Sydney Particle Characterisation Study

$PM_{2.5}$ Source Apportionment in the Sydney Region between 2000 and 2014

David D Cohen, Armand J Atanacio, Eduard Stelcer, David Garton

Australian Nuclear Science and Technology Organisation

DISCLAIMER

This report was prepared by ANSTO in good faith exercising all due care and attention, but no representation or warranty, express or implied, is made as to the relevance, accuracy, completeness or fitness for purpose of this document in respect of any particular user's circumstances. Users of this document should satisfy themselves concerning its application to, and where necessary seek expert advice in respect of, their situation. The views expressed within are not necessarily the views of the NSW EPA and may not represent NSW EPA policy.

The NSW Environment Protection Authority (NSW EPA) commissioned this report from the Australian Nuclear Science and Technology Organisation (ANSTO).

Contents

1. Key findings at a glance

This Sydney Fine Particle Study applied positive matrix factorisation (PMF) source apportionment techniques on existing daily $PM_{2.5}$ datasets at four sites in the greater Sydney region. The study period was from 1 January 2000 to 31 December 2014. The four sites were at Lucas Heights, Richmond, Mascot and Liverpool. The datasets, collected and analysed for $PM_{2.5}$ mass and over 23 different elemental and chemical species over 15 years were used as inputs to the PMF codes producing seven different source fingerprints and quantifying their contributions to the total $PM_{2.5}$ mass at each of the sites.

The average $PM_{2.5}$ mass across all sites of 6.823µgm^{-3} was split into seven component source fingerprints summarised in **Table 1-1.**

Table 1-1. Average $PM_{2.5}$ source fingerprints across all sites.

Source fingerprint	Description	Av. $PM_{2.5}$ mass (µgm^{-3})	%
Soil	To represent fine windblown dust	0.245	(4±5)%
Sea	To represent sea spray transported from the coast across the sampling sites	0.507	(10±11)%
Mixed-2ndryS	To represent secondary sulfates from coal fired power stations, oil refineries, motor vehicles and industry	1.63	(24±16)%
Mixed-Ind-Saged	To represent industrial sources with components of aged secondary sulfates and sea spray	0.945	(15±13)%
Mixed-Smoke-Auto	To represent smoke from biomass burning, particularly domestic wood heaters in the winter months with components of smoke from light, medium and heavy diesel vehicles, particularly at Mascot	2.08	(24±20)%
Auto1	To represent the primary automobile source	1.22	(20±10)%
Auto2	To represent a second minor automobile source, associated with the use of leaded petrol which ceased early in the study in 2001	0.234	(3±2)%

The Mixed-2ndryS and the Mixed-Ind-Saged fingerprints contributed 50–70% in the summer months while the Mixed-Smoke-Auto from domestic wood combustion contributed 60–80% in the winter months. The table shows that roughly 85% of the fine $PM_{2.5}$ mass across the four sampling sites is anthropogenic.

The two automobile factors together contributed 1.45µg/m^3 or around 23% of the total fine mass across all sites, whereas secondary sulfates contributed around 24% to the total fine mass across all sites.

2. Executive summary

The Australian Nuclear Science and Technology Organisation (ANSTO) has been applying accelerator based nuclear techniques to the characterisation of fine $PM_{2.5}$ ambient air pollution since the early 1990s. Over the decades large long-term databases have been acquired at dozens of sites both in Australia and internationally on the $PM_{2.5}$ mass together with over 23 different elemental and chemical species that make up this fine particle pollution.

In this study we used data previously collected by ANSTO from four of our long-term sampling sites covering the period from 1 January 2000 to 31 December 2014. Positive matrix factorisation (PMF) source apportionment techniques were applied to this data to identify seven different source components or fingerprints that make up the measured total $PM_{2.5}$ mass at each of these four sites.

The primary aim of this study was to:

- convert the existing 15-year $PM_{2.5}$ mass and elemental datasets for four given sites in the Sydney basin into identifiable source fingerprints
- quantify the absolute and the percentage contribution of each of these fingerprints to the total fine $PM_{2.5}$ mass
- provide seasonal and annual variations for each of the source fingerprints
- provide a readily accessible database containing the daily source fingerprints and their contributions covering the 15-year period from 2000–2014 for four given sites in the Sydney basin
- if possible, identify and quantify the major contributors of fine particle pollution to the ambient air quality in Sydney.

Typically fine particles were collected over 24-hour periods twice a week (104 filters per year) at Lucas Heights, Richmond, Mascot and Liverpool sites over a 15-year period from 2000 to 2014. In all, around 6000 sampling days are represented by this study. Each of these filters was analysed for the 23 elemental and chemical species: hydrogen (H), sodium (Na), aluminium (Al), silicon (Si), phosphorous (P), sulfur (S), chlorine (Cl), potassium (K), calcium (Ca), titanium (Ti), vanadium (V), chromium (Cr), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), selenium (Se), bromium (Br), lead (Pb), black carbon (BC) and total nitrogen (TotN) to concentrations down to $1 ngm^{-3}$ of air sampled. TotN is the total nitrogen from ammonium and nitrate ions.

$PM_{2.5}$ mass and composition

Box and whisker plots are a convenient way to look at the distribution of daily $PM_{2.5}$ masses at each of the sites. **Figure 2-1** shows such plots for the four sampling sites for each of the sampling years between 2000 and 2014. The (+) symbols are the yearly averages, the horizontal bars in the boxes are the annual median values, the shaded rectangular boxes represent 25–75% of the measured masses for each year, the vertical whiskers on each box represent the 95% confidence intervals for the annual masses and the circular dots represent the extreme or outlier events beyond the whiskers. The dashed horizontal lines at $8 \mu gm^{-3}$ and $25 \mu gm^{-3}$ represent the standards in the National Environment Protection (Ambient Air Quality) Measure (Air NEPM) for an annual average mass and a maximum 24-hour period. When the annual (+) symbols lie above the $8 \mu gm^{-3}$ line then the Air NEPM annual average mass has been exceeded, and when the circular dots lie above the $25 \mu gm^{-3}$ line the Air NEPM 24-hour maximum mass level has been exceeded. The vertical mass scale on each of the plots

has been set to 50µgm^{-3} so the extent of the box and whiskers for each year can be readily seen even though there are a few points above 50µgm^{-3} at each site during the 15-year study period.

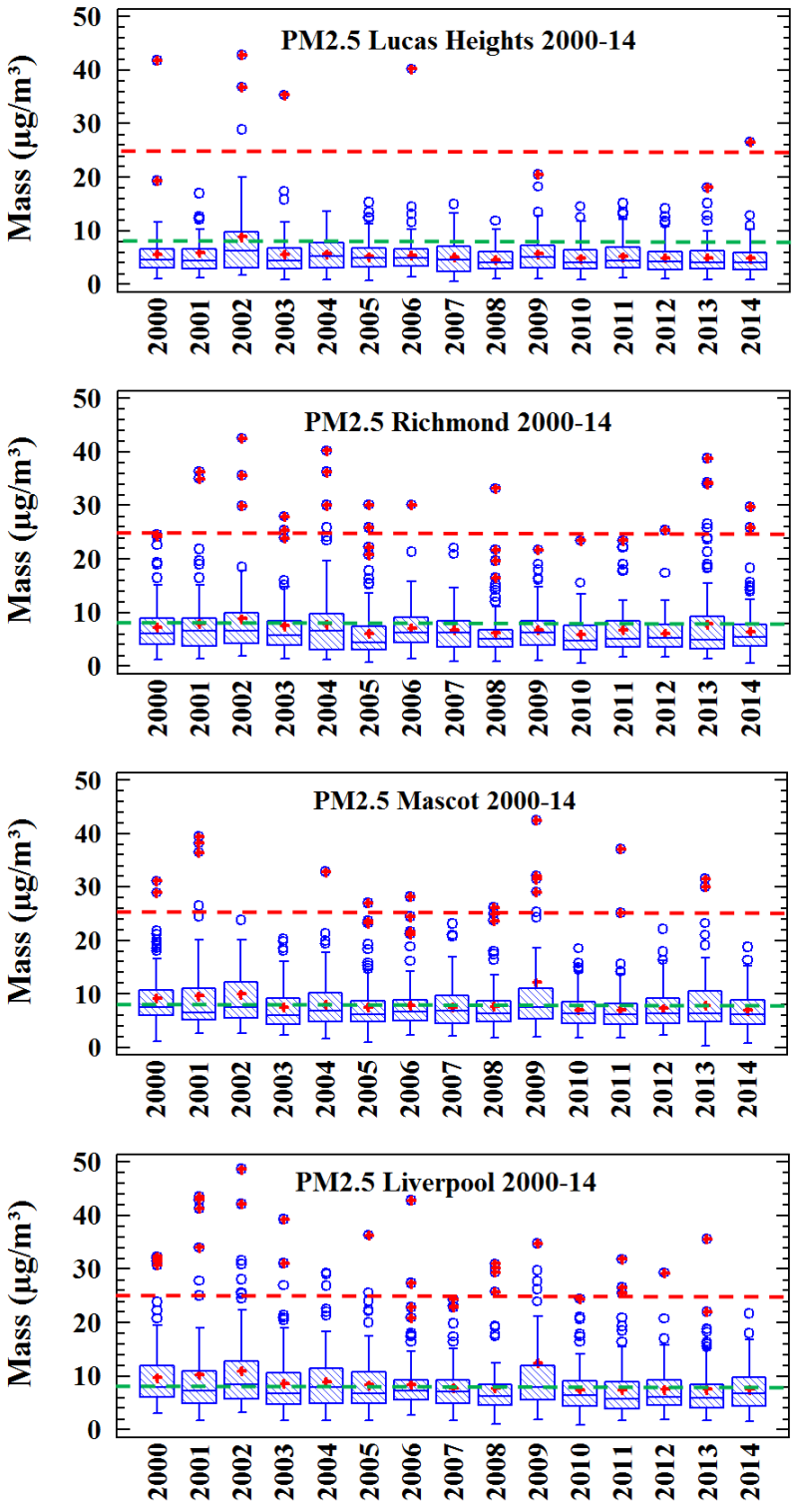

Figure 2-1. Fine $PM_{2.5}$ mass at each of the four sites, 2000–14.

