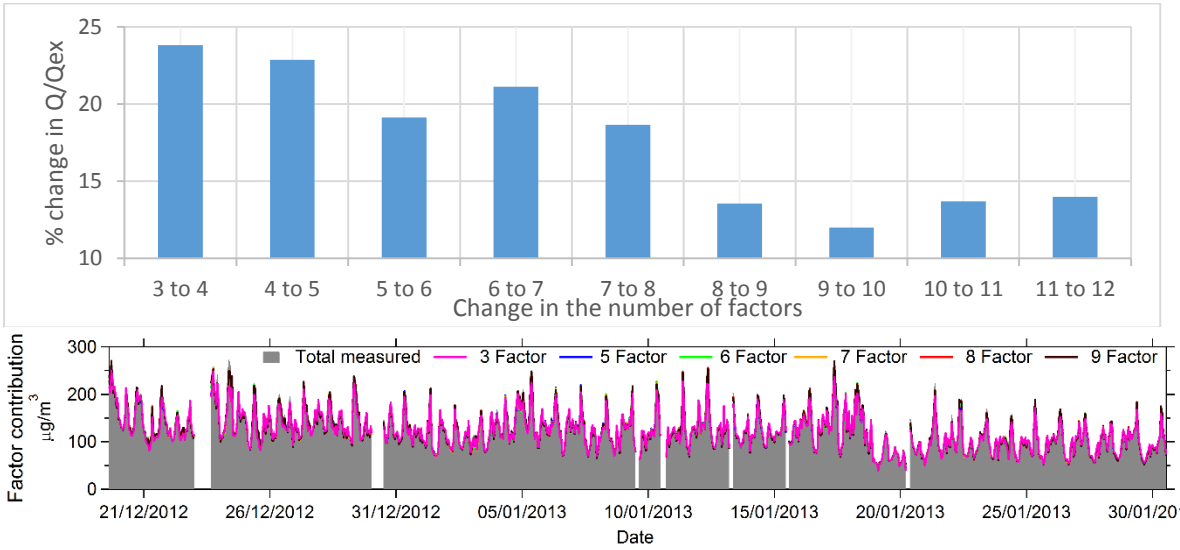


Figure S2. Relative change in the Q/Qexpected ratio with change in factor number (top) and time series of the total measured VOC mass (grey filled) and the modelled VOC mass for different number of factors in the PMF solution (bottom).

Regarding the variation of PMF derived factors with number of factors, we have already provided a Table in the supplementary information (Table S2) we do not consider it necessary to replace this table with a Figure containing the same information. However, we have added a supplementary Figure S3 showing how factor profiles, the mass of each species explained by each factor and the factor contributions evolve from the 5 Factor solution to the 9 Factor solution.

In the factor profiles it can be seen that individual compounds move to the newly generated factor from several previous profiles when the number of factors is increased. More specifically, aromatics and several other compounds from a mixed combustion factor, brick kiln and residential burning factor move into two new profiles, the traffic factor and mixed industrial emission factor when the number of factors allowed is increased to 6. When the number of factors is increased to 7 Aldehydes and acids, previously distributed among the residential biofuel and waste disposal factor, brick kiln factor, mixed daytime and the biogenic

factor move to the solvent evaporation factor. At the same time, the time series of the mixed industrial factor shows, that brick kiln emissions which got pushed into the mixed industrial factor while compromising the brick kiln source profile to accommodate more acetaldehyde and acetic acid, shift to the brick kiln factor. When the number of factors is increased to 8 several compounds previously accommodated in the mixed industrial, solvent evaporation and mixed daytime factor, most notably 1,3-butadiyne, shift to the new profile. The eight factor solution is the first one, where the factor contribution for residual night-time primary emissions in all daytime factors drop below 10 µg/m^{3}for the full period.

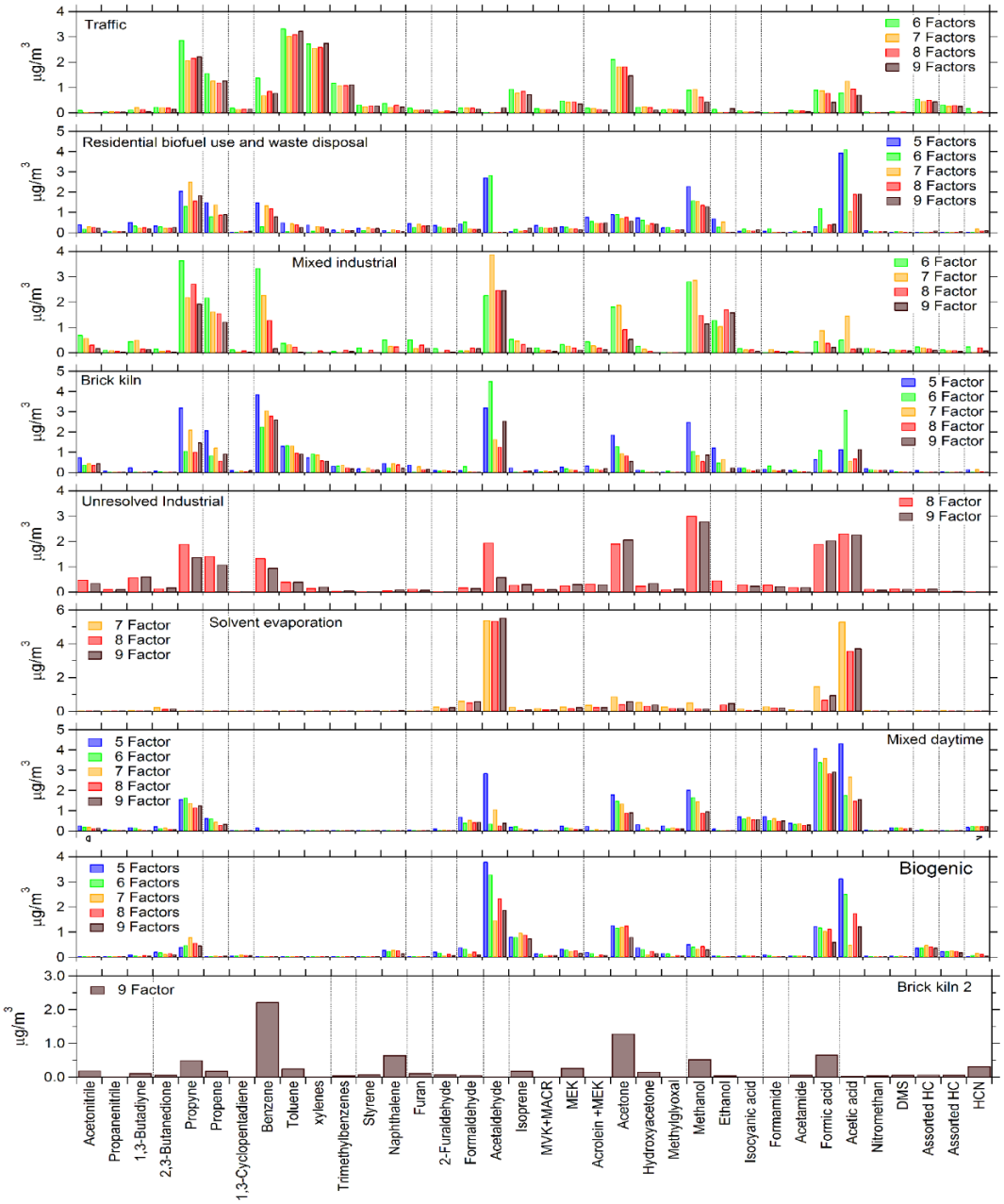


Figure S3a. Evolution of the factor profiles of the eight sources identified and the 9th source which is considered to arise due to splitting of the brick kiln factor from the 5 Factor to the 9 Factor solution.

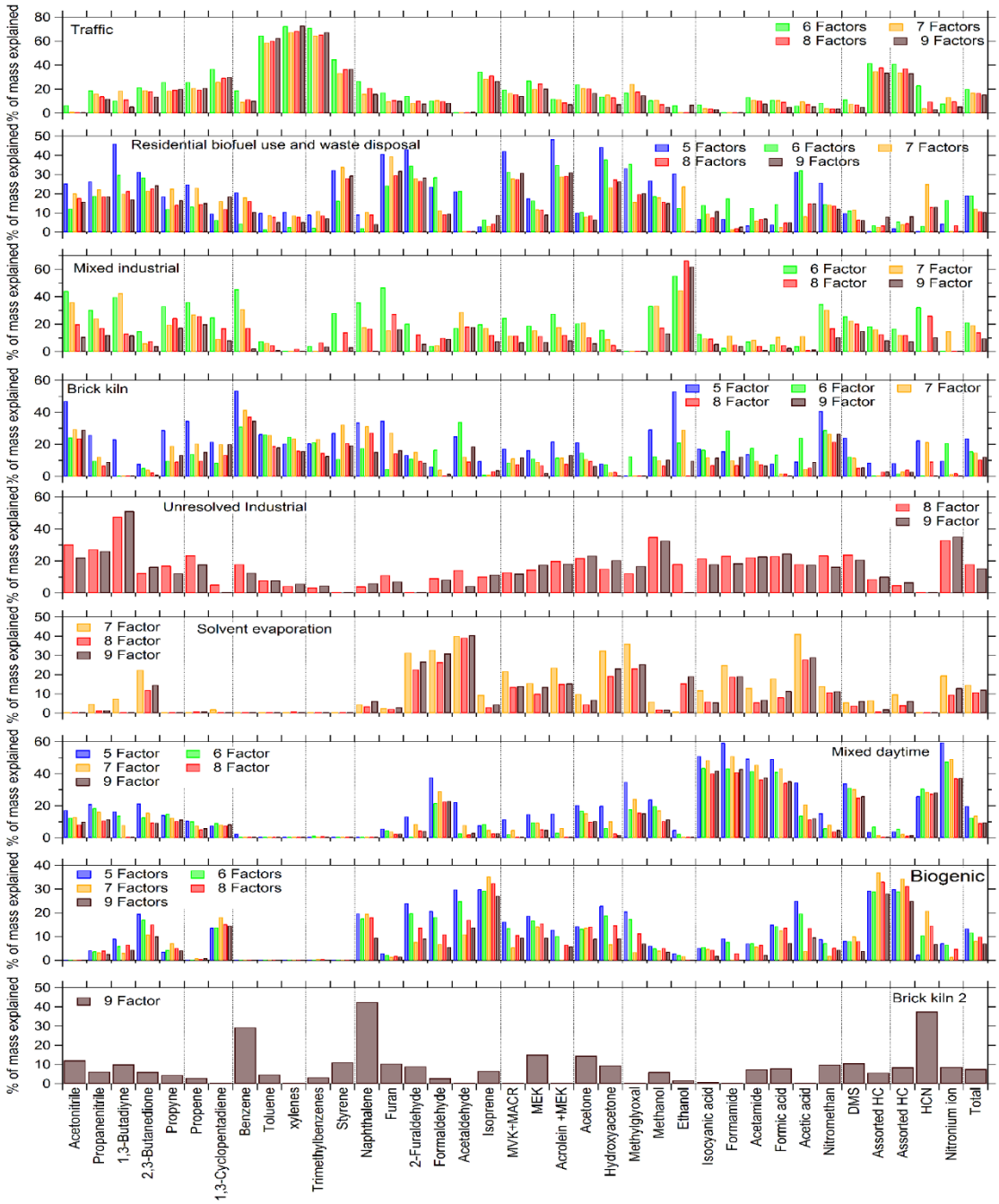


Figure S3b. Evolution of the percentage of the mass of each compound explained by the eight sources identified and the 9th source which is considered to arise due to splitting of the brick kiln factor from the 5 Factor to the 9 Factor solution.

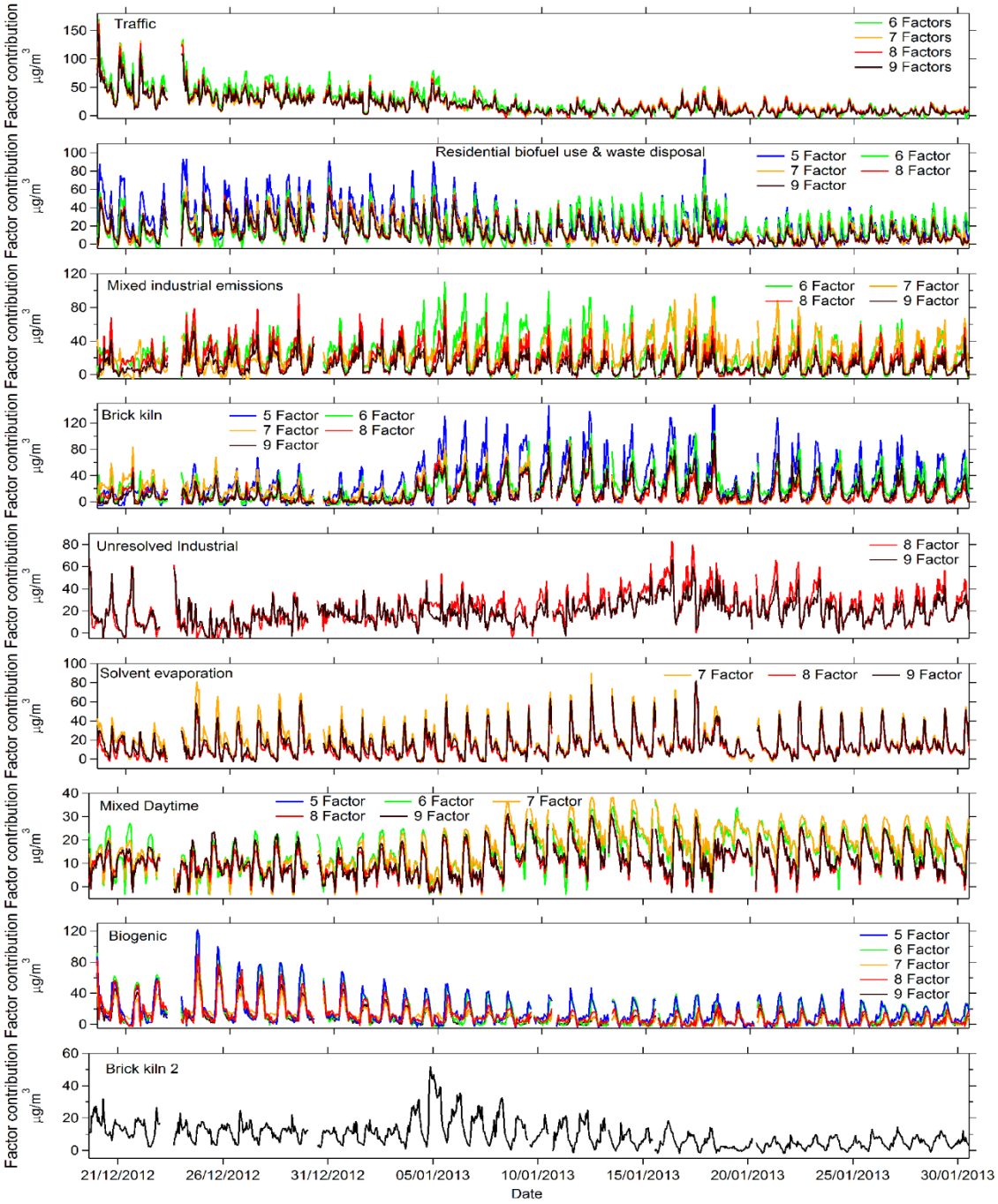


Figure S3c Evolution of the factor contribution of the eight sources identified and the 9th source which is considered to arise due to splitting of the brick kiln factor from the 5 Factor to the 9 Factor solution.

Eight factors is a quite large number compared to PMF solutions reported in many other studies. Without strong supporting evidence for each factor, I find it hard to believe that PMF can robustly separate this many distinct sources. It is especially hard to believe when each factor, on average, accounts for 12.5% of signal – but you have stated that the overall measurement uncertainty is 20% (line 190). Any additional information that you can provide to show that an 8 factor solution is physically plausible would be very helpful in convincing the reader of your solution. This could include additional PMF diagnostics (as

suggested above) or other information. For example, it seems that prior to running PMF, you had some idea of what important emission sources to expect, perhaps from your previous paper or from an emissions inventory (I see four listed in the introduction). Can you make a stronger connection between the a priori knowledge and the PMF solution?

We thank the referee for this suggestion. It is true that we had some prior information and we have made a stronger connection in the text now. Before we performed the source apportionment using PMF model to the NMVOCs measured in the Kathmandu valley, we undertook a thorough analysis of the primary NMVOCs dataset as reported in our companion paper to this Special issue Sarkar et al., 2016. In addition information about local sources and emission activity periods (e.g. operation of brick kilns, change in meteorological conditions, was provided by co-authors. Finally we also relied on previous studies and information from existing emission inventories, for a good background. We summarize these below:

1) Biogenic emission sources (characterized by high daytime concentrations of isoprene emitted due to the presence of nearby forested areas). The biogenic factor is strongly defined by emissions from deciduous trees, which were present and responded quickly to changes in solar radiation during the first two weeks of the campaign (See Sarkar et al. 2016 and the response to reviewer #2 for a plot demonstrating how rapidly the vegetation responded to changes in cloud cover). Since the deciduous trees shed their leaves during the second part of the campaign and biogenic emissions from evergreen trees and needle leaved trees are much lower during that time, the overall penalty on Q for combining biogenic emissions with traffic emissions and residential burning into one "higher during the first part of the campaign" factor with no clear diel profile is low.

2) Biomass co-fired brick kilns emissions (characterized by the emissions of high concentrations of benzene from the nearby brick kilns and its excellent correlation with acetonitrile, presumably due to co-combustion of crop residue and other biomass). Brick kilns, were not operational during the first part of the campaign. This perturbation is sufficient to confidently separate their emissions from combustion emissions of other industrial units, which continued their operations throughout our study time period. At the same time all the industrial sources including brick kilns are spatially co-located and follow similar temporal patterns. Therefore, brick kilns and other industrial emissions can be combined into one "high during the second part of the campaign and high night-time emission factor" with a relatively low penalty on Q. The price to pay is that certain compounds that are high in the mixed industrial emissions but not in the brick kiln emissions are pushed into the residential burning factor.

3) Biomass/residential burning emission sources (characterised by the presence of the high concentrations of several oxygenated VOCs, acetonitrile and aromatic compounds) were expected to be an important source of VOCs based on a priori knowledge from the existing emission inventories (for the REAS inventory, this is the single most important source of VOCs).

4) Traffic emission sources (characterized by the presence of high concentrations toluene, C8- and C9-aromatics), were considered to be extremely important based on earlier studies conducted in the Kathmandu valley (Shrestha et al. 2013). However, since the measurement site was rarely downwind of Kathmandu valley during evening traffic rush hours, the traffic factor is defined by a few strong plumes with mass loadings of ~120 $\mu g/m^3$. As a consequence,

the overall impact of the traffic source on Q is low and the emissions from this source can be combined with the residential burning factor with a relatively low penalty on Q.

5) Mixed daytime/photochemical sources are characterized by the presence of compounds such as isocyanic acid during daytime which is produced as a result of photooxidation of precursor amides. This source was already demonstrated to be an important source for several compounds in the companion paper. Since photochemical formation of secondary pollutants is clearly important in the Kathmandu valley it is highly desirable to segregated this secondary pollution from primary emissions which proves to be challenging in a less than 8 factor solution.

6) Contribution from industrial sources located in nearby industrial estates were expected and could be separated from brick kilns thanks to the fact that brick kilns were closed during the beginning of the campaign.

7) Solvent usage is an important source of VOCs according to several emission inventories and the single most important source according to the EDGAR emission inventory. Due to the foggy/hazy conditions at night this source has a very peculiar time series. Soluble compounds tend to partition into the fog/haze aqueous phase at night and rapidly partitions into the gas phase during morning hours. This is particularly true for compounds with a high temperature dependence of their solubility in water. This, however, means non-water soluble solvents and water soluble solvents do get separated by the PMF and the later land in the mixed industrial and unresolved industrial factor.

While based on the preliminary information we had prior to initializing the model it seemed best to run the model with only 7 factors the raw time series data (input data) strongly supports an 8 factor solution. We have added the following text to clarify this:

“The primary data strongly supports an 8 factor solution. The top 2-3 compounds explained by each of the factors have a much higher R when their input time series is correlated compared to the R obtained when their time series is correlated with the time series of any other compound (Supplementary table S5).

“The traffic explains more than 60% of the variability of Toluene, C-8 and C9 aromatics. The time series of Toluene, C-8 and C9 aromatics correlates with R >0.96 for all possible pairs when the original time series of these compounds are correlated with each other. The R of the time series of these same compounds with the time series of styrene is lower (0.81-0.85) while a correlation of their time series with all other compounds yields R <0.78. This indicates toluene, sum of C-8 and C9 aromatics share a major common source with each other which is not shared by other compounds, namely the traffic source. Hence a less than 6 factor PMF solution which is incapable of capturing the traffic source is not a better representation of the reality.

For styrene the highest correlation is with furan R=0.87 indicating that the two compounds have a significant source in common, which styrene also shares with higher aromatics and propyne (R=0.86), but the lower R of styrene with the aromatic compounds indicates that styrene has at least two dominant sources with distinct emission ratios. These sources are the traffic source (explaining roughly 40% of the styrene) and the residential burning source which explains 30% of the styrene and furan variability. These two sources are separated only with a 6 factor solution.

Benzene has a strong source in the form of biomass co-fired brick kilns which results in a distinct increase in emission at the time the brick kilns restart their operations. This source is shared with acetonitrile (R=0.89), nitromethane (R=0.82) and naphthalene (R=0.81) but all of these compounds also have other sources which are either not shared with benzene or have different emission ratios. This source appears in the 3 factor solution but its source profile is contaminated with mixed industrial emission. The closure period of brick kilns is only fully captured and restricted to the brick kiln factor after the number of factors is increased to 7.

The mixed industrial source explains 66% of the ethanol variability, but this compound has a relatively low R with all other compounds (0.73 with propene and 0.7 with nitromethane and acetonitrile <0.66 with the rest) indicating that there must be at least two distinct ethanol sources with different source fingerprints. A second distinct ethanol source in the form of solvent evaporation, however, separates from the mixed daytime factor only in the 7 factor solution.

The mixed daytime factor primarily contains photo-chemically formed compounds most notably isocyanic acid, which shows a strong correlation with its own precursors formamide (R=0.85) and acetamide (R=0.82). Figure S8 presents reaction schematic for the formation of formamide and isocyanic acid. This compound has a much weaker correlation with other compounds, which have other sources in addition to the photochemical source (R=0.5 to 0.58 for formaldehyde, acetaldehyde, the nitronium ion, formic acid and acetic acid). This factor should ideally be restricted to photo-chemically formed secondary compounds, however, it remains heavily contaminate with night-time primary emissions during the second half of the campaign till the number of factors is increased to 8 (Figure S3c). Even the 8 and 9 Factor solution still contain some minor contamination from primary emissions. Hence the name of the source is retained as “mixed daytime source”.

The solvent evaporation factor is characterised by acetaldehyde and acetic acid which have their strongest correlation with each other (R=0.82). Apart from this, the defining compound, acetaldehyde, shows moderate correlation with formaldehyde (R=0.72) and acetone (R=0.68) but only the former correlates with acetic acid (R=0.85) as it shares both the solvent evaporation source and the photo-oxidation source with acetaldehyde, while the later (acetone) correlates much stronger with methyl ethyl ketone (R=0.95) and methyl vinyl ketone (R=0.86) and isoprene (R=0.79) and hence shares the biogenic emission source in addition to the the solvent evaporation factor. While these three daytime sources are resolved in the 7 factor solution their source profiles continue to be contaminated with primary emissions. While the same can be pushed around from the biogenic factor into the mixed daytime factor using rotational tools, they cannot be sufficiently removed from all three till an 8^{th} factor is allowed.
The unresolved industrial emission factor explains a significant fraction of the 1,3-butadiyne which shares most source with methanol (R=0.9). The source profile also captures several other compounds with a lower correlation with 1,3-butadiyne including propanenitrile (R=0.86), acrolein + methylketene (R=0.82) and propene (R=0.8). The R for cross correlating the time series with that of ethanol, the defining compound of the mixed industrial source profile is only 0.73 and ethanol correlates only weakly with acrolein + methylketene (R=0.59) indicating that these mixed industrial emissions and unresolved industrial emissions represent distinct sources, which can only be resolved in an 8 factor solution. “

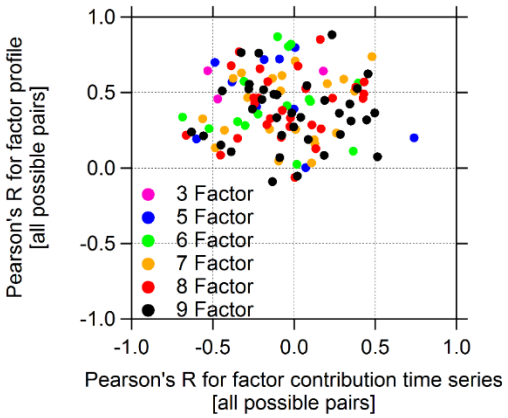
Line 204 You discarded solutions with 7 or fewer factors because there appeared to be "mixing" of sources. But, this could also be due to rotational ambiguity. Did you attempt to unmix solutions with ≤7 factors by exploring FPEAK, for example?

FPEAK does not help removing primary emissions from the daytime profiles with 7 factors only. The constraint mode achieves superior results when the two options are compared for an equal number of factors. It should be noted that FPeak (available in the previous version of the PMF) and the constraint mode (new feature of the 5.0 version) are two alternate rotational tools which cannot be used in combination. We opted for the constraint mode, which allows us to use source profile fingerprints from samples collected at the source to refine the solution. The constraint mode performs is superior , as it exploiting rotational ambiguity to push the solution into a physically more realist space using pre-existing knowledge and the user decides how much price in the form of a "higher than the local minimum Q" he/she is willing to pay for a solution that corresponds better to the real world. FPeak simply explores how much the solution can be changed due to rotational ambiguity and the user usually goes for the solution with minimum Q. It has recently been recognized that a minimum Q represents the mathematical minimum but the same does not correspond to the most plausible real world solution (Paatero et al. 2014). We have added a statement that FPeak and the constraint mode are two alternate rotational tools and cannot be combined with each other to line 222 of the ACPD version, which reads:

"The constraint mode is a new rotational tool introduced in the 5.0 version of the EPA PMF as an alternative to the FPeak module. The constraint mode allows to exploit the rotational ambiguity of the model to push the PMF solution into a physically more realistic space. It uses pre-existing knowledge such as source fingerprints, source emission ratios or activity data. We found that when the two modules were compared for an equal number of factors the constraint mode performance was superior to the FPeak module."

Line 226 Can you provide a scatterplot showing R^2 between each factor time series, vs R^2 between each factor profile?

We have added the plot to this response, however, we are showing R vs R, since R^2 can have a positive 1 both for perfect correlation and anti-correlation. The information provided by this plot seems to be more difficult to interpret compared to the information provided by the plot presently available in the supplementary material so we have retained the old plot.



It isn't clear to me if you are discussing correlation between time series or profile here, and it would help to see the "whole picture".

We are discussing the correlation between the time series. We have now clarified it…

"The original model output showed positive correlations between **the factor contribution time series** ~~factors such as~~ **of** the biomass co-fired brick kilns and mixed industrial emissions (R^2 = 0.27) f**actor** as well as the residential biofuel use and waste disposal factor with traffic factor (R^2 = 0.42)."

Lines 237-240 Not sure I understand this. Is "mixed daytime" a photochemistry tracer? Why wouldn't solvent evaporation contain acetonitrile or aromatics? Both acetonitrile and aromatics are commonly found in solvents. Why use a ratio to acetic acid as the nudging control?

Yes, "mixed daytime" is primarily a photochemistry tracer, although we were not able to completely remove primary emissions from the photochemistry source profile and, therefore, continue to call it "mixed daytime".

The primary problem is that once resolved as separate, the solvent evaporation factor with its sharp solubility driven peak is not at all amenable to accepting aromatic compounds, which are not water soluble, into its factor profile. The factor profile contains no acetonitrile and only <1% contribution from other aromatic compounds to start with. Instead, the PMF prefers to deposit these compounds with the photochemistry source or biogenic emissions. If constraints are only placed on these two factors, the PMF will simply execute a complete factor swap while running the bootstrap runs. It will shift all compounds that were in the biogenic factor to the solvent evaporation factor (which has no problems in accepting the constraints on aromatic compounds with no penalty on Q) and all the compounds that were in the solvent evaporation factor to the biogenic factor without making changes to the source profile of the biogenic emissions during the constraint runs. Primary emissions can only be restricted in the biogenic and mixed daytime factors by placing constraints on all daytime factors and constraining with the ratio to acetic acid as well, rather than with ratios to isoprene only. The reason for the second constraint is, that when a compound is completely removed from a factor profile, the constraining emission ratio no longer applies (as the constraining equation is not defined for 0 in the denominator). If constraints are placed on isoprene alone, but for all daytime factors, the model ends up remove isoprene completely from the mixed daytime factor during the constraint run. This allows the model to shift all the aromatics all other primary emissions into the photochemistry factor, rather than shifting them to one of the combustion sources or the solvent evaporation factor which costs a higher Q.

Figure 2 To which axis do the gray bars and red lines-and-markers belong, respectively?

In Figure 2, the left axis corresponds to the grey bar and the right axis corresponds to the red lines and markers, respectively.:

"Figure 3 represents the factor profiles of all the eight factors resolved by the PMF model in which grey bars (left axis) indicate the mass concentrations and red lines with markers (right axis) show the percentage of a species in the respective factor."

Line 370 You are interpreting m/z 69 $C_5H_8H^+$ as isoprene, correct? I suggest that this ion mass is actually a cycloalkane or alkane fragment, which seems far more plausible for vehicle exhaust. See Gueneron et al. (2015).

We thank the reviewer for pointing us to Gueneron et al., 2015 this fragment is indeed a potential explanation for $C_5H_8H^+$ in vehicle exhaust, although Borbon et al. 2001 (already cited in the paper) identified isoprene in the emissions of petrol vehicles that were not equipped with catalytic converter using GC FID not PTR-MS and found the emission factors were equivalent to those of pentenes an butenes. Considering the old fleet plying in Kathmandu valley, isoprene is a plausible contributor to the traffic source. Even so, if we consider that the fragmentation of cycloalkanes and cycloalkenes should also result in product ions at m/z 111 and/or m/z 125 and the signal at those masses at ~135 Td should be e above 0.2 ppb. However, in the observed mass spectra, there was no significant signal at these m/z values. Therefore, we conclude that isoprene is the more plausible assignment. As this is an important point, we have included this discussion in the revised version along with a citation to Gueneron et al., 2015.

We inserted the following text (insertion bold) into line 370f: "Few previous studies **employing GC-FID** have reported traffic related sources of isoprene in urban areas (Borbon et al., 2001; Hellèn et al., 2012)" and have added to line 372ff "**A recent study suggested m/z 69 $C_5H_8H^+$could also result from the fragmentation of cycloalkanes and cycloalkenes (Gueneron et al., 2015). Fragmentation of these compounds should also result in product ions at m/z 111 and/or m/z 125 and the signal at those masses at ~135 Td should be above 0.2 ppb. However, in the observed mass spectra, there was no significant signal at these m/z values. Therefore, we conclude that isoprene is the more plausible assignment."**

Line 477 Why would solvent evaporation correlate with the rate of change of temperature/sunlight, and not directly with temperature?

Assuming that the gas phase is in constant equilibrium with the aqueous phase at all times, mixing ratios should correlate with the change in temperature and not the absolute temperature. Considering a case where the water solubility of a compound or the saturation vapour pressure changes by a fixed factor for a fixed temperature difference. The spike in the gas phase mixing ratios of the compound would be sharper, if the temperature change occurred in a shorter period of time and the increase would be more gradual, if the same temperature changes occurred more slowly. Also the increase of the mixing ratios during daytime is counteracted by the dilution effect. When the rate of the temperature increase per unit time decreases in the late morning, the compounds no longer partition into the gas phase fast enough to overcome the dilution effect, hence the mixing ratios start dropping when the rate of change slows, even before it becomes negative.

We have shifted the following text from L536 to L477 to make the reasons more clear and also added the citations for the change in the solubility.

"However, the change of the saturation vapor pressure for a temperature change from 5^{o}C to 20^{o}C for the dominant compounds (acetaldehyde and acetic acid) present in the solvent evaporation factor is small (less than a factor of 1.3; Betterton and Hoffmann (1988); Johnson

et al. (1996)) and, therefore, does not account for the observed magnitude of increase (by a factor of ~5) from 06:00 - 09:00 LT. Instead, the temperature dependence of the solubility of these compounds in an aqueous solution would explain a change of this magnitude." This information is already provided in P21, L536 – P22, L546 of the ACPD version of the manuscript we have now shifted it to Line 477. The temperature change drives the compound from the aerosol aqueous phase into the gas phase.

Lines 545-548 Can you state (or reiterate, possibly I missed this above) why it cannot be that the solvent evaporation and unresolved industrial factors are "split" from a single source by the PMF? This section also seems fairly complex and highly speculative. Can you cite an example where such a situation has been shown to occur?

