AIRBORNE PARTICULATE MATTER POLLUTION IN THE AUSTRALIAN CAPITAL TERRITORY

Technical Report – Chemical Composition and Source Contribution Analysis

M.G. Froehlich, S. Chatterjee, M. Li, T. DeSilva, J. McDonald, S. Rockliff, I. Firkins-Fox[*]

Environmental Chemistry, ACT Government Analytical Laboratory, Health Protection Service, ACT Health
25 Mulley Street, Holder, ACT 2611.

SCOPE

This document is a *technical report* on both the chemical elemental analysis and source-factor contribution analysis of airborne fine- and course particulate matter (PM) sampled at the Monash air quality monitoring station located in the Tuggeranong valley, ACT. Its main scope is to summarise and provide any necessary procedural details, numerical parameters used in the mathematical models, and their immediate results in a technical/scientific manner to allow interested readers to reproduce our findings. An *executive summary* outlining the most important details and results in a general comprehensible way will be published separately.

INTRODUCTION AND AIM

Atmospheric aerosols and airborne particulate matter (PM) are highly variable components of the lower atmosphere and play a number of important roles in environmental issues related to climate, air quality, visibility degradation, atmospheric chemistry, and (adverse) human health effects.[1] The health risks posed by PM are based on a number of factors, but mainly due to particle size, surface area, and chemical composition.

Elevated concentrations of PM (both PM_{10} aerodynamic diameter of less than 10 μm, and $PM_{2.5}$ aerodynamic diameter of less than 2.5 μm,) are of particular concern in Canberra. Composition of the PM is of community concern especially when elevated concentrations occur. It is known that the size and composition of both coarse aerosols (PM_{10}) and fine particulate matter ($PM_{2.5}$) are directly linked to their sources – which are mainly fuel combustion processes (transportation, energy production, residential heating), secondary aerosols (referred to as secondary sulfates, comprised of ammonium sulfate and nitrate resulting from atmospheric photochemistry), traffic related pollution (tyre abrasion,

[*] Corresponding author: Ian Firkins-Fox, ian.fox@act.gov.au, Locked Bag 5005, Weston Creek, ACT 2611.

resuspended road dust), and natural sources (crustal dust, sea salt, plant material). Therefore, source identification and composition analysis of size-segregated particulate matter is crucial in establishing and monitoring source-specific health risks for an exposed population as well as to inform policy makers to aid in the development of suitable legislation for improved air quality management within an airshed.

A previous report (Preliminary Assessment of Wintertime Air Quality in the Tuggeranong Valley, ACT)[2] recommended that elemental analysis be undertaken to help confirm the sources of PM pollution. The scope of this project is to provide detailed information not only on the chemical composition of coarse and fine airborne particulate matter present in the Canberra airshed and the Australian Capital Territory, but also, if possible, to derive information on the origins and contributing factors to these pollutants.

DATASET AND METHOD

For the purpose of this study a small number of $PM_{2.5}$ and PM_{10} filters (Teflon® 2 µm pore size membrane filters, monthly counts summarised in the table on the right) – which were routinely collected within the **ACT Ambient Air Quality Monitoring Program** framework using Thermo Scientific Partisol™ 2025Plus/2025i Sequential Air Samplers at the **Monash** monitoring station (35°25'05.8"S, 149°05'38.5"E) in Canberra, ACT – have been sent off to the Institute for Environmental Research at the Australian Nuclear Science and Technology Organisation (ANSTO, Sydney) for accelerator based Ion Beam Analysis (IBA).[3]

Month	# $PM_{2.5}$	# PM_{10}
May-14	5	
Jun-14	5	
Jul-14	1	1
Aug-14	4	4
Sep-14		
Oct-14	7	7
Nov-14	5	
Dec-14	1	
Jan-15		
Feb-15	1	
Mar-15	2	
Apr-15		
May-15		
Jun-15	3	
Jul-15	2	
Aug-15	1	
Σ	**37**	**12**

Accelerator based IBA is a collective term for techniques in which high energy ion beams, typically in the range of MeV, are used to non-invasively probe the elemental compositions (and depth profiles) in near-surface layers of solid samples. A detailed description of ion beam analysis is outside the scope of this report; information and technical details on accelerator based ion beam analysis itself,[4] and on its applications for PM elemental analysis, can be found in the literature.[5]

In this particular case the collected membrane filters have been subjected to **particle-induced X-ray emission** (PIXE)[6] spectrometry for the detection and quantitative analyses of the elements F, Na, Al, Si, P, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Se, Br, Sr, and Pb, and subsequently to **particle-induced gamma-ray emission** (PIGE)[7] spectrometry for the additional quantification of H which serves as a measure for the concentration of organic matter. The total $PM_{2.5}$ and PM_{10} mass concentrations have been determined gravimetrically by weighing the filters pre and post sampling, using a digital balance (Mettler Toledo® XP6 Microbalance, capacity 0.08 mg - 6.1 g, sensitivity ± 1 µg) under strictly controlled conditions at the ACT Government Analytical Laboratory. The sampling period for each filter was 24 h at a nominal flow rate of ~16.7 L/min. Exact sample volumes (~21-

24 m^3) were automatically measured by the Partisol™ 2025Plus/2025i air samplers, and individual filter (normalised to 0°C) readings were used to express all concentrations in mass per volume.

Using elemental and total mass concentrations as input, two mathematical procedures have been applied to determine PM compositions and furthermore to estimate pollutant sources and their seasonal contributions:

- Correlation analysis and known stoichiometric relations, using the equations given in the *interagency monitoring of protected visual environments* (IMPROVE) program.
- *Positive matrix factorisation* (PMF) for a more detailed and comprehensive statistical analysis of source factors and source factor contributions.

More information on the individual methods will be presented together with a discussion of their results in the next section.

RESULTS AND DISCUSSION

CHEMICAL COMPOSITION ANALYSIS – ELEMENT CONCENTRATIONS

Primary results of the chemical (PIXE & PIGE) analysis, *i.e.*, average element concentrations over all samples, are summarised on the next two pages in Table 1 for fine aerosols ($PM_{2.5}$) and Table 2 for coarse aerosols (PM_{10}). All element concentrations (given in the upper sections of the tables) are reported in **ng/m³**. The sample mean values and their respective 95% confidence intervals (second column) are highlighted; element species for which a majority of individual results are significantly below the method detection limits (MDL, right columns) are greyed out – special care should be taken when interpreting these numbers. Additionally, the population median (Med), population standard deviation (SD), and population maxima (Max) are reported for each element.

It is evident that average concentrations of crustal elements (Si, Al, Fe, Ca, K, Na, Ti) are significantly higher for the (coarse) PM_{10} fraction – hinting to a (not unexpectedly) distinct contribution of **wind-blown soil**. In general, (aged) **sea salt** (NaCl) from sea spray is a significant factor only in marine environments, but there are Cl concentrations in the PM_{10} fraction indicating a sea salt contribution to the aerosols in the ACT. The amount of sulfur is about the same in both the fine and coarse PM fractions. This could be attributed to the small size of **secondary sulfate** aerosols (around 0.1–1 μm),[8] which either pass through the filters or adsorb to primary aerosols in about the same ratio. The hydrogen on the filters is associated with sulfates, nitrates, water, and mainly **organic material** – and again, average quantities are roughly on the same level for $PM_{2.5}$ and PM_{10}.

The lower sections in Table 1 and Table 2 summarise the results of composition analyses in **μg/m³** based on the IMPROVE program, those will be discussed in detail later on.

