Source characterization of volatile organic compounds measured by PTR-ToF-MS in Delhi, India

Liwei Wang[1], Jay G. Slowik[1], Nidhi Tripathi[2], Deepika Bhattu[1], Pragati Rai[1], Varun Kumar[1], Pawan Vats[4], Rangu Satish[5,a], Urs Baltensperger[1], Dilip Ganguly[4], Neeraj Rastogi[5], Lokesh K. Sahu[2], Sachchida N. Tripathi[3,*], and André S. H. Prévôt[1,*]

[1] Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, 5232, Switzerland

[2] Space and Atmospheric Sciences Division, Physical Research Laboratory, Ahmedabad 380009, India

[3] Department of Civil Engineering and Centre for Environmental Science and Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

[4] Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi 110016, India

[5] Geosciences Division, Physical Research Laboratory, Ahmedabad 380009, India

[a] Now at: Department of Environmental Science and Analytical Chemistry, Stockholm University, Stockholm 11419, Sweden

[*] Corresponding authors

Abstract. Characteristics and sources of volatile organic compounds (VOCs) were investigated with highly time-resolved simultaneous measurements by two proton-transfer-reaction time-of-flight mass spectrometers (PTR-ToF-MS) at an urban and a suburban site in New Delhi, India from January to March 2018. During the measurement period, high mixing ratios of VOCs and trace gases were observed, with high nocturnal mixing ratios and strong day-night variations. The positive matrix factorization (PMF) receptor model was applied separately to the two sites, and six major factors of VOCs were identified at both sites, i.e., two factors related to traffic emissions, two to solid fuel combustion, and two secondary factors. At the urban site, traffic-related emissions comprising mostly mono-aromatic compounds were the dominant sources, contributing 56.6% of the total mixing ratio, compared to 36.0% at the suburban site. Emissions from various solid fuel combustion processes, particularly in the night, were identified as a significant source of aromatics, phenols and furans at both sites. The secondary factors accounted for 15.9% of the total VOC concentration at the urban site and for 33.6% at the suburban site. They were dominated by oxygenated VOCs and exhibited substantially higher contributions during daytime.

Key words: VOCs, Source apportionment, traffic emission, solid fuel combustion, secondary formation, Delhi

1 Introduction

Volatile organic compounds (VOCs) are important trace gas constituents in the troposphere, impacting local and regional air quality, human health and climate both directly and indirectly (IPCC, 2013). With the participation of NO_x, oxidation of VOCs leads to the formation of tropospheric O_3, causing regional photochemical smog (Atkinson, 2000; de Gouw et al., 2005). The chemical transformation of VOCs forms less-volatile compounds and can contribute to gas-to-particle partitioning either by new particle formation or condensation on existing particles (Hallquist et al., 2009; Ehn et al., 2014). In addition, many VOCs are toxic, such as aromatic compounds, and exposure to large amounts of VOCs may adversely affect human health, including acute and chronic effects on different systems and organs, and even cancer (Kampa and Castanas, 2008; Nurmatov et al., 2013).

VOCs originate from both natural and anthropogenic sources. Biogenic VOCs (BVOCs), mainly emitted by plants, are regarded as the largest source of VOCs globally (Atkinson and Arey, 2003; Hallquist et al., 2009). However, in urban areas, anthropogenic VOCs can be dominant (Borbon et al., 2013). Vehicular exhaust emissions have long been regarded as the dominant source of VOCs in many urban areas. Many of these compounds are reactive and thus contribute significantly to urban O_3 pollution, photochemical smog, and secondary organic aerosol (SOA) formation (Fraser et al., 1998; Derwent et al., 2010; Müller et al., 2012). Biomass burning is considered as the second largest source of VOCs worldwide (Crutzen and Andreae, 1990) and could be a major VOC source in some urban areas during biomass burning events (Karl et al., 2007; Yokelson et al., 2009; Baudic et al., 2016; Languille et al., 2019). Compared to vehicular emissions, biomass burning sources emit more oxygenated and high molecular-weight VOCs such as furans and phenols. Due to high atmospheric reactivity and higher SOA yield, these compounds can also contribute substantially to SOA formation (Sekimoto et al., 2018; Koss et al., 2018).

Air pollution in South Asia has attracted more and more attention in recent years. This region is regarded as one of the most polluted regions in the world (e.g., Monks et al., 2009). Due to rapid urbanization and the lack of widespread advanced pollution control technologies in the industrial, energy, and transportation sectors (Mahata et al., 2018), air pollution has become an increasingly important issue across the region, particularly in large urban areas. The increasing emissions of air pollutants not only impact the local/regional air quality and human health, but also influences distant and pristine areas through transport (Bonasoni et al., 2010; Lawrence and Lelieveld, 2010; Mahata et al., 2018). Delhi, the capital city of India, with a population of 18.98 million (2012) people, is facing an air quality problem ranked as the worst among 1600 major cities in the world (WHO, 2014). The critical air quality problems have left India with high death rates from chronic obstructive pulmonary disease (COPD) and respiratory disease e.g., asthma (WHO, 2014). Despite the severe air pollution in Delhi, information on pollution levels, as well as an emission inventory of VOCs, and their sources in Delhi is still lacking (Monks et al., 2009; Kumar et al., 2015). A few studies on BTEX (i.e. benzene, toluene, ethylbenzene, and xylene) in Delhi reported that vehicular emissions could be the dominant source in Delhi (Srivastava et al., 2005; Hoque et al., 2008). Several studies also identified motor vehicle emissions as a significant source of particulate matter (PM), with high contributions from solid fuel combustion and industrial emissions (Sahu et al., 2011; Sharma et al., 2014; Kumar et al., 2015). These studies reveal the importance of local anthropogenic sources. However, since previous studies mainly focused on certain families of VOCs or a few VOCs species, a

73   comprehensive investigation of the VOC pollution levels, specific emission sources, as well as their roles in the

74   local tropospheric chemistry has not been reported previously. VOC source apportionment studies have been

75   conducted only in a few Indian cities. For example, Srivastava studied 23 VOC compounds in Mumbai city, and

76   pointed out the lack of evaporative and oceanic emissions (Srivastava, 2004). Sahu and Saxena (2015) studied

77   15 VOC species measured by a PTR-ToF-MS in urban Ahmedabad in winter. A recent study on 32 VOCs at

78   Mohali (a suburban site in northwest Indo-Gangetic Plain) identified biofuel usage, biomass burning and

79   vehicular emissions as important primary sources, and pointed out the lacking of present emission inventories

80   (Pallavi et al., 2019).

81   In this study, we report simultaneous on-line measurements of VOCs using two PTR-ToF-MS instruments at an

82   urban and a suburban site of Delhi, India. The level, composition and source characteristics of different VOCs

83   (158 ions at IITD and 90 ions at MRIU) were analyzed with the aid of the positive matrix factorization (PMF)

84   model. Spatial and temporal comparisons between the two sites are discussed for both the sources and selected

85   ions.

86   **2 Experiments and methods**

87   **2.1 Measurement sites**

88   Figure 1 shows the studied region and locations of the two measurement sites. The measurements were

89   conducted from 18 January to 10 March 2018 at an urban site at the Indian Institute of Technology (IITD), New

90   Delhi, and from 16 January to 8 March 2018 at a suburban site at Manav Rachna International University

91   (MRIU), Faridabad (Fig. 1).

92   At the urban site (28° 33'N, 77° 12'E), the inlet system was installed on the rooftop (~20 m above the ground) of

93   block VI (a four-story building), Centre for Atmospheric Sciences, on the campus of IITD (Gani et al., 2019;

94   Rai et al., 2020). The IITD campus is located in the southern part of the city center, and is surrounded by

95   educational, commercial, and residential districts. The study site is approximately 80 m north of a busy street

96   and is surrounded by several streets inside the campus as well. Thus, in addition to vehicular emissions,

97   commercial and residential activities may also produce VOC emissions in the immediate vicinity. The sampling

98   line of the PTR-ToF-MS was approximately 1.5 m long, and the inlet consisted of polyether ether ketone

99   (PEEK) tubing, with an inner diameter of 0.075 mm.

100   The other measurement site, at MRIU, was located in a relatively open area in suburban Delhi, about 20 km

101   southeast of the IITD site. Besides, the northeast territory of MRIU is of slightly higher elevation compared to

102   the sampling site, as shown in Fig. 1. The suburban site is located inside a big campus, and it is surrounded by

103   several small parks, with only a few narrow roads nearby. Although the site is not far from the main road, the

104   traffic load and other anthropogenic activities nearby may be much less than that at IITD. The instruments were

105   located on the first floor of a teaching building on the campus of MRIU. A similar PEEK inlet as at IITD was

106   used, with the sampling line approximately 2 m long.

107   **2.2 Online instruments**

108   Data collected in this study included the mixing ratios of VOCs, NO_x, CO, and meteorological parameters. The

109   involved monitors, analyzers, and sensors are described in detail in the following. Two PTR-ToF-MS 8000

110   (Ionicon Analytical G.m.b.H, Innsbruck, Austria) were simultaneously deployed at the two sites. At both sites,

111   the PTR-ToF-MS instruments were operated in the H_3O^+ mode, where the sampled VOCs are protonated via

112   non-dissociative proton transfer from H_3O^+ ions (Eq.1):

113    $H_3O^+ + R \rightarrow RH^+ + H_2O$ (1)

114   The PTR-ToF-MS measures non-methane organic gases (NMOG) with a proton affinity higher than water, i.e.,

115   most of the common VOCs such as carbonyls, acids, and aromatic hydrocarbons, as well as alkanes with more

116   than eight carbons and alkenes with more than two carbons. A detailed description of the instrument is found in

117   Jordan et al. (2009) and Graus et al. (2010). For both sites, the time resolution was set to 30 s, the drift tube

118   voltage was set at 600 V, with a drift temperature of 60 ℃, and a pressure of 2.2-2.3 mbar, resulting in a

119   reduced electric field (*E/N*) value of about 130 Td. Therefore, a similar fragmentation pattern is expected.

120   Calibrations were performed twice at the IITD site and three times at the MRIU site by dynamic dilution of

121   VOCs using a certified 15-compound gas standard (Ionimed Analytik GmbH, Austria at ∼ 1 ppmv; with stated

122   accuracy better than 8 %). The calibration components were methanol, acetonitrile, acetaldehyde, ethanol,

123   acrolein, acetone, isoprene, crotonaldehyde, 2-butanone, benzene, toluene, o-xylene, chlorobenzene, α-pinene,

124   and 1, 2- dichlorobenzene. The background measurements were performed using a dry zero air cylinder every

125   two weeks. The raw data were processed using the Tofware post processing software (version 2.5.11,

126   TOFWERK AG, Thun, Switzerland) with the PTR module as distributed by Ionicon Analytik GmbH

127   (Innsbruck, Austria), running in the Igor Pro 6.37 environment (Wavemetrics Inc., Lake Oswego, OR, U.S.A.).

128   Volume mixing ratios (in ppbv) were calculated based on the method described by de Gouw and Warneke

129   (2007) and the literature reaction rates (*k*) of the ion with the H_3O^+ ion were applied when available (Cappellin

130   et al., 2012). For ions where the reaction rate had not been measured, a rate constant of 2×10^{-9} cm^3 s^{-1} was

131   assumed.

