Atmospheric Environment 45 (2011) 439–449

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Using multiple type composition data and wind data in PMF analysis to apportion and locate sources of air pollutants

Yiu-chung Chan [a,*], Olga Hawas [b], Darryl Hawker [a], Peter Vowles [a], David D. Cohen [c], Eduard Stelcer [c], Rod Simpson [d], Gary Golding [e], Elizabeth Christensen [e]

[a] Griffith School of Environment, Griffith University, Nathan Campus, Queensland 4111, Australia
[b] Department of Environment and Resource Management, GPO Box 2771, Queensland 4001, Australia
[c] Institute for Environmental Research, ANSTO, Locked Bag 2001 Kirrawee DC, NSW 2232, Australia
[d] Faculty of Science, Health and Education, University of the Sunshine Coast, Maroochydore DC, Queensland 4558, Australia
[e] Queensland Health Forensic and Scientific Services, 39 Kessels Road, Coopers Plains, Queensland 4108, Australia

A R T I C L E I N F O

Article history:
Received 24 May 2010
Received in revised form
7 September 2010
Accepted 28 September 2010

Keywords:
Particulate matter
VOC
Wind direction
Source apportionment
Positive matrix factorisation

A B S T R A C T

In this study a small but comprehensive data set from a 24-hourly sampling program carried out during June 2001 in an industrial area in Brisbane was chosen to investigate the effect of inclusion of multiple type composition data and wind data on source apportionment of air pollutants using the Positive Matrix Factorisation model, EPA PMF 3.0. The combined use of aerosol, VOC, main gaseous pollutants composition data and wind data resulted in better values of statistical indicators and diagnostic plots, and source factors which could be more easily related to known sources. The number of source factors resolved was similar to those reported in the literature where larger data sets were used. Three source factors were identified for the coarse particle samples, including ‘crustal matter’, ‘vehicle emissions’ and ‘sea spray’. Seven source factors were identified for the fine particle and VOC samples, including ‘secondary and biogenic’, ‘petroleum refining’, ‘vehicle emissions’, ‘petroleum product wholesaling’, ‘evaporative emissions’, ‘sea spray’ and ‘crustal matter’. The factor loadings of the 16 wind sectors and the calm wind sector from the PMF analysis were also used to quantify the directional contribution of the source factors. While the contributions were higher in the prevailing wind directions as expected, calm winds were also found to contribute up to 17% of the pollutant mass on average. The factor loadings, normalised by the overall abundance of the wind sectors, were also used to assess the directional dependences of the source factors. The results matched well with the location of known sources in the area. There was also a higher contribution potential from calm winds for local sources compared to that for distant sources. The results of directional effect using the PMF factor loading approach were similar to those by using the other approaches. This approach, however, also provides estimates of the mass contribution of source factors by wind sector and also the uncertainty of the results.

1. Introduction

The identification of emission sources and quantification of their contribution to the ambient concentration of pollutants has been one of the major foci in urban air quality research. Examples of the commonly used receptor model software which have been developed for this purpose include the USEPA’s Positive Matrix Factorisation ((PMF 3.0) USEPA, 2008), Unmix 6.0 (USEPA, 2007) and Chemical Mass Balance ((CMB 8.2) USEPA, 2004). Except for CMB, the use of receptor models requires the input of a composition data set which consists of a sufficient number of samples collected from the receptor site. The number of samples in the data set employed in the published literature ranges from tens (Laupsa et al., 2009) to tens of thousands (Yue et al., 2008), with the majority of analyses done on more than 100 samples (Pandolfi et al., 2008; USEPA, 2008). The receptor models generally do not provide a strict guideline on the minimum number of samples required. This is because the number of samples required also depends on the number of chemical species (in particular the tracers) analysed, the quality of the data and the number of contributing sources. In fact, in theory, the PMF model will work on a minimum data set where sample size equals the number of contributing sources. This minimum data set should consist of samples where each has only one source (USEPA, 2008). In reality, samples entirely contributed

* Corresponding author. Tel.: +617 37353558; fax: +617 37357459.
E-mail address: a.chan@griffith.edu.au (Y.-c. Chan).

1352-2310/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2010.09.060

by one source or entirely lacking a contribution from one source are rare, therefore a large data set is desirable.

Such a large data set can be obtained from a long term sampling program (Cohen et al., 2010), highly time-resolved measurements (Dreyfus et al., 2009) or individual particle analysis using techniques such as scanning electron microscopy and aerosol spectrometry (Yue et al., 2008). This may imply high sampling and/or analysis costs which can compromise feasibility of small scale projects such as the investigation of localised pollution problems by local authorities. Hopke (2010) advises getting the best and the most one can from the data available. A possible way of extracting more source information from a composition data set is to include data of other air pollutants such as volatile organic compounds (VOCs) (Wu et al., 2007) and major gaseous pollutants such as NO_x, SO_2 and CO (Rizzo and Scheff, 2007), particle size distribution data (Zhou et al., 2005) and air trajectory and meteorological data (Buzcu-Guven et al., 2007) in the analysis. The combination of particulate matter and gaseous pollutant composition data for source apportionment is currently a matter of debate. For example, the Unmix model does not recommend the use of multiple data types (USEPA, 2007), while the CMB and PMF models suggest this may assist the identification of sources of fine particles ($PM_{2.5}$) (USEPA, 2004,2008) given the similar aerodynamic properties of fine particles and gaseous pollutants.

The high correlation between wind direction and air pollution has been widely reported (Yue et al., 2008), but as far as is known the direct combination of wind data with composition data in PMF modelling has not been reported in the literature. There have been studies done on the assessment of the directional dependence of sources, such as the Directional Relative Strength (DRS) approach and the Conditional Probability Function (CPF) approach. In the DRS approach, the ratio of the source contribution weighted-wind sector abundance to the overall wind abundance in a wind sector is used to assess the pollution potential of the source in that sector (Lau et al., 2010). In the CPF approach, the ratio of number of hours in which the source contribution exceeds the set threshold to the number of hours in which the wind comes from a wind sector is used to assess the directional dependence of the source in that sector (Lestari and Mauliadi, 2009; Yue et al., 2008). When applying the CPF approach to a 24-hourly composition data set, the source contribution of each hour is assumed as the same as the daily averaged contribution (Lestari and Mauliadi, 2009). Also the contribution threshold is usually set to represent the higher contribution events (for example, values from the upper 50th percentile (Buzcu-Guven et al., 2007) to the upper 25th percentile (Lestari and Mauliadi, 2009) of the source contribution have been used in the reported literature), therefore not all the contribution events are included in this approach. The calm wind periods are also usually excluded from the DRS and CPF analyses. In addition, while these approaches have been found useful in locating the sources, they do not provide estimates of the uncertainty of the results and estimates of the mass contribution of sources by wind sector. The direct combination of wind sector abundance data with composition data in PMF modelling could be used to overcome these problems.

In this study, a composition data set from a 24-hourly sampling program carried out over one month (June) in 2001 in an industrial area (Eagle Farm) in Brisbane, Australia (Hawas, 2001; Hawas et al., 2002), was used to investigate the effect of inclusion of multiple type composition data and wind data on source apportionment of air pollutants. This small data set was chosen because simultaneous samplings were carried out to obtain information on fine particles, coarse particles ($PM_{2.5-10}$) and VOCs. The routine monitoring carried out by the responsible government agency at the site also provided hourly data on meteorology and the concentration of pollutants including NO, NO_2, SO_2 and O_3. The wide range of industrial activities and the heavy traffic in the surrounding area also provide a variety of contributing sources and make the data set suitable for the purpose of this study. Emission data for the surrounding area are also available from the Australian National Pollutant Inventory database (NPI) (DEWHA, 2010) for comparison with the modelled results.

The objectives of this study then were: (1) to investigate the effect of inclusion of multiple type composition data and wind sector abundance data on source apportionment of particulate matter, VOC and gaseous pollutants using the PMF model; (2) to investigate the use of factor loading of the wind sectors from the PMF analysis to estimate the mass contribution of source factors from different wind directions; and (3) to investigate the use of factor loadings, normalised by the overall wind sector abundances, to estimate the directional dependences of source factors, and to compare the results with those from the DRS and CPF analyses.

