VOC source–receptor relationships in Houston during TexAQS-II

Michael Leuchner [1], Bernhard Rappenglück[*]

Department of Earth and Atmospheric Sciences, University of Houston, 4800 Calhoun Rd, Houston, TX 77204-5007, United States

ABSTRACT

During the TRAMP field campaign in August–September 2006, C_2–C_{10} volatile organic compounds (VOCs) were measured continuously and online at the urban Moody Tower (MT) site. This dataset was compared to corresponding VOC data sets obtained at six sites located in the highly industrialized Houston Ship Channel area (HSC). Receptor modeling was performed by positive matrix factorization (PMF) at all sites. Conditional probability functions (CPF) were used to determine the origin of the polluted air masses in the Houston area. A subdivision into daytime and nighttime was carried out to discriminate photo-chemical influences. Eight main source categories of industrial, mobile, and biogenic emissions were identified at the urban receptor site, seven and six, respectively, at the different HSC sites. At MT natural gas/crude oil contributed most to the VOC mass (27.4%), followed by liquefied petroleum gas (16.7%), vehicular exhaust (15.3%), fuel evaporation (14.3%), and aromatics (13.4%). Also petrochemical sources from ethylene (4.7%) and propylene (3.6%) play an important role. A minor fraction of the VOC mass can be attributed to biogenic sources mainly from isoprene (4.4%). Based on PMF analyses of different wind sectors, the total VOC mass was estimated to be twofold at MT with wind directions from HSC compared to air from a typical urban sector, for petrochemical compounds more than threefold. Despite the strong impact of air masses influenced by industrial sources at HSC, still about a third of the total mass contributions at MT can be apportioned to other sources, mainly motor vehicles and aromatic solvents. The investigation of diurnal variation in combination with wind directional frequencies revealed the greatest HSC impact at the urban site during the morning, and the least during the evening.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Volatile organic compounds (VOCs) are important in the formation of secondary air pollutants such as ozone, peroxy-carboxylic nitric anhydrides, aldehydes, and secondary organic aerosols in photochemical processes. In addition, many VOC species act directly as air toxics or hazardous air pollutants. Thus, monitoring and modeling of the spatiotemporal distribution of those species, in particular in industrialized urban areas with high emission rates, is important for mitigation strategies concerning human health.

The Houston metropolitan area is a highly industrialized urban region with the largest number of petrochemical facilities within the United States, large number of other industrial, and significant mobile emission sources due to a very limited public transportation system. Extensive areas of deciduous and coniferous vegetation cover along with subtropical climate conditions make biogenic

sources another significant VOC emission source in this region. Previous studies (e.g. Jobson et al., 2004; Zhao et al., 2004; Buzcu and Fraser, 2006; Xie and Berkowitz, 2006) have shown that the VOC mixture in Houston is very complex, due to numerous different sources such as evaporation of oil, gasoline and solvents, petrochemical manufacturing and refining, and motor vehicle exhaust. In order to identify the different source categories and estimate their respective contributions, advanced multivariate receptor models have been developed and applied successfully in many air pollution studies. Out of the four most widely applied source apportionment techniques that are principal component analysis/absolute principal component scores (PCA/APCS) (e.g. Guo et al., 2004, 2006; Chan and Mozurkewich, 2007), chemical mass balance (CMB) (e.g. Na and Kim, 2007; Badol et al., 2008), UNMIX (e.g. Jorquera and Rappenglück, 2004; Olson et al., 2007), and positive matrix factorization (PMF) (e.g. Paatero, 1997; Lingwall and Christensen, 2007), the latter has been tested successfully in comparison with the other receptor models for both VOCs (Anderson et al., 2002; Miller et al., 2002) and PM (Viana et al., 2008a) and has shown to provide robust results for VOCs in the Houston industrial areas (Kim et al., 2005; Buzcu and Fraser, 2006; Xie and Berkowitz, 2006).

Within the framework of the comprehensive Texas Air Quality Study-II (TexAQS-II) which covered a wide range of simultaneous

* Corresponding author. Tel.: +1 713 743 1834; fax: +1 713 748 7906.
E-mail address: brappenglueck@uh.edu (B. Rappenglück).
[1] Present address: Fachgebiet für Ökoklimatologie, Technische Universität München, Am Hochanger 13, 85354 Freising, Germany

measurements of meteorological parameters and air chemical species in the Houston–Galveston area during August and September 2006, the TexAQS-II Radical and Aerosol Measurement Project (TRAMP) comprised the only ground-based supersite (further details see Lefer and Rappenglück, in this issue). It was located at the Moody Tower (MT) on the University of Houston campus, approximately 5 km southeast of Downtown Houston and 10 km west from the western edge of the Houston Ship Channel (HSC) (Fig. 1). The monitoring site was built on the rooftop of a 60 m tall student dormitory building without major local emission sources. The site is unique, because depending on meteorological conditions it may be affected by different air masses from urban, biogenic, and industrial sources throughout the Houston area in contrast to previous receptor modeling studies that focused exclusively on highly industrialized sites located within HSC (Buzcu-Guven and Fraser, 2008; Buzcu and Fraser, 2006; Xie and Berkowitz, 2006; Kim et al., 2005).

For the determination of different source categories, the whole set of C_2–C_{10} VOCs was measured at this urban site. In addition, VOC data from six sites within the HSC were analyzed to identify industrial source contributions.

This study focuses on a comparative analysis and apportionment of industrial, urban, and biogenic source contributions at the urban and industrial sites, with the objective of assessing the impact of HSC emissions on the urban Houston air, utilizing multivariate receptor modeling of continuous online measurements of VOCs in the summer of 2006.

2. Methods

2.1. Experimental setup

During TRAMP (Aug. 07–Sept. 30, 2006), continuous online measurements of C_2–C_{10} VOCs were performed at MT. In addition, VOC data were obtained by the TCEQ and Enhanced Industry Sponsored Monitoring auto-GC networks, which form part of the photochemical assessment monitoring stations (PAMS), within HSC where similar automated gas chromatographic (GC) systems were used. These sites included Channelview, Wallisville Road (Wallisville), Clinton Drive (Clinton), Haden Road (HRM3), Milby Park, and Lynchburg Ferry (Lynchburg) (Fig. 1). The standard methods for instrument calibration, operation, and data quality assurance at these sites follow the PAMS procedures as described by USEPA (1999).

Here, we briefly report the GC instrumentation deployed at MT: the analysis of the VOCs is performed based on developments by Veillerot et al. (1998) using gas chromatography coupled to flame ionization detectors (FID). It includes a Perkin–Elmer VOC-system consisting of a Clarus 500 GC equipped with two FIDs and two columns for multi-dimensional gas chromatography. VOCs are sampled every hour for 40 min at 15 ml min^{-1} flow through a 6 m long 1/8″ glass-lined stainless steel tube and preconcentrated on a Peltier-cooled cold trap at −30 °C. Subsequently the sample is desorbed at 325 °C in a backflush mode by a Turbomatrix 650 Automatic Thermal Desorber and injected into the GC with high purity helium (99.995%) as the carrier gas. Prior to preconcentration, water is removed through a Nafion® dryer. For separation two columns run in series are used: high molecular weight VOCs are separated by a BP-1 column (50 m × 0.22 mm I.D.; methyl silicone, SGE International, Australia), low molecular weight VOCs are transferred into a PLOT fused silica column (50 m × 0.32 mm I.D.; AL_2O_3/Na_2SO_4, Varian, USA). Both are placed in the same oven with a 47.8 min multiramp temperature program: 46 °C (15 min hold), 5 °C min^{-1} to 170 °C, 15 °C min^{-1} to 200 °C (6 min hold). FIDs are kept at 250 °C and supplied with high purity hydrogen (PG-H$_2$ 250, Perkin–Elmer, USA) and synthetic air (TOC gas generator 1250, Parker Balston, USA).