132   The mixing ratio of NO_x was measured by chemiluminesence using the Serinus 40 Oxides of Nitrogen analyzer

133   (Ecotech) at IITD and a Model 42i (TEC, USA) at MRIU. The CO mixing ratio was analyzed with infrared

134   radiation absorption method using a CO Analyzer (Serinus 30, Ecotech) at IITD and a Model 48i (TEC, USA) at

135   MRIU.

136   **2.3 Source apportionment**

137   Source apportionment of VOCs was performed using the positive matrix factorization (PMF) receptor model

138   (Paatero and Tapper, 1994), which represents the measured PTR-ToF-MS mass spectral time series as a linear

139   combination of static factor profiles (characteristic of particular sources and/or atmospheric processes) and their

140   time-dependent concentrations. This can be represented in matrix notation as follows:

141    $\mathbf{X} = \mathbf{GF} + \mathbf{E}$ (2)

142   Where $\mathbf{X}$, $\mathbf{G}$, $\mathbf{F}$, and $\mathbf{E}$ are matrices corresponding to the measured mass spectral time series (time x *m/z*), factor

143   time series (time x factor), factor profiles (factor x *m/z*), and model residuals (time x *m/z*), respectively. In this

144   study, we used a total of 158 ions at IITD and 90 ions at MRIU measured by the PTR-ToF-MS (list of ions

145 utilized are shown in Table S1 and Table S2). Equation (2) is solved by minimizing the objective function, Q

146 (Eq. 3), using a weighted least-squares algorithm:

147 $$Q = \sum_i^n \sum_j^m \left(e_{ij}/s_{ij}\right)^2 \qquad (3)$$

148 Here, e_{ij} is the residuals (elements of **E**) and s_{ij} represents the corresponding measurement uncertainty.

149 As recommended by Paatero (2003), ions with a signal-to-noise ratio (SNR) lower than 0.2 were excluded from

150 the input matrix, and ions with a signal-to-noise ratio (SNR) between 0.2 and 2 were down-weighted by

151 increasing their uncertainties by a factor of 2. In addition, in this study, a few ions such as methanol,

152 acetaldehyde, acetone, and acetic acid with mixing ratios 3-4 times higher than the other ions were excluded

153 from the input. Due to extremely high SNR compared to other compounds, inclusion of these ions in PMF leads

154 to solution where only these ions are well-explained, and useful source information is not retrieved.

155 PMF was implemented as the Multilinear Engine (ME-2) (Paatero, 1999), with the Source Finder (SoFi) toolkit

156 for Igor Pro (Wavemetrics, Inc., Portland; Canonaco et al., 2013) used for model configuration and post-

157 analysis. As different combinations of **G** and **F** can yield mathematically similar solutions (i.e., similar Q), ME-

158 2 enables intelligent rotational control to achieve reasonable solutions by involving constraints and external

159 data. In this study, constraints were applied combining a scalar a (usually between 0 and 1) and a reference

160 profile (Canonaco et al., 2013). The a-value determines the extent to which the resolved factors f' and g' are

161 allowed to vary from the input reference elements of f and g according to Eq. 4:

162 $$f' = f \pm a \cdot f \quad \text{and} \quad g' = g \pm a \cdot g \qquad (4)$$

163 **3 Results and discussion**

164 **3.1 Temporal and spatial variation**

165 Figure 2 presents time series of the mixing ratios of CO and NO_x as well as of the sum of all VOCs analyzed in

166 the PMF model and temperature at the two sites. The fractional composition of VOCs in terms of the chemical

167 formula-derived families is included as well. The mean mixing ratios of CO were 1.29 ppmv, and 0.95 ppmv at

168 IITD and MRIU, respectively. The average mixing ratio of NO_x at IITD was 148.0 ppbv, 5 times higher than at

169 MRIU (29.6 ppbv). Similarly, the summed VOC mixing ratio (i.e., all the ions included in PMF analysis, see

170 Table S1 and S2) was 27.6 ppbv at IITD, significantly higher than at MRIU (19.4 ppbv). The ambient

171 temperature showed large variations during the campaign with an overall increasing trend after 21 February.

172 Thus, the whole study period can be divided into two separate periods of relatively cold and warm temperatures.

173 Fig S1 shows the box-whisker plots of temperature during the two periods at the two sites. As shown in Fig. S1,

174 the average temperature was 17 ºC during the cold period and 23 ºC during the warm period at IITD. At MRIU,

175 the average temperature was very similar when VOCs were measured, with the cold days averaging at 16 ºC and

176 the warm periods at 23 ºC. During the warm period, the mean temperature reached the minimum at 6:00 LT, 1

177 hour earlier than in the cold period. These differences in temperature along with other meteorological

178 parameters are expected to have an impact on the emission profiles (e.g. Sekimoto et al., 2018) as well as the

179 dilution and chemical transformation processes.

180 To investigate temporal and site-dependent changes in the relative composition of the VOCs, the measured

181 signals of different ions were classified into seven families based on their identified chemical formula, namely

182 aromatic C_xH_y, non-aromatic C_xH_y, $C_xH_yO_1$, $C_xH_yO_2$, $C_xH_yO_3$, C_xH_yN, and $C_xH_yNO_z$ compounds. Aromatic

183 C_xH_y were classified due to their low H:C ratio, and oxidized aromatics are classified into the three $C_xH_yO_z$

184 families. Aromatic C_xH_y were the largest fraction at both sites constituting about 45.4% at IITD (averaged 13.9

185 ppbv) and 34.3% at MRIU (averaged 7.0 ppbv). The high fractions of aromatic compounds indicate strong

186 influences from anthropogenic emissions, mainly vehicle exhaust, at both sites. At IITD, the contributions of

187 $C_xH_yO_1$ (21.9%, 5.6 ppbv on average) and $C_xH_yO_2$ (13.6%, 3.6 ppbv) were also significant, followed by non-

188 aromatic C_xH_y (11.3%, 3.5 ppbv), C_xH_yN (5.2%, 1.2 ppbv), $C_xH_yO_3$ (2.3%, 0.6 ppbv), and C_xH_yNO (0.4%, 0.5

189 ppbv). At MRIU, the VOC family composition was similar, with the exception of a higher fraction and

190 concentration of non-aromatic C_xH_y (23.2%, 4.2 ppbv), which was dominated by high C_5H_8 and C_6H_{10} in the

191 daytime compared to that at IITD. Besides, the averaged mixing ratios of the majority of the families were lower

192 at MRIU than at IITD, except for non-aromatic C_xH_y and $C_xH_yO_2$. The averaged mixing ratios of $C_xH_yO_1$,

193 $C_xH_yO_2$, C_xH_yN, $C_xH_yO_3$, and C_xH_yNO were 3.6 ppbv, 3.4 ppbv, 0.5 ppbv, 0.2 ppbv, and 0.09 ppbv,

194 respectively.

195 Owing to the variation of planetary boundary layer height (PBLH), as well as emission sources, temperature,

196 and solar radiation, the time series of the mixing ratios of VOCs, NO_x and CO exhibited substantial differences

197 between day and night periods. Fig. S2 shows the diurnal patterns of NO_x, CO, and different VOC families at

198 both sites. The nocturnal mixing ratios of most of the VOC families, as well as those of CO and NO_x were

199 higher than during daytime, with much greater diurnal variation at IITD than at MRIU. The spatial difference in

200 the diurnal variation may be due to a lower influence of local emissions at the suburban MRIU site because of

201 lower population density and fuel consumption. In addition, the relative proportions of the VOC families varied

202 over time, indicating different emission patterns and oxidation chemistry. For instance, substantial contributions

203 and concentrations of aromatic compounds and $C_xH_yNO_1$ were observed at night at IITD, indicating strong

204 anthropogenic emissions in the urban area, such as traffic-related emissions and solid fuel combustion.

205 Meanwhile, higher daytime contributions were found for the $C_xH_yO_1$ and $C_xH_yO_2$ families; and in particular the

206 $C_xH_yO_3$ family peaked around midday at IITD, indicating tropospheric aging and secondary formation during

207 daytime. Moreover, the nocturnal concentrations of all the gas phase species were higher at IITD compared to

208 MRIU. During daytime, however, the concentrations of non-aromatic C_xH_y, $C_xH_yO_2$, as well as CO at MRIU

209 were higher compared to those at IITD.

210 **3.2 VOC source apportionment**

211 **3.2.1 Solution selection**

212 We evaluated unconstrained solutions from 4 to 11 factors individually for both sites, and solutions were

213 selected based on interpretation of the spectra, reference information from emission inventory, correlation of

214 VOC factors with ancillary measurements, and mathematical diagnostics describing PMF performance. A six-

215 factor solution was selected as the best representation of the data measured at both sites, which includes two

216 traffic factors, two solid fuel combustion factors, and two secondary factors. An industrial factor is not presented

217 in the results, as marker compounds typically originating from industrial emissions are either not measured

218   (such as halogen compounds) nor included in this PMF analysis (such as methanol and acetone as described in

219   Section 2.3). In terms of PMF performance, the Q/Q_{exp} did not show any step change with increasing factor

220   number (Fig. S3). Solutions with less than six factors failed to provide reasonably separated sources, while

221   solutions with more factors led to additional traffic or solid fuel combustion factors that did not improve the

222   interpretability of the results. These additional factors were likely the result of minor variations across discrete

223   instances of a given source (e.g. vehicle-to-vehicle variations in VOC composition) and slight differences in

224   reactive history.

225   At IITD, the concentrations of several aromatics were very high, being 10-20 times higher than those of the

226   major phenols, and over 100 times higher than for the compounds with the lowest concentration. As a result,

227   initial PMF results were over-weighted towards explaining aromatic variability. Thus, aromatics were

228   apportioned to all the resolved factors, including factors that otherwise appear secondary and where such high

229   attribution is not reasonable. To this mathematical artifact and guide the model towards environmentally

230   reasonable results, constraints were applied to the IITD secondary factors as follows. First, PMF was performed

231   with a new input matrix excluding the most abundant aromatic ions, namely $C_6H_6H^+$, $C_7H_8H^+$, $C_8H_{10}H^+$, and

232   $C_9H_{12}H^+$ (PMF results with the new input are shown in Fig. S4). This yielded a reasonable five-factor solution,

233   from which the spectra of the two secondary factors were selected for use as reference profiles in the original

234   dataset. PMF was then executed again on the full dataset (i.e., all 158 ions including the aromatics), with

235   constraints applied to the two secondary factors using the a-value approach. Since we obtained the reference

236   from the same dataset, the a-value was set to a smaller range than the usual case. Here, solutions with the a-

237   value ranging from 0.1 to 0.3 were evaluated and no significant differences in terms of the spectra and temporal

238   variations were observed among all the results. In the following analysis, the solution with an a-value=0.1 for

239   both factors was selected as the IITD result, and the result with an a-value=0.3 is presented in Figure S5.

240   At MRIU, the lower aromatic concentrations meant that that even the unconstrained PMF did not apportion

241   significant aromatic mass to resolved secondary factors. For consistency, we tested also the constraint-based

242   method described above, but the two methods did not yield significant differences at MRIU (see Figure S6).

243   Therefore, at MRIU the unconstrained solution was selected for further interpretation.

3.2.2 Factor identification

245   Figures 3 and 4 show the factor profiles (a), time series (b), and diurnal patterns (c) of the selected PMF solution

246   at IITD and MRIU, respectively. In the following, we present a detailed discussion of the factor characteristics

247   at IITD. The MRIU factors are qualitatively similar and therefore a detailed discussion is not repeated, but rather

248   a site comparison is presented in Section 3.3.

249   The first two factors are related to vehicle emissions and denoted Traffic1 and Traffic2. Both are enhanced

250   during the night and relatively low during daytime. The mass spectra of both Traffic1 and Traffic2 are

251   dominated by aromatic C_xH_y compounds, namely C_6H_6 (*m/z* 79.054), C_7H_8 (*m/z* 93.070), C_8H_{10} (*m/z* 107.086),

252   C_9H_{12} (*m/z* 121.101), and $C_{10}H_{14}$ (*m/z* 135.117). These ions are tentatively attributed to benzene, toluene, C8-

253   aromatics, C9-aromatics, and C10-aromatics, which are well-studied markers for vehicular emissions (Yao et

254   al., 2015; Cao et al., 2016). While Traffic1 is mostly composed of pure aromatics, Traffic2 includes relatively

255 high contributions from non-aromatic hydrocarbons and some oxygenated compounds. As shown in Fig. 3 (b),

256 both traffic factors exhibit temporal variations similar to NO_x, which originates mainly from vehicle emissions

257 in an urban area. The correlation of Traffic 2 with NO_x ($R^2=0.76$) is better than either Traffic1 with NO_x

258 ($R^2=0.55$) or their sum (Traffic1+Traffic2) with NO_x ($R^2=0.67$). To alleviate urban traffic congestion, heavy-

259 duty vehicles are banned in Delhi during the rush hours and daytime (7:00-21:00 LT). As a result, most heavy-

260 duty traffic in Delhi occur overnight rather than during daytime. This difference in traffic patterns corresponds

261 to the temporal differences in Traffic1 and Traffic2. As shown in Fig. S7, the ratio of Traffic 2 to Traffic 1 is

262 very low during the day time and starts to increase slightly from 16:00 LT. Although both Traffic1 and Traffic2

263 are high during the night with their maximum concentrations around 21:00 LT, the Traffic2/Traffic1 ratio is as

264 low as 0.6 at that time. The ratio increases overnight with a sharp increase during the early morning, reaching a

265 maximum value of 1.4 at 07:00 LT, suggesting that Traffic1 is the dominant traffic source both day and night.

