

Comparison of receptor models for source apportionment of volatile organic compounds in Beijing, China

Yu Song^a, Wei Dai^a, Min Shao^{b,*}, Ying Liu^b, Sihua Lu^b,
William Kuster^c, Paul Goldan^c

^a Department of Environmental Sciences, Peking University, Beijing 100871, China

^b State Joint Key Laboratory of Environmental Simulation and Pollution Control, Peking University, Beijing 100871, China

^c Chemical Sciences Division, NOAA Earth System Research Laboratory, Boulder, CO 80305, USA

Received 13 June 2007; received in revised form 28 September 2007; accepted 12 December 2007

VOCs sources were similar for three models with CMB showing a higher estimate for vehicles.

Abstract

Identifying the sources of volatile organic compounds (VOCs) is key to reducing ground-level ozone and secondary organic aerosols (SOAs). Several receptor models have been developed to apportion sources, but an intercomparison of these models had not been performed for VOCs in China. In the present study, we compared VOC sources based on chemical mass balance (CMB), UNMIX, and positive matrix factorization (PMF) models. Gasoline-related sources, petrochemical production, and liquefied petroleum gas (LPG) were identified by all three models as the major contributors, with UNMIX and PMF producing quite similar results. The contributions of gasoline-related sources and LPG estimated by the CMB model were higher, and petrochemical emissions were lower than in the UNMIX and PMF results, possibly because the VOC profiles used in the CMB model were for fresh emissions and the profiles extracted from ambient measurements by the two-factor analysis models were “aged”.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: VOC; Source apportionment; UNMIX; PMF; CMB; Beijing

1. Introduction

Many ambient volatile organic compounds (VOCs) have been found to have adverse effects on public health (1990 Clean Air Act Amendments, Section 112). Chemical reactions of VOCs with nitrogen oxides under sunlight lead to the production of secondary air pollutants, resulting in tropospheric ozone and secondary organic aerosols (SOAs) (Seinfeld and Pandis, 1998).

The air quality in Beijing is of great concern as the 2008 Olympic Games are approaching. Episodes of high surface

ozone concentrations can occur frequently there during the summer and autumn. From 1999 to 2004, the national ozone standard was exceeded on 110, 109, 77, 45, 57, and 67 days, respectively (Beijing Environmental Protection Bureau, 1999–2004). Monitoring data also indicate that PM10 pollution is a serious problem in Beijing. Moreover, recent research has found that in summer, secondary organic carbon (SOC) accounted for more than 50% of the total organic carbon in PM10 pollution in Beijing (Duan et al., 2005). Surface ozone and particulate matter abatement are needed urgently in Beijing, and because VOCs are important precursors to ozone and SOA, their sources must be identified and quantified to develop effective control measures. Although emission sources can be quantified by means of source inventory establishment, dispersion models, and receptor models, establishing an emissions inventory is often too time- and resource-consuming. Moreover, the performance of dispersion models depends

Abbreviations: 2M-propene, 2-methylpropene; 2M-1-butene, 2-methyl-1-butene; 3M-1-butene, 3-methyl-1-butene; 2M-2-butene, 2-methyl-2-butene; MTBE, methyl tert-butyl ether.

* Corresponding author. Tel.: +86 10 62757973; fax: +86 10 62751927.

E-mail address: mshao@pku.edu.cn (M. Shao).

strongly on the accuracy of the emissions inventory, and complex atmospheric motion could degrade the modeling results. Receptor models based on chemical composition, however, have been used intensively for source apportionment.

Several studies have apportioned VOC sources in Beijing using receptor models. Applying a chemical mass balance (CMB) model, both Lu (2004) and Liu et al. (2005) identified gasoline-powered vehicle exhaust and gasoline vapor as the main sources, contributing about 70% of the ambient VOCs in summer. Song et al. (2007) extracted eight sources using positive matrix factorization (PMF); gasoline-related emissions, petrochemicals, and liquefied petroleum gas (LPG) contributed 52, 20, and 11%, respectively, of the total ambient VOCs.

To better clarify VOC sources, results from multi-receptor models should be compared because each model has its own advantages and disadvantages. Factor analysis models can extract source contributions from ambient samples without source profiles. PMF, an advanced factor method, constrains all the elements in the factorized matrices to nonnegative (Hopke, 2003), but PMF can be very complicated and time-consuming. As for CMB models, the availability of accurate source profiles is a limitation. Previously established profiles may not be suitable for current use and some profiles may be region-specific (Watson et al., 2001). UNMIX, another advanced multivariate receptor model, has undergone intense development and been applied in source apportionment of VOCs and particulate matter (Hopke, 2003). UNMIX is relatively simple and easy to perform using common software packages (Henry, 2003). A few comparisons of particulate matter among different receptor models have been performed (Poirot et al., 2001; Maykut et al., 2003; Ito et al., 2004). Hopke et al. (2006) recently used six different models to analyze particulate composition data sets for Washington, DC, and Phoenix, AZ. Few studies, however, have compared models using the same VOC data set (Miller et al., 2002; Mukerjee et al., 2004).

In this study, both the UNMIX and CMB models were applied to VOC source apportionment in Beijing using the same data set applied to the PMF model by Song et al. (2007). A comparison of results from the three models may provide insight into model performance and a better understanding of VOC sources in Beijing.

2. Methods

2.1. Sampling and chemical analysis

The sampling site was on the roof of a five-story building (~20 m above ground) on the campus of Peking University (PKU) (Supplemental materials, Fig. S1). The campus is surrounded by heavily trafficked roads. To the east, the sampling site faces one main road, which is oriented north–south. Another main road is close to the site, to the south. This site is also surrounded by residential apartments and electronics manufacturers.

The concentrations of VOC species were quantified using a custom-made GC-FID/MS system with two channels. Samples were collected through a dual coaxial Teflon line system, and C2–C5 alkanes, C2–C4 alkenes, and acetylene were separated on an Al₂O₃/KCl column and quantified with a flame ionization detector. The C5–;C10 alkanes, C5–C9 alkenes, C6–C9 aromatics, C1–C5 alcohols, C2–C7 aldehydes and ketones, C1–C5 alkylnitrates, isoprene, monoterpenes, and several chlorofluorocarbons were separated on

a semipolar stationary phase column (DB624) on the second channel, and a subset of these were quantified with a linear quadrupole mass spectrometer. The detection limit for most of the compounds measured was near 0.5 pptv (signal to noise ratio of 2). More detailed descriptions of the sampling pre-concentration procedure and instrument parameters can be found in Goldan et al. (2004). Up to 45 VOC species were quantified at a time resolution of 30 min. The sampling duration was from 1 to 27 August 2005.

In this study, as in the PMF study (Song et al., 2007), 31 species and 1019 samples without any missing data were used. Table 1 lists the concentration statistics. The average mixing ratio of all 31 species was 101.8 µg m⁻³. The alkanes, alkenes, and aromatics contributed 42.1, 16.7, and 32.5%, respectively, to the total VOC mass concentration. The analytical uncertainties were estimated as 10% of the quantification for benzene, toluene, ethylbenzene, *m,p*-xylene, *o*-xylene, isoprene, α -pinene, β -pinene, and limonene, and as 15% of the measurements for the other chemical species.