The MT GC system is calibrated using a 30 component EU Directive ozone precursor mixture provided by the UK National Physical Laboratory (NPL, UK). This certified mixture contains
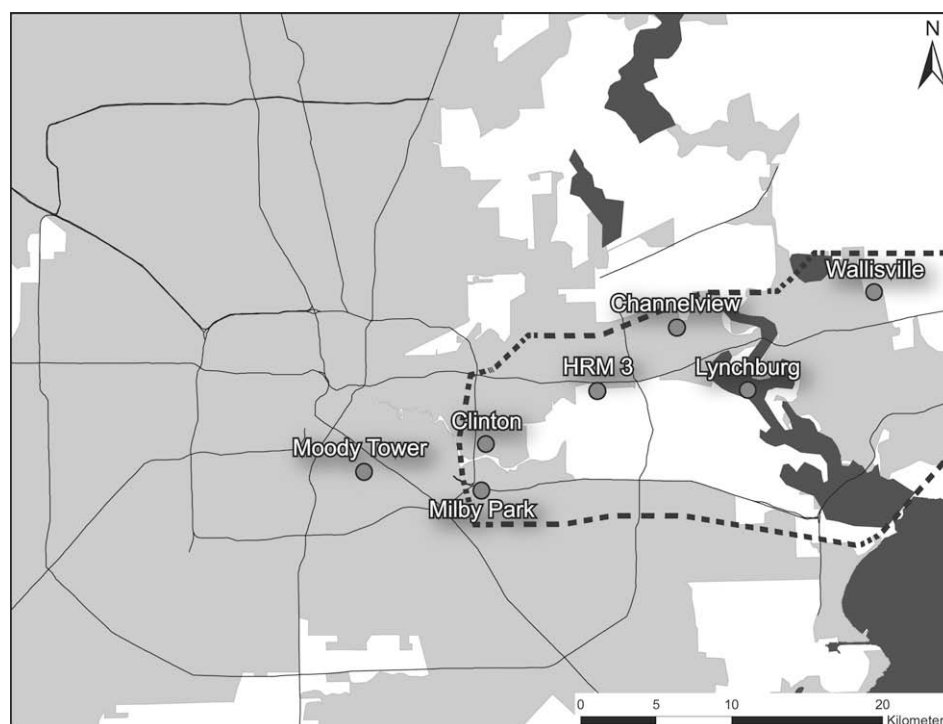

Fig. 1. Locations of the investigated sites in the Houston area during the TexAQS-II campaign. The dashed line shows roughly the extended industrial area surrounding the Houston Ship Channel.

VOCs in mixing ratios of about 4 ppbv and has been found to be very stable over long time periods. Uncertainties are stated to be below 2% (95% confidence limit). The NPL standard is used to verify the stability of the linear carbon response of the GC. A deviation of $\pm 5\%$ from the average carbon response is acceptable with the exception of acetylene which has shown systematically larger deviations. For identification purposes a 72 VOC standard in the range of 0.2–11.3 ppbv (stated uncertainty below 5%) provided by the National Center for Atmospheric Research (NCAR, USA) is used. This standard allows optimum identification and has been successfully used in recent intercomparisons (Rappenglück et al., 2006). Blank samples are generated using a "dry compressed air" gas cylinder from Matheson-Trigas equipped with a Supelco Supelcarb HC Hydrocarbon Trap (volume 120 cm^3). Detection limits (DL) for a sample volume of 600 ml are 5 pptv for compounds $>C_7$ and 10 pptv for other compounds with exception of 2,4-dimethylpentane, 2,3-dimethyl-2-pentene (DL: 10 pptv), and acetylene (DL: 20 pptv). The estimated overall accuracy was $\pm 20\%$ and the precision depending on the specific hydrocarbon compound ± 5 to $\pm 15\%$ and about $\pm 20\%$ for acetylene. DLs for the PAMS sites were stated 0.4 ppbC for all measured species.

67 VOCs were quantified at MT, 49 at Channelview, Wallisville, HRM3, and Lynchburg, 47 at Clinton, and 46 at Milby Park.

2.2. Positive matrix factorization (PMF) model description

In this study the EPA PMF 1.1 receptor model based on the multilinear engine ME-2 (Paatero, 1999) was applied. A detailed description of the PMF method can be found in Paatero and Tapper (1994) and Paatero (1997). Here, we briefly provide a short and simplified overview of PMF.

The basic objectives of PMF are the determination of the number of VOC source factors, the chemical composition profile of each factor, the contribution of each factor to each sample, and residuals out of a data matrix that consists of mixing ratio measurements of a certain number of chemical species in a certain number of samples. The minimized sum of the squared residuals that is weighted inversely by the variation of data points is determined. Some natural physical constraints must be considered: the original data must be reproduced by the model, the predicted source compositions and contributions must be non-negative, and the sum of the predicted mass contributions for each source must be less than or equal to the total measured mass for each compound.

There are several other receptor models besides PMF. The most commonly used are UNMIX, PCA/APCS, and CMB. Each approach has advantages and limitations as described in several comparison studies (e.g. Willis, 2000; Anderson et al., 2002; Miller et al., 2002; Song et al., 2008; Viana et al., 2008a). In general, the PMF method showed good agreement with the UNMIX model (e.g. Anderson et al., 2002; Jorquera and Rappenglück, 2004) and performed very well in comparison to CMB and PCA/APCS. In contrast to UNMIX, PMF, in particular the flexible multilinear engine, can solve multilinear problems with the possibility of implementing many kinds of constraints, e.g. individual data points can be weighted (Willis, 2000; Viana et al., 2008b), despite the increasing risk of putting additional distance between statistical model and physical reality (Willis, 2000). Miller et al. (2002) found out that in comparison to the other three model approaches, the extracted factors from the PMF analysis represented the major sources that were used to generate the simulated data most closely. The lack of the non-negativity constraint is another significant limitation of PCA/APCS and CMB (Anderson et al., 2002).

In this work, PCA and UNMIX models were also applied to the data set. The source profiles obtained by the PMF were the physically most reasonable results. Additionally, the modeled profile compositions show good agreement with canister samples taken within the Houston area representing crude oil handling, traffic, and an oak forest (profiles not shown).

In general, the PMF model provides robust results for the identification of sources in complex atmospheric environments, e.g. Houston, which was shown in previous studies (Zhao et al., 2004; Kim et al., 2005; Buzcu and Fraser, 2006; Xie and Berkowitz, 2006).

2.3. PMF implementation

Out of the 67 compounds at MT, eight had to be excluded from the dataset due to coelution problems. Another 21 that showed a significantly smaller signal/noise ratio or had >20% of missing values or values below detection limit, were downweighted by a factor three. The code in Table 1 shows the status of the different species in the PMF. The selection of the appropriate number of factors was based on standard criteria such as the value of the minimized sum of squared residuals, residual distribution, and the correspondence to realistic physical phenomena. Missing values were substituted by the median as described by Brown et al. (2007) and their respective uncertainties were increased accordingly. The obtained results were confirmed and supported by estimating the PMF model's uncertainties using a bootstrap technique with 200 runs combined with a method to account for the rotational freedom in the solution.

Corresponding procedures were applied to the HSC data. The following numbers of compounds were included into the PMF analyses: Clinton 47, Milby Park 44, Channelview 43, Wallisville 47, HRM3 45, Lynchburg 39.

Uncertainties for the HSC sites were set to a constant value of 20% as described by Buzcu and Fraser (2006). For the MT data, uncertainties between 5 and 30% were applied, according to the signal/noise ratio. In general, lighter compounds and species with generally higher mixing ratios were associated with smaller uncertainties. For comparison, other methods of handling uncertainties, missing values, and values below detection limit were also applied according to Jorquera and Rappenglück (2004) and Xie and Berkowitz (2006), with no significant effect on the results of the PMF.

In addition to PMF of the whole dataset, temporal and spatial subsets were distinguished. In order to investigate influences of photochemical processes on the PMF results, a classification between daytime (06:00–21:00 h CDT) and nighttime (21:00–06:00 h CDT) was performed. Another analysis with the objective to determine the impact of HSC emissions onto the Houston urban air mass compared different wind sectors, namely a sector that can be assigned primarily to industrial contributions from the HSC area (wind directions NNE to E, 11.25–101.25°) and a sector with generally higher fractions of urban sources (W to N, 270–360°).

2.4. Conditional probability functions

The conditional probability function (CPF) analysis is a well documented tool to identify physical locations associated with different source categories without applying back trajectory analysis (Ashbaugh et al., 1985; Xie and Berkowitz, 2006; Buzcu-Guven and Fraser, 2008). The CPF is derived from directional surface wind observations and factorial time series data from the PMF. The CPF is the fraction of total samples in wind sector $\Delta\theta$ that have a higher source contribution than a threshold criterion value and defined as $CPF_{\Delta\theta} = m_{\Delta\theta}/n_{\Delta\theta}$, where $m_{\Delta\theta}$ is the number of

Table 1
Statistical parameters and PMF status code for inclusion of species in the PMF analysis.