2. Methodology

2.1. Sampling and chemical analysis

The sampling was carried out at the Queensland Department of Environment and Resource Management (DERM) monitoring station at Eagle Farm (27.44°S 153.08°E), Brisbane. The surrounding area is characterised as light to heavy industrial with intensive movement of traffic. A map of the sampling site and the significant industrial sources and major roads within 10 km of the station is shown in Fig. 1, which is compiled from information from the National Pollutant Inventory for 2000–2001 (DEWHA, 2010). 24-hourly (7 am–7 am) samples were collected during 2–30 June 2001. A dichotomous sampler was used to collect both fine and coarse particle samples on Teflon filters at the flow rate of 16.7 L min^{-1}. VOC and carbonyl compound samples were simultaneously collected on sorbent tubes and 2,4-dinitrophenylhydrazine (DNPH) cartridges, respectively, using personal sampling pumps. The sorbents used in the VOC sampling were Tenax TA (for the sampling of C_7–C_{26} VOCs) and Carbosieve III (for C_4–C_6 VOCs). Based on the results of breakthrough tests carried out at the site, the volume of air sampled was 12 L and 280 L for the VOC and carbonyl sampling, respectively. The sampling height was approximately 1.5 m. The dichotomous samples were analysed for black carbon (mainly elemental carbon, EC) and 20 elements using the Laser Integrated Plate method and the ion beam analysis techniques. The VOC samples were analysed for 36C_4–C_{10} VOCs using Thermal Desorption–Gas Chromatography–Mass Spectroscopy (TD–GC–MS), and the carbonyl samples were extracted and analysed for 12 carbonyls using High Performance Liquid Chromatography (HPLC). The sampling and analysis details have been reported elsewhere (Hawas, 2001; Hawas et al., 2002). The DERM station also collected hourly meteorology data and data of ambient concentrations of O_3, NO_x and SO_2.

The laboratory and field blank levels of the analysed species in the particle samples were below 0.005 μg m^{-3}, except P (0.007 μg m^{-3}) and Si (0.015 μg m^{-3}). The detection limit (DL) of the particulate matter species ranged from 0.002 to 1 μg m^{-3} (EC), with the majority of the DL values below 0.01 μg m^{-3}. The blank levels of the VOCs ranged from below 0.001 to 15.7 μg m^{-3} (styrene), with the majority of the blank level values below 0.5 μg m^{-3}. The DL of the VOCs analysed was estimated as 0.083 μg m^{-3} while that of the carbonyl compounds was estimated as 0.18 μg m^{-3} (Hawas, 2001). The chemical species categorised as poor in data quality were excluded from the source apportionment analysis (Section 2.2). The concentration values (blank subtracted) which were below DL were assumed to be 0.5 DL. In total, 28 days of samples were included in

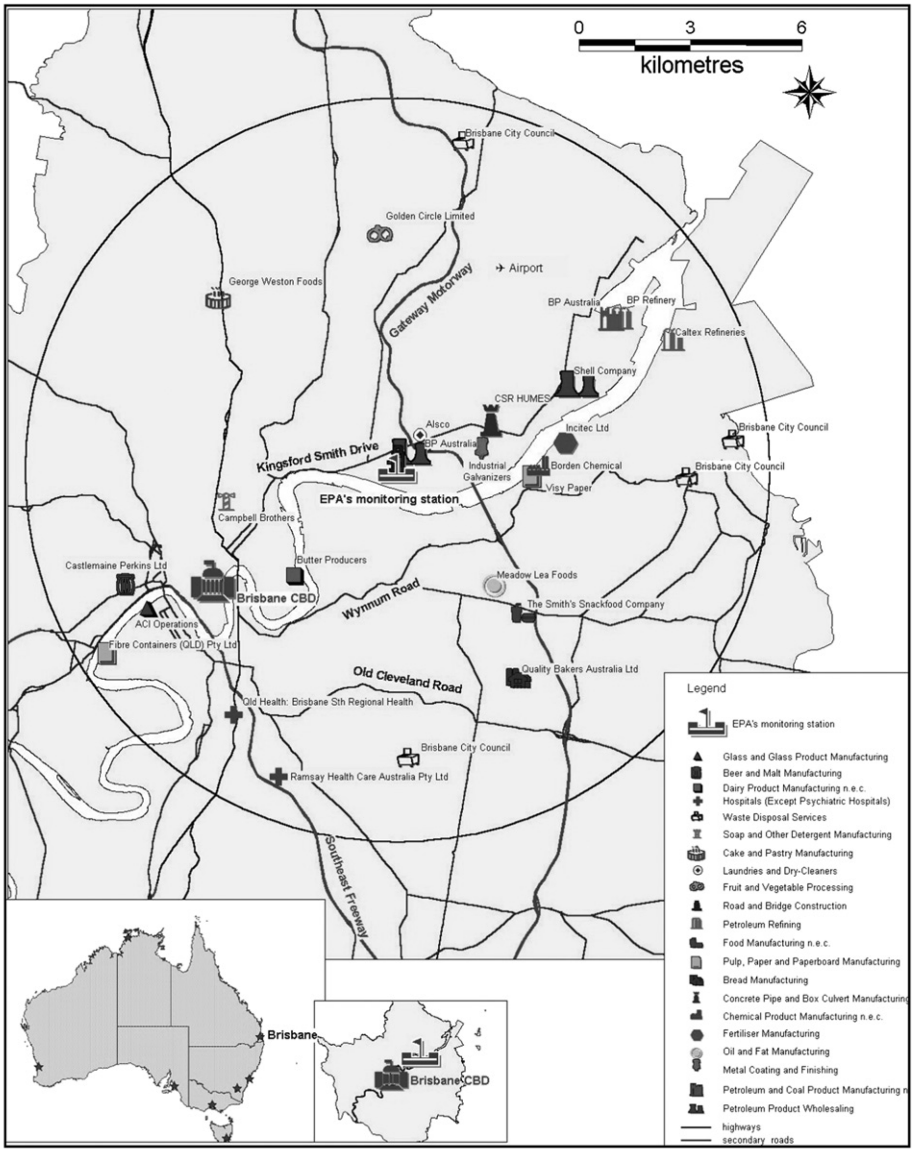

Fig. 1. Major industries and roads within 10 km of the monitoring station.

the analysis. A summary of the composition data is shown in Table 1. As shown here, the total mass of the analysed species explained 41% and 27% of the mass of fine and coarse particles in the samples on average, respectively. The remaining mass was probably due to species not analysed in this study including organic carbon, ammonium salts and nitrates. Also among the VOCs analysed in this study, emission data for twelve of them are also available from the Australian National Pollutant Inventory database (NPI) (DEWHA, 2010) for comparison with the modelled results.

2.2. PMF analysis

PMF 3.0 (USEPA, 2008) was used to apportion contribution from emission sources. The principles and usage of the model are detailed in the user manual and in the literature and so will not be repeated here. The guidelines specified in the user manual were closely followed in this study. The PMF 3.0 model requires two input files: one for the measured concentration of the species and one for the estimated uncertainty of the concentration. In the concentration file, missing values (denoted by the value of −999) were replaced by the median value of the species and assigned an uncertainty value of 4 × standard deviation (sd) by this model. The uncertainty of the concentration was estimated by (USEPA, 2008):

$$\text{For concentrations} < \text{DL}: \ \text{Uncertainty} = 5 \times \text{DL}/6$$
$$\text{For concentrations} \geq \text{DL}: \ \text{Uncertainty} = \left(\text{DL}^2 + \text{Precision}^2\right)^{0.5} \quad (1)$$

The precision of the concentrations of coarse particles, fine particles, coarse EC and fine EC, in terms of RSD%, were estimated as ±2.2%, ±5.5%, ±5% and ±1%, respectively. The precision of the concentrations of the other particulate matter species ranged from below ±0.001 to ±0.12 $\mu g\ m^{-3}$ (coarse H), with the majority of the precision values below ±0.05 $\mu g\ m^{-3}$ (Hawas, 2001). The precision of the concentrations of the VOCs was estimated as ±25% and ±15% for the VOC and the carbonyl compounds, respectively (Chan et al., 2008b). Concentrations beyond 4 × sd from the average concentration of the species were assigned a large uncertainty value of 10 × sd to minimise the distortion of data set by these outliers. Categorisation of quality of data was based on the signal to noise

Table 1
Composition of particulate matter and VOC samples (in μg m^{-3}).