266 This can be explained as cold start emissions from gasoline vehicles emit high amounts of VOCs while heavy-

267 duty vehicles contribute much less to VOCs but more to BC and NO_x (Platt et al., 2017). Further, the spectrum

268 of Traffic2 is characterized by high fractions of high mass aromatic compounds. For instance, the ratio of C8/C7

269 aromatics is tripled in Traffic2 compared to that in Traffic1, which is similar to the ratio of the emission factors

270 from previous studies (3.5 times (Gentner et al., 2013)). Since the vehicle number in Delhi has increased

271 dramatically in the past two decades (the registered vehicular population has tripled since 1994 and has reached

272 7.6 million), the related pollution is considered as an increasingly significant source of atmospheric pollution.

273 Overall, the sum of the two traffic factors is found to be the dominant VOC source in Delhi, with contributions

274 of 33.8% from Traffic1 and 22.8% from Traffic2 to the total analyzed VOCs.

275 Two other factors are found to be related to solid fuel combustion (SFC) and are named SFC1 and SFC2. SFC1

276 is characterized by high loadings of aromatics, as well as oxygenated ions, such as $C_6H_6(C_nH_{2n})O_1$,

277 $C_6H_6(C_nH_{2n})O_2$, $C_4H_4(C_nH_{2n})O_1$, $C_4H_4(C_nH_{2n})O_2$, and $C_5H_4O_{1-2}$. These ions are tentatively attributed to phenolic

278 compounds and furans (Stockwell et al., 2015; Bruns et al., 2017). The relative fractions of emission factors

279 (EF, in mg per kg fuel) of $C_5H_4O_2$ (100 mg kg^{-1}), C_6H_6O (110 mg kg^{-1}), and $C_6H_6O_2$ (60 100 mg kg^{-1}) are

280 similar to their relative concentrations in SFC1 (i.e. $C_5H_4O_2$ (m/z 97.028, 0.019), C_6H_6O (m/z 95.049, 0.021),

281 and $C_6H_6O_2$ (m/z 111.044, 0.012)), consistent with the identification of the factor. Besides, previous studies

282 have reported that furans and phenols are emitted in high quantities by wood/biomass burning and coal

283 combustion (Bruns et al., 2017; Klein et al., 2018) and thus are regarded as important markers for these

284 combustion processes. In urban Delhi, the use of solid fuels is not limited to wood and coal, but includes open

285 combustion of many types of biomass and even waste. In addition, traditional stoves remain popular in some

286 residential households and restaurants, and their VOC emission profiles are still not clear. SFC1 is identified as

287 being more related to primary emissions from solid fuel combustion, which might include many types of

288 biomass burning, coal combustion, and even trash combustion. The spectrum of the SFC2 factor is dominated

289 by benzene and to a lesser extent toluene, while the more reactive phenols and furans are only present in small

290 amounts, suggesting increased age. In addition, it includes a high mass fraction of N-containing compounds,

291 such as C_3H_3N, C_4H_5N, and C_7H_5N. Although the chemical identification of these ions is not confirmed, these

292 ions were reported in biomass burning as well (Stockwell et al., 2015; Sekimoto et al., 2018) and may be

293 attributed to nitriles, which have longer OH lifetimes compared to furans and phenols (Meylan and Howard,

294 1993; Atkinson et al., 1989). SFC2 also explains a higher fraction of higher molecular-weight molecules, e.g.

295 $C_6H_5NO_3$ and $C_7H_7NO_3$. These ions are attributed to nitrophenols, which are semivolatile compounds found in

296 biomass burning-related secondary organic aerosol (Mohr et al., 2013). As shown in Fig 3(c), SFC1 exhibits a

297 diurnal pattern with a maximum value at 20:00 LT and a smaller peak at 08:00 LT, consistent with residential

298 heating and/or cooking activities. SFC2 is also high at night, although it peaks several hours later than SFC1 (at

299 5:00 LT), and decreases strongly during daytime without any morning peak. This suggests that the nocturnal

300 peak of SFC2 may be due to dark oxidation of emitted precursor VOCs, e.g. by nitrate radicals. Therefore,

301 considering emissions, chemical transformation and reactivity, SFC1 represents more primary emissions from

302 various types of solid fuel combustion in central Delhi, while SFC2 is associated with more aged emissions

303 from solid fuel combustion.

304 Finally, two secondary VOC factors are identified, the spectra of which are distinguished from the other factors

305 by a strong presence of oxygenated compounds. Secondary VOC1 (denoted SecVOC1 in the following)

306 contributes a large fraction of $C_2H_4O_3$ and $C_4H_2O_3$, and the spectrum is dominated by C_4H_8O, $C_3H_4O_2$, $C_3H_6O_2$,

307 and $C_3H_6O_3$, although the relative contributions to these ions are comparably lower than to $C_2H_4O_3$ and $C_4H_2O_3$.

308 These ions are likely oxidation products from various photochemical processes and are not unique to specific

309 precursors and oxidation processes. The time series of SecVOC1 follows that of the solar radiation, which has a

310 regular contribution cycle during daytime. The diurnal of SecVOC1 shows a rapid enhancement starting from

311 around 7:00-8:00 LT and declines continuously after 13:00 LT. This indicates that while many of these ions can

312 be formed rapidly during daytime, they may have a short lifetime owing to partitioning to the condensed phase

313 and/or heterogeneous processes. Secondary VOC2 (SecVOC2) has significant contributions from $C_3H_4O_2$ and

314 $C_3H_6O_2$ as well, but the relative contributions to $C_3H_4O_2$, $C_4H_6O_{2-3}$, $C_6H_6O_2$, $C_8H_8O_3$, and $C_6H_5NO_3$ are

315 substantially higher than for SecVOC1. It also includes many oxygenated compounds with higher molecular

316 weight, e.g. $C_5H_{10}O_{1-2}$, $C_6H_{10}O_{1-2}$, and $C_6H_{12}O_{1-2}$. SecVOC2 (Fig. 3(c)) is not only high during daytime, with a

317 slow rise in the morning and a peak around 10:00-12:00 LT, but is also high at night, reaching its nocturnal

318 maximum from 20:00 LT to midnight. Major fractions of alkyl nitrates ($RONO_2$) are detected as ($ROH·H^+$)

319 fragment ions by the PTR-ToF-MS (Aoki et al., 2007). Therefore, species such as $C_5H_{10}O_{1-2}$, $C_6H_{10}O_{1-2}$, and

320 $C_6H_{12}O_{1-2}$, may be partially attributed to fragments from organic nitrates, which contribute to the high nocturnal

321 SecVOC2 mixing ratio. Therefore, it is likely that SecVOC1 is a mix of first generation products and later-

322 generation oxidation products, while SecVOC2 is possibly associated with secondary generation products and

323 nighttime chemistry.

324 **3.3 Comparison of VOC sources**

325 Figure 5 compares the diurnal cycles and mass spectra of the two Traffic and the two SFC factors between the

326 two sites. While the factor profiles are similar for the primary factors (i.e. Traffic1, Traffic 2 and SFC1 at both

327 sites share the same major ions) their diurnal patterns are different. Higher nocturnal concentrations and day-

328 night variations are observed at IITD than at MRIU for all three primary factors. For example, for Traffic1 the

329 nocturnal maximum is 5.3 times higher than the daytime minimum at IITD and only 2.7 times higher at MRIU.

330 For SFC1, the corresponding numbers are 36.8 at IITD and 6.9 at MRIU. This is possibly due to increased

331 traffic density and more intense combustion activities in the populated urban areas. Besides, Fig. S8 presents

332 concentration weighted trajectory (CWT) plots of the six factors at both sites. Details of the back-trajectory and

333 CWT analysis are shown in the Supplement. As shown in Fig. S8, high concentrations of primary factors are

334   found both north and northwest of IITD, whereas high primary factors originate from southeast of MRIU,

335   different from the direction of IITD.

336   Unlike the primary factors, the mass spectra of SFC2 differ between the two sites. At IITD, SFC2 is not only

337   rich in benzene, but also in toluene. However, as shown in Fig. 5, toluene is much lower at MRIU, while the

338   relative benzene concentration remains high. In addition, SFC2 at MRIU is richer in some oxygenated

339   compounds, such as $C_6H_{12}O_{1-2}$, $C_5H_{10}O_2$, $C_6H_{10}O_2$, $C_4H_{4/6/8}O_2$. These ions are the major ions found in the

340   SecVOC2 factor at IITD. Because toluene has a much shorter photochemical lifetime than benzene, this

341   suggests that SFC2 comprises less freshly emitted but more aged compounds at MRIU than at IITD. Although

342   the SFC2 factors exhibit similar diurnal trends at the two sites, the time of daily maximum is different. At

343   MRIU, SFC2 reaches the maximum at 07:00 LT, which is approximately 2 hours later than that at IITD.

344   Further, the maximum averaged SFC2 mixing ratio at MRIU is 4.41 ppbv, lower than the peak value of 6.20

345   ppbv at IITD, and consistent with the higher primary emission levels at IITD.

346   The two secondary VOC factors, in particular, show major differences between the two locations in terms of

347   diurnal variations and spectra. On a simple level, the two SecVOC2 are considered to represent related factors

348   because of their similar long-term trends with local temperature (Fig. 3 and 4). However, the mass spectra are

349   different, indicating different aging mechanisms. As illustrated in Fig. 3 and 4, the spectrum of SecVOC1 is

350   dominated by $C_4H_2O_3$ and $C_2H_4O_3$ at IITD but less so at MRIU for both SecVOC factors, especially in the case

351   of $C_4H_2O_3$. Further, the IITD SecVOC1 shows additional origins apart from the directions of primary factors,

352   indicating influences from transport (Fig. S8). The SecVOC2 at IITD, however, is comparable to the MRIU

353   SecVOC1 in terms of factor fingerprint (Fig. S9). The spectra of the two factors are both characterized by high

354   contributions to $C_3H_4O_2$, $C_4H_6O_{2-3}$, $C_5H_{10}O_{1-2}$, $C_6H_{10}O_{1-2}$, and $C_6H_{12}O_{1-2}$, with a relatively higher contribution to

355    $C_4H_2O_3$ and $C_2H_4O_3$ by IITD SecVOC2. However, their time trends and diurnal variations are rather different.

356   Although they both increase during daytime, the maximum is found about 4 hours later at MRIU than at IITD,

357   possibly due to less primary emission in the morning. The elevated periods of IITD SecVOC2 are not only

358   related to higher temperature, but also consistent with the periods with much stronger primary emissions (e.g.

359   Traffic factors). Similarly, the time series of MRIU SecVOC1 is also high in January when all the primary

360   factors are high compared to the rest of the periods. Besides, as shown in Fig. S8, SecVOC2 at IITD and

361   SecVOC1 at MRIU exhibit similar geographical origins as the primary factors at the respective sites, suggesting

362   local oxidation of primary emissions. The spectrum of MRIU SecVOC2 is also loaded with $C_4H_2O_3$, $C_2H_4O_3$, as

363   well as high C_4H_8O and $C_3H_4O_2$, but the contributions to these ions are much lower than SecVOC1. In addition,

364   MRIU SecVOC2 exhibits a similar time series as that of local temperature, which increases in the warmer

365   period of the campaign. The overall trend is quite different from the local primary factors, but has some

366   similarities with that of IITD SecVOC2. As shown in Fig. S8, high concentrations of MRIU SecVOC2 originate

367   mainly from north and northwest directions, consistent with the location of the IITD site. Therefore, it is

368   possible that MRIU SecVOC2 represents oxidized VOCs on a relatively regional scale.

369   Figure 6 presents the stacked factor diurnal patterns at the two sites. Much higher nocturnal mixing ratios of

370   VOCs are observed at IITD than at MRIU, leading to pronounced day-night differences, while during daytime

371   higher mixing ratios are found at MRIU than at IITD (with 15 ppbv at MRIU in the mid-afternoon, compared to

372   10 ppbv at IITD). This can be explained by the fact that at IITD the VOC mixing ratio is dominated by primary

373 emissions, especially during the night, while at MRIU the VOC mixing ratio is dominated by secondary

374 formation, at least during the day. At IITD, primary factors contribute to around 85% before midnight, with

375 large fractions coming from Traffic factors and SFC1. Afterwards, SFC2 gradually increases and contributes to

376 over 40% of the total mixing ratio before sunrise, when other primary factors decrease to about 50%. At the

377 suburban MRIU site, although the nocturnal mixing ratio of primary VOCs are about half as high as that at

378 IITD, they contribute to around 70% of the total VOCs concentrations. In the daytime, however, the SecVOC

379 contribution is as high as 54%, compared to the maximum of 41% at IITD. Generally, the MRIU SecVOC

380 contributes about 15-20% more than that at the IITD site throughout the day, mostly due to increased SecVOC1.