Index	VOC species name	Max[a]	Min[b]	Median[c]	Mean[d]	N[e]	Nbdl[f]	Code[g]
1	ethane	89.95	1.06	9.64	12.41	1111	0	0
2	ethylene	41.89	0.29	2.59	4.20	1111	0	0
3	propane	182.17	0.73	10.18	16.18	1111	0	0
4	propylene	85.29	0.08	1.79	3.35	1111	0	0
5	isobutane	472.11	0.20	4.95	13.32	1111	0	0
6	*n*-butane	216.29	0.36	4.73	9.22	1111	0	0
7	acetylene	8.55	0.06	1.00	1.33	1111	0	0
8	*trans*-2-butene	1.12	0.05	0.20	0.24	1111	0	1
9	1-butene	9.36	0.07	0.33	0.60	1111	0	0
10	isobutene	7.43	0.18	0.87	1.10	1111	0	1
11	*cis*-2-butene	1.03	bdl	0.11	0.15	1111	33	0
12	cyclopentane	5.32	bdl	0.50	0.69	1111	5	0
13	isopentane	77.28	0.72	5.87	8.69	1111	0	0
14	*n*-pentane	54.04	0.31	3.10	4.61	1109	0	0
15	1,3-butadiene	52.84	bdl	0.27	0.69	1109	1	2
16	2-methyl-2-butene & cyclopentene	1.55	bdl	0.11	0.16	1111	157	0
17	*trans*-2-pentene	2.17	bdl	0.20	0.29	1111	13	0
18	2-methyl-1-butene	2.58	0.06	0.26	0.36	1111	0	0
19	1-pentene	3.31	bdl	0.23	0.28	1111	3	0
20	3-methyl-1-butene	1.26	bdl	0.13	0.18	1111	74	0
21	*cis*-2-pentene	1.37	bdl	0.10	0.17	1111	216	0
22	2,2-dimethylbutane	3.06	bdl	0.27	0.37	1111	63	1
23	2,3-dimethylbutane	4.34	bdl	0.52	0.68	1111	4	0
24	2-methylpentane	21.76	bdl	1.79	2.33	1111	1	0
25	3-methylpentane	21.78	bdl	1.23	1.65	1111	1	0
26	isoprene	14.51	bdl	1.74	2.68	1111	21	0
27	*n*-hexane	41.21	0.21	2.06	2.82	1111	0	0
28	*cis*-3-hexene	0.56	bdl	bdl	0.10	1111	978	2
29	*trans*-2-hexene	0.45	bdl	bdl	0.10	1111	615	2
30	2-methyl-1-pentene	0.44	bdl	0.09	0.11	1111	196	2
31	*cis*-2-hexene	2.47	bdl	bdl	0.23	1111	994	2
32	methylcyclopentane	21.17	0.08	0.96	1.33	1111	0	1
33	2,4-dimethylpentane	5.41	bdl	0.20	0.34	1111	112	1
34	benzene	69.36	bdl	1.54	2.01	1104	1	0
35	cyclohexane	6.68	bdl	0.43	0.72	1111	23	0
36	2-methylhexane & 2,3-dimethylpentane	13.75	0.04	1.33	1.63	1110	0	1
37	3-methylhexane	9.41	bdl	1.32	1.57	1111	2	0
38	trichloroethylene	0.44	bdl	0.08	0.09	1006	4	2
39	2,2,4-trimethylpentane & 1-heptene	11.49	0.31	1.98	2.45	1111	0	0
40	*n*-heptane	11.45	0.12	1.03	1.25	1111	0	0
41	2,3-dimethyl-2-pentene	16.69	bdl	0.18	0.24	1111	18	2
42	methylcyclohexane	22.17	0.04	1.11	1.53	926	0	2
43	2,3,4-trimethylpentane	3.29	0.06	0.47	0.61	1111	0	0
44	toluene	28.30	0.57	4.60	5.57	1111	0	0
45	2-methylheptane & 4-methylheptane	3.12	bdl	0.23	0.34	1068	12	1
46	3-methylheptane	2.28	bdl	0.31	0.37	1111	5	0
47	*n*-octane	5.77	0.06	0.34	0.47	1111	0	0
48	ethylbenzene	5.96	bdl	1.16	1.31	1111	1	0
49	*m*- & *p*-xylene	18.94	0.35	3.72	4.22	1111	0	0
50	styrene & unknown	13.88	bdl	2.24	2.37	1111	1	1
51	*o*-xylene	6.26	bdl	1.42	1.58	1111	1	0
52	*n*-nonane	13.34	0.07	0.54	0.87	1109	0	1
53	isopropylbenzene	2.23	bdl	0.12	0.18	1111	177	1
54	*n*-propylbenzene	1.11	bdl	0.23	0.27	1111	51	1
55	*m*-ethyltoluene	5.24	bdl	0.43	0.72	1111	225	1
56	1,3,5-trimethylbenzene	7.01	bdl	0.53	0.74	1111	12	1
57	1,2,4-trimethylbenzene & tertbutylbenzene	11.59	bdl	1.43	1.79	1111	21	1
58	*m*-diethylbenzene	2.06	bdl	0.28	0.34	1111	70	1
59	*p*-diethylbenzene	6.34	bdl	0.39	0.57	1111	41	1
60	*n*-butylbenzene & *o*-diethylbenzene	3.69	bdl	0.39	0.46	1111	16	1

[a] Maximum of mixing ratios in ppbC.
[b] Minimum of mixing ratios in ppbC (*bdl* = below detection limit).
[c] Median of mixing ratios above detection limit in ppbC.
[d] Average of mixing ratios above detection limit in ppbC.
[e] Number of valid hourly VOC samples.
[f] Number of samples below detection limit.
[g] Code for inclusion of species in PMF analysis: 0 = normal; 1 = weak with three times downweighting; 2 = exclusion.

samples in a specific wind sector ($\Delta\theta = 22.5°$) that exceeded the criterion value and $n_{\Delta\theta}$ is the total number of observations in the same wind sector. The criterion value was defined as the 75th percentile of all data to ensure that only 'high' mixing ratios are used for the determination of source directionality. Calm winds were excluded from the analysis (<1 ms^{-1}). The CPF analysis assumes a rather straight path of the air parcel from the source to the receptor site.
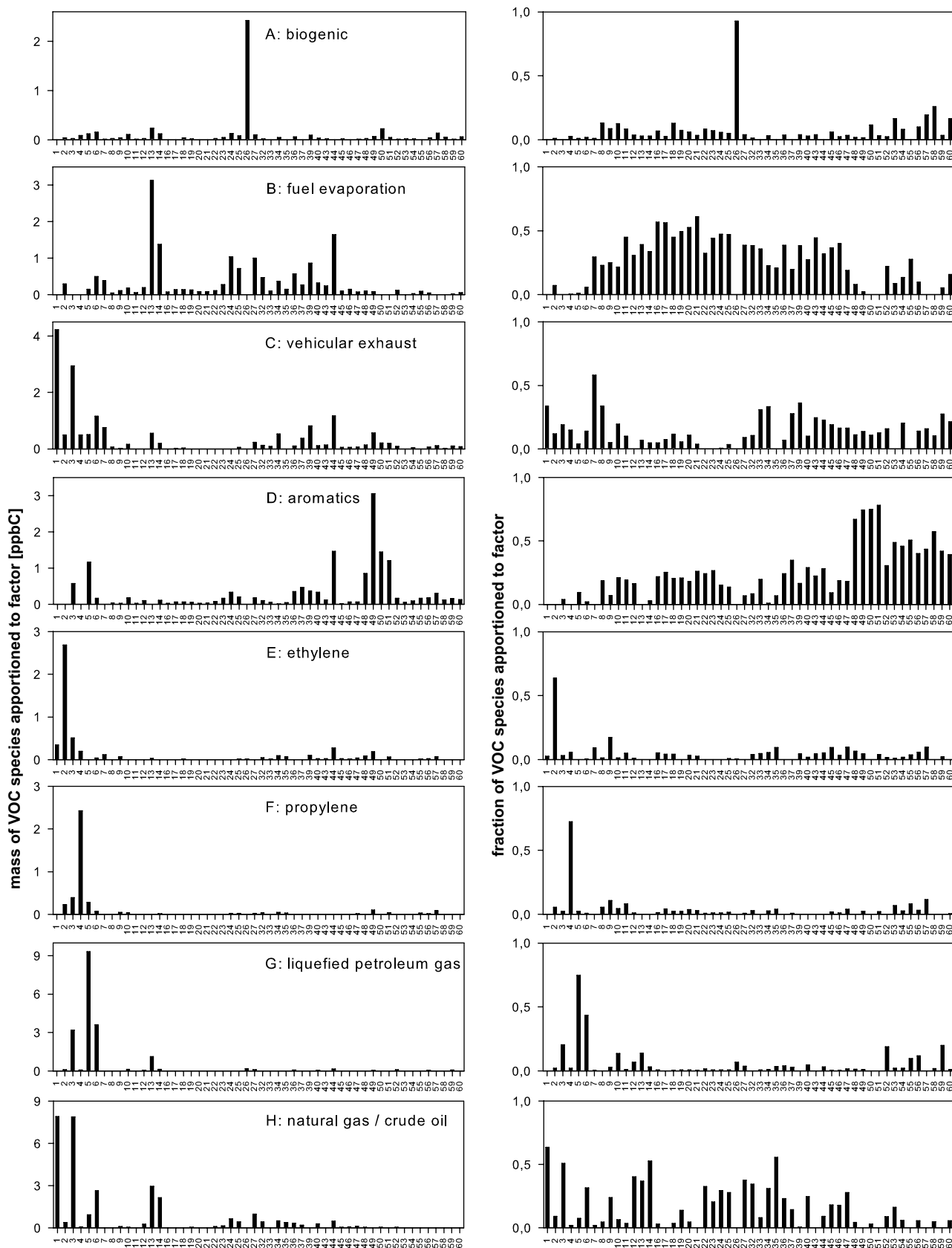
Fig. 2. Source composition profiles for the eight factors resolved by the PMF analysis at the urban receptor site MT: mass [pptC] of each species apportioned to each factor (left); relative fraction of each species to the overall species mass (right). The numbers on the abscissa represent the VOC index values shown in Table 1.
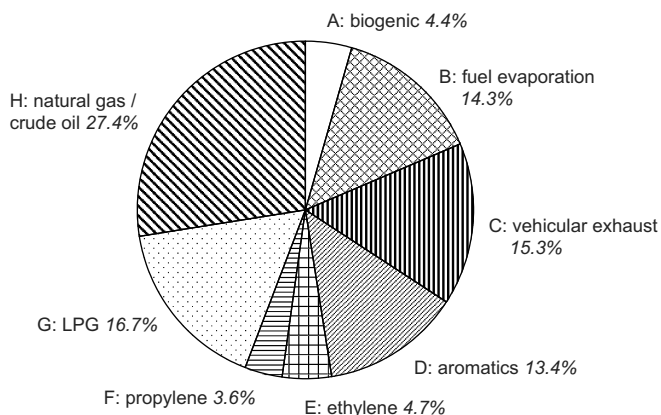
Fig. 3. Source contributions to the VOC mass [in pptC] at the urban receptor site MT.