$PM_{2.5}$	Average			Med	SD	Max	MDL
Na	**131**	±	**115**		346	1197	703
Al	**35**	±	**10**	26	30	146	20
Si	**120**	±	**29**	101	86	410	13
P	**6**	±	**1**	6.4	2.9	16.5	11
S	**243**	±	**50**	192	149	692	10
Cl	**89**	±	**30**	78	89	366	10
K	**114**	±	**25**	120	76	315	10
Ca	**9**	±	**3**	7	9	37	9
Ti	**3**	±	**1**	2	3	10	6.9
V	**0.9**	±	**0.4**	0.5	1.2	5.6	5.8
Cr	**2.1**	±	**0.5**	1.6	1.6	7.5	4.6
Mn	**1.8**	±	**0.5**	1.7	1.5	5.8	3.8
Fe	**48**	±	**9**	43	26	132	3
Co	**0.7**	±	**0.3**	0.0	0.9	2.9	4.5
Ni	**0.5**	±	**0.3**	0.5	0.8	3.3	3.1
Cu	**2.8**	±	**0.6**	2.7	1.9	9.4	3.4
Zn	**7**	±	**2**	5	6	23	3
Se	**2.0**	±	**0.8**	1.1	2.5	11.8	10
Br	**5**	±	**2**	5	5	24	14
Pb	**9**	±	**2**	10	6	20	23
H	**556**	±	**119**	567	357	1702	11
Soil	**0.5**	±	**0.1**	0.4	0.3	1.6	
AS	**1.0**	±	**0.2**	0.8	0.6	2.8	
OMH	**5.7**	±	**1.3**	5.6	3.9	17.6	
NaCl	**0.3**	±	**0.3**		0.9	3.1	
K_{non}	**0.20**	±	**0.06**	0.24	0.16	0.58	
RCM	**7.6**	±	**1.1**	8.0	4.1	20.6	
% RCM	**53**	±	**4**	50	11	90	
Mass	**15.1**	±	**2.9**	15.3	8.9	41.6	

Table 1 *Top section*: Total average $PM_{2.5}$ element concentrations and their most important statistical parameters in ng/m^3; columns from left: average value ± 2 × standard error of the mean, population median (Med), population standard deviation (SD), population maximum (Max), and method detection limit (MDL). Shaded data indicate average concentrations significantly below the respective method detection limits, thus care should be taken when interpreting these numbers.
Bottom section: *S*ource contributions and their statistical parameters estimated from element concentrations, for details see later section "Source contribution – IMPROVE algorithm". Rows from top: soil, ammonium sulfate (AS, secondary sulfate), organic matter from hydrogen (OMH), sea salt aerosol from sea spray (NaCl), non-soil potassium (K_{non}, mass calculated as K_2O), reconstructed mass (RCM), overall RCM percentage, and gravimetric filter mass, all in μg/m^3.

PM_{10}	Average			Med	SD	Max	MDL
Na	72	±	159		251	869	672
Al	**125**	±	**55**	103	86	276	19
Si	**481**	±	**179**	423	282	971	12
P	**5**	±	**1**	4	2	8	10
S	**214**	±	**74**	207	116	442	9
Cl	**222**	±	**145**	160	228	677	9
K	**112**	±	**39**	111	62	221	8
Ca	**51**	±	**17**	49	26	100	8
Ti	**10**	±	**4**	10	6	21	7
V	**0.6**	±	**0.5**	0.2	0.7	2.0	5.9
Cr	**1.8**	±	**0.8**	1.7	1.3	4.9	4.4
Mn	**4**	±	**1**	4	2	9	4
Fe	**143**	±	**51**	135	80	281	3
Co	**0.8**	±	**0.7**	0.2	1.1	3.0	3.5
Ni	**0.3**	±	**0.3**		0.4	1.0	3.0
Cu	**6**	±	**3**	5	4	13	3
Zn	**7**	±	**3**	4	5	16	3
Se	**0.9**	±	**0.9**		1.5	3.9	8.9
Br	**7**	±	**3**	6	5	16	12
Pb	**10**	±	**5**	7	9	30	21
H	**334**	±	**139**	259	219	784	9
Soil	**1.9**	±	**0.7**	1.8	1.1	3.7	
AS	**0.9**	±	**0.3**	0.9	0.5	1.8	
OMH	**3.1**	±	**1.5**	1.8	2.3	7.7	
NaCl	**0.2**	±	**0.4**	0.0	0.6	2.2	
K_{non}	**0.06**	±	**0.06**	0.03	0.09	0.26	
RCM	**6.1**	±	**2.0**	6.3	3.1	11.2	
% RCM	**56**	±	**18**	50	25	55	
Mass	**11.9**	±	**5.0**	11.7	7.4	22.4	

Table 2 PM_{10} element concentrations and sample statistics in ng/m^3 (top section); results of source contribution analysis in µg/m^3 (bottom section). For a detailed description of columns and rows see caption of Table 1.
For some species in this table (Na, Ni, Se) not enough data was available to calculate any meaningful medians; this also applies to Na and NaCl in Table 1.

	H	Na	Al	Si	P	S	Cl	K	Ca	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Se	Br	Pb
H	1																				
Na	0.01	1																			
Al	**-0.27**	-0.15	1																		
Si	-0.23	-0.14	**0.84**	1																	
P	**-0.42**	-0.14	0.15	0.07	1																
S	0.06	**-0.30**	**0.58**	**0.56**	0.08	1															
Cl	**0.28**	0.12	**-0.32**	**-0.25**	**-0.49**	**-0.45**	1														
K	**0.87**	0.13	**-0.27**	-0.24	**-0.52**	-0.07	**0.51**	1													
Ca	**-0.55**	-0.02	**0.73**	**0.69**	0.11	**0.32**	-0.18	**-0.54**	1												
Ti	0.24	0.17	0.14	0.16	**-0.33**	-0.07	0.07	**0.35**	-0.09	1											
V	-0.18	-0.20	0.01	0.19	0.18	0.14	-0.16	-0.22	0.18	-0.17	1										
Cr	**0.26**	-0.24	-0.12	-0.12	-0.02	0.04	0.16	0.22	-0.08	-0.07	-0.15	1									
Mn	**-0.33**	**-0.32**	**0.37**	**0.51**	0.10	**0.36**	**-0.38**	**-0.37**	**0.48**	0.19	**0.42**	-0.11	1								
Fe	**0.49**	0.13	0.21	**0.35**	**-0.43**	0.06	**0.29**	**0.65**	-0.02	**0.52**	0.10	0.14	0.01	1							
Co	0.04	0.07	**0.31**	**0.31**	-0.13	0.06	0.22	0.18	0.16	**0.32**	**-0.27**	0.09	0.10	**0.32**	1						
Ni	0.04	0.00	-0.08	0.18	-0.09	-0.16	**0.26**	0.10	-0.02	0.18	-0.15	0.03	-0.14	0.12	-0.04	1					
Cu	**0.47**	-0.08	-0.11	0.03	-0.24	0.08	0.21	**0.53**	**-0.31**	0.07	**0.27**	-0.03	-0.09	**0.52**	-0.14	-0.14	1				
Zn	**0.79**	0.00	**-0.31**	-0.20	**-0.42**	-0.15	**0.42**	**0.82**	**-0.46**	**0.28**	-0.11	**0.34**	-0.22	**0.59**	0.08	0.18	**0.59**	1			
Se	-0.11	-0.10	0.13	0.20	0.15	-0.01	-0.18	-0.11	0.18	0.22	0.19	-0.07	0.23	0.18	-0.02	0.24	-0.07	0.05	1		
Br	0.18	0.07	0.13	0.02	0.04	0.09	0.04	0.13	-0.15	-0.06	**-0.34**	**0.26**	-0.17	0.04	0.01	-0.06	0.08	0.19	**-0.31**	1	
Pb	0.05	**-0.29**	0.06	0.01	0.04	**0.25**	0.17	0.11	-0.06	**-0.29**	-0.02	**0.28**	-0.21	-0.03	0.05	-0.08	0.16	0.10	-0.10	**0.25**	1

Table 3 Spearman correlation coefficients between all analysed $PM_{2.5}$ element concentrations. Values denoting significant correlations $|\rho| > 0.5$ are highlighted in bold blue, while values for medium correlations $0.25 \leq |\rho| \leq 0.5$ are marked bold.

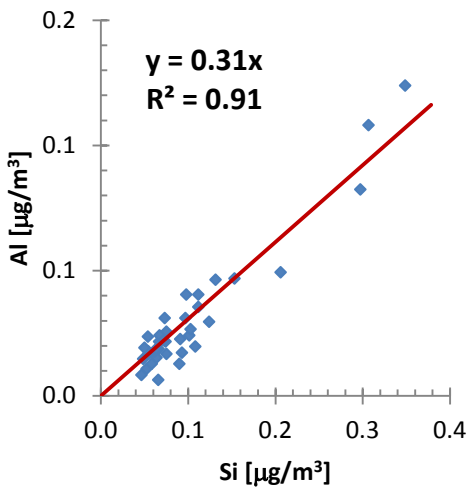
It should be noted that it was originally intended to submit each one PM_{10} filter and one $PM_{2.5}$ filter for 1 in 6 days over a two month period across each season. However there were a series of technical problems that interfered with the collection of PM_{10} samples. This lead to the actual number of PM_{10} filters submitted for element analysis being rather limited. To keep the project on schedule, and because fine particulate matter is of much more interest in terms of human health effects,[1] missed PM_{10} samples were replaced with additional $PM_{2.5}$ samples – any further discussions and results presented in this report are almost exclusively referring to $PM_{2.5}$ data.