Species	Ave concentration	Max concentration	PMF category
Coarse particles	10.2	19.7	Weak
Coarse EC	0.32	0.61	Strong
Coarse Al	0.08	0.23	Strong
Coarse Si	0.43	1.13	Strong
Coarse S	0.08	0.17	Strong
Coarse Cl	0.62	1.67	Strong
Coarse K	0.09	0.28	Strong
Coarse Ca	0.16	0.48	Strong
Coarse Ti	0.05	0.15	Strong
Coarse V	0.001	0.002	Weak
Coarse Cr	0.001	0.003	Weak
Coarse Mn	0.005	0.011	Strong
Coarse Fe	0.31	0.91	Strong
Coarse Co	0.002	0.005	Weak
Coarse Cu	0.003	0.007	Weak
Coarse Zn	0.01	0.03	Strong
Coarse Na	0.47	2.32	Weak
Coarse H	0.12	0.22	Weak
Fine particles	7.2	14.9	Weak
Fine EC	2.28	3.67	Strong
Fine Al	0.004	0.020	Weak
Fine Si	0.04	0.11	Strong
Fine S	0.18	0.32	Strong
Fine Cl	0.09	0.40	Strong
Fine K	0.04	0.11	Strong
Fine Ca	0.02	0.04	Strong
Fine Ti	0.02	0.09	Strong
Fine V	0.002	0.010	Weak
Fine Cr	0.001	0.003	Weak
Fine Mn	0.003	0.009	Weak
Fine Fe	0.08	0.22	Strong
Fine Zn	0.02	0.08	Weak
Fine Br	0.002	0.009	Weak
Fine Pb	0.01	0.02	Weak
Fine H	0.18	0.40	Weak
Isoprene[b]	0.69	5.56	Weak
α-Pinene	2.08	10.8	Strong
β-Pinene	1.60	6.37	Strong
Cymene	0.42	1.80	Strong
Cineole	0.73	2.09	Strong
Benzene[a]	14.0	94.4	Strong
Toluene[a]	41.1	319	Strong
Ethylbenzene[a]	6.05	41.5	Strong
m,p-Xylene[a]	21.8	125	Strong
o-Xylene[a]	8.54	55.4	Strong
Styrene[a]	15.5	95.6	Strong
n-Propylbenzene	0.53	1.69	Strong
1,3,5-Trimethylbenzene	1.29	3.70	Strong
1,2,4-Trimethylbenzene	4.38	13.1	Strong
Pentane	459	4110	Strong
n-Hexane[a]	98.5	901	Strong
Heptane	289	6980	Strong
Octane	6.38	17.3	Strong
Nonane	10.3	27.8	Strong
Decane	9.63	60.2	Strong
2,3-Dimethylbutane	460	4160	Strong
2-Methylpentane	110	2600	Strong
3-Methypentane	7.31	41.8	Strong
Methylcyclopentane	5.12	34.2	Strong
Cyclohexane[a]	2.33	13.6	Strong
2-Methylhexane	97.3	1810	Strong
3-Methylhexane	97.7	1750	Strong
Methylcyclohexane	5.81	22.8	Strong
1,1,1-Trichloroethane	0.86	2.01	Strong
Carbon tetrachloride	2.55	4.63	Strong
Trichloroethylene[a]	0.20	0.69	Weak
Tetrachloroethylene[a]	26.5	307	Strong
p-Dichlorobenzene	3.08	13.6	Strong
Formaldehyde[a]	1.82	2.87	Strong
Acetaldehyde[a]	3.04	4.14	Strong
Acetone[a]	0.47	2.13	Strong
Propionaldehyde	0.34	0.58	Weak

[a] VOCs included in the NPI database; the xylenes were grouped together in the NPI database.
[b] Isoprene was only analysed in the 9 samples collected after 20/6/2001.

ratio (S/N) and the percentage of samples above DL (Buzcu-Guven et al., 2007). Those species which have S/N ≥ 2 were categorised as strong in data quality. Those with S/N between 0.2 and 2, or with a 50th percentile value of 0 (i.e. with 50% or more of the samples below DL), were categorised as weak in quality. These species are not likely to provide enough variability in concentration and therefore contribute to the noise in the results (USEPA, 2008). In the PMF model the provided uncertainty of species categorised as weak in quality are tripled. Those with S/N ≤ 0.2 or a 75th percentile value of 0 were categorised as bad in quality and were excluded from the PMF analysis. The list of particulate matter and VOC species included in PMF analysis and the quality of the data are shown in Table 1. The O_3, NO and NO_2 data were categorised as strong in quality, while the SO_2 and wind data were categorised as weak in quality. The hourly wind direction data was compiled and input as 24-hourly average abundance of wind in the 16 wind sectors (N to NNW; expressed as a fraction) and the calm wind sector. The overall abundance of wind in the wind sectors during June 2001 is shown in Fig. 2. The quality of wind sector abundance data were categorised as weak in the PMF analysis.

The objective of PMF analysis is to minimise the value of Q, which is defined as (USEPA, 2008):

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{1}{U_{ij}} \left(C_{ij} - \sum_{k=1}^{p} G_{ik} F_{kj} \right) \right]^2 \quad (2)$$

where C_{ij}'s are the measured concentration values (in μg m^{-3}), U_{ij}'s are the estimated uncertainty values (in μg m^{-3}), G_{ik}'s are the factor score (source contribution) values, F_{kj}'s are the factor loading (source profile) values, n is the number of samples, m is the number of species and p is the number of sources included in the analysis. Two criteria for the target Q value have been suggested in the literature. One is for target Q = nm (PMF 1.1) (USEPA, 2005) and the other one is for target Q = nm − np − mp (PMF2 and PMF 3.0) (Paatero, 2004 and USEPA, 2008). The former criterion is based on the definition of Q (Eq. (2)) while the latter is based on the degrees of freedom of the multivariate analysis. In this study it was found that for a small data set a large number of factors would be needed

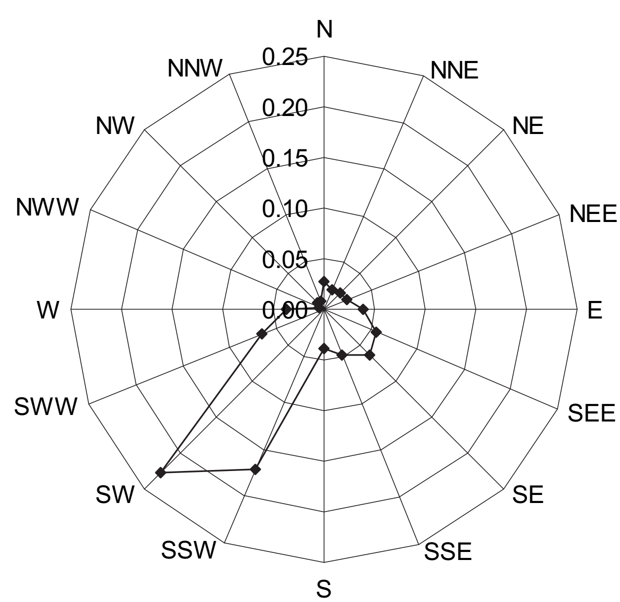

Fig. 2. Overall wind sector abundance during June 2001.

to fulfil the latter criterion, and also the extracted factor profiles were difficult to relate to physical sources. The former criterion was found to result in a more realistic number and profile of factors, and was therefore used in the analysis. In addition to the Q value, other fit diagnostic guidelines specified in the user manual were also employed to determine the appropriate number of factors. These guidelines include that:

- a robust Q value be obtained close to the theoretical Q value.
- the factors identified can be related to the known physical sources.
- O/P scatter plots show good correlation between the observed concentration and the measured concentration of species, and with the majority of standardised residuals of species between −3 and +3.
- G-space plots show data points lying within the source axes.
- source profiles have small Discrete Difference Percentile (DDP) values, in particular for tracer species.

A total of 100 random runs were computed for each data set to ensure no local minima of Q values were observed. In order to compare the results from run to run, the initial seed value was set to 1 in all trials. The convergent run with the Q value closest to the target Q value was used as the base run. Then 100 bootstrap runs were performed to assess the uncertainty of the factor loadings and factor scores in the base run. The PMF 3.0 model also provides the rotational freedom parameter (F_{peak}) function that can control whether more extreme values are assumed for the factor loadings (by assigning positive F_{peak} values) or the factor scores (by assigning negative F_{peak} values). In this study, altering the F_{peak} value was not found to result in substantially better source profiles. Consequently, base run results ($F_{peak} = 0$) are reported in this paper.