381 This may be partially explained by the differences in rotation techniques we applied in running the model for the

382 two secondary factors. As illustrated in Figure S10, by applying the similar a-value approach to MRIU as used

383 at IITD, the constrained MRIU SecVOC is around 1.5 ppb lower than the raw PMF result, and contributes a

384 maximum of 10% less at around 14:00-15:00 LT. However, SecVOC contributes 31% to the total analyzed

385 VOCs mixing ratio in the constrained MRIU solution, instead of a contribution of 34% from the unconstrained

386 result as illustrated above. Both PMF solutions resolved at MRIU show much higher fractions of SecVOC

387 compared to that at IITD (16%). Therefore, the rotation techniques only play a minor part on the discrepancies

388 between the two sites. More importantly, the difference in SecVOC is probably due to different oxidation

389 conditions at the two sites. Owing to the high mixing ratio of NO_x and the suppression of oxidants like OH

390 radicals, the chemical oxidation of primary VOCs occurs to a greater extent downwind of urban emission

391 sources. As shown in Fig. S8, high mixing ratios of SecVOC at MRIU originate from northwest directions,

392 different from the origins of local primary factors (southeast). Besides, the longer oxidation time may be another

393 reason of the higher SecVOC mixing ratios at MRIU. In addition, several ions below m/z 60 that can be detected

394 by the PTR-ToF-MS are excluded from PMF as discussed in Sec. 2.3, such as methanol, acetaldehyde, acetone

395 and acetic acids. These ions are 3-4 times higher than the dominant ions in the PMF analysis, possibly owing to

396 much higher emission rates and natural abundance. Besides, other excluded compounds such as C_1-C_4 alkanes

397 and C_1-C_4 alkenes which are not detectable by the PTR-ToF-MS are substantial contributors to the total VOC

398 mixing ratio as well. However, these ions are minor contributors to SOA formation and only substantially

399 contribute to the formation of ozone, which is a major issue in summer. Although the mixing ratio of the sum of

400 VOCs in the PMF only accounts for 39.6 % at IITD and 24.2 % at MRIU (Fig. S11), many of these compounds

401 are the dominant precursors in terms of SOA formation (Wu and Xie, 2017, 2018).

3.4 Evaluation of biogenic signatures

403 Biogenic VOCs, i.e. isoprene and monoterpenes, are not separated into a specific factor in this study. This is in

404 part due to the small number of ions that can be unambiguously assigned to these sources. Specifically, the

405 structure assignment and quantification of non-aromatic C_xH_y can be uncertain not only because of isomers, but

406 also because of fragments from aldehydes, alcohols, and other long-chain aliphatic hydrocarbons. Therefore,

407 C_5H_8 (m/z 69.070) and the sum of $C_{10}H_{16}$ (m/z 137.132) and C_6H_8 (m/z 81.070, a major fragment of $C_{10}H_{16}$) can

408 only serve as upper estimates of the actual isoprene and monoterpenes, respectively. The averaged mixing ratios

409 of isoprene and monoterpenes are low during the campaign, with 0.8 ppbv and 0.46 ppbv, respectively at IITD,

410 and 1.2 ppbv and 0.32 ppbv, respectively at MRIU. Figure 7 illustrates the explained variation of the two major

411 biogenic markers (i.e. isoprene and monoterpenes) and several other ions possibly related to their oxidation

412 products. Among them, $C_4H_6O_{1-2}$ (*m/z 71.049, m/z 87.044*) and $C_5H_8O_{1-2}$ (*m/z 85.065, m/z 101.060*) are found in

413 the photooxidation of isoprene, and $C_9H_{14}O$ (*m/z 139.112*) and $C_{10}H_{14/16}O$ (*m/z 151.112, m/z 153.127*) are

414 potential photochemical products from monoterpenes. Instead of being related to a specific secondary factor, the

415 biogenic markers are largely explained by primary factors. There are several possible explanations. First, both

416 isoprene and monoterpenes have been found in biomass burning emissions from numerous laboratory and field

417 studies (e.g. Bruns et al., 2017). Second, previous studies also indicated anthropogenic sources of isoprene from

418 vehicular emissions (Borbon et al., 2001;Wagner and Kuttler, 2014). One recent paper showed that fragrances

419 and personal care products may be an important emission source of urban monoterpenes and are correlated with

420 traffic emissions (McDonald et al., 2018). Third, considerable amounts of these ions may be affected by isomers

421 and fragments of larger ions originating from primary emissions. Finally, the two BVOCs are very reactive

422 compounds, which have very short lifetimes especially during daytime, and also during nighttime if NO_3 is

423 present. Therefore, large amounts of BVOCs can be degraded very quickly via photochemical oxidation during

424 daytime, resulting in lower daytime BVOC mixing ratios despite of stronger natural emissions. As shown in

425 Figure S12, both isoprene and monoterpene are high at night and very low during daytime, similar to the diurnal

426 trends of the primary factors. The only exception is isoprene at MRIU, which may be due to strong daytime

427 emissions or contributions from isomers/fragments.

428 As shown in Figure 7, the $C_xH_yO_z$ ions are explained partially by secondary factors, indicating possible

429 secondary formation from biogenic precursors. However, at both sites, considerable amounts of these ions are

430 also explained by emissions from SFC. Previous studies showed that the emission factors of $C_4H_6O_{1-2}$ are

431 comparable to that of benzene (200 mg kg^{-1}), and the EFs for $C_5H_8O_{1-2}$ and $C_{10}H_{16}O$ are comparable to that of

432 toluene (27 mg kg^{-1}) in biomass/wood burning emissions (Bruns et al., 2017; Koss et al., 2018). Besides, as

433 these small molecular weight ions are possible products from various precursors and chemical pathways, their

434 structures and precursor definitions are uncertain. However, the explained variations can still provide

435 information on the aging processes. In general, the fraction explained by secondary factors increases as the ions

436 become more oxygenated. In addition, at IITD SecVOC1 explains much higher fractions of the $C_xH_yO_1$ ions

437 compared to SecVOC2, while the $C_xH_yO_{2-3}$ ions are to a great extent explained by SecVOC2. As some of these

438 less oxygenated ions originate from both fast photochemical formation and later-generation production, they are

439 explained by both secondary factors. For example C_4H_6O, which can be attributed to MVK (methyl vinyl

440 ketone) and MACR (methacrolein), is formed rapidly via isoprene OH oxidation, but can also be found as

441 second generation products via several other oxidation pathways. Generally, both biogenic markers and their

442 oxidation products are low during the measurement period, and these ions contribute only small amounts to the

443 two secondary factors, indicating minor contributions from biogenic emissions at the two sites during the

444 campaign.

445 **3.5 Characteristics of selected ions**

446 Figure 8 shows the diurnal patterns of the explained variation of each PMF factor to selected ions measured at

447 the two sites, stacked such that the height is the total modelled concentration. The red line in each plot

448 represents the mean diurnal pattern of the measured mixing ratio. The model represents over 80% of most of

449 these selected ions, and thus in general explains the variations of these ions well. The comparison of these

450 diurnal patterns illustrates that distinct differences exist in the mixing ratios, variations, and source compositions

451 of certain VOCs. The assignment of a specific compound structure to each ion can be ambiguous, particularly

452 for the spectra obtained from ambient air due to the possibility of isomers and/or fragmentation. Still, we

453 provide tentative candidates for the nine ions based on their molecular formulas and compounds found in

454 previous fuel combustion and oxidation simulation studies. Structural assignments for aromatics and PAHs are

455 more certain because the low H to C atomic ratio reduces the number of reasonable structures, therefore, C_6H_6,

456 C_7H_8 and $C_{10}H_8$ are mainly attributed to benzene, toluene, and naphthalene, respectively. Also included in Fig. 8

457 are $C_5H_4O_2$, C_6H_6O, $C_2H_4O_3$, $C_4H_2O_3$, and $C_6H_5NO_3$; we suggest that the most likely structural assignments are

458 furfural, phenol, a fragment of peroxyacetyl nitrate (PAN), maleic anhydride, and nitrophenol, respectively.

459 Benzene and toluene are well-interpreted aromatics, which contribute significantly to emissions from several

460 combustion processes, including vehicle emissions, biomass burning, coal combustion, etc. (Pieber et al., 2018;

461 Bruns et al., 2017; Klein et al., 2018). As shown in Fig. 8 (a), at IITD benzene originates to a significant extent

462 from traffic emissions (53% on average) over the whole day and in addition from the SFC factors (47% on

463 average), largely at night. The traffic fraction is lowest (29%) in the early morning and increases during

464 daytime, with a maximum of 74 % around 17:00-18:00 LT, with the opposite trend for the SFC fraction. This

465 suggests that the IITD site is dominated by pollution from vehicular emissions because of the highly trafficked

466 streets nearby, and also influenced by strong solid fuel combustion due to the high population density. At

467 MRIU, however, there are comparable benzene contributions by the Traffic1, SFC1, and SFC2 factors. Even

468 though the sum of the SFC mixing ratios is lower at MRIU compared to IITD, the relative SFC contribution is

469 higher at MRIU. Although the traffic fraction increases during daytime at both sites, it explains a maximum of

470 74% at IITD and 49% at MRIU around 17:00-18:00 LT, with the opposite behavior for the SFC fraction.

471 Toluene displays a similar diurnal cycle and source contributions as benzene, characterized by significant

472 contributions from primary sources, especially Traffic1. The traffic fraction is as high as 90 % around 18:00 LT

473 and reaches a minimum of 67 % around 10:00 LT at IITD, which is much higher than that of benzene. The

474 benzene/toluene (B/T) of the traffic factor ranges from 0.34 to 0.40, which is comparable to the emission factors

475 of benzene/toluene for gasoline emissions (0.58, Gentner et al., 2013). Besides, the traffic fraction is lowest in

476 the morning, which is associated with the strong emissions of SFC in the morning and high traffic emissions

477 during the rush hour in the late afternoon. The B/T is 3.4 for the SFC factors, which is in the reported range of

478 2-7 for residential biomass/wood burning (Bruns et al., 2017; Koss et al., 2018). Similarly, traffic is also the

479 highest source of toluene at MRIU, ranging from 67%-77%.

480 As mentioned above, phenolic compounds and substituted furans are found to be significant oxygen-containing

481 compounds at the two sites. Furfural is an important marker from biomass burning, which arises from pyrolysis

482 of cellulose and hemicellulose. Phenols are detected in biomass burning and coal combustion emissions as the

483 most abundant oxygenated aromatic compounds. (Stockwell et al., 2015; Klein et al., 2018; Bruns et al., 2017)

484 Indeed, as shown in Fig. 8 (c,f), the majority of furfural and phenol are explained by the two SFC factors at both

485 sites. The two SFC factors contribute 91% to furfural and 85% to phenol at IITD, while the corresponding

486 contributions are 70% and 57% at MRIU. SFC1 contributes around 54% and 39% during the night and morning,

487 respectively, with a higher night-to-day ratio at IITD, which captures both the characteristics of residential

488 heating styles and urban/suburban usage differences due to population density. The SFC2-to-SFC1 ratio

489 generally increases during the night. As described earlier, one possible explanation is that SFC2 results from the

490   fast nocturnal evolution of primary SFC1 emissions probably driven by night chemistry under high NO_x

491   conditions (Tiitta et al., 2016). In addition to the SFC factors, at IITD around 8% of furfural and 4% of phenol

492   are explained by the two SecVOC factors, compared to around 11% and 25% at MRIU, respectively.