This study was not specifically designed to assess compliance with the Air NEPM standards. Nevertheless we can make minimum estimates of the number of exceedances since we do measure total mass over a 24-hour period during the study period.

Table 2-1 summarises the number of Air NEPM annual and maximum daily mass exceedances for each site over the 15-year study period. As sampling only occurred twice a week on Wednesdays and Sundays, and only 104 filters were collected each year, this is a minimum estimate for both the annual exceedances and the 24-hour maximum exceedances.

Table 2-1. A summary of the number of Air NEPM annual and maximum daily mass exceedances for each site over the 15-year study period.

Site	Annual Air NEPM mass exceedances over 15 years	24-hour max. Air NEPM mass exceedances over 15 years
Lucas Heights	1	11
Richmond	1	28
Mascot	2	23
Liverpool	3	43

Increased annual averages were measured at each site in 2009, due to the September dust storms, where the daily $PM_{2.5}$ levels rose to over 300µgm^{-3} at some sites.

The $PM_{2.5}$ mass was analysed for 23 different elements and chemical species using the accelerator based ion beam analysis (IBA) techniques developed over many decades at ANSTO (Cohen et al. 1996, 2004a, 2004b). These IBA techniques are ideally suited to analysing small microgram samples obtained from $PM_{2.5}$ ambient air pollution.

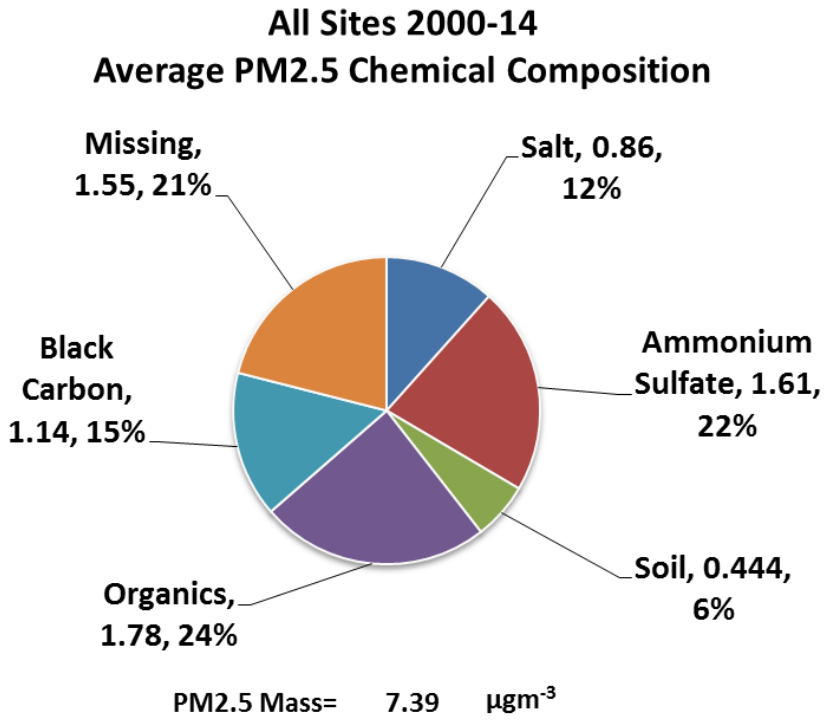

Figure 2-2. Average chemical composition for all four sites, 2000–14.

The average chemical composition is shown in the pie chart in **Figure 2-2**. The average $PM_{2.5}$ mass was 7.39µgm^{-3} and average chemical composition across all four sampling sites was 12% sea salt, 22% ammonium sulfate, 6% windblown soil, 24% organics and 15% BC. The 21% missing mass was composed of water vapour and nitrates which were not measured in this study.

Figures 2-1 shows that the average $PM_{2.5}$ mass across all four sites does not vary a lot with time. The sampling sites are different, some being rural and others heavily urbanised, so the total fine mass varies but the chemical composition turns out to be surprisingly similar. Yet this is not too unexpected as all four sites are generally within the common airshed of the major Sydney metropolitan region, so fine $PM_{2.5}$ ambient air composition is expected to vary little. Generally fine particles, once airborne, can stay in the atmosphere for days and weeks and hence have a better chance of mixing and dispersing, as well as the ability to be carried large distances by weather systems.

Source fingerprint contributions to $PM_{2.5}$ mass

The PMF source apportionment codes used here were developed by Paatero and Tapper (1994). Source apportionment is a one step process that looks at the inter-element correlations in multi-dimensional space and produces a set of correlated elements in a source fingerprint, as well as the contribution of each fingerprint to the total measured $PM_{2.5}$ mass.

For this analysis each site was optimally fitted by the seven fingerprints related to windblown soils, sea spray, industrial emissions, secondary sulfates, smoke from biomass burning and automobiles. These fingerprints were given the names Soil, Sea, Mixed-Ind-Saged, Mixed-2ndryS, Mixed-Smoke-Auto, Auto1 and Auto2. The term 'Mixed' was used to show that these are not fixed defined emission sources related to just one source but, because we are analysing at the receptor site, any one fingerprint can be a mixture of several different emission sources. This occurs because air moving over different sources on its way to the receptor site will pick up signatures from each of the emission sources it passes over.

The pie charts for the seven factor PMF fits to the full 15-year elemental dataset for the four sites, Lucas Heights, Richmond, Mascot and Liverpool, are shown in **Figure 2-3**.

At the Lucas Heights site, the $PM_{2.5}$ mass of $(4.93\pm2.64)\mu gm^{-3}$ was composed of 3% Soil, 17% Sea spray, 9% Mixed industrial and aged sulfate, 30% secondary sulfates, 12% smoke from biomass burning and diesel motor vehicles and 26% motor vehicles from two different sources.

At the Richmond site, the $PM_{2.5}$ mass of $(6.50\pm4.55)\mu gm^{-3}$ was composed of 4% Soil, 4% Sea spray, 15% Mixed industrial and aged sulfate, 26% secondary sulfates, 30% smoke from biomass burning and diesel motor vehicles and 20% motor vehicles from two different sources.

At the Mascot site, the $PM_{2.5}$ mass of $(7.59\pm4.27)\mu gm^{-3}$ was composed of 3% Soil, 8% Sea spray, 23% Mixed industrial and aged sulfate, 20% secondary sulfates, 23% smoke from biomass burning and diesel motor vehicles and 23% motor vehicles from two different sources.

At the Liverpool site, the $PM_{2.5}$ mass of $(8.17\pm5.36)\mu gm^{-3}$ was composed of 4% Soil, 9% Sea spray, 14% Mixed industrial and aged sulfate, 22% secondary sulfates, 32% smoke from biomass burning and diesel motor vehicles and 20% motor vehicles from two different sources.

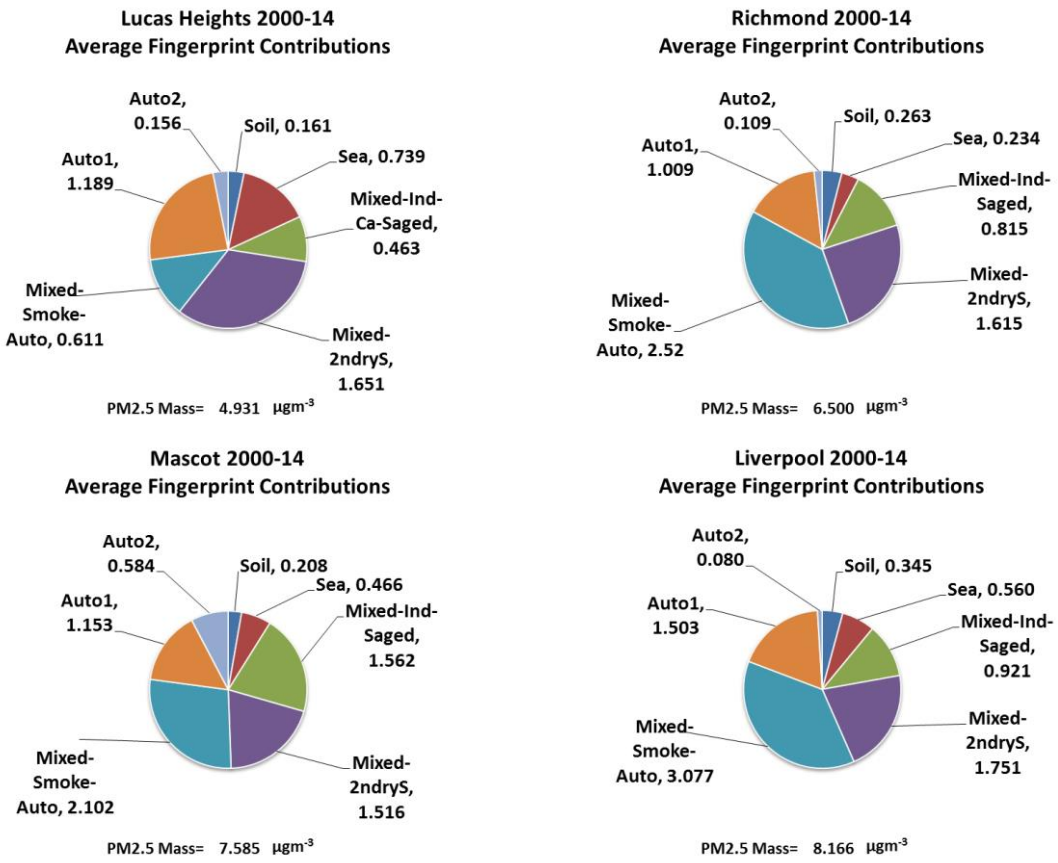

Figure 2-3. Pie charts for the seven factor PMF fits to the full 15-year elemental dataset for the four sites, Lucas Heights, Richmond, Mascot and Liverpool in the greater Sydney metropolitan region.

It should be noted that for all sites the Mixed-Ind-Saged and the Mixed-Smoke-Auto sources were a combination of both primary and secondary particles, whereas the 2ndryS source was mainly composed of secondary particles.