For a more in-depth analysis, once the element concentrations for each individual filter were known, the first step was to look for species-specific correlations over the whole sample population. For this purpose **Spearman's** correlation coefficients ρ

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

between each species were calculated, where $d_i = x_i - y_i$ is the difference between the ranked variables. In other words, ρ is essentially the Pearson product-moment correlation of the data **ranks**, thus not specifically measuring *linearity*, but more generally quantifying *monotonic* relations. The resulting correlation coefficients (for $PM_{2.5}$) can be found in Table 3. Strong correlations ($\rho > 0.5$) are highlighted in bold blue, while medium correlations ($0.25 \leq \rho \leq 0.5$) are only marked bold.

Strong positive correlations, indicating the same source, can be seen for the sets {Al, Si, S, Ca} and {K, Fe, Cu, Zn, H}, while strong negative correlations are found for {K, Ca} and {Ca, Zn, H}. These results make sense when looking at the chemical composition of the major soil parent materials, which are Quartz SiO_2, Calcite $CaCO_3$, Feldspar $KAlSi_3O_8$, and Mica $K(Mg,Fe)_3AlSi_3O_{10}(OH)_2$.[9] Therefore, finding strong correlation between Al, Si, Ca, and to a certain extent Fe, K, and Ti, points to windblown **soil/dust** as one source for $PM_{2.5}$. On the other hand, a fair amount of K is also known to stem from **biomass burning**;[10] *i.e.*, it can be used as a trace element for that very source, which also explains the negative K–Ca correlation found in this study (more soil/dust and less wood burning during summer and vice versa during winter months).



In addition to the correlation analysis it is also useful to look at species-specific concentration scatter plots in order to check in which form the respective elements are found. For that purpose Figure 1 (right) and Figure 2 (below) show both the [Al]–[Si] and the [H]–[S] scatter plots.

Figure 1 [Al]–[Si] scatter plot including a linear regression (red line). Regression equation and correlation coefficient are shown top left.

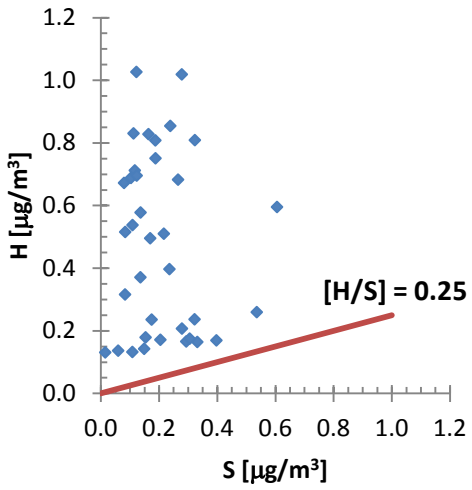


Figure 2 [H]–[S] scatter plot, the red line indicates the [H/S] = 0.25 ratio.

The data in Figure 1 confirms that there is a strong [Al]–[Si] correlation, which is not only monotonic (Spearman's $\rho = 0.84$, see Table 3) but also linear (Pearson's $r = 0.91$). The observed [Al/Si] = 0.31 ratio is very close to Earth's crustal value (≈ 0.28) and thus a good indicator for clay alumina-silicates as a source. This ratio has also been reported for dust aerosols from Australia's Lake Eyre Basin previously.[11] The plot in Figure 2 is used to check if all S from **secondary sulfate** aerosols is found in form of *fully neutralised* ammonium sulfate $(NH_4)_2SO_4$ – which is the case as all [H/S] ratios are well above 0.25.

To further verify the source appointments established and discussed so far, one can also look at the time-series plots of element concentrations, especially for any known source-characteristic tracer elements. Hence, both Figure 3 and Figure 4 (right) depict the (monthly based) time-series for elements characteristically found in crustal soil/dust and elements more likely to be found in wood smoke.

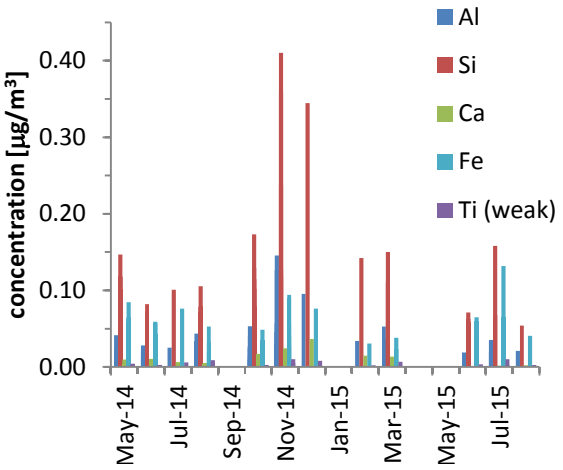


Figure 3 Time series of elements characteristic of soil and mineral dust.

It is evident that the concentrations of Al, Si, Ca, and Fe are significantly higher on $PM_{2.5}$ filters collected during the summer months (November to January). The main source of these elements (windblown soil and dust) has already been established and agrees with the fact that dust aerosols are more likely found during the hot season. On the contrary, the amount of K, which is a well known tracer element for residential wood combustion and biomass burning,[12] is found to be elevated during the cold season, *i.e.*, during May to August.

The green bars labelled "OMH" in Figure 4 indicate the concentration of organic matter estimated from H, however, for convenient scaling the concentration was divided by a factor of 100. As expected the carbon levels correlate with K, even though this is more of a rough estimation because

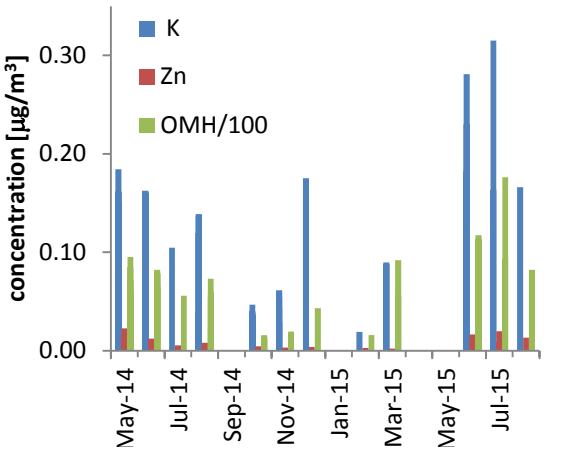


Figure 4 Time series of elements characteristic of wood smoke / biomass burning.

individual elemental and soot-carbon concentrations were not directly measured. Zn concentrations (red bars) are shown in Figure 4 because of the strong [K] – [Zn] correlation observed in Table 3. This correlation has been reported in connection with biomass burning and wood smoke in previous papers.[13]

The next step, after directly identifying some of the pollutant sources, will be to use the chemical composition data in combination with some mathematical models to further refine source appointments and/or identify additional source contributions.

SOURCE CONTRIBUTION ANALYSIS – IMPROVE ALGORITHM

IMPROVE (interagency monitoring of protected visual environments)[14] is a visibility and aerosol monitoring program in the United States which was implemented in 1985 in order to detect and predict haze formation in national parks and protected wilderness areas. Besides using photometric data as main input, the program also utilises data from elemental analyses for calculations of aerosol concentrations contributing to haze and air pollution – such as ammonium sulfate, ammonium nitrate, (total) organic carbon, and fine soil dust.

The calculation of these source contributions is based on simple stoichiometric relationships and some additional assumptions – details can be found in the scientific literature[15] and only a short summary will be given here:

- The soil mass concentration is estimated by summing all elements predominantly associated with crust minerals; all elements are assumed to be present in form of their oxides; FeO and Fe_2O_3 are assumed equally abundant. An overall factor of 1.16 is used on all stoichiometric ratios to account for minor soil constituents like Na_2O, MgO, H_2O, and CO_2. The final equation reads:
  [SOIL] = 2.20 [Al] + 2.49 [Si] + 1.63 [Ca] + 2.42 [Fe] + 1.94 [Ti].

- K has a soil (K_2O) and non-soil component (K_{non} from smoke), thus [Fe] is used as a surrogate for soil potassium, *i.e.*, a factor of 0.6 [Fe] is taken into account for [SOIL]. For tracing smoke **$[K_{non}]$ = [K] – 0.6 [Fe]** is used.

- A general problem with Teflon filter substrates is that Chlorine can be volatilised during filter collection, thus salt is calculated by **[NaCl] = 2.5 [Na]**, rather than simply [Na] + [Cl].

- All elemental S is assumed to be from sulfate; if it is fully neutralised as ammonium sulfate $(NH_4)_2SO_4$ it can be calculated by **[AS] = 4.125 [S]**.

- Organic matter from hydrogen **[OMH] = 11 × ([H] – 0.25 [S])**, assuming that the average organic composition was 71% C, 20% O, and 9% H.