2.2. PMF, UNMIX, and CMB models

The receptor model can be generally expressed in terms of the contribution from p independent sources to m chemical species in n samples as follows:

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

where x_{ij} is the measured concentration of the j th species in the i th sample, g_{ik} is the concentration from the k th source contributing to the i th sample, and f_{kj} is the j th species fraction from the k th source.

Both PMF and UNMIX are advanced multivariate receptor models that determine the number of sources and their chemical compositions and contributions without source profiles. The data in PMF are weighted by the inverse

Table 1
Measurement statistics from August 2005, Beijing (data in µg m⁻³, number of samples = 1019)

Compound	Mean	Standard deviation	Maximum	Minimum
Ethane	4.47	2.00	13.55	1.26
Ethylene	5.18	3.34	25.63	0.66
Acetylene	5.64	3.06	23.66	0.57
Propane	6.24	3.78	24.49	0.62
Propylene	1.97	1.24	11.10	0.34
2M-propene	1.29	0.87	6.62	0.19
<i>i</i> -Butane	5.36	3.17	21.97	0.46
<i>n</i> -Butane	6.36	3.82	25.61	0.50
<i>t</i> -2-Butene	0.86	0.83	7.77	0.05
1-Butene	2.21	1.57	11.08	0.17
<i>c</i> -2-Butene	0.75	0.70	7.81	0.05
2M-1-Butene	0.87	0.62	7.54	0.12
3M-1-Butene	0.13	0.06	0.53	0.02
2M-2-Butene	0.86	1.02	14.82	0.05
<i>i</i> -Pentane	11.84	7.32	83.06	1.09
<i>n</i> -Pentane	4.89	2.80	21.09	0.43
1-Pentene	0.34	0.17	1.29	0.05
<i>c</i> -2-Pentene	0.30	0.25	3.38	0.03
<i>t</i> -2-Pentene	0.51	0.43	5.14	0.04
<i>n</i> -Hexane	2.22	1.28	10.22	0.30
<i>n</i> -Decane	1.50	1.08	10.48	0.19
Benzene	5.43	2.74	18.03	0.70
Toluene	11.14	6.01	40.69	1.37
Ethylbenzene	4.08	2.41	14.37	0.37
<i>m,p</i> -Xylene	8.54	5.14	31.02	0.79
<i>o</i> -Xylene	3.91	2.34	15.26	0.42
Isoprene	1.12	0.78	4.98	0.08
α -Pinene	0.50	0.60	3.50	0.01
β -Pinene	0.08	0.08	0.74	0.01
Limonene	0.11	0.12	0.99	0.01
MTBE	3.07	1.77	13.02	0.30

of the measurement errors for each observation. Factors in PMF are constrained to be nonnegative. Then PMF incorporates error estimates of the data to solve matrix factorization as a constrained, weighted least-squares problem (Miller et al., 2002). More detail on PMF can be found in Paatero (2004).

UNMIX solves the problem using a geometrical approach. If the data consist of n observations of m species, then the data can be plotted in an m -dimensional data space, where the coordinates of a data point are the observed concentrations of the species during a sampling period. If p sources exist, the data space can be reduced to a $(p - 1)$ -dimensional space. It is assumed that for each source, some data points exist for which the contribution of the source is not present or small compared to the other sources. These are called edge points and UNMIX works by finding these points and fitting a hyperplane through them; this hyperplane is called an edge. If p sources exist, then the intersection of $p - 1$ of these edges defines a point that has only one contributing source. Thus, this point gives the source composition. In this way, we determine the compositions of the p sources from which the source contributions are calculated (Henry, 2003).

UNMIX 2.4 provided by Dr Ronald Henry was used in this study. UNMIX works fundamentally by searching for edges. To produce good edges, it is necessary to explore the data to check for influential points. These points are usually those that lie too far below or along the edge, enabling them to affect the location of the edge considerably more than the other points. The UNMIX software removed 17 samples as influential points. The software suggested excluding the marker in diesel-vehicle exhaust, n -decane, because it presented more than 50% of the variance due to error and might be detrimental to the model. We plotted all retained species against the sum of all 31 species to identify those compounds with well-defined edges (Henry, 2003). Species with good edges, along with common marker species in source characterization, were applied in the initial run to generate a realistic solution indicative of local sources. Based on the species in the initial run, we attempted to incorporate as many species as possible to provide an unbiased comparison with the PMF and CMB results.

Unlike PMF and UNMIX, CMB does not need a large sample number, but does require prior information about the sources to produce a solution. A CMB model yields a weighted least-squares solution to the above equation and produces the source apportionment. The weighting process minimizes the weighted sum of the squared differences between the measured and modeled concentrations (Watson et al., 1990). Usually, however, the sample number for VOCs (sampling by canister is still dominant; online sampling methods with high time-resolution are not yet popular) is not high enough for apportioning sources by factor analysis methods, and the CMB model is still widely used.

Although VOC source profiles for Beijing were not fully established, we tried to incorporate as many local profiles as possible in our analysis. Profiles measured elsewhere in China were used if profiles were not available for Beijing, and profiles measured outside China were used if profiles were not available from China. Finally, seven source profiles were selected for CMB in this study. The profiles of LPG and paint were measured in Beijing (Lu, 2004). China's LPG was found to be rich in butane and butene in addition to propane, whereas LPG in North America was dominated by propane (Fujita, 2001; Watson et al., 2001; Lu, 2004; Tsai et al., 2006). The petrochemical production profile was collected in the Pearl River Delta in China (Wang, 2002), as a local profile was not available. No local published profile for natural gas was found and we used a profile from the United States (Fujita, 2001) since all natural gas worldwide is of similar composition. The profiles for gasoline-powered vehicle exhaust and gasoline evaporation were from North American cities (Fujita, 2001; Watson et al., 2001) because the local profiles suffered from strong collinearity with the petrochemical profile. The biogenic profile, which was solely dominated by isoprene, was also included in the model.

Notably, chemical reactive species such as 2M-propene, 2M-1-butene, 3M-1-butene, 2M-2-butene, α -pinene, β -pinene, limonene, and MTBE were excluded from the 31 species because very little was known about their abundances in source profiles. The arithmetic averages of the observations at the same hour on different days were used in the analysis.

The EPA-CMB (version 8.2; US EPA, 2004) software was used in this study. Major performance measures for the model were χ^2 , R^2 , and percent of attributable mass (%Mass). The standard fitting criteria for a CMB model run were checked for each sample: $\chi^2 < 4$, $R^2 > 0.8$, and %Mass between 80 and 120%.