Fingerprint variations with time

In the four figures, **Figure 2-4** to **Figure 2-7** below, we plot the annual median concentrations for 15 years for each of the seven fingerprints for each of the sampling sites. We use the annual median values, as the median values are not as affected by extreme outlier events as the annual averages and hence are more representative of annual changes over time. For example, the average annual values for 2009 are affected by the major dust storm of September 2009 whereas the annual median value is not so affected.

The three major anthropogenic source fingerprints, Mixed-2ndryS, Mixed-Smoke-Auto and Auto1+2 (where Auto1+2 is the sum of Auto1 and Auto2), on average show significant declines in concentrations of 46%, 49% and 40% respectively over time from 2000 to 2014. Between 2007 and 2014, there seems to be no significant upward or downward trend in Mixed-Smoke-Auto and Auto1+2 fingerprints at Lucas Heights, Richmond and Mascot. Mixed-2ndryS appears to be generally reducing. The natural source fingerprints of Soil and Sea have much smaller concentrations and appear relatively stable over the 15-year study period.

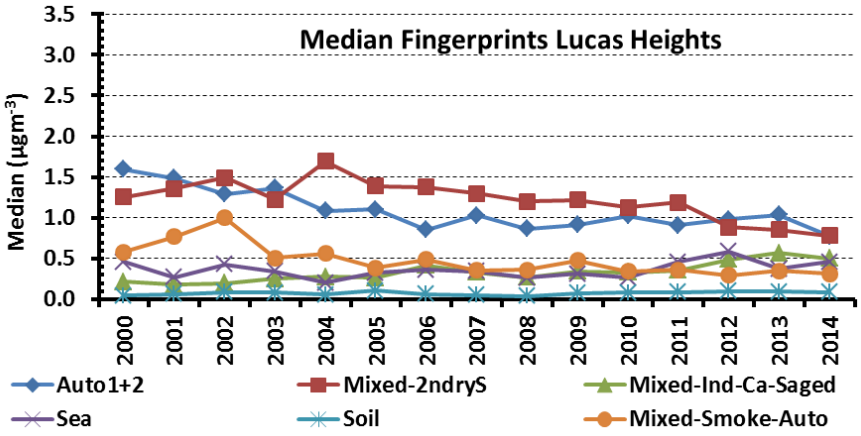

Figure 2-4. Annual median fingerprint contributions by year for Lucas Heights, 2000–14.

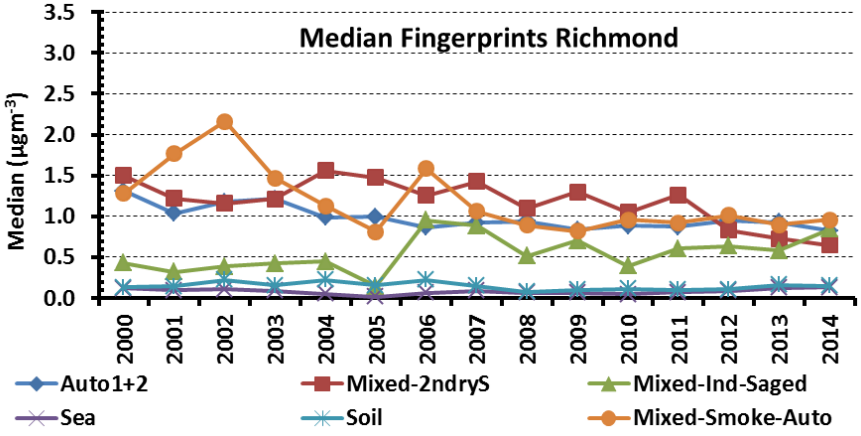

Figure 2-5. Annual median fingerprint contributions by year for Richmond, 2000–14.

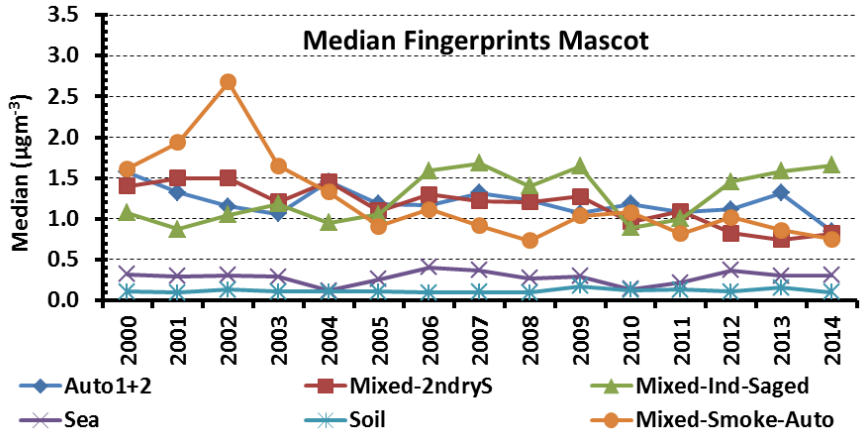

Figure 2-6. Annual median fingerprint contributions by year for Mascot, 2000–14.

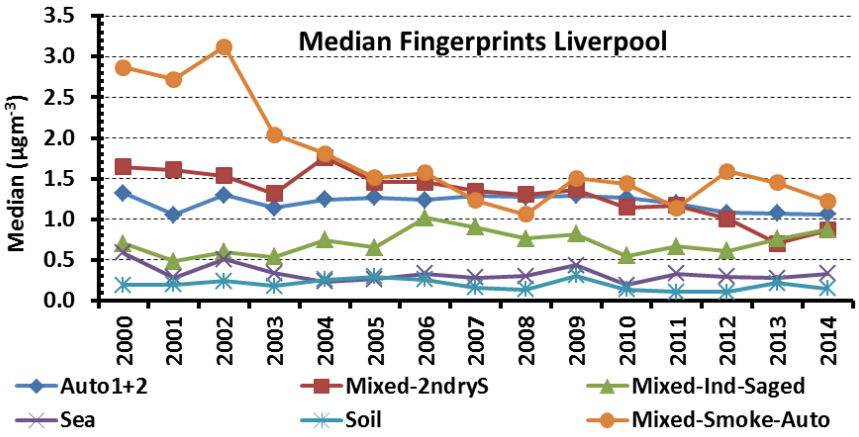

Figure 2-7. Annual median fingerprint contributions by year for Liverpool, 2000–14.

The Mixed-Ind-Saged fingerprint, representing industrial emissions at each site, varied significantly from year to year, probably depending on the prevailing meteorological conditions each year as well as changes in source contributions. The increase in this fingerprint contribution at Richmond, Mascot and Liverpool after 2006 is particularly notable as it coincided with the opening of the M7 motorway. The reason for this increase at several sites is unclear and warrants further investigation.

Domestic wood burning

Domestic wood heaters have a significant effect on ambient fine particle air quality, particularly in western parts of Sydney in the winter months.

- An obvious feature of the pie chart plots of **Figure 2-3** is the higher Mixed-Smoke-Auto factor contributions for Richmond and Liverpool compared with Lucas Heights and Mascot sites.
- These pie charts are for averages over the 15-year sampling period. Consideration of the time series data (Section 12) showed winter peaks clearly visible against the summer minima. These wintertime peaks reach 60–80% of the total measured gravimetric mass. In the summertime they fall to less than 10% of the total $PM_{2.5}$ mass.
- The Mixed-Smoke-Auto contributions at Liverpool during the winter months were on average about 5µgm^{-3} higher than at Mascot, and this is anticipated to be due to wood smoke and biomass burning not associated with controlled burning or bushfire events.
- Similar data were obtained for the Richmond site, showing that western Sydney is more significantly impacted by wood smoke and biomass burning not associated with controlled burning or bushfire events in the winter months.

Automobiles

The PMF seven fingerprint fits to the total $PM_{2.5}$ measured mass all gave two distinct automobile fingerprints, Auto1 and Auto2. These consistently occurred in approximately the same proportions whether 6, 7 or 8 source factors were used in the PMF fits. The Auto1 fingerprint was always much higher in concentration than the Auto2 fingerprint. At Mascot,

where the number of motor vehicles was higher than at the other sites, both Auto fingerprints were driven by between 15% and 20% of the total measured lead (Pb).

Leaded petrol was used in NSW up until January 2001 so we believe the Auto2 fingerprint being much smaller than the Auto1 fingerprint is probably directly related to the use of leaded petrol in motor vehicles. The lead associated with the Auto1 fingerprint is probably part of the re-entrained road dust kicked up by motor vehicle movements. Lead is still present in the Auto2 fingerprint in small quantities beyond 2001 because it is present in the retrained road dust kicked up by motor vehicles.

It is convenient to add these two fingerprints together to produce the Auto1+2 fingerprint, as the Auto2 contributions are generally small and the seasonal variations are similar.

The average across all four sites for the annual median Auto1+2 fingerprint concentration fell from 1.45µgm^{-3} in 2000 to 0.875µgm^{-3} in 2014. This was a decrease of 40% even though the average number of registered motor vehicles in NSW rose by more than 40% over this same time period. This can only mean that motor vehicles in NSW have become cleaner and less polluting over the past 15 years.

Over the 15-year study period Auto1+2 fingerprints at Liverpool, Lucas Heights, Mascot and Richmond contributed 10–25%, 15–40%, 10–30% and 10–25%, respectively.

It should be pointed out these Auto1+2 percentages are probably an underestimate of the total contributions from motor vehicles since there are further automobile components in the Mixed-Smoke-Auto and the Mixed-2ndryS fingerprints which also originate from motor vehicle emissions. Automobiles are still a significant source of ambient air pollution in the Sydney region. Light, medium and heavy diesel motor vehicles are major contributors to this type of fine particle air pollution.

Secondary sulfates

Analysis showed that, in Sydney, on average about 24% of the $PM_{2.5}$ mass is ammonium sulfate. The PMF analysis, across all the Sydney sites, showed that the total sulfur (S) mass was mostly associated with two of the seven fingerprints: Mixed-2ndryS (~80% of total S) and Mixed-Ind-Saged (~20% of total S).