- Reconstructed mass (RCM) is then the sum over all the individual source factors, and is ideally very close to the gravimetrically measured filter mass.

It should be noted that the latter statement about RCM being close to the measured filter mass will inevitably not be met, as the original IMPROVE algorithm also accounts for

additional contributions from ammonium nitrate, total organic carbon, elemental carbon, and light absorbing carbon – all data which were unfortunately not available in the present study. Thus, a certain deviation from the total chemical mass balance (CMB) must be expected and the following results of the composition analyses shown in Figure 5 and Figure 6, as well as those in the bottom sections of Table 1 and Table 2, should be interpreted as approximations rather than exact contributions.

Figure 5 now shows the IMPROVE source contribution breakdown for all individual $PM_{2.5}$ filters averaged over respective sampling months, with sources being soil/dust, salt (NaCl), ammonium sulphate (AS), organic matter from hydrogen (OMH), and potassium (K) from smoke as contributing factors. The "unspecified" fraction represents the remainder from filter mass minus reconstructed mass (RCM). The average mass closure is slightly above 50% which is, considering the limited size of the data set and available chemical analyses, not too bad. The monthly breakdown for the PM_{10} filters is depicted in Figure 6 and Table 2.

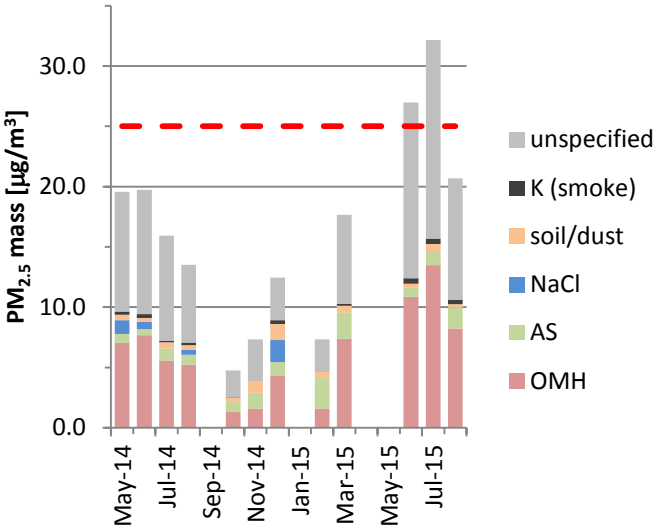


Figure 5 Monthly averaged total $PM_{2.5}$ levels and their source appointments based on the IMPROVE algorithm. The dashed red line indicates the 24h

In general, monthly $PM_{2.5}$ levels observed at the Monash air quality monitoring station are well below the advisory standard outlined by the NEPM for Ambient Air Quality[16] – with two exceedances observable in Figure 5 for June and July 2015. It should, however, be noted that these two averages are based on a very small number of analysed filters and thus do not necessarily reflect the overall $PM_{2.5}$ pollution level for these months.

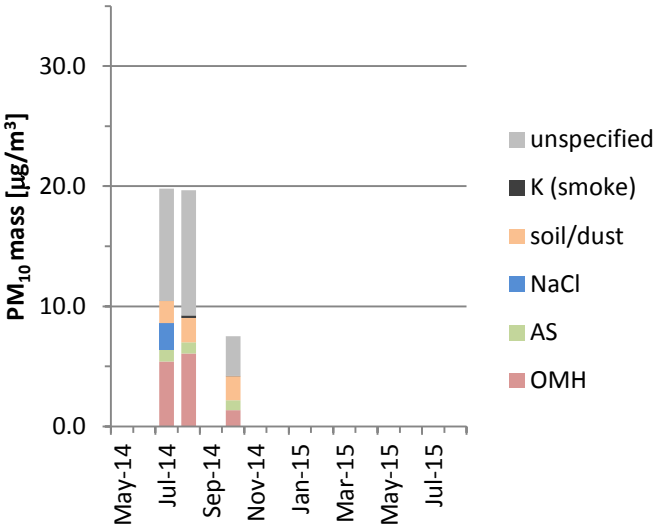


Figure 6 Monthly averaged PM_{10} levels and source factors based on the IMPROVE method.

It can clearly be seen that the total $PM_{2.5}$ levels are significantly higher during the winter months. This can be attributed to the increased residential wood burning activity, evident from the increased amounts of potassium from smoke and organic matter detected during these months. No elemental carbon data set is available however, so this finding is not fully represented in the above introduced source factors.

During the warmer months an increased contribution of windblown soil/dust to the $PM_{2.5}$ levels is visible. This also holds true for the total PM_{10} levels, for which the measured amount of soil and dust is generally found to be significantly higher than for $PM_{2.5}$ (which is expected as most dust particles exhibit aero-dynamic diameter well above 2.5 μm).

Even though the potential of the IMPROVE calculations is limited in the framework of this report (mainly due to small data sets), they do provide rough indications and thus viable information on seasonal fluctuations in the composition of airborne particulate matter pollution. In order to further refine the above discussed source contribution analysis as well as to provide well-established error estimations, a more sophisticated statistical model needs to be utilised.

SOURCE CONTRIBUTION ANALYSIS – POSITIVE MATRIX FACTORISATION

Positive matrix factorisation (PMF) is an advanced one-step receptor model based on least-squares techniques that uses error estimates of the measured data in order to provide meaningful weights during the fitting process.[17] It is a mathematical approach for quantifying the contribution of sources to samples based on the composition or fingerprints of the sources.[15,17,18] A speciated data set can be viewed as a data matrix X of $i \times j$ dimensions, in which i number of samples and j chemical species were measured, with uncertainties u. The overall aim of the receptor model is solving the chemical mass balance between measured species concentrations and a number of proposed source factors:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij}$$

with number of factors p, species profiles f of each source, the particulate mass contributions g by each factor to each individual sample, and the residuals e_{ij} for each sample/species. PMF will now decompose the sample data into two matrices, namely factor contributions (G) and factor profiles (F). In the end the factor profiles have to be interpreted by the user in order to identify the source types – which will require sampling site knowledge and experience in identifying the source associated with elemental fingerprints of each factor.

Factor contributions and profiles are then derived by minimising the objective function Q

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{u_{ij}} \right]^2$$

which is a critical parameter for PMF. Two versions will be obtained for each solution – namely Q(true), a goodness-of-fit parameter calculated including all points, and Q(robust), a goodness-of-fit parameter calculated excluding any points that cannot be fit by the model (defined as samples for which the uncertainty-scaled residual is greater than 4).

The difference between Q(true) and Q(robust) is a measure of the impact of data points with high scaled residuals, *i.e.*, data points that are not consistently present during the

sampling period. On the other hand, if uncertainties are too high very similar results in both Q values will be obtained as the residuals are scaled by the uncertainty.

Species	S/N	Category
H	10.0	Strong
Si	9.9	Strong
S	9.8	Strong
Fe	9.4	Strong
K	8.8	Strong
Cl	6.8	Strong
Al	2.5	Strong
Zn	1.8	Strong
Ca	1.4	Strong
Cu	0.7	Strong
Mn	0.4	Strong
Ti	0.3	Weak
P	0.3	Weak
Cr	0.2	Weak
Na	0.2	Weak
Pb	0.1	Weak
Ni	0.1	Weak
Br	0.1	Weak
Co	0.0	Weak
Se	0.0	Weak
V	0.0	Weak
Mass	10.0	Weak

Table 4 Element, signal to noise ratio, assigned category.

Parameter	
Data type; averaging timeframe	$PM_{2.5}$ [µg/m^3]; 24-h
N non-weak species (m)	11 (22 total)
N samples (n)	37
N factors (p)	5
Treatment of missing data	No missing data included
Treatment of concentrations ≤ MDL	Data used as reported, no modification or censoring
Treatment of concentrations ≤ 0.0	Data used as reported, no modification or censoring
Lower limit for normalised factor contributions g_{ik}	–0.2
Robust mode	Yes
Constraints	None
Extra modelling uncertainty (%)	10
Seed value	55
N bootstraps in BS; r^2 for BS	400; 0.8
BS block size	1
DISP dQ^{max}	4, 8, 15, 25
DISP active species	all non-weak
N bootstraps; r^2 for BS in BS-DISP	100; 0.8
BS-DISP active species	H, Al, S, Cl, K
BS-DISP dQ^{max}	0.5, 1, 2, 4
CPU runtimes* for DISP, BS-DISP	< 1 h, < 2 h

* Windows 7, 64-bit,running in Oracle VirtualBox VM on (early 2013) MacBook Pro, OS X 10.10.5 with Intel Core i7-3840QM CPU @ 2.8 GHz, 16 GB RAM

Table 5 Summary of EPA PMF input/settings

The freely available software package **EPA PMF 5.0** (maintained and distributed by the United States Environmental Protection Agency) was used for the analysis presented in this report. Further technical details as well as a user guide can be found online.[19] In the following, the most important modelling parameters will be provided.