3. Results and discussion

3.1. Comparison between UNMIX and PMF

UNMIX extracted eight sources. The minimum R^2 was 0.98 for the selected species and the minimum signal to noise ratio was 3.37. After several tests, t -2-butene, c -2-butene, 3M-1-butene, 2M-2-butene, n -decane, o -xylene, α -pinene, β -pinene, and limonene were excluded because efforts to incorporate more species led to no feasible solution. Table 2 presents the

Table 2
Source profiles extracted by UNMIX (in ppbV; the values in parentheses are estimates of uncertainties; bold values denote that they are greater than twice the estimated standard deviation)

Compound	Source 1	Source 2	Source 3	Source 4	Source 5	Source 6	Source 7	Source 8
Ethane	0.077 (0.020)	0.065 (0.036)	0.003 (0.100)	0.175 (0.064)	0.291 (0.068)	-0.034 (0.069)	0.229 (0.704)	-0.746 (45.235)
Ethylene	0.137 (0.019)	0.214 (0.069)	-0.014 (0.153)	0.028 (0.043)	0.090 (0.021)	0.119 (0.024)	-0.278 (1.780)	0.581 (33.935)
Acetylene	0.144 (0.028)	0.074 (0.063)	0.120 (0.067)	0.007 (0.063)	0.232 (0.026)	0.169 (0.043)	-0.072 (1.219)	0.339 (11.842)
Propane	0.089 (0.033)	0.055 (0.047)	-0.059 (0.175)	0.281 (0.084)	0.196 (0.040)	-0.064 (0.105)	-0.057 (1.192)	-0.892 (56.033)
Propylene	0.013 (0.002)	0.051 (0.018)	0.022 (0.023)	0.013 (0.008)	0.003 (0.006)	0.030 (0.012)	0.050 (0.270)	0.344 (18.414)
2M-Propene	0.005 (0.002)	-0.006 (0.007)	0.044 (0.044)	0.000 (0.006)	-0.004 (0.004)	0.018 (0.008)	0.042 (0.236)	0.241 (9.515)
<i>i</i> -Butane	0.055 (0.014)	0.038 (0.022)	0.016 (0.048)	0.132 (0.032)	0.051 (0.008)	0.003 (0.038)	0.001 (0.434)	-0.177 (17.671)
<i>n</i> -Butane	0.071 (0.021)	0.071 (0.025)	0.031 (0.058)	0.187 (0.055)	0.078 (0.013)	-0.013 (0.064)	-0.035 (0.690)	-0.571 (35.222)
1-Butene	0.006 (0.005)	0.000 (0.011)	0.036 (0.030)	0.030 (0.007)	-0.004 (0.007)	0.029 (0.010)	-0.005 (0.090)	0.310 (12.867)
2M-1-Butene	0.001 (0.001)	-0.003 (0.003)	0.054 (0.063)	0.001 (0.003)	0.001 (0.001)	0.007 (0.004)	0.022 (0.111)	0.076 (3.412)
<i>i</i> -Pentane	0.123 (0.004)	0.098 (0.016)	0.283 (0.217)	0.108 (0.015)	0.025 (0.018)	0.041 (0.034)	0.112 (0.248)	-0.243 (23.181)
<i>n</i> -Pentane	0.051 (0.003)	0.048 (0.010)	0.083 (0.040)	0.047 (0.008)	0.027 (0.007)	0.032 (0.012)	-0.044 (0.403)	-0.154 (9.881)
1-Pentene	0.001 (0.000)	0.002 (0.001)	0.015 (0.017)	0.002 (0.001)	0.002 (0.001)	0.002 (0.001)	0.010 (0.047)	0.012 (0.704)
<i>c</i> -2-Pentene	0.000 (0.000)	-0.002 (0.002)	0.019 (0.022)	-0.001 (0.001)	-0.001 (0.001)	0.004 (0.002)	0.007 (0.036)	0.040 (2.026)
<i>t</i> -2-Pentene	0.000 (0.000)	-0.004 (0.003)	0.034 (0.040)	0.000 (0.002)	-0.001 (0.001)	0.006 (0.003)	0.009 (0.049)	0.067 (3.303)
<i>n</i> -Hexane	0.013 (0.002)	0.084 (0.026)	0.009 (0.011)	0.000 (0.007)	0.004 (0.003)	0.020 (0.005)	0.015 (0.049)	0.034 (0.403)
Benzene	0.042 (0.003)	0.055 (0.013)	0.017 (0.030)	0.022 (0.014)	0.063 (0.007)	0.048 (0.013)	0.056 (0.146)	0.006 (1.404)
Toluene	0.084 (0.011)	0.065 (0.018)	0.039 (0.078)	0.025 (0.025)	0.049 (0.014)	0.124 (0.032)	0.046 (0.174)	0.136 (4.106)
Ethylbenzene	0.022 (0.008)	0.028 (0.013)	-0.006 (0.050)	-0.019 (0.022)	-0.014 (0.013)	0.095 (0.041)	0.033 (0.228)	0.172 (10.411)
<i>m,p</i> -Xylene	0.030 (0.015)	0.057 (0.029)	0.021 (0.069)	-0.042 (0.046)	-0.038 (0.029)	0.202 (0.094)	0.035 (0.511)	0.557 (35.258)
Isoprene	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.550 (2.876)	0.000 (0.000)
MTBE	0.026 (0.002)	0.027 (0.007)	0.047 (0.033)	0.013 (0.007)	-0.011 (0.008)	0.016 (0.005)	0.090 (0.374)	0.052 (1.182)

Table 3

The six source profiles that were similar among PMF, UNMIX, and CMB (volume percentage; unavailable values are left blank)