We are aware of previous papers exploring the impact of atmospheric conditions on the PMF output (e.g. Yuan et al. 2012), however, we are not aware of any other case where the gas phase mixing ratios were affected by the presence of a large aerosol aqueous phase. In previously reported studies the complications were caused by photochemistry. The two factors cannot be combined because the two correlate only during the day (R=0.55) and not during the night (R=0.29). When day and night are clubbed together R drops to 0.42. At night the solvent/evaporation factor anti-correlates with RH (R= -0.59) while the unresolved industrial factor has only a mild positive correlation with RH (R=0.29). During the day the solvent/evaporation shows the highest correlation with ΔT (R=0.64) while the unresolved industrial factor shows no significant correlation with ΔT (R=0.28). The raw data, now added as Table S5, also suggests against combining these two factors. The time series of measured acetaldehyde and acetic acid show a rather weak correlation with 1,3-butadiyne and methanol (R<0.54). On the other hand, the measured time series of 1,3-butadiyne and methanol correlates extremely strongly (R=0.9), indicating there is a strong and unique common source which causes sharp spikes in these two compounds which has very different emission ratios of 1,3-butadiyne to acetaldehyde, acetic acid and formic acid compared to the solvent evaporation factor (which is not a significant source of 1,3-butadiyne and methanol). The fact that the correlation of 1,3-butadiyne with ethanol, the defining compound of the mixed industrial factor, ethanol, is equally poor, speaks against combining the mixed industrial factor with the unresolved industrial factor.

The referee is correct that this section was poorly supported by data and we now address this valid concern by adding supplementary table S5 and the following text which replaces the original:

"While the correlation of the solvent evaporation factor with the unresolved industrial factor during daytime seems to suggest the two should be combined into one factor profile, several facts suggest against it. Firstly, the two do not correlate at night since the unresolved industrial factor shows a mild positive correlation rather than anti-correlation with RH at night (R=0.29) and no strong correlation with ΔT during the day (R=0.28). Secondly, the raw time series of 1,3-butadiyne and methanol (Supplementary table S5) correlates extremely strongly (R=0.9), indicating there is a strong and unique common source which causes sharp spikes in these two compounds. The fact that the time series of 1,3-butadiyne correlates poorly with acetaldehyde, acetic acid and formic acid indicates that the solvent evaporation factor (which is not a significant source of 1,3-butadiyne and methanol), has very different emission ratios of 1,3-

butadiyne to acetaldehyde, acetic acid and formic acid compared to the unresolved industrial emissions factor to explain the raw data. The fact that the time series of 1,3-butadiyne correlates equally poorly with that of ethanol, the defining compound of the mixed industrial factor, suggests against combining the mixed industrial factor with the unresolved industrial factor. It, therefore, seems, that the unresolved industrial factor is related to primary emissions from a distinct source, while the source profile and diel cycle of the solvent evaporation factor may be strongly confounded by meteorology and chemistry. Confounding factors have been reported to affect PMF solutions previously (Yuan et al. 2012)".

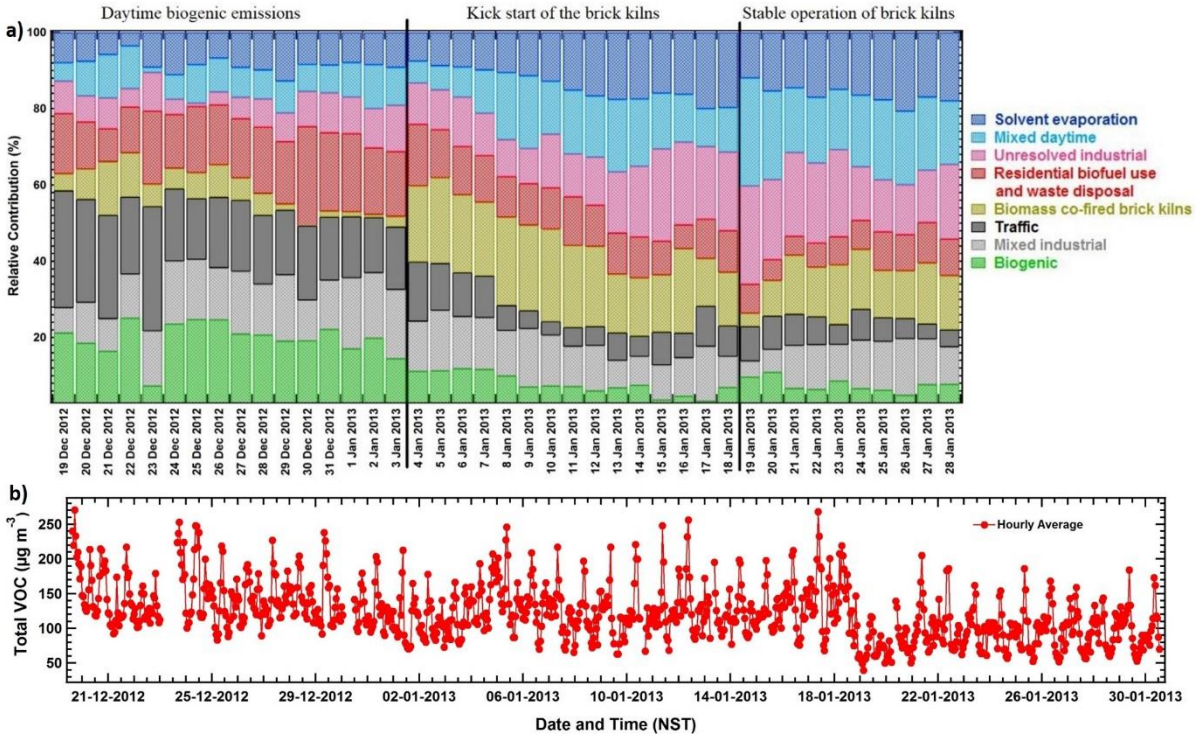
Lines 600-610 This also seems to point to a "splitting" of a single source into the unresolved industrial and solvent evaporation factors.

In the context of our study a "single source" would be a specific industrial point source or a specific well constrained sector which can be targeted by policy makers to reduce pollution. The fact that the industrial units responsible for the emissions associated with the solvent evaporation factor and the unresolved industrial factor are likely located in the same two industrial estates does not necessarily mean, the same plant or even the same sector is to be blamed for both types of emissions. The source profile of the unresolved industrial emissions is very specific and points towards plastic/adhesives/pharmaceutical industries. The source profile of the solvent evaporation factor is so strongly confounded by meteorology, that the origin of the emissions cannot be determined with great confidence. However, the fact that the acetic acid and acetaldehyde mass attributed to this factor is primarily distributed between three different combustion sources (brick kiln, residential fuel use and waste disposal and mixed industrial) when the number of factors is reduced to 6 (Figure S3a & b) indicates that multiple (combustion) sources contribute to the primary emission. The mass of the unresolved industrial emissions, on the other hand, gets distributed between the mixed industrial and mixed daytime factor when the number of factors is reduced to 7 (Figure S3a&b). The removal of a significant fraction of the mass of certain compounds from the mixed industrial into the unresolved industrial factor is accompanied by an almost complete separation of the conditional probability functions of these two factors. This means genuine sources are split from each other when the number of factors is increased to 8 and except for a conditional probability function pointing to two specific industrial estates and a late morning peak in emissions, the solvent evaporation and the unresolved industrial factor do not have much in common.

Figure 17 Can you also include the time series of the total VOC mass loading?

Done.

We have now included the timeseries of the total VOC mass loading in the revised Figure 17 (now 18) of the manuscript



Line 818 Many of the oxygenated VOCs are direct emissions from solvents, industry etc (Figure 16). So I do not think it is correct to say that photochemically produced VOCs are a dominant source of O_3 potential, especially when Figure 18a shows that the mixed daytime source contributed only about 5%.

We thank the reviewer for drawing attention to the confusion caused by the choice of words at Line 818. We completely agree that OVOCs have considerable anthropogenic sources in the Kathmandu Valley too. In fact this point was made strongly at Line 821 just three lines after L818 of the original submission, and in L818 we were only trying to make the point that without the PMF analyses, measurements of OVOCs and isoprene, which in several ambient environments are primarily controlled by photochemistry and biogenic sources, could have led to the premature assumption that these natural sources are more important for the daytime ozone formation potential in the Kathmandu Valley, whereas in fact anthropogenic sources are more important by collectively contributing 70% to the mass loading as noted in Line 821 of original submission .

In the revised version, we rephrased L 818 as follows to avoid potential confusion by adding the word "presumptuously" as follows::

"The distribution of the daytime O_3 production potential obtained from the measurements (Figure 19b) shows that 78% of the total daytime O3 production potential was due to the contribution from isoprene and oxygenated NMVOCs which could *presumptuously* indicate dominance of biogenic emissions and photochemistry in the Kathmandu Valley even in the winter."

Lines 847-848 Can you clarify how this is related to a result of your work.

Done, The paragraph now reads:

"Speciation of NMVOCs in the emission inventory for Nepal only includes compound classes (e.g. alkanes, alkenes etc.) and not specific compounds. This imposes certain limitations while comparing emission inventories with the compounds measured in our study. However, the existing emission inventories …"

Conclusion The conclusion is heavily weighted towards comparison with the emission inventories. While this is an important result, it is not the only finding discussed in the paper. The conclusion could be improved by an assessment of the major findings related to the source contributions to different categories of VOCs, specific VOCs, and O_3 and SOA formation potential.

Done.

We have now included a paragraph listing all the other important findings of this study in the conclusion.

"Eight different NMVOC sources were identified by the PMF model using the new "constrained model operation" mode. Unresolved industrial emissions (17.8%), traffic (16.8%) and mixed industrial emissions (14.0%) contributed most to the total measured NMVOC mass loading while biogenic emissions (24.2%), solvent evaporation (20.2%), traffic (15.0%) and unresolved industrial emissions (14.3%) were the most important contributors to the ozone formation potential. Biomass co-fired brick kilns and traffic contributed approximately equally to the secondary organic aerosol (SOA) production (28.9% and 28.2%, respectively), while the most important contributors to the mass loadings of carcinogenic benzene were brick kilns (37.3%), unresolved industrial (17.8% and mixed industrial (17.2%) sources. Photo-oxidation (mixed daytime factor) contributed majorly to two newly identified ambient compounds namely, formamide (41.1%) and acetamide (36.5%) along with their photooxidation product isocyanic acid (40.2%).

Minor comments (typographical corrections):

Lines 140-145 Put all verbs in present tense for consistency ("will provide" → "provides")

Done

Line 203 "Fewer" than 7 factors

Done

Line 235 "constraints"

Done

Line 289 "FCBTBK": what does this acronym stand for?

FCBTBK stands for fixed chimney bull's trench brick kiln. This has already been mentioned in P8, L245 of the ACPD version of the manuscript.

Figures 15, 16 Can you include the explanation for the different source acronyms (e.g. "MD, SE, UI") in the caption.

Done.

We have now added the full form of all the acronyms for Figure 15 to Figure 17 in the revised version of the manuscript.

References

Gueneron, M., Erickson, M. H., VanderSchelden, G. S., and Jobson, B. T.: PTR-MS fragmentation patterns of gasoline hydrocarbons, Int. J. Mass Spectrom., 379, 97-109, http://dx.doi.org/10.1016/j.ijms.2015.01.001, 2015.

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, 10.5194/acp-16-3979-2016, 2016.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891-2918, 10.5194/acp-9-2891-2009, 2009.

REFEREE 2:

General comment: The manuscript shows results of a source apportionment study of NMVOCs measured by PTR-TOF-MS in the Kathmandu Valley in Nepal during winter 2013. Positive matrix factorization analysis was conducted to identify possible emission sources for 37 m/z measured by PTR-MS. The sources were identified from the chemical fingerprint of each PMF factor and their diurnal profiles. Conditional probability functions plots were used to determine the directions of the sources and attribute the chemical emissions to specific spatial areas in the region and specific activities. The sources found by the authors through PMF were compared with the results of current emission inventories used for Nepal, which, in contrast to the authors results, rely on sources emission factors measured in other regions of the world and are not supported by in-situ collected measurements. Sources and species contributions differ among the authors results and the current inventories as well as between different inventories. Finally, the atmospheric impact as daytime ozone production and SOA formation based on the measured compounds and PMF sources contributions is briefly discussed. I found the manuscript very interesting, of high quality and of high impact as it presents several new findings which can help mitigating the emissions in the region under study. The presented topic also follows in the scope of ACP. The article is overall well written, and results are presented clearly with figures and tables. I highly recommend the manuscript publication, once these specific comments have been addressed:

We thank the referee for the kind words appreciating the importance of the work and for highly recommending the manuscript for publication in ACP. We have found several of Referee 2's specific comments very helpful and these are now reflected in the revised submission (changes are specified in replies and manuscript version with "tracked changes" given at the end of the responses here).

Specific comments:

-It is a bit confusing how the methods section is presented. There should be a small section introducing the measurement site, the PTR-MS data used, and the grab samples, before any PMF discussion. This would be helpful to follow better section 2.2 and support the nudging tool. Could you list the m/z from PTR-MS you used for implementing the PMF and why? Could you also provide some references of the nudging procedure?

Done.

We have now reorganized the Materials and Methods section (section 2) as per the suggestion of both the referees.

We have now included a column to Table S1 of the supplementary information to show the m/z ratios corresponding to the NMVOCs used for PMF. The detailed description concerning selection of these NMVOCs for the PMF run has now been added to the revised Materials and Methods section (Section 2).

The nudging procedure described in this work was performed using the priori knowledge of the emission sources in the Kathmandu valley and the emission ratios (ERs) obtained from the analysis of the grab samples collected from the point sources. This is the first ever study to use such nudging procedure to obtain robust solution using PMF. An earlier work of Baudic et al.,

2016 has mentioned the need of using the nudging procedure/constrains using priori knowledge of the emission sources and ERs obtained from point sources to obtain robust solution using PMF but did not implement the same.

-PTR-TOF-MS usually provides an unambiguous identification of chemical species, however, it would be interesting to know briefly, the operational settings of PTR-MS, which m/z were selected for running the PMF and how the m/z were attributed to the chemical compounds. Were the grab samples measured with the same PTR-MS? Could you provide some information about these data: m/z selected and how the compounds were identified. Line 331, could you provide the standard deviation for the signal stability?

We have now included a section on PTR-TOF-MS measurements in the revised manuscript (Section 2.2 of the revised manuscript) that briefly discusses about the operational settings of the PTR-TOF-MS, sampling of ambient air and the identification procedure of the NMVOCs. The m/z ratios to the corresponding NMVOCs used for the PMF run is now provided in Table S1 of the supplementary information.

No, the grab samples were measured with a PTR-QMS which is installed at IISER Mohali, India. The analytical details, calibration procedure and information regarding the identification of NMVOCs using this PTR-MS are available in Sinha et al., 2014.

For the grab samples we only reported 7 compounds which we have tested to be stable in glass flasks. These compounds are: acetonitrile, benzene, toluene, sum of C8 aromatics, sum of C9 aromatics, styrene and naphthalene.

The zero air background for the m/z reported was 0.04±0.05 ppb, 0.04±0.04 ppb, 0.04±0.06 ppb, 0.07±0.08 ppb, 0.10±0.11 ppb, 0.02±0.06 ppb and 0.02±0.05 ppb for acetonitrile, benzene, toluene, sum of C8 aromatics, sum of C9 aromatics, styrene and naphthalene, respectively. The concentration range in the grab samples was 4±0.3 to 323±8 ppb for acetonitrile, 27±4 to 339±19 ppb for benzene, 32±5 to 150±14 ppb for toluene, 40±6 to 113±8 ppb for C8 aromatics, 33±6 to 62±12 ppb for C9 aromatics, 11±1.3 to 95±17 ppb for styrene and 11±1.5 to 64±9 ppb for naphthalene.

We have now included this information in the Section describing the grab sampling and included a citation to the article that details the storage stability and validation of the glass flask sampling procedure and thank the referee for the excellent suggestion.

Citation:

Chandra, P., Sinha, V., Hakkim, H. 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, International Journal of Mass Spectrometry, https://doi.org/10.1016/j.ijms.2017.05.008, 2017.

-It would be interesting to provide some details about the calculations for ozone and SOA formation, you could do this with a short section in the methods after section 2.4.

The ozone formation potential of individual NMVOCs was calculated as described by the following equation (Sinha et al., 2012):

$$Ozone\ production\ potential = (\Sigma k_{(VOC_i+OH)}\ [VOC_i]) \times OH \times n$$

For the ozone production potential calculation, the average hydroxyl radical concentration was assumed to be [OH] = 1×10^6 molecules cm^{-3} with n = 2 and only data pertaining to the mid-daytime period were considered (11:00 - 14:00 LT).

This information is now included in Section 2.7

"The ozone formation potential of individual NMVOCs was calculated as described by the following equation (Sinha et al., 2012):

$$Ozone\ production\ potential = (\Sigma k_{(VOC_i+OH)}\ [VOC_i]) \times OH \times n$$

For the ozone production potential calculation, the average hydroxyl radical concentration was assumed to be [OH] = 1×10^6 molecules cm^{-3} with n = 2 and only data pertaining to the mid-daytime period were considered (11:00 - 14:00 LT)."

Secondary organic aerosol (SOA) production was calculated using the concentrations and the known SOA yields for benzene, toluene, styrene, xylene, trimethylbenzenes, naphthalene and isoprene (Ng et al., 2007; Chan et al., 2009; Yuan et al., 2013; Kroll et al., 2006). SOA yield of a particular NMVOC depends on the NO_x conditions and Pudasainee et al. (2006) previously reported NO_x-rich conditions in the Kathmandu valley. Therefore, SOA production was calculated by using reported SOA yield at high NO_x conditions according to the following equation:

$$SOA\ production = [VOC_i] \times SOA\ yield\ of\ VOC_i$$

This information is now included in Section 2.7

"SOA yield of a particular NMVOC depends on the NO_x conditions and Pudasainee et al. (2006) previously reported NO_x-rich conditions in the Kathmandu valley. Therefore, SOA production was calculated by using reported SOA yield at high NO_x conditions according to the following equation:

$$SOA\ production = [VOC_i] \times SOA\ yield\ of\ VOC_i$$"

Pudasainee, D., Sapkota, B., Shrestha, M. L., Kaga, A., Kondo, A., and Inoue, Y.: Ground level ozone concentrations and its association with NO_x and meteorological parameters in Kathmandu Valley, Nepal, Atmos. Environ., 40, 8081–8087, doi:10.1016/j.atmosenv.2006.07.011, 2006

-Figure 2. The contribution of propyne compared to isoprene for the biogenic factor is quite high, can you comment it?

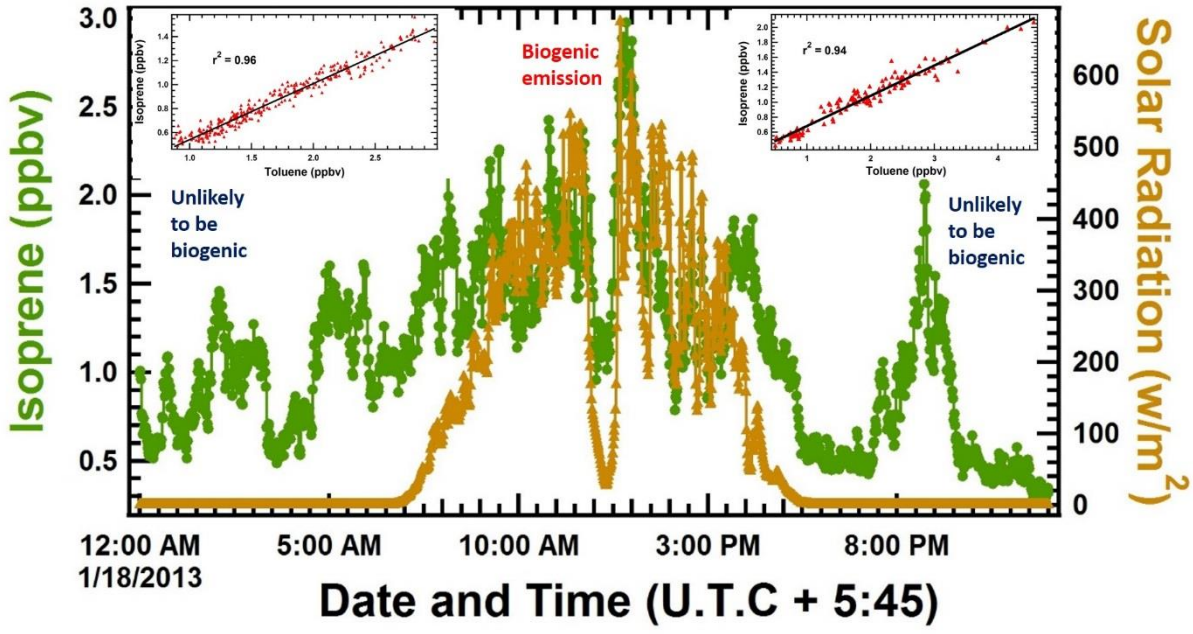
Since the source fingerprint of the primary source (traffic) is determined by night time emissions and the traffic factor profile during the daytime is different due to photochemical loss (between the source and the receptor downwind). As a consequence, some of the co-

emitted compounds (in particular those on which we placed no specific constraint) remain in the biogenic source profile even after constraints are imposed to remove combustion derived isoprene and other associated primary emissions such as propyne.

-Figure 3. How do you explain the higher background and general higher peaks during the first part of the campaign? L. 370: Could you provide more information about isoprene emission from traffic? Could you have any interference on the PTR-MS m/z attributed to isoprene?

The higher peaks during the first part of the campaign are due to emissions from deciduous trees which shed their leaves during the latter part of the campaign. The highest peaks during this period occur during the daytime and not at night. We have provided additional references for isoprene emissions from traffic and have discussed them in more detail also in response to reviewer 1.

During the measurement period, significant isoprene concentrations (~ 0.5-2 ppb) were observed during evening and night time which are likely from biomass combustion and traffic emission sources (Sarkar et al., 2016) as the evening and night time isoprene has a strong correlation with vehicular emission tracer toluene. The following figure (Figure S9 of the supplementary material) shows an illustrative day's (18 January 2013) isoprene data against solar radiation. It can be observed from the figure that the daytime isoprene emission correlates very nicely with solar radiation which indicates biogenic emission while during evening hours and night time, isoprene showed high peaks that show good correlation (r>0.9) with toluene.



The issue of possible interferences to the isoprene signal has already been discussed above while addressing the comment of referee 1 and revisions have been made as outlined in the reply to reviewer 1's comment.

Could be there a connection between oxidation products of traffic emission and the unresolved industrial emissions, as for mixed daytime emissions as oxidation products from biogenic emissions?

The traffic factor is dominated by the contribution of toluene and higher aromatics (C8- and C9-aromatics). The oxidation of all these aromatics produces phenols and cresols. However, we did not observe phenols and cresols above 200 ppt in the Kathmandu valley. The unresolved industrial emissions factor is dominated by propene, propyne, methanol, acetone, acetic acid, formic acid and 1,3 butadiyne. Propene, propyne and 1,3 butadiyne cannot be formed due to oxidation of higher aromatics. Consequently the methanol, acetone, acetic acid and formic acid found in the same factor profile cannot be from photo-oxidation either. Furthermore, a bimodal diel profile as observed for the unresolved industrial emissions profile is not characteristic of photochemically emitted compounds.

The mixed daytime emissions profile is dominated by nitrogen containing compounds, most notably isocyanic acid, and its precursors formamide and acetamide. In addition, the profile contains photochemically formed methanol, acetone, acetaldehyde, formaldehyde, formic acid and acetic acid. One of its features is, that the mass loadings of the photo-oxidation products present in the mixed daytime increases after the brick kilns resume operation. The biogenic emissions, on the other hand, decrease in the second half of the campaign when the deciduous trees shed their leaves. As a consequence, the primary oxidation products of isoprene, MVK+MACR (methyl vinyl ketone + methacrolein) and MEK (methyl ethyl ketone). MVK+MACR are associated with the biogenic emission factor itself and do not enter the photo-oxidation factor.

-Why biogenic emissions are higher during the first part of the campaign? Were temperature and solar radiation also higher for this part of the campaign?

Primarily because deciduous trees shed their leaves in early January (Sarkar et al., 2016) and no longer contribute to biogenic emissions during the second part of the campaign. We do not have measurements of temperature and solar radiation from the first part of the campaign due to a software glitch but in general conditions were warmer in the first part with reduced fog relative to the second part of the campaign.

-Section 3.2 would be easier to follow with a map of the measurement site and mentioned cities, industrial estates and forests.

Done.

We have now added a map of the measurement site to the revised manuscript (Figure 1).

-Section 3.3, The differences between the current inventories used in Nepal are briefly mentioned in the text, however, it would be interesting to write a few lines at the beginning of this section to introduce the inventories and on which data and assumptions they are based on. Is EDGAR v4.2 also considered for the winter season?

For EDGAR v4.2 spatially resolved seasonal data is not available. We have mentioned in the text and in the figure that the EDGAR v4.2 emission inventory are for the year 2008, while REAS 2.1 emissions are for December and January 2008. To make this clearer we have inserted "full" before "year" and added the following text in the paragraph on the EDGAR inventory: "EDGAR v4.2 inventory provides only spatially resolved data, not seasonally resolved data."

- Lines 786-808, Did you compare your NMVOCs data with the wind directions? How was the wind direction affecting the sources emissions captured at the measurements site?

The conditional probability function (CPF) plots shown in Figure 12 shows the wind directional dependency of different source categories reported in this study. The figure is discussed in Section 3.2.

Lines 786-808 describe the time series of the total VOC mass. The raw NMVOC data and its dependency on wind direction was analysed in Sarkar et al. (2016) already. As stressed in Sarkar et al. (2016) and in the materials and methods section of the current paper, during this time of the year, wind direction and speed in the Kathmandu valley usually followed predictable diurnal cycles. Such behaviour is typical for a site heavily influence by mountain meteorology. Hence the changes in the source strength of emission sources are not caused by changes in the wind direction, between the first and the second part of the campaign. They are due to genuine changes in the activity/emission strength.

-Line 815 and 840, please give the equations used for O_3 and SOA formation with respective references. It is not easy to understand figure 18 without any specification on the compounds used for each pie chart. Were the measured data used for pie b) the same data sets used to run the PMF?

Done. We have now included the equations used to calculate the O_3 and SOA formation in the materials and methods section of the revised manuscript.

Yes. The measured data used in Figure 19.b) is the same data set used to run the PMF model. As suggested by reviewer 1 we have improved the discussion of this figure.

-Line 828-832, much information is provided, please rephrase the period. Conclusions: Please include a short summary of the main findings here.

Done.

Earlier the sentence was:

"Based on measured methane and 63 non methane hydrocarbon measurements in the city of Lahore which is much larger and by all indications more polluted than Kathmandu (Barletta et al. 2016)) the authors reported a maximum contribution of about 14% due to all alkanes including methane to the total measured OH reactivity."

We have now modified this as follows:

“For the city of Lahore, Barletta et al.2016 , reported the maximum contribution of methane and 63 non methane hydrocarbons to the total measured OH reactivity as 14%. Lahore, is much larger and by all indications more polluted city than Kathmandu.”

Done. We have now included a short summary of the main findings in a paragraph in the conclusions as mentioned while addressing the comments of referee 1.

Technical comments: -Some acronyms are not explained, or only explained once in the whole manuscript. Could you also provide the extended form of all acronyms used for tables and figures in their captions?

Done.

The extended form of the acronyms are now provided in the figure and table captions/footnotes.

-L. 698, ca. 30%.

Done

Manuscript prepared for J. Name
with version 2015/04/24 7.83 Copernicus papers of the LaTeX class copernicus.cls.
Date: 23 May 2017

Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization

Chinmoy Sarkar[1], Vinayak Sinha[1], Baerbel Sinha[1], Arnico K. Panday[2], Maheswar Rupakheti[3,4], and Mark G. Lawrence[3]

[1]Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research (IISER) Mohali, Sector 81, S. A. S. Nagar, Manauli PO, Punjab, 140306, India
[2]International Centre for Integrated Mountain Development (ICIMOD), Khumaltar, Lalitpur, Nepal
[3]Institute for Advanced Sustainability Studies (IASS), Berliner Str. 130, 14467 Potsdam, Germany
[4]Himalayan Sustainability Institute (HIMSI), Kathmandu, Nepal

Correspondence to: V. Sinha (vsinha@iisermohali.ac.in)

Abstract. A positive matrix factorization model (US EPA PMF version 5.0) was applied for the source apportionment of the dataset of 37 NMVOCs measured over a period of 19 December 2012 – 30 January 2013 during the SusKat-ABC international air pollution measurement campaign using a Proton Transfer Reaction Time of Flight Mass Spectrometer in the Kathmandu Valley. In all, eight source categories were identified with the PMF model using the new "constrained model operation" mode. Unresolved industrial emissions and traffic source factors were the major contributors to the total measured NMVOC mass loading (17.9 % and 16.8 %, respectively) followed by mixed industrial emissions (14.0 %), while the remainder of the source was split approximately evenly between residential biofuel use and waste disposal (10.9 %), solvent evaporation (10.8 %), biomass co-fired brick kilns (10.4 %), biogenic emissions (10.0 %) and mixed daytime factor (9.2 %). Conditional probability function (CPF) analyses were performed to identify the physical locations associated with different sources. Source contributions to individual NMVOCs showed biomass co-fired brick kilns significantly contribute to the elevated concentrations of several health relevant NMVOCs such as benzene. Despite the highly polluted conditions, biogenic emissions had largest contribution (24.2 %) to the total daytime ozone production potential, even in winter, followed by solvent evaporation (20.2 %), traffic (15.0 %) and unresolved industrial emissions (14.3 %). Secondary organic aerosol (SOA) production had approximately equal contributions from biomass co-fired brick kilns (28.9 %) and traffic (28.2 %). Comparison of PMF results based on the in-situ data versus REAS v2.1 and EDGAR v4.2 emission inventories showed that both the inventories underestimate the contribution of traffic and do not take the contribution of brick kilns into account. In addition, the REAS inventory overestimates the contribution of residential biofuel use and underestimates the contribution of solvent use and industrial sources in the Kathmandu Valley. The quantitative source

apportionment of major NMVOC sources in the Kathmandu Valley based on this study will aid in improving hitherto largely un-validated bottom up NMVOC emission inventories, enabling more focused mitigation measures and improved parameterizations in chemical-transport models.

1 Introduction

Non-methane volatile organic compounds (NMVOCs) are important atmospheric constituents and are emitted from both natural and anthropogenic sources (Hewitt, 1999). They are important as precursors of surface ozone and secondary organic aerosols (SOA) and affect atmospheric oxidation capacity, climate and human health (IPCC, 2013). Thus, identification of NMVOC sources is necessary for devising appropriate mitigation strategies to improve air quality and reduce undesired impacts of secondary pollutants such as tropospheric ozone and secondary organic aerosol.

Source apportionment of NMVOCs can be achieved by applying source–receptor models to measured ambient datasets. Ambient NMVOC mixing ratios depend on the emission profiles of the sources contributing to the ambient mixture, their relative source strengths, transport, mixing and removal processes in the atmosphere. Source receptor models perform statistical analyses on the dataset to identify and quantify the contribution of different sources to the measured NMVOC concentrations (Watson et al., 2001). Positive matrix factorization (PMF) is currently among the most widely applied receptor models for the source apportionment of NMVOCs, in particular for datasets with high temporal resolution (Anderson et al., 2002; Miller et al., 2002; Kim et al., 2005; Buzcu and Fraser., 2006; Brown et al., 2007; Vlasenko et al., 2009; Slowik et al., 2010; Yuan et al., 2012; Crippa et al., 2013; Kaltsonoudis et al., 2016). In comparison to other receptor models based on principal component analysis/absolute principal component scores (PCA/APCS) (Guo et al., 2004, 2006), chemical mass balance (CMB) (Na and Pyo Kim., 2007; Morino et al., 2011) and UNMIX (Jorquera and Rappenglück., 2004; Olson et al., 2007), PMF provides more robust results as it does not permit negative source contributions. Moreover, a priori knowledge about the number and signature of NMVOC source profiles are not required, which is particularly useful and apt for NMVOC source apportionment studies in a new or understudied atmospheric chemical environment. The recently developed PMF version 5.0 also allows further refining the solution and reducing rotational ambiguity of the solutions using pre-existing knowledge of emission ratios from known point sources. Source apportionment of non-methane hydrocarbons (NMHCs) and oxygenated VOCs (OVOCs) using PMF source–receptor models has been carried out in several previous studies (Shim et al., 2007; Leuchner and Rappenglück , 2010; Gaimoz et al., 2011; Bon et al., 2011; Chen et al., 2014).

NMVOC emission inventories are frequently associated with large uncertainties (Zhang et al., 2009). This is particularly true for metropolitan cities in the developing world. Emission inventories can be evaluated using the results obtained from source receptor models such as the PMF model. This evaluation is important to improve the accuracy of the existing emission inventories and therefore to

develop effective air pollution control strategies. In this study, we report the application of the PMF model for source apportionment of NMVOCs using the NMVOC data measured in the Kathmandu Valley, Nepal, which has been reported and analyzed in detail in Sarkar et al. (2016).