3. Results and discussion

3.1. Data overview at the urban receptor site

Table 1 lists the measured compounds, an index value, maximum, minimum, arithmetic mean, median, number of valid samples, number of samples below detection limit, and the code for level of inclusion into the PMF at MT. The most abundant VOCs were the alkanes ethane, propane, isobutane, n-butane, and isopentane with median mixing ratios of 9.64, 10.18, 4.95, 4.73, and 5.87 ppbC, respectively. Among alkenes ethylene (2.59 ppbC), propylene (1.79), and isoprene (1.74), among aromatics toluene (4.60) and m-&p-xylene (3.72) showed the highest medians.

3.2. Source apportionment at the urban receptor site

3.2.1. Source profiles identified by the PMF analysis

The PMF conducted for MT resolved eight factors. Fig. 2 displays the source composition profiles determined by the receptor model in mass of the respective VOC species apportioned to each factor and relative fractions of a species to the overall mass of the respective species. Out of the determined factors, only one (profile A) can be attributed to biogenic sources and consists mainly of isoprene. About 93% of the total isoprene accounting for 2.42 ppbC is apportioned to this factor. Very small quantities of other compounds add to the total factor mass.

All other factors can be ascribed to anthropogenic sources. The composition of profile B with isopentane (3.13 ppbC), n-pentane (1.37), and toluene (1.64) of which 39%, 34%, and 31% of their total masses are explained, indicates fuel evaporation as the main source. Over half of the total mass of some alkenes is contained by this factor, mainly cis-2-pentene (61%), trans-2-pentene (56%), 2-methyl-2-pentene & cyclopentene (57%), and 3-methyl-1-butene (53%).

Most of the acetylene mass (58%, 0.75 ppbC) is explained by profile C. With significant contributions of further compounds typically associated with vehicular combustion such as toluene (23%, 1.17), m-&p-xylene (14%, 0.56), other aromatics, and alkenes as well as ethane (34%, 4.22) and propane (19%, 2.94), it can be attributed to vehicular exhaust. This is confirmed by a comparison with typical VOC ratios determined for vehicular exhaust. The mean ratio of toluene/benzene in factor C is 2.20, while Jobson et al. (2004) observe a ratio of 2.00 for vehicular emissions in the Houston Washburn Tunnel, and 1.59 for averaged vehicular exhaust data from several other tunnel studies. The same good agreement stands for important ratios: o-xylene/m-&p-xylene 0.35 (Washburn 0.38, tunnel average 0.38), ethylbenzene/toluene 0.12 (0.16, 0.17), isopentane/n-pentane 2.90 (3.25, 2.97).
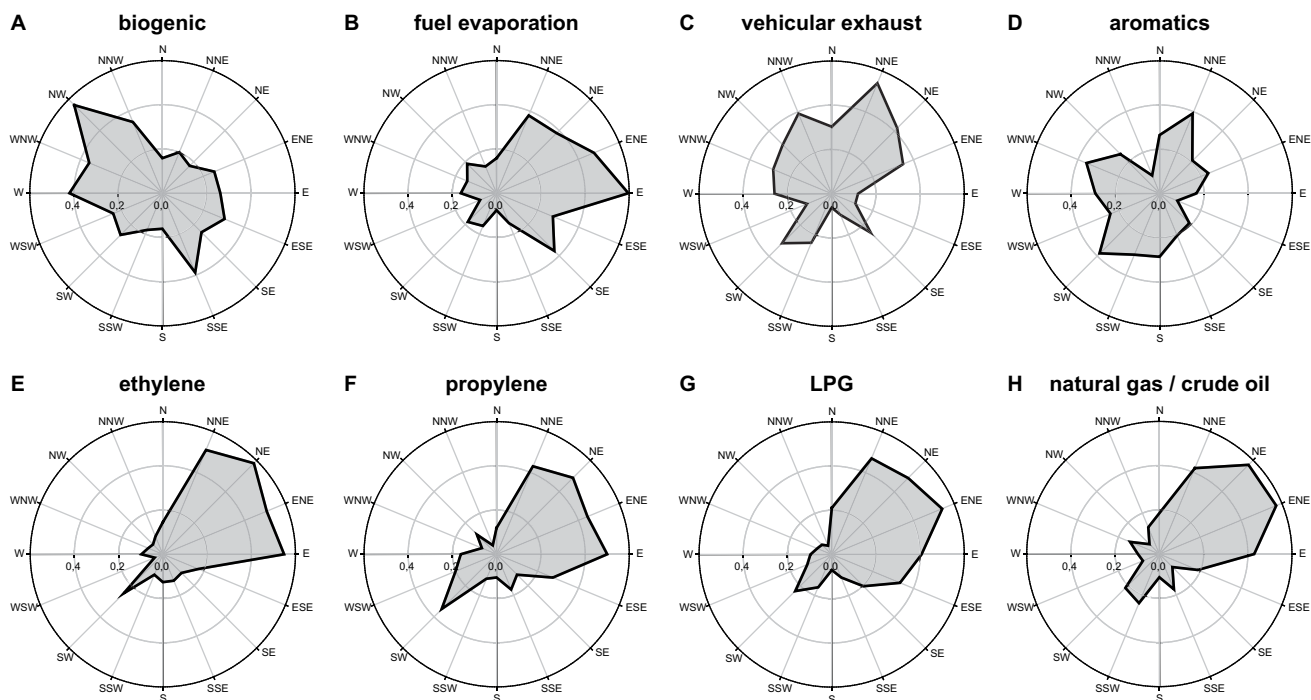

Fig. 4. CPF plots of the PMF source profiles at the MT site, showing the wind directional frequency of mixing ratios above the 75th percentile of all data. Factors B, E, F, G, and H show strong contributions from directions associated with the HSC area.

Profile D consists mainly of aromatics and smaller amounts of some alkanes and alkenes, most of them found in industrial solvent and painting sources. The most abundant species are *m*-&*p*-xylene (74%, 3.05) and toluene (28%, 1.46). *O*-xylene (78%, 1.20), styrene & unknown (75%, 1.44), and ethylbenzene (67%, 0.85) have a fraction of more than 60% of their mass contained in this factor.

Profiles E and F are each mostly represented by a single compound: ethylene and propylene, respectively. Both compounds can be associated with petrochemistry, but are obviously related to different processes in these facilities. Ethylene contributes 2.69 ppbC (64%). Besides ethylene only small amounts of 1-butene (17%, 0.07), propane (3%, 0.51), and ethane (3%, 0.34) are present.

Factor F contains 2.42 ppbC or 72% of the total propylene mass. Propane (3%, 0.39), isobutane (2%, 0.28), and ethylene (5%, 0.23) also appear in this profile.