2.3. Estimation of mass contribution of source factors by wind sector

The factor loading value in the wind sector profile of the sources was used to estimate the average mass contribution of the sources from the 17 wind sectors (including the calm wind sector). Basically,

$$C_{kl} = F_k \times F_{kl} / \sum_{l=1}^{17} F_{kl} \tag{3}$$

where C_{kl} is the average contribution of source k in wind sector l (in $\mu g\ m^{-3}$), F_k is the factor loading of source k in aerosol mass or VOC mass (i.e. the average contribution of source k to aerosol mass or VOC mass) (in $\mu g\ m^{-3}$), and F_{kl} is the factor loading of source k in sector l (in $\mu g\ m^{-3}$). The total contribution of all the identified sources due to wind in sector l can also be estimated by summing the contribution of all the source factors from sector l together.

2.4. Estimation of directional dependences of source factors

The directional mass contributions of a source factor, which were estimated from the wind sector factor loading values in Section 2.2, are the combined result of the locations of the source factor and the wind sector abundance over the sampling period (Fig. 2). Therefore by dividing the wind sector factor loading values by the overall abundance of wind in each sector, the normalised wind sector factor loading values can then be used to assess the directional dependences of the source factor and to locate the directions of the source factor:

$$\text{Normalised } F_{kl} = F_{kl}/P_l \tag{4}$$

where F_{kl} is the factor loading of source k in sector l (in $\mu g\ m^{-3}$) and P_l is the abundance of wind in sector l (expressed as a fraction). The

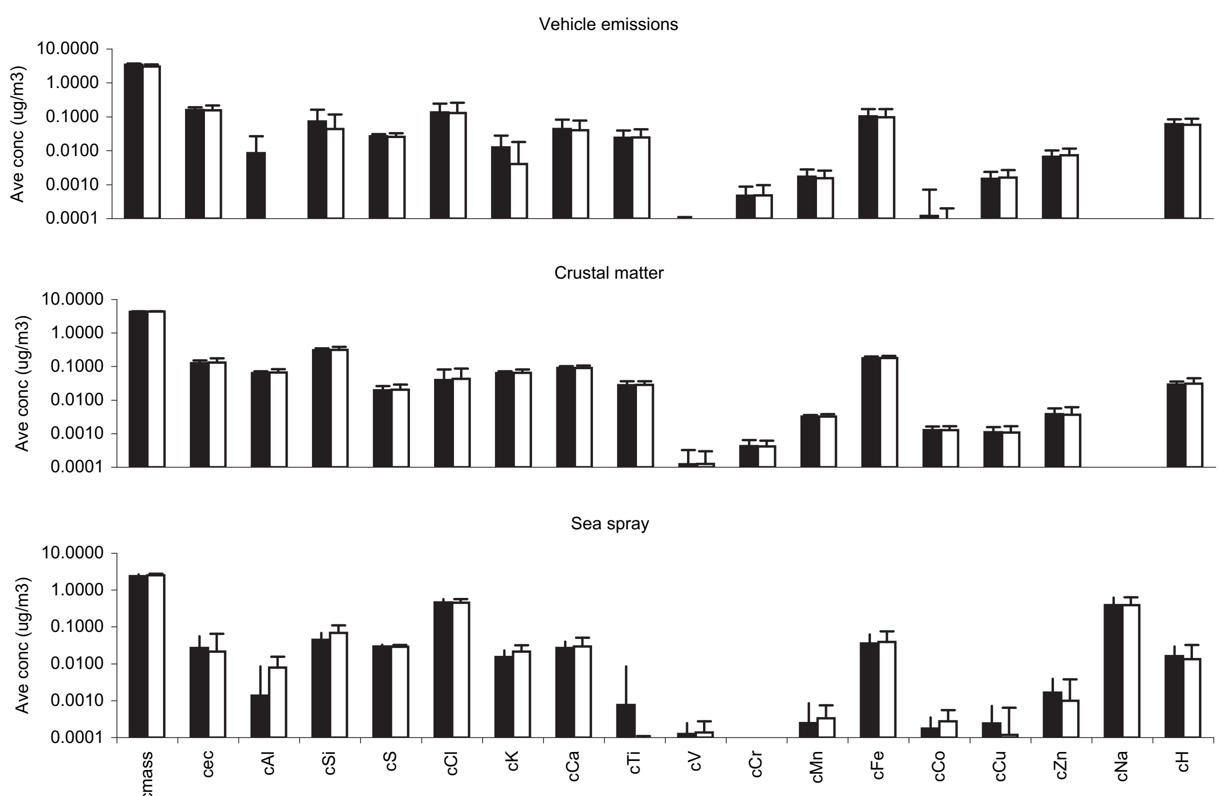

Fig. 3. Source profiles of coarse particles from PMF model analysis. Remarks: Solid bars are for the PMF model analysis with wind data, blank bars are for the without wind data analysis. 'cmass' represents the concentration of coarse particles in the samples.

possibility of having the source k in sector l can then be assessed by comparing the normalised factor loading value from this sector with those from the other wind sectors.

In this study, the normalised factor loading values were also compared with the Directional Relative Strength and Conditional Probability Function indices. The DRS index was calculated by Lau et al. (2010) as follows:

$$E_{kl} = \sum_{i=1}^{16} (C_{ki} P_{il}) / \sum_{i=1}^{16} C_{ki} \text{ and } DRS_{kl} = E_{kl}/P_l \tag{5}$$

where E_{kl} is the source contribution weighted-wind sector abundance of source k in sector l (expressed as a fraction), C_{ki} is the contribution from source k to sample i (in μg m^{-3}), P_{il} is the abundance of wind in sector l during the sampling period of sample i (expressed as a fraction), P_l is the overall abundance of wind in sector l (in fraction), and DRS_{kl} is the directional relative strength of source k in sector l (no units).

The CPF value was calculated by Lestari and Mauliadi (2009):

$$CPF_{kl} = m_{kl}/n_l \tag{6}$$

where CPF_{kl} is the conditional probability function of source k in sector l (no units), m_{kl} is the number of hours that the contribution of source k from sector l exceeded the threshold criterion, and n_l is the number of hours that wind is from sector l. The hourly contribution from the source was assumed as the same as the 24-hourly average contribution. The contribution threshold was chosen as the 75th percentile value of the contribution from the source. Calm wind periods were excluded from the DRS and CPF analyses.

3. Results and discussions

3.1. Identified source types and source profiles

In order to assess the effect of inclusion of multiple type composition data and wind abundance data on source apportionment of air pollutants, source apportionment using PMF was trialled using the data sets consisting of coarse particles only, fine particles only, VOC only, fine particles + VOC and fine particles + VOC + gaseous pollutants. In addition, the inclusion of data of wind abundance in the 16 sectors and the calm wind sector was also trialled. The source profiles from the trials were compared against each other. The values of statistical indicators and diagnostic plots were also compared against the fit diagnostic guidelines.

For the PMF runs, the results were found to be better when multiple type data were used, based on the values of the statistical indicators, the diagnostic plots and the relevance of the resolved factors to known sources in the area. The results were further improved when wind data were also included. For example, the ratio of the robust Q value to the theoretical Q value was 1.12 (3 factors were resolved) in the coarse particle with wind data set run, compared to a ratio of 1.47 (3 similar factors resolved) in the run without wind data. For the fine particles + VOC + gaseous pollutants and wind data set run, this ratio was 1.00 (7 factors resolved), compared to those of 1.11 (7 factors resolved but with one of them