493   Nitrated phenols are important biomass burning tracers, in particular for secondary aerosol formation (Mohr et

494   al., 2013; Bertrand et al., 2018). Here, the diurnal cycles of $C_6H_5NO_3$ at the two sites illustrate contributions

495   from both emission and atmospheric transformation. At both sites, SFC2 accounts for a large percentage of the

496   nocturnal nitrophenol mixing ratios while the SecVOC factors explain the remaining fraction. This is consistent

497   with SFC2 representing nighttime aging of SFC emissions. It is worth noting that during the night there is a

498   much larger unexplained fraction (around 35%) at IITD than at MRIU, with a corresponding uncertainty in the

499   source contributions. Still, the main difference at the two sites relates to the contributions of the SecVOC

500   factors. At IITD, the SecVOC2 factor contributes more at night than during daytime, probably due to nocturnal

501   formation including NO_3 chemistry and the photolysis reactions of $C_6H_5NO_3$ during daytime. At MRIU, both

502   SecVOC1 and SecVOC2 contribute to $C_6H_5NO_3$, indicating the combined influence from local and regional

503   chemistry.

504   Maleic anhydride and PAN display distinct variation patterns at the two sites. Maleic anhydride is the product

505   from the photo-oxidation of furans, unsaturated carbonyls, and aromatics (Bierbach et al., 1994; Yuan et al.,

506   2017). The mixing ratio of maleic anhydride is about 5 times higher at IITD than that at MRIU during daytime

507   and is only explained by SecVOC1. As shown in Fig S13, high mixing ratios of maleic anhydride originate from

508   north (primary emissions) and northwest of IITD, the same directions as SecVOC1. Similarly, high mixing

509   ratios of maleic anhydride at MRIU originate both from southeast (primary emissions) and northwest (IITD site)

510   directions. During nighttime, the maleic anhydride mixing ratio at MRIU is similar to the one at IITD, with

511   considerable amounts explained by SecVOC2 and SFC2. A recent study indicates that furans contributed to over

512   90% of maleic anhydride within the first 4 hours of OH oxidation in biomass burning plumes (Coggon et al.,

513   2019). Therefore, the higher morning peak and nocturnal mixing ratios of furans may be the driver of much

514   higher maleic anhydride at IITD. Besides, it is possible that under the conditions of high humidity and aerosol

515   mass loading, maleic anhydride could react and/or condense away before the city plume reaching the suburban

516   area. In contrast to maleic anhydride, $C_2H_4O_3$ shows a higher mixing at MRIU than at IITD. $C_2H_4O_3$ is mainly

517   attributed to a fragment of PAN, however, a small fraction is also related to primary factors at both sites (Fig. 8).

518   This ion was reported in laboratory studies on primary biomass burning emissions (Bruns et al., 2017). Possibly,

519   small amounts of this ion might be attributed to unknown isomers, which are directly emitted or rapidly formed

520   in combustion processes. PAN is formed in VOC oxidation with NO_x (de Gouw and Warneke, 2007). Due to its

521   long lifetime at low temperature, PAN can undergo long-range transport and then play an important role in

522   tropospheric ozone formation in the remote areas when it decomposes (Fischer et al., 2014). As illustrated in

523   Fig. S13, high mixing ratios of $C_2H_4O_3$ originate from north and far northwest, indicating influence of both local

524   emission/oxidation formation and long-range transportation. At MRIU, however, high mixing ratios of $C_2H_4O_3$

525   mainly originate from northwest, suggesting the influence of transportation from the urban areas. Besides, PAN

526   is mainly explained by SecVOC1 at IITD, and a much higher PAN mixing ratio is observed at the suburban

527   MRIU, originating from both secondary VOC factors, indicating a longer oxidation time at this site.

528   **4 Conclusions**

Measurements of volatile organic compounds (VOCs) were performed simultaneously at two different sites in Delhi, India using two PTR-ToF-MS instruments. Much higher mixing ratios of VOCs were observed at the urban IITD site, with higher nocturnal contributions from anthropogenic emissions than at the suburban MRIU site. Using positive matrix factorization, we found a 6-factor solution at both sites, each consisting of two traffic factors, two solid fuel combustion factors, and two secondary factors. Anthropogenic activities were shown to be important VOCs sources at both sites. Among them, traffic related emissions comprised the dominant source, contributing 56.6% at the urban IITD site and 36.0% at the suburban MRIU site to the total analyzed VOC mixing ratio (excluding high-intensity VOCs as described in Section 2.3). Solid fuel combustion contributed 27.5% at the urban IITD site, and even 30.4% at the suburban MRIU site. Secondary formation was the most important source of VOCs during daytime, and contributed 15.9% at IITD and 33.6% at MRIU to the total concentration. Higher mixing ratios of oxygenated VOCs were found at the suburban site, likely due to suppression of oxidant levels in the urban atmosphere and longer aging time at the suburban site.

Comparison of the factor diurnals and profiles indicated that the anthropogenic sources were similar at two sites, even though the VOC mixing ratios at MRIU were much lower than IITD. The secondary factors, however, were different at the two sites both in terms of diurnal variations and of spectra. The two secondary VOC factors exhibited higher mixing ratios and contributions at the suburban site, in particular during daytime. At IITD, the two SecVOC factors indicated different aging process under strong primary emissions, while at the suburban MRIU they suggested local oxidation and regional aging. Besides, evaluation of biogenic markers indicated very small influence from biogenic emissions and their oxidation products.

This work highlights the crucial role that anthropogenic sources play in the pollution levels and variation characteristics of VOCs in the ambient atmosphere of Delhi, India. Further control measures to reduce emissions from traffic exhaust and solid fuel combustion are urgently needed to mitigate the severe pollution and the environmental impact of VOCs as well as aerosols (not investigated in this paper) in this region. Significant differences in the concentrations as well as the pollution sources stress the complexity of both emission and chemistry in this region. Long-term measurements of a set of VOCs, aerosols, and other oxidants would be ideal to obtain a more detailed understanding of the formation mechanisms during various conditions.

Data availability

Data will be available at a zenodo repository (at the time of publication).

565 **Reference**

566 Atkinson, R., Baulch, D. L., Cox, R. A., Jr., R. F. H., Kerr, J. A., and Troe, J.: Evaluated kinetic and Photochemical

567   Data for Atmospheric Chemistry: Supplement III. IUPAC Subcommittee on Gas Kinetic Data Evaluation for

568   Atmospheric Chemistry, Journal of Physical and Chemical Reference Data, 18, 881-1097, 10.1063/1.555832,

569   1989.

570 Atkinson, R.: Atmospheric chemistry of VOCs and NOx, Atmospheric Environment, 34, 2063-2101,

571   https://doi.org/10.1016/S1352-2310(99)00460-4, 2000.

572 Atkinson, R., and Arey, J.: Atmospheric degradation of volatile organic compounds, Chemical reviews, 103, 4605-

573   4638, 2003.

574 Aoki, N., Inomata, S., and Tanimoto, H.: Detection of C1–C5 alkyl nitrates by proton transfer reaction time-of-

575   flight mass spectrometry, International Journal of Mass Spectrometry, 263, 12-21,

576   https://doi.org/10.1016/j.ijms.2006.11.018, 2007.

577 Baudic, A., Gros, V., Sauvage, S., Locoge, N., Sanchez, O., Sarda-Estève, R., Kalogridis, C., Petit, J. E., Bonnaire,

578   N., Baisnée, D., Favez, O., Albinet, A., Sciare, J., and Bonsang, B.: Seasonal variability and source

579   apportionment of volatile organic compounds (VOCs) in the Paris megacity (France), Atmos. Chem. Phys., 16,

580   11961-11989, 10.5194/acp-16-11961-2016, 2016.

581 Bertrand, A., Stefenelli, G., Jen, C. N., Pieber, S. M., Bruns, E. A., Ni, H. Y., Temime-Roussel, B., Slowik, J. G.,

582   Goldstein, A. H., El Haddad, I., Baltensperger, U., Prevot, A. S. H., Wortham, H., and Marchand, N.: Evolution

583   of the chemical fingerprint of biomass burning organic aerosol during aging, Atmos. Chem. Phys., 10, 7607-

584   7624, 10.5194/acp-18-7607-2018, 2018.

585 Bierbach, A., Barnes, I., Becker, K. H., and Wiesen, E.: Atmospheric Chemistry of Unsaturated Carbonyls:

586   Butenedial, 4-Oxo-2-pentenal, 3-Hexene-2,5-dione, Maleic Anhydride, 3H-Furan-2-one, and 5-Methyl-3H-

587   furan-2-one, Environmental Science & Technology, 28, 715-729, 10.1021/es00053a028, 1994.

588 Bonasoni, P., Laj, P., Marinoni, A., Sprenger, M., Angelini, F., Arduini, J., Bonafè, U., Calzolari, F., Colombo, T.,

589   Decesari, S., Di Biagio, C., di Sarra, A. G., Evangelisti, F., Duchi, R., Facchini, M. C., Fuzzi, S., Gobbi, G. P.,

590   Maione, M., Panday, A., Roccato, F., Sellegri, K., Venzac, H., Verza, G. P., Villani, P., Vuillermoz, E., and

591   Cristofanelli, P.: Atmospheric brown clouds in the Himalayas: first two years of continuous observations at the

592   Nepal Climate Observatory-Pyramid (5079 m), Atmos. Chem. Phys., 10, 7515-7531, 10.5194/acp-10-7515-

593   2010, 2010.

594 Borbon, A., Fontaine, H., Veillerot, M., Locoge, N., Galloo, J. C., and Guillermo, R.: An investigation into the

595   traffic-related fraction of isoprene at an urban location, Atmospheric Environment, 35, 3749-3760,

596   https://doi.org/10.1016/S1352-2310(01)00170-4, 2001.

597 Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez,

598   M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D.

599   D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude

600   megacities: Observations versus emission inventories in Los Angeles and Paris, Journal of Geophysical

601   Research: Atmospheres, 118, 2041-2057, 10.1002/jgrd.50059, 2013.

602 Bruns, E. A., Slowik, J. G., El Haddad, I., Kilic, D., Klein, F., Dommen, J., Temime-Roussel, B., Marchand, N.,

603   Baltensperger, U., and Prévôt, A. S. H.: Characterization of gas-phase organics using proton transfer reaction

604   time-of-flight mass spectrometry: fresh and aged residential wood combustion emissions, Atmos. Chem. Phys.,

605   17, 705-720, 10.5194/acp-17-705-2017, 2017.

606   Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: SoFi, an IGOR-based interface

607   for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2

608   application to aerosol mass spectrometer data, Atmospheric Measurement Techniques, 6, 3649-3661,

609   10.5194/amt-6-3649-2013, 2013.

610   Cao, X., Yao, Z., Shen, X., Ye, Y., and Jiang, X.: On-road emission characteristics of VOCs from light-duty

611   gasoline vehicles in Beijing, China, Atmospheric Environment, 124, 146-155,

612   https://doi.org/10.1016/j.atmosenv.2015.06.019, 2016.

613   Cappellin, L., Karl, T., Probst, M., Ismailova, O., Winkler, P. M., Soukoulis, C., Aprea, E., Märk, T. D., Gasperi,

614   F., and Biasioli, F.: On quantitative determination of volatile organic compound concentrations using proton

615   transfer reaction time-of-flight mass spectrometry, Environmental Science & Technology, 46, 2283-2290,

616   10.1021/es203985t, 2012.

617   Coggon, M. M., Lim, C. Y., Koss, A. R., Sekimoto, K., Yuan, B., Gilman, J. B., Hagan, D. H., Selimovic, V.,

618   Zarzana, K., Brown, S. S., Roberts, J. M., Müller, M., Yokelson, R., Wisthaler, A., Krechmer, J. E., Jimenez, J.

619   L., Cappa, C., Kroll, J., de Gouw, J., and Warneke, C.: OH-chemistry of non-methane organic gases (NMOG)

620   emitted from laboratory and ambient biomass burning smoke: evaluating the influence of furans and oxygenated

621   aromatics on ozone and secondary NMOG formation, Atmos. Chem. Phys. Discuss., 2019, 1-42, 10.5194/acp-

622   2019-516, 2019.

623   Crutzen, P. J., and Andreae, M. O.: Biomass Burning in the Tropics: Impact on Atmospheric Chemistry and

624   Biogeochemical Cycles, Science, 250, 1669-1678, 10.1126/science.250.4988.1669, 1990.

625   de Gouw, J., and Warneke, C.: Measurements of volatile organic compounds in the earth's atmosphere using

626   proton-transfer-reaction mass spectrometry, Mass Spectrom Rev, 26, 223-257, 10.1002/mas.20119, 2007.