Compound	Gasoline exhaust			Liquid/evaporated/exhaust gasoline			LPG			Natural gas			Petrochemical production			Biogenic		
	PMF	UMX ^{b,c}	CMB	PMF	UMX	CMB ^d	PMF	UMX	CMB	PMF	UMX	CMB	PMF	UMX	CMB	PMF	UMX	CMB
Ethane* ^a	0.117	0.077	0.042	0.017	0.003	0.000	0.023	0.175	0.009	0.385	0.291	0.763	0.000	-0.034	0.056	0.002	0.229	0.000
Ethylene	0.099	0.137	0.201	0.116	-0.014	0.000	0.071	0.028	0.010	0.000	0.090	0.000	0.174	0.119	0.525	0.005	-0.278	0.000
Acetylene*	0.168	0.144	0.239	0.185	0.120	0.000	0.009	0.007	0.002	0.095	0.232	0.000	0.060	0.169	0.000	0.000	-0.072	0.000
Propane*	0.120	0.089	0.006	0.000	-0.059	0.022	0.179	0.281	0.248	0.017	0.196	0.154	0.000	-0.064	0.011	0.000	-0.057	0.000
Propylene	0.009	0.013	0.032	0.021	0.022	0.000	0.071	0.013	0.050	0.023	0.003	0.000	0.049	0.030	0.000	0.090	0.050	0.000
2M-Propene	0.003	0.005		0.033	0.044		0.029	0.000		0.027	-0.004		0.014	0.018		0.026	0.042	
<i>i</i> -Butane*	0.062	0.055	0.006	0.013	0.016	0.056	0.160	0.132	0.160	0.002	0.051	0.022	0.016	0.003	0.000	0.020	0.001	0.000
<i>n</i> -Butane*	0.084	0.071	0.022	0.025	0.031	0.163	0.142	0.187	0.083	0.001	0.078	0.037	0.010	-0.013	0.015	0.000	-0.035	0.000
<i>t</i> -2-Butene	0.000		0.007	0.021		0.015	0.046	0.000	0.086	0.000		0.000	0.000		0.009	0.009		0.000
1-Butene	0.004	0.006	0.049	0.005	0.036	0.014	0.122	0.030	0.236	0.022	-0.004	0.000	0.027	0.029	0.016	0.017	-0.005	0.000
<i>c</i> -2-Butene	0.000		0.004	0.019		0.014	0.033	0.000	0.076	0.007		0.000	0.002		0.007	0.005		0.000
2M-1-Butene	0.001	0.001		0.027	0.054		0.003	0.001		0.043	0.001		0.004	0.007		0.004	0.022	
3M-1-Butene	0.000			0.001			0.000	0.000		0.016			0.000			0.002		
2M-2-Butene	0.001			0.040			0.000	0.000		0.000			0.000			0.014		
<i>i</i> -Pentane*	0.119	0.123	0.106	0.218	0.283	0.360	0.031	0.108	0.006	0.029	0.025	0.013	0.013	0.041	0.018	0.029	0.112	0.000
<i>n</i> -Pentane*	0.047	0.051	0.040	0.063	0.083	0.090	0.004	0.047	0.008	0.036	0.027	0.012	0.026	0.032	0.041	0.000	-0.044	0.000
1-Pentene	0.000	0.001	0.004	0.005	0.015	0.011	0.002	0.002	0.002	0.023	0.002	0.000	0.002	0.002	0.027	0.000	0.010	0.000
<i>c</i> -2-Pentene	0.000	0.000	0.004	0.010	0.019	0.011	0.003	-0.001	0.001	0.010	-0.001	0.000	0.001	0.004	0.000	0.001	0.007	0.000
<i>t</i> -2-Pentene	0.000	0.000	0.006	0.017	0.034	0.022	0.006	0.000	0.001	0.016	-0.001	0.000	0.001	0.006	0.006	0.000	0.009	0.000
<i>n</i> -Hexane*	0.013	0.013	0.021	0.013	0.009	0.023	0.006	0.000	0.001	0.017	0.004	0.000	0.024	0.020	0.010	0.007	0.015	0.000
<i>n</i> -Decane	0.000		0.001	0.000		0.001	0.000	0.000	0.001	0.007		0.000	0.010		0.000	0.008		0.000
Benzene*	0.052	0.042	0.045	0.038	0.017	0.017	0.003	0.022	0.006	0.057	0.063	0.000	0.028	0.048	0.051	0.026	0.056	0.000
Toluene*	0.066	0.084	0.081	0.058	0.039	0.091	0.027	0.025	0.004	0.094	0.049	0.000	0.128	0.124	0.138	0.035	0.046	0.000
Ethylbenzene	0.009	0.022	0.011	0.000	-0.006	0.013	0.000	-0.019	0.001	0.000	-0.014	0.000	0.091	0.095	0.029	0.036	0.033	0.000
<i>m,p</i> -Xylene	0.000	0.030	0.051	0.000	0.021	0.058	0.018	-0.042	0.003	0.055	-0.038	0.000	0.209	0.202	0.022	0.052	0.035	0.000
<i>o</i> -Xylene	0.005		0.018	0.007		0.020	0.009	0.000	0.002	0.003		0.000	0.087		0.014	0.018		0.000
Isoprene	0.001	0.000	0.004	0.001	0.000	0.000	0.001	0.000	0.005	0.006	0.000	0.000	0.003	0.000	0.004	0.580	0.550	1.000
α -Pinene	0.000			0.001			0.000	0.000		0.000			0.000			0.002		
β -Pinene	0.000			0.000			0.000	0.000		0.001			0.000			0.000		
Limonene	0.000			0.000			0.000	0.000		0.000			0.000			0.000		
MTBE	0.019	0.026		0.046	0.047		0.003	0.013		0.006	-0.011		0.024	0.016		0.009	0.090	

^a Asterisk indicates a fitting species in the CMB.^b UMX is UNMIX.^c Source 1 in Table 2.^d Only evaporated gasoline, not including liquid/exhaust gasoline.

source profiles resolved by UNMIX and the estimates of uncertainties. Six common sources were found in the three models (Table 3). Table 4 lists the source contributions estimated by the three models, which resolved almost 100% of the total selected VOCs. The negative compositions were less than half of the uncertainties; therefore, we concluded that these values were not significant and attributable to error.

The first three sources (Table 2) may be gasoline-related emissions because of their high loadings in *i/n*-pentane and MTBE (Watson et al., 2001). Compared to Source 3, Sources 1 and 2 had lower *i*-pentane loadings, but higher values for ethane, ethylene, and acetylene, which are typical by-products of gasoline combustion. Therefore, Sources 1 and 2 were identified as gasoline-powered vehicle emissions, and Source 3 as likely liquid/evaporated/exhaust gasoline. PMF also extracted gasoline exhaust and liquid/evaporated/exhaust gasoline (Table 3). The chemical compositions were similar, for example, abundant alkanes/alkenes with low carbon number during gasoline combustion and high *i*-pentane and MTBE during liquid gasoline splashing and gasoline evaporation. The source contributions from UNMIX were comparable with those from PMF: 42.6 and 39.7% for gasoline exhaust emissions and 5.7 and 11.8% for liquid/evaporated/exhaust gasoline, respectively (Table 4).

Source 4 was rich in propane, butane, and butene, which were representative species in China's LPG samples (Lu, 2004; Tsai et al., 2006) (Table 2) and also abundant in PMF results (Table 3). This profile had relatively high loadings in ethane, ethylene, *i/n*-pentane, and MTBE in UNMIX, suggesting that some emissions from gasoline exhaust were mixed in this source. Thus, the estimated contribution from LPG in UNMIX (19.9%) was higher than that in PMF (11.0%) (Table 4).

Source 5 was characterized by high ethane, which is the marker species of natural gas, and a similar source appeared in PMF results (Watson et al., 2001) (Tables 2 and 3). This profile exhibited an abundance of acetylene and propane, suggesting mixing with vehicular emissions, and led to a higher

Table 4

A comparison of source contributions estimated by UNMIX, PMF, and CMB (percentage in mass concentration; the values in parentheses are uncertainties; unavailable values are blank)

Source	UNMIX	PMF	CMB
Gasoline exhaust	42.6 = 35.6 (11.2) + 7.0 (3.0)	39.7 (0.9)	46.6 (1.4)
Liquid/ evaporated/ exhaust gasoline	5.7 (2.5)	11.8 (0.6)	14.1 (0.9) ^a
Gasoline-related	48.3	54.7	60.7
LPG	19.9 (13.4)	11.0 (0.8)	26.4 (0.8)
Natural gas	9.7 (1.9)	4.6 (0.7)	3.4 (0.5)
Petrochemical	17.4 (5.0)	19.9 (0.9)	9.8 (1.2)
Biogenic	2.0 (9.4)	1.6 (0.2)	0.6 (0.0)
Painting		4.7 (0.4)	2.9 (0.5)
Diesel exhaust		3.2 (0.3)	
Total mass concentration ^b	92.1	101.8	94.9

^a Only evaporated gasoline.