Kathmandu is considered to be amongst the most polluted cities in Asia (Panday et al., 2009). According to the existing Nepalese emission inventory (International Centre for Integrated Mountain Development's (ICIMOD) database) and the REAS v2.1 (Kurokawa et al., 2013) emission inventories residential biofuel use is considered to be the most important anthropogenic source of NMVOCs in the Kathmandu Valley. It is considered to contribute $\sim 67\,\%$ (REAS) to $\sim 83\,\%$ (Nepalese inventory), towards the total NMVOC mass loadings. In contrast,EDGAR v4. (Olivier et al., 1994) attributes 66 % of the emissions in the Kathmandu Valley to solvent use and a recent emission inventory study conducted by the International Centre for Integrated Mountain Development (ICIMOD) which relied on measurement of particulate matter (Figure S7) suggested that traffic is the dominant source (69 %) of air pollution in a part of the Kathmandu Valley within the Ring Road (i.e. the Kathmandu Metropolitan City (KMC) and Lalitpur Sub-metropolitan City) and some nearby sub-urban rural areas outside the Ring Road (Pradhan et al., 2012).

The objective of the current study is to identify and quantify the contributions of different emission sources to the ambient wintertime NMVOC concentrations in the Kathmandu Valley using a positive matrix factorization (US EPA PMF 5.0; Brown et al. (2015)) receptor model. NMVOC measurements were carried out at Bode, a suburban site in the Kathmandu Valley over a period of 19 December 2012 – 30 January 2013 during the SusKat-ABC field campaign. The NMVOC measurements, new findings and qualitative analyses of sources have been presented and discussed in Sarkar et al. (2016). The NMVOC measurements suggested significant contribution of varied emission sources such as traffic (associated with high toluene, xylenes and trimethylbenzenes), biomass co-fired brick kilns (associated with high acetonitrile and benzene), industries and wintertime biogenic sources (as characterized by high daytime isoprene). Based on the NMVOCs emission profiles, two distinct periods were identified in the dataset: the first period (19 December 2012 – 3 January 2013) was associated with high daytime isoprene concentrations whereas the second period (4 – 18 January 2013) was associated with sudden increase in acetonitrile and benzene concentrations which was attributed to the start in operations of biomass co-fired brick kilns in the Kathmandu Valley (Sarkar et al., 2016). For quantitative source apportionment, hourly mean measured concentrations of all 37 NMVOCs measured during the instrumental deployment (19 December 2012 – 30 January 2013), were used for the PMF analysis. Sensitivity tests were conducted for the PMF 5.0 model version to evaluate how the new rotational tool called "constrained model operation feature" improves the representation of source profiles in the PMF model output. To identify the physical locations for the identified sources, an important prerequisite for targeted mitigation, conditional probability function (CPF) analyses were also performed. The results obtained from the PMF analyses were compared with three emission inventories – the existing Nepalese inventory, REAS v2.1 (Regional Emission

inventory in ASia) and the EDGAR v4.2 (Emissions Database for Global Atmospheric Research) emission inventory. Additionally, the contributions of each source category to individual NMVOC mass concentrations, ozone formation potential and formation of secondary organic aerosol (SOA) were also analyzed.

2 Materials and Methods

2.1 Site Description

BS: NMVOC measurements during this study were performed in the winter season from 19 December 2012 until 30 January 2013 at Bode (27.689° BS: N, 85.395° BS: E, 1345 m a.m.s.l.) BS: in Bhaktapur district, which is a suburban site located in the westerly outflow of the Kathmandu Metropolitan City. The land use in the vicinity of the measurement site consisted of the following cities - Kathmandu Metropolitan City ($\sim$ 10 km BS: to the west), Lalitpur Sub-Metropolitan City ($\sim$ 12 km BS: south-west of the site) and Bhaktapur Municipality ($\sim$ 5 km BS: south-east of the site). The site is located in the Madhyapur-Thimi Municipality. In addition, the region north of the site had a small forested area (Nilbarahi Jungle; $\sim$ 0.5 km^2 BS: area) and a reserve forest (Gokarna Reserve Forest; $\sim$ 1.8 km^2 BS: area) at approximately 1.5 km BS: and 7 km BS: from the measurement site, respectively. Several brick kilns were located in the south-east of the site within a distance of 1 km BS:. Major industries were located mainly in the Kathmandu and Patan cities whereas Bhaktapur industrial estate was located at around 2 km BS: from the measurement site (in the south-eastern direction). A substantial number of small industries were also located in the south-eastern direction. The Tribhuvan International Airport is located about 4 km BS: to the west of the Bode site. A detailed description of the measurement site and prevalent meteorology is already provided in the companion paper to this special issue Sarkar et al. (2016) BS:. A zoomed view of the land use in the vicinity of the measurement site is provided in Figure 1.

2.2 PTR-TOF-MS measurements

BS: NMVOC measurements were performed using a high-sensitivity PTR-TOF-MS (model 8000; Ionicon Analytic GmbH, Innsbruck, Austria) over a mass range of 21-210 amu BS:. The PTR-TOF-MS instrument works on the basic principle of soft chemical ionization (CI) where reagent hydronium ions ($H3O^+$ BS:) react with analyte NMVOC molecules having proton affinity (P.A) greater than that of water vapour (165 Kcal/mol BS:) to form protonated molecular ions (with m/z ratio = molecular ion + 1), enabling the identification of NMVOCs (Lindiger et al., 1998) BS:. As all the relevant analytical details pertaining to the PTR-TOF-MS instrument, ambient air sampling and the quality assurance of the NMVOC dataset has already been provided in detail in Sarkar et al. (2016) BS:, only a brief description of the ambient air sampling and the analytical operating conditions is provided here. Ambient air sampling was performed continuously through a Teflon inlet line protected from

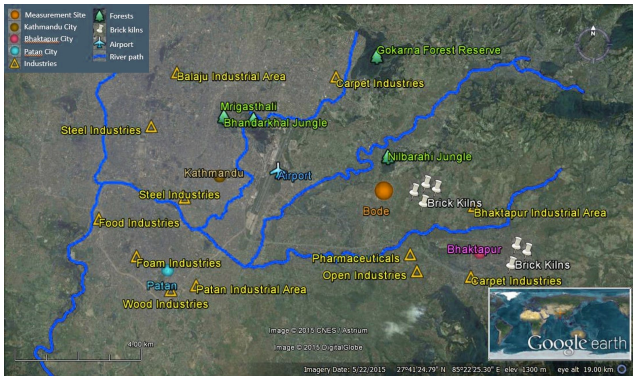


Figure 1. Location of the measurement site (Bode, orange circle) along with surrounding cities (Kathmandu, brown circle; Patan, turqoise circle; Bhaktapur, pink circle), brick kilns (white markers), major industries (yellow triangles), forest areas (green tree symbols), airport (blue marker) and major river paths (sky blue) in the Google Earth image of the Kathmandu Valley (obtained on 22 May 2015 at 14:55 LT).

floating dust and debris using an in-line Teflon membrane particle filter. The PTR-TOF-MS was operated at a drift tube pressure of 2.2 mbar[BS:], a drift tube temperature of 60°C[BS:] and a drift tube voltage of 600 V[BS:] which resulted in an operating E/N ratio of ∼135 Td (E = electrical field strength in V cm^{-1}[BS:]; N = buffer gas number density in molecule cm^{-3}[BS:]; 1 Td = 10^{-17} V cm^{-2}[BS:]). Identification of several previously unmeasured and rarely measured NMVOCs were achieved due to the high mass resolution (m/Δm > 4000) and low detection limit (few tens of ppt) of the instrument. For the quality assurance of the measured NMVOC dataset, the instrument was calibrated twice during the measurement period and regular instrumental background checks were performed using zero air at frequent intervals. Detailed description of the sensitivity characterization of the instrument and the quality assurance of the primary dataset is available in Sarkar et al. (2016).

[BS:]During the measurement period, a total of 37 NMVOC signals (m/z) were observed in the PTR-TOF-MS mass spectra that had an average concentration of > 200 ppt. The cut-off of an average concentration of > 200 ppt was employed keeping in mind the highest instrumental background signals observed during the campaign, so as to have complete confidence that the ions signals were attributable to ambient compounds. For mass identifications at a particular m/z ratio, further quality control was applied. Firstly, only those ion peaks were considered for the mass assignments for which there were no contribution from the major shoulder ion peaks within a mass width bin of 0.005 amu. Next, ion peaks devoid of any variability (that is the time series profile was flat) were not considered for mass assignments at all. Further details including some known interferences that were identified and taken into account are available in Sarkar et al. (2016)[BS:]. Table S1 in the supplementary information lists the identified 37 NMVOCs the corresponding m/z attributions (with references to few previous works which reported the same compound assignment, wherever applicable), and the elemental molecular formula.

2.3 Collection of grab samples

BS: Grab samples from garbage fires (termed garbage burning) were collected near the measurement site (∼ 200 m BS: in the northern direction, upwind of Bode; 27.690° BS: N, 85.395° BS: E) on 7 December 2014 between 15:00 - 15:03 LT BS: . A "brick kiln" grab sample was collected on 6 December 2014 from a fixed chimney bull's trench brick kiln (FCBTK) co-fired using coal, wood dust and sugarcane extracts. Figure S1 of the supplementary information shows pictures of the grab sample collection and the instrumental setup for the analysis. The whole air samples were collected in 2 litre glass flasks that had been validated for the stability of NMVOCs (Chandra et al., 2017) BS: and were analyzed within 38 hours of the collection (BS: on 9 December 2014 between 03:42 - 04:05 LT BS:). The whole air samples (WAS) were diluted (dilution factor of 9.93) using zero air for the quantification of NMVOCs present in the grab samples using a PTR-QMS instrument (Sinha et al., 2014) BS: . The average background signals (zero air) were subtracted from each m/z channel and stable data of at least 10 cycles (∼ 10 minutes BS:) were considered for the calculation of mixing ratios as per the protocol described by Sinha et al. (2014). BS: The zero air background for the m/z reported was 0.04±0.05 ppb, 0.04±0.04 ppb, 0.04±0.06 ppb, 0.07±0.08 ppb, 0.10±0.11 ppb, 0.02±0.06 ppb and 0.02±0.05 ppb for acetonitrile, benzene, toluene, sum of C8 aromatics, sum of C9 aromatics, styrene and naphthalene, respectively. The concentration range in the grab samples was 4±0.3 to 323±8 ppb for acetonitrile, 27±4 to 339±19 ppb for benzene, 32±5 to 150±14 ppb for toluene, 40±6 to 113±8 ppb for C8 aromatics, 33±6 to 62±12 ppb for C9 aromatics, 11±1.3 to 95±17 ppb for styrene and 11±1.5 to 64±9 ppb for naphthalene.

2.4 Positive Matrix Factorization (PMF)

The United States Environmental Protection Agency's (US EPA) Positive Matrix Factorization (PMF) receptor model version 5.0 (Norris et al., 2014) was used for source apportionment of NMVOCs in the Kathmandu Valley. The model is based on the multi-linear engine (ME-2) approach and has been described in detail by Paatero (1997, 1999). From a data matrix of a number of NMVOCs in a given number of samples, the PMF model helps to determine the total number of possible NMVOC source factors, the chemical fingerprint (source profile) for each factor, the contribution of each factor to each sample, and the residuals of the dataset using the following equation (Paatero and Tapper , 1994),

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

Where, X_{ij} is the NMVOC data matrix with i number of samples and j number of measured NMVOCs which are resolved by the PMF to provide p number of possible source factors with the source profile f of each source and mass g contributed by each factor to each individual sample, leaving the residuals e for each sample. To obtain the solution of equation (1), sum of the squared

residuals (e^2) and variation of data points (σ^2) are inversely weighted in PMF as expressed by the following equation (Paatero and Tapper , 1994),

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(\frac{e_{ij}}{\sigma_{ij}})^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}(\frac{X_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{\sigma_{ij}})^2 \quad (2)$$

Where, Q is the object function and a critical parameter for PMF, n is the number of samples, and m is the number of considered species. The original data should always be reproduced by the PMF model within the uncertainty considering the non-negativity constraint for both the predicted source profile and the predicted source contributions. The explained variability (EV) as given below demonstrates the relative contribution of each factor to the individual compound and can be expressed as (Gaimoz et al., 2011),

$$EV_{kj} = \frac{\sum_{i=1}^{n} |g_{ik} f_{kj}|/\sigma_{ij}}{\sum_{i=1}^{n}(\sum_{k=1}^{p} |g_{ik} f_{kj}| + |e_{ij}|)/\sigma_{ij}} \quad (3)$$

The explained variability is most useful to policy makers. If the observed mass loading of a compound that is known to be harmful to human health is high, the explained variability will indicate which sources are responsible for most of its emissions and what fraction of the total observed mass is contributed by each source. Therefore, this allows planning mitigation strategies.

BS: Bootstrap runs were performed to ascertain the magnitude of random errors of the dataset (Norris et al., 2014; Paatero et al., 2014). BS: ~~To ascertain the magnitude of~~ BS: ~~r~~BS: Random errors BS: ~~that~~ can be caused BS: ~~due to the use of random seeds followed by the selection of the run with the lowest Q~~ due to the existence of infinite solutions with different g_{ik}, f_{kj} and e_{ij} matrices but identical $Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(e_{ij}/\sigma_{ij})^2$.BS: ~~,~~ ~~bootstrap runs were performed~~. In the bootstrap runs, the timeseries is partitioned into smaller segments of a user specified length and the PMF is run on each of these smaller segments, for the same number of factors as the original model run. The model output of each bootstrap run is mapped onto the original solution using a cross correlation matrix of the factor contributions g_{ik} of a given bootstrap run with the factor contributions g_{ik} of the same time segment of the original solution using a threshold of the Pearson's correlation coefficient (R) > 0.6 BS: as suggested by Norris et al. (2008, 2014). The bootstrap factor is assigned to the factor with which it is most strongly positively correlated, as long as the value of R is greater than 0.6. If it cannot be attributed to any factor of the original solution it will be termed unmapped. The presence of a high fraction unmapped factor ($> 20\,\%$) is a clear indication of large random errors BS: (introduced by a few critical observations that drastically impact factor profiles) and should be investigated carefully (Norris et al., 2014). In our analysis, no unmapped factors were present.

For each factor, the factor profile of all bootstrap runs combined is compared with the profile of the original model output. The model BS: ~~will~~ provide BS: s a box and whisker plot for the mass loading ($\mu g\,m^{-3}$) and percentage of each compound attributed to the factor profile of each of the factors during the bootstrap runs. It BS: ~~will~~ also ascertain BS: s for each compound whether or not the original

solution for that factor falls into the interquartile range of the bootstrap results and provide[BS:s] this information in a table format.

When all sources are equally strong throughout the entire period, this bootstrap model provides a robust estimate of the total random error. However, if one of the sources is completely absent for a significant fraction of the total hours (like the brick kiln source throughout the first 13 days of the SusKat-ABC campaign), the bootstrap model may overestimate the random error substantially. For such a source, mass loading of all the compounds that contribute strongly to the factor profile of the source will typically be outside the interquartile range. For the same set of compounds, similar behavior could also be seen for the factor profile of several other factors. In such a situation, the error estimate of the bootstrap runs should only be considered as the upper limit of the potential random error.

In addition to the random error, the PMF model [BS:also] has rotational ambiguity (Ulbrich et al., 2009; Paatero et al., 2014). [BS:This rotational ambiguity is caused due to the existence of multiple solutions which have a Q similar to the solution produced by the PMF model but different factor profiles and factor contributions. Thus, the model will find different local minima of the residual matrix, while determining the factor contribution matrix ($g_{ik}f_{kj}$). The coexistence of different solutions for the factor contribution matrix ($g_{ik}f_{kj}$)] [BS:~~There can be multiple solutions with a different factor profile for all factors for which the model will find a different local minimum of the residual matrix while determining the factor contribution matrix. This fact that different solution for $g_{ik}f_{kj}$~~] with the same sum of the scaled residuals $Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(e_{ij}/\sigma_{ij})^2$ [BS:~~exist~~] is called the rotational ambiguity of the model. The PMF 5.0 has a new feature named as “the constrained model operation" in which the rotational ambiguity of the model can be constrained using external knowledge of the source composition (f_{kj}) or contribution (g_{ik}) matrix. For instance, if a source was inactive for a particular period, then the contribution due to that factor during that time period could be pulled to zero in the model to provide more robust output. Alternatively, the emission ratios obtained from a particular source through samples collected at the source can also be used to constrain the model. Constraining the PMF model using such external knowledge gives rise to a penalty in Q (the object function) and a maximum penalty of 5 % is recommended as a reasonable threshold (Paatero and Hopke, 2009). A detailed discussion of the use of constraints to a receptor model has been provided in previous studies (Paatero et al., 2002; Rizzo and Scheff , 2007; Paatero and Hopke, 2009; Norris et al., 2008).

2.5 Implementation of PMF

PMF was applied to the hourly averaged dataset of 37 ions measured using a Proton Transfer Reaction Time of Flight Mass Spectrometer (PTR-TOF-MS). All relevant analytical details pertaining to

the site description, meteorology, sampling and quality assurance of the NMVOC dataset has already been described in detail in the companion paper to this special issue (Sarkar et al., 2016).

BS: ~~Briefly, NMVOC measurements during this study were performed in the winter season from 19 December 2012 until 30 January 2013 at Bode (27.689°~~ BS: ~~N, 85.395°~~ BS: ~~E, 1345 m a.m.s.l.) in Bhaktapur district, which is a suburban site located in the westerly outflow of the Kathmandu Metropolitan City. The land use in the vicinity of the measurement site consisted of the following cities - Kathmandu Metropolitan City (~ 10 km to the west), Lalitpur Sub-Metropolitan City (~ 12 km south-west of the site) and Bhaktapur Municipality (~ 5 km south-east of the site). The site is located in the Madhyapur-Thimi Municipality. In addition, the region north of the site had a small forested area (Nilbarahi Jungle; ~ 0.5 km^2 BS: ~~area) and a reserve forest (Gokarna Reserve Forest; ~ 1.8 km^2 BS: area) at approximately 1.5 km and 7 km from the measurement site, respectively. The Tribhuvan International Airport is located at about 4 km to the west of the Bode site.~~

All the available data BS: ~~during this aforementioned study Sarkar et al. 2016~~ were used for the PMF analysis and the missing values were replaced by a missing value indicator (-999). To ensure that differential uncertainties do not drive the object function Q and give undue weighting to calibrated organic ions while constructing source profiles, we followed the procedure used by Leuchner and Rappenglück (2010) for source apportionment of NMVOCs in the Houston Ship Channel area, assigning a constant uncertainty of 20 % for all the ions. BS: ~~The attribution of ions to parent compounds and corresponding detection limits were as described in Sarkar et al. 2016~~. Due to its erratic timeseries profile, HCN ($m/z =$ 28.007) was classified as a weak species in the PMF input while all other ions were classified as strong species. BS: For weak species, the stated uncertainty is tripled, to reduce their impact on the scaled residual and hence Q. All the input data was converted from mixing ratios BS: of ppb to mass concentrations ($\mu\mathrm{g\,m}^{-3}$) using the relevant temperature, pressure and molecular weight BS: and BS: ~~. T~~ BS: t he total measured NMVOC concentration was calculated by adding the mass concentrations of all measured NMVOCs BS: . This conversion allows calculating the explained variability (Garimoz et al. 2011) for the total VOC mass and comparing the results with emission inventories. The conversion does not introduce significant additional uncertainty and the variability induced by the temperature (average range observed was: 5-20 °, C BS:) has largely been taken into account by running the model with a 5% BS: extra modelling uncertainty. The total VOC mass BS: ~~and~~ is classified as a weak species in the PMF input (Norris et al., 2014). All the measured ions had a signal to noise (S/N) ratio greater than 2. Table S2 of the supplementary information shows the signal to noise (S/N) ratios for all input NMVOC species used in the PMF along with other statistical parameters of the dataset.

PMF model runs ranging from 5 to 12-factor numbers were carried out to ascertain the best solution for this study, consistent with the chemical environment of the Kathmandu Valley. Based on the $Q/Q_{theoretical}$ ratio, the physical plausibility of the factors and BS: constraints imposed by the rotational ambiguity of the solution, an 8-factor solution was deemed to be the best for this dataset. BS: For the data presented in this study, the $Q/Q_{theoretical}$ ratio is <1 even for a 3 factor solution with no physical plausibility and and hence the absolute number does not help to decide the optimum number of factors. Supplementary Figure S2 shows clearly, that the number of factors has almost no impact

on how well the total mass is reproduced by the model, but the last distinct drop in the Q/$Q_{theoretical}$ ratio is seen when the number of factors is increased to 8. When BS: ~~less~~ BS: fewer than 7-factors were employed, several source profiles appeared to be mixed BS: (Figure S3a,b), indicating inadequate resolution of sources. The solution incorporating 7-factors BS: ~~caused strong overlap of mixed industrial emissions with the unresolved industrial emissions factor. This solution~~ was considered inappropriate, as the daytime biogenic emissions BS: and photochemical sources could not be separated from the nighttime combustion source of isoprene in the 7-factor solution. Even when the model was nudged towards separating the biogenic emissions and the anthropogenic combustion sources of isoprene using the constraint mode, this separation could only be accomplished with a large penalty on Q in the 7-factor solution. The 9-factor solution had too much rotational ambiguity and assigned brick kiln emissions to two largely co-linear factors, both of which had an incomplete source profile with respect to aromatic compounds and were essentially created to better account for minor variations in the emission ratios associated with brick kiln emissions during the firing up period and the continuous operation later in the campaign BS: (Figure S3c).

The diagnostics for the 8-factor solution are summarized in Table 1. The eight factors were - 1) traffic, 2) residential biofuel use and waste disposal, 3) mixed industrial emissions, 4) biomass co-fired brick kilns, 5) unresolved industrial emissions, 6) solvent evaporation, 7) mixed daytime source and 8) biogenic emissions. BS: A detailed description for the identification and the attribution of the 8-factor solutions is provided later in section 3.1. BS: The primary data strongly supports an 8 factor solution. The top 2-3 compounds explained by each of the 8 factors have a much higher R when their input time series is correlated compared to the R obtained when their time series is correlated with the time series of any other compound (Supplementary Table S5).

BS: The traffic factor explains more than 60 % BS: of the variability of Toluene, C-8 and C9 aromatics. The time series of Toluene, C8 and C9 aromatics correlates with R >0.96 for all possible pairs when the original time series of these compounds are correlated with each other. The R of the time series of these same compounds with the time series of styrene is lower 0.81-0.85 while a correlation of their time series with all other compounds yields R <0.78. This indicates toluene, sum of C-8 and C9 aromatics share a major common source with each other which is not shared by other compounds, namely the traffic source. Hence a less than 6 factor PMF solution which is incapable of capturing the traffic source is not a better representation of the reality.

BS: For styrene the highest correlation is with furan R=0.87 indicating that the two compounds have a significant source in common, which styrene also shares with higher aromatics and propyne (R=0.86), but the lower R of styrene with the aromatic compounds indicates that styrene has at least two dominant sources with distinct emission ratios. These sources are the traffic source (explaining

roughly 40 % BS: of the styrene) and the residential burning source which explains 30 % BS: of the styrene and furan variability. These two sources are separated only with a 6 factor solution.

BS: Benzene has a strong source in the form of biomass co-fired brick kilns which results in a distinct increase in emission at the time the brick kilns restart their operations. This source is shared with acetonitrile (R=0.89), nitromethane (R=0.82) and naphthalene (R=0.81) but all of these compounds also have other sources which are either not shared with benzene or have different emission ratios. This source appears in the 3 factor solution but its source profile is contaminated with mixed industrial emission. The closure period of brick kilns is only fully captured and restricted to the brick kiln factor after the number of factors is increased to 7.

BS: The mixed industrial source explains 66 % BS: of the ethanol variability, but this compound has a relatively low R with all other compounds (0.73 with propene and 0.7 with nitromethane and acetonitrile <0.66 with the rest) indicating that there must be at least two distinct ethanol sources with different source fingerprints. A second distinct ethanol source in the form of solvent evaporation, however, separates from the mixed daytime factor only in the 7 factor solution.

BS: The mixed daytime factor primarily contains photo-chemically formed compounds most notably isocyanic acid, which shows a strong correlation with its own precursors formamide (R=0.85) and acetamide (R=0.82). Figure S8 presents reaction schematic for the formation of formamide and isocyanic acid. This compound has a much weaker correlation with other compounds, which have other sources in addition to the photochemical source (R=0.5 to 0.58 for formaldehyde, acetaldehyde, the nitronium ion, formic acid and acetic acid). This factor should ideally be restricted to photo-chemically formed secondary compounds, however, it remains heavily contaminate with night-time primary emissions during the second half of the campaign till the number of factors is increased to 8 (Figure S3c). Even the 8 and 9 Factor solution still contain some minor contamination from primary emissions. Hence the name of the source is retained as mixed daytime source.

BS: The solvent evaporation factor is characterised by acetaldehyde and acetic acid which have their strongest correlation with each other (R=0.82). Apart from this, the defining compound, acetaldehyde, shows moderate correlation with formaldehyde (R=0.72) and acetone (R=0.68) but only the former correlates with acetic acid (R=0.85) as it shares both the solvent evaporation source and the photo-oxidation source with acetaldehyde, while the later (acetone) correlates much stronger with methyl ethyl ketone (R=0.95) and methyl vinyl ketone (R=0.86) and isoprene (R=0.79) and hence shares the biogenic emission source in addition to the the solvent evaporation factor. While these three daytime sources are resolved in the 7 factor solution their source profiles continue to be contaminated with primary emissions. While the same can be pushed around from the biogenic factor into the mixed daytime factor using rotational tools, they cannot be sufficiently removed from both till an 8th factor is allowed.

BS: The unresolved industrial emission factor explains a significant fraction of the 1,3-butadiyne which shares most of its sources with methanol (R=0.9). The source profile also captures several

Table 1. Diagnostic for the results of the positive matrix factorization (PMF) model run

n (samples)	1006
m (species)	37
k (factors)	8
Q (theoretical)	4480.37
Q (model)	4562.89
Mean ratio NMVOC(estimated)/NMVOC(observed)	0.999

other compounds with a lower correlation with 1,3-butadiyne including propanenitrile (R=0.86), acrolein + methylketene (R=0.82) and propene (R=0.8). The R obtained while cross correlating the time series of 1,3-butadiyne with that of ethanol, the defining compound of the mixed industrial source profile, is only 0.73 and ethanol correlates only weakly with Acrolein + Methylketene (R=0.59) indicating that these mixed industrial emissions and unresolved industrial emissions represent distinct sources, which can only be resolved in a 8 factor solution.

To identify the uncertainty associated with the PMF solution, bootstrap runs were performed 100 times taking 96 hours as the segment length. BS: This is slightly shorter than the recommended length based on the equation of Politis and White (2004)BS:, of 108 hours but represents a multiple of 24 hours and hence ensures each bootstrap run contains four full days' worth of data.There were no unmapped factors in the bootstrap runs.

Figure 2 shows the correlation between the estimated total measured NMVOC concentrations calculated using the contributions from all factors (vertical axis) with measured total measured NMVOC concentrations (horizontal axis). An excellent correlation (r^2 = 0.99) indicates that PMF model can explain almost all variance in the total measured NMVOC concentrations.

The constrained model mode was used to further improve the 8-factor solution. BS: The constraint mode is a new rotational tool introduced in the 5.0 version of the EPA PMF as an alternative to the FPeak module. The constraint mode allows to exploit the rotational ambiguity of the model to push the PMF solution into a physically more realistic space. It uses pre-existing knowledge such as source fingerprints, source emission ratios or activity data. We found that when the two modules were compared for an equal number of factors the constraint mode performance was superior to the FPeak module. The original model output showed positive correlations between BS: the factor contribution time series of BS: ~~factors such as~~ the biomass co-fired brick kilns and mixed industrial emissions (r^2 = 0.27) BS: factors as well as the residential biofuel use and waste disposal factor with traffic factor (r^2 = 0.42). Since this is a new feature and has only recently been used by Brown et al. (2015) for ambient air data, a detailed description of the implementation procedure and an analysis of how the constraints affected the model output is provided here. Several constraints were used to obtain a more robust PMF solution.

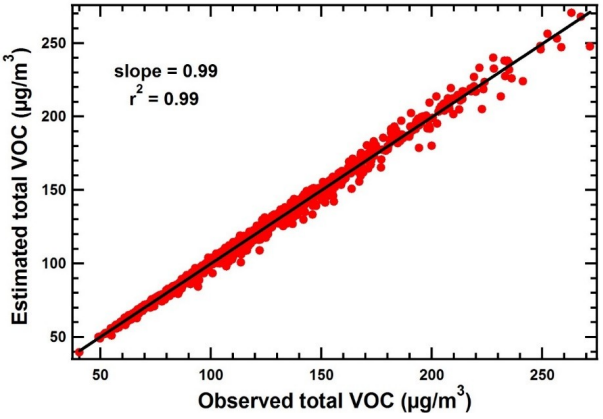


Figure 2. Correlation between estimated and observed NMVOC concentrations

First, the upper limit for the emission ratio of the individual aromatic compounds to isoprene as reported by Misztal et al. (2015) were used to constrain the factor profile of primary biogenic emissions. As a small fraction of the biogenic isoprene gets attributed to other daytime factor (mixed daytime) by the PMF model, the same constraint[BS: s] were used on mixed daytime factor and the solvent evaporation factor as well.

Second, it was assumed that aromatic compounds and acetonitrile are not photochemically produced. Acetic acid is associated with both mixed daytime and solvent evaporation, and so the ratios of aromatic compounds and acetonitrile to acetic acid were nudged towards 0.0001 for these two factors.

Third, to improve the representation of brick kiln emissions, and the residential biofuel use and waste disposal in the model, the respective factors, which were clearly identified in the original model solution, were nudged using the emission ratios of aromatic compounds to benzene from grab samples of domestic waste burning (garbage burning grab sample) and fixed chinmney bull's trench brick kiln emissions (FCBTBK grab sample) collected directly at the point source. This was required, because in the original model output, the residential biofuel use and waste disposal factor correlated with the traffic factor (r^2 = 0.42) while the brick kilns emission factor correlated with the mixed industrial emissions factor (r^2 = 0.27). This indicates that there was substantial rotational ambiguity for these two factor pairs.

Nudging was performed by exerting a soft pull allowing for a maximum 0.2 % change in Q for each constraint. A soft pull allows the change in the Q value up to a certain limit by pulling the values to a target value for an expression of elements (the emission ratio). If no minima can be found for which the change in $Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(e_{ij}/\sigma_{ij})^2$ is less than 0.2 % in the $g_{ik}f_{kj}$ matrix after f_{kj} has been constrained, no change was made and the original solution was retained. If the condition

Table 2. Inter NMVOC emission ratios used for biogenic, solvent evaporation and mixed daytime factors to nudge the PMF model and the corresponding emission ratios before and after nudging

ERs/Isoprene	ERs used to nudge	BG		SE		MD	
		before nudging	after nudging	before nudging	after nudging	before nudging	after nudging
Acetonitrile	0.002	0.06	0.00	0.00	0.004	2.78	1.75
Benzene	0.002	0.29	0.00	0.52	0.00	0.15	0.00
Toluene	0.012	0.10	0.01	0.39	0.00	4.82	0.00
Styrene	0.002	0.02	0.00	0.06	0.00	0.00	0.002
Xylenes	0.002	0.00	0.0002	0.35	0.41	4.65	0.00
Trimetylbenzenes	0.002	0.06	0.01	0.09	0.00	1.85	0.20
Naphthalene	0.002	0.31	0.30	0.36	0.60	0.00	0.002
ERs/Acetic acid	ERs used to nudge	BG		SE		MD	
		before nudging	after nudging	before nudging	after nudging	before nudging	after nudging
Acetonitrile	0.0001	0.57	0.00	0.00	0.0001	0.07	0.09
Benzene	0.002	1.48	0.00	0.04	0.00	0.01	0.00
Toluene	0.0001	1.01	0.004	0.05	0.00	0.12	0.00
Styrene	0.0001	0.15	0.00	0.01	0.00	0.00	0.0001
Xylenes	0.0001	0.00	0.0001	0.04	0.01	0.12	0.00
Trimetylbenzenes	0.0001	0.59	0.004	0.01	0.00	0.05	0.01
Naphthalene	0.0001	3.08	0.15	0.04	0.01	0.00	0.0001

BG = Biogenic; SE = Solvent evaporation; MD = Mixed daytime

can be met without changing Q by more than the threshold, the revised factor profiles will be used as the base upon which the next constraint in the list of constraints will be executed.

Implementing the constraints mentioned above, significantly improved the representation of biogenic emissions, mixed daytime and solvent evaporation factors. Figure S4 of the supplementary information shows a comparison of the box and whisker plots of the biogenic emissions, mixed daytime and solvent evaporation factors before and after nudging and demonstrates the significant improvement after applying constraints.