The compounds isobutane (75%, 9.33), *n*-butane (43%, 3.62), and propane (21%, 3.18) primarily assemble G. The combination of these species is typically found in liquefied petroleum gas (LPG). Other contributing VOCs are isopentane (14%, 1.12) and isobutene (14%, 0.12).

Natural gas and evaporation, in particular from crude oil and refining compounds, represent H with the main alkane constituents ethane (64%, 7.90), propane (51%, 7.87), *n*-butane (32%, 2.63),

isobutane (7%, 0.92), *n*-pentane (53%, 2.15), and isopentane (37%, 2.95). Smaller amounts come from cyclohexane (56%, 0.38), *n*-hexane (37%, 0.97), and cyclopentane (40%, 0.25).

3.2.2. Mass contributions at the urban receptor site

Fig. 3 shows the contributions of the different factors to the total measured VOC mass at MT. The most abundant sources representing 27.4% of the VOC mass are natural gas/crude oil (H). Industrial LPG emissions account for 16.7% (G), vehicular exhaust for 15.3% (C), fuel evaporation for 14.3% (B), aromatics for 13.4% (D), and petrochemical emissions of ethylene (E) and propylene (F) for 4.7% and 3.6%, respectively. Biogenic emissions (A) play only a minor role at the urban site with 4.4% of the total VOC mass, despite the abundance of large areas of photosynthetically active biomass in the Houston area. However, during the campaign isoprene mixing ratios already tended to decrease in September following the seasonal trend of solar radiation.

3.2.3. Determination of the source directionality with conditional probability functions at the urban receptor site

Using CPF analysis, certain physical locations for the different source categories can be identified. Fig. 4 exhibits the eight CPF plots associated with the respective factors. The five factors that by
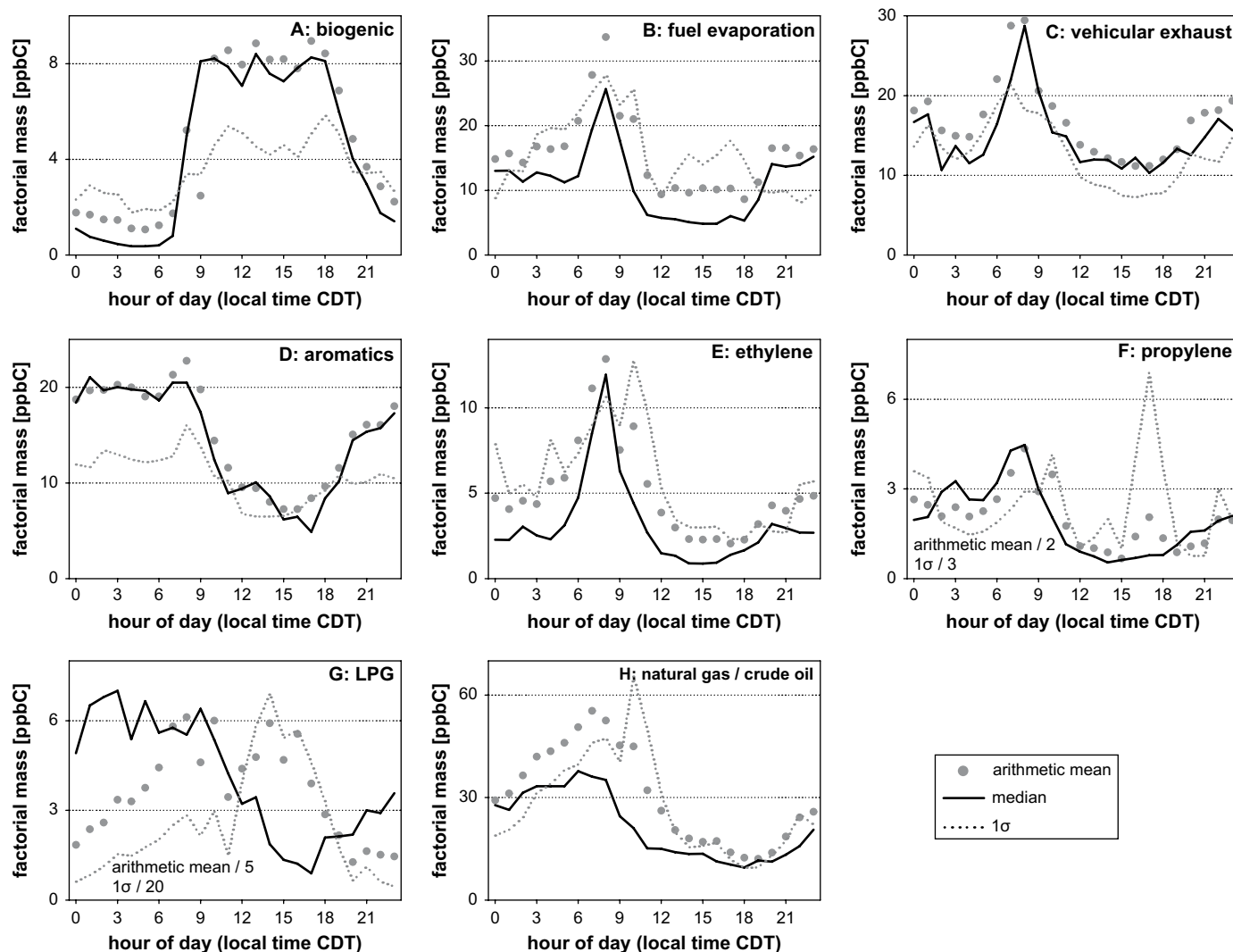

Fig. 5. Diurnal variation of the median, mean, and standard deviation of the factor mass [ppbC] of the respective source category.

their profile composition can be associated clearly with petrochemical, refining, and evaporative sources (B, E–H), show significantly higher contributions with wind directions pointing towards easterly and northeasterly regions, identifying the HSC area as their principal source. The overall contribution of these factors is 66.7% of the total VOC mass at the urban site. Other sources, e.g. mobile and urban/industrial do not constitute such a strong directional dependency, however, the factor mainly associated with vehicular combustion indicates elevated values from all northern directions, pointing towards major freeways and the Downtown Houston area. The aromatic factor displays a very heterogeneous directional pattern with highest values from SW, WNW, and NNE. The lowest contributions can be found from ESE and NNW. Biogenic sources show higher concentrations with wind correlated to northwesterly directions potentially pointing to Memorial Park, the largest park area within the Inner Loop area of Houston. Most factors (C–H) contain another smaller peak of high contributions from southwestern directions indicating other industrial sources.

3.2.4. Diurnal variation of different source categories

Apart from emissions, ambient VOC concentrations are largely determined by photochemistry and dilution processes, in particular boundary layer variations in the course of a day. Compared to these processes, VOC deposition plays a minor role, at least in urban areas. Fig. 5 exhibits hourly statistical parameters of the source categories at MT. The factors show a prominent peak in the morning hours (06:00–10:00 h CDT) when traffic rush hour and industrial emissions accumulate in the shallow boundary layer. Photochemical processes and a growing boundary layer lead to decreasing and minimal values in the late afternoon. During

Table 2
Ratios of different results from PMF analyses with subsets of data: (left) ratios between daytime (06:00–21:00 h CDT) and nighttime (21:00–06:00 h CDT), and (right) HSC sector (11.25–101.25°) and an urban sector (270–360°).

Source category	Ratio day/night	Ratio HSC/urban
A: biogenic	1.91	1.72
B: fuel evaporation	1.15	1.58
C: vehicular exhaust	0.77	1.99
D: aromatics	0.99	1.80
E: ethylene	1.59	*n.a.*
F: propylene	0.72	*n.a.*
E/F: petrochemical	*n.a.*	3.33
G: LPG	1.51	2.01
H: natural gas/crude oil	0.83	2.18

nighttime, values are increasing due to the lack of photochemical degradation and decreasing boundary layer height with additional traffic and industrial emissions. Elevated medians at night are found for aromatics, LPG, and natural gas/crude oil. The standard deviations follow the course of the means in general, for some periods a drastic increase that is accompanied by higher means in comparison to the medians reveals the impact of certain plumes with very high concentrations. For example, LPG is influenced essentially by those plumes in particular in the afternoon, but also propylene, natural gas/crude oil, and ethylene. Basically all plots experience higher means than medians as an indicator for plumes.