^b Unit is $\mu\text{g m}^{-3}$.

contribution estimation in UNMIX (9.7%) than in PMF (4.6%) (Table 4).

Source 6 had high levels of benzene, toluene, ethylbenzene, *m,p*-xylene (BTEX), ethylene, and acetylene. The *i/n*-pentane level was relatively low, which ruled out gasoline-related sources (Table 2). The conditional probability function (CPF) could be used to identify the orientation of this source from the extracted contributions and the wind direction data (Ashbaugh et al., 1985, see the Supplemental materials for more details). Fig. 1 shows a CPF plot of this source. It had high values in the southeast and west, where the petrochemical plants are located (Supplemental materials, Fig. S1). Three important chemical plants are located in Beijing: Yanshan Petrochemical Corporation Limited (YSPCL) in Fangshan, Beijing Chemical Plants in Chaoyang, and Eastern Chemical Works in Tongzhou. Their products include ethylene, polyethylene, polystyrene, and biphenyl. High BTEX concentrations were found in Tongzhou (Lu, 2004; Liu et al., 2005). The VOC emissions from the chemical plants, especially from YSPCL, may be transported to the sampling site by the common daytime southerly summer winds (Beijing Meteorological Center, 1987). Thus, this source may be related to petrochemical production in Beijing. A similar source was also extracted by PMF (Table 3). The contributions estimated by UNMIX and PMF were similar, 17.4 and 19.9%, respectively (Table 4). The abundance of acetylene (16.9%) in the UNMIX petrochemical profile implied mixing of vehicular exhaust. Moreover, the BTEX emitted from gasoline exhaust or paint that was not extracted by UNMIX may have been mixed into this petrochemical source.

Source 7 was dominated by isoprene, which is an indicator of a biogenic source (Watson et al., 2001) (Table 2). Its contribution was relatively low: 2.0% in UNMIX and 1.6% in PMF (Table 4). However, these may be underestimates because

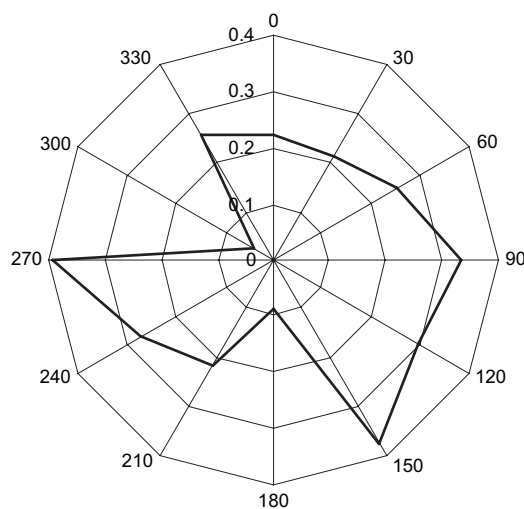


Fig. 1. Conditional probability function (CPF) plot of the petrochemical source for the highest 25% of the relative volume concentrations (wind data were collected on the top of a building about 100 m from the sampling site, the wind sector was 30°, and calm winds with wind speeds less than 0.5 m s⁻¹ were excluded).

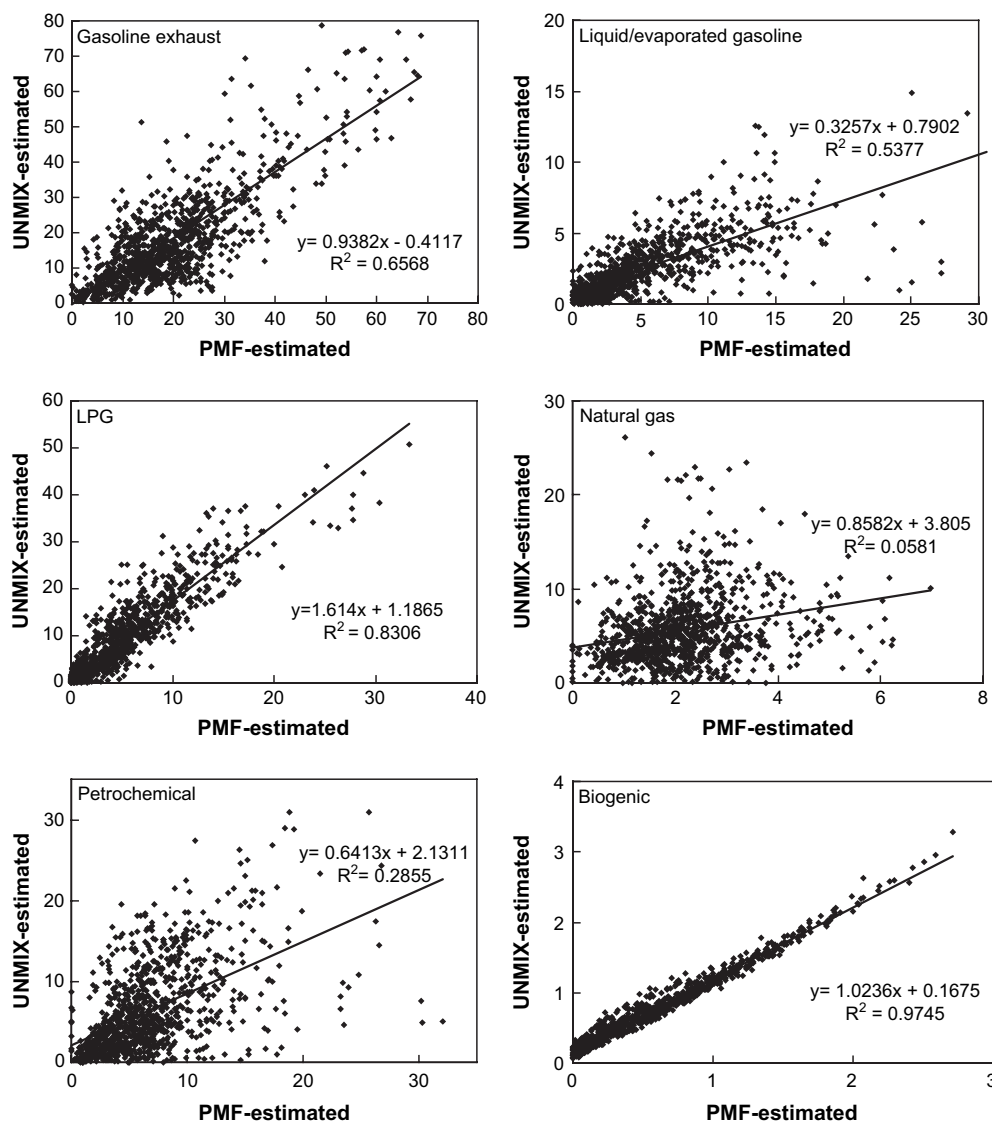


Fig. 2. Comparisons of the contributions (ppbV) for six common sources estimated by UNMIX and PMF (sampler number = 1002).

isoprene is very active in the atmosphere, and the measurements reflect the isoprene remaining after photochemical loss.