After nudging, the contribution of the biogenic factor correlated better with solar radiation (r^2 = 0.48) while the mixed daytime factor correlated better with ambient temperature (r^2 = 0.42). The factor profile of the solvent evaporation correlates better with the rise in solar radiation and temperature after sunrise (07:00 - 09:00 LT; r^2 = 0.53). Table 2 represents the emission ratios used to nudge the biogenic, mixed daytime and solvent evaporation factors and provides the corresponding emission ratios (ERs) before and after nudging.

It can be seen that most constraints on the aromatic to isoprene ratio could be executed without exceeding the penalty on Q. In the biogenic factor, only the naphthalene/isoprene ratio could not be constrained. The solvent evaporation and mixed daytime factors contain only a small fraction of the total daytime isoprene (8 % and 7 %, respectively). Given the very small overall isoprene mass in these two factor profiles, few additional ratios did not meet the constraining criteria in these factor profiles (namely acetonitrile/isoprene and trimethylbenzenes/isoprene ratio in the mixed daytime factor and the xylenes/isoprene and naphthalene/isoprene ratio in the solvent evaporation factor). Some of these compounds (such as naphthalene) could not be constrained in the same factors while constraining the ERs with respect to acetic acid.

The fact that the constrained run was incapable of removing naphthalene from the source profiles of the biogenic and the solvent evaporation source and the fact that the diel profiles of both these factors show a weak secondary peak between 17:00 - 22:00 LT, seems to indicate that an additional weak combustion source with a high naphthalene emission ratio is possibly poorly represented by the current 8-factor solution. Cooking on 3-stone fires is known to emit large amounts of benzene and naphthalene (Stockwell et al., 2015) and the temporal profile of such a cooking source could overlap with that of the garbage fires. It can be noted that 3-stone fires is still a common way to cook for construction workers and brick kiln workers staying in temporary camps in the Kathmandu Valley. This would make it challenging for the model to separate these two sources. We will henceforth refer to the garbage burning factor as the residential biofuel use and waste disposal factor.

Figure S5a of the supplementary information shows the G-space plots for two factors, namely biomass co-fired brick kilns and mixed industrial emissions. A stronger correlation (r^2 = 0.42) existed in the original solution prior to nudging with ERs of FCBTBK grab samples, which reduced to r^2 = 0.18. Similarly, after nudging with ERs of the garbage burning grab sample the correlations between residential biofuel use and waste disposal was reduced from 0.27 to 0.18, as shown in Figure S5b. Thus, the new solution fills the solution space better.

Table 3 summarizes the aromatics/benzene emission ratios derived from the PMF (before and after nudging) and its comparison with the emission ratios obtained from grab samples for biomass co-fired brick kilns and residential biofuel use and waste disposal sources. These emission ratios are also compared with the ERs reported for 3-stone firewood stoves in Stockwell et al. (2015) and the mixed garbage burning and open cooking fire sources reported for Nepal in Stockwell et al. (2016).

For the residential biofuel use and waste disposal source, the original model run already had emission ratios very similar to the garbage burning grab samples of the garbage burning fire. The constrained run improved the agreement further for styrene, trimethylbenzenes and naphthalene. Constraining this factor with the ERs of 3-stone firewood stoves from Stockwell et al. (2015) instead of our garbage burning grab samples resulted in a larger penalty on Q and did not improve the representation of the biogenic, mixed daytime and solvent evaporation factors.

Table 3. Comparison of aromatics/benzene ERs (emission ratios) obtained from PMF (before and after nudging), respective grab samples, the 3-stone firewood source reported in Stockwell et al. (2015) and the mixed garbage burning and open cooking fire sources reported in Stockwell et al. (2016)

ERs/Benzene	FCBTBK grab samples	BK PMF (before nudging)	BK PMF (after nudging)	garbage burning grab samples	RB+WD PMF (before nudging)	RB+WD PMF (after nudging)	3-stone firewood[1]	Mixed garbage[2]	Open hardwood cooking[2]
Toluene	0.80	0.28	0.35	0.34	0.33	0.34	0.11	0.37	0.27
Styrene	0.08	0.05	0.06	0.16	0.22	0.18	0.09	0.19	0.11
Xylenes	0.58	0.16	0.22	0.25	0.28	0.25	0.10	0.18	0.12
Trimethylbenzenes	0.31	0.06	0.09	0.08	0.16	0.12	0.03	0.02	0.03
Naphthalene	0.09	0.14	0.15	0.09	0.16	0.11	0.40	-	-

1. Stockwell et al. (2015); 2. Stockwell et al. (2016); BK = Biomass co-fired brick kilns; RB+WD = Residential biofuel use and waste disposal

For brick kilns, the emission ratios of the constrained model output runs diverged from the emission ratios of the FCBTBK grab samples. However, the temporal profile of the activity, especially the closure of the brick kilns during the first part of the campaign is better captured by the constrained run and the correlation with mixed industrial emission sources reduced significantly. The FCBTBK grab samples were collected on 6 December 2014, two years after the SusKat study, so differences from the emission profiles observed during the SusKat-ABC campaign are a possibility. Alternatively, the differences could also stem from the inherently variable nature of this source. In particular, naphthalene and benzene were higher in the source profiles of the SusKat-ABC campaign compared to their relative abundances in the FCBTBK grab samples. At the time the FCBTBK grab samples were collected (on 6 December 2014), brick kilns were co-fired using coal, wood dust and sugarcane extracts. It is possible that in January, during peak winter season, a different type of biomass, one associated with higher benzene and naphthalene emissions (e.g. wood) was used in these biomass co-fired brick kilns, resulting in the slight disagreement between the PMF source profile and FCBTBK grab sample signature for this factor. Table S3 of the supplementary information shows the percentage contribution of PMF derived factors obtained from constrained runs with 5, 6, 7, 8 and 9-Factors.

2.6 Conditional probability function (CPF) analyses

For identifying the physical locations associated with different local sources, conditional probability function (CPF) analyses were performed. CPF is a well-established method to identify source locations of local sources based on the measured wind (Fleming et al., 2012). In CPF, the probability of a particular source contribution from a specific wind direction bin exceeding a certain threshold is

employed which is calculated as follows:

$$CPF = \frac{m_{\Delta\theta}}{n_{\Delta\theta}} \quad (4)$$

Where $m_{\Delta\theta}$ represents the number of data points in the wind direction bin $\Delta\theta$ which exceeded the threshold criterion and $n_{\Delta\theta}$ represents the total number of data points from the same wind direction bin. For this study, $\Delta\theta$ was chosen as 30° and data for wind speed $> 0.5\,\mathrm{m}^{-1}$ were used.

2.7 Calculation of ozone and SOA formation potential

BS: The ozone formation potential of individual NMVOCs was calculated as described by the following equation (Sinha et al., 2012):

$$Ozone\ production\ potential = (\sum k_{(VOC_i+OH)}[VOC]_i) \times OH \times n \quad (5)$$

BS: For the ozone production potential calculation, the average hydroxyl radical concentration was assumed to be [OH] = 1×10^6 BS: molecules cm^{-3} BS: with n = 2 and only data pertaining to the mid-daytime period were considered (11:00 - 14:00 LT).

SOA yield of a particular NMVOC depends on the NOx conditions and Pudasainee et al. (2006) BS: previously reported NOx-rich conditions in the Kathmandu valley. Therefore, SOA production was calculated by using reported SOA yield at high NOx conditions according to the following equation:

$$SOA\ production = [VOC]_i \times SOA\ yield\ of\ VOC_i \quad (6)$$

3 Results and Discussion

3.1 Identification of PMF Factors

Figure 3 represents the factor profiles of all the eight factors resolved by the PMF model in which grey bars BS: left axis indicate the mass concentrations and red lines with markers BS: right axis show the percentage of a species in the respective factor.

Identification and attribution of these factors is discussed in detail in the following sections.

3.1.1 Factor 1 - Traffic

More than 60 % of the total toluene, sum of C8-aromatics, sum of C9-aromatics and ∼ 37 % of the total assorted hydrocarbons ($m/z =$ 97.102 and 83.085) were explained by Factor 1. Toluene and C8-aromatics contributed most (∼ 16 % and ∼ 13 %, respectively) to the total measured NMVOC mass of Factor 1. In addition four other compounds also contributed ≥ 5 % to the total mass of this factor (propyne (∼ 11 %), acetone (∼ 9 %), propene (∼ 6 %) and sum of C9-aromatics (∼ 5 %)). The other 31 NMVOCs contributed ∼ 40 % of the total measured NMVOC mass to this factor but their individual contributions were ≤ 5 % each. The diel profile of Factor 1 (Figure 4) showed characteristic evening peak at 17:00 LT with an average concentration of ∼ 40 µg m^{-3}. This evening

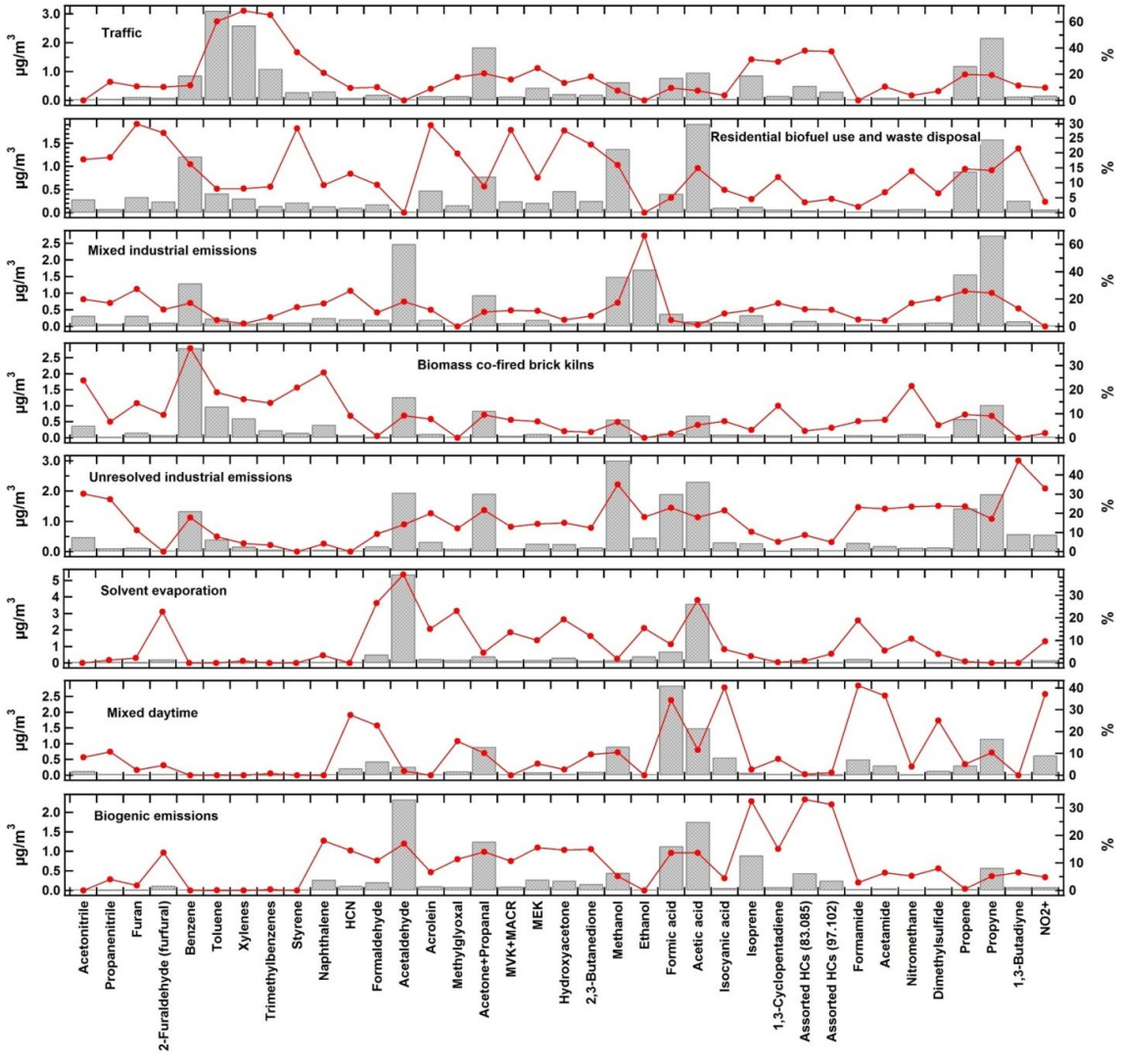


Figure 3. Factor profiles of the eight sources obtained by PMF analysis

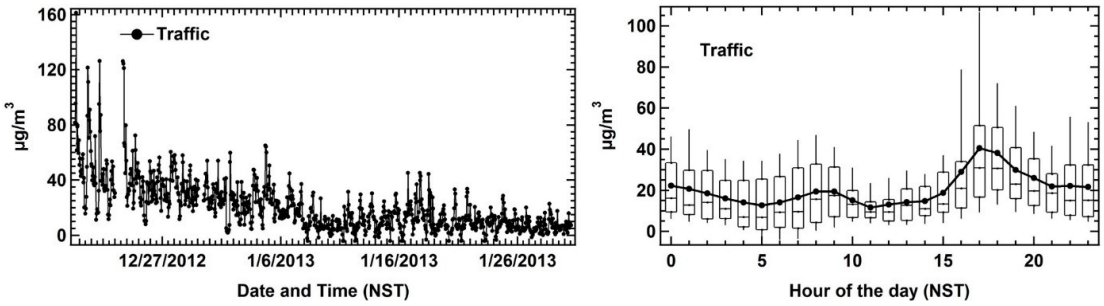


Figure 4. Timeseries and diel box and whisker plot for Factor 1 (Traffic)

peak showed large variability and plume-like characteristics as the average and median diverged frequently. Occasionally, the mass contribution of this factor amounted to $\sim 100\ \mu g\,m^{-3}$. The high variability during the evening peak hour indicates that the source strength is not equal for all wind directions, but varies with fetch region.

Table 4 shows that the aromatics/benzene emission ratios for this factor are in good agreement with the emission ratios reported by previous studies for vehicular emissions in tunnel experiments and in metropolitan sites/megacities. In view of the diel profile and observed chemical signatures, Factor 1 was attributed to traffic. It can be noted that in winter, rush-hour in the city starts at 16:00 LT, while westerly winds still bring urban air to the measurement site. The morning rush hour in the city takes place in calmer winds which leads to a less sharp peak. It is interesting to note that $\sim 37\,\%$ of the total styrene was present in this factor and $\sim 31\,\%$ of the total isoprene was also explained by this factor. Few previous studies *BS:* employing GC-FID have reported traffic related sources of isoprene in urban areas (Borbon et al., 2001; Hellèn et al., 2012) and also estimated isoprene as one of the top 10 contributors to OH reactivity from traffic (Nakashima et al., 2010). *BS:* A recent study suggested m/z 69 $C_5H_8H^+$ *BS:* could also result from the fragmentation of cycloalkanes and cycloalkenes (Gueneron et al., 2015)*BS:*. Fragmentation of these compounds should also result in product ions at m/z 111 and/or m/z 125 and the signal at those masses at 135 Td should be above 200 ppt considering the measured $C_5H_8H^+$ *BS:* ion signal in the Kathmandu valley during our study. However, in the observed mass spectra, there was no significant signal at these m/z values. Therefore, we conclude that isoprene is the more plausible assignment. *BS:* ~~Our results indicate that traffic can be a significant source of nighttime isoprene in the Kathmandu Valley.~~

3.1.2 Factor 2 - Residential biofuel use and waste disposal

Factor 2, too, showed regular evening hour peaks and a bimodal profile (Figure 5). However, the evening peak of average concentrations as high as $\sim 40\,\mu g\,m^{-3}$ occurred after the traffic peak (at 19:00 LT) and had less variability, indicating that this source is an area source that is spatially spread throughout the Katmandu Valley. The diel box and whisker plot also has a relatively weak morning

Table 4. Emission ratios of NMVOCs/benzene for aromatic hydrocarbons derived from the PMF model for factor attributed to traffic and comparison of ERs with previous studies for traffic source profiles

ERs/Benzene	Kathmandu PMF	Tunnel study, Stockholm[1]	Tunnel study, Hong Kong[2]	Tunnel study, Taipei[3]	Mexico City[4]	Los Angeles[5]
Toluene	3.41	3.89	2.27	2.38	3.47	2.45
C8-aromatics	2.89	2.81	0.87	1.86	3.55	1.38
C9-aromatics	1.20	-	0.77	1.36	2.31	0.48
Styrene	0.30	-	-	0.39	0.17	-
Naphthalene	0.19	-	0.10	-	-	-

1. Kristensson et al. (2004); 2. Ho et al. (2009); 3. Hwa et al. (2002) ; 4. Bon et al. (2011) ; 5. Borbon et al. (2013)

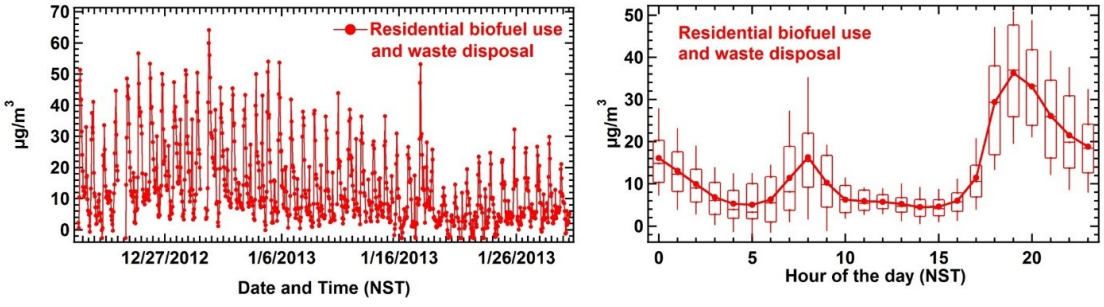


Figure 5. Timeseries and diel box and whisker plot for Factor 2 (Residential biofuel use and waste disposal)

peak (at 08:00 LT) with average concentrations of $\sim 18\ \mu g m^{-3}$. Figure 3 shows, that this factor explains 30 % of the total styrene, furan, 2-furaldehyde and acrolein.

Most of the measured NMVOC mass in this factor was contributed by acetic acid, propyne, methanol, benzene, propene and acetone + propanal ($\sim 14\,\%$, $\sim 12\,\%$, $\sim 10\,\%$, $\sim 9\,\%$, $\sim 7\,\%$ and $\sim 6\,\%$ respectively). The other 31 measured NMVOCs contributed $\sim 42\,\%$ to this factor, but their individual contributions were $\leq 5\,\%$ each (Figure 3). It was observed that garbage/trash burning activities were more intense during evening hours in winter in the Kathmandu Valley. Table 5 shows a comparison of the aromatics/benzene emission ratios obtained from the PMF, with previously reported aromatics/benzene ratios for waste and trash burning, and with the emission ratios of garbage burning grab samples that were collected in the Kathmandu Valley near the point source (a household waste fire). It can be seen that the aromatics/benzene emission ratios of the PMF output are in excellent agreement with the values obtained for garbage burning grab samples collected in the Kathmandu Valley.

There is some agreement with the emission ratios reported in previous studies, though all of these previous studies found higher emission ratios for styrene. This could indicate that the composition of household waste in the Kathmandu Valley is different (less polystyrene, plastic and more biomass)

Table 5. Emission ratios of NMVOCs/benzene for acetonitrile and aromatic hydrocarbons derived from the PMF model for the factor attributed to Residential biofuel use and burning household waste and comparison with previously reported studies and the garbage burning grab samples collected at the point source

ERs/Benzene	Kathmandu PMF	Kathmandu garbage burning grab samples	Mixed garbage burning[1]	Household waste burning[2]	Open hardwood cooking[1]	Trash burning[3]	Scrap tires burning[2]
Acetonitrile	0.23	0.77	-	-	-	0.06	-
Toluene	0.34	0.34	0.37	0.38	0.27	0.41	0.63
C8-aromatics	0.25	0.25	0.19	0.22	0.11	0.10	0.43
C9-aromatics	0.12	0.08	0.18	-	0.12	0.03	0.03
Styrene	0.18	0.16	0.02	0.54	0.03	0.86	0.30
Naphthalene	0.11	0.09	-	0.01	-	0.10	0.30

1.Stockwell et al. (2016) ; 2. Lemieux et al. (2004) ; 3. Stockwell et al. (2015)

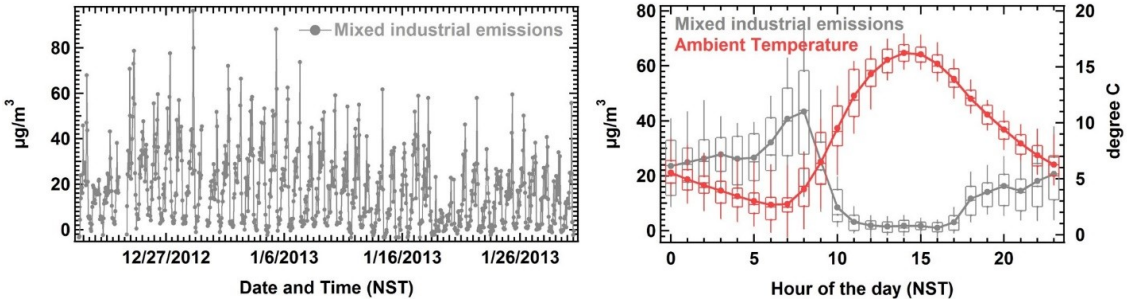


Figure 6. Timeseries and diel box and whisker plot for Factor 3 (Mixed industrial emissions)

or that the source profile is mixed with that of a second source, with similar spatial and temporal characteristics. Residential biofuel use is expected to have a similar temporal profile and did not appear as a separate factor in the PMF solution. Therefore, Factor 2 was attributed to residential biofuel use and waste disposal sources collectively.

3.1.3 Factor 3 - Mixed industrial emissions

This factor explained 66 % of the total ethanol, which is used as an industrial solvent. Moreover, $\sim 20-25$ % of the total propyne, propene, acetonitrile, dimethyl sulfide (DMS) and furan were also present in this factor. All these compounds have industrial sources (Karl et al., 2003; Kim et al., 2008) as they are widely used as solvents/reactants in various industrial processes and can be emitted during combustion processes. Therefore, Factor 3 was attributed to mixed industrial emissions. Most of the measured NMVOC mass in this factor was contributed by propyne (~ 16 %), acetaldehyde (~ 15 %), ethanol (~ 10 %), propene (~ 9 %), methanol (~ 9 %), benzene (~ 8 %) and acetone + propanal ($\sim$

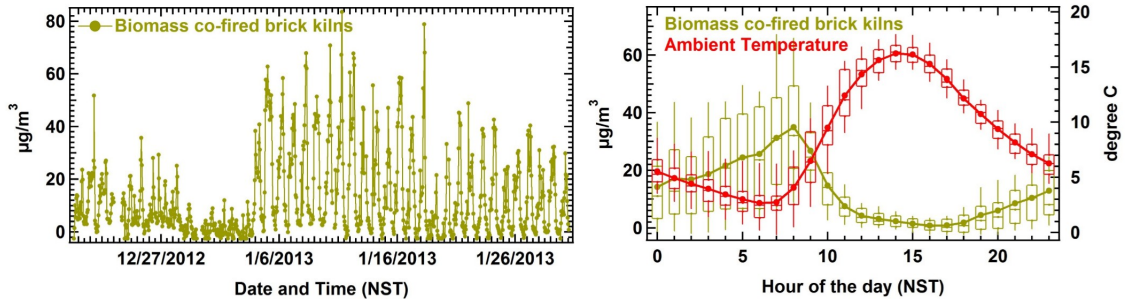


Figure 7. Timeseries and diel box and whisker plot for Factor 4 (Biomass co-fired brick kilns)

5 %). The emissions reflect both release of chemicals used in the industrial units as well as emissions associated with combustion of a variety of fuels including biofuels. The other 30 NMVOCs jointly contributed only ∼ 28 % of the total measured NMVOC mass and their individual contribution were ≤ 5 % each. The emission strength of industrial sources is typically constant throughout the day and hence the observed mass concentrations are driven by boundary layer dynamics. The diel box and whisker plot (Figure 6) shows a gradual increase in the mass concentrations throughout the night. The highest mass concentration are observed just after sunrise, when the inversion in the mountain Valley is most shallow. This shallow early morning boundary layer is caused by the cold pooling of air at night, which results in an accumulation of cold air at the Valley bottom. The rising sun first warms the upper part of the Valley's atmosphere, while the Valley bottom is still in the shade of the surrounding mountains. Once direct sunlight reaches the Valley bottom, warming and thermally driven convection breaks the shallow boundary layer and wind speeds increase, increasing turbulent mixing under a growing boundary layer. The daytime mass concentrations of the mixed industrial emissions are hence an inverse of the temperature and wind speed profile (Figure 6).

3.1.4 Factor 4 - Biomass co-fired brick kilns

The diel box and whisker plot of factor 4 (Figure 7) shows a profile that is similar to the profile of mixed industrial emissions, indicating that this factor should be attributed to a source that operates 24/7, as its mass loadings, too, represent an inverse of the temperature and wind speed profile. The timeseries of Factor 4 showed sudden increase from 4 January 2013 at exactly the time when brick kilns in the Kathmandu Valley became operational (Sarkar et al., 2016).

Benzene (∼ 23 %) contributed most to the total measured NMVOC mass of Factor 4. In addition acetaldehyde (∼ 10 %), propyne (∼ 8 %), toluene (∼ 8 %), acetone (∼ 7 %), acetic acid (∼ 5 %) and xylenes (∼ 5 %) also contributed significantly to the total measured NMVOC mass. The other 30 NMVOCs contributed ∼ 34 % to the total measured NMVOC mass of this factor, but their individual contribution were ≤ 5 % each. Overall, factor 4 explained ∼ 37 % of the total benzene and ∼ 24 % of the total acetonitrile mass loading.

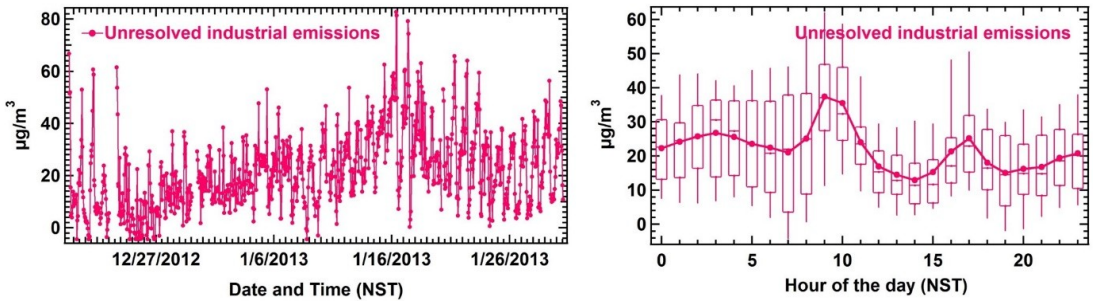


Figure 8. Timeseries and diel box and whisker plot for Factor 5 (Unresolved industrial emissions)

It is reported that brick kilns in the Kathmandu Valley burn large quantity of biomass, wood and crop residues along with coal (Stone et al., 2010; Sarkar et al., 2016) that can lead to significant emission of aromatics and acetonitrile (Akagi et al., 2011; Yokelson et al., 2013; Sarkar et al., 2013). Therefore Factor 4 was attributed to the biomass co-fired brick kilns and the conditional probability function analysis (section 3.2) is consistent with this assignment.

3.1.5 Factor 5 - Unresolved industrial emissions

Factor 5 explained $\sim 48\,\%$ of the total 1,3-butadiyne, $\sim 35\,\%$ of the total methanol, $\sim 30\,\%$ of the total acetonitrile and $27\,\%$ of the total propanenitrile and $24\,\%$ of the total nitromethane. 1,3-butadiyne is used in the production of several polymers and acetonitrile and propene can be side products in this process. Propanenitrile is used to start acrylic polymerization reactions in industrial processes. The largest use of methanol worldwide is as feedstock for the plastic industry and nitromethane is used in the synthesis of several important pharmaceutical drugs. It can be noted that several pharmaceutical industries are located in the Thimi area which is only $\sim 2\,\mathrm{km}$ away from the measurement site. Nitromethane is also emitted from combustion of diesel fired generators (Inomata et al., 2013, 2014; Sekimoto et al., 2013) which are used as a back-up power source by both small and large industrial units in the Kathmandu Valley. It is, therefore, likely that miscellaneous nearby industries contributed significantly to the unresolved factor. The diel profile of Factor 5 (Figure 8) showed morning and evening peaks (at 09:00 - 10:00 LT and 17:00 LT, respectively), which is not typical for industrial emissions, but this factor always had a high background with average mass loadings of $\sim 20\ \mu\mathrm{g\,m^{-3}}$ throughout. The timeseries and diel profile (Figure 8) of this factor did not reveal characteristics that could be related uniquely to a known emission source.

Figure 8 displayed elevated daytime mass concentrations and an evening peak for this factor that occurs slightly before the traffic peak in the early evening during the first part of the SusKat-ABC campaign (until 25 December). Towards the end of the campaign (from 10 January onwards), the same factor had diurnal variations that showed some similarity to profiles of both the solvent evaporation (morning peak) and mixed industrial emissions (slow rise throughout evening and nighttime)

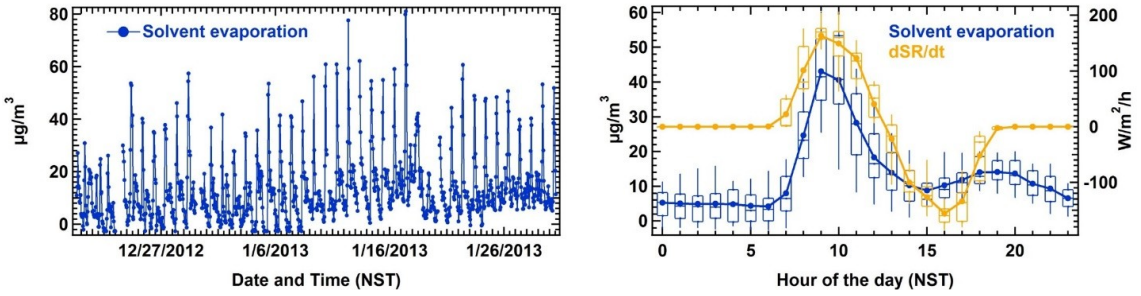


Figure 9. Timeseries and diel box and whisker plot for Factor 6 (Solvent evaporation)

factors. Between 25 December and 10 January, diurnal patterns are weak and peaks in the unresolved factor seem to coincide with peaks in the solvent evaporation factor. This comparison of the diel profiles is shown in Figure S6 of the supplementary information. Since this factor seems to contain contributions of multiple sources and potentially the photooxidation products of their emissions, this factor was termed as the unresolved industrial emissions factor.

Most of the total measured NMVOC mass of Factor 5 was due to oxygenated NMVOCs like methanol ($\sim 14\,\%$), acetic acid ($\sim 11\,\%$), acetaldehyde ($\sim 9\,\%$), acetone ($\sim 9\,\%$) and formic acid ($\sim 9\,\%$) but benzene, propyne and propene also contributed $> 5\,\%$ ($\sim 9\,\%$, $\sim 6\,\%$ and $\sim 6\,\%$, respectively) to the total measured NMVOC mass of this factor. The other 29 NMVOCs together contributed only $\sim 27\,\%$ to this factor and their individual contributions were less than $5\,\%$.

3.1.6 Factor 6 - Solvent evaporation

Factor 6 explains approximately 25-40 % of the compounds containing the aldehyde functional group. It explained $\sim 39\,\%$ of the total acetaldehyde, $\sim 27\,\%$ of the total formaldehyde and $\sim 23\,\%$ of 2-furaldehyde. Moreover, $\sim 28\,\%$ of the total acetic acid and $\sim 23\,\%$ of the total methylglyoxal were explained by this factor. Acetaldehyde and acetic acid contributed $\sim 40\,\%$ and $\sim 27\,\%$ respectively to the total measured NMVOC mass of Factor 6 while formic acid, formaldehyde, acetone and ethanol together contributed $\sim 15\,\%$ ($\sim 5\,\%$, $\sim 4\,\%$ and $\sim 3\,\%$, respectively) to the total measured NMVOC mass of this factor. The other 31 species contributed only $\sim 18\,\%$. The diel profile (Figure 9) of this factor correlates best with the increase in rates of temperature (dT/dt, R^2 = 0.41) and solar radiation (dSR/dt, R^2 = 0.38) during the daytime hours (between 06:00 – 17:00 LT; as can be seen in Table S4 of the supplementary information). Factor 6 showed a sharp peak directly after sunrise between 08:00 – 10:00 LT. This time coincides with the maximum increase in both temperature and solar radiation. Average mass loadings of ~ 45 µg m^{-3} were observed during this period. BS: However, the change of the saturation vapor pressure for a temperature change from 5 °C BS: to 20 °C BS: for the dominant compounds (acetaldehyde and acetic acid) is small (less than a factor of 1.3; Betterton and Hoffmann (1988); Johnson et al. (1996)) BS: and, therefore, does not account for the observed

magnitude of increase (by a factor of ~ 5) from 06:00 - 09:00 LT[BS:] . Instead, the temperature dependence of the solubility of these compounds in an aqueous solution (factor 5-7) would explain a change of this magnitude. The sharp peaks observed in this factor during morning hours could be explained by the Kathmandu Valley meteorology. After sunrise when air temperatures start to rise, the boundary layer continues to be shallow until direct sunlight reaches the Valley bottom. The accumulation of compounds in a shallow boundary layer contributes to high ambient concentrations. The dilution due to the rising boundary layer and daytime westerly winds in the Valley reduces the concentrations subsequently. Therefore, this factor is attributed as solvent evaporation.