The biogenic factor has a diurnal behavior that mostly reflects the stomatal condition of the plants. These emissions depend mainly on solar radiation and ambient temperatures, the mixing
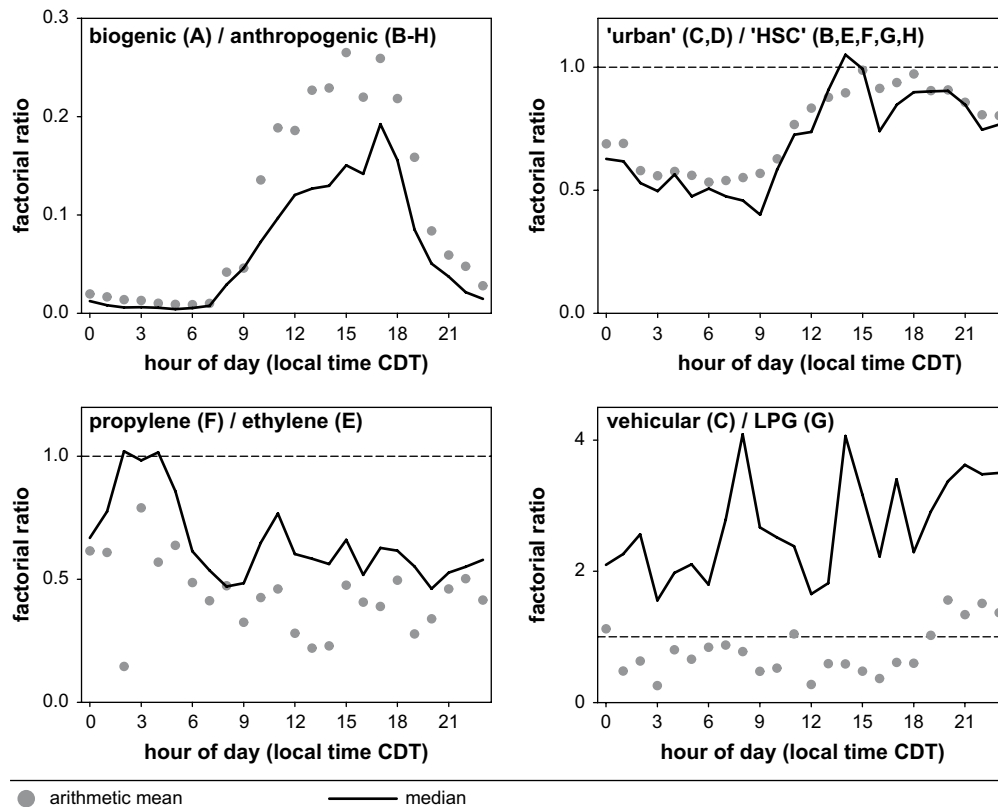
Fig. 6. Diurnal variation of the median and mean of the factorial mass ratios at MT. Four ratios of mass contributions of different factors and combinations of factors are plotted: the biogenic factor (A) in relation to the sum of all anthropogenic factors (B–H) (upper left); the sum of the vehicular factor (C) and the aromatic factor (D) referred to as 'urban' due to the distribution of the conditional probability function (Fig. 4), in relation to the five factors (B, E–H) associated with processes found mainly in the HSC (upper right); propylene (F) in relation to ethylene (E) (lower left); vehicular exhaust (C) in relation to LPG (G) (lower right).

ratios additionally on the diurnal variability of the mixing layer height and the photochemical degradation of the compounds. Thus, the highest values can be observed during the day and the lowest during the night.

The diurnal variability of selected factorial ratios is displayed in Fig. 6. Despite the relatively low contributions of biogenic emissions at MT, the ratio of biogenic to anthropogenic factors emphasizes the importance of this source during the day. In the afternoon, when most anthropogenically derived VOC contributions are lower, the mean ratio can be as high as 0.26. During the night, biogenic contributions are negligible. The relation of the rather urban factors C and D to the constantly higher HSC factors (B, E–H) shows even larger HSC contributions at night and in the morning. Increased contributions of E can be observed when compared with the other petrochemical factor F during the entire day. Single plumes of ethylene are expressed by a lower mean than median ratio. A large impact of plumes to the LPG contributions is emphasized by the comparison to vehicular exhaust. While the median is always above one, the mean ratio is mostly below one except in the evening.

3.2.5. Daytime and nighttime analysis

In order to further quantify possible impacts of photochemistry and boundary layer evolution, two separate PMF analyses were carried out at MT for daytime (06:00–21:00 h CDT) and nighttime only (21:00–06:00 h CDT).

In both cases eight similar factors were extracted. Table 2 lists the ratios of the different factorial mass contributions between daytime and nighttime results. The biogenic factor shows the highest ratio with almost twice the contributions to the total VOC mass during daytime which is due to the photosynthetical plant activity during the diurnal photoperiod. Ethylene and LPG exhibit elevated values during daytime as well. This can be mainly ascribed

to morning emissions in the shallow boundary layer, temperature dependent evaporative losses of LPG and significant impacts of plumes in the afternoon. Fuel evaporation and aromatics contribute equally during both periods. Despite the morning rush hour, vehicular exhaust shows higher contributions at night that can be attributed to boundary layer development and photochemical degradation during the day, which are also responsible for higher propylene and natural gas/crude oil contributions during nighttime.

3.3. Analysis at the industrial sources

In comparison to the urban receptor site MT, the selected HSC sites are located close to several different industrial emission sources within HSC. Their location E to NE of MT is also the direction associated with those air masses containing the highest conditional probabilities and median mixing ratios at MT.

For Channelview, Wallisville, Milby Park, and Lynchburg six factors were extracted, while seven were resolved at Clinton and HRM3. In the following, a brief description of the most important results from the PMF analyses at the HSC sites is given and compared to MT. Fig. 7 shows the mass contributions of the different source categories at the six HSC sites.

Most source profiles that could be extracted for the HSC sites can also be found at MT. Natural gas/crude oil emissions (H) are abundant at all sites and with the exception of Milby Park, account for most of the total measured VOC mass. The relative contributions range from 28.4% at Milby Park to 49.6% at Channelview, whereas at MT an impact of 27.4% was observable. LPG (G) could not be retrieved as a separate factor at HRM3 and Lynchburg and contributed only 7.2% at Clinton. All other sites including MT are influenced substantially by LPG (20–30%). At Wallisville, HRM3, Lynchburg, and MT two separate factors were found for
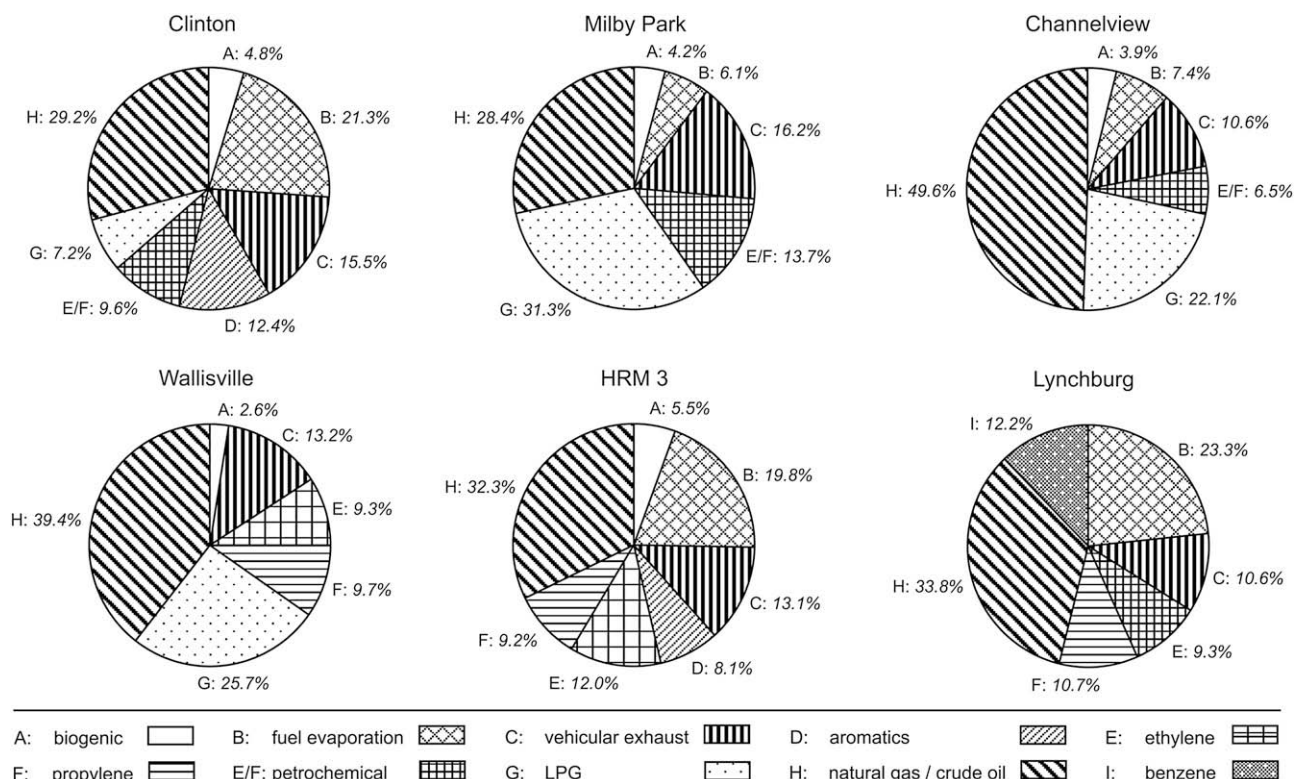