Negative concentrations may be generated by UNMIX because of errors, and Source 8 had large negative values. Its contribution uncertainty was three times its contribution estimation, suggesting that this source was statistically nonsignificant. Moreover, its source contribution was relatively low, 2.5%. Thus, this source was excluded from further consideration.

In Fig. 2, the contributions of the six common sources estimated by UNMIX and PMF for all 1002 samples are compared. The relationships between the two model results were good for gasoline exhaust, liquid/evaporated gasoline, LPG, petrochemical emissions, and biogenic emissions, but not for natural gas. As pointed out earlier, this may be due to the mixing of gasoline exhaust emissions in the natural gas source profile. Moreover, note that PMF extracted diesel vehicle exhaust and paint, whereas UNMIX did not. In general, however, the results from the two multivariate methods, UNMIX and PMF, were in good agreement.

3.2. Comparison between the CMB and factor models

CMB and UNMIX shared six common sources (Table 3). Paint was extracted by both CMB and PMF, but not by UNMIX. The chemical species used as markers to identify the six sources were abundant in the profiles (Table 3), although some differences occurred in the profiles for the other VOC species among PMF, UNMIX, and CMB.

The apportionment results based on the CMB model are listed in Table 4 and Fig. 3. Model performance parameters were good, with all R^2 equal to or greater than 0.99, χ^2 less than 0.67, and Mass% ranging from 96 to 115%. The seven contributors were gasoline exhaust (47%), LPG (26%), evaporated gasoline (14%), petrochemical production (10%), natural gas (3%), paint (3%), and biogenic emissions (2%).

The gasoline exhaust source profile in UNMIX and PMF was lower in ethylene and acetylene and higher in C2–C4 alkane than that in CMB. Ethane, ethylene, and acetylene levels presented in the UNMIX and PMF profiles for liquid/

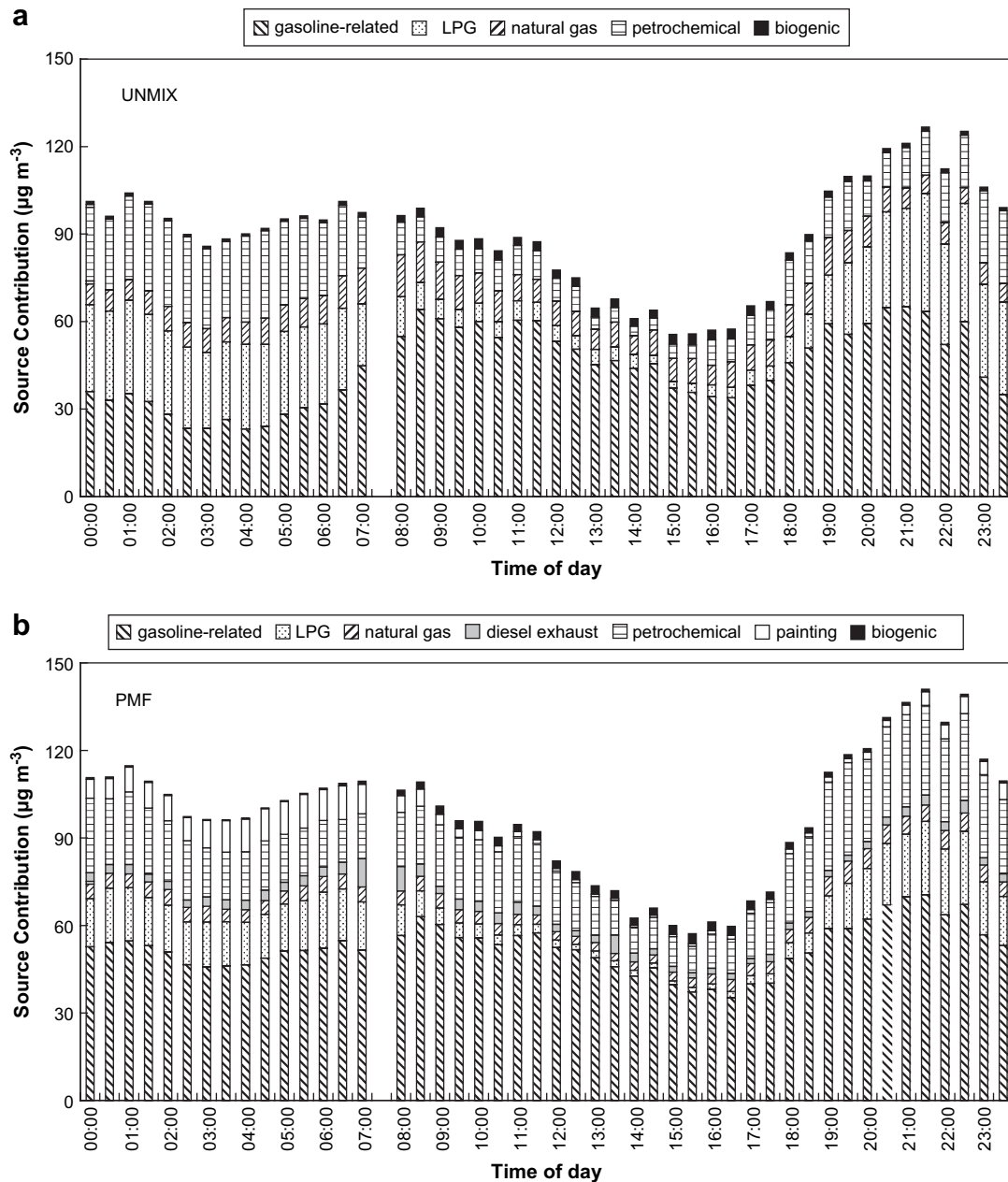


Fig. 3. Comparisons of the diurnal patterns of source contributions estimated by UNMIX (a), PMF (b), and CMB (c) (the data for 07:30 were excluded because too many measurements were missing for this time period).

evaporated/exhaust gasoline were higher than in the CMB evaporated gasoline profile due to the incorporation of gasoline exhaust. However, lower levels of *i/n*-butane and BTEX were also found in the profiles from the two-factor models. The gasoline-related contribution estimated by CMB was 60.7%, which was 12.4 and 9.2% higher than that in UNMIX and PMF, respectively.