627   de Gouw, J. A., Middlebrook, A. M., Warneke, C., Goldan, P. D., Kuster, W. C., Roberts, J. M., Fehsenfeld, F. C.,

628   Worsnop, D. R., Canagaratna, M. R., Pszenny, A. A. P., Keene, W. C., Marchewka, M., Bertman, S. B., and

629   Bates, T. S.: Budget of organic carbon in a polluted atmosphere: Results from the New England Air Quality

630   Study in 2002, Journal of Geophysical Research: Atmospheres, 110, doi:10.1029/2004JD005623, 2005.

631   Derwent, R. G., Jenkin, M. E., Utembe, S. R., Shallcross, D. E., Murrells, T. P., and Passant, N. R.: Secondary

632   organic aerosol formation from a large number of reactive man-made organic compounds, Science of The Total

633   Environment, 408, 3374-3381, https://doi.org/10.1016/j.scitotenv.2010.04.013, 2010.

634   Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann, R.,

635   Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I.-H., Rissanen, M., Jokinen, T., Schobesberger, S., Kangasluoma,

636   J., Kontkanen, J., Nieminen, T., Kurtén, T., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., Canagaratna, M.,

637   Maso, M. D., Berndt, T., Petäjä, T., Wahner, A., Kerminen, V.-M., Kulmala, M., Worsnop, D. R., Wildt, J., and

638   Mentel, T. F.: A large source of low-volatility secondary organic aerosol, Nature, 506, 476,

639   10.1038/nature13032, 2014.

640   Fischer, E. V., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Millet, D. B., Mao, J., Paulot, F., Singh, H. B.,

641   Roiger, A., Ries, L., Talbot, R. W., Dzepina, K., and Pandey Deolal, S.: Atmospheric peroxyacetyl nitrate (PAN):

642   a global budget and source attribution, Atmos. Chem. Phys., 14, 2679-2698, 10.5194/acp-14-2679-2014, 2014.

643   Fraser, M. P., Cass, G. R., and Simoneit, B. R. T.: Gas-Phase and Particle-Phase Organic Compounds Emitted

644   from Motor Vehicle Traffic in a Los Angeles Roadway Tunnel, Environmental Science & Technology, 32, 2051-

645 2060, 10.1021/es970916e, 1998.

646 Gani, S., Bhandari, S., Seraj, S., Wang, D. S., Patel, K., Soni, P., Arub, Z., Habib, G., Hildebrandt Ruiz, L., and

647 Apte, J. S.: Submicron aerosol composition in the world's most polluted megacity: the Delhi Aerosol Supersite

648 study, Atmos. Chem. Phys., 19, 6843-6859, 10.5194/acp-19-6843-2019, 2019.

649 Gentner, D. R., Worton, D. R., Isaacman, G., Davis, L. C., Dallmann, T. R., Wood, E. C., Herndon, S. C., Goldstein,

650 A. H., and Harley, R. A.: Chemical Composition of Gas-Phase Organic Carbon Emissions from Motor Vehicles

651 and Implications for Ozone Production, Environmental Science & Technology, 47, 11837-11848,

652 10.1021/es401470e, 2013.

653 Graus, M., Muller, M., and Hansel, A.: High resolution PTR-TOF: quantification and formula confirmation of

654 VOC in real time, J Am Soc Mass Spectrom, 21, 1037-1044, 10.1016/j.jasms.2010.02.006, 2010.

655 Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N.

656 M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M.

657 E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S.

658 H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of

659 secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9, 5155-5236, 10.5194/acp-9-

660 5155-2009, 2009.

661 Hoque, R. R., Khillare, P. S., Agarwal, T., Shridhar, V., and Balachandran, S.: Spatial and temporal variation of

662 BTEX in the urban atmosphere of Delhi, India, Science of The Total Environment, 392, 30-40,

663 https://doi.org/10.1016/j.scitotenv.2007.08.036, 2008.

664 IPCC: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth

665 Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D.,

666 Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.,

667 Cambridge University Press, Cambridge, UK and New York, USA, 2013.

668 Jordan, A., Haidacher, S., Hanel, G., Hartungen, E., Märk, L., Seehauser, H., Schottkowsky, R., Sulzer, P., and

669 Märk, T. D.: A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer

670 (PTR-TOF-MS), International Journal of Mass Spectrometry, 286, 122-128, 10.1016/j.ijms.2009.07.005, 2009.

671 Kampa, M., and Castanas, E.: Human health effects of air pollution, Environmental Pollution, 151, 362-367,

672 https://doi.org/10.1016/j.envpol.2007.06.012, 2008.

673 Karl, T. G., Christian, T. J., Yokelson, R. J., Artaxo, P., Hao, W. M., and Guenther, A.: The Tropical Forest and

674 Fire Emissions Experiment: method evaluation of volatile organic compound emissions measured by PTR-MS,

675 FTIR, and GC from tropical biomass burning, Atmos. Chem. Phys., 7, 5883-5897, 10.5194/acp-7-5883-2007,

676 2007.

677 Klein, F., Pieber, S. M., Ni, H., Stefenelli, G., Bertrand, A., Kilic, D., Pospisilova, V., Temime-Roussel, B.,

678 Marchand, N., El Haddad, I., Slowik, J. G., Baltensperger, U., Cao, J., Huang, R.-j., and Prévôt, A. S. H.:

679 Characterization of Gas-Phase Organics Using Proton Transfer Reaction Time-of-Flight Mass Spectrometry:

680 Residential Coal Combustion, Environmental Science & Technology, 52, 2612-2617, 10.1021/acs.est.7b03960,

681 2018.

682 Koss, A. R., Sekimoto, K., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M.,

683 Brown, S. S., Jimenez, J. L., Krechmer, J., Roberts, J. M., Warneke, C., Yokelson, R. J., and de Gouw, J.: Non-

684 methane organic gas emissions from biomass burning: identification, quantification, and emission factors from

685   PTR-ToF during the FIREX 2016 laboratory experiment, Atmos. Chem. Phys., 18, 3299-3319, 10.5194/acp-
686   18-3299-2018, 2018.

687   Kumar, P., Khare, M., Harrison, R. M., Bloss, W. J., Lewis, A. C., Coe, H., and Morawska, L.: New directions:
688   Air pollution challenges for developing megacities like Delhi, Atmospheric Environment, 122, 657-661,
689   https://doi.org/10.1016/j.atmosenv.2015.10.032, 2015.

690   Languille, B., Gros, V., Petit, J.-E., Honoré, C., Baudic, A., Perrussel, O., Foret, G., Michoud, V., Truong, F.,
691   Bonnaire, N., Sarda-Estève, R., Delmotte, M., Feron, A., Maisonneuve, F., Gaimoz, C., Formenti, P., Kotthaus,
692   S., Haeffelin, M., and Favez, O.: Wood burning: A major source of Volatile Organic Compounds during
693   wintertime in the Paris region, Science of The Total Environment, 135055,
694   https://doi.org/10.1016/j.scitotenv.2019.135055, 2019.

695   Lawrence, M. G., and Lelieveld, J.: Atmospheric pollutant outflow from southern Asia: a review, Atmos. Chem.
696   Phys., 10, 11017-11096, 10.5194/acp-10-11017-2010, 2010.

697   Mahata, K. S., Rupakheti, M., Panday, A. K., Bhardwaj, P., Naja, M., Singh, A., Mues, A., Cristofanelli, P.,
698   Pudasainee, D., Bonasoni, P., and Lawrence, M. G.: Observation and analysis of spatiotemporal characteristics
699   of surface ozone and carbon monoxide at multiple sites in the Kathmandu Valley, Nepal, Atmos. Chem. Phys.,
700   18, 14113-14132, 10.5194/acp-18-14113-2018, 2018.

701   Meylan, W. M., and Howard, P. H.: Computer estimation of the Atmospheric gas-phase reaction rate of organic
702   compounds with hydroxyl radicals and ozone, Chemosphere, 26, 2293-2299, https://doi.org/10.1016/0045-
703   6535(93)90355-9, 1993.

704   McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-
705   Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G.,
706   Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B., and Trainer, M.: Volatile chemical
707   products emerging as largest petrochemical source of urban organic emissions, Science, 359, 760-764,
708   10.1126/science.aaq0524, 2018.

709   Mohr, C., Lopez-Hilfiker, F. D., Zotter, P., Prévôt, A. S. H., Xu, L., Ng, N. L., Herndon, S. C., Williams, L. R.,
710   Franklin, J. P., Zahniser, M. S., Worsnop, D. R., Knighton, W. B., Aiken, A. C., Gorkowski, K. J., Dubey, M.
711   K., Allan, J. D., and Thornton, J. A.: Contribution of Nitrated Phenols to Wood Burning Brown Carbon Light
712   Absorption in Detling, United Kingdom during Winter Time, Environmental Science & Technology, 47, 6316-
713   6324, 10.1021/es400683v, 2013.

714   Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A.,
715   Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou,
716   E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J.,
717   Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A., Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z.,
718   Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Liousse, C., Maione, M., McFiggans, G., Metzger, A.,
719   Mieville, A., Moussiopoulos, N., Orlando, J. J., O'Dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt,
720   U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson,
721   D., ten Brink, H., Theloke, J., van der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, C., and von Glasow,
722   R.: Atmospheric composition change – global and regional air quality, Atmospheric Environment, 43, 5268-
723   5350, https://doi.org/10.1016/j.atmosenv.2009.08.021, 2009.

724   Müller, M., Graus, M., Wisthaler, A., Hansel, A., Metzger, A., Dommen, J., and Baltensperger, U.: Analysis of

725   high mass resolution PTR-TOF mass spectra from 1,3,5-trimethylbenzene (TMB) environmental chamber
726   experiments, Atmospheric Chemistry and Physics, 12, 829-843, 10.5194/acp-12-829-2012, 2012.

727   Nurmatov, U. B., Tagieva, N., Semple, S., Devereux, G., and Sheikh, A.: Volatile organic compounds and risk of
728   asthma and allergy: a systematic review and meta-analysis of observational and interventional studies, Primary
729   Care Respiratory Journal, 22, PS9, 10.4104/pcrj.2013.00010, 2013.

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

732   Paatero, P.: The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems,
733   Including the n-Way Parallel Factor Analysis Model, Journal of Computational and Graphical Statistics, 8, 854-
734   888, 10.1080/10618600.1999.10474853, 1999.

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

737   Pallavi, Sinha, B., and Sinha, V.: Source apportionment of volatile organic compounds in the northwest Indo-
738   Gangetic Plain using a positive matrix factorization model, Atmos. Chem. Phys., 19, 15467-15482,
739   10.5194/acp-19-15467-2019, 2019.

740   Pieber, S. M., Kumar, N. K., Klein, F., Comte, P., Bhattu, D., Dommen, J., Bruns, E. A., Kılıç, D., El Haddad, I.,
741   Keller, A., Czerwinski, J., Heeb, N., Baltensperger, U., Slowik, J. G., and Prévôt, A. S. H.: Gas-phase
742   composition and secondary organic aerosol formation from standard and particle filter-retrofitted gasoline
743   direct injection vehicles investigated in a batch and flow reactor, Atmos. Chem. Phys., 18, 9929-9954,
744   10.5194/acp-18-9929-2018, 2018.

745   Platt, S. M., El Haddad, I., Pieber, S. M., Zardini, A. A., Suarez-Bertoa, R., Clairotte, M., Daellenbach, K. R.,
746   Huang, R. J., Slowik, J. G., Hellebust, S., Temime-Roussel, B., Marchand, N., de Gouw, J., Jimenez, J. L.,
747   Hayes, P. L., Robinson, A. L., Baltensperger, U., Astorga, C., and Prévôt, A. S. H.: Gasoline cars produce more
748   carbonaceous particulate matter than modern filter-equipped diesel cars, Scientific Reports, 7, 4926,
749   10.1038/s41598-017-03714-9, 2017.

750   Rai, P., Furger, M., Haddad, I. E., Kumar, V., Wang, L., Singh, A., Dixit, K., Bhattu, D., Petit, J.-E., Ganguly, D.,
751   Rastogi, N., Baltensperger, U., Tripathi, S. N., Slowik, J. G., and Prevot, A. S. H.: Real-time measurement and
752   source apportionment of elements in Delhi's atmosphere, under review in Science of the Total Environment,
753   2020.

754   Sahu, L. K., and Saxena, P.: High time and mass resolved PTR-TOF-MS measurements of VOCs at an urban site
755   of India during winter: Role of anthropogenic, biomass burning, biogenic and photochemical sources,
756   Atmospheric Research, 164-165, 84-94, https://doi.org/10.1016/j.atmosres.2015.04.021, 2015.