As a first step each element species had to be categorised regarding its signal to noise ratio (S/N) within the data matrix. Because of either large measurement uncertainties or concentrations significantly below the method detection limits (MDL) not all element concentrations are sufficient to be used as model input. Table 4 (above left) provides a short overview of all elements present in the PMF data set, their signal to noise ratios, and the user-defined category of each species. The S/N ratio is automatically calculated by the PMF software taking concentrations as well as measurement uncertainties (or optional MDL) into account. The assigned category will determine how strong a species will contribute to the best-fit solution – "strong" species are fully taken into account, "weak" will triple the species' uncertainty, and "bad" will exclude the species from the model. Based on the overall small number of available filters in the present data set no elements were excluded from the PMF model (except for F, originating from the Teflon® filters rather than any natural source), but most of the species with very low S/N ratios (< 0.5) have been

categorised as weak. The gravimetric $PM_{2.5}$ filter mass (bottom row in Table 4) is specified as a "total variable" which is used by the program only for post-processing of results and thus automatically set to *weak* to minimise its influence on model solutions. In addition, an overall +10% model uncertainty was added to the species-specific measurement uncertainties in order to allow large enough errors to ensure reasonable and stable PMF solutions.

Using this categorisation, after several short PMF test runs, a stable final solution for five factors has been found starting from 250 base runs and using the input parameters summarised in Table 5. Base run #229 resulted in the lowest $Q(\text{robust}) = 139.2$ value, which is satisfactorily lower than the theoretical $Q_{exp} = m \times n - p(m+n) = 167$ for the present data set (EPA PMF only uses strong species to calculate Q_{exp}). The residual analysis for this particular run revealed that the scaled residuals for all strong species are well-modelled, meaning they are within the required ± 3 standard deviations and all normally distributed, as shown in Figure 7 on the right. In fact, the scaled residuals of *all* species in the data set are within these limits, but for clarity, only strong species are shown in the figure. This finding is also confirmed by looking at the $Q/Q_{expected}$ ratios (Figure 8 to the right) for all elements, showing that only Zn exhibits a ratio ≥ 2, and Ca and K slightly below that value. These three elements also have the broadest scaled-residuals distributions and together with their Q-ratios at around 2 this indicates that they might be present in infrequent sources – a fact which most likely originates from the low number of filters. More information on this ratio (and how it is calculated) can be found in the EPA PMF user's guide.[20]

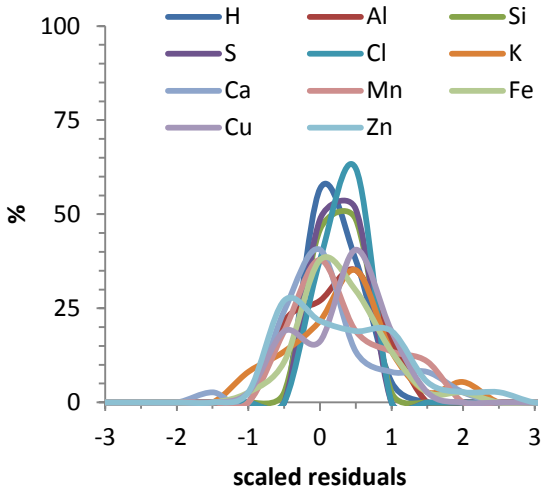


Figure 7 Scaled residual analysis for the stable PMF solution showing only strong species included in the model.

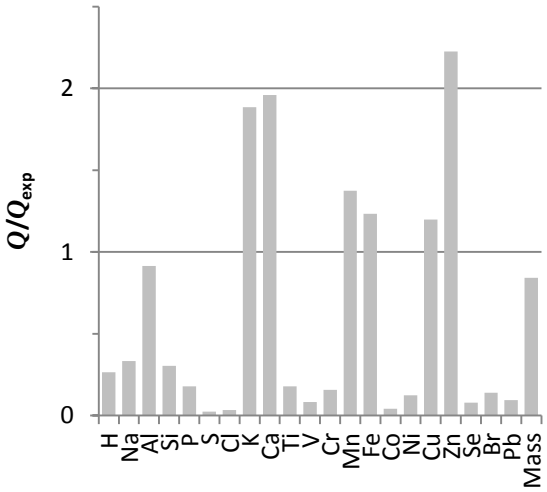


Figure 8 Q/Q_{exp} values for each species included in the base model.

As already mentioned, the rather broad residuals distribution and the $Q/Q_{exp} > 2$ value for Zn indicate that this species is not well modelled; however, test runs with both smaller and larger uncertainties and runs with more than 5 source factors did not result in any stable solutions – an in-depth error analysis (discussed in detail below) revealed that this is in fact the most stable solution, and any changes to the input parameters would lead to rotational ambiguity. It was expected though, that the PMF approach would not result in a perfect solution, as this method is foremost a statistical tool and will perform much better for data sets including at least hundred(s) of individual samples.

For the error estimation of the stable solution all three possible methods available within the EPA PMF 5.0 software package have been used, which are the displacement (DISP), the bootstrap (BS), and a combination of both methods (BS-DISP). DISP intervals include rotational ambiguity effects, BS intervals include effects from random errors and partially include effects of rotational ambiguity (while being generally robust if data uncertainties are misspecified), and BS-DISP intervals fully include both random errors and rotational ambiguity (but less robust in case of misspecified data uncertainty). Detailed descriptions of these methods can be found in the user manual as well as in scientific literature.[21]

The first step in the base error estimation procedure was to perform the DISP calculation, followed by the BS calculation (with input parameters for both methods given in Table 5) allowing all 11 strong species to be displaced. There were no decreases in Q for the DIPS runs, nor any swap counts in the lower dQ^{max} ranges. For the final BS-DISP run (which needs both DISP and BS error bars as input) only key elements of source factors were allowed to be displaced (see Table 5). No cases with a *significant* decrease in Q could be identified, as the largest decrease was only dQ = –0.59 (–0.42%), even though 114 possible swaps in the best fit and 31 swaps in the DISP result were found. In general this means that all source profiles and source factors in the initial base solution are reasonably well modelled and no additional rotations or species displacements would significantly improve the solution based on the current data set and data uncertainties.

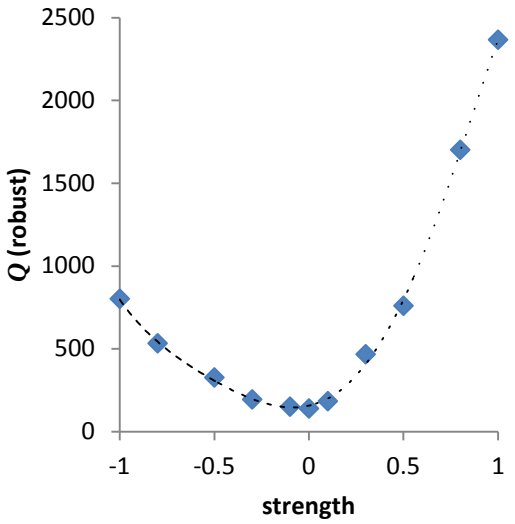


Figure 9 Q vs. Fpeak strength for rotational analysis. Base run at strength = 0.00

In order to further confirm that the found solution is mathematically unique, *i.e.*, no rotations and/or transformations of the factor matrices G and F exist that would lead to the same Q-values, an Fpeak was performed using the implemented rotational tools. A positive Fpeak strength is used to sharpen the F-matrix and smear the G-matrix, and vice-versa for a negative strength. The resulting Q-values for the explored range of strength-parameters are plotted in Figure 9 on the left, and it can easily be seen that the stable base solution (at zero strength) already represents the minimum, thus appearing to be unique.

The most important results of the PMF solution are the source factor *profile* matrix F and the source factor *contribution* matrix G, which are shown in Figure 10 and Figure 11 on the following pages. The first figure shows the element concentration profiles within each individual source factor in μg/m^3 (often called source factor composition or factor fingerprints), while the latter one depicts the percentage contribution of each element to a given source factor.