CMB yielded a relatively high estimate for LPG contribution (26.4%), which was higher than the PMF and UNMIX results: 11.0 and 19.9%, respectively. As noted above, the overestimation of UNMIX LPG was caused by the incorporation of some gasoline-related emissions. The butene level was much higher in CMB than in PMF and UNMIX. This discrepancy could have been caused by the differences in the three

LPG profiles, but further study is needed for confirmation. The use of bottled LPG is still popular in Beijing, for both transportation and cooking, and many LPG stations are distributed in suburban and urban areas of Beijing. Moreover, the gas-filling process is poorly controlled in Beijing, and thus some LPG may be emitted into the surrounding air. This may be an important contributor to the high LPG levels estimated by CMB and UNMIX.

The CMB natural gas was dominated by ethane and propane. As noted, vehicular exhaust was incorporated in the UNMIX natural gas profile; thus, the contribution of natural gas in UNMIX was higher than that in CMB.

The petrochemical emissions estimated by CMB accounted for 9.8%, nearly 10% less than the contributions estimated by

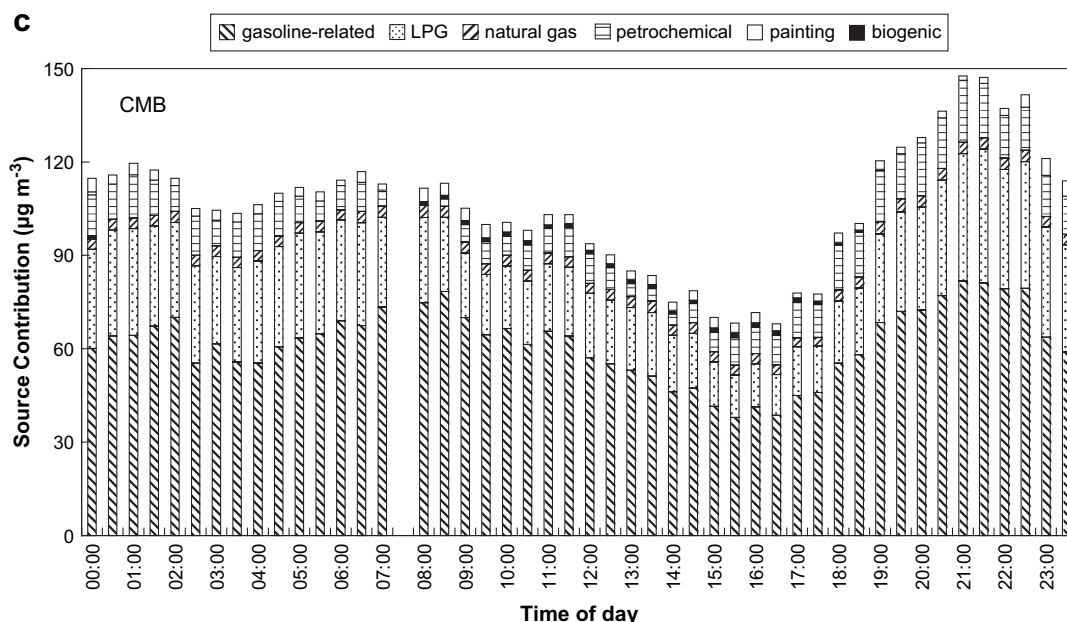


Fig. 3. (continued).

UNMIX and PMF. A large discrepancy was observed between the UNMIX/PMF-extracted petrochemical profile and the CMB profile. The former was dominated by BTEX and ethylene, while the latter was dominated by ethylene (52.5%) and toluene (13.8%). As stated by Song et al. (2007), some BTEX might be shifted from gasoline-related emissions to petrochemical emissions in the factor analysis methods. Thus, UNMIX and PMF estimated lower gasoline-related contributions and higher petrochemical contributions.

The three models yielded similar contribution estimates for biogenic emissions because their profiles were dominated by isoprene.

Because of the reactivity of VOCs, some discrepancies will always occur when using different receptor models to apportion VOC sources. The ambient VOCs that have reacted with hydroxyl and hydroperoxy radicals before their arrival at the receptor sites are “aged.” However, CMB uses “fresh” profiles measured at the initial point of emission to estimate contributions. While not requiring source profiles, factor analysis methods apportion contributions using a large number of samples in which each VOC species may vary in age and time available for undergoing photochemical reactions. Even when using only nighttime data, this problem cannot be avoided because the “aged” VOCs from the daytime remain. Thus, these inherent differences contributed to the discrepancies between the CMB and UNMIX/PMF model results.

3.3. Diurnal patterns of source contributions

The average diurnal patterns of source contributions estimated by the three models are shown in Fig. 3. The total VOC concentrations were higher at night than in the day, which may be explained by two factors. First, during summer in Beijing, daytime winds are stronger than those at night

(Beijing Meteorological Center, 1987), facilitating the dilution of air pollutants during the day. Fig. 4 shows the diurnal wind pattern measured on the top of a building about 100 m from the sampling site. The wind speed in the afternoon was twice that at midnight. Meanwhile, a higher atmospheric mixing boundary layer in the daytime will also result in better dilution of air pollutants. Second, photochemical reactions are more active during the day; thus, considerable amounts of alkenes, ethylbenzene, and xylene are depleted. However, the activity of photochemical reactions is lower at night, resulting in the accumulation of certain VOC species such as olefins and alkenes. The low wind speed (less than 1.0 m s^{-1}) may be the primary reason for high VOC concentrations at night.

Contributions from gasoline-related sources were high in the morning (07:00–09:00), low in the afternoon (14:00–16:00), and then high again in the evening (20:00–22:00). This could be explained by the strong traffic emissions associated with morning and evening rush hours and the weak dispersion and advection. Moreover, the VOCs were relatively fresh during these periods, as solar radiation was weak. The contributions of LPG were higher at night than during the day, a pattern shared with total VOCs, and may be explained by the above reasons. The peak concentrations of natural gas occurred around 06:00–08:00 and 19:00–20:00, when most people are cooking and bathing and therefore consuming large amounts of natural gas. Petrochemical emissions were also higher at night than during the day, probably because the winds from the chemical plants were more frequent at night and the chemical loss was less. Moreover, ethylene, ethylbenzene, *m,p*-xylene, and *o*-xylene, as the main components of the petrochemical emissions, are consumed by photochemical reactions during the day. The biogenic source followed a diurnal pattern very different from the other sources; this source began to rise rapidly after sunrise, and then its contribution dropped

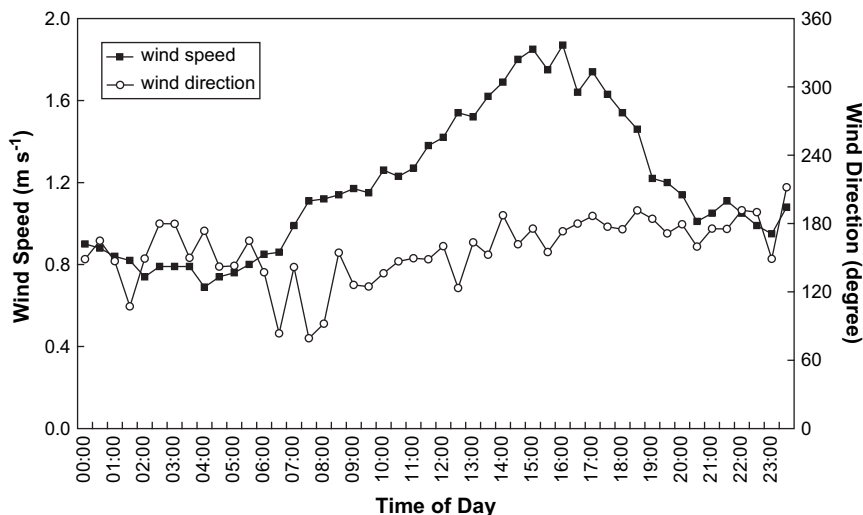


Fig. 4. Diurnal pattern of wind speed and direction.

after sunset and remained low throughout the night. This pattern is attributable to photosynthesis-related processes that depend strongly on sunlight.