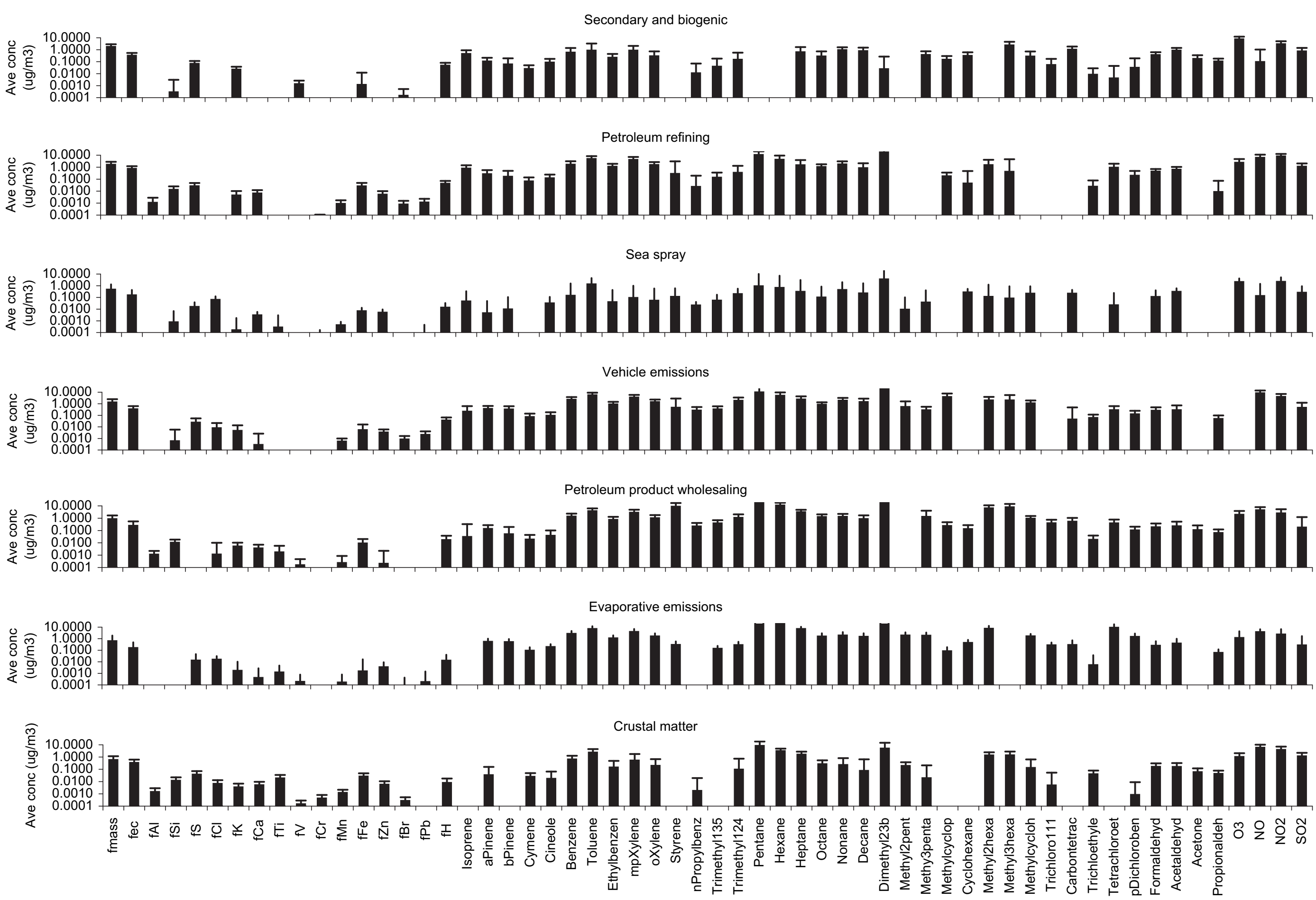

Fig. 4. Source profiles of fine particles, VOC and gaseous pollutants from PMF model analysis.

Table 2
Average contribution (in μg m^{-3}) and emission loading (%) of sources.

Source types	PMF 3.0 Coarse particles	PMF 3.0 Fine particles	PMF 3.0 PM_{10} [a]	NPI PM_{10} emission loading % [b]	PMF 3.0 12 NPI VOCs	NPI 12 VOCs emission loading % [b]
Secondary and biogenic	–	1.7 (23.6%)	1.7 (9.8%)	2.3%	4.5 (2.7%)	0.6%
Evaporative emissions	–	0.6 (8.5%)	0.6 (3.5%)		47.6 (28.2%)	63.9%
Vehicle emissions	3.5 (33.9%)	1.3 (18.3%)	4.8 (27.4%)	12.9%	19.8 (11.8%)	
Petroleum refining	–	1.6 (22.3%)	1.6 (9.3%)	57.9%	19.7 (11.7%)	16.6%
Petroleum product wholesaling	–	0.8 (11.6%)	0.8 (4.8%)	0.2%	29.9 (17.7%)	9.6%
Others sources	0.1 (1.2%)	0.1 (1.1%)	0.2 (1.1%)	12.7%	36.4 (21.6%)	9.3%
Sea spray	2.4 (23.3%)	0.5 (6.5%)	2.9 (16.3%)	–	3.2 (1.9%)	–
Crustal matter	4.3 (41.7%)	0.6 (8.1%)	4.9 (27.8%)	13.2%	7.3 (4.3%)	–
Total mass	10.2 (100%)	7.2 (100%)	17.5 (100%)	100%	168 (100%)	100%

[a] Compiled from the coarse particle and fine particle results.
[b] Based on the NPI PM_{10} and VOC emission loading data (DEWHA, 2010).

representing mixed source types) in the run without wind data, 1.19 (7 factors resolved with 2 of them representing mixed factors) in the fine particle + VOC data only run and 1.45 (only 4 factors resolved and all were mixed factors) in the fine particle data only run.

The source profiles from the PMF analysis of the coarse particles and wind data set and the fine particles + VOC + gaseous pollutants and wind data set are presented in Fig. 3 and Fig. 4, respectively. The source profiles from the coarse particle data only run are also included in Fig. 3 for comparison. The bars in the figures represent the estimated uncertainty of each species in the source profiles. The Discrete Difference Percentile (DDP) method recommended by the PMF and Unmix models (USEPA, 2007,2008) for small data sets was adopted in this study to estimate the uncertainty of source profiles. In this method the uncertainty of a factor loading value is taken as the difference between the 90th percentile value of the factor loading from the 100 bootstrap runs, and the factor loading value from the base run.

Three source factors were identified for the coarse particle samples. The first factor has high factor loading values for EC, H, Fe, Cu and Zn, indicating a 'vehicle emissions' origin. The high EC content indicates substantial contribution from the diesel fuelled vehicles prevalent in this industrial area. The 'crustal matter' factor is characterised by enrichment with Si, Al, Ca, Ti, Fe and EC. The abundance of EC in this profile suggests mixing of soil dust with road side dust. Some of the Ca, Ti and Fe content could also originate from the cement manufacturing and ilmenite/TiO_2 refining industries in the area. The 'sea spray' factor has elevated values of Na, Cl and S. The lower Cl/Na ratio of 1.23 in the 'sea spray' factor compared to that in surface sea water of 1.8 (Goldberg, 1963) indicates partial loss of Cl. The S content in the source profile also indicates replacement of chloride by sulfate.

PMF analysis of the fine particle/VOC/gaseous pollutants and wind data set resolved seven source factors. In this case, the 'crustal matter' factor has high loading values for Si, Ca, Ti, Fe, EC, SO_2 and NO but low values for the VOCs. This source profile indicates contributions from road side dust (indicated by the high EC and NO content) and the cement manufacturing and ilmenite/TiO_2 refining industries (due to the high Ca, Ti, Fe and SO_2 content, with the SO_2

Table 3
Average contribution (in %) from major source types, compiled from the results reported in recent source apportionment studies using PMF techniques.

Study	No. of samples	No. of species analysed	No. of factors	Ave Concentration (μg m^{-3})	Marine %	Crustal %	Vehicles/ evaporative %	Other combustion/ industry %	Secondary/ biogenic %	Unexplained %	Reference
Eagle Farm, Brisbane, Australia, June 2001	28 coarse	1	3	10.2	2	40	32	–	–	3	This study
	28 fine	817	7	7.2	58	8	30	32	22	0	
4 Australian cities, 8 sites, 2003–2004	437 coarse	23	4	9.8	38	38	5	–	19	0	Chan et al., 2008a
	437 fine	23	8	5.9	20	6	24	22	25	3	
Hanoi, Vietnam, 2001–2008	784 fine	21	6	54.0	–	3	40	49	8	–	Cohen et al., 2010
Bandung, Indonesia, dry season 2001–2007	180 coarse	10	5	19.0	23	19	–	40	18	–	Lestari and Mauliadi, 2009
	180 fine	10	7	48.0	13	–	22	40	25	–	
Athens, Greece, 3 sites, 2002	62 coarse	17	3	18.0	20	30	50	–	–	–	Karanasiou et al., 2009
	67 fine	17	5	41.0	24	18	34	24	–	–	
Oslo, Norway, 1 site, 2004–2005	78 fine	37	6	26.0	–	18	32	27	23	–	Laupsa et al., 2009
Barcelona, Spain, 2003–2007	243 coarse	26	8	12.6	29	67	28	2	3	−29	Amato et al., 2009
	279 fine	26	8	27.7	3	14	29	10	45	−1	
Chicago, Illinois, US, 2 sites, 2001–2003	372 fine	57	10	15.5	3	6	23	12	58	−2	Rizzo and Scheff, 2007
Southwestern Oregon, US, 2000–2004	493 fine	32	9	3.2	10	9	3	38	40	–	Hwang and Hopke, 2007
Midwest US, 5 sites, 2002–2005	1207 fine	26	7–9	14.5	–	5	30	14	51	–	Buzcu-Guven et al., 2007
Eagle Farm, Brisbane, Australia, June 2001	28 VOC	38	7	168	2	4	45	27	2	20	This study
3 Australian cities, 6 sites, 2003–2004	353 VOC	14	5	21.9	–	–	61	18	14	7	Chan et al., 2008b
Hong Kong, China, 4 sites, 2002–2003	497 VOC	51	9	19.3 (ppbv)	–	–	64	13	23	–	Lau et al., 2010
Los Angeles, US, 2 sites, 2001–2003	1100 VOC	30	6	98.8	–	–	83	15	2	–	Brown et al., 2007

possibly originating from the use of fuel oil). The ‘sea spray’ factor is characterised by a high Cl content.