Based on monthly averages, the contributions to the Mixed-2ndryS fingerprints are between 10% and 40% across all sites but are higher at the inland sites of Liverpool, Lucas Heights and Richmond than at the coastal Mascot site. For the Mixed-Ind-Saged fingerprints, contributions lie between 5% and 40% across all sites but are higher at Mascot, which is impacted more by industry around the international airport and Port Botany shipping and container depot.

Significant secondary sulfate sources in the study area include coal fired power stations, oil refineries at Rosehill and Kurnell, metals industry in Port Kembla to the south of Sydney, shipping in ports, diesel vehicles and multiple light to medium industries. The Clyde refinery ceased operation in 2012 and the Kurnell refinery in 2013. Both refineries are still used as major fuel storage depots.

The NSW Environment Protection Authority (NSW EPA) air emission inventory for 2008 showed that 98% of the total 246kT of sulfur dioxide emissions in NSW were associated with the coal fired power stations to the north and west of Sydney, burning 25 million tonnes of coal annually. This coal typically contains at least 0.5% sulfur by weight. We have shown in previous publications (Cohen et al. 2014 and Appendix C) that this secondary sulfate from coal combustion is transported into the Sydney basin to sites like Liverpool and Richmond, so we

expect much of this Mixed-2ndryS and part of the Mixed-Ind-Saged to be associated with coal burning for power generation.

Key signature elements for coal burning are sulfur (S), arsenic (As) (not measured here) and selenium (Se). Arsenic cannot be measured by x-ray techniques used here because it overlaps with lead, which occurs at concentrations an order of magnitude higher than arsenic. We see around 20% of the total measured selenium occurs in the Mixed-2ndryS at each of the sampling sites, further evidence that coal burning is a significant contributor to secondary sulfates in the Sydney basin.

Linking of hourly back trajectory wind data to high sulfate days has been used by Cohen et al. 2014 (see Appendix C) to link sulfate concentrations in the Sydney basin with coal fired power stations emissions in NSW. These techniques should be further utilised to identify and quantify the origins of secondary sulfate emissions in NSW. Chemical transport modelling could also be used to provide predictions of source contributions to secondary sulfate within the Sydney basin.

Conclusions

In summary, this study identified four main pollution sources or source factors in the Sydney airshed between 2000 and 2014 which significantly contributed to the $PM_{2.5}$ mass affecting all four sites. These were:

- smoke from domestic wood heaters which contribute significantly to fine particle levels particularly in western Sydney in winter months, peaking between 60% and 80% in the wintertime at some sites
- secondary sulfates from the burning of large amounts of coal for power generation and fine particles from industry and motor vehicles mainly generated by light, medium and heavy diesel vehicles, peaking between 50% and 70% in the summer months at some sites
- significant contributions to the $PM_{2.5}$ mass loading from industrial sources represented by the Mixed–Ind-Saged source fingerprints at all of the four sites studied during the 15 years from 2000–2014, peaking between 30% and 50% in the summer months at some sites
- vehicle emissions which contribute to the Auto1, Auto2 and Mixed-Smoke-Auto source fingerprints; reductions in vehicle contributions occurred during the period despite increases in the vehicle fleet, but such reductions seem to have tailed off more recently with Auto1+2 contributions of about $1\mu gm^{-3}$ or between 10% and 15% of the annual $PM_{2.5}$ mass in 2014.

The approach adopted focused on identifying sources which contribute significantly to daily, seasonal and annual average concentrations. During the analysis, smoke from biomass burning was identified as a significant intermittent source contributing to short-term spikes in fine particle ambient mass concentrations.

A database of each of the seven source fingerprints for each sampling day for each of the four sites has been produced and has been made readily available. An instruction manual to do this has been reproduced in full in Appendix B. All the data in this database is fully accessible and extractable by the user.

Recommendations

Taking into consideration the general findings from this report and what needs to be followed up in the future, we make the following recommendations:

- The number of elemental and chemical species should be expanded from the basic 23 currently used. The present study did not directly measure magnesium (Mg), arsenic (As), ammonium ions, nitrate ions, elemental carbon (EC) or organic carbon (OC) which are key signatures for several sources.
- The measurement of chemical species like levoglucosan and mannosan tracers for biomass burning, methanosulfonate (MSA–) and oxalate ($C_2O_4^{2-}$) ions would help to better differentiate biomass burning, fossil fuel combustion, automobile and individual industrial sources. The measurement of these species, together with the missing species mentioned above, could bring the total elemental and chemical species list to over 35, making PMF analysis much more effective in identifying and quantifying more source types.
- The combination of PMF sources with wind and back trajectory modelling for every hour of every sampling day over long timeframes and tied to major known point source locations (as done in recent papers by Cohen et al. 2012 and 2014, Appendix C) should be undertaken to identify and quantify $PM_{2.5}$ windblown soil and secondary sulfate sources at other existing long-term sampling sites in NSW.
- The PMF particle characterisation data provided by this study should be used in the development and validation of chemical transport models which can then be used to provide a refined projection of source contributions to both primary and secondary particles.

It should be pointed out the IBA analysis techniques used at ANSTO are non-destructive, whereas the additional analyses suggested in the recommendations above such as ion chromatography (IC) and (EC/OC) analyses destroy the filters.

Also, historical filters from this study are not available for any further destructive analysis. The recommendations above therefore relate to the collection of future samples.

Glossary

2ndryS	Secondary sulfate aerosol obtained after the conversion of sulfur dioxide gas to sulfate particles
ABS	Australian Bureau of Statistics, Canberra
Air NEPM	National Environment Protection (Ambient Air Quality) Measure; it includes ambient air standards for $PM_{2.5}$ of 8µgm^{-3} annual average, and 25µgm^{-3} 24-hour maximum
Al	Aluminium
ANSTO	Australian Nuclear Science and Technology Organisation, Lucas Heights, Sydney
ASP	ANSTO's Aerosol Sampling Program, in operation since the 1990s
BC	Black carbon, usually sourced from diesel motor vehicles and low temperature burning of carbonaceous materials
Br	Bromine
Ca	Calcium
Cl	Chlorine
Co	Cobalt
CMass	Name given to $PM_{2.5}$ gravimetric mass in ngm^{-3} or µgm^{-3}
CMB	Chemical mass balance codes for source apportionment
Coarse particles	Those particles with aerodynamic diameters between 2.5µm and 10µm
Cr	Chromium
Cu	Copper
F	Fluorine
Fe	Iron
IBA	Ion beam analysis
FitMass	Sum of all the calculated PMF source masses to be compared with the total $PM_{2.5}$ gravimetric mass
K	Potassium
K_{non}	Non soil potassium (K_{tot}-0.6*Fe), an estimator for smoke from biomass burning
LIPM	Laser integrated plate method for measuring black carbon concentrations
MDL	Minimum detectable limit for a measurement
Mg	Magnesium
Mixed-Ind-Ca-Saged	Name given to a mixed industry, calcium (probably cement/ building) aged sulfate aerosol source
Mixed-2ndryS	Name given to the mixed secondary sulfate fingerprint generated from sources such as coal burning for power generation, heavy fuel oil burning, oil refining and automobiles
Mixed-Smoke-Auto	Name given to a mixed smoke from biomass burning (high K) and smoke from automobile (probably diesel trucks or buses) source
Mn	Manganese
Na	Sodium
NATA	National Association of Testing Authorities
ng	Nanogram (1ng =0.000 000 001 gram =10^{-9} gram)
nssSulfate	Non sea salt sulfate, sulfate ions not associated with sea salt
Ni	Nickel
NO	Nitric oxide
NO_2	Nitrogen dioxide
NOx	Oxides of nitrogen (commonly NOx = NO + NO_2)
NSW EPA	NSW Environment Protection Authority
O_3	Ozone
OEH	Office of Environment and Heritage
OMH	Organic matter estimated from the total hydrogen content of a filter

Organics	Organic aerosol, assumed to be 9% hydrogen, 20% oxygen and 71% carbon as per Malm et al. 1994
P	Phosphorous
Pb	Lead
PCA	Principal Components Analysis method
PESA	Proton elastic scattering analysis
PIGE	Proton induced gamma ray emission
PIXE	Proton induced x-ray emission
$PM_{2.5}$	Particulate matter of aerodynamic diameter <2.5µm
PM_{10}	Particulate matter of aerodynamic diameter <10µm
PMF	Positive matrix factorisation – a source apportionment code developed in the mid-1990s
RCM	Reconstructed mass obtained by adding together all the individual analysed masses within one filter
RBS	Rutherford backscattering
S	Sulfur
SD	Standard deviation of a measured result
Se	Selenium
Si	Silicon
Soil	Name given to a soil source usually dominated by oxides of Al, Si, Ti, Ca and Fe
Ti	Titanium
TotN	Total nitrogen measurement, usually the sum of nitrogen in ammonium and nitrate ions
TSP	Total suspended particulate matter
V	Vanadium
Zn	Zinc
µm	Micrometre (1µm =0.000001 metre =10^{-6} metre); also called a micron
µg	Microgram (1µg =0.000001 gram =10^{-6} gram); one millionth of a gram
μgm^{-3}	Microgram per cubic metre

3. Introduction

ANSTO has been involved in fine airborne particulate matter research since the 1990s under its Aerosol Sampling Program (ASP). Over that time ASP has continuously collected and characterised 24-hour $PM_{2.5}$ ambient air filters twice weekly at dozens of sites across Australia and Asia.

Fine particles are defined as particulate matter with aerodynamic diameters less than 2.5µm ($PM_{2.5}$). These small particles have significant health effects, can travel hundreds of kilometres once airborne and are just the right diameter to scatter and absorb white light, making the pollution they produce clearly visible to the human eye.

Each ASP filter has been analysed for fine mass as well as the concentration of over 20 key elements using standard accelerator based ion beam-analysis (IBA) techniques. At ANSTO we currently have a well characterised $PM_{2.5}$ database consisting of tens of thousands of filters, representing nearly 100 different sampling sites and decades of $PM_{2.5}$ data and analyses.