4. Conclusions

Although several receptor models have been used for source apportionment, each model has physical assumptions and constraints. Intercomparisons among the results of different models are important to better understand the sources. In this study, three advanced receptor models, UNMIX, PMF, and CMB, were used to apportion ambient VOCs in Beijing, and their results were compared. The results of all models showed that gasoline-powered vehicle emissions contributed most to VOCs, and the contributions from petrochemical emissions and LPG were also important.

The results from UNMIX and PMF were in good agreement, and the discrepancies between CMB and the two-factor analysis methods may be attributable to inherent model weaknesses, more specifically the unavailability of accurate source profiles for CMB. Considering the complexity and variability of the actual conditions, we highly recommend applying different models to the same data set in source apportionment. The overlapping conclusions should be seriously considered and used to guide efforts in improving air quality.

The source apportionment results from the present study may be useful in shaping policy to abate surface ozone pollution, regardless of whether the ozone formation is controlled by the VOCs. Furthermore, note that reducing VOC emissions may also decrease SOA pollution in the ambient air.

Acknowledgments

We thank Dr Ronald Henry for providing the UNMIX model. This study was funded by the China National Natural Science Foundation program (grant nos. 40575059 and

20637001) and the National Key Basic Research and Development Program of China (grant no. 2005CB422204).

Appendix A. Supplementary information

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.envpol.2007.12.014](https://doi.org/10.1016/j.envpol.2007.12.014).

References

- 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.
- Beijing Environmental Protection Bureau, 1999–2004. Annual Report. <http://www.bjepb.gov.cn>.
- Beijing Meteorological Center, 1987. *Climatology of Beijing*. Beijing Press, 185.
- Duan, F., He, K., Ma, Y., Jia, Y., Yang, F., Lei, Y., Tanaka, S., Okuta, T., 2005. Characteristics of carbonaceous aerosols in Beijing, China. *Chemosphere* 60, 355–364.
- Fujita, E.M., 2001. Hydrocarbon source apportionment for the 1996 Paso del Norte Ozone Study. *Science of the Total Environment* 276, 171–184.
- Goldan, P.D., Kuster, W.C., Williams, E., Murphy, P.C., Fehsenfeld, F.C., Meagher, J., 2004. Nonmethane hydrocarbon and oxy hydrocarbon measurements during the 2002 New England Air Quality Study. *Journal of Geophysical Research-Atmospheres* 109 (D21), D21309.
- Henry, R.C., 2003. Multivariate receptor modeling by N-dimensional edge detection. *Chemometrics and Intelligent Laboratory Systems* 65, 179–189.
- Hopke, P.K., 2003. Recent developments in receptor modeling. *Journal of Chemometrics* 17 (5), 255–265.
- Hopke, P.K., Ito, K., Mar, T.F., Christensen, W.F., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Larson, T.V., Liu, H., Neas, L., Pinto, J., Stolzel, M., Suh, H., Paatero, P., Thurston, G.D., 2006. PM source apportionment and health effects. 1. Intercomparison of source apportionment results. *Journal of Exposure Analysis and Environmental Epidemiology* 16, 275–286.
- Ito, K., Xue, N., Thurston, G., 2004. Spatial variation of PM_{2.5} chemical species and source-apportioned mass concentrations in New York City. *Atmospheric Environment* 38, 5269–5282.
- Liu, Y., Shao, M., Zhang, J., Fu, L., Lu, S., 2005. Distributions and source apportionment of ambient volatile organic compounds in Beijing city, China. *Journal of Environmental Science and Health* 40, 1843–1860.

- Lu, S.H., 2004. Source apportionment of anthropogenic emissions of volatile organic compounds. MSc Thesis, Peking University.
- Maykut, N.N., Lewtas, J., Kim, E., Larson, T.V., 2003. Source apportionment of PM_{2.5} at an urban IMPROVE site in Seattle, Washington. *Environmental Science & Technology* 37, 5135–5142.
- 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.
- Mukerjee, S., Norris, G.A., Smith, L.A., Noble, C.A., Neas, L.M., Ozkaynak, A.H., Gonzales, M., 2004. Receptor model comparisons and wind direction analysis of volatile organic compounds and submicrometer particles in an arid, binational, urban air shed. *Environmental Science & Technology* 38, 2317–2327.
- Paatero, P., 2004. User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, Part 1: Tutorial. University of Helsinki, Finland.
- Poirot, R.L., Wishinski, P.R., Hopke, P.K., Polissar, A.V., 2001. Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont. *Environmental Science & Technology* 35, 4622–4636.
- Seinfeld, J.H., Pandis, S.N., 1998. *Atmospheric Chemistry and Physics*. Wiley, New York.
- Song, Y., Shao, M., Liu, Y., Lu, S.H., Kuster, W., Goldan, P., Xie, S.D., 2007. Source apportionment of ambient volatile organic compounds in Beijing. *Environmental Science & Technology* 41, 4348–4353.
- Tsai, W.Y., Chan, L.Y., Blake, D.R., Chu, K.W., 2006. Vehicular fuel composition and atmospheric emissions in South China: Hong Kong, Macau, Guangzhou, and Zhuhai. *Atmospheric Chemistry and Physics* 6, 3281–3288.
- United States Environmental Protection Agency (US EPA), 2004. EPA-CMB8.2 Users Manual. EPA, Washington, DC.
- Wang, B.G., 2002. Study on chemical speciation and source apportionment of atmospheric volatile organic compounds in Pearl River Delta. PhD dissertation, Peking University.
- Watson, J.G., Robinson, N.F., Chow, J.C., Henry, R.C., Kim, B.M., Pace, T.G., Meyer, E.L., Nguyen, Q., 1990. The USEPA/DRI chemical mass balance receptor model. *Environmental Software* 5, 38–49.
- Watson, J.G., Chow, J.C., Fujita, E.M., 2001. Review of volatile organic compound source apportionment by chemical mass balance. *Atmospheric Environment* 35, 1567–1584.