757   Sahu, S. K., Beig, G., and Parkhi, N. S.: Emissions inventory of anthropogenic PM2.5 and PM10 in Delhi during
758   Commonwealth   Games   2010,   Atmospheric   Environment,   45,   6180-6190,
759   https://doi.org/10.1016/j.atmosenv.2011.08.014, 2011.

760   Sekimoto, K., Koss, A. R., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M.,
761   Brown, S. S., Warneke, C., Yokelson, R. J., Roberts, J. M., and de Gouw, J.: High- and low-temperature
762   pyrolysis profiles describe volatile organic compound emissions from western US wildfire fuels, Atmos. Chem.
763   Phys., 18, 9263-9281, 10.5194/acp-18-9263-2018, 2018.

764   Sharma, S. K., Mandal, T. K., Saxena, M., Rashmi, Rohtash, Sharma, A., and Gautam, R.: Source apportionment

765   of PM10 by using positive matrix factorization at an urban site of Delhi, India, Urban Climate, 10, 656-670,

766   https://doi.org/10.1016/j.uclim.2013.11.002, 2014.

767   Srivastava, A.: Source apportionment of ambient VOCS in Mumbai city, Atmospheric Environment, 38, 6829-

768   6843, https://doi.org/10.1016/j.atmosenv.2004.09.009, 2004.

769   Srivastava, A., Sengupta, B., and Dutta, S. A.: Source apportionment of ambient VOCs in Delhi City, Science of

770   The Total Environment, 343, 207-220, https://doi.org/10.1016/j.scitotenv.2004.10.008, 2005.

771   Stockwell, C. E., Veres, P. R., Williams, J., and Yokelson, R. J.: Characterization of biomass burning emissions

772   from cooking fires, peat, crop residue, and other fuels with high-resolution proton-transfer-reaction time-of-

773   flight mass spectrometry, Atmos. Chem. Phys., 15, 845-865, 10.5194/acp-15-845-2015, 2015.

774   Tiitta, P., Leskinen, A., Hao, L., Yli-Pirilä, P., Kortelainen, M., Grigonyte, J., Tissari, J., Lamberg, H., Hartikainen,

775   A., Kuuspalo, K., Kortelainen, A. M., Virtanen, A., Lehtinen, K. E. J., Komppula, M., Pieber, S., Prévôt, A. S.

776   H., Onasch, T. B., Worsnop, D. R., Czech, H., Zimmermann, R., Jokiniemi, J., and Sippula, O.: Transformation

777   of logwood combustion emissions in a smog chamber: formation of secondary organic aerosol and changes in

778   the primary organic aerosol upon daytime and nighttime aging, Atmos. Chem. Phys., 16, 13251-13269,

779   10.5194/acp-16-13251-2016, 2016.

780   Wagner, P., and Kuttler, W.: Biogenic and anthropogenic isoprene in the near-surface urban atmosphere — A case

781   study in Essen, Germany, Science of The Total Environment, 475, 104-115,

782   https://doi.org/10.1016/j.scitotenv.2013.12.026, 2014.

783   WHO, Ambient (outdoor) air pollution in cities database 2014, report, World Health Organization, Geneva, 2014.

784   Wu, R., and Xie, S.: Spatial Distribution of Ozone Formation in China Derived from Emissions of Speciated

785   Volatile Organic Compounds, Environmental Science & Technology, 51, 2574-2583, 10.1021/acs.est.6b03634,

786   2017.

787   Wu, R., and Xie, S.: Spatial Distribution of Secondary Organic Aerosol Formation Potential in China Derived

788   from Speciated Anthropogenic Volatile Organic Compound Emissions, Environmental Science & Technology,

789   52, 8146-8156, 10.1021/acs.est.8b01269, 2018.

790   Yao, Z., Shen, X., Ye, Y., Cao, X., Jiang, X., Zhang, Y., and He, K.: On-road emission characteristics of VOCs

791   from diesel trucks in Beijing, China, Atmospheric Environment, 103, 87-93,

792   https://doi.org/10.1016/j.atmosenv.2014.12.028, 2015.

793   Yokelson, R. J., Crounse, J. D., DeCarlo, P. F., Karl, T., Urbanski, S., Atlas, E., Campos, T., Shinozuka, Y.,

794   Kapustin, V., Clarke, A. D., Weinheimer, A., Knapp, D. J., Montzka, D. D., Holloway, J., Weibring, P., Flocke,

795   F., Zheng, W., Toohey, D., Wennberg, P. O., Wiedinmyer, C., Mauldin, L., Fried, A., Richter, D., Walega, J.,

796   Jimenez, J. L., Adachi, K., Buseck, P. R., Hall, S. R., and Shetter, R.: Emissions from biomass burning in the

797   Yucatan, Atmos. Chem. Phys., 9, 5785-5812, 10.5194/acp-9-5785-2009, 2009.

798   Yuan, Y., Zhao, X., Wang, S., and Wang, L.: Atmospheric Oxidation of Furan and Methyl-Substituted Furans

799   Initiated by Hydroxyl Radicals, The Journal of Physical Chemistry A, 121, 9306-9319,

800   0.1021/acs.jpca.7b09741, 2017.

801

Figure captions:

Fig. 1 Maps of the study region and the two sampling sites (from Google Maps). The black circle denotes the urban site at IITD and the black square the suburban site at MRIU.

Fig. 2 Temporal variations of CO, NO_x, analyzed VOC mixing ratios and temperature as well as the contributions of seven VOC families at the two sites.

Fig. 3 PMF results at IITD, showing (a) factor profiles (b) time series and (c) diurnal patterns. In (a), the left axis for each factor profile is the relative composition of each factor (i.e., horizontal sum is 1) and the right axis is the relative contribution of each factor to a given ion. Ions are colored based on the seven family classes, as described in Section 3.1. (b) Temporal evolution of resolved factors at IITD, together with external reference data, e.g. NOx, CO, solar radiation and temperature. (c) Medians of diurnal cycles of factors at IITD, shaded with interquartile ranges as well as the 10th and 90th percentiles.

Fig. 4 PMF results at MRIU, showing (a) factor profiles (b) time series, and (c) diurnal patterns. (a) Relative composition (left axis) and relative contribution (right axis) of each factor to a given ion. Ions are colored based on the seven family classes, as described in Section 3.1. (b) Temporal evolution of resolved factors at MRIU, together with external reference data, e.g. NO_x, CO, solar radiation and temperature. (c) Medians of diurnal cycles of factors at MRIU, shaded with interquartile ranges as well as the 10th and 90th percentiles.

Fig. 5 Comparisons of averaged factor diurnal patterns and factor profiles at the two sites. The left panels present the factor profiles with the IITD spectrum on top and the MRIU spectrum on bottom, color coded by the VOC families described in Section 3.1. The right panels show the box and whisker plots of diurnal cycles at the two sites, with the blue panel representing IITD and the red panel MRIU.

Fig. 6 (a, c) Diurnal patterns of factor mixing ratios at the two sites, color-coded by the six retrieved factors. (b, d) Diurnal patterns of the fractional contributions of the factors at the two sites.

Fig. 7 Explained variations of selected ions at the two sites, stacked such the sum is the total explained variation, color coded by the six factors (a) at IITD and (b) at MRIU. Missing ions at MRIU were excluded from PMF analysis due to low SNR. The possible candidates for these ions are isoprene (C_5H_8), MVK+MACR (C_4H_6O), 2,3-butanedione ($C_4H_6O_2$), 3-methyl-3-butene-2-one (C_5H_8O), methyl methacrylate ($C_5H_8O_2$), monoterpenes ($C_{10}H_{16}$), and camphor ($C_{10}H_{16}O$), respectively.
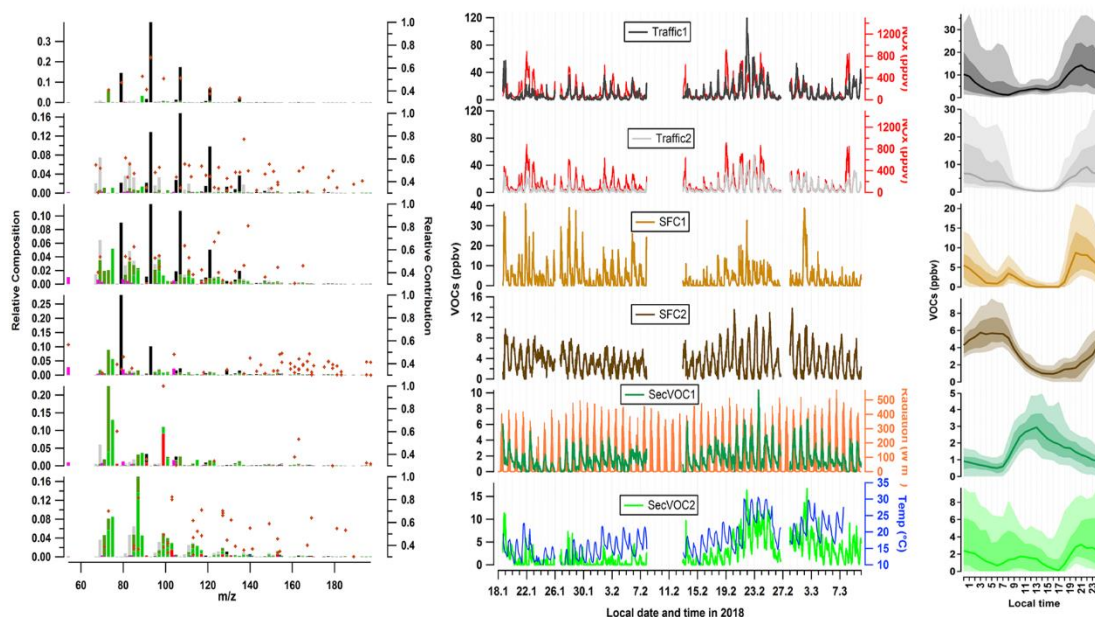
Fig. 8 Diurnal cycles of selected marker ions, stacked with explained variations by each factor, and the red line representing the measured average concentrations. Tentative structures assigned to the individual ions are denoted in each figure.

832

833

834   **Fig. 1** Maps of the study region and the two sampling sites (from Google Maps). The black circle denotes the

835   urban site at IITD and the black square the suburban site at MRIU.
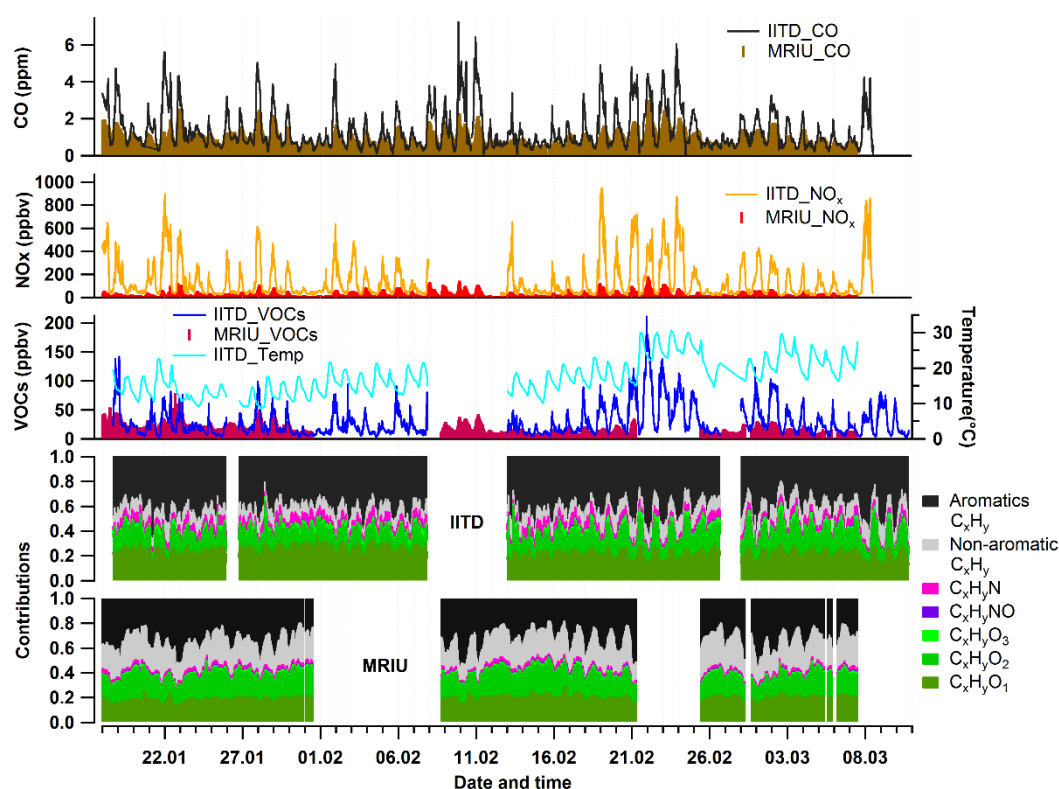
836

837   **Fig. 2** Temporal variations of CO, NO_x, analyzed VOC mixing ratios and temperature as well as the

838   contributions of seven VOC families at the two sites.