The ‘evaporative emissions’ and the ‘vehicle emissions’ factors both contain enhanced loading values for benzene, toluene, xylenes, C_5–C_{10} alkanes and NO. The high EC content and the relatively large content of the higher molecular weight alkanes and NO in the latter factor suggest a tail pipe exhaust origin of VOCs. The high content of cyclohexane and tetrachloroethylene in the ‘evaporative emissions’ factor also indicates evaporation of VOCs from architectural surface coatings and from the dry cleaning industry (DEWHA, 2010). The high content of trichloroethylene in the ‘vehicle emissions’ factor also is an indication of contribution from solvent use (DEWHA, 2010).

The ‘petroleum refining’ factor has elevated loadings of SO_2, hydrocarbons and EC. In addition, the high content of formaldehyde and acetaldehyde indicates mixing with emissions from airport activities. Petroleum refining facilities and the Brisbane Airport are both located in an NE direction to the sampling station (DEWHA, 2010) (Fig. 1). The ‘petroleum product wholesaling’ factor is enriched with hydrocarbons and styrene (DEWHA, 2010). There is a major petroleum product storage and wholesaling facility located close to the sampling station (DEWHA, 2010) (Fig. 1) and a couple of service stations in the NW and W directions. The ‘secondary and biogenic’ factor is characterised by S, H, O_3, propionaldehyde and isoprene, suggesting biogenic emissions as well as the formation of secondary sulfates and smog products as possible sources (Queensland EPA, 2003). The ratio of H/S of 0.71 in the ‘secondary and biogenic’ factor is much higher than the theoretical value of 0.25 based on $(NH_4)_2SO_4$. This may be due to the fact that the hydrogen content of urban particulate matter also exists in the form of ammonium nitrate and organic compounds other than those analysed in this study.

3.2. Contribution of sources to particulate matter and gaseous pollutants

The average contributions of the source types to the ambient concentration of coarse particles, fine particles and the 12 NPI VOCs in the samples from the PMF analysis are listed in Table 2. For the coarse particle samples, PMF analysis shows that ‘crustal matter’ contributed to about 42% of the aerosol mass on average, followed by ‘vehicle emissions’ (34%), ‘sea spray’ (23%) and unexplained mass (1%). For the fine particle samples, ‘secondary and biogenic’ contributed to about 24% of the aerosol mass on average, followed by ‘petroleum refining’ (22%), ‘vehicle emissions’ (18%), ‘petroleum product wholesaling’ (12%), ‘evaporative emissions’ (9%), ‘crustal matter’ (8%), ‘sea spray’ (6%) and unexplained mass (1%). For the VOC samples, ‘evaporative emissions’ VOCs contributed to about

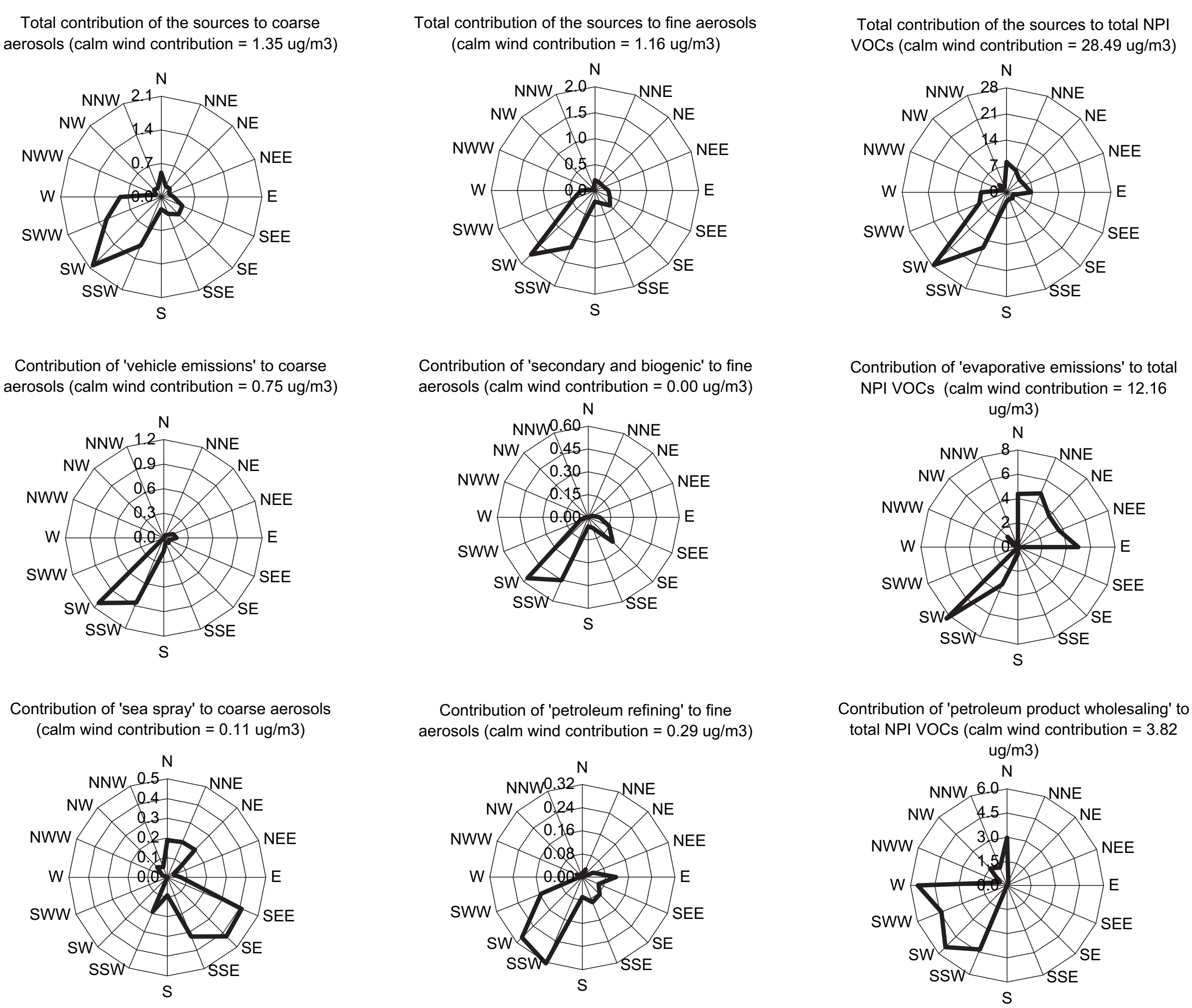

Fig. 5. Average contribution of all the identified sources in total and major source factors to the ambient concentration of coarse particles, fine particles and the 12 NPI VOCs by wind sector.

28% of the total mass of the 12 NPI VOCs on average, followed by 'petroleum product wholesaling' (18%), 'vehicle emissions' (12%), 'petroleum refining' (12%), 'crustal matter' (4%), 'secondary and biogenic' (3%), 'sea spray' (2%). A higher percentage of the VOC mass was unexplained (22%), compared to that for the particulate matter. Unlike particulate matter, the tracer VOC species often exist abundantly in more than one source. For example, Na and Cl are mostly from sea spray contribution, while Si and Al mostly originate from crustal matter in urban aerosols. However, VOC tracers such as aromatic compounds and hydrocarbons can originate from various sources including vehicle emissions and industrial activities. The inclusion of more tracer species such as methane, C_2–C_4 compounds and gaseous phase polycyclic aromatic hydrocarbons (PAHs) in the sampling program may be useful in obtaining more accurate VOC source apportionment results.