$PM_{2.5}$ particles are produced mainly by combustion processes such as combustion of fossil fuels in motor vehicles, coal burning in power stations, oil refineries and biomass burning during bushfire episodes or in domestic wood heaters. Coarser particles, such as PM_{10} or total suspended particulate matter (TSP) are mainly produced by mechanical processes such as windblown soils and sea spray and grinding or cutting of materials. This study is exclusively dedicated to $PM_{2.5}$ ambient air pollution; its mass, composition and sources.

In this study we apply positive matrix factorisation (PMF) analysis to our multi-elemental $PM_{2.5}$ data obtained at four Sydney sites: Richmond, Mascot, Liverpool and Lucas Heights, covering the 15-year period from 1 January 2000 to 31 December 2014. The PMF analysis is used to determine average receptor source fingerprints at each site, as well as any trends that may be associated with those receptor sources over a range of time scales including daily, monthly and yearly.

The primary aims of this study were to:

- convert the existing 15-year $PM_{2.5}$ mass and elemental datasets for four given sites in the Sydney basin into identifiable source fingerprints
- quantify the absolute and the percentage contribution of each of these fingerprints to the total fine $PM_{2.5}$ mass
- provide seasonal and annual variations for each of the source fingerprints
- provide a readily accessible database containing the daily source fingerprints together with their contributions covering the 15-year period from 2000 to 2014 for four given sites in the Sydney basin
- if possible, identify and quantify the major contributors of fine particle pollution to the ambient air quality in Sydney.

It is important to point out that this study is a receptor site modelling study not an emissions inventory study. Filters were collected at the corresponding receptor sites. These filters contain mixed information about a range of emissions sources arriving at the site over a 24-hour period. Some of these source emissions will have travelled hundreds of kilometres before being collected at our sites. During this period the original source emission profiles have been

changed by atmospheric chemistry as well as mixed with other source emissions passed over on the way to our sites. For example, pure sea spray from coastal regions, driven inland by afternoon sea breezes, can be dramatically altered by sunlight, chemistry and other emissions by the time it reaches sampling sites like Richmond in western Sydney. It is no longer mostly sodium chloride (NaCl) but has large chlorine losses in the presence of sulfates from other sources and has picked up significant black carbon (BC) from diesel motor vehicles. Consequently the source fingerprints and their contributions we report on here are receptor site fingerprints and not profiles of individual emission sources. This is an important distinction to make.

This document describes the four sampling sites and sampling conditions as well as the methodology used to obtain receptor source fingerprints and their contributions to the total measured $PM_{2.5}$ mass. The fingerprints obtained at each sampling site are unique to that site and represent a range of mixed sources arriving at that site over the sampling period of 2000 to 2014. The contribution of each of these fingerprints is also presented on a daily, monthly and annual basis at each of these sites for the sampling period of 2000 to 2014.

$PM_{2.5}$ sampling sites and sampling conditions

$PM_{2.5}$ samples were collected on Teflon filters at each of the four Sydney sites (see **Figure 3-1**): Lucas Heights, Richmond, Mascot, and Liverpool, using ANSTO ASP cyclone units. These units were constructed at ANSTO and are based on a standard US EPA approved IMPROVE $PM_{2.5}$ cyclone unit (see **Figure 3-1**) with a mass-flow controlled rated at $22 Lmin^{-1}$ which equates to a volume of approximately $32 m^3$ over a 24-hour sampling period. Samples were collected at each site from midnight to midnight, on 25mm diameter $250 \mu g cm^{-2}$ thick stretched Teflon filters every Wednesday and Sunday.

Figure 3-1. Left: ASP $PM_{2.5}$ cyclone sampling unit at Lucas Heights site. Right: An exposed 25mm stretched Teflon filter in its container. Photos: C Thompson/ANSTO

A map of the sampling site locations in relation to the Sydney basin area is presented in **Figure 3-2**. The longitude and latitude of each of the sites is given in **Table 3-1**.

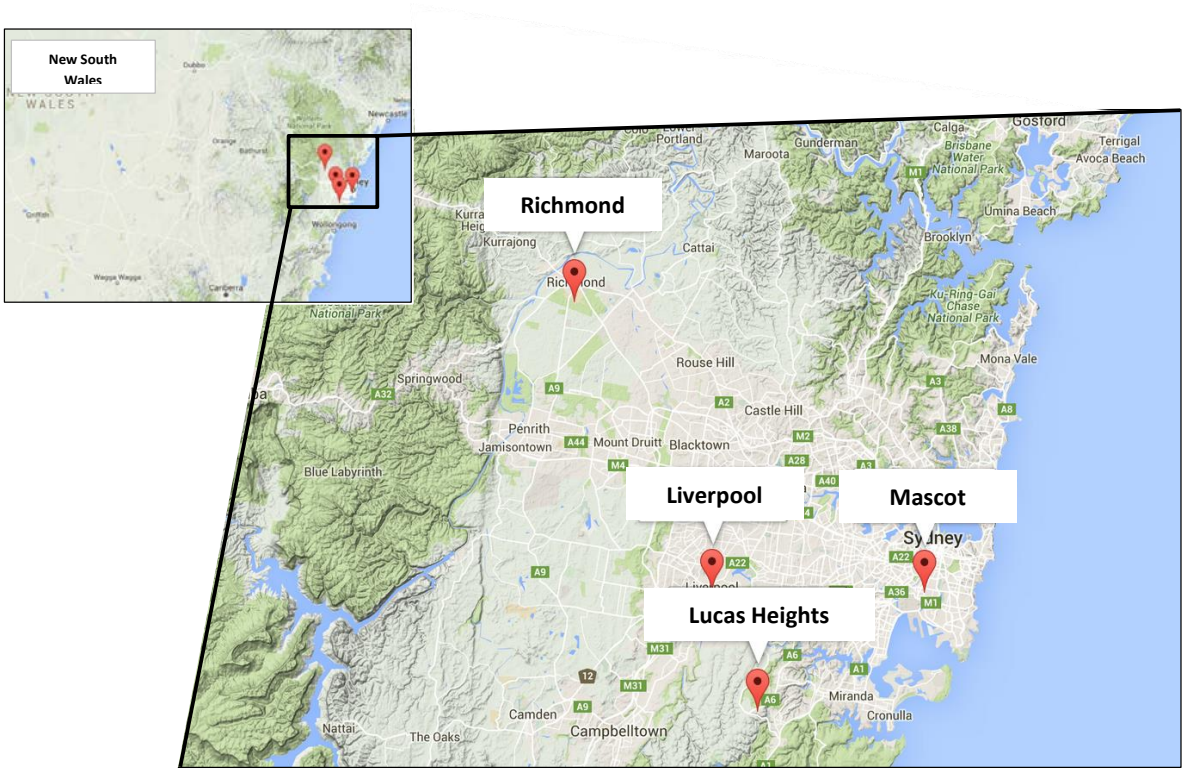

Figure 3-2. Map of the Sydney basin and locations of the sampling sites.

Table 3-1. Longitude and latitude of each of the four Sydney sampling sites.

Sites	Long (x)	Lat (y)
Lucas Heights	150.9825	–34.0517
Richmond	150.7483	–33.6181
Mascot	151.1956	–33.9260
Liverpool	150.9079	–33.9315

Lucas Heights

This is a semirural/urban site with some small and medium industries as well as a significant landfill garbage site within 3km of the sampler. There is a major busy road running north–south within one kilometre of the site on its western side. The site was also impacted by significant construction works, involving roads and new buildings during the period 2012–13 inclusive.

Richmond

This is a rural/semiurban site in north western Sydney within the grounds of the University of Western Sydney. It has the town of Richmond 2km to the north and the town of Londonderry 3km to the south. It is 52km NW from the Sydney central business district (CBD). It is known to be impacted by smoke from domestic wood burning for heating in the winter months.

Mascot

This site is the closest to the Sydney CBD being 7km south of the CBD, 7km north west of major container shipping operations at Port Botany and 2.5km north east of Sydney's main international airport. It is located on the first floor roof of the Botany City Council building, adjacent to Botany Road, which carries significant local traffic as well as major diesel truck and bus movements associated with the Botany Port and the international airport. Botany Road has significant city traffic flows associated with the morning and afternoon peak hours. This site is expected to be heavily impacted by heavy, medium and light motor vehicle emissions.

Liverpool

This site is co-located in western Sydney within a standard NSW Office of Environment and Heritage (OEH) ambient air monitoring site. Liverpool is considered a significant western Sydney urban area with known pollution from the Sydney coast and the CBD being transported into the area during afternoon sea breeze events especially in the summer months.

Samples and related data collected between 1 January 2000 and 31 December 2014 were included in this source apportionment study. This represents around 1500 daily filters collected at each sampling site. It should be noted that data associated with a relatively small number of filters in this time period have been excluded from the PMF analysis. This is described in detail later, but generally related to issues associated with sampling such as incorrect exposure time, excessively high or low pump flow rates or damaged filters during transport and handling. Consequently there may be a few small gaps in the standard Wednesday/ Sunday data time series over the study period.

4. Sample analysis and $PM_{2.5}$ mass

Each stretched Teflon filter containing sampled $PM_{2.5}$ ambient air was run for 24 hours. If the average Sydney air contains around $10\mu gm^{-3}$ of $PM_{2.5}$ particles and the pump flow rates were $21Lmin^{-1}$, then the average expected mass collected on each filter would be only around 300µg. Hence very sensitive measurement techniques are required to analyse samples as small as this for many different chemical species.

Gravimetric mass

Filters for each site were weighed, on a microbalance, pre and post exposure to determine the gravimetric $PM_{2.5}$ particulate mass to better than ±2µg. Gravimetric mass measurements were performed on a NATA calibrated Mettler Toledo MX5 microbalance (readability ±1µg) under our standard laboratory conditions of (22 ± 5)°C and relative humidity of (50 ± 10)%.

Black carbon

Black carbon (BC) was determined using the standard laser integrating plate (LIPM) absorption method with a HeNe 633nm wavelength laser and assuming a mass absorption coefficient of $7m^2g^{-1}$ (Taha et al. 2007) for the fine particle fraction (see Equation (8) below for further details).