839

840

Fig. 3 PMF results at IITD, showing (a) factor profiles (b) time series and (c) diurnal patterns. In (a), the left axis for each factor profile is the relative composition of each factor (i.e., horizontal sum is 1) and the right axis is the relative contribution of each factor to a given ion. Ions are colored based on the seven family classes, as described in Section 3.1. (b) Temporal evolution of resolved factors at IITD, together with external reference data, e.g. NOx, CO, solar radiation and temperature. (c) Medians of diurnal cycles of factors at IITD, shaded with interquartile ranges as well as the 10th and 90th percentiles.


847

Fig. 4 PMF results at MRIU, showing (a) factor profiles (b) time series, and (c) diurnal patterns. (a) Relative composition (left axis) and relative contribution (right axis) of each factor to a given ion. Ions are colored based on the seven family classes, as described in Section 3.1. (b) Temporal evolution of resolved factors at MRIU, together with external reference data, e.g. NOx, CO, solar radiation and temperature. (c) Medians of diurnal

852   cycles of factors at MRIU, shaded with interquartile ranges as well as <mark>the</mark> 10th and 90th percentiles.
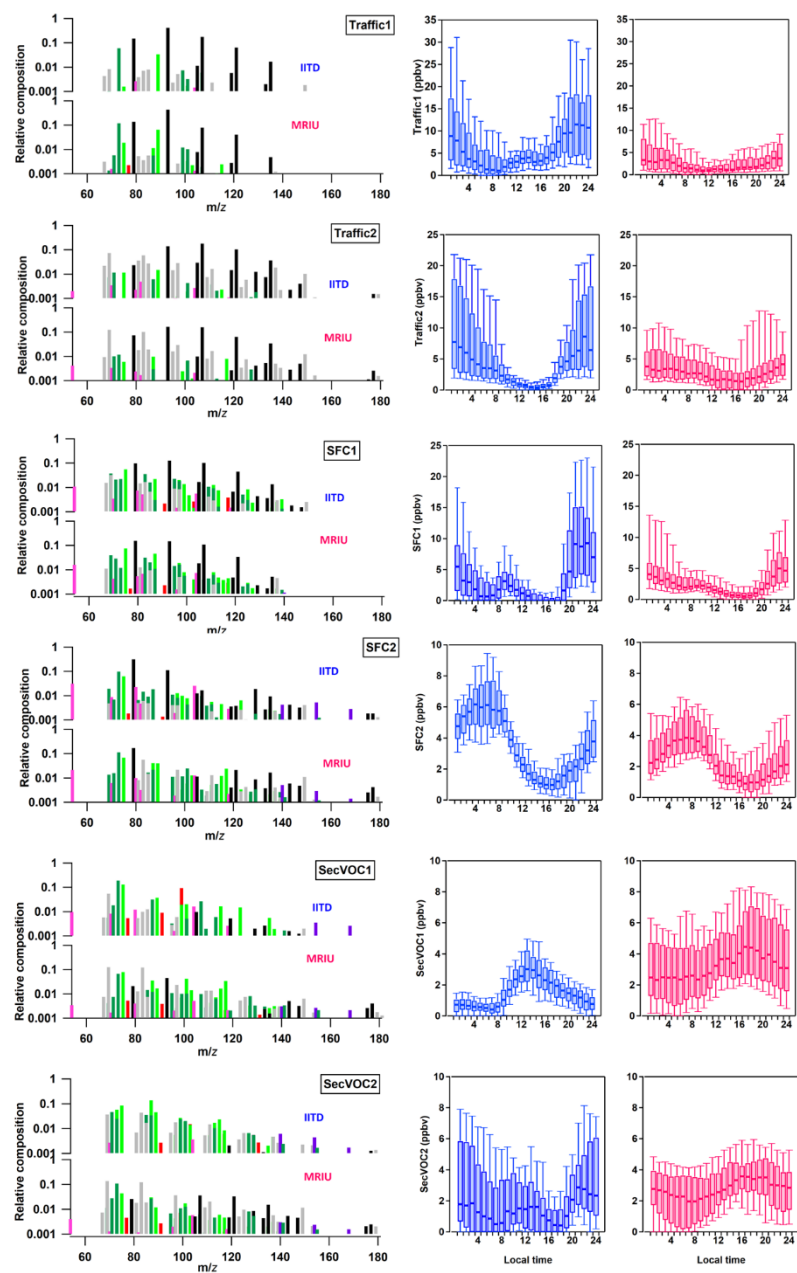
853

854   **Fig. 5** Comparisons of averaged factor diurnal patterns and factor profiles at the two sites. <mark>The left panels</mark>

855   <mark>present the factor profiles with the IITD spectrum on top and the MRIU spectrum on bottom, color coded by the</mark>

856   <mark>VOC families described in Section 3.1. The right panels show the box and whisker plots of diurnal cycles at the</mark>

857   <mark>two sites, with the blue panel representing IITD and the red panel MRIU.</mark>
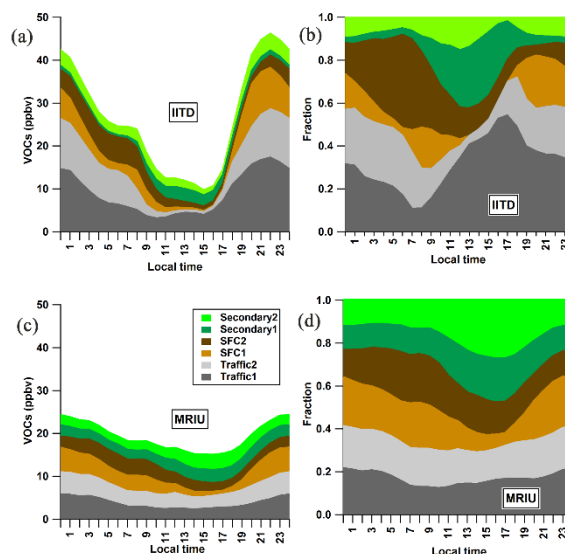
858

859

Fig. 6 (a, c) Diurnal patterns of factor mixing ratios at the two sites, color-coded by the six retrieved factors. (b, d) Diurnal patterns of the fractional contributions of the factors at the two sites.
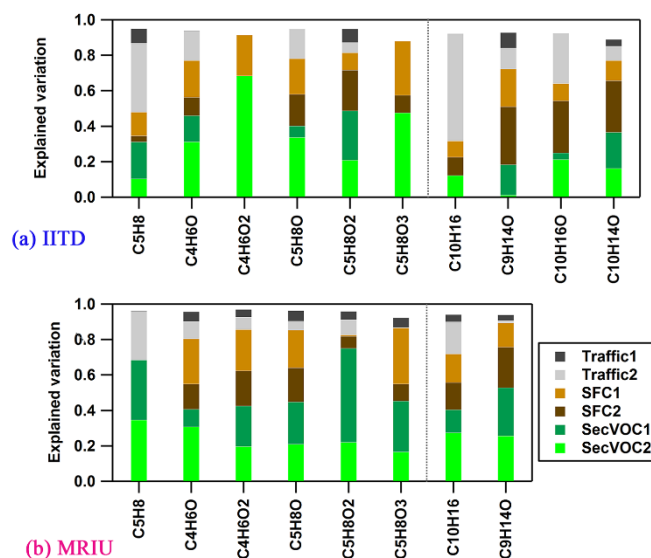

862

Explained variations of selected ions at the two sites, stacked such the sum is the total explained variation, color coded by the six factors (a) at IITD and (b) at MRIU. Missing ions at MRIU were excluded from PMF analysis due to low SNR. <mark>The possible candidates for these ions are isoprene (C_5H_8), MVK+MACR (C_4H_6O), 2,3-butanedione ($C_4H_6O_2$), 3-methyl-3-butene-2-one (C_5H_8O), methyl methacrylate ($C_5H_8O_2$), monoterpenes ($C_{10}H_{16}$), and camphor ($C_{10}H_{16}O$), respectively.</mark>
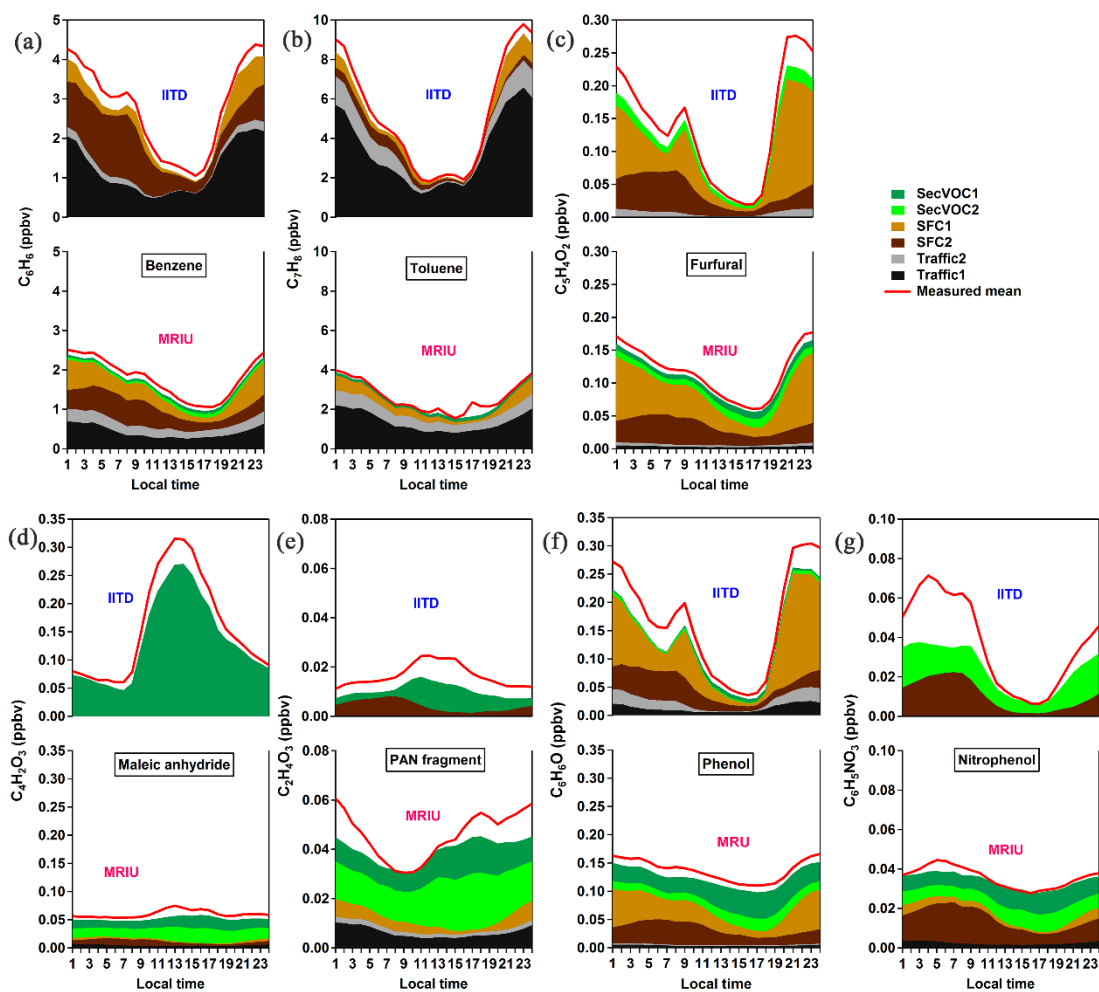
868

869 **Fig. 8** Diurnal cycles of selected marker ions, stacked with explained variations by each factor, and the red line

870 representing the measured average concentrations. Tentative structures assigned to the individual ions are

871 denoted in each figure.