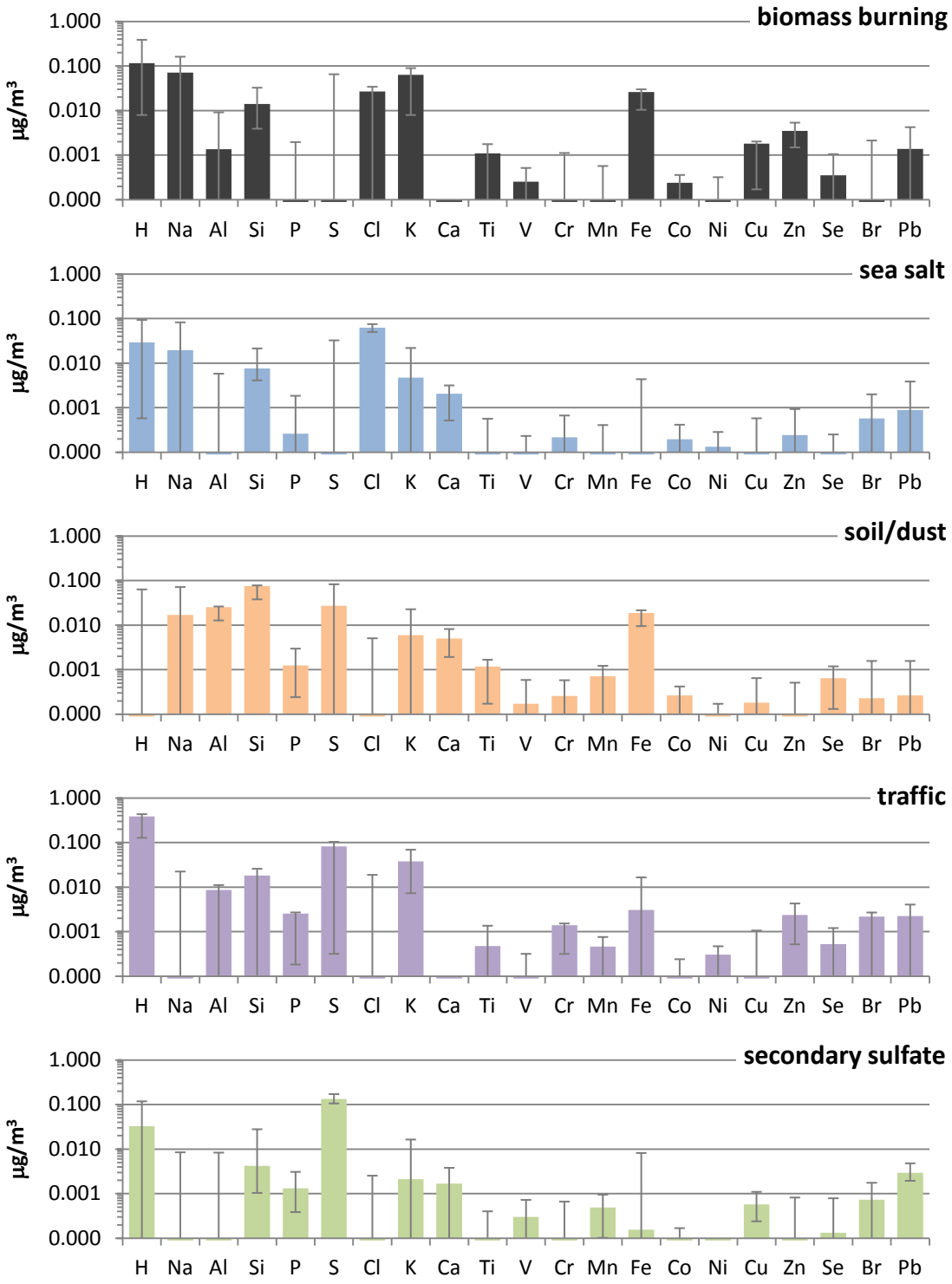


Figure 10 Concentration profiles (fingerprints) of the five identified PMF source factors including error bars (5th and 95th percentiles from bootstrap error estimations). Due to the concentration ranges the abscissa are shown on logarithmic scales.

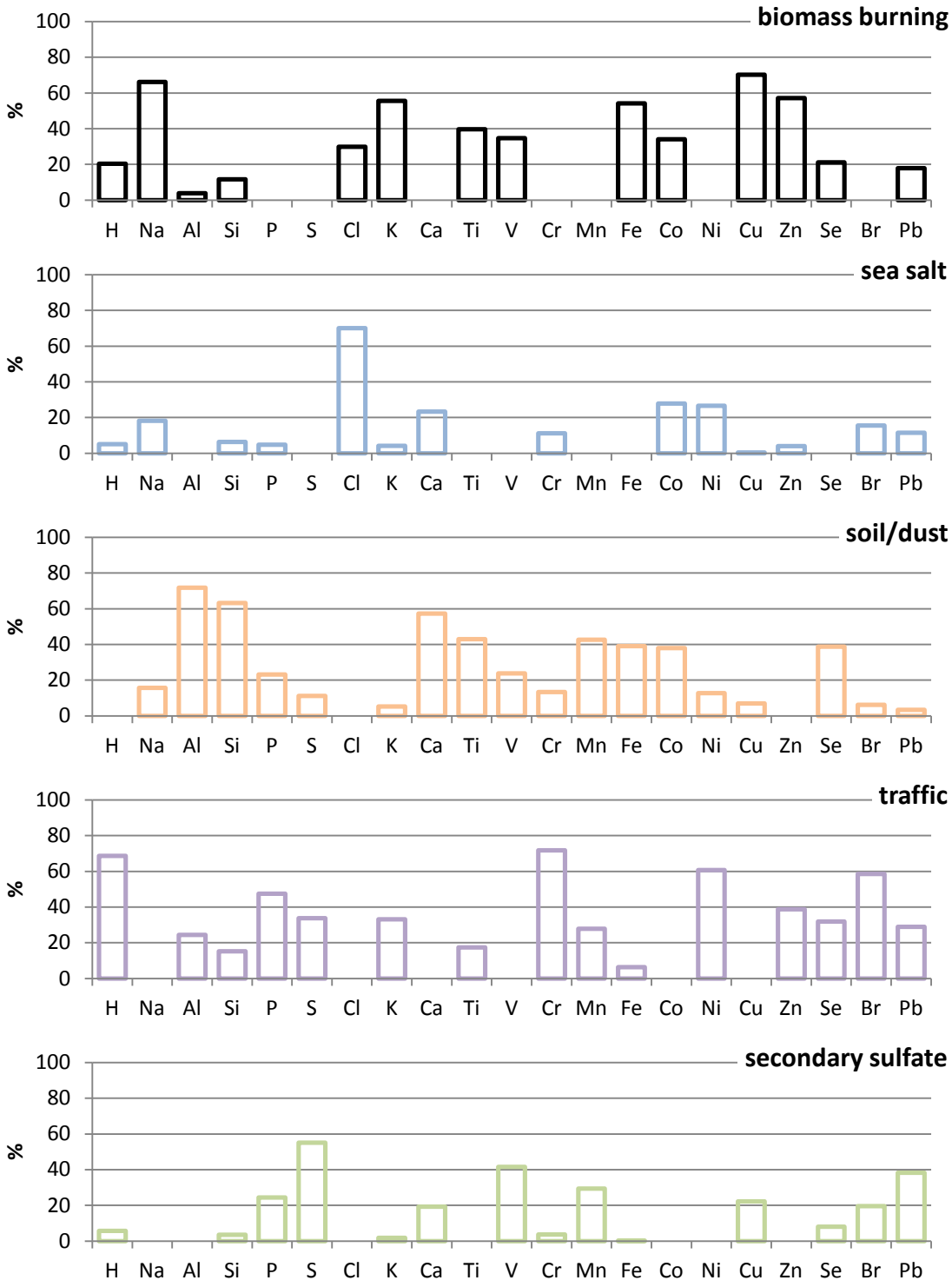


Figure 11 Percentage contribution of each element to the five identified source factors.

Assigning meaningful sources to each of the five factors in the current PMF solution is now (as already mentioned) based on site knowledge and mostly on tracer elements for known pollution sources. For example, the first factor in Figure 11 (top, black) contains almost 60% of all K and about as much of Zn – both elements known to be key indicators for **biomass burning** (*i.e.*, wood smoke either from residential heating and/or bush fires). The source with the highest concentrations of Al, Si, Fe, and Ti is most likely to be windblown **soil/dust**, which is found for the third factor (orange), while **sea salt** (second factor, blue) can easily be identified by its high concentration of Cl. (Note: one would of course also expect to find a high Na contribution to this factor, however, either the Na data set was incomplete or Na is predominantly present as $Na^{\oplus}$, but no water-soluble ions have been measured.) The remaining two factors can be assigned to **secondary sulfate** aerosols – due to the [H]/[S] concentration ratio of ≈ 0.25 found in the bottom factor in Figure 10 (green), and to (mainly) road **traffic** related pollution (fourth factor in the figures, purple) – due to its high content of organic matter in combination with Br, Pb, P and Zn (from engine emissions), S (from tyre abrasion), or Cr and Ni (from oil combustion).