Elemental analysis

Each filter was analysed using a combination of multi-elemental accelerator based ion beam analysis (IBA) techniques (Cohen et al. 1996, 2004a, 2004b); proton induced x-ray emission analysis (PIXE), proton induced gamma-ray emission analysis (PIGE), Rutherford backscattering (RBS) and proton elastic scattering analysis (PESA). Calibration of these IBA techniques was performed using thin film reference standards (Micromatter Pty Ltd) certified by mass to ±5%. These techniques, applied simultaneously and non-destructively, utilise an 8–10mm diameter proton beam of 2.6MeV energy with ion beam currents of around 10nA for an analysis time of around 3–5 minutes. This combination of applied IBA techniques measures the following most commonly occurring elements in airborne fine particles: H, Na, Al, Si, P, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Se, Br, Pb, BC and TotN to concentrations down to 1ngm^{-3} of air sampled. TotN is the total nitrogen content as measured by RBS. Note also that IBA measurements measure the total elemental concentration, not just the soluble part or the ionic part. For example, H is the total hydrogen measurement which includes hydrogen in ammonium ions, in organic compounds and any water of crystalisation chemically attached to compounds.

The resulting fine particle elemental database, which provides the input data for PMF analysis, contains the following data related to each analysed filter: the $PM_{2.5}$ mass (ngm^{-3}), elemental concentrations (ngm^{-3}), error and minimum detectable limit (MDL) values (ngm^{-3}).

Errors

The error value assigned in PMF analysis for each elemental concentration value is comprised of the system calibration errors (*Calib*), the experimental measurement errors (*Expt*) and statistical counting errors (*Stat*) which are added in quadrature:

$$Err = \sqrt{Calib^2 + Expt^2 + Stat^2} \qquad (1)$$

For example, these three component errors might be ±5%, ±3% and ±10%, respectively. The statistical counting error depends on the square root of the counting time. Generally for major peaks in our IBA analysis such as sulfur we acquire an area of more than 10,000 counts which would have a statistical counting error of ±3%. Smaller trace element peaks like nickel might have an area of only 100 counts for which Stat =±30%. Note the Err term should not be zero for the PMF analysis to function properly.

Minimum detectable limit (MDL)

The minimum detectable limit (MDL) assigned to each concentration value represents an estimate of the experimental minimum mass concentration for each of the measured elemental species in (ngm^{-3}). The *MDL* is determined as three standard deviations above the background area (converted to a concentration) times the background under each peak in the IBA spectra used to determine the corresponding elemental concentrations, that is:

$$MDL = 3(Bkg)^{1/2} \qquad (2)$$

This is a very conservative estimate as we could easily work with one or two standard deviations above the background. Note the *MDL* term should not be zero for the PMF analysis to function properly.

$PM_{2.5}$ mass at each site

Figure 4-1 shows the daily $PM_{2.5}$ mass measured at each of the four sites during the study period from 1 January 2000 to 31 December 2014. The horizontal dashed line represents the recommended 24-hour Air NEPM maximum value for $PM_{2.5}$ of 25µgm^{-3}. Clearly there are several mass exceedances of the Air NEPM maximum at each site over the study period. These are primarily due to dust storms, bushfire events or controlled burning for bushfire control. These are discussed and listed in more detail in the sections below. Note the vertical axis in the plots of **Figure 4-1** have been fixed to 50µgm^{-3} so daily variations are clearly visible. This means some outlier events above 50µgm^{-3} will not be plotted.

The average $PM_{2.5}$ masses at Lucas Heights, Richmond, Mascot and Liverpool sites during the study period were 4.93µgm^{-3}, 6.50µgm^{-3}, 7.59µgm^{-3} and 8.17µgm^{-3}, respectively.

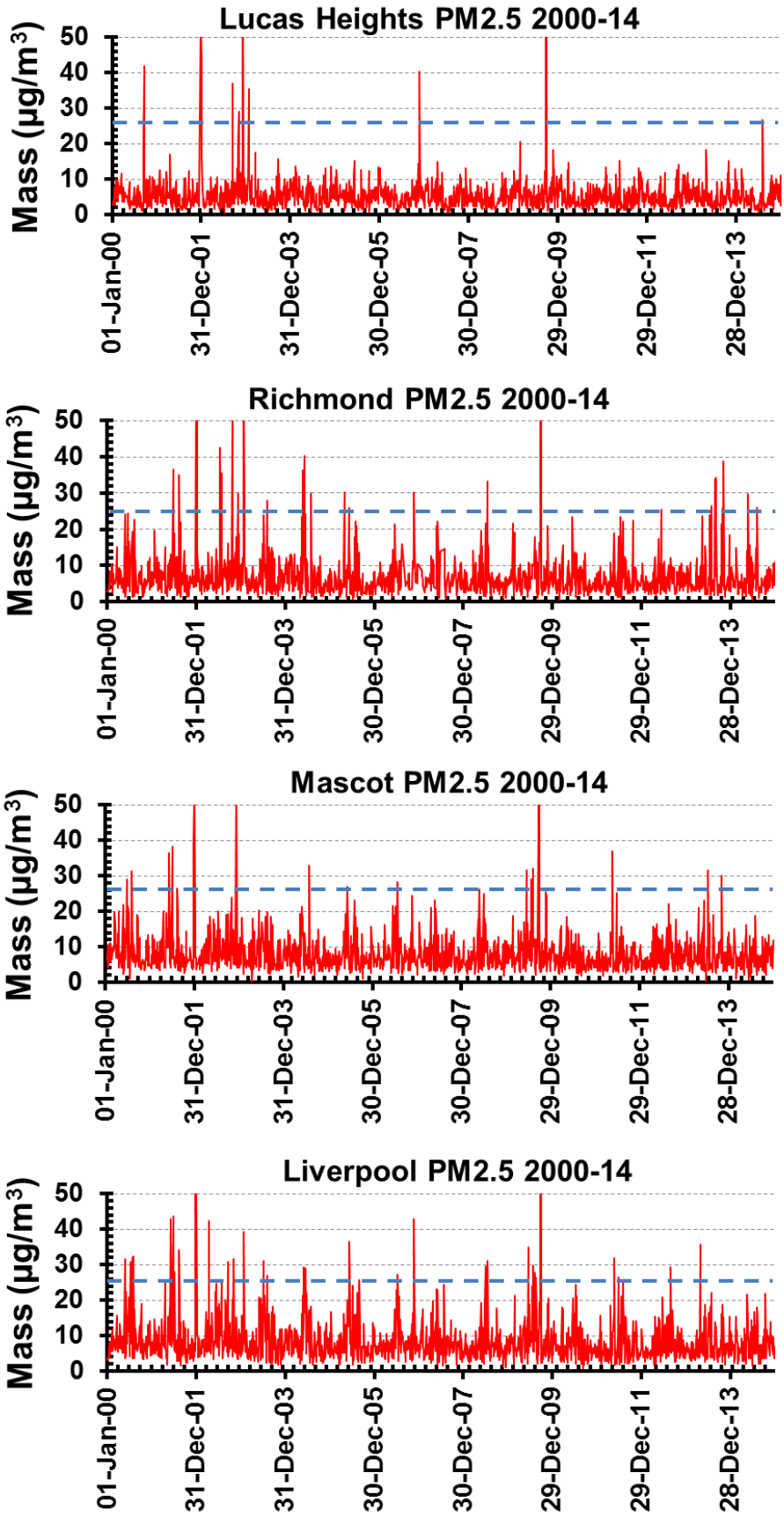

Figure 4-1. Fine $PM_{2.5}$ mass at each of the four sites, 2000–14. The dashed horizontal line shows the Air NEPM 24-hr exceedance level of 25µgm^{-3}.

$PM_{2.5}$ lead at each site

Figure 4-2 shows the daily $PM_{2.5}$ lead measured at each of the four sites during the study period from 1 January 2000 to 31 December 2014. Note that the higher lead levels at Lucas Heights, Richmond and Liverpool pre December 2001 were due to the use of leaded petrol in the Sydney area. Leaded petrol sales ceased in Sydney in January 2001. Once leaded petrol sales ceased the lead in Sydney's air dropped to levels below 50ngm^{-3} from their winter peaks of over 1500ngm^{-3} in earlier decades (Cohen et al. 2005).