The PMF modelled result of percentage contribution of the 'petroleum refining' factor to fine particle mass (9%) was much lower than the percentage of emission loading of the petroleum refining industry (58%) (Table 2). This is probably due to the effective dispersion of fine particles emitted from the tall stacks of the refineries. When the tall stack sources and the natural sources were excluded from the comparison, the PMF modelled results were roughly comparable to the emission inventory data.

Table 3 summarises the average contribution from major source types in this study and some other recent source apportionment studies using PMF techniques. PMF software including PMF 2 (Paatero, 2004), PMF 1.1 (USEPA, 2005) and PMF 3.0 (USEPA, 2008) have been used in the studies. As shown in Table 3, the number and type of source factors resolved from PMF analysis in this work is similar to those reported in the other studies. Most of the studies have included hundreds of samples and more than 10 species in the analysis. In general, 3–8 source factors were resolved for coarse particles, 5–10 factors for fine particles and 5–9 factors for VOCs. Despite the differences in site characteristics and sampling periods in the studies, a similar set of major source types have been consistently identified. For example, crustal matter, sea spray (for coastal sites), vehicle emissions and secondary aerosols were the major source types identified for particulate matter. On the other hand, evaporative and vehicular emissions and industrial emissions were the major source types identified for VOCs. These findings show that the combined use of multiple type composition data and wind data in PMF analysis resolves source factors which are similar to those reported in the literature where larger data sets were used. In this way, the PMF model can be applied to small data sets such as the one in this study and enable useful results to be obtained from small scale investigations of localised pollution problems.

3.3. Mass contribution of source factors by wind sector

In this study, the estimation of average mass contributions of source factors by wind sector were based on the values of wind sector factor loading from the PMF analysis (Section 2.3). The results are shown in Fig. 5 for the major source factors as well as for the total contribution from the source factors. Since the prevailing winds were in the SW to SSW sectors (Fig. 2), the directional contributions of most of the source factors were also dominated by

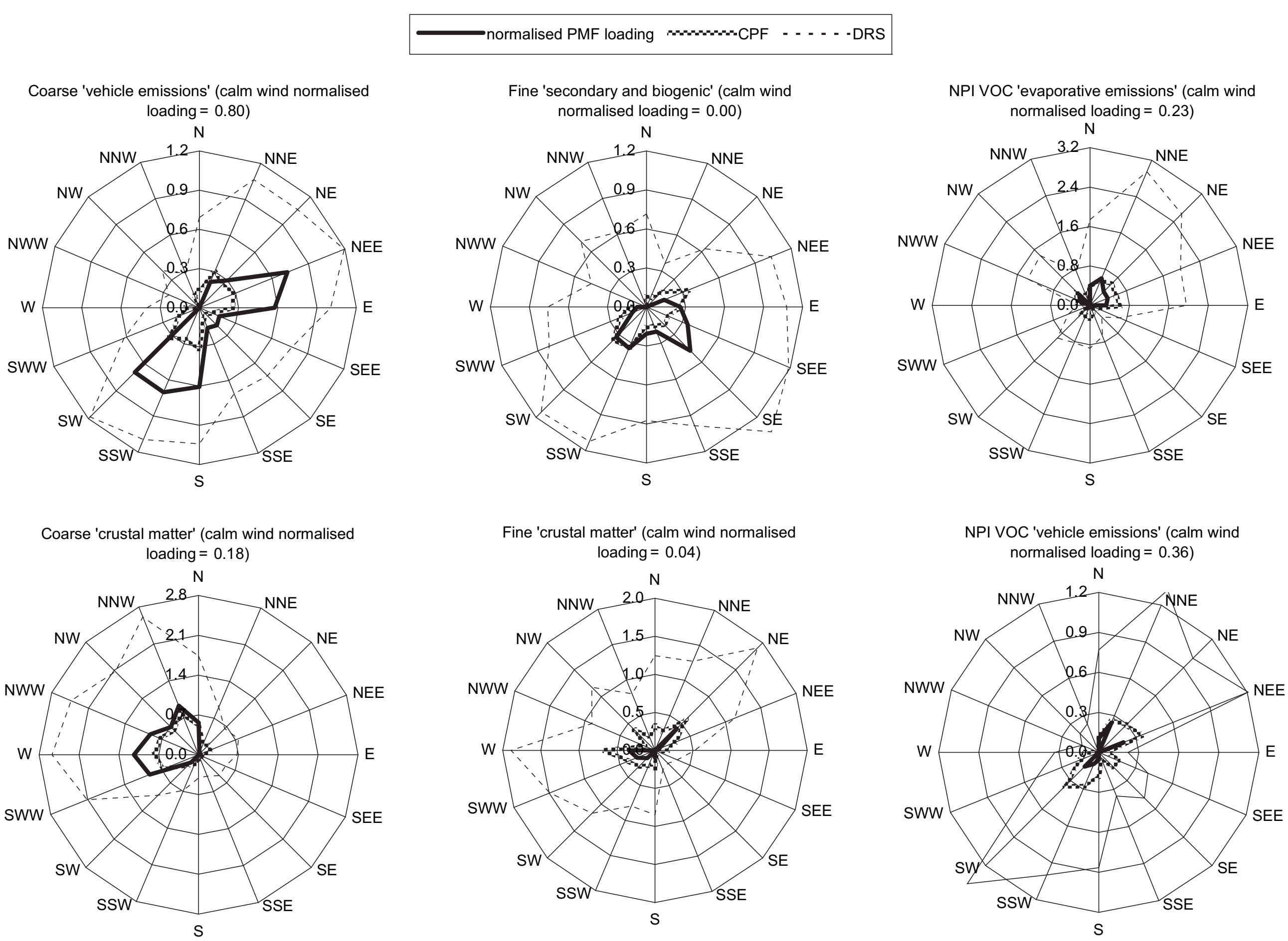

Fig. 6. Directional dependences of example source factors to the ambient concentration of coarse particles, fine particles and the 12 NPI VOCs.

these wind sectors as expected. The only exception was the 'sea spray' factor for which the dominant directional contributions were in the NE and SE sectors. This confirms easterly sea breezes as the major contributing factor for sea spray. Although the prevailing south-westerly winds might recirculate part of the sea spray back into the area, the amount was still small comparing to the easterly wind contribution. Fig. 5 also shows that on average, the calm wind sector contributed 1.4 μg m^{-3} (13%) of the coarse particle mass, 1.2 μg m^{-3} (16%) of the fine particle mass and 29 μg m^{-3} (17%) of the NPI VOC mass.

The DDP values of the wind sector factor loadings also show that the uncertainty of the predicted directional contribution for the small data set in this study ranged from about 60% for the prevailing wind sectors, to about 300% for the other wind sectors.

3.4. Directional dependences of source factors

In this study the normalised wind sector factor loadings were used to assess the directional dependences of the source factors and to locate the source factors. The normalised wind sector factor loadings were also compared to the DRS and CPF indices which have been used in recent research for the same purpose (Section 2.4). The normalised wind sector factor loadings are plotted against the DRS and CPF indices for example source factors in Fig. 6. As shown here, the three indicators basically indicate a similar pattern of directional dependences of the source factors.

The directional dependences of the source factors (Fig. 6) matched well with the location of known sources in the area (Fig. 1). The coarse 'crustal matter' factor was related to westerly winds, due to westerly land breezes bringing in soil dust and re-entraining road side dust from the city. On the other hand, the fine 'crustal matter' factor was also related to winds from NE, possibly due to contributions from the mineral refinery, concrete manufacturing and fertiliser manufacturing facilities in that direction. This factor was also related to winds from W, probably due to contribution from the glass manufacturing facility to that direction. The 'vehicle emissions' factors were related to winds from NE and SW directions – there is heavy city traffic to the SW and a major road nearby to the NE. The 'secondary and biogenic' factor was related to winds from SW and SE, probably due to emissions from forest and vegetation to the west, and the recirculation of secondary sulfates and O_3 formed from previous days. The 'evaporative emissions' factor was related to winds broadly from the N to E directions – there is a dry cleaning and textile facility nearby to the NE and an adhesive manufacturing facility to the E. This factor was also related to winds from NW, suggesting a contribution from the food manufacturing facility to that direction.

There was also a higher contribution potential from calm winds for local sources compared to that for distant sources. For example, the normalised factor loading of the calm wind sector for coarse particles was 0.80 for 'vehicle emissions', compared to 0.18 for 'crustal matter' and 0.08 for 'sea spray'. These findings show that the PMF analysis can also be used to estimate the mass contribution of sources by wind sector, to investigate the locations of the sources and to estimate the uncertainty of the results.