Fig. 7. Source contributions to the VOC mass [in pptC] at the six investigated industrial sites located in the HSC (Clinton, Milby Park, Channelview, Wallisville, HRM3, Lynchburg).

petrochemical processes, consisting mainly of ethylene (E) and propylene (F). At all other sites a combined factor (E/F) was resolved. The aromatic factor D was only found at Clinton, HRM3, and MT. The largely abundant vehicular emissions (C) are found at fractions of 10–16% at all sites. Fuel evaporation (B) was present everywhere except Wallisville, although the profile varies among the different sites substantially in its composition and contribution. A factor that mainly contains benzene (I) was only found at Lynchburg. Biogenic contributions (A) range from 2.6% (Wallisville) to 5.5% (HRM3). The Lynchburg site did not show A. Some fractions of A at some HSC sites are also due to smaller amounts of other constituents that indicate industrial contributions of isoprene.

Previous studies of source apportionment in HSC have found similar profiles. Xie and Berkowitz (2006) obtained six to eight factors at five HSC sites for nighttime data only that resemble very well profiles from this study (B, C, E/F, H). Kim et al. (2005) extracted seven factors at both investigated sites; six at Clinton and five at HRM3 were in good accordance to this analysis. Buzcu and Fraser (2006) focused on weekly profiles of several HSC sites, thus, a direct comparison cannot be made without restrictions. Nevertheless, similar factors were calculated. Zhao et al. (2004) used a proton transfer reaction mass spectrometer (PTR-MS) and focused on different and clumped compounds, thus a direct comparison is not possible.

3.4. Industrial, urban, and biogenic contributions to the Houston air mass

The retrieved source profiles, the CPFs at MT, and the mass contributions at all industrial sites clearly indicate the large impact of refining, crude oil and fuel handling, evaporation and other chemically and petrochemically derived processes on the air in Houston. These emissions are mostly reflected in profiles B, E, F, G, and H. This HSC fingerprint can be clearly detected in the urban area. An east–west gradient can be observed when it comes to the contributions of those sources. The highest fractions with 89.4%, 85.6%, and 84.1% were extracted at Lynchburg, Channelview, and Wallisville, respectively, that are located in the eastern parts of HSC. The sites located in the western parts Milby Park, HRM3, and Clinton show fractions of 79.5%, 73.3%, and 67.3%. At MT 66.7% were still observed. The CPFs at MT show distinctly that these factors are transported almost exclusively from the HSC accounting for this high fraction of the VOC mass.
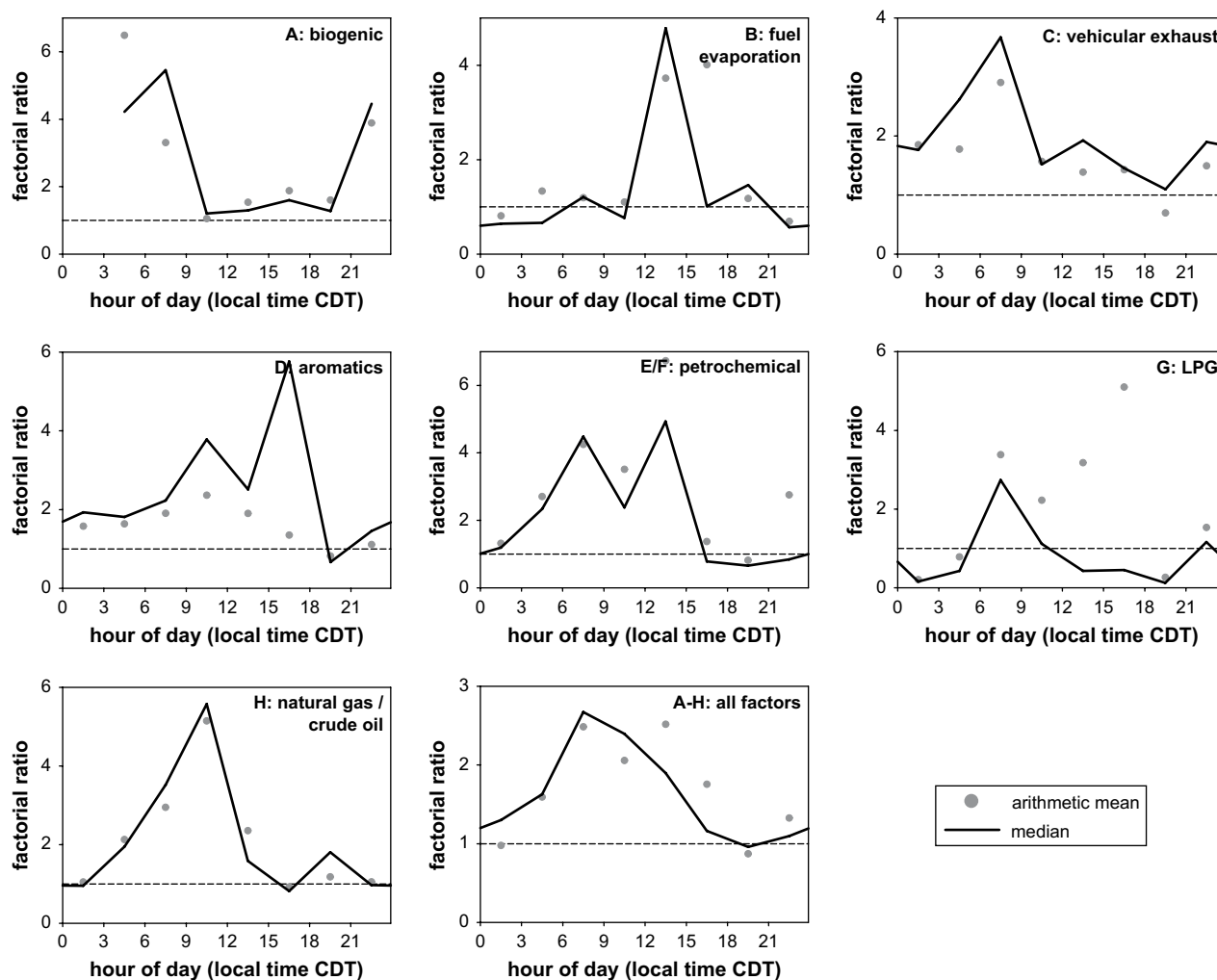
Fig. 8. Diurnal variation of the median and mean of the factorial mass ratios at MT modeled by the PMF analyses using exclusively data from selected wind sectors. The ratio is calculated by the factorial VOC mass from the HSC sector (NNE to E) divided by the VOC mass of the respective source category from the urban sector (W to N). Plot A–H shows the ratio of the total VOC contribution from the HSC sector to the total contribution from the urban sector. The biogenic contribution of the urban sector is zero from 00:00 h to 03:00 h and could thus not be represented in plot A.

The highest absolute contributions were observed at Channelview (2.41 times the average mass for a VOC at MT) and Lynchburg (1.69) followed by Clinton (1.33), Wallisville (1.26), and HRM3 (1.22). Only Milby Park faced slightly lower values than MT (0.95).

In order to quantify contributions of HSC to urban air masses, an additional model run for MT considered data with wind directions exclusively from NNE to E (11.25–101.25°). The results are compared with another PMF run with data from mostly urban wind directions. A sector from W to N (270–360°) was selected for this analysis. In 26% of the cases measured at MT, the air mass was coming from the HSC sector with some variability during the day. From early morning to noon the frequency of HSC wind directions ranged between 35% and 47%, while in the late evening and at night less than 20% (minimum 12%) arrived from these directions.