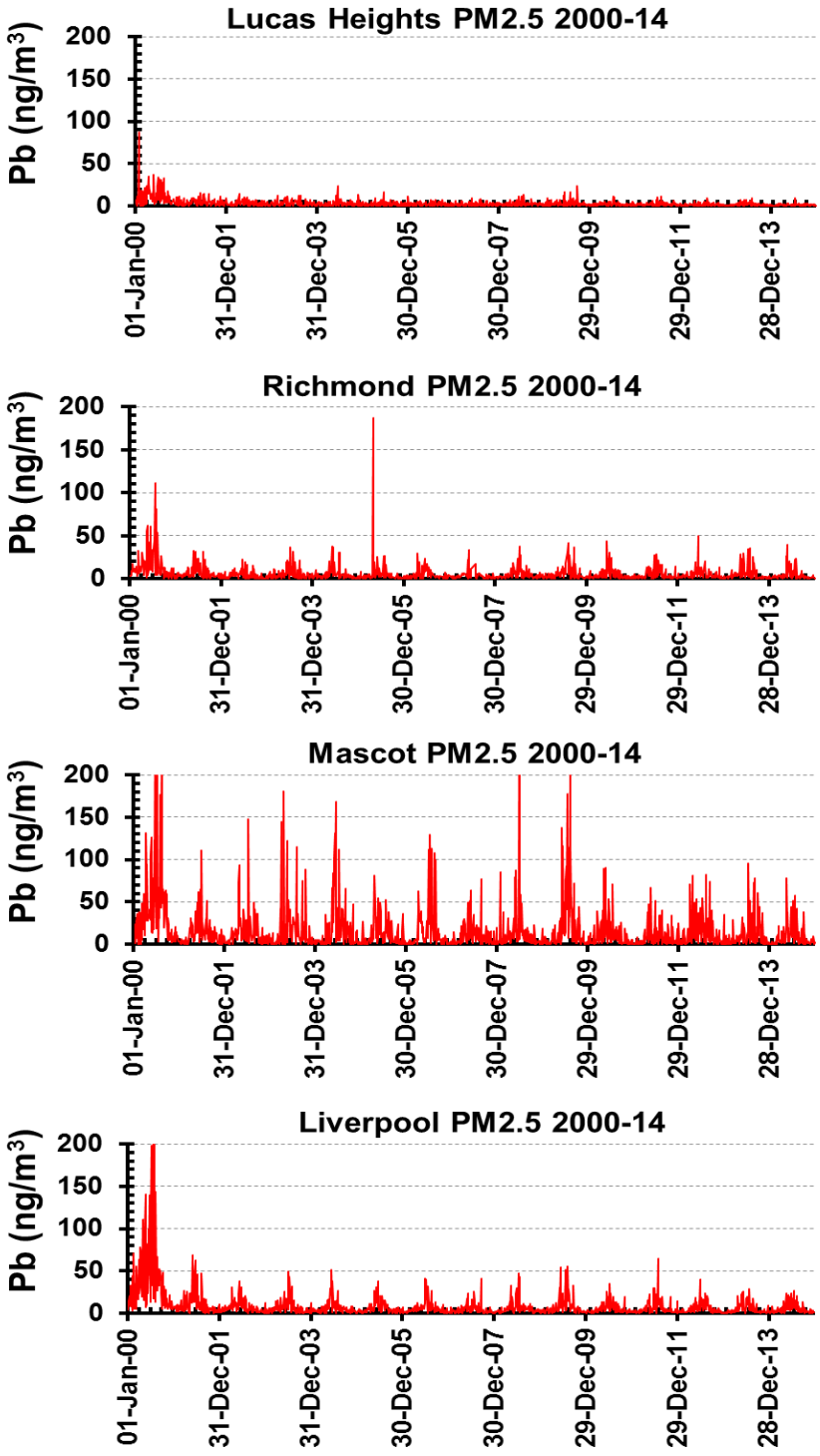

Figure 4-2. Fine $PM_{2.5}$ lead (Pb) at each of the four sites, 2000–14.

As Mascot is dominated by motor vehicles, its lead levels in air stayed higher for longer than the other sites post January 2001, especially in the winter months. Bromine is an additive in leaded petrol and so should correlate with lead.

A plot of the bromine lead concentrations at Mascot for 2000–01, shown in **Figure 4-3**, demonstrates that for this earlier period leaded petrol was the major source of lead. However, at Mascot for the period 2002–14 this bromine lead correlation no longer exists so the increased lead levels at Mascot in the winter periods compared to the other three Sydney sites must be due to a lead source other than leaded petrol.

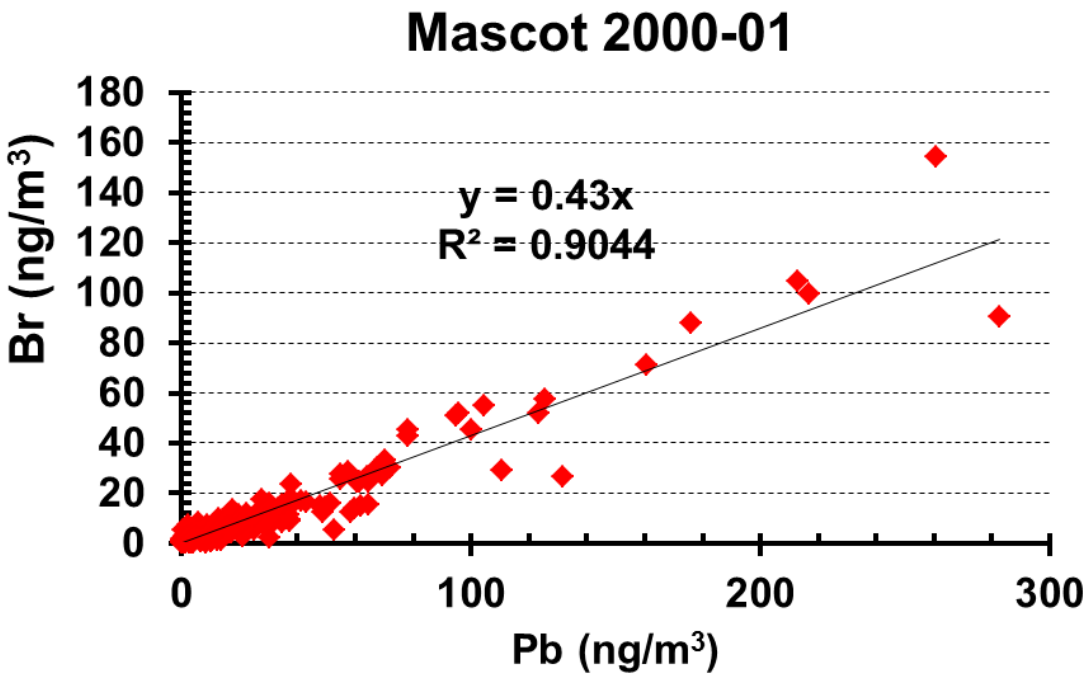

Figure 4-3. Plot of bromine versus lead for the period 2000–01 at Mascot.

5. Average elemental composition all sites

For the PMF source apportionment techniques to function it is important that the set of elements or chemical species measured should span the expected sources to be determined. There is always a compromise between what we can measure easily, with sufficient sensitivity and precision, and what might be a key signature species for a given source. For $PM_{2.5}$ fine particles we require elemental or chemical species measured down to around $1ngm^{-3}$ of air sampled. For our Teflon filters with a collection area of $2.27cm^2$ and total volume of air sampled in 24 hours this corresponds to measurement sensitivities on a filter of around $13ngcm^{-2}$ or less.

For this 15-year study we have used accelerator based IBA techniques which have the required measured sensitivities and can non-destructively measure selected elements from hydrogen to uranium in the periodic table in a few minutes of machine running.

Table 5-1 shows the 23 species used in this study, their measured average, median, standard deviation (SD) and maximum values over all four sites and covering the study period from 1 January 2000 to 31 December 2014. This represents 5776 sampling days from the four sites.

Table 5-1. The average, median, standard deviation and maximum values of the 23 elemental and chemical species used in the 15-year study.

All sites	Species	Average (ngm^{-3})	Median (ngm^{-3})	SD (ngm^{-3})	Max. (ngm^{-3})
$PM_{2.5}$	**H**	257	171	351	14,940
5776 days	**Na**	338	236	374	3,492
	Al	22.2	8.7	111	5,260
	Si	78.3	39.3	315	15,059
	P	4.5	1.7	36.1	2,696
	S	391	309	294	2,750
	Cl	286	135	388	3,816
	K	61.3	40.0	85.7	3,086
	Ca	30.3	23.3	39.7	1,660
	Ti	3.8	2.0	22.9	1,230
	V	0.8	0.5	1.3	30
	Cr	0.5	0.3	1.1	54
	Mn	2.0	1.2	5.3	267
	Fe	59.3	34.4	242	12,815
	Co	0.4	0.3	1.7	92
	Ni	0.6	0.3	1.1	40
	Cu	2.4	1.5	3.1	59
	Zn	11.5	6.1	17.0	251
	Se	0.8	0.4	1.5	34
	Br	3.8	2.5	7.8	329
	Pb	8.5	3.4	18.0	399
	BC	1,136	857	938	8,973
	TotN	351	201	501	6,777

Soils are generally well represented by the oxides of Al, Si, Ti, Ca and Fe. They can be soils blown in by dust storms, from agricultural activities and from retrained road dusts. Sea spray from coastal regions can travel hundreds of kilometres inland depending on meteorological conditions. Secondary sulfates arise from the conversion of sulfur dioxide gas to sulfate particles and their subsequent neutralisation to ammonium sulfate. Industry sources have a very broad range of key signatures depending on the types of industries contributing to the pollution. Ferrous metal, Cu, Zn, Fe, industrial heavy oil combustion from shipping and the like contains S, V and Ni. Cement and building industries have high Ca, K and BC components.

Table 5-2. The elemental species that might be associated with some fine particle sources; this list is not exhaustive, just representative.

Source	Key signature species
Soil	Al, Si, Ti, Ca, Fe, O
Sea spray	Na, Cl, S, Mg, Br
Secondary sulfate	S, H, V, Ni, BC, TotN
Industry	Na, S, Ca, Fe, Cu, Zn, Mn, V, Ni, Cr, Co, Pb, BC, TotN
Smoke	H, Na, Al, Si, S, K, Ca, Fe, Zn, BC, TotN
Automobiles	H, S, Ca, K, Fe, Cu, Zn, Pb, Br, BC, TotN

Fine particle smoke from biomass burning is well characterised by K, Ca, S, Zn and BC. Before 2001 Australian petrol driven vehicles had a clear Pb, Br fingerprint. Since the withdrawal of leaded petrol we have to now rely on less well defined signatures like Zn from tyre wear, Cu from brake pads or H, S, K, Ca from engine oils.

Clearly the species we are measuring span most of these source types and so should be easily identifiable.

Figure 5-1 shows the minimum detectable limits and the typical measured errors for our IBA techniques for the set of elements and species given in **Table 5-2** above.

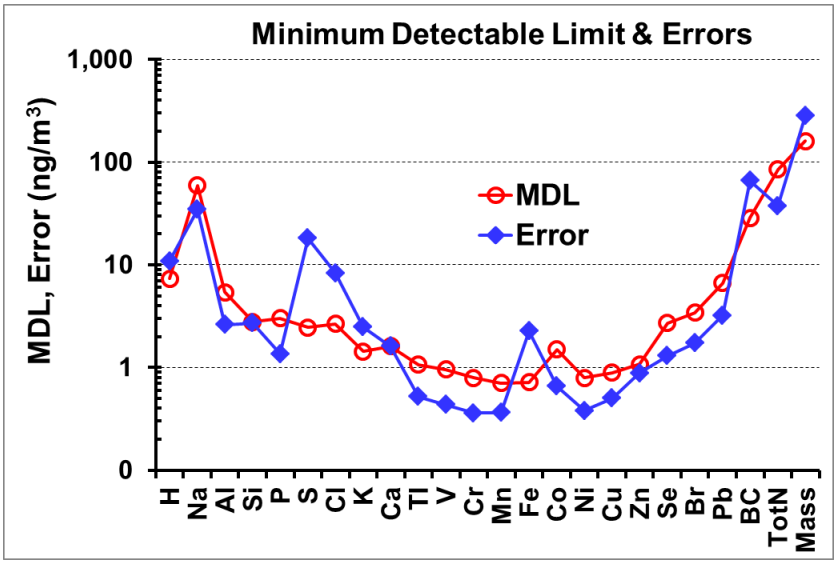

Figure 5-1. Median MDLs and errors for 23 measured elemental species and the gravimetric mass measurement across all four sampling sites.