3.1.7 Factor 7 - Mixed daytime

Formic acid and acetic acid contributed most to the total measured NMVOC mass of Factor 7 ($\sim 25\,\%$ and $\sim 13\,\%$, respectively) while propyne, methanol and acetone together contributed $\sim 26\,\%$ ($\sim 10\,\%$, $\sim 8\,\%$ and $\sim 8\,\%$, respectively). The other 32 species collectively contributed $\sim 36\,\%$ to this factor but their individual contributions were $\leq 5\,\%$. Like factor 6, this factor, too, has a predominance of oxygenated compounds (that could be due to photooxidation) with a minor contribution from NMVOCs such as acetonitrile and propyne which can be emitted from primary emission sources such as biomass burning and industrial emissions (Hao et al., 1996; Andreae and Merlet , 2001; Akagi et al., 2011). The diel profile of this factor (Figure 10) is similar to that of the ambient temperature and solar radiation with an average mass concentration of ~ 20 $\mu g\,m^{-3}$ between 12:00 - 14:00 LT.

Approximately $41\,\%$ of the total formamide, $\sim 37\,\%$ of the total acetamide and $\sim 40\,\%$ of the total isocyanic acid are explained by this factor. Both formamide and acetamide can be produced by hydroxyl radical initiated photooxidation of primary amines (such as methyl amine) and in turn can photochemically form isocyanic acid through hydroxyl radical mediated oxidation (Roberts et al., 2014; Ge et al., 2011; Sarkar et al., 2016). In addition $34\,\%$ of the formic acid and $23\,\%$ of the formaldehyde mass was explained by this factor. The timeseries (Figure 10) of this factor showed higher baseline concentrations during second part of the measurement period when primary emissions were higher due to both biomass burning and biomass co-fired brick kiln emissions as described in Sarkar et al. (2016). During this period, influenced strongly by biomass burning sources, specific NMVOCs such as isocyanic acid, formamide and acetamide showed enhancement in their background concentrations. This is likely due to the higher emissions of precursor alkyl amines and other N-containing compounds from the incomplete combustion of biomass (Stockwell et al., 2015) which can form formamide and acetamide via photooxidation. Due to the contribution from both photooxidation and primary emissions, this factor was attributed as the mixed daytime factor.

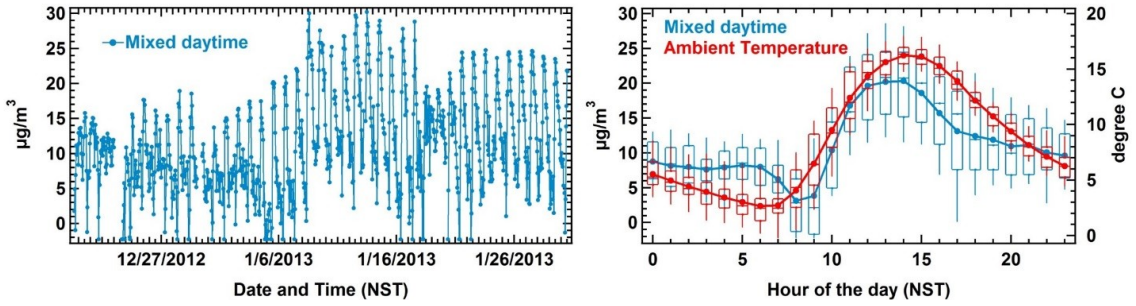


Figure 10. Timeseries and diel box and whisker plot for Factor 7 (Mixed daytime)

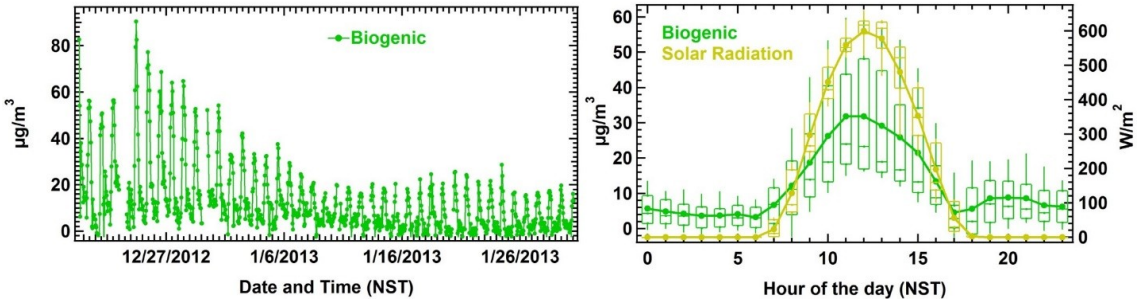


Figure 11. Timeseries and diel box and whisker plot for Factor 8 (Biogenic emissions)

3.1.8 Factor 8 - Biogenic emissions

Factor 8 explains more of the total isoprene mass than any of the other factors ($\sim 33\,\%$) and shows a distinct daytime peak with the highest mass loadings of $\sim 32\,\mu g\,m^{-3}$ observed between 11:00 - 12:00 LT (Figure 11). The diel profile (Figure 11) of this factor correlates best with solar radiation (R^2 = 0.33; as can be seen in Table S4 of the supplementary information and Figure S9) during the daytime hours (between 06:00 - 17:00 LT). Average nighttime concentrations of this factor were always less than $10\,\mu g\,m^{-3}$. The timeseries profile showed very high daytime mass loadings up to $\sim 80\,\mu g\,m^{-3}$ for the first part of the campaign (19 December 2012 – 2 January 2013) and lower mass loadings as the campaign progressed. This is also consistent with the observation of deciduous trees in the Kathmandu Valley shedding their leaves during peak winter (Sarkar et al., 2016). Therefore, the factor was attributed to biogenic emissions.

Most of the total measured NMVOC mass in this factor was associated with oxygenated NMVOCs namely acetaldehyde, acetic acid, acetone and formic acid which contributed $\sim 21\,\%$, $\sim 15\,\%$, $\sim 11\,\%$ and $\sim 10\,\%$, respectively to Factor 8. Isoprene contributed $\sim 8\,\%$ to the total NMVOC mass. The other 32 NMVOCs together contributed $\sim 35\,\%$.

To summarize, based on the characteristics observed in the factor profiles, factor timeseries and diel plots, Factor 1 was attributed to traffic (TR), Factor 2 was attributed to residential biofuel use

and waste disposal (RB+WD), Factor 3 was attributed to mixed industrial emissions (MI), Factor 4 was attributed to biomass co-fired brick kilns (BK), Factor 5 to unresolved industrial emissions (UI), Factor 6 was attributed to solvent evaporation (SE), Factor 7 was attributed to mixed daytime source (MD) and Factor 8 was attributed to biogenic NMVOC emissions (BG). Table S4 of the supplementary information shows the calculated correlation coefficients between the PMF resolved source factors and the independent meteorological parameters.

It can be seen from Table S4 of the supplementary information that during daytime, the solvent evaporation (SE) factor correlated best with the rate of change in solar radiation and the rate of change in ambient temperature (r = 0.62 and 0.64, respectively). This supports the assignment of the solvent evaporation factor as evaporation depends on temperature. BS: ~~However, the change of the saturation vapor pressure for a temperature change from 5 °C BS: to 20 °C BS: for the dominant compounds (acetaldehyde and acetic acid) is small (less than a factor of 1.3;~~ Betterton and Hoffmann (1988); Johnson et al. (1996)) BS: ~~and, therefore, does not account for the observed magnitude of increase (by a factor of ~ 5) from 06:00 - 09:00~~ LT BS: ~~. Instead, the temperature dependence of the solubility of these compounds in an aqueous solution (factor 5-7) would explain a change of this magnitude.~~ The solvent evaporation factor strongly anti-correlated with RH during the nighttime BS: (R=-0.59) and correlated well with the unresolved industrial (UI) factor (R = 0.55) BS: changes in solar radiation (R=0.62) and ΔT (R=0.64) during daytime. BS: ~~It is, therefore, possible that the sources of the solvent evaporation and unresolved industrial emission may be identical or at least spatially co-located. We hypothesize, that compounds, that firstly display a significant solubility in aqueous solution and secondly a strong temperature dependence of the solubility are attributed to this separate factor. At night, soluble compounds partition into the aqueous phase of the fog aerosol and hence their mixing ratios will not build up in the nocturnal boundary layer to the same extent as those of less soluble compounds despite continuous emissions from industrial units. Those compounds with a high temperature dependence on solubility like acids and aldehydes will rapidly shift to the gas phase from their nocturnal fog water reservoir when temperatures increase in the morning and their solubility decreases, which manifests itself in a disproportionate (considering only evaporation) increase of their mixing ratios at that time.~~ BS: While the correlation of the solvent evaporation factor with the unresolved industrial factor during daytime seems to suggest the two should be combined into one factor profile, several facts suggest against it. Firstly, the two do not correlate at night since the unresolved industrial factor shows a mild positive correlation rather than anti-correlation with RH at night (R=0.29) and no strong correlation with ΔT during the day (R=0.28). Secondly, the raw time series of 1,3-butadiyne and methanol (Supplementary Table S5) correlates extremely strongly (R=0.9), indicating there is a strong and unique common source which causes sharp spikes in these two compounds. The fact that the time series of 1,3-butadiyne correlates poorly with acetaldehyde, acetic acid and formic acid indicates that the solvent evaporation factor (which is not a significant source of 1,3-butadiyne and methanol), has very different emission ratios of 1,3-butadiyne to acetaldehyde, acetic acid and formic acid compared to the unresolved industrial emissions factor to explain the raw data. The fact that the time series of 1,3-butadiyne correlates equally poorly with that of ethanol, the defining compound of the mixed industrial factor, suggest against combining the

mixed industrial factor with the unresolved industrial factor. It, therefore, seems, that the unresolved industrial factor is related to primary emissions from a distinct source, while the source profile of the solvent evaporation factor may be strongly confounded by meteorology and chemistry. Confounding factors have been reported to affect PMF solutions previously (Yuan et al. 2012).

The mixed daytime factor (MD) correlated with solar radiation, ambient temperature and wind speed (r = 0.58, 0.74 and 0.57, respectively). The biogenic factor (BG) had the best correlation with solar radiation (r = 0.57) during daytime, consistent with its attribution to biogenic emissions. During daytime, the mixed industrial emissions and biomass co-fired brick kiln emissions had very low mass concentration due to the boundary layer dilution and ventilation effect of high westerly winds in the Kathmandu Valley (Sarkar et al., 2016). The ambient RH was also lower during the daytime. Therefore, both the mixed industrial emissions and brick kilns emission showed positive correlations with ambient RH (r = 0.65 and 0.74, respectively). During nighttime, no significant correlation was observed between the PMF resolved factors except the correlation of the biogenic factor with the residential biofuel use and waste disposal (RB+WD) factor (r = 0.58) which indicates that the high emissions of oxygenated NMVOCs and isoprene from RB+WD sources could result in a minor mis-attribution of the combustion derived emissions to the biogenic factor.

3.2 Conditional probability functions (CPF) to determine source directionality

Figure 12 shows the Conditional Probability Function (CPF) plots that were used to examine the spatial profile of the eight different PMF source factors. For the CPF plots, only data with wind speed $> 0.5\,\mathrm{ms}^{-1}$ were considered. Six factors namely traffic, residential biofuel use and waste disposal, mixed industrial emissions, unresolved industrial emissions, solvent evaporation and biomass co-fired brick kilns could be associated clearly with anthropogenic activities and are, therefore, likely to be impacted by spatially fixed sources, while one factor (mixed daytime) was related to photochemistry. One factor, biogenic emissions, is natural but can also be attributed to spatially fixed sources such as forests.

The CPF plot for the traffic factor showed maximum conditional probability (0.4 - 0.7) from the W-NW direction where the Kathmandu city center and the busiest traffic intersections were located. The conditional probability for the SW and NE wind direction ranged from 0.2-0.4. Two cities, namely Lalitpur (Patan) and Bhaktapur, respectively, are located upwind of the site in these directions. The lowest conditional probability was observed for the SE wind direction.

The residential biofuel use and waste disposal factor showed a high conditional probability of emissions exceeding the mean for air masses reaching the site from most wind directions (0.5 - 0.7 for NW-N, ~ 0.4 for N-NE and S-SW and 0.2 for E-S), indicating that this source is spatially distributed throughout the Kathmandu Valley. Only for the wind sector from SW-NW the conditional probability of this source is low. The reason for this low conditional probability is that every day in the afternoon, winds from the western mountain passes reach the receptor site. The same wind

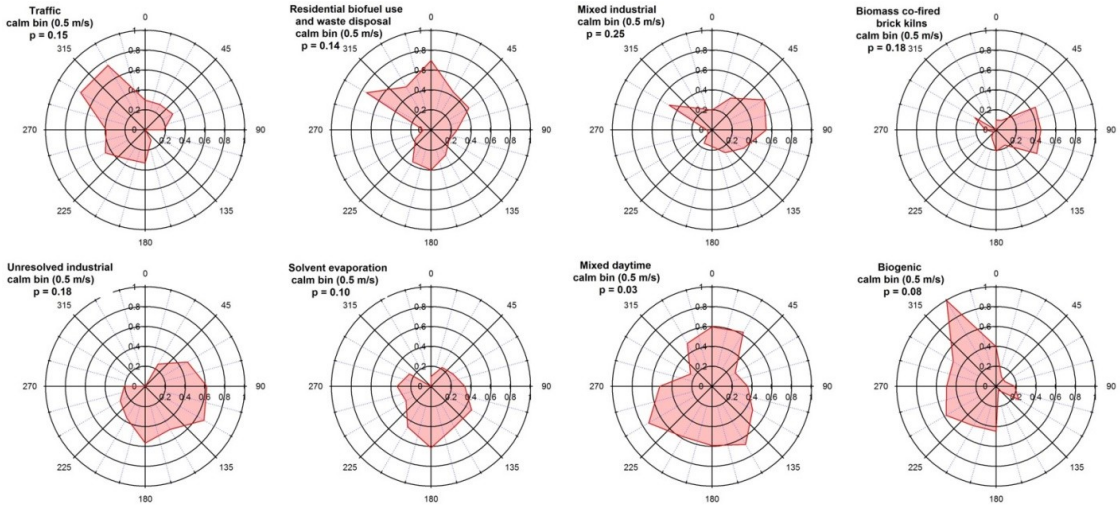


Figure 12. Conditional probability functions (CPF) plots for all source factors resolved by PMF showing wind directional dependency of different source categories

direction is extremely rare after sunset and during the early morning hours, when residential biofuel use and waste disposal mostly occur. Consequently, the conditional probability plot shows low conditional probabilities for this wind sector.

The mixed industrial emissions factor showed the highest conditional probability of air masses with above average mass loadings reaching the receptor site from the NE to SE wind sector (p = 0.4-0.6), where Bhaktapur industrial BS: ~~area~~ BS: estate is located within a distance of 3-4 km upwind of the receptor site. Conditional probabilities of 0.2-0.4 were observed for the NW wind direction where several industries are located.

For brick kilns the highest conditional probability was observed for air masses reaching the receptor site from the NE-SE ($p \sim 0.4$), which had several active brick kilns near the Bhaktapur Industrial Estate, which was $\sim 4\,\mathrm{km}$ upwind of the receptor site.

It is interesting to note that the unresolved industrial emissions factor shows a clear directional dependence (p = 0.5-0.7 for the NE-SW wind sector) indicating that this factor, too, can be attributed to spatially fixed sources in Bhaktapur Industrial Estate and Patan Industrial Estate. Polymer production, manufacturing industries for adhesives, paints and/or pharmaceuticals upwind of the site likely contributed towards the measured NMVOC mass of the unresolved industrial factor.

The solvent evaporation factor, too, shows high conditional probabilities for the SE-SW wind direction (Patan Industrial Estate) and low conditional probabilities for the NW-NE wind direction. The conditional probability function shows significant overlap with that of the unresolved industrial emissions factor. It therefore highlights the plausibility that solvent/chemical evaporation or emissions from industrial units are the primary source for this factor although the temperature changes after sunrise drives the partitioning into the gas phase.

Within the bin of calm wind speeds ($< 0.5\,\mathrm{ms}^{-1}$) the maximum conditional probabilities were observed for mixed industrial emissions, unresolved industrial emissions and brick kilns (0.25, 0.18 and 0.18, respectively) which indicates that emissions from these sources tended to accumulate in a shallow boundary layer during stagnant conditions in the Kathmandu Valley. Therefore, using taller chimney stacks, at least for combustion sources, to prevent accumulation of emissions in a shallow boundary layer could potentially improve the air quality of the Valley during foggy nights.

The mixed daytime factor shows no obvious directional dependence for the conditional probability of recording values above the average at the receptor site ($p > 0.3$ for all directions). Slightly higher conditional probabilities ($p \sim 0.6$) are recorded for air masses reaching the receptor site from the N-NE and S-SW wind direction.

The biogenic factor showed high conditional probabilities for air masses reaching the receptor site from the SW to N direction (p = 0.5 to 1) where few forested areas such as Nilbarahi jungle and Gokarna forest were located. Also forested areas in mountain slopes in the SW and NW direction and the midday fetch region being frequently from this sector explains the directional dependency of the biogenic factor.

The CPF analysis of the PMF model output clearly indicates that spatially fixed sources are responsible for a significant fraction of the overall measured NMVOC mass loadings and opens up the possibility to identify and mitigate emissions or at least the build-up of pollutants in a shallow inversion.

3.3 Source contribution to the total measured NMVOC mass loading and comparison with emission inventories

Figure 13 shows a pie chart summarizing contributions of individual sources to the total measured NMVOC mass loading. Total measured NMVOC mass loading was calculated by summing up the concentrations of individual measured NMVOCs (in $\mu\mathrm{g\,m}^{-3}$). The distribution shows that biogenic sources and the mixed daytime factor contributed only $10\,\%$ and $9.2\,\%$, respectively, to the total measured NMVOC mass loading while all the anthropogenic sources collectively contributed $\sim 80\,\%$ to the total measured NMVOC mass loading.

According to two widely used emission inventories, namely REAS v2.1 (Regional Emission inventory in ASia) and EDGAR v4.2 (Emissions Database for Global Atmospheric Research) (Kurokawa et al., 2013; Olivier et al., 1994) and the existing Nepalese inventory obtained from the International Centre for Integrated Mountain Development's (ICIMOD) database, residential biofuel use is considered to be the pre-dominant source of anthropogenic NMVOC emissions in Nepal. When the analysis is spatially restricted to the Kathmandu Valley for those inventories that provide gridded emissions (as shown in Figure 14), differences between EDGAR v4.2 and REAS v2.1 appear.

The EDGAR v4.2 inventory (for the BS: full year 2008) attributes only 10.6% of the total anthropogenic NMVOC emissions in the Kathmandu Valley (85.2-85.5 Longitude and 27.6-27.8 Latitude)

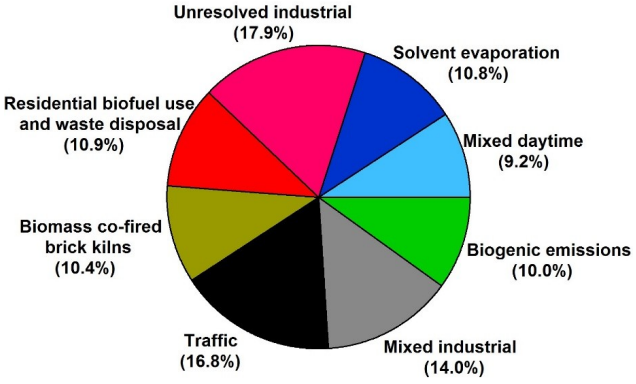


Figure 13. Contributions of various sources to the total NMVOC mass loading observed at Bode, a semi-urban site in the Kathmandu Valley

to be due to residential biofuel use and an additional 8.9% to solid waste disposal. These numbers are in reasonable agreement with our PMF output, which attributes 13.5% instead of 19.5% of the total measured NMVOC mass to these two sources combined. BS: EDGAR v4.2 inventory provides only spatially resolved data, not seasonally resolved data.

The REAS v2.1 inventory (for the year 2008) estimates that 67.2 % of the total wintertime (December and January) anthropogenic NMVOC emissions in the Kathmandu Valley (85.25–85.5 Longitude and 27.5–27.75 Latitude) originate from residential and commercial biofuel use — a significant overestimation when the numbers are compared to our PMF output and the EDGAR v4.1 inventory. The national Nepali emission inventory, too apportions a large share of the total national annual NMVOCs emissions to residential and commercial biofuel use (83.1 %). It, therefore, appears, that while apportioning the emissions spatially, the REAS v2.1 emission inventory does not fully account for the socio-economic differences between rural and urban areas. The EDGAR v4.2 emission inventory, on the other hand, seems to apportion most of the national consumption of LPG cooking gas to the highly urbanised Kathmandu valley and correspondingly scales down the emission from biofuel use within the Kathmandu valley. In absolute terms the annual NMVOC emissions attributed to domestic fuel usage within the Kathmandu valley by EDGAR v4.2 are a factor of 3.6 lower compared to the annual NMVOC emissions attributed to this sector by REAS v2.1.

The EDGAR inventory considers solvent use (66 %) and mixed industrial emissions to represent the second most important source of NMVOCs. Solvent use and other industrial emissions (8.5 %) combined account for 74.5 %. Collectively they are considered to contribute $\sim$ 10,% to the total anthropogenic NMVOC mass in the EDGAR v4.2 inventory, while our PMF results attribute 52.8 % of the measured NMVOCs to solvent use and industrial emissions combined.It should be noted, that solvent use and other factors related to industrial emissions (mixed industrial and unresolved

industrial) must be combined while comparing our PMF output with emission inventories. Both the mixed industrial emission factor and the unresolved industrial emission factor contain a significant NMVOC mass fraction from industrial solvent use, but also combustion related emissions from industrial units. Unfortunately, industrial solvent use and industrial combustion emissions from co-located units cannot be cleanly segregated using the PMF model, which relies on spatio-temporal patterns while building factor profiles. Overall, our PMF output agrees with the EDGAR v4.2 inventory, that industries are the dominant source of NMVOCs in the Kathmandu valley. According to the REAS v2.1 inventory, solvent use is considered to be the second most dominant contributor (29.8 %) to wintertime NMVOC emissions in the Kathmandu valley. Solvents and other industrial emissions (0.9 %) combined account for 30.7 % of the total wintertime NMVOC emissions in the REAS v2.1 emission inventory. Since, most of the national consumption of solvents and a significant share of Nepal's industrial production is concentrated in the Kathmandu valley, the discrepancies between the REAS v2.1 emission inventory and our results indicate, that the REAS v2.1 emission inventory does not sufficiently account for the special status of the Kathmandu valley while spatially apportioning emissions. The emissions that EDGAR v4.2 attributed to solid waste disposal, industries, the transport sector, and solvent use within the Kathmandu valley are a factor of 17.4, 14.0, 7.4 and 3.3 times higher compared to what the REAS v2.1 inventory attributes to the same sectors for the same geographical area.

The annual Nepalese inventory (for the year 2000) considers solvent and paint use to be the second largest contributor to the anthropogenic NMVOC emissions in Nepal, while industries are considered to make an insignificant overall contribution (0.7 %). These numbers cannot be compared to our results in a meaningful manner, as the national emissions in particular for sectors such as domestic fuel usage and agricultural waste burning may be dominated by the rural hinterland, while our PMF results apply to the largest urban agglomeration in Nepal.

Traffic was considered to contribute only between $\sim$ 1.3 % (in the REAS v2.1 inventory) to a maximum of $\sim$2.6 % (in EDGAR v4.2 inventory) of the total anthropogenic NMVOC emissions in the Kathmandu valley. This stands in stark contrast to the results of our PMF analyses, which indicate traffic contributes ca. 20 %, solvent evaporation and industrial solvent/chemical usage accounts for ca. 36 % (unresolved industrial emissions + solvent evaporation) and other industrial emissions (mixed industrial emissions + brick kilns) account for ca. 30 % of the total measured anthropogenic NMVOC mass loading in the Kathmandu valley. According to the recent study of the vehicle fleet in Kathmandu valley Shrestha et al. (2013), transport sector NMVOC emissions in the Kathmandu valley for the year 2010 amounted to 7654 $\mathrm{t\,y^{-1}}$, a number that is 10 times higher than the number currently in the EDGAR v4.2 inventory and 72 times higher than the number currently in the REAS v2.1 inventory. If the emission estimate of (Shrestha et al., 2013) was incorporated into EDGAR v4.2 inventory without any further changes, the percentage share of transport sector emissions to the total NMVOC emissions would increase to 38.7 %, while the contribution of domestic fuel usage

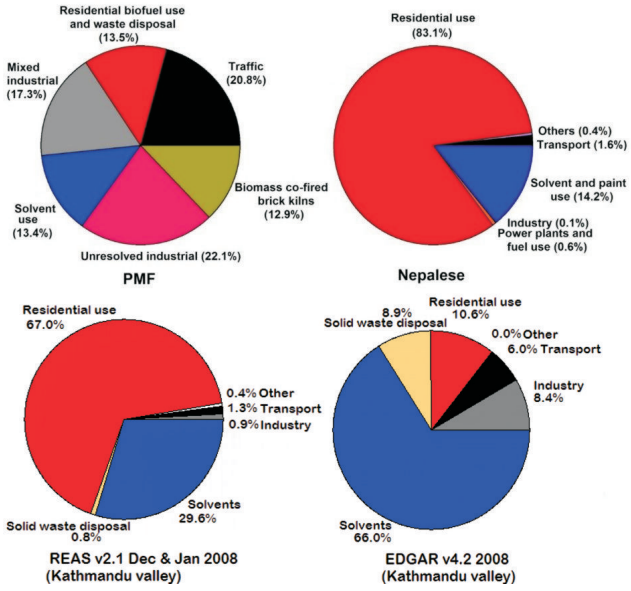


Figure 14. Comparison of the PMF derived contribution of anthropogenic sources with NMVOCs source contribution according to the existing Nepalese, REAS and EDGAR emission inventory

and waste disposal would drop to 12.7 %(PMF 13.5 %) and the contribution of industrial emissions and solvent use would drop to 48.6 % (PMF 52.8 %). Our PMF results, however, seem to suggest, that 2012 transport sector emissions have decreased by ∼50 % compared to the 2010 emissions presented in (Shrestha et al., 2013), possibly due to a reduction of the number older vehicles in the fleet.

Inefficient biomass co-fired brick kilns are a unique industrial source in the Kathmandu Valley, and contributed significantly (∼ 15 %) to the total measured anthropogenic NMVOC mass loading. The existing Nepalese inventory considers contributions of brick kilns only to the emission of particulate matter (PM_{10} and $PM_{2.5}$)), while the two other emission inventories do not include emissions from brick kilns in the Kathmandu Valley at all. If transport sector NMVOC emissions of ∼3800 $\mathrm{t\,y^{-1}}$ and an additional ∼2400 $\mathrm{t\,y^{-1}}$ NMVOC emissions from brick kilns, were included in the EDGAR v4.2 emission inventory, the EGAR emission inventory and our PMF output would agree perfectly (within ± 0.2 %) on the relative contribution of all sources, without changing the contribution from any of the other sources.

Only two sources, domestic fuel usage (on account of the changed heating demand) and agricultural waste burning are expected to have significant seasonality. Jointly, they account for less than 10 % of the total NMVOC emissions. Since cooking needs persist throughout the year and the decrease in agricultural waste burning outside harvest season may be partially offset by leaf-litter burning (a source currently not in the model), it is likely that the failure to account for seasonal effects imparts an uncertainty of less than 1 % on the overall result of our analysis.

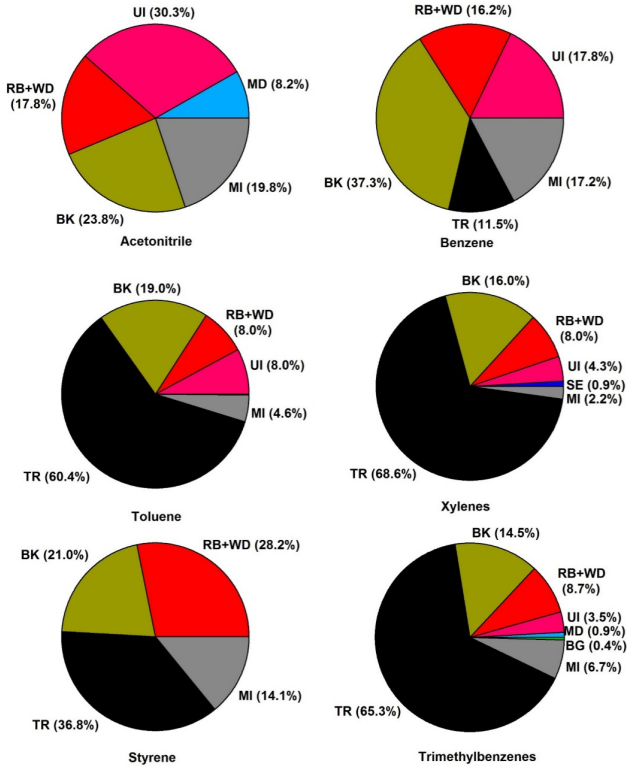


Figure 15. Contribution of PMF derived source factors to acetonitrile and aromatic NMVOCs. Source names are abreviated as follows: MD=mixed daytime, MI= mixed industrial, UI = unresolved industrial, BK = brick kiln, TR = traffic, RB+WD = residential burning and waste disposal, SE = solvent evaporation, BG = biogenic

The REAS v2.1 emission inventory for the Kathmandu valley, on the other hand, seems to require large corrections. While our analysis of the REAS inventory was restricted to December and January, annual averages of individual sources differ by less than $\pm$ 10 % from the winter values. Therefore, the difference in the time window selected for the analysis cannot explain the observed discrepancies to the EDGAR emission inventory.

3.4 Source contribution to individual NMVOCs

Figure 15 represents the pie charts showing contribution of the eight source factors to individual NMVOCs such as acetonitrile, benzene, styrene, toluene, sum of C8-aromatics (xylenes and ethylbenzene) and sum of C9-aromatics (trimethylbenzenes and propylbenzene). Maximum contribution to the acetonitrile mass concentration was observed from the unresolved industrial emission sources ($\sim$ 30 %) followed by the biomass co-fired brick kilns emission ($\sim$ 24 %) and mixed industrial emission ($\sim$ 20 %) factors. Residential biofuel use and waste disposal features only fourth ($\sim$ 18 %). The same sources also contribute most to benzene emissions, indicating that fuel usage, rather than its

Table 6. Emission ratios of NMVOCs/benzene for acetonitrile and aromatic hydrocarbons derived from the PMF model for different sources and comparison with the ratios for different source categories reported in previous studies.

ERs/Benzene	RB+WD	BK	MI	UI	Garbage burning grab samples	Waste burning[1]	Wood burning[2]	Charcoal burning[2]
Acetonitrile	0.23	0.14	0.25	0.36	0.77	0.06	-	-
Toluene	0.34	0.35	0.18	0.30	0.34	0.41	0.05	0.50
C8-aromatics	0.18	0.06	0.08	0.00	0.25	0.10	-	0.46
C9-aromatics	0.25	0.22	0.06	0.12	0.08	0.03	-	-
Styrene	0.12	0.09	0.09	0.04	0.16	0.86	-	-
Naphthalene	0.11	0.15	0.20	0.05	0.09	0.10	-	-

1. Stockwell et al. (2015); 2. Tsai et al. (2003); RB+WD = Residential biofuel use and waste disposal; BK = Biomass co-fired brick kilns; MI = Mixed industrial emissions; UI = Unresolved industrial emissions

application as solvent/chemical reagents in industrial processes is responsible for most of the industrial acetonitrile emissions. It also indicates that industrial rather than residential biofuel usage contributes more towards outdoor NMVOC air pollution. Most of the benzene (which is a human carcinogen) can be attributed to biomass co-fired brick kilns ($\sim 37\,\%$), mixed industrial ($\sim 17\,\%$) and unresolved industrial ($\sim 18\,\%$) sources. Residential biofuel use again featured only fourth as far as the contribution towards mixing ratios of this compound in the outdoor environment is concerned. Table 6 shows a comparison of NMVOCs/benzene emission ratios for four PMF derived sources (residential biofuel use and waste disposal, biomass co-fired brick kilns, mixed industrial and unresolved industrial sources) to the emission ratios obtained from the grab samples collected for garbage burning in the Kathmandu Valley and the previously reported emission ratios for waste burning, wood burning and charcoal burning sources.