All the error bars shown on the source profiles (Figure 10) are those which resulted from the BS calculations – and even though the small size of the available data set imposed limitations on the overall PMF model quality, the resulting profiles could be appointed straightforwardly and seem to be very reasonable regarding possible air pollution sources in the ACT.

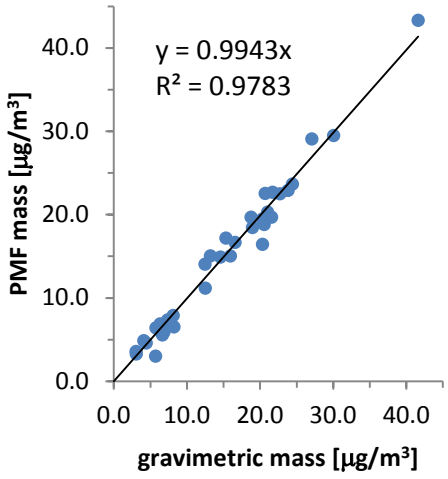


Figure 12 PMF mass *vs.* observed mass.

Additional quality assessment of the current PMF solution is provided by both the good correlation between the (observed) gravimetric filter mass and the (calculated) PMF mass, and the two well coinciding timeseries of these masses – depicted in Figures 12 and 13 on the right. Again, it is evident that despite the rather small numbers of samples and some missing months, the software does a fairly decent job reproducing the observed total filter mass.

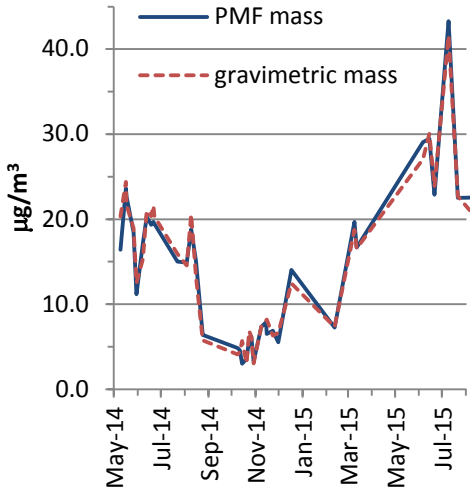


Figure 13 Timeseries of both PMF and observed $PM_{2.5}$ mass.

Finally, with the species contributions known and the source appointments done, the last step is to break down the overall sample compositions into individual (percent) factor contributions to unveil seasonal fluctuations of pollution sources. This is done in Figure 14 on the next page, which shows the monthly-averaged contribution of each source to the $PM_{2.5}$ levels. As already mentioned, not all months are covered in the data set and thus caution should be taken when interpreting this figure – some basic facts can however be easily deduced. Apparently, the two main

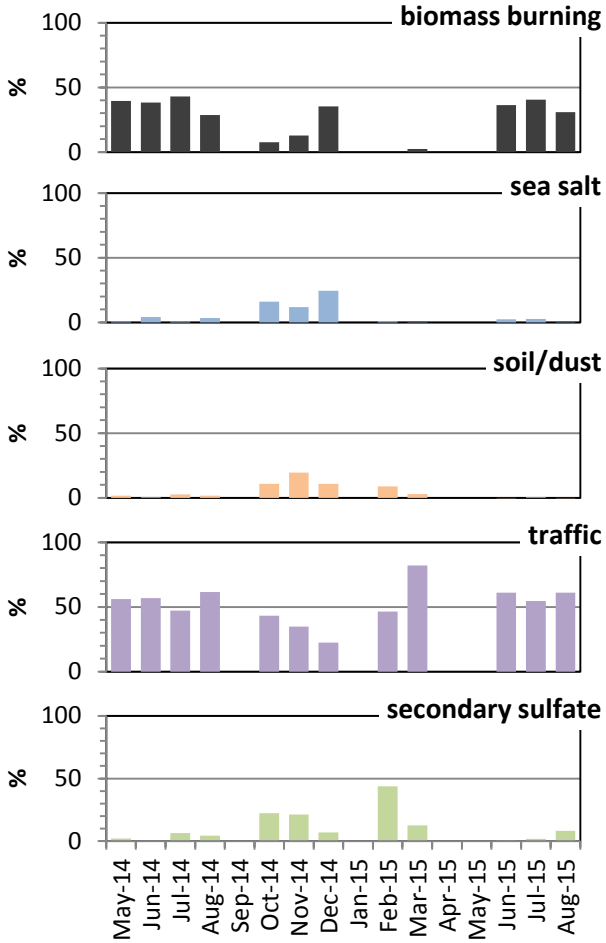


Figure 14 Percentage of source factor contributions to the PMF fine particulate matter level.

sources of the airborne fine particulate matter levels monitored in the Monash air shed are road traffic and biomass combustion. Traffic constituted a fairly constant factor over the whole sampling period, reading a bit above 50% (with small fluctuations and one exceptional peak for March 2015). In general, wood combustion made up about 30-40% of the levels during the cold season and dropped to around 10% during warmer months. The visible peak during December 2014 is likely to come from bush fires and/or hazard reduction burns conducted both in and around the ACT. Secondary sulfate aerosols can be found all year long, but naturally exhibit higher concentrations during the longer summer days, due to the underlying atmospheric photochemistry. The remaining two factors, sea salt and windblown soil dust, are mostly inferior contributors to *fine* particulate matter, and only show significant levels between Oct-2014 and Dec-2014 – which, according to available climate data,[22] was a rather windy season. (Especially for sea salt as a contributing factor in the ACT wind is a prerequisite – but it can travel inland up to several hundred kilometres.[23]) Similar to the IMPROVE mass-composition analysis (see Figure 5), Figure 15 below shows the composition of the monthly averaged PMF mass concentrations in µg/m^3. Any recently discussed findings can also be seen in this figure. As with Figure 5, the June and July 2015 peaks are based on a small number of filters and do not necessarily represent the total levels for both months.

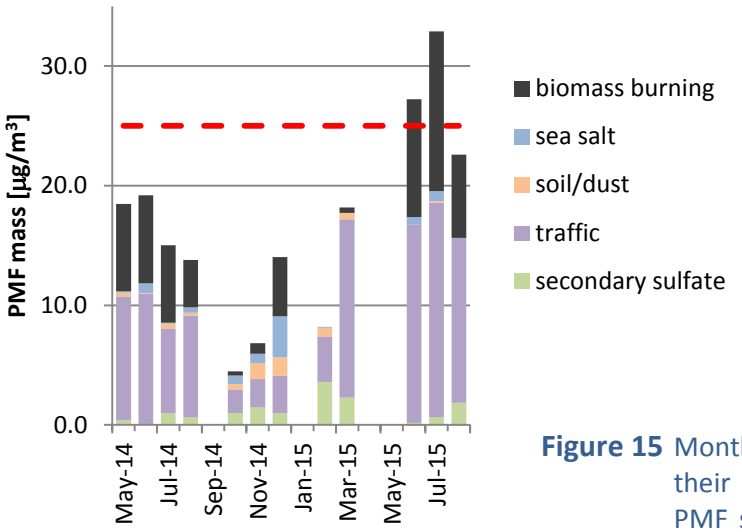


Figure 15 Monthly averaged total $PM_{2.5}$ levels and their source appointments based on the PMF solution; red line: Air Quality NEPM 24-h standard.

Figure 16 (below) illustrates the overall fine particulate matter composition in the Monash air shed, showing the total source contributions averaged over all individual filters. A very distinct ranking of the sources can be observed, with (as already mentioned) road traffic with 56% and biomass burning with 31% clearly being the top factors, followed by secondary sulfate aerosols with about 6%. Sea salt and windblown soil are the two factors with the lowest contributions, together adding up to around 7%.

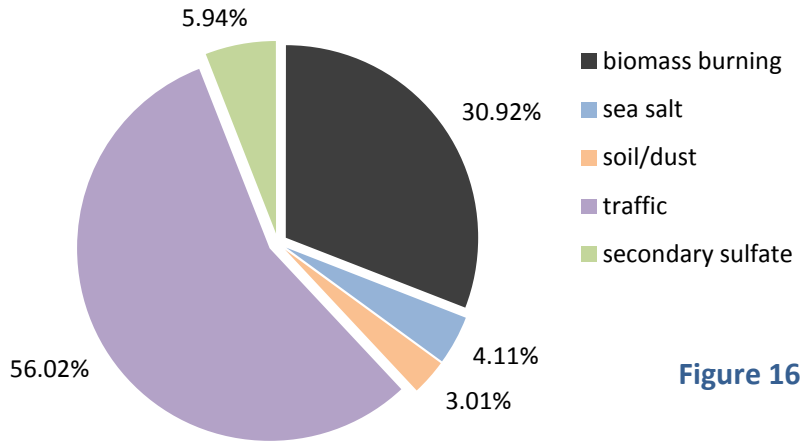


Figure 16 PMF total average source factor contributions over all analysed samples.