We see that for most species we have excellent sensitivity. The main area of weakness is for species above Zn, like arsenic (As) and selenium (Se), which occur in ambient fine particle air pollution at levels generally below 1ngm^{-3} and are key signatures for coal burning. Currently we have to rely on secondary sulfate, selenium and other associated species to detect a coal signature.

Database error codes

It is important to note that over the 15-year study period it is reasonable to expect problems with sampling/analysis to occur from time-to-time. These may include: a filter exposed for longer than the 24-hour period, a filter which was exposed for less than 24 hours or failed to be exposed at all, a fault with the sampling unit, excessively high (or low) air flow rates, or a filter damaged during transport/handling. Some sites have even been struck by lightning. In these instances, the filter (where possible) is still analysed, however a data '*Error Code*' denoted by the numbers '0 to 9' is included in a designated column alongside the related data. Generally all *Error Code* =0 data are good and included in the analysis, *Error Code* =1–5 warns that something was not normal but the data is probably acceptable, while *Error Code* =6–9 indicates problematic data that should not be included in the PMF source apportionment analysis.

6. Pseudo-element concentrations

The range of measured elemental species provides an opportunity to estimate several chemical forms of the $PM_{2.5}$ fine mass. We call these *pseudo-elements*. Based on Malm et al. 1994, the pseudo-elements are defined below as:

Salt

Salt is generally only a significant factor in and around marine environments, in the form of sea spray, however sea salt may also be transported hundreds of kilometres inland, so we expect it to be present at all four of our sampling sites. Salt is estimated from the measurement of sodium assuming it occurs as sodium chloride crystals in the $PM_{2.5}$ mass, hence:

$$\text{Salt} = 2.54[\text{Na}] \quad (3)$$

A sea salt factor can often show the seasonal variations in wind direction for coastal or near coastal localities.

Ammonium sulfate

This is an estimate of the amount of particulate ammonium sulfate [$(NH_4)_2SO_4$] contained on a filter. It is determined from the measurement of sulfur using the formula:

$$\text{Ammonium sulfate} = 4.125[\text{S}] \quad (4)$$

Ammonium sulfate originates from the conversion of sulfur dioxide gas (SO_2), typically from coal burning, oil refining, industry and motor vehicles, to sulfuric acid (H_2SO_4) in the presence of water, which forms secondary particles when neutralised by ammonia present in the atmosphere. Note Equation (4) assumes that the sulfate is fully neutralised and does not occur as ammonium bisulfate or sulfuric acid which may occur in reduced ammonia atmospheres.

Organics

Particulate organic compounds generally contain carbon (C), hydrogen (H) and oxygen (O). In our case, it is estimated mainly through measurements of the hydrogen content by subtracting the hydrogen associated with ammonium compounds using the formula:

$$\text{Organics} = 11([\text{H}] - 0.25[\text{S}]) \quad (5)$$

This assumes the average organic compound in the fine fraction is 9% hydrogen, 20% oxygen and 71% carbon as per Malm et al. 1994. Note this will overestimate the aerosol organic content if large amounts of ammonium nitrate are present.

Soil

The fine soil component is estimated from the summation of the main oxides commonly found in soil, such as silicon oxide (SiO_2), aluminium oxide (Al_2O_3), iron oxides (FeO, Fe_2O_3), calcium oxide (CaO), and titanium oxide (TiO_2) using the formula:

$$[\text{Soil}] = 2.2[\text{Al}] + 2.49[\text{Si}] + 1.63[\text{Ca}] + 1.94[\text{Ti}] + 2.42[\text{Fe}] \quad (6)$$

where the brackets [ ] represent the concentration of that pseudo-element in ngm^{-3}. Fine soil in the atmosphere occurs typically from natural windblown dust, agriculture and industries such as quarrying.

Smoke estimate (K_{non})

Fine potassium is a key indicator of biomass burning including the burning of wood in domestic heaters and bushfires and controlled burning of bush for bushfire mitigation. Potassium also occurs as an oxide in fine soil so following Malm et al. (1994) we define non-soil fine potassium as:

$$[K_{non}] = ([K] - 0.6*[Fe]) \quad (7)$$

K_{non} can then be used as an estimator of smoke from biomass burning.

Black carbon

Black carbon (BC) is estimated by the widely used laser integrating plate method (LIPM) as described by Taha et al. 2007. The mass concentration of BC is calculated using the formula:

$$BC = \frac{100}{\varepsilon}\left(\frac{A}{V}\right) ln\left(\frac{I_0}{I}\right) \quad (8)$$

where I_0 is the unexposed filter laser transmission intensity, I is the exposed filter laser transmission intensity, V is the volume of sampled air (m^3), A is the area of exposed filter (cm^2), and ε is the mass absorption coefficient for fine BC particles. In our case, we assume a mass attenuation coefficient of $7m^2g^{-1}$ from Taha et al. 2007, which is appropriate for fine carbon particles.

Reconstructed mass

The reconstructed mass (RCM) is determined by the summation of the pseudo-elements ammonium sulfate, organics, soil, salt and BC as described by the formula:

$$RCM = salt + ammonium\ sulfate + soil + organics + BC \quad (9)$$

The RCM value is used to assess mass closure. Typically, we obtain a reconstructed mass of between 70% and 80% of the gravimetric mass. This represents good mass closure as our analysis techniques do not measure nitrates or water vapour, which would account for the remaining 20–30%.

7. Average chemical composition

Although the average chemical composition of the $PM_{2.5}$ size fraction at the four sampling sites is not the focus of the current analyses we include it here for completeness and insight into what sources one might expect are contributing to the total $PM_{2.5}$ mass. The tables below provide the average, median, standard deviation (SD) and maximum of the daily measured values of the common chemical species at each of the sites for fine ambient air over the whole study period of 2000–14. We included the parameter K_{non} defined above in Equation (7) as an estimator of biomass burning. Values of K_{non} between zero and $100ngm^{-3}$ probably represent fine potassium values from motor vehicles including diesel trucks and buses, while higher K_{non} values between 100 and $300ngm^{-3}$ are most likely associated with domestic wood heaters in the winter months. Values of K_{non} above $300ngm^{-3}$ are more likely to be associated with bushfire events and controlled burning episodes. Negative values of K_{non} represent fine potassium events not related to biomass burning and are most often related to dust storm

events. All values in the tables are in μgm^{-3} except for K_{non} which is in ngm^{-3}. The standard deviation (SD) in the tables is calculated on the average and represents the seasonal (summer to winter) variations seen at all sites, not the deviation on the measurement areas.

Lucas Heights

At Lucas Heights between 1 January and 31 December 2014 there were 1474 sampling days with *Error Code* =0. The average fine mass was (5.43±4.4)μgm^{-3} with a median value of 4.58μgm^{-3} and a 24-hour maximum value of 70μgm^{-3}. The average $PM_{2.5}$ chemical composition over the 15-year study period was 16% salt, 28% ammonium sulfate, 6% soil, 16% organics and 12% BC. The remaining percentage mass was made up of water vapour and nitrates which we do not measure on our Teflon filters.

Table 7-1. The average, median, standard deviation (SD) and maximum of the daily measured values of the common chemical species at Lucas Heights, 2000–14.

Lucas Heights	Species	Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
1474 days	Mass	5.43	4.58	4.4	70
$PM_{2.5}$	RCM	4.28	3.62	3.2	42
	RCM%	80.1	79.3	13.7	146
	Salt	0.86	0.64	0.9	8
	Ammonium sulfate	1.53	1.22	1.2	10
	Soil	0.32	0.19	0.6	12
	Organics	0.88	0.38	2.2	38
	Black carbon	0.65	0.55	0.4	3
	K_{non} (ngm^{-3})	21.5	13.8	35.7	499

The average value of K_{non} was only 22ngm^{-3} however **Figure 7-1** shows that at this site K_{non} varied from –100 to 500ngm^{-3} during the sampling period 2000–14. The negative K_{non} values were associated with dust storm events.

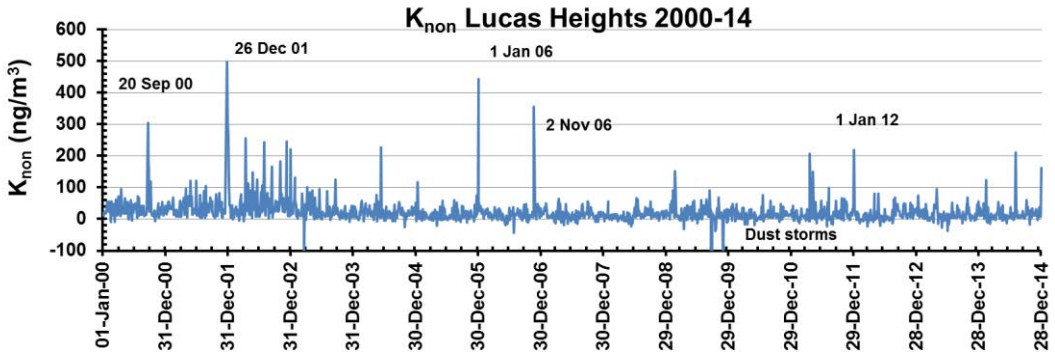

Figure 7-1. Time series plot of the smoke estimator, K_{non}, 2000–14.

There is very little summer–winter seasonal variations in K_{non} for Lucas Heights so we concluded that it is not significantly impacted by the burning of wood from domestic wood heaters. K_{non} values