Residential biofuel use and waste disposal contributed $\sim 28\,\%$ of the total styrene which were emitted significantly from waste burning. However, traffic was found to be equally important as a styrene source ($\sim 37\,\%$) in the Kathmandu Valley. Recently, styrene has been detected from traffic and was found to have high emission ratios with respect to benzene after cold startup of engines and in LPG fuel (Alves et al., 2015). Biomass co-fired brick kilns and mixed industrial emissions also contribute significantly ($\sim 21\,\%$ and $\sim 14\,\%$, respectively) towards styrene mass loadings. Traffic was found to be the most important source of higher aromatics including toluene, C8-aromatics, and C9-aromatics ($> 60\,\%$). Biomass co-fired brick kilns were the second largest contributors towards their mass loadings, while residential biofuel usage and waste disposal ranked third.

Figure 16 shows the pie charts summarizing contributions of PMF derived sources to two newly quantified compounds in the Kathmandu Valley, namely formamide and acetamide along with isocyanic acid and formic acid. All these compounds showed maximum contribution from the mixed

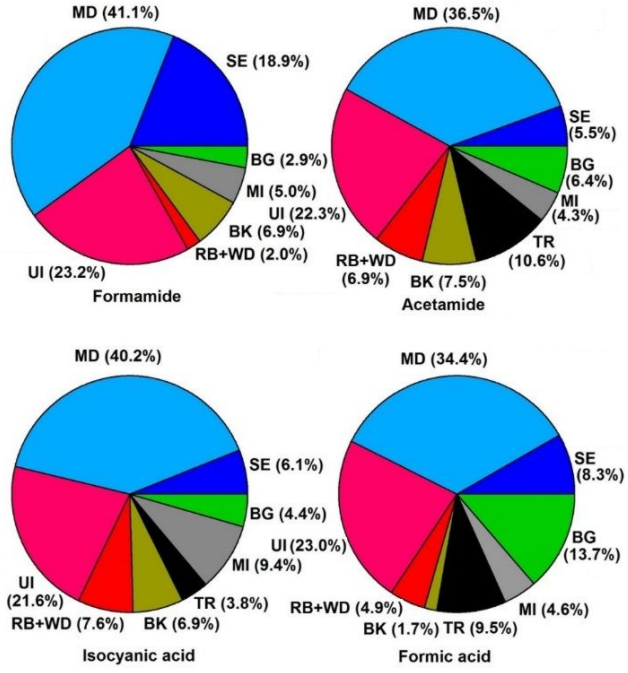


Figure 16. Contribution of PMF derived sources to formamide, acetamide, isocyanic acid and formic acid. Source names are abreviated as follows: MD=mixed daytime, MI= mixed industrial, UI = unresolved industrial, BK = brick kiln, TR = traffic, RB+WD = residential burning and waste disposal, SE = solvent evaporation, BG = biogenic

daytime factor ($\sim 34\,\%$ to $\sim 41\,\%$) due to the photo-oxidation source. As discussed previously in Sarkar et al. (2016) and in section 3.1.7, both formamide and acetamide are formed primarily as a result of photooxidation of amine compounds and N-containing compounds. These can be emitted from the various inefficient combustion processes in the Kathmandu Valley. Photooxidation of these amides further forms isocyanic acid (reaction schematic is shown in Figure S8 of the supplementary information). Apart from the mixed daytime source, unresolved industrial emissions factor also contributed significantly to all these compounds ($\sim 22\,\%$ to $\sim 23\,\%$) as they are used as reactants (e.g. formic acid is used as reactant to produce formamide in industries) or produced during different industrial processes (such as formamide is produced in pharmaceuticals and plastic industries). Solvent evaporation factor contributed $\sim 19\,\%$ to formamide while biogenic factor contributed $\sim 14\,\%$ to formic acid. Contributions from all the other sources to these NMVOCs were $< 10\,\%$.

Figure 17 represents the pie charts showing contribution of the eight sources derived from PMF to 1,3-butadiyne and oxygenated compounds namely methanol, acetone, acetaldehyde, ethanol and acetic acid. It can be seen from Figure 17 that emissions of all these compounds in the Kathmandu Valley were dominated by different industrial activities. The total unresolved industrial emis-

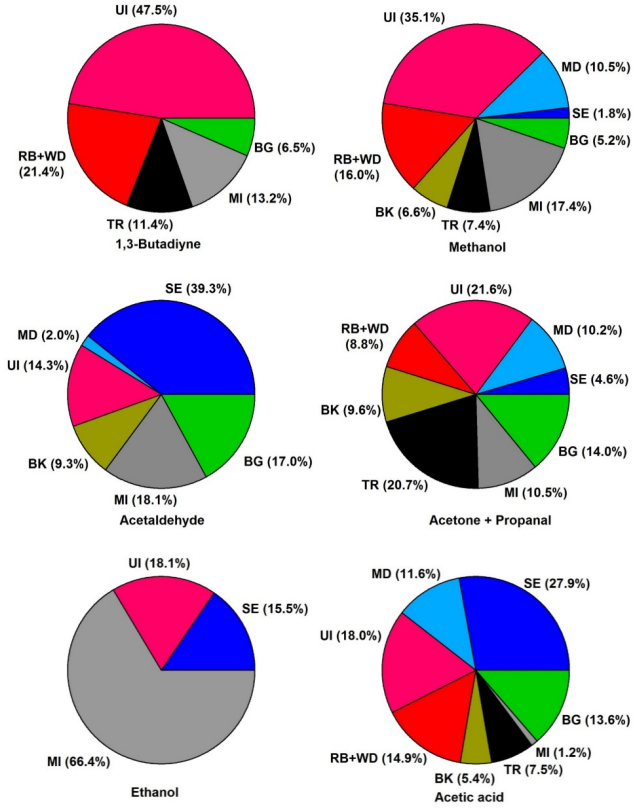


Figure 17. Contribution of PMF derived sources to 1,3-butadiyne and oxygenated NMVOCs such as methanol, acetone, acetaldehyde, ethanol and acetic acid. Source names are abreviated as follows: MD=mixed daytime, MI= mixed industrial, UI = unresolved industrial, BK = brick kiln, TR = traffic, RB+WD = residential burning and waste disposal, SE = solvent evaporation, BG = biogenic

sions factor dominated the contribution to 1,3-butadiyne ($\sim 48\,\%$), methanol ($\sim 35\,\%$) and acetone ($\sim 22\,\%$). Residential biofuel use and waste disposal also contributed significantly to 1,3-butadiyne ($\sim 21\,\%$) and methanol ($\sim 16\,\%$). Traffic was found to have significant contribution to acetone ($\sim 21\,\%$). It is known that acetaldehyde, ethanol and acetic acid are used as solvents in different industries and it was found that industrial sources obtained from PMF (mixed industrial + unresolved industrial + solvent evaporation) together contributed $\sim 72\,\%$ of the total acetaldehyde, $100\,\%$ of the total ethanol and $\sim 47\,\%$ of the total acetic acid. Biogenic sources also had significant contribution to acetaldehyde and acetic acid ($\sim 17\,\%$ and $\sim 14\,\%$, respectively) whereas residential biofuel use and waste disposal contributed to $\sim 15\,\%$ of the total acetic acid.

Figure 18 represents a timeseries of daily mean relative contribution of the PMF derived sources during SusKat-ABC campaign. As discussed in Sarkar et al. (2016), the whole campaign can be divided into three different periods based on the measurements – first period (from the start of the campaign until 3 January 2013) was associated with high daytime isoprene emissions due to strong

biogenic emissions, the second period (4 – 18 January 2013) was marked by enhancements in acetonitrile and benzene concentrations due to the kick start of the biomass co-fired brick kilns in the Kathmandu Valley and in the third period (19 January until the end of the campaign), more oxygenated NMVOCs were observed which was believed to be due to the stable operation of the brick kilns and more contribution from the industrial sources. PMF derived results also supports these observation as can be seen in Figure 18. It can be seen that from the start of the campaign until 3 January 2013 contribution of PMF derived biogenic sources were $> 20\,\%$ for most of the time while contribution from the brick kilns emission factor was negligible ($\leq 5\,\%$). From 4 January until 18 January 2013, the contribution of brick kilns increased significantly ($\sim 20\,\%$ to $\sim 40\,\%$) as almost all brick kilns in the Kathmandu Valley became operational. After 18 January until the end of the campaign, the contribution of brick kilns become lower due to its stable operation.

During the first period, contribution of traffic was found to be higher ($\sim 20\,\%$ to $\sim 30\,\%$) compared to the rest of the campaign. The higher contribution of the mixed daytime source during the second and third part of the campaign was due to the early morning and daytime photooxidation of the precursor compounds which were emitted as a result of biomass co-fired brick kilns and other biomass burning emissions during these periods. The mixed industrial emissions factor contributed almost equally throughout the campaign (contributing $\sim 10\,\%$ to $\sim 15\,\%$) but the solvent evaporation and the unresolved industrial emissions factor contributed more during the second and third part of the campaign (increase of $\sim 10\,\%$).

3.5 Source contribution to daytime ozone production potential and SOA formation

Figure 19a represents the source contribution to daytime O_3 production potential while Figure 19b represents the contribution of different classes of compounds measured in the Kathmandu Valley to the daytime O_3 production potential as discussed in Sarkar et al. (2016). The daytime O_3 production potential for individual sources was calculated by summing up the O_3 production potential for the individual compounds which was calculated according to the method described by Sinha et al. (2012). The distribution of the daytime O_3 production potential obtained from the measurements (Figure 19b) shows that $\sim 70\,\%$ of the total daytime O_3 production potential was due to the contribution from isoprene and oxygenated NMVOCs which [BS:] could presumptuously indicated dominance of biogenic emissions and photochemistry in the Kathmandu Valley even in the winter. But the distribution of different sources obtained from PMF to daytime O_3 production potential shows that the biogenic factor together with the photochemistry factor (mixed daytime) contributed only $\sim 30\,\%$ of the total O_3 production potential. The remaining $\sim 70\,\%$ was contributed by anthropogenic sources. While solvent evaporation contributed most ($\sim 20\,\%$) to the total daytime O_3 production potential, traffic and unresolved industrial emission stood second and third, respectively, in terms of anthropogenic ozone precursor emissions. Residential biofuel use and waste disposal, and biomass co-fired

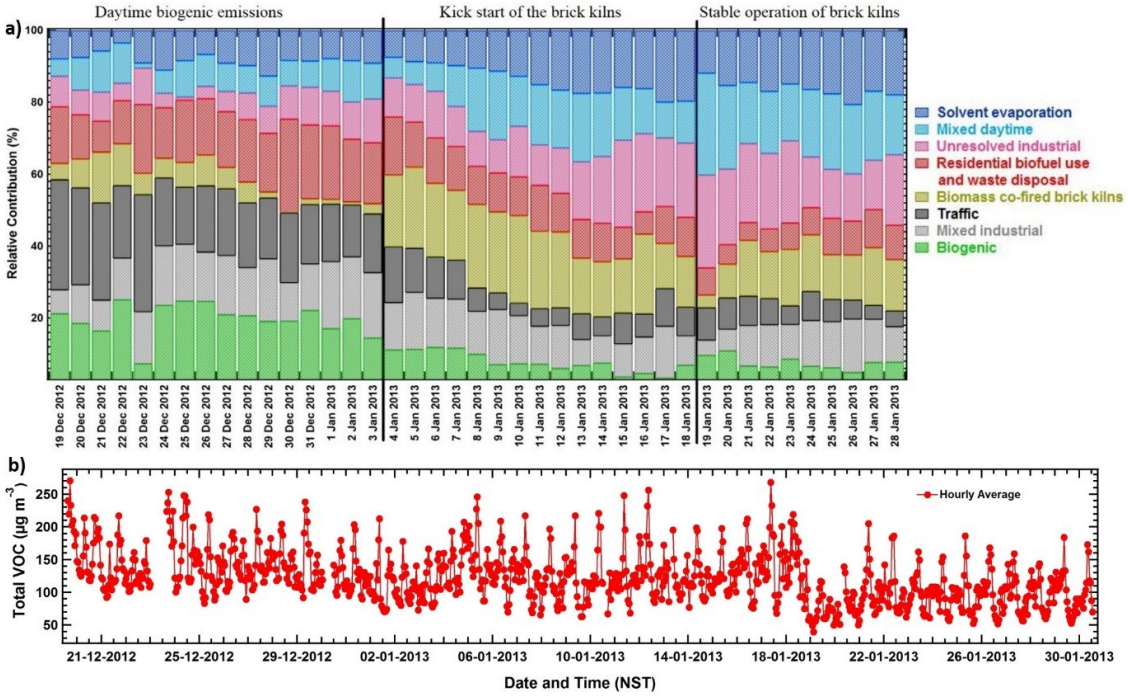


Figure 18. Daily mean relative contribution of PMF derived eight sources during SusKat-ABC campaign

brick kilns while potentially important from a human health perspective, contributed only a minor fraction of the total anthropogenically emitted ozone precursors.

The consequence of including only a subset of NMVOCs is an underestimation of the OH reactivity and hence ozone production potential, which scales directly with the OH reactivity. BS: For the city of Lahore, Barletta et al. (2016) BS: , reported the maximum contribution of BS: ~~Based on measured~~ methane and 63 non methane hydrocarbon BS: to the measured OH reactivity as 14%. BS: ~~measurements in the city of~~ Lahore BS: ~~which~~ is much larger, and by all indications more polluted BS: city, than Kathmandu. Barletta et al. (2016) BS: ~~the authors reported a maximum contribution of about~~ 14% BS: ~~due to all alkanes including methane to the total measured OH reactivity~~. Despite high concentration abundances in urban atmospheric environments, the rate constants of these species are typically 100 times lower than compounds like isoprene, and hence their contribution to the total OH reactivity is much lower. For example, even 3 ppm methane (observed only in plumes) would contribute only $\sim 0.5\ s^{-1}$ to the total OH reactivity and hence make an insignificant contribution to the ozone production potential. Hence, our analyses of the ozone production potential may underestimate the total ozone production potential by 15–25%, if we can extrapolate the observations from another South Asian city like Lahore.

Secondary organic aerosol (SOA) production was calculated using the concentrations and the known SOA yields for benzene, toluene, styrene, xylene, trimethylbenzenes, naphthalene and isoprene (Ng et al., 2007; Chan et al., 2009; Yuan et al., 2013; Kroll et al., 2006). As the biomass co-fired brick kilns and the traffic factor contains most of the reactive aromatic compounds, they ap-

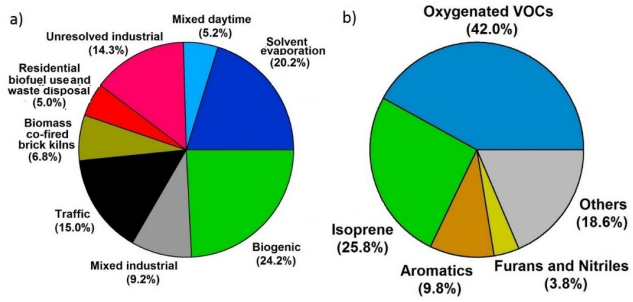


Figure 19. Daytime O_3 production potential obtained a) from the source contribution using PMF and b) from the measurements performed in the Kathmandu Valley

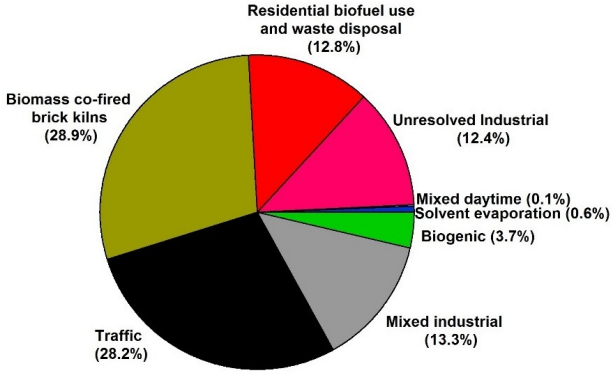


Figure 20. Contribution of PMF derived eight sources to the SOA formation in the Kathmandu Valley

peared to be the dominant contributors to SOA production (as shown in Figure 20) in the Kathmandu Valley.

4 Conclusions

The PMF model results reveal several new results regarding the source apportionment of NMVOCs in the Kathmandu Valley. Speciation of NMVOCs in the emission inventory for Nepal only includes compound classes (e.g. alkanes, alkenes etc.) and not specific compounds. BS: This imposes certain limitations while comparing emission inventories with the compounds measured in our study. However, BS: ~~Also,~~ the existing emission inventories (e.g. REAS v2.1, EDGAR v4.2; Kurokawa et al. (2013); Olivier et al. (1994) and Nepalese inventory (ICIMOD)) are highly uncertain as there has been no validation using in-situ measurements of these mostly bottom up inventories which rely on fuel and source emission factors measured in other technologically different regions of the world (primarily the US and Europe). By using the specific NMVOC emission tracer data measured in the Kathmandu Valley and constraining the PMF with measured source profiles of complex sources (e.g.

biomass co-fired brick kilns, residential solid biofuel use and waste disposal), it is shown that the contribution from sources such as residential solid biofuel use and waste disposal is overestimated in the REAS v2.1 emission inventory. At the same time, the emissions from industrial sources are underestimated. Both REAS v2.1 and EDGAR v4.2 underestimate the contribution of traffic and do not include brick kiln emissions. BS: ~~The presence of elevated concentrations of several health relevant NMVOCs (e.g. benzene) could be attributed to the biomass co-fired brick kiln sources.~~ BS: Eight different NMVOC sources were identified by the PMF model using the new "constrained model operation" mode. Unresolved industrial emissions (17.8 %BS:), traffic (16.8 %BS:) and mixed industrial emissions (14.0 %BS:) contributed most to the total measured NMVOC mass loading, while biogenic emissions (24.2 %BS:), solvent evaporation (20.2 %BS:), traffic (15.0 %BS:) and unresolved industrial emissions (14.3 %BS:) were the most important contributors to the ozone formation potential. Biomass co-fired brick kilns and traffic contributed approximately equally to the secondary organic aerosol (SOA) production (28.9 %BS: and 28.2 %BS:, respectively), while the most important contributors to the mass loadings of carcinogenic benzene were brick kilns (37.3 %BS:), unresolved industrial (17.8 %BS: and mixed industrial (17.2 %BS:) sources. Photo-oxidation (mixed daytime factor) contributed majorly to two newly identified ambient compounds namely, formamide (41.1 %BS:) and acetamide (36.5 %BS:) along with their photooxidation product isocyanic acid (40.2 %).

This study has provided quantitative information regarding the contributions of the major NMVOC sources in the Kathmandu Valley. This will enable focused mitigation efforts by policy makers and practitioners to improve the air quality of the Kathmandu Valley by reducing emissions of both toxic NMVOCs and formation of secondary pollutants. The results will also enable significant improvements in existing NMVOC emission inventories so that chemical-transport models can be parameterized more accurately over the South Asian region and the air quality-climate predictions by models can become more reliable.

Authors' contributions

Sections of this study were submitted in part fulfilment of the PhD work of C.S. carried out under the supervision of V.S. at IISER Mohali. The VOC dataset QA/QC and analyses were performed by C.S. and V.S. whereas B.S. designed and set up the PMF model and ensured QA/QC of PMF output which was performed by C.S. A.P. helped with interpretation of PMF results and suggested grab sampling experiments at an early stage. C.S., V.S. and B.S. wrote the paper and all co-authors discussed the results and commented on the paper.

Acknowledgements. Chinmoy Sarkar and Vinayak Sinha acknowledge the support extended by the Founding Director of IISER Mohali, Professor N. Sathyamurthy to enable participation of the IISER Mohali team in the SusKat-ABC campaign. Chinmoy Sarkar acknowledges the Ministry of Human Resources and Development

(MHRD), India and IASS Potsdam, Germany for funding with a service contract. IASS Potsdam funded the deployment of the PTR-TOF-MS by the IISER Mohali team in Kathmandu and local logistical support was provided by Khadak S. Mahata, Dipesh Rupakheti, Bhogendra Kathayat at the Bode site.

This study was partially supported by core funds of ICIMOD contributed by the governments of Afghanistan, Australia, Austria, Bangladesh, Bhutan, China, India, Myanmar, Nepal, Norway, Pakistan, Switzerland, and the United Kingdom.

All the data reported in this article can be obtained from the corresponding author by sending an email to vsinha@iisermohali.ac.in.

References

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, Atmos. Chem. Phys., 11, 4039-4072, doi:10.5194/acp-11-4039-2011, 2011.

Alves, C. A., Lopes, D. J., Calvo, A. I., Evtyugina, M., Rocha, S., and Nunes, T.: Emissions from Light-Duty Diesel and Gasoline in-use Vehicles Measured on Chassis Dynamometer Test Cycles, Aerosol and Air Quality Research, doi:10.4209/aaqr.2014.01.0006, 2015.

Anderson, M. J., Daly, E. P., Miller, S. L., and Milford, J. B.: Source apportionment of exposures to volatile organic compounds: II. Application of receptor models to TEAM study data, Atmospheric Environment, 36, 3643-3658, doi:10.1016/S1352-2310(02)00280-7, 2002.

Andreae, M. O., and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochemical Cycles, 15, 955-966, doi:10.1029/2000gb001382, 2001.

Barletta, B., Simpson, I. J., Blake, N. J., Meinardi, S., Emmons, L. K., Aburizaiza, O. S., Siddique, A., Zeb, J., Yu, L. E., Khwaja, H. A., Farrukh, M. A. and Blake, D. R.: Characterization of carbon monoxide, methane and nonmethane hydrocarbons in emerging cities of Saudi Arabia and Pakistan and in Singapore, J. Atmos. Chem., 11, 2399-2421, doi:10.1007/s10874-016-9343-7, 2016.

Betterton, E. A., and Hoffmann, M. R.: Henry's law constants of some environmentally important aldehydes, Environmental Science & Technology, 22, 1415-1418, doi:10.1021/es00177a004, 1988.

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, doi:10.5194/acp-11-2399-2011, 2011.

Borbon, A., Fontaine, H., Veillerot, M., Locoge, N., Galloo, J. C., and Guillermo, R.: An investigation into the traffic-related fraction of isoprene at an urban location, Atmospheric Environment, 35, 3749-3760, doi:10.1016/S1352-2310(01)00170-4, 2001.

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, Journal of Geophysical Research: Atmospheres, 118, 2041-2057, doi:10.1002/jgrd.50059, 2013.

Brown, S. G., Frankel, A., and Hafner, H. R.: Source apportionment of VOCs in the Los Angeles area using positive matrix factorization, Atmospheric Environment, 41, 227-237, doi: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, Science of The Total Environment, 518-519, 626-635, doi:10.1016/j.scitotenv.2015.01.022, 2015.

Buzcu, B., and Fraser, M. P.: Source identification and apportionment of volatile organic compounds in Houston, TX, Atmospheric Environment, 40, 2385-2400, doi:10.1016/j.atmosenv.2005.12.020, 2006.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, 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, doi:10.5194/acp-9-3049-2009, 2009.

Chandra, 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, International Journal of Mass Spectrometry, doi:10.1016/j.ijms.2017.05.008, 2017.

Chen, W. T., Shao, M., Lu, S. H., Wang, M., Zeng, L. M., Yuan, B., and Liu, Y.: Understanding primary and secondary sources of ambient carbonyl compounds in Beijing using the PMF model, Atmos. Chem. Phys., 14, 3047-3062, 10.5194/acp-14-3047-2014, 2014.

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 Prevot, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, Atmos. Chem. Phys., 13, 8411-8426, 10.5194/acp-13-8411-2013, 2013.

Fleming, Z. L., Monks, P. S., and Manning, A. J.: Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition, Atmospheric Research, 104-105, 1-39, doi:10.1016/j.atmosres.2011.09.009, 2012.

Gaimoz, C., Sauvage, S., Gros, V., Herrmann, F., Williams, J., Locoge, N., Perrussel, O., Bonsang, B., dÁrgouges, O., Sarda-Estéve, R., and Sciare, J.: Volatile organic compounds sources in Paris in spring 2007. Part II: source apportionment using positive matrix factorisation, Environmental Chemistry, 8, 91-103, doi:dx.doi.org/10.1071/EN10067, 2011.

Ge, X., Wexler, A. S., and Clegg, S. L.: Atmospheric amines - Part I. A review, Atmospheric Environment, 45, 524-546, doi:10.1016/j.atmosenv.2010.10.012, 2011.

Gueneron, M., Erickson, M. H., VanderSchelden, G. S., and Jobson, B. T.: PTR-MS fragmentation patterns of gasoline hydrocarbons, International Journal of Mass Spectrometry, 379, 97–109 doi:10.1016/j.ijms.2015.01.001, 2015.

Guo, H., Wang, T., and Louie, P. K. K.: Source apportionment of ambient non-methane hydrocarbons in Hong Kong: Application of a principal component analysis/absolute principal component scores (PCA/APCS) receptor model, Environmental Pollution, 129, 489-498, doi:10.1016/j.envpol.2003.11.006, 2004.

Guo, H., Wang, T., Blake, D. R., Simpson, I. J., Kwok, Y. H., and Li, Y. S.: Regional and local contributions to ambient non-methane volatile organic compounds at a polluted rural/coastal site in Pearl River Delta, China, Atmospheric Environment, 40, 2345-2359, doi:10.1016/j.atmosenv.2005.12.011, 2006.

Hao, W. M., Ward, D. E., Olbu, G., and Baker, S. P.: Emissions of CO_2, CO, and hydrocarbons from fires in diverse African savanna ecosystems, Journal of Geophysical Research: Atmospheres, 101, 23577-23584, doi:10.1029/95JD02198, 1996.

Hellèn, H., Tykkä, T., and Hakola, H.: Importance of monoterpenes and isoprene in urban air in northern Europe, Atmospheric Environment, 59, 59-66, doi:10.1016/j.atmosenv.2012.04.049, 2012.

Hewitt, C. N.: Reactive Hydrocarbons in the Atmosphere, Academic Press, ISBN: 978-0-12-346240-4, San Diego, 1999.

Ho, K. F., Lee, S. C., Ho, W. K., Blake, D. R., Cheng, Y., Li, Y. S., Ho, S. S. H., Fung, K., Louie, P. K. K., and Park, D.: Vehicular emission of volatile organic compounds (VOCs) from a tunnel study in Hong Kong, Atmos. Chem. Phys., 9, 7491-7504, doi:10.5194/acp-9-7491-2009, 2009.

Hwa, M. Y., Hsieh, C. C., Wu, T. C., and Chang, L. F. W.: Real-world vehicle emissions and VOCs profile in the Taipei tunnel located at Taiwan Taipei area, Atmospheric Environment, 36, 1993-2002, doi:10.1016/S1352-2310(02)00148-6, 2002.

Inomata, S., Tanimoto, H., Fujitani, Y., Sekimoto, K., Sato, K., Fushimi, A., Yamada, H., Hori, S., Kumazawa, Y., Shimono, A., and Hikida, T.: On-line measurements of gaseous nitro-organic compounds in diesel vehicle exhaust by proton-transfer-reaction mass spectrometry, Atmospheric Environment, 73, 195-203, doi:10.1016/j.atmosenv.2013.03.035, 2013.

Inomata, S., Fujitani, Y., Fushimi, A., Tanimoto, H., Sekimoto, K., and Yamada, H.: Field measurement of nitromethane from automotive emissions at a busy intersection using proton-transfer-reaction mass spectrometry, Atmospheric Environment, 96, 301-309, doi:10.1016/j.atmosenv.2014.07.058, 2014.

IPCC: Impacts, Adaptation and Vulnerability : Working Group II Contribution to the Intergovernmental Panel on Climate Change : Fifth Assessment Report (AR5): Summary for Policymakers. , Intergovernmental Panel on Climate Change. Working Group Impacts, 2013.

Johnson, B., Betterton, E., and Craig, D.: Henry's law coefficients of formic and acetic acids, Journal of Atmospheric Chemistry, 24, 113-119, doi:10.1007/bf00162406, 1996.

Jorquera, H., and Rappenglück, B.: Receptor modeling of ambient VOC at Santiago, Chile, Atmospheric Environment, 38, 4243-4263, doi:10.1016/j.atmosenv.2004.04.030, 2004.

Kaltsonoudis, C., Kostenidou, E., Florou, K., Psichoudaki, M., and Pandis, S. N.: Temporal variability and sources of VOCs in urban areas of Eastern Mediterranean, Atmos. Chem. Phys., 16, 14825-14842, 10.5194/acp-16-14825-2016, 2016.

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 proton-transfer-reaction mass spectrometry to characterize volatile organic compound sources at the La Porte super site during the Texas Air Quality Study 2000, Journal of Geophysical Research: Atmospheres, 108, 4508, doi:10.1029/2002jd003333, 2003.

Kim, E., Brown, S. G., Hafner, H. R., and Hopke, P. K.: Characterization of non-methane volatile organic compounds sources in Houston during 2001 using positive matrix factorization, Atmospheric Environment, 39, 5934-5946, doi:10.1016/j.atmosenv.2005.06.045, 2005.

Kim, K. H., Hong, Y. J., Pal, R., Jeon, E. C., Koo, Y. S., and Sunwoo, Y.: Investigation of carbonyl compounds in air from various industrial emission sources, Chemosphere, 70, 807-820, doi:10.1016/j.chemosphere.2007.07.025, 2008.

Kristensson, A., Johansson, C., Westerholm, R., Swietlicki, E., Gidhagen, L., Wideqvist, U., and Vesely, V.: Real-world traffic emission factors of gases and particles measured in a road tunnel in Stockholm, Sweden, Atmospheric Environment, 38, 657-673, doi:10.1016/j.atmosenv.2003.10.030, 2004.

Kroll, J. H., Ng, N. L., Murphy, S. M., Flagan, R. C., and Seinfeld, J. H.: Secondary Organic Aerosol Formation from Isoprene Photooxidation, Environmental Science & Technology, 40, 1869-1877, doi:10.1021/es0524301, 2006.

Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima, K., and Akimoto, H.: Emissions of air pollutants and greenhouse gases over Asian regions during 2000-2008: Regional Emission inventory in ASia (REAS) version 2, Atmos. Chem. Phys., 13, 11019-11058, doi:10.5194/acp-13-11019-2013, 2013.

Lemieux, P. M., Lutes, C. C., and Santoianni, D. A.: Emissions of organic air toxics from open burning: a comprehensive review, Progress in Energy and Combustion Science, 30, 1-32, doi:10.1016/j.pecs.2003.08.001, 2004.

Leuchner, M., and Rappenglück, B.: VOC source-receptor relationships in Houston during TexAQS-II, Atmospheric Environment, 44, 4056-4067, doi:10.1016/j.atmosenv.2009.02.029, 2010.

Lindinger, W., Hansel, A., and Jordan, A.: On-line monitoring of volatile organic compounds at pptv levels by means of proton transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research, Int. J. Mass Spectrom, 173, 191–241, doi:10.1016/s0168-1176(97)00281-4, 1998.

Miller, S. L., Anderson, M. J., Daly, E. P., and Milford, J. B.: Source apportionment of exposures to volatile organic compounds. I. Evaluation of receptor models using simulated exposure data, Atmospheric Environment, 36, 3629-3641, doi:10.1016/S1352-2310(02)00279-0, 2002.

Misztal, P. K., Hewitt, C. N., Wildt, J., Blande, J. D., Eller, A. S. D., Fares, S., Gentner, D. R., Gilman, J. B., Graus, M., Greenberg, J., Guenther, A. B., Hansel, A., Harley, P., Huang, M., Jardine, K., Karl, T., Kaser, L., Keutsch, F. N., Kiendler-Scharr, A., Kleist, E., Lerner, B. M., Li, T., Mak, J., Nölscher, A. C., Schnitzhofer, R., Sinha, V., Thornton, B., Warneke, C., Wegener, F., Werner, C., Williams, J., Worton, D. R., Yassaa, N., and Goldstein, A. H.: Atmospheric benzenoid emissions from plants rival those from fossil fuels, Scientific Reports, 5, 12064, doi:10.1038/srep12064, 2015.

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, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010jd014762, 2011.

Na, K., and Pyo Kim, Y.: Chemical mass balance receptor model applied to ambient C2-C9 VOC concentration in Seoul, Korea: Effect of chemical reaction losses, Atmospheric Environment, 41, 6715-6728, doi:10.1016/j.atmosenv.2007.04.054, 2007.

Nakashima, Y., Kamei, N., Kobayashi, S., and Kajii, Y.: Total OH reactivity and VOC analyses for gasoline vehicular exhaust with a chassis dynamometer, Atmospheric Environment, 44, 468-475, doi:10.1016/j.atmosenv.2009.11.006, 2010.

Ng, N. L., Kroll, J. H., Chan, A. W. H., Chhabra, P. S., Flagan, R. C., and Seinfeld, J. H.: Secondary organic aerosol formation from m-xylene, toluene, and benzene, Atmos. Chem. Phys., 7, 3909-3922, doi:10.5194/acp-7-3909-2007, 2007.

Norris, G., Vedantham, R., Wade, K., Zahn, P., Brown, S., Prouty, J. D., and Foley, C.: EPA positive matrix factorization (PMF) 3.0 fundamentals and user guide. Prepared for the US. Environmental Protection Agency, Washington, DC, by the National Exposure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma, CA; and Lockheed Martin Systems Engineering Center, Arlington, VA, EP-D-05-004; STI-907045.05- 3347-UG, October, 2008.