4. Conclusion

A possible way of extracting more source information from an aerosol composition data set is to include data of other air pollutants and wind data in the analysis. The use of multiple type composition data for source apportionment is currently a matter of debate, while the direct combination of wind data with composition data in receptor modelling has not been reported in the literature. In this study a small but comprehensive data set from a 24-hourly sampling program carried out during June 2001 in an industrial area in Brisbane was chosen to investigate the use of multiple type composition data (aerosols, VOCs and major gaseous pollutants) and wind data in source apportionment of air pollutants using the USEPA's PMF 3.0 model. The values of statistical indicators and diagnostic plots were compared against the fit diagnostic guidelines specified in the user manual of the model. The use of combined data was found to result in more well defined source factors and better fit diagnostics, compared to when non-combined data were used.

Three source factors were identified by the PMF model for the coarse particle samples viz. 'crustal matter', 'vehicle emissions' and 'sea spray'. Seven factors were identified for the fine particle and VOC samples. These were 'secondary and biogenic', 'petroleum refining', 'vehicle emissions', 'petroleum product wholesaling', 'evaporative emissions', 'sea spray' and 'crustal matter'. The number of source factors resolved was similar to those reported in the literature where larger data sets were used.

The factor loadings of the wind sectors from the PMF analysis were used to estimate the mass contribution of source factors by wind sector and to estimate the uncertainty of the results. While the contributions were dominant in the prevailing winds as expected, calm winds were found to also contribute up to 17% of the pollutant mass on average. The normalised factor loadings of the wind sectors were also used to assess the directional dependences of the source factors. The results of directional dependences from the PMF analysis were similar to those from use of the DRS and CPF indices. The results of directional dependences matched well with the location of known sources in the area. There was also a higher contribution potential from calm winds for local sources compared to that for distant sources.

The findings from this study indicate that with the use of multiple type composition data and wind data, PMF analysis could be applied to small data sets to enable source location and apportionment for air pollutants in small scale investigations of localised pollution problems.

Acknowledgments

We would like to acknowledge the assistance of staff from Queensland Department of Environment and Resource Management and the Department of Primary Industry with sampling and data collection. The project was funded by an Australian Institute of Nuclear Science and Engineering Grant 01/029.

References

Amato, F., Pandolfi, M., Escrig, A., Querol, X., Alastuey, A., Pey, J., Perez, N., Hopke, P.K., 2009. Quantifying road dust resuspension in urban environment by multilinear engine: a comparison with PMF2. Atmospheric Environment 43, 2770–2780.

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

Buzcu-Guven, B., Brown, S.G., Frankel, A., Hafner, H.R., Roberts, P.T., 2007. Analysis and apportionment of organic carbon and fine particulate matter sources at multiple sites in the midwestern United States. Journal of Air and Waste Management Association 57, 606–619.

Chan, Y.C., Christensen, E., Golding, G., King, G., Gore, W., Cohen, D.D., Hawas, O., Stelcer, E., Simpson, R., Denison, L., Wong, N., 2008b. Source apportionment of ambient volatile organic compounds in major cities in Australia by positive matrix factorisation. Clean Air and Environmental Quality 2/42, 22–29.

Chan, Y.C., Cohen, D.D., Hawas, O., Stelcer, E., Simpson, R., Denison, L., Wong, N., Hodge, M., Comino, E., Carswell, S., 2008a. Apportionment of sources of fine and coarse particles in four major Australian cities by positive matrix factorisation. Atmospheric Environment 42, 374–389.

Cohen, D.D., Crawford, J., Stelcer, E., Bac, V.T., 2010. Characterisation and source apportionment of fine particulate sources at Hanoi from 2001 to 2008. Atmospheric Environment 44, 3320–3328.

DEWHA, 2010. National Pollutant Inventory. Accessible from the Department of the Environment, Water, Heritage and the Arts web site: http://www.npi.gov.au. Accessed in April 2010.

Dreyfus, M.A., Adou, K., Zucker, S.M., Johnston, M.V., 2009. Organic aerosol source apportionment from highly time-resolved molecular compound measurements. Atmospheric Environment 443, 2901–2910.
Goldberg, E., 1963. Chemistry – the oceans as a chemical systems. In: Hill, H. (Ed.), Composition of Sea Water, Comparative and Descriptive Oceanography. The Sea, vol. 2. Wiley-Interscience, New York, pp. 3–25.
Hawas, O., 2001. Chemical composition of air pollutants in an industrial area in Brisbane. Honours thesis. Australian School of Environmental Studies, Griffith University, Australia.
Hawas, O., Hawker, D., Chan, A., Cohen, D., Christensen, E., Golding, G., Vowles, P., 2002. Sampling and analysis of ambient volatile organic compounds (VOCs) in an industrial area in Brisbane, Australia. Clean Air and Environmental Quality 36/4, 40–45.
Hopke, P.K., 2010. Discussion of "Sensitivity of a molecular marker based positive matrix factorisation model to the number of receptor observations" by Yuan-Xun Zhang, Rebecca J. Sheesley, Min-Suk Bae and James J. Schauer. Atmospheric Environment 44, 1138.
Hwang, I., Hopke, P.K., 2007. Estimation of source apportionment and potential source locations of $PM_{2.5}$ at a west coastal IMPROVE site. Atmospheric Environment 41, 506–518.
Karanasiou, A.A., Siskos, P.A., Eleftheriadis, K., 2009. Assessment of source apportionment by positive matrix factorisation analysis on fine and coarse urban aerosol size fractions. Atmospheric Environment 43, 3385–3395.
Lau, A.K.H., Yuan, Z., Yu, J.Z., Louie, P.K.K., 2010. Source apportionment of ambient volatile organic compounds in Hong Kong. Science of the Total Environment 408, 4138–4149.
Laupsa, H., Denby, B., Larssen, S., Schaug, J., 2009. Source apportionment of particulate matter ($PM_{2.5}$) in an urban area using dispersion, receptor and inverse modelling. Atmospheric Environment 43, 4733–4744.
Lestari, P., Mauliadi, Y.D., 2009. Source apportionment of particular matter at urban mixed sites in Indonesia using PMF. Atmospheric Environment 43, 1760–1770.
Paatero, P., 2004. User's Guide for Positive Matrix Factorisation Programs PMF2. YP-Tekniika KY Company.
Pandolfi, M., Viana, M., Minguillon, M.C., Querol, X., Alastuey, A., Amato, F., Celades, I., Escrig, A., Monfort, E., 2008. Receptor models application to multi-year ambient PM_{10} measurements in an industrialized ceramic area: comparison of source apportionment results. Atmospheric Environment 42, 9007–9017.
Queensland EPA, 2003. Air Emission Inventory South-east Queensland Region. Queensland Environmental Protection Agency and Brisbane City Council.
Rizzo, M., Scheff, P.A., 2007. Fine particulate source apportionment using data from the USEPA speciation trends network in Chicago, Illinois: comparison of two source apportionment models. Atmospheric Environment 41, 6276–6288.
USEPA, 2004. EPA CMB 8.2 Users Manual. USEPA Office of Air Quality Planning and Standards.
USEPA, 2005. EPA PMF 1.1 User's Guide. USEPA National Exposure Research Laboratory.
USEPA, 2007. EPA Unmix 6.0 Fundamentals and User Guide. USEPA Office of Research and Development.
USEPA, 2008. EPA Positive Matrix Factorisation (PMF) 3.0 Fundamentals and User Guide. USEPA Office of Research and Development.
Wu, C., Larson, T.V., Wu, S., Williamson, J., Westberg, H.H., Liu, L.J.S., 2007. Source apportionment of $PM_{2.5}$ and selected hazardous air pollutants in Seattle. Science of the Total Environment 386, 42–52.
Yue, W., Stolzel, M., Cyrys, J., Pitz, M., Heinrich, J., Kreyling, W.G., Wichmann, H.E., Peters, A., Wang, S., Hopke, P.K., 2008. Source apportionment of ambient fine particle size distribution using positive matrix factorisation in Erfurt, Germany. Science of the Total Environment 398, 133–144.
Zhou, L., Hopke, P.K., Stanier, C.O., Pandis, S.N., Ondov, J.M., Pancras, J.P., 2005. Investigation of the relationship between chemical composition and size distribution of airborne particles by partial least squares and positive matrix factorisation. Journal of Geophysical Research 110 (D07S18). doi:10.1029/2004JD005050.