The analysis with HSC wind directions revealed the identical eight categories as identified for the entire dataset. However, the PMF with wind directions from predominantly urban sources only showed seven factors, due to the combined factor E/F. Table 2 lists the ratios of the mass contributions between the PMF results of the HSC and the urban sector at MT. All contributions are larger with wind directions from HSC. The contributions of the two petrochemical factors E and F are added up for comparison with the combined E/F factor from the urban sector. The ratio for these profiles is 3.33 indicating that most of the petrochemically related emissions originate from HSC. The other categories that are associated with chemical and petrochemical production processes also have approximately doubled contributions (natural gas/crude oil 2.18, LPG 2.01) from the HSC sector; fuel evaporation is increased by 58% with HSC wind directions. The relatively lower ratio is due to a more homogeneous distribution of fuel evaporation sources throughout the entire metropolitan area, thus reducing a specific HSC impact at MT. Despite the large influence of all directions for aromatic solvent input at MT, the HSC impact is 80% higher than from the urban sector. The almost doubled contribution of mobile sources (1.99) is not due to HSC, but rather to major freeways that are within the same sector that is analyzed for HSC emissions. The ratio for the biogenic factor is ambiguous: while the isoprene contribution to this factor is around 2.5 ppbC at both sectors, the entire source profile deriving from the HSC has larger contributions of other compounds such as the light C_2–C_4 alkanes due to some non-biogenic origin of isoprene in the HSC leading to a ratio of 1.72. On average, the VOC contributions are doubled (2.03) when air masses arrive from HSC in comparison to the urban sector.

The diurnal variation of the ratio of contributions from the HSC sector to the urban sector is exhibited in Fig. 8. The highest HSC influence can generally be observed from morning to early afternoon with some exceptions amongst the different source categories such as ratio increases in the afternoon, e.g. for fuel evaporation. Higher HSC contributions in the morning coincide well with the highest frequency of wind directions from this sector enhancing the absolute HSC impact during this time. The high mean values of the LPG ratio compared to the medians show again the great impact of single plumes from HSC in the afternoon. The biogenic ratio is rather constant between morning and evening, during the night no biogenic contributions from the urban sector were calculated.

The results confirm the major impact of the industrial HSC on the urban Houston air as observed at MT, due to increased contributions of most factors. Despite the strong influence with approximately two-thirds of the total measured VOC mass deriving from HSC air masses to the urban receptor site, still another third of the mass contributions can be apportioned to other sources, mainly motor vehicles and aromatic solvents. Biogenic emissions, despite large areas of biomass in the region, seem to play a minor role – at least during this time of the year – still contributing significantly to the total VOC mass in particular during the afternoon.

4. Conclusions

During TRAMP in August–September 2006 continuous online measurements of C_2–C_{10} VOCs were performed at the urban MT supersite and additionally at six auto-GC sites in the HSC area. A comparative analysis and apportionment of source contributions with the objective of assessing the impact of industrial HSC emissions and other anthropogenic and biogenic sources on the urban site was performed with multivariate receptor modeling. The PMF analyses indicate complex mixtures from traffic, industrial, and biogenic sources at all investigated sites.

The results show that the urban site is influenced largely by HSC emissions that account for approximately two-thirds of the VOC mass contributions, but also by mobile, evaporative, and industrial emissions from other locations in the Houston area as shown by CPF and analysis of different wind sectors. An average air parcel from HSC has about a twofold contribution to a typical urban air mass, varying greatly temporally and categorically. Diurnal analysis revealed the largest impact of the HSC to the urban air during the morning coinciding with the highest frequency of air masses coming from HSC wind directions.

References

Anderson, M.J., Daly, E.P., Miller, S.L., Milford, J.B., 2002. Source apportionment of exposures to volatile organic compounds: II. Application of receptor models to TEAM study data. Atmospheric Environment 36, 3643–3658.

Ashbaugh, L.L., Malm, W.C., Sadeh, W.Z., 1985. A residence time probability analysis of sulfur concentrations at Grand Canyon National Park. Atmospheric Environment 19, 1263–1270.

Badol, C., Locoge, N., Galloo, J.C., 2008. Using a source–receptor approach to characterise VOC behaviour in a French urban area influenced by industrial emissions. Part II: source contribution assessment using the chemical mass balance (CMB) model. Science of the Total Environment 389, 429–440.

Brown, S.G., Frankel, A., Hafner, H.R., 2007. Source apportionment of VOCs in the Los Angeles area using positive matrix factorization. Atmospheric Environment 41, 227–237.

Buzcu, B., Fraser, M.P., 2006. Source identification and apportionment of volatile organic compounds in Houston, TX. Atmospheric Environment 40, 2385–2400.

Buzcu-Guven, B., Fraser, M.P., 2008. Comparison of VOC emissions inventory data with source apportionment results for Houston, TX. Atmospheric Environment 42, 5032–5043.

Chan, T.W., Mozurkewich, M., 2007. Application of absolute principal component analysis to size distribution data: identification of particle origins. Atmospheric Chemistry and Physics 7, 887–897.

Guo, H., Wang, T., Louie, P.K.K., 2004. Source apportionment of ambient non-methane hydrocarbons in Hong Kong: application of a principle component analysis/absolute principle component scores (PCA/APCS) receptor model. Environmental Pollution 129, 489–498.

Guo, H., Wang, T., Blake, D.R., Simpson, I.J., Kwok, Y.H., Li, Y.S., 2006. 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.

Jobson, B.T., Berkowitz, C.M., Kuster, W.C., Goldan, P.D., Williams, E.J., Fehsenfeld, F.C., Apel, E.C., Karl, T., Lonneman, W.A., Riemer, D., 2004. Hydrocarbon source signatures in Houston, Texas: influence of the petrochemical industry. Journal of Geophysical Research 109 Art. No. D24305.

Jorquera, H., Rappenglück, B., 2004. Receptor modeling of ambient VOC at Santiago, Chile. Atmospheric Environment 38, 4243–4263.

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

Lefer, B., Rappenglück, B., 2008. TRAMP Overview, in this issue.

Lingwall, J.W., Christensen, W.F., 2007. Pollution source apportionment using a priori information and positive matrix factorization. Chemometrics and Intelligent Laboratory Systems 87, 281–294.

Miller, S.L., Anderson, M.J., Daly, E.P., Milford, J.B., 2002. Source apportionment of exposures to volatile organic compounds: I. Evaluation of receptor models using simulated exposure data. Atmospheric Environment 36, 3629–3641.

Na, K., Kim, Y.P., 2007. 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.

Olson, D.A., Norris, G.A., Seila, R.L., Landis, M.S., Vette, A.F., 2007. Chemical characterization of volatile organic compounds near the World Trade Center: ambient concentrations and source apportionment. Atmospheric Environment 41, 5673–5683.

Paatero, P., Tapper, U., 1994. Analysis of positive matrix factorization: a nonnegative factor model with optimal utilization of error estimates of data values. Environmetrics 5, 111–126.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems 37, 23–35.

Paatero, P., 1999. 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 1/4, 854–888.

Rappenglück, B., Apel, E., Bauerfeind, M., Bottenheim, J., Brickell, P., Čavolka, P., Cech, J., Gatti, L., Hakola, H., Honzak, J., Junek, R., Martin, D., Noone, C., Plass-Dülmer, C., Travers, D., Wang, T., 2006. The first VOC intercalibration exercise within the Global Atmosphere Watch (GAW). Atmospheric Environment 40, 7508–7527.

Song, Y., Dai, W., Shao, M., Liu, Y., Lu, S., Kuster, W., Goldan, P., 2008. Comparison of receptor models for source apportionment of volatile organic compounds in Beijing, China. Environmental Pollution 156, 174–183.

US Environmental Protection Agency (USEPA), 1999. Enhanced Ozone Monitoring (PAMS). http://www.epa.gov/air/oaqps/pams/.

Veillerot, M., Locoge, N., Galloo, J.C., Guillermo, R., 1998. Multidimensional capillary gas chromatography for the monitoring of individual non-methane hydrocarbons in air. Analysis Magazine 26, M38–M43.

Viana, M., Pandolfi, M., Minguillón, M.C., Querol, X., Alastuey, A., Monfort, E., Celades, I., 2008a. Inter-comparison of receptor models for PM source apportionment: case study in an industrial area. Atmospheric Environment 42, 3820–3832.

Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A.S.H., Hueglin, C., Bloemen, H., Wåhlin, P., Vecchi, R., Miranda, A.I., Kasper-Giebl, A., Maenhaut, W., Hitzenberger, R., 2008b. Source apportionment of particulate matter in Europe: a review of methods and results. Aerosol Science 39, 827–849.

Willis, R.D., 2000. Final Report, vol. 5. Workshop on UNMIX and PMF as Applied to PM2.5, pp. 14–16 February 2000.

Xie, Y., Berkowitz, C.M., 2006. The use of positive matrix factorization with conditional probability functions in air quality studies: an application to hydrocarbon emissions in Houston, Texas. Atmospheric Environment 40, 3070–3091.

Zhao, W., Hopke, P.K., Karl, T., 2004. Source identification of volatile organic compounds in Houston, Texas. Environmental Science and Technology 38, 1338–1347.