Norris, G., Duvall, R., Brow, S., and Bai, S.: EPA positive matrix factorization (PMF) 3.0 fundamentals and user guide. Prepared for the US. Environmental Protection Agency, Washington, DC, by the National Expo-

sure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma, CA; and Lockheed Martin Systems Engineering Center, Arlington, VA, EPA/600/R-14/108, 2014.

Olivier, J. G. J., Bouwman, A. F., van der Maas, C. W. M., and Berdowski, J. J. M.: Emission Database for Global Atmospheric Research (EDGAR), in: Non-CO_2 Greenhouse Gases: Why and How to Control?, edited by: van Ham, J., Janssen, L. J. H. M., and Swart, R. J., Springer Netherlands, 93-106, 1994.

Olson, D. A., Norris, G. A., Seila, R. L., Landis, M. S., and Vette, A. F.: Chemical characterization of volatile organic compounds near the World Trade Center: Ambient concentrations and source apportionment, Atmospheric Environment, 41, 5673-5683, doi:10.1016/j.atmosenv.2007.02.047, 2007.

Paatero, P., and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111-126, doi:10.1002/env.3170050203, 1994.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometrics and Intelligent Laboratory Systems, 37, 23-35, doi:10.1016/S0169-7439(96)00044-5, 1997.

Paatero, P.: The Multilinear Engine-A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n-Way Parallel Factor Analysis Model, Journal of Computational and Graphical Statistics, 8, 854-888, doi:10.1080/10618600.1999.10474853, 1999.

Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, Chemometrics and Intelligent Laboratory Systems, 60, 253-264, doi:10.1016/S0169-7439(01)00200-3, 2002.

Paatero, P., and Hopke, P. K.: Rotational tools for factor analytic models, Journal of Chemometrics, 23, 91-100, doi:10.1002/cem.1197, 2009.

Paatero, P., Eberly, S., Brown, S. G., and Norris, G.A.: Methods for estimating uncertainty in factor analytic solutions, Atmos. Meas. Tech., 7, 781–797, doi:10.5194/amt-7-781-2014, 2014.

Panday, A. K., Prinn, R. G., and Schär, C.: Diurnal cycle of air pollution in the Kathmandu Valley, Nepal: 2. Modeling results, Journal of Geophysical Research: Atmospheres, 114, D21308, doi:10.1029/2008jd009808, 2009.

Politis, D. N., and White, h.: Automatic block-length selection for the dependent bootstrap, Econometrics Reviews, 23, 53–70, doi:10.1081/ETC-120028836, 2004.

Pradhan, B. B., Dangol, P. M., Bhaunju, P. M., and Pradhan, S.: Rapid Urban Assessment of Air Quality for Kathmandu, Nepal: Summary, ICIMOD, 2012.

Pudasainee, D., Sapkota, B., Shrestha, M. L., Kaga, A., Kondo, A., and Inoue, Y.: Ground level ozone concentrations and its association with NOx and meteorological parameters in Kathmandu Valley, Nepal, Atmospheric Environment, 40, 8081–8087, doi:10.1016/j.atmosenv.2006.07.011, 2006.

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, Atmospheric Environment, 41, 6986-6998, doi:10.1016/j.atmosenv.2007.05.008, 2007.

Roberts, J. M., Veres, P., VandenBoer, T. C., Warneke, C., Graus, M., Williams, E. J., Lefer, B. L., Brock, C. A., Bahreini, R., Öztürk, F., Middlebrook, A. M., Wagner, N. L., Dubè, W. P. A., and de Gouw, J. A.: New insights into atmospheric sources and sinks of isocyanic acid, HNCO, from recent urban and regional observations, Journal of Geophysical Research: Atmospheres, 119, 1060-1072, doi:10.1002/2013JD019931, 2014.

Sarkar, C., Kumar, V., and Sinha, V.: Massive emissions of carcinogenic benzenoids from paddy residue burning in North India, Current Science, 104, 1703 - 1709, 2013.

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, doi: 10.5194/acp-16-3979-2016, 2016.

Sekimoto, K., Inomata, S., Tanimoto, H., Fushimi, A., Fujitani, Y., Sato, K., and Yamada, H.: Characterization of nitromethane emission from automotive exhaust, Atmospheric Environment, 81, 523-531, doi:10.1016/j.atmosenv.2013.09.031, 2013.

Shim, C., Wang, Y., Singh, H. B., Blake, D. R., and Guenther, A. B.: Source characteristics of oxygenated volatile organic compounds and hydrogen cyanide, Journal of Geophysical Research: Atmospheres, 112, doi:10.1029/2006jd007543, 2007.

Shrestha, S. R., Oanh, N. T. K., Xu, Q., Rupakheti, M., and Lawrence, M.: Analysis of the vehicle fleet in the Kathmandu Valley for estimation of environment and climate co-benefits of technology intrusions, Atmospheric Environment, 81, 579-590, doi:10.1016/j.atmosenv.2013.09.050, 2013.

Sinha, V., Williams, J., Diesch, J. M., Drewnick, F., Martinez, M., Harder, H., Regelin, E., Kubistin, D., Bozem, H., Hosaynali-Beygi, Z., Fischer, H., Andrès-Hernàndez, M. D., Kartal, D., Adame, J. A., and Lelieveld, J.: Constraints on instantaneous ozone production rates and regimes during DOMINO derived using in-situ OH reactivity measurements, Atmos. Chem. Phys., 12, 7269-7283, doi: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, doi:10.5194/acp-14-5921-2014, 2014.

Slowik, J. G., Vlasenko, A., McGuire, M., Evans, G. J., and Abbatt, J. P. D.: Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site, Atmos. Chem. Phys., 10, 1969-1988, doi:10.5194/acp-10-1969-2010, 2010.

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, doi:10.5194/acp-15-845-2015, 2015.

Stockwell, C. E., Christian, T. J., Goetz, J. D., Jayarathne, T., Bhave, P. V., Praveen, P. S., Adhikari, S., Maharjan, R., DeCarlo, P. F., Stone, E. A., Saikawa, E., Blake, D. R., Simpson, I. J., Yokelson, R. J., and Panday, A. K.: 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, 11043-11081, doi:10.5194/acp-16-11043-2016, 2016.

Stone, E. A., Schauer, J. J., Pradhan, B. B., Dangol, P. M., Habib, G., Venkataraman, C., and Ramanathan, V.: Characterization of emissions from South Asian biofuels and application to source apportionment of carbonaceous aerosol in the Himalayas, Journal of Geophysical Research: Atmospheres, 115, D06301, doi:10.1029/2009jd011881, 2010.

Tsai, S. M., Zhang, J., Smith, K. R., Ma, Y., Rasmussen, R. A., and Khalil, M. A. K.: Characterization of Non-Methane Hydrocarbons emitted from various Cookstoves used in China, Environmental Science & Technology, 37, 2869-2877, doi:10.1021/es026232a, 2003.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891-2918, doi:10.5194/acp-9-2891-2009, 2009.

Vlasenko, A., Slowik, J. G., Bottenheim, J. W., Brickell, P. C., Chang, R. Y. W., Macdonald, A. M., Shantz, N. C., Sjostedt, S. J., Wiebe, H. A., Leaitch, W. R., and Abbatt, J. P. D.: Measurements of VOCs by proton transfer reaction mass spectrometry at a rural Ontario site: Sources and correlation to aerosol composition, Journal of Geophysical Research: Atmospheres, 114, D21305, doi:10.1029/2009JD012025, 2009.

Watson, J. G., Chow, J. C., and Fujita, E. M.: Review of volatile organic compound source apportionment by chemical mass balance, Atmospheric Environment, 35, 1567-1584, doi:10.1016/S1352-2310(00)00461-1, 2001.

Yokelson, R. J., Burling, I. R., Gilman, J. B., Warneke, C., Stockwell, C. E., de Gouw, J., Akagi, S. K., Urbanski, S. P., Veres, P., Roberts, J. M., Kuster, W. C., Reardon, J., Griffith, D. W. T., Johnson, T. J., Hosseini, S., Miller, J. W., Cocker Iii, D. R., Jung, H., and Weise, D. R.: Coupling field and laboratory measurements to estimate the emission factors of identified and unidentified trace gases for prescribed fires, Atmos. Chem. Phys., 13, 89-116, doi:10.5194/acp-13-89-2013, 2013.

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, Journal of Geophysical Research: Atmospheres, 117, D24302, doi:10.1029/2012jd018236, 2012.

Yuan, B., Hu, W. W., Shao, M., Wang, M., Chen, W. T., Lu, S. H., Zeng, L. M., and Hu, M.: VOC emissions, evolutions and contributions to SOA formation at a receptor site in Eastern China, Atmos. Chem. Phys., 13, 8815-8832, doi:10.5194/acp-13-8815-2013, 2013.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131-5153, doi:10.5194/acp-9-5131-2009, 2009.

Supplementary Information of

Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization

Chinmoy Sarkar et al.

Correspondence to: Vinayak Sinha (vsinha@iisermohali.ac.in)

Table S1. Most likely identity of VOCs (having average mixing ratios > 0.2 ppb) detected at specific protonated m/z ratios, molecular formula, likely mass assignment, reference of previous mass assignment, sensitivity, limit of detection (LOD), average ambient mixing ratios (±1 σ)

Protonated m/z or ion	Formula	Most Likely Identity	References of some previously reported studies	Sensitivity (ncps/ppb)	LOD (ppb)	Average (sdev) mixing ratio (ppb)
28.007	HCN	Hydrogen Cyanide	Stockwell et al., 2015; Karl et al., 2003	18.48	0.241	1.56 (0.24)
31.018	HCHO	Formaldehyde	Inomata et al., 2010; Stockwell et al., 2015	18.88	0.103	1.78 (0.50)
33.034	CH_3OH	Methanol	Seco et al., 2011; de Gouw et al., 2003	19.16	0.090	7.42 (1.28)
41.039	C_3H_4	Propyne	Akagi et al., 2011; Stockwell et al., 2015	7.167	0.080	7.67 (1.80)
42.034	CH_3CN	Acetonitrile*	Seco et al., 2011; de Gouw et al., 2003	20.91	0.043	1.08 (0.38)
43.055	C_3H_6	Propene	Stockwell et al., 2015; Park et al., 2013	7.45	0.048	3.98 (1.21)
44.014	NHCO	Isocyanic acid	Warneke et al., 2011	20.64	0.067	0.90 (0.08)
45.033	C_2H_4O	Acetaldehyde*	De Gouw et al., 2003; Seco et al., 2011	20.04	0.262	8.81 (4.58)
45.990	NO_2^+	Nitronium ion from fragmentation of C1-C5 alkyl nitrates	Aoki et al., 2007	20.91	0.094	1.08 (0.24)
46.029	CH_3NO	Methanamide		20.91	0.093	0.76 (0.16)
47.013	CH_2O_2	Formic acid	Jordan et al., 2009; Williams et al., 2001	21.04	0.041	4.96 (1.02)
47.049	C_2H_6O	Ethanol	Park et al., 2013; Seco et al., 2011	21.05	0.361	1.59 (0.85)
51.044	C_4H_2	1,3-Butadiyne$	Yokelson et al., 2013	8.56	0.013	0.67 (0.14)
56.060	C_3H_5N	Propanenitrile$	Yokelson et al., 2013	22.27	0.022	0.21 (0.05)
57.034	C_3H_4O	Acrolein* + Methylketene	Stockwell et al., 2015; Jordan et al., 2009	22.26	0.034	0.80 (0.26)
59.049	C_3H_6O	Acetone* + Propanal	de Gouw et al., 2003; Seco et al., 2011	23.47	0.074	4.21 (0.65)
60.051	C_2H_5NO	Acetamide		22.80	0.069	0.39 (0.05)
61.027	$C_2H_4O_2$	Acetic acid	de Gouw et al., 2007; Stockwell et al., 2015; Seco et al., 2011	22.94	0.440	4.24 (1.21)
62.026	CH_3NO_2	Nitromethane@	Inomata et al., 2014; Akagi et al., 2013	23.07	0.020	0.24 (0.08)
63.026	C_2H_6S	Dimethyl Sulfide	Akagi et al., 2011; Park et al., 2013	23.21	0.049	0.26 (0.03)
67.054	C_5H_6	1,3-Cyclopentadiene	Stockwell et al., 2015	10.78	0.008	0.23 (0.06)
69.033	C_4H_4O	Furan	Stockwell et al., 2015; Jordan et al., 2009	24.02	0.009	0.46 (0.17)
69.070	C_5H_8	Isoprene*	Stockwell et al., 2015; de Gouw et al., 2003; Seco et al., 2011	10.02	0.013	1.11 (0.24)
71.049	C_4H_6O	Methyl vinyl ketone; Methacrolein; Crotonaldehyde*	Seco et al., 2011; Stockwell et al., 2015; de Gouw et al., 2007	27.17	0.017	0.35 (0.10)

73.027	$C_3H_4O_2$	Methylglyoxal	Stockwell et al., 2015; Muller et al., 2012	24.56	0.021	0.31 (0.10)
73.063	C_4H_8O	Methyl ethyl ketone*	de Gouw et al., 2003; Stockwell et al., 2015; Park et al., 2013	21.91	0.036	0.69 (0.12)
75.042	$C_3H_6O_2$	Hydroxyacetone	Christian et al., 2003; Heigenmoser et al., 2013; Stockwell et al., 2015	24.83	0.066	0.63 (0.18)
79.054	C_6H_6	Benzene*	Jordan et al., 2009; de Gouw et al., 2003	13.43	0.013	2.71 (1.17)
83.085	C_6H_{10}	Assorted Hydrocarbons	Stockwell et al., 2015	13.01	0.008	0.45 (0.09)
87.042	$C_4H_6O_2$	2,3-Butanedione	Stockwell et al., 2015; Karl et al., 2007	26.45	0.028	0.35 (0.08)
93.070	C_7H_8	Toluene*	Seco et al., 2011; Jordan et al., 2009	15.78	0.006	1.53 (0.38)
97.031	$C_5H_4O_2$	2-Furaldehyde (furfural)	Ruuskanen et al., 2011; Liu et al., 2012; Li et al., 2013	27.80	0.010	0.26 (0.07)
97.102	C_7H_{12}	Assorted Hydrocarbons	Stockwell et al., 2015	14.96	0.006	0.23 (0.05)
105.070	C_8H_8	Styrene	Jordan et al., 2009; Stockwell et al., 2015	16.07	0.004	0.21 (0.08)
107.086	C_8H_{10}	Xylenes*	Jordan et al., 2009; Stockwell et al., 2015	15.36	0.004	0.97 (0.27)
121.101	C_9H_{12}	Trimethylbenzenes	Muller et al., 2012; Jordan et al., 2009	18.30	0.004	0.38 (0.10)
129.070	$C_{10}H_8$	Naphthalene	Jordan et al., 2009; Stockwell et al., 2015	19.40	0.009	0.33 (0.09)

* VOC sensitivities determined using VOC gas standards in calibration experiments

$ Observed mass accuracy for 1,3-Butadiyene and Propanenitrile were 21 mDa and 10 mDa, respectively

@ Corrected for the ^{13}C isotopologues of acetic acid

Table S2. Statistical parameters for the measured species used as PMF input

VOC Species	Category	S/N	Min	25th	Median	75th	Max	N[1]	Nbdl[2]
Hydrogen Cyanide	Weak	3.10	0.29	0.62	0.90	1.79	12.43	969	0
Formaldehyde	Strong	3.93	0.60	1.43	1.80	2.34	5.39	969	0
Methanol	Strong	4.00	3.31	6.69	8.51	10.32	21.79	969	0
Propyne	Strong	4.00	3.37	8.27	10.95	13.73	37.58	969	0
Acetonitrile	Strong	3.95	0.50	1.00	1.47	2.04	5.60	969	0
Propene	Strong	4.00	1.37	3.99	6.03	7.78	18.34	969	0
Isocyanic acid	Strong	3.91	0.94	1.22	1.36	1.61	2.39	969	0
Acetaldehyde	Strong	3.97	3.46	8.83	11.52	15.46	61.96	969	0
Nitronium ion	Strong	3.86	0.81	1.37	1.65	1.97	5.42	969	0
Formamide	Strong	3.71	0.55	0.93	1.20	1.51	2.67	969	0
Formic acid	Strong	4.00	2.42	6.85	7.92	9.61	17.74	969	0
Ethanol	Strong	2.78	bdl	1.16	2.21	3.49	18.63	969	51
1,3-Butadiyne	Strong	3.99	0.13	0.89	1.20	1.51	3.43	969	0
Propanenitrile	Strong	3.77	0.09	0.32	0.40	0.50	0.98	969	0
Acrolein + Methylketene	Strong	3.95	0.40	1.14	1.51	2.06	4.09	969	0
Acetone + Propanal	Strong	3.99	3.55	6.99	8.71	10.38	23.48	969	0
Acetamide	Strong	3.53	0.51	0.69	0.82	0.97	1.53	969	0
Acetic acid	Strong	3.88	4.69	9.31	11.99	15.56	35.04	969	0
Nitromethane	Strong	3.83	0.16	0.37	0.50	0.69	1.85	969	0
Dimethyl Sulfide	Strong	3.46	0.19	0.50	0.56	0.63	2.35	969	0
1,3-Cyclopentadiene	Strong	3.97	0.14	0.36	0.50	0.70	4.42	969	0
Furan	Strong	3.98	0.19	0.66	1.14	1.52	3.38	969	0
Isoprene	Strong	4.00	0.28	1.76	2.39	3.43	11.25	969	0
Methyl vinyl Ketone + Methacrolein	Strong	3.94	0.18	0.61	0.81	1.11	3.08	969	0
Methylglyoxal	Strong	3.89	0.18	0.56	0.76	1.01	2.27	969	0
Methyl ethyl ketone	Strong	3.93	0.37	1.30	1.73	2.18	4.93	969	0
Hydroxyacetone	Strong	3.74	0.27	1.17	1.54	2.15	4.47	969	0
Benzene	Strong	4.00	0.95	3.98	6.79	10.11	37.35	969	0
Assorted Hydrocarbons	Strong	3.99	0.10	0.78	1.10	1.66	6.88	969	0
2,3-Butanedione	Strong	3.86	0.16	0.80	1.04	1.36	3.35	969	0
Toluene	Strong	4.00	0.51	3.09	4.51	6.54	30.71	969	0
2-Furaldehyde (furfural)	Strong	3.96	0.18	0.61	0.83	1.15	2.23	969	0
Assorted Hydrocarbons	Strong	3.97	0.08	0.48	0.70	1.03	3.03	969	0
Styrene	Strong	3.98	0.09	0.41	0.67	1.06	3.32	969	0
Xylenes	Strong	4.00	0.31	2.12	3.19	4.79	24.73	969	0
Trimethylbenzenes	Strong	3.99	0.17	0.94	1.39	2.14	10.44	969	0
Naphthalene	Strong	3.97	0.28	0.92	1.44	2.00	6.94	969	0

1. Number of valid hourly VOC samples; 2. Number of samples below detection limit

Table S3: Percentage contribution of PMF derived factors obtained from constrained runs with 5-, 6-, 7-, 8- and 9-Factors

PMF solutions	RB + WD	BK	BK2	MCS	TR	MI	UI	BG	MD	SE
5-Factor	19.2	23.7	-	23.9	-	-	-	13.4	19.8	-
6-Factor	15.6	18.4	-	-	20.1	21.5	-	11.7	12.7	-
7-Factor	12.3	14.7	-	-	17	19.5	-	8.2	13.8	14.7
8-Factor	10.9	10.4	-	-	16.8	14	17.9	10	9.2	10.8
9-Factor	10.5	12.2	7.7	-	15.6	9.5	15.5	7	9.8	12.2

RB + WD = Residential biofuel use and waste disposal; BK = Biomass co-fired brick kilns; BK2 = Second brick kiln factor; MCS = Mixed combustion sources; TR = Traffic; MI = Mixed industrial; UI = Unresolved industrial; BG = Biogenic; MD = Mixed daytime; SE = Solvent evaporation

Table S4. Correlation coefficient (r) between the time series of PMF resolved factors and other independent meteorological parameters (solar radiation, ambient temperature, change in solar radiation, change in ambient temperature, wind speed, wind direction, relative humidity and absolute humidity) for a) daytime period (06:00 – 17:00 LT) and b) nighttime period (17:00 – 06:00 LT)

a)	SE	MD	BG	RB+WD	BK	MI	UI	TR	SR	AT	ΔSR	ΔAT	WS	RH	AH
SE	1.00	-													
MD	-0.22	1.00	-												
BG	-0.01	-0.01	1.00	-											
RB+WD	-0.01	-0.35	0.10	1.00	-										
BK	0.23	-0.55	-0.31	0.18	1.00	-									
MI	0.07	-0.68	-0.28	**0.51**	**0.54**	1.00	-								
UI	**0.55**	-0.49	-0.17	-0.24	0.37	0.06	1.00	-							
TR	0.05	-0.33	0.27	0.21	-0.19	0.17	-0.02	1.00	-						
SR	0.23	**0.58**	**0.57**	-0.35	-0.51	-0.65	-0.16	-0.10	1.00	-					
AT	-0.11	**0.74**	0.31	-0.41	-0.74	-0.79	-0.23	-0.07	**0.71**	1.00	-				
ΔSR	**0.62**	-0.47	0.15	0.14	0.38	0.29	0.46	-0.08	0.11	-0.46	1.00	-			
ΔAT	**0.64**	-0.06	0.33	-0.06	0.03	-0.07	0.28	-0.04	**0.50**	0.02	**0.67**	1.00	-		
WS	-0.35	**0.57**	-0.05	-0.23	-0.50	-0.48	-0.39	0.00	0.13	**0.63**	-0.71	-0.45	1.00	-	
RH	0.10	-0.82	-0.31	0.42	**0.65**	**0.74**	0.37	0.10	-0.75	-0.91	0.45	-0.05	-0.62	1.00	-
AH	0.21	-0.27	0.08	0.02	-0.17	-0.11	0.49	0.11	-0.02	0.13	0.20	0.16	-0.11	0.25	1.00

b)	SE	MD	BG	RB+WD	BK	MI	UI	TR	SR	AT	ΔSR	ΔAT	WS	RH	AH
SE	1.00	-													
MD	0.14	1.00	-												
BG	-0.25	-0.18	1.00	-											
RB+WD	0.34	-0.10	**0.58**	1.00	-										
BK	-0.20	-0.38	-0.07	-0.33	1.00	-									
MI	-0.25	-0.60	0.08	0.04	0.32	1.00	-								
UI	0.29	-0.29	-0.58	-0.47	0.38	-0.03	1.00	-							
TR	0.17	-0.36	0.36	0.46	-0.24	0.14	-0.24	1.00	-						
SR	0.10	0.11	0.05	-0.04	-0.17	-0.28	-0.01	0.32	1.00	-					
AT	0.47	0.36	0.15	0.48	-0.59	-0.58	-0.18	0.37	0.44	1.00	-				
ΔSR	-0.17	-0.22	-0.03	-0.02	0.25	0.38	0.04	-0.34	-0.83	-0.59	1.00	-			
ΔAT	-0.12	0.00	-0.06	-0.21	0.09	0.08	0.08	-0.11	-0.09	-0.14	0.21	1.00	-		
WS	0.30	0.43	-0.02	0.07	-0.35	-0.53	-0.09	0.15	**0.55**	**0.70**	-0.78	-0.15	1.00	-	
RH	-0.59	-0.49	-0.07	-0.44	0.40	0.48	0.25	-0.26	-0.44	-0.80	**0.59**	0.16	-0.66	1.00	-
AH	0.04	-0.06	0.18	0.26	-0.43	-0.33	0.03	0.21	0.05	0.55	-0.08	-0.05	0.21	0.04	1.00

SE = Solvent evaporation; MD = Mixed daytime; BG = Biogenic; RB+WD = Residential biofuel use and waste disposal; BK = Biomass co-fired brick kilns; MI = Mixed industrial emissions; UI = Unresolved industrial emissions; TR = Traffic; SR = Solar radiation; AT = Ambient temperature; ΔSR = Rate of change in solar radiation (dSR/dt); ΔAT = Rate of change in ambient temperature (dT/dt); WS = Wind speed; RH = Relative humidity; AH = Absolute humidity

Table S5. Correlation coefficient (r) between the time series of the input parameters for the PMF model

	m28.007	*m31.018*	*m33.034*	*m41.039*	*m42.034*	*m43.055*	*m44.014*	*m45.033*	*m45.990*	*m46.029*	*m47.013*	*m47.049*	*m51.044*	*m56.060*	*m57.034*	*m59.049*	*m60.051*	*m61.027*	*m62.026*	*m63.026*	*m67.054*	*m69.033*	*m69.070*	*m71.049*	*m73.027*	*m73.063*	*m75.042*	*m79.054*	*m83.085*	*m87.042*	*m93.070*	*m97.031*	*m97.102*	*m105.070*	*m107.086*	*m121.101*	*m129.070*
m28.007	1.00																																				
m31.018	0.04	1.00																																			
m33.034	0.16	0.35	1.00																																		
m41.039	0.25	0.24	0.79	1.00																																	
m42.034	0.09	0.06	0.79	0.65	1.00																																
m43.055	0.17	0.18	0.84	0.93	0.78	1.00																															
m44.014	-0.02	0.55	0.23	-0.01	0.27	0.03	1.00																														
m45.033	0.04	0.72	0.52	0.37	0.44	0.39	0.51	1.00																													
m45.990	-0.19	0.65	0.05	-0.24	-0.13	-0.18	0.51	0.37	1.00																												
m46.029	-0.22	0.60	0.09	-0.22	0.12	-0.14	0.85	0.59	0.69	1.00																											
m47.013	-0.09	0.77	0.03	-0.14	-0.18	-0.16	0.54	0.52	0.70	0.65	1.00																										
m47.049	0.15	0.15	0.65	0.66	0.70	0.73	0.05	0.43	-0.22	-0.04	-0.20	1.00																									
m51.044	0.09	0.36	0.90	0.73	0.64	0.80	0.12	0.48	0.10	0.05	0.14	0.54	1.00																								
m56.060	0.15	0.36	0.84	0.88	0.71	0.91	0.18	0.47	-0.02	0.02	0.03	0.63	0.86	1.00																							
m57.034	0.15	0.55	0.81	0.80	0.64	0.81	0.25	0.64	0.06	0.11	0.18	0.59	0.82	0.87	1.00																						
m59.049	0.17	0.58	0.68	0.70	0.45	0.66	0.22	0.68	0.24	0.13	0.38	0.42	0.74	0.75	0.79	1.00																					
m60.051	-0.09	0.64	0.24	0.03	0.13	0.05	0.82	0.48	0.68	0.80	0.68	-0.05	0.26	0.28	0.29	0.46	1.00																				
m61.027	0.01	0.86	0.50	0.32	0.26	0.31	0.58	0.82	0.53	0.57	0.75	0.19	0.53	0.46	0.69	0.69	0.63	1.00																			
m62.026	0.08	0.30	0.80	0.65	0.91	0.76	0.38	0.66	0.02	0.25	0.09	0.70	0.68	0.72	0.76	0.61	0.25	0.54	1.00																		
m63.026	0.09	0.52	0.61	0.53	0.54	0.60	0.33	0.65	0.20	0.30	0.39	0.52	0.65	0.66	0.64	0.69	0.38	0.58	0.70	1.00																	
m67.054	0.21	0.27	0.53	0.75	0.39	0.65	0.00	0.34	-0.14	-0.18	0.07	0.43	0.51	0.63	0.65	0.67	0.11	0.35	0.46	0.43	1.00																
m69.033	0.22	0.13	0.77	0.90	0.73	0.90	-0.04	0.29	-0.34	-0.25	-0.25	0.68	0.70	0.85	0.81	0.60	-0.02	0.24	0.69	0.50	0.66	1.00															
m69.070	0.19	0.53	0.37	0.51	0.03	0.37	0.03	0.46	0.15	-0.05	0.46	0.15	0.46	0.46	0.55	0.79	0.27	0.56	0.23	0.49	0.70	0.34	1.00														
m71.049	0.17	0.56	0.70	0.78	0.46	0.73	0.16	0.61	0.03	0.01	0.24	0.47	0.75	0.81	0.95	0.86	0.28	0.68	0.61	0.62	0.73	0.75	0.73	1.00													
m73.027	-0.02	0.84	0.35	0.33	0.00	0.27	0.36	0.55	0.48	0.36	0.70	0.06	0.50	0.48	0.65	0.67	0.57	0.83	0.24	0.49	0.40	0.26	0.66	0.73	1.00												
m73.063	0.21	0.66	0.64	0.70	0.33	0.63	0.17	0.62	0.26	0.08	0.41	0.38	0.71	0.72	0.79	0.95	0.42	0.70	0.50	0.64	0.65	0.59	0.83	0.87	0.75	1.00											
m75.042	0.17	0.74	0.60	0.57	0.26	0.52	0.22	0.60	0.30	0.16	0.50	0.28	0.69	0.66	0.84	0.79	0.41	0.82	0.45	0.56	0.55	0.54	0.72	0.87	0.86	0.87	1.00										
m79.054	0.20	0.01	0.73	0.70	0.89	0.78	0.10	0.36	-0.18	-0.05	-0.22	0.64	0.59	0.69	0.62	0.51	0.05	0.19	0.82	0.50	0.51	0.76	0.17	0.50	-0.01	0.41	0.28	1.00									
m83.085	0.16	0.46	0.32	0.48	-0.01	0.35	-0.03	0.39	0.10	-0.12	0.41	0.13	0.44	0.44	0.51	0.78	0.23	0.50	0.18	0.47	0.67	0.33	0.98	0.71	0.64	0.80	0.66	0.12	1.00								
m87.042	0.13	0.66	0.52	0.60	0.17	0.52	0.14	0.45	0.22	0.04	0.49	0.24	0.65	0.65	0.78	0.78	0.38	0.71	0.36	0.56	0.64	0.56	0.79	0.87	0.87	0.86	0.93	0.22	0.78	1.00							
m93.070	0.17	0.28	0.62	0.77	0.37	0.74	-0.08	0.23	-0.01	-0.23	0.07	0.37	0.69	0.73	0.67	0.76	0.16	0.35	0.45	0.54	0.74	0.68	0.67	0.75	0.50	0.78	0.62	0.52	0.69	0.72	1.00						
m97.031	0.17	0.70	0.61	0.65	0.37	0.58	0.27	0.63	0.13	0.15	0.39	0.39	0.61	0.68	0.86	0.77	0.39	0.75	0.54	0.57	0.63	0.66	0.69	0.89	0.76	0.83	0.91	0.39	0.64	0.89	0.60	1.00					
m97.102	0.23	0.52	0.38	0.55	0.03	0.42	-0.02	0.43	0.09	-0.12	0.42	0.18	0.49	0.50	0.58	0.81	0.24	0.54	0.23	0.50	0.71	0.41	0.98	0.76	0.68	0.86	0.75	0.17	0.97	0.84	0.74	0.73	1.00				
m105.070	0.23	0.18	0.64	0.86	0.52	0.82	-0.10	0.22	-0.27	-0.31	-0.15	0.49	0.64	0.79	0.76	0.66	0.02	0.26	0.52	0.43	0.75	0.87	0.51	0.79	0.39	0.67	0.59	0.66	0.51	0.66	0.85	0.67	0.59	1.00			
m107.086	0.15	0.27	0.54	0.71	0.25	0.67	-0.12	0.15	-0.01	-0.27	0.07	0.28	0.62	0.66	0.61	0.69	0.13	0.32	0.32	0.40	0.68	0.61	0.64	0.70	0.51	0.74	0.61	0.40	0.67	0.71	0.97	0.57	0.73	0.81	1.00		
m121.101	0.16	0.22	0.51	0.73	0.24	0.67	-0.16	0.12	-0.10	-0.33	0.03	0.30	0.60	0.65	0.59	0.68	0.08	0.27	0.31	0.40	0.69	0.64	0.64	0.69	0.48	0.72	0.58	0.39	0.68	0.71	0.96	0.57	0.73	0.84	0.99	1.00	
m129.070	0.29	0.28	0.63	0.73	0.61	0.71	0.02	0.48	-0.12	-0.10	0.06	0.54	0.60	0.69	0.68	0.75	0.14	0.39	0.68	0.63	0.71	0.72	0.59	0.71	0.32	0.71	0.56	0.81	0.56	0.57	0.71	0.65	0.62	0.76	0.59	0.61	1.00

Figure S1. Collection of garbage burning grab samples in the Kathmandu Valley (on left) and the instrumental setup for the analysis (on right)

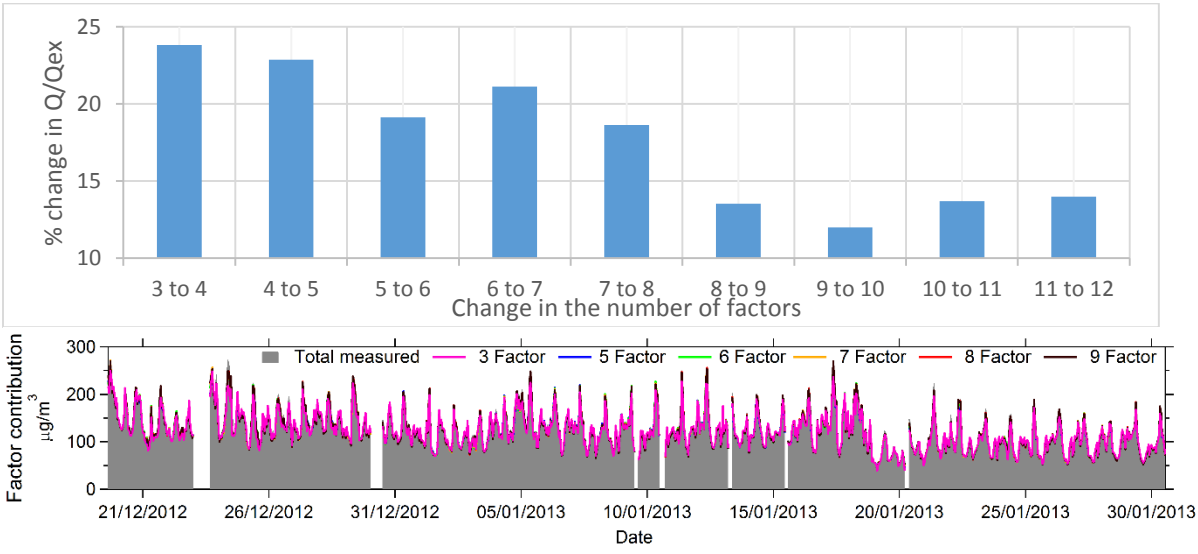


Figure S2. Relative change in the Q/Qexpected ratio with change in factor number (top) and time series of the total measured VOC mass (grey filled) and the modelled VOC mass for different number of factors in the PMF solution (bottom).