Considering the imposed limitations on the PMF model these numbers are of course not totally accurate. Additionally, the PA PMF software package only shows (the 5th and 95th) percentiles and due to the small (and thus not normally distributed) data set, a statement about error bars on this figure is not constructive. As already discussed however, the final solution appeared to be reasonably stable and self-contained and thus we expect these numbers to be a reasonable approximation of source factors in the ACT.

SUMMARY AND CONCLUSION

The elemental composition of both fine and coarse airborne particulate matter ($PM_{2.5}$ and PM_{10}) – collected over a period of 16 months at the Monash air quality monitoring station in Canberra, Australian Capital Territory – was analysed using the two non-invasive ion beam techniques particle induced X-ray emission (PIXE) and particle induced gamma emission (PIGE) spectrometry. Based on element concentrations and statistical modelling, five major pollutant sources (biomass burning, sea salt, soil/dust, traffic, and secondary sulphates) for $PM_{2.5}$ have been successfully appointed and the seasonal fluctuations in their contributions to the overall fine aerosol levels in the Tuggeranong valley have been examined.

The most important findings and results of this study are as follows:

- In general, the daily average particulate matter and aerosol pollution levels in the monitored airshed (Tuggeranong valley) were **satisfactorily below** the respective 24-

hour ambient air quality NEPM standards of 50.0 μg/m^3 for PM_{10} and 25.0 μg/m^3 for $PM_{2.5}$.

- For **PM_{10}** the average concentration of windblown crustal **soil/dust** is about 4–5 times greater than in $PM_{2.5}$.
- For **$PM_{2.5}$** the average concentration of **wood-smoke** related K was found to be about three times as high as in PM_{10}.
- **$PM_{2.5}$** levels, which usually have more impact on human health issues, were significantly **elevated during winter** months (May – August); Exceedances of the 24-hour NEPM standard for $PM_{2.5}$ are more likely to occur (and have occurred) during the cold period.
- For $PM_{2.5}$ distinct **seasonal fluctuations** were observed for both **biomass combustion** and **soil/dust** related contributions. During winter 30–40% of the fine aerosol pollution can be appointed to wood smoke and less than 5–10% to windblown soil, whilst in summer this ratio is inverted (up to 20% from soil and less than 10% from wood smoke – if no exceptional bushfire events occur).
- Both (aged) sea salt and secondary sulfate aerosols contribute to the PM levels in the investigated airshed, and are subjected to seasonal fluctuations. Naturally, due to atmospheric transport/chemistry higher levels for both were found during hot months.

In conclusion it should be noted that both the elemental analysis of airborne particulate matter within a specified airshed, and the accompanying source-contribution analysis, are extremely viable methods to not only monitor the concentrations of potentially harmful species in air pollutants, but also to help measure the impact of health and environmental policies over time. In this regard a continuation and/or implementation of similar projects is recommended, including the following suggestions for improved outcomes:

- Extending the chemical analysis for direct measurement of water soluble ions (Na^+, K^+, NH_4^+, SO_4^{2-}, NO_3^-) and carbon (elemental/soot carbon EC, organic carbon OC).
- If technically feasible, extend filter sampling times from 24 hours to (e.g.) 72 hours, in order to improve quantitative analysis of low-concentration species.
- In order to effectively study potential policy impacts, any future monitoring should be conducted over a minimum period of 2 years to include at least two sets of each season, the longer the better.
- Include the Florey monitoring station as a second sampling location in order to increase data coverage for statistical assessment and/or spatial breakdown of air pollution if required, especially for fine particulate matter.

ACKNOWLEDGEMENTS

The authors would like to thank the Australian Nuclear Science and Technology Organisation (ANSTO) for conducting the sample analyses.

Valuable comments and helpful discussions during the preparation of this report by Prof. D. Cohen and E. Stelcer (both ANSTO), as well as Prof. H. Bridgman (University of Newcastle) are greatly appreciated.

NOTES AND REFERENCES

[1] J. Schwartz et al., *Journal of the Air and Waste Management Association*, 46, 927, **1996**.
B. Brunekreef et al., *Epidemiology*, 8, 298, **1997**.
H. Duhme et al., *Toxicology Letters*, 102-103, 307, **1998**.
V. Ramanathan et al., *Science*, 294, 2119, **2001**.
A. Nel, *Science*, 308, 804, **2005**.
A.G. Russell, B. Brunekreef, *Environmental Science and Technology*, 43, 4620, **2009**.
J.Z. Yu et al., *Aerosol and Air Quality Research*, 14, 237, **2014**.

[2] http://www.health.act.gov.au/public-information/public-health/act-air-quality-monitoring/wintertime-air-quality-tuggeranong

[3] http://www.ansto.gov.au/ResearchHub/IER/Capabilities/IBA/index.htm

[4] *Ion beam techniques for the analysis of light elements in thin films, including depth profiling*, IAEA-TECDOC-1409, International Atomic Energy Agency, Vienna, **2004**.
D.D. Cohen et al., *Nuclear Instruments and Methods in Physics Research*, B109, 218, **1996**.
D.D. Cohen et al., *Nuclear Instruments and Methods in Physics Research*, B161, 775, **2000**.
D.D. Cohen et al., *Nuclear Instruments and Methods in Physics Research*, B219, 145, **2004**.

[5] Y.C. Chan et al., *Proceedings of the 17th International Clean Air and Environment Conference*, Hobart, 3-6 May **2005**.
Y.C. Chan et al., *Atmospheric Environment*, 42, 374, **2008**.
D.D. Cohen et al., *Atmospheric Environment*, 44, 320, **2010**.
D.D. Cohen et al., *Nuclear Instruments and Methods in Physics Research*, B318, 204, **2014**.
E. Stelcer et al., *Environmental Chemistry*, 11, 644, **2014**.

[6] http://www.ansto.gov.au/ResearchHub/IER/Capabilities/IBA/IBACapabilities/PIXE/index.htm

[7] http://www.ansto.gov.au/ResearchHub/IER/Capabilities/IBA/IBACapabilities/PIGE/index.htm

[8] Y. Hazi et al., *Atmospheric Environment*, 37, 5403, **2003**.

[9] F.W. Clarke, H.S. Washington, *The composition of the earth's crust*, 127, US Government Printing Office, Washington D.C., **1924**.

[10] M.O. Andreae, *Science*, 220, 1148, **1983**.

[11] M. Radhi et al., *Atmospheric Environment*, 44, 3519, **2010**.

[12] C.P. Calloway et al., *Atmospheric Environment*, 23, 67, **1967**.

[13] Hedberg et al., *Atmospheric Environment*, 36, 4823, **2002**.
Saarikoski et al., *Atmospheric Environment*, 41, 3577, **2006**.
A.K. Frey et al., *Boreal Environment Research*, 14, 255, **2009**.

[14] http://vista.cira.colostate.edu/improve/Default.htm

[15] R.A. Eldred et al., *Nuclear Instruments and Methods in Physics Research*, B22, 289, **1987**.
W.C. Malm et al., *Journal of Geophysical Research*, 99, 1347, **1994**.
J.G. Watson, *Journal of the Air and Waste Management Association*, 52, 628, **2002**.

[16] As of February 2016 the advisory reporting standard has been implemented as a proper 25 $\mu g/m^3$ - 24h compliance standard.

[17] P. Pateero et al., *Environmetrics*, 5, 111, **1994**.
P. Pateero, *Chemometrics and Intelligent Laboratory Systems*, 37, 23, **1997**.

[18] Y.C. Chan et al., *Atmospheric Environment*, 45, 439, **2011**.
D.D. Cohen et al., *Atmospheric Environment*, 61, 204, **2012**.

[19] http://www.epa.gov/air-research/positive-matrix-factorization-model-environmental-data-analyses

[20] http://www.epa.gov/air-research/epa-positive-matrix-factorization-50-fundamentals-and-user-guide

[21] S.G. Brown et al., *Science of the Total Environment*, 518-519, 626, **2015**.

[22] http://www.bom.gov.au/climate/data/

[23] J.A. Thornton et al., *Nature*, 464, 271, **2010**.