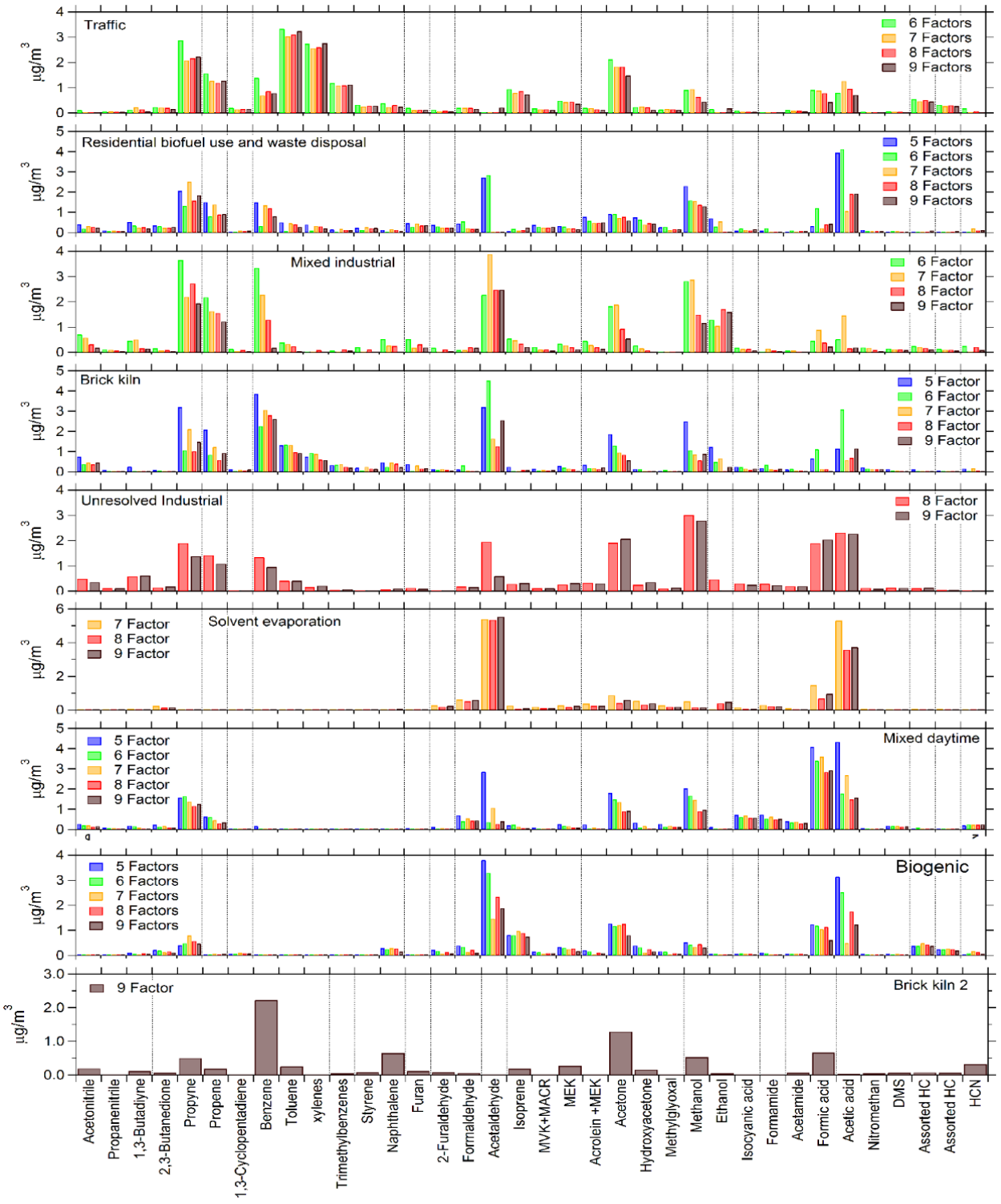


Figure S3a. Evolution of the factor profiles of the eight sources identified, and the 9th source which is considered to arise due to splitting of the brick kiln factor, from the 5 Factor to the 9 Factor solution.

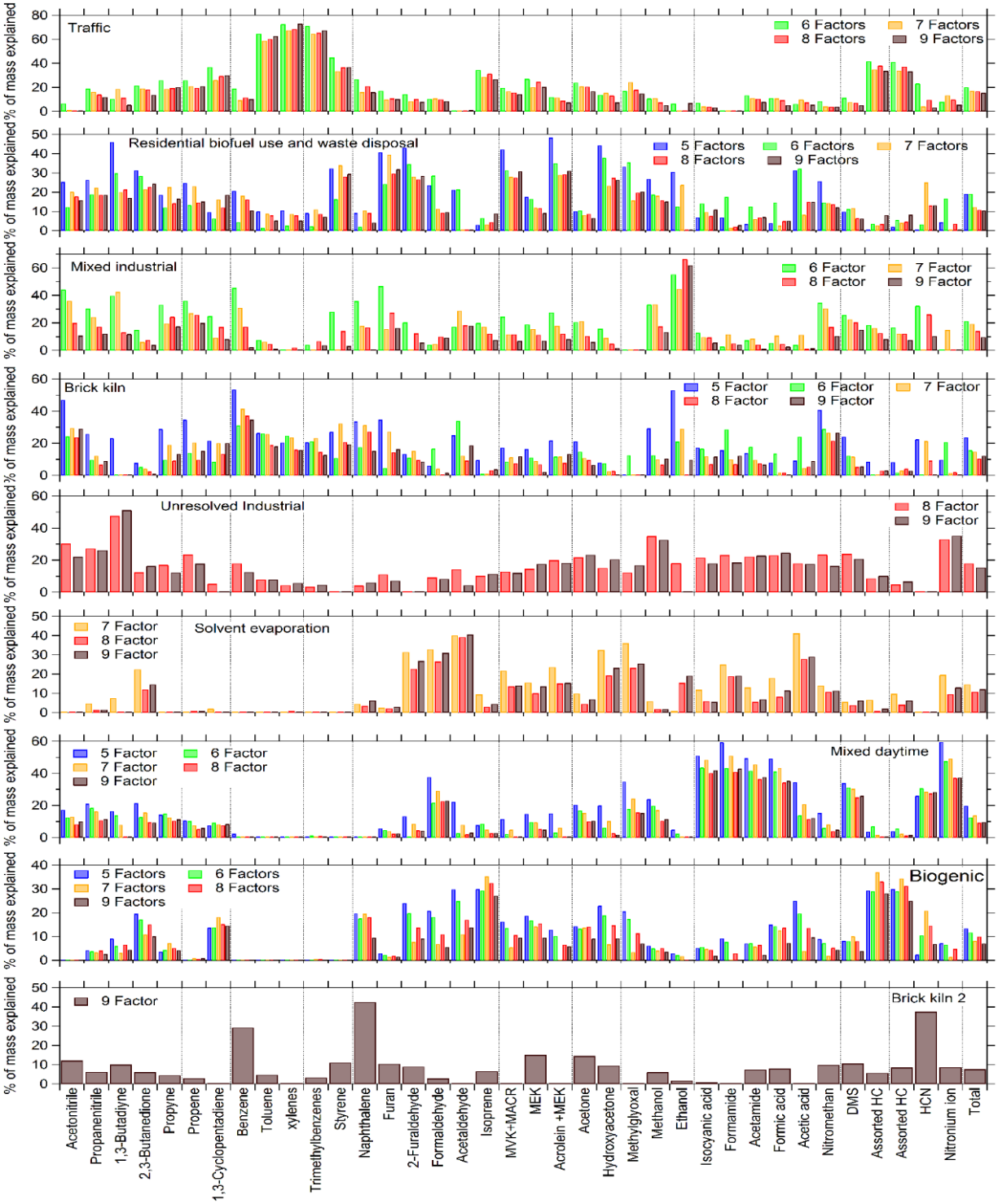


Figure S3b. Evolution of the percentage of the mass of each compound explained by the eight sources identified, and the 9th source which is considered to arise due to splitting of the brick kiln factor, from the 5 Factor to the 9 Factor solution.

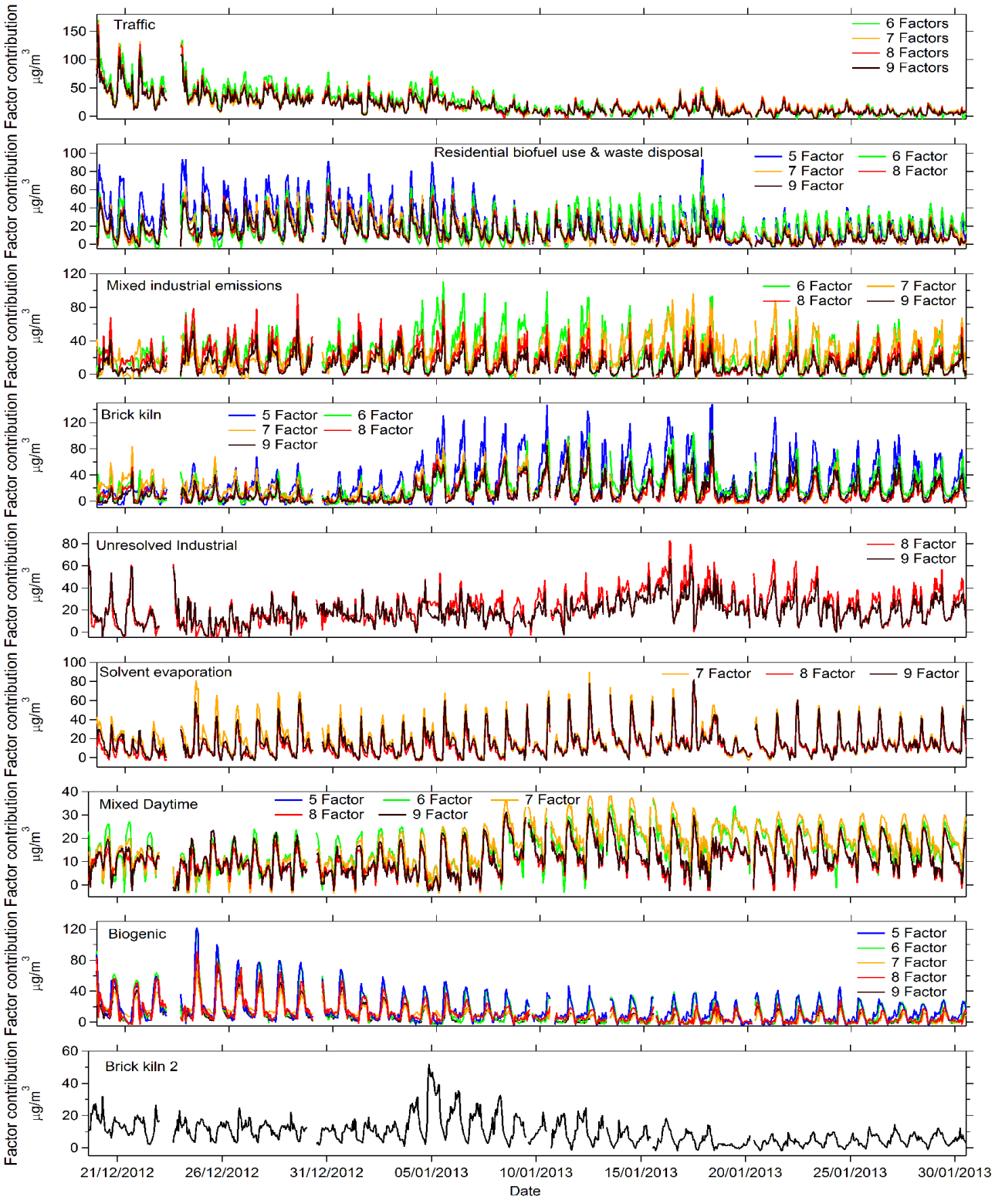


Figure S3c Evolution of the factor contribution of the eight sources identified, and the 9th source which is considered to arise due to splitting of the brick kiln factor, from the 5 Factor to the 9 Factor solution.

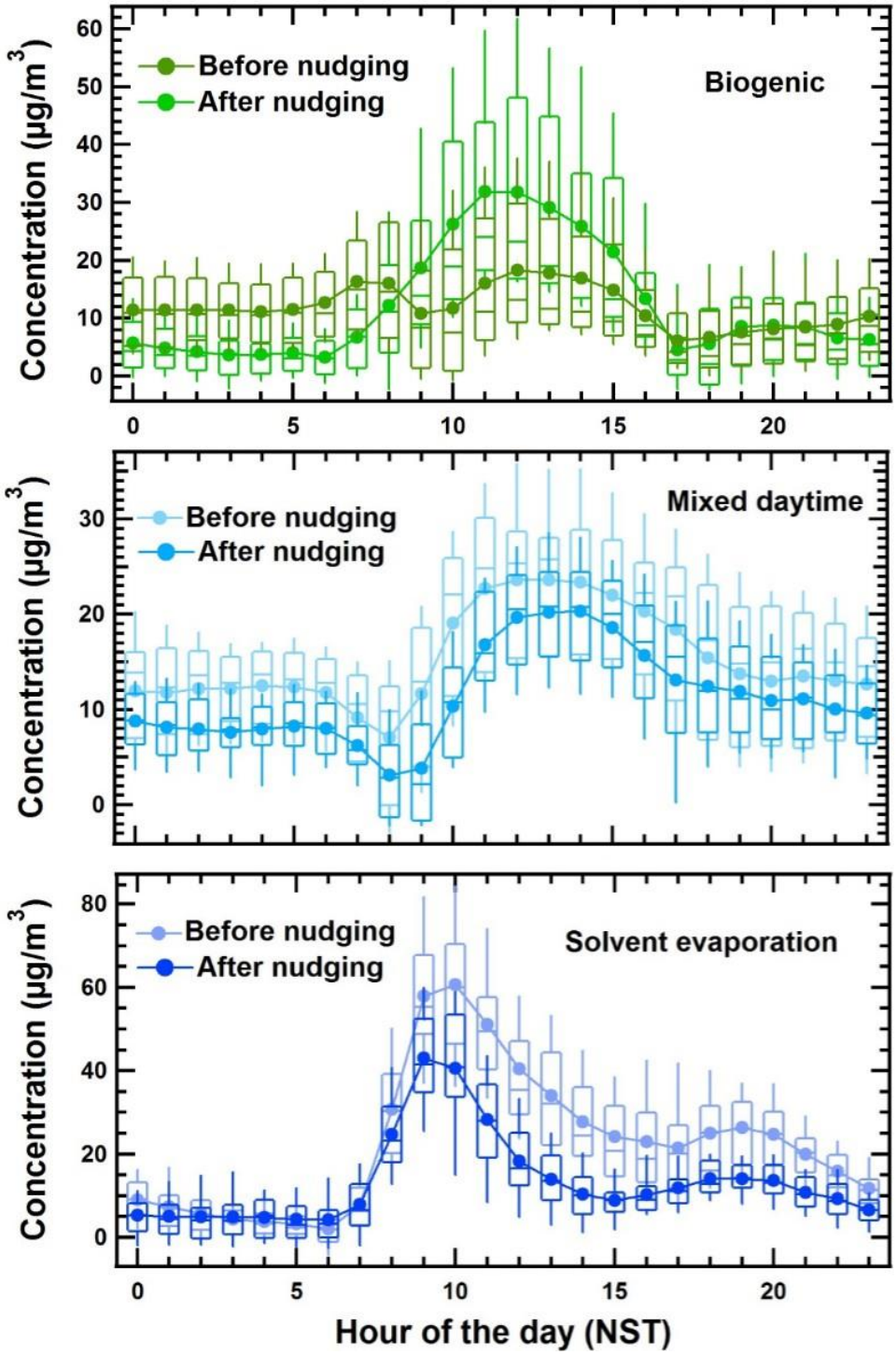


Figure S4. Comparison of the diel profiles of biogenic emissions, mixed daytime and solvent evaporation factors before and after nudging

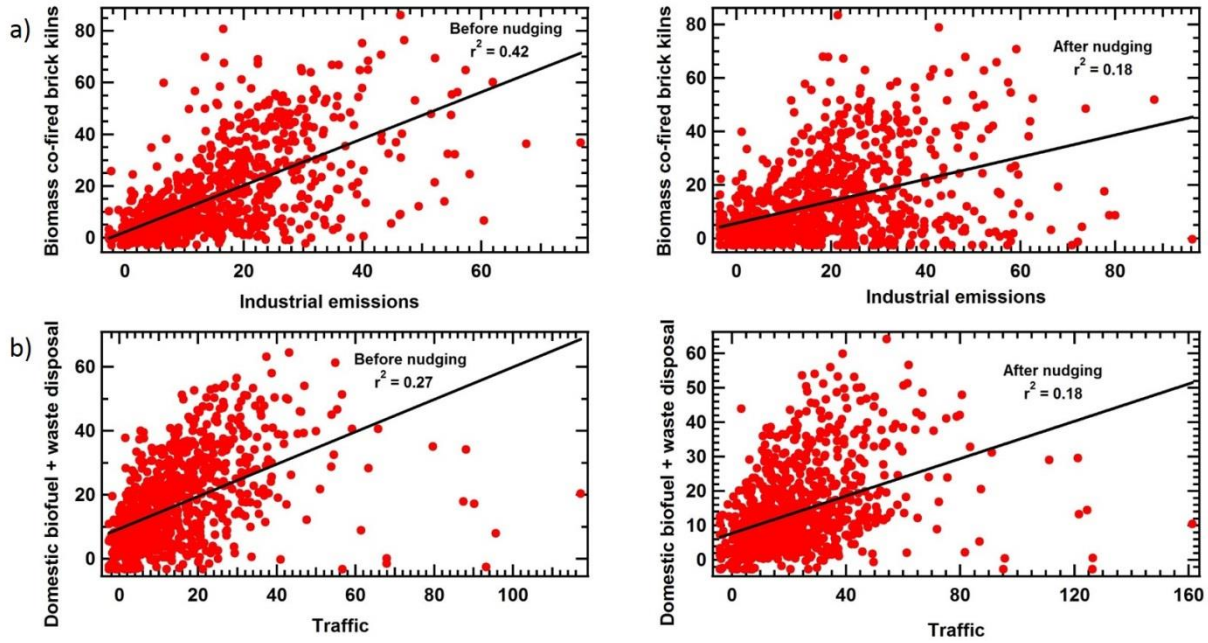


Figure S5. Comparison of the G-space plots between a) biomass co-fired brick kilns and mixed industrial emissions and b) residential biofuel use and waste disposal and traffic before and after nudging

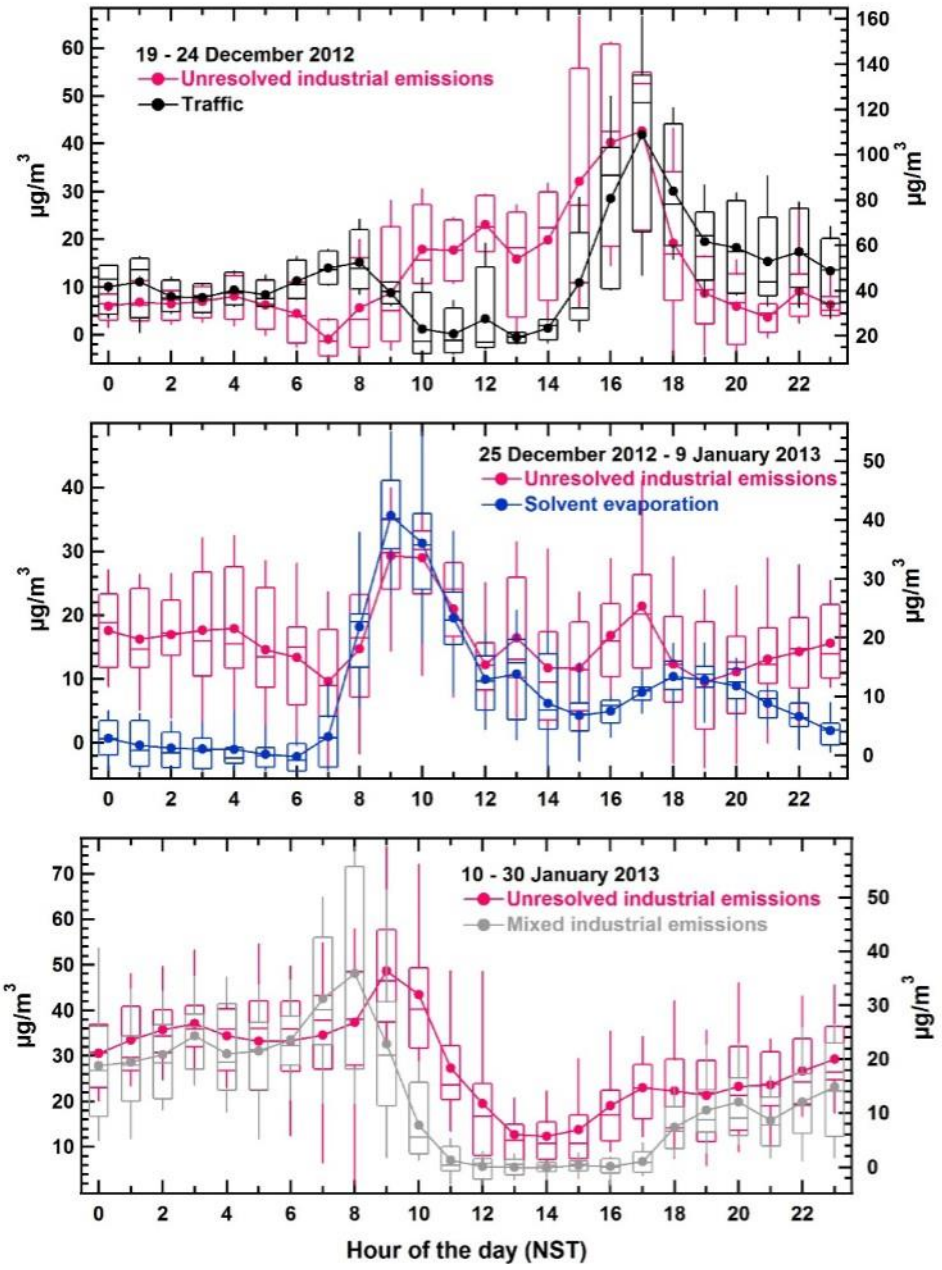


Figure S6. Comparison of the diel profile of the unresolved industrial emissions with that of traffic (19 – 24 December 2012), solvent evaporation (25 December 2012 – 9 January 2013) and mixed industrial emissions (10 – 30 January 2013)

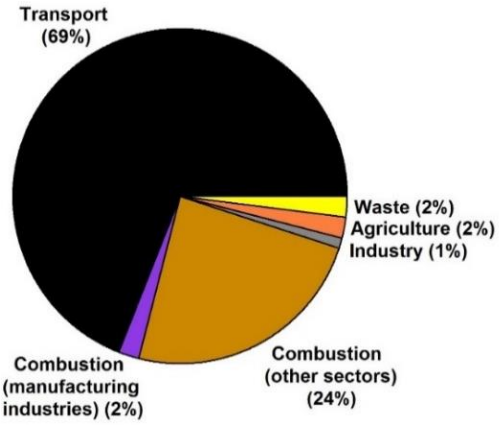


Figure S7. Emissions of particulate matter (sum of PM10 and PM2.5) and CO from different sectors in Kathmandu and Lalitpur (Pradhan et al, 2012)

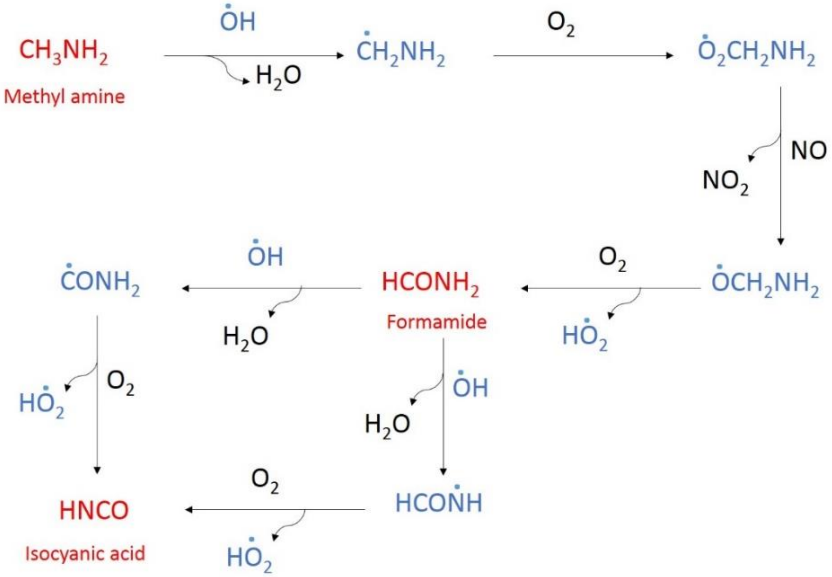


Figure S8. Reaction schematic for the formation of formamide and isocyanic acid (blue colored species represents radicals)

Figure S8 represents the reaction schematic of the proposed mechanism for the formation of formamide, acetamide and isocyanic acid based on the previous laboratory experiments which shows that photooxidation of alkyl amines leads to the formation of formamide and acetamide which undergoes further photooxidation to form isocyanic acid which can have severe health impact at concentration thresholds above 1 ppb (Roberts et al., 2011;Roberts et al., 2014). This study provides the first ever ambient evidence of the photochemical source of isocyanic acid by quantification of both the amides (the precursor) and isocyanic acid (the product) collectively and the source apportionment of these compounds in the Kathmandu Valley.

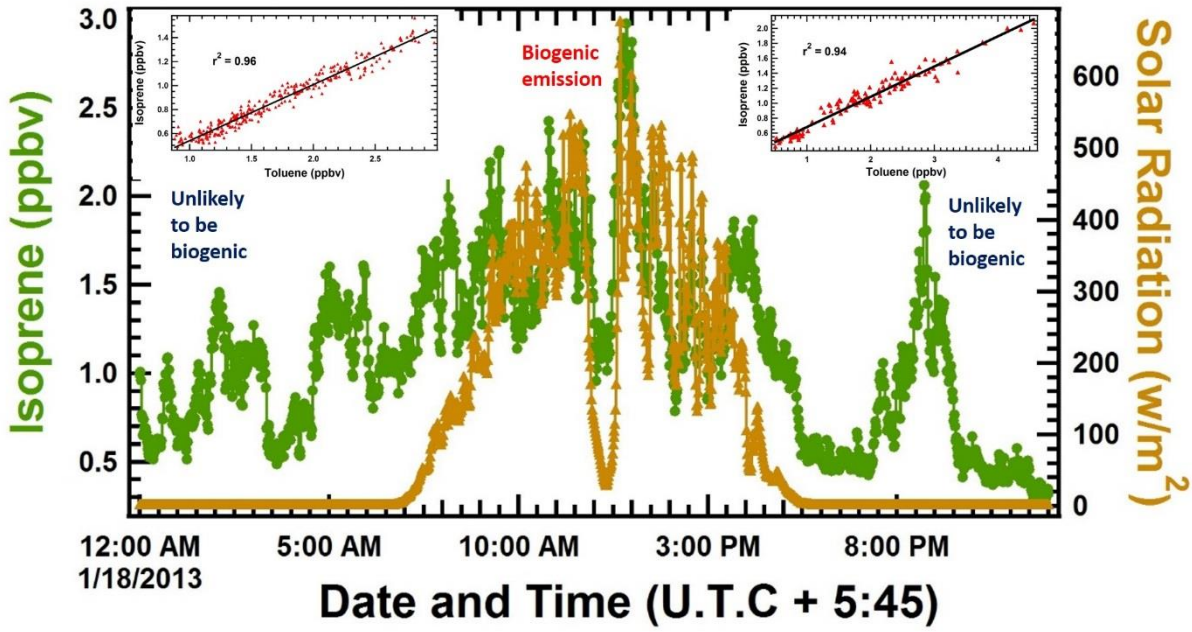


Figure S9. Shows a representative day's (18 January 2013) isoprene data against solar radiation. It can be observed from the figure that the daytime isoprene emission correlates very nicely with solar radiation which indicates biogenic emission while during evening hours and night time, isoprene showed high peaks that shows excellent correlation with toluene.

References

Akagi, S. K., Yokelson, R. J., Burling, I. R., Meinardi, S., Simpson, I., Blake, D. R., McMeeking, G. R., Sullivan, A., Lee, T., Kreidenweis, S., Urbanski, S., Reardon, J., Griffith, D. W. T., Johnson, T. J., and Weise, D. R.: Measurements of reactive trace gases and variable O3 formation rates in some South Carolina biomass burning plumes, Atmos. Chem. Phys., 13, 1141-1165, 10.5194/acp-13-1141-2013, 2013.

Aoki, N., Inomata, S., Tanimoto, H.: Detection of C1–C5 alkyl nitrates by proton transfer reaction time-of-flight mass spectrometry, International Journal of Mass Spectrometry, https://doi.org/10.1016/j.ijms.2006.11.018, 2007

de Gouw, J. A., Goldan, P. D., Warneke, C., Kuster, W. C., Roberts, J. M., Marchewka, M., Bertman, S. B., Pszenny, A. A. P., and Keene, W. C.: Validation of proton transfer reaction-mass spectrometry (PTR-MS) measurements of gas-phase organic compounds in

the atmosphere during the New England Air Quality Study (NEAQS) in 2002, Journal of Geophysical Research-Atmosphere, 108, 10.1029/2003jd003863, 2003.

Inomata, S., Tanimoto, H., Kato, S., Suthawaree, J., Kanaya, Y., Pochanart, P., Liu, Y., and Wang, Z.: PTR-MS measurements of non-methane volatile organic compounds during an intensive field campaign at the summit of Mount Tai, China, in June 2006, Atmos. Chem. Phys., 10, 7085-7099, doi:10.5194/acp-10-7085-2010, 2010.

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 proton-transfer-reaction mass spectrometry to characterize volatile organic compound sources at the La Porte super site during the Texas Air Quality Study 2000, Journal of Geophysical Research: Atmospheres, 108, 4508, 10.1029/2002jd003333, 2003.

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 Gouw, J.: Isocyanic acid in the atmosphere and its possible link to smoke-related health effects, Proceedings of the National Academy of Sciences, 10.1073/pnas.1103352108, 2011.

Roberts, J. M., Veres, P., VandenBoer, T. C., Warneke, C., Graus, M., Williams, E. J., Lefer, B. L., Brock, C. A., Bahreini, R., Öztürk, F., Middlebrook, A. M., Wagner, N. L., Dubé, W. P. a., and de Gouw, J. A.: New insights into atmospheric sources and sinks of isocyanic acid, HNCO, from recent urban and regional observations, Journal of Geophysical Research: Atmospheres, 119, 1060–1072, 10.1002/2013JD019931, 2014.

Seco, R., Peñuelas, J., Filella, I., Llusià, J., Molowny-Horas, R., Schallhart, S., Metzger, A., Müller, M., and Hansel, A.: Contrasting winter and summer VOC mixing ratios at a forest site in the Western Mediterranean Basin: the effect of local biogenic emissions, Atmos. Chem. Phys., 11, 13161-13179, 10.5194/acp-11-13161-2011, 2011.

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, 10.5194/acp-15-845-2015, 2015.

Warneke, C., Roberts, J. M., Veres, P., Gilman, J., Kuster, W. C., Burling, I., Yokelson, R., and de Gouw, J. A.: VOC identification and inter-comparison from laboratory biomass

burning using PTR-MS and PIT-MS, International Journal of Mass Spectrometry, 303, 6-14, http://dx.doi.org/10.1016/j.ijms.2010.12.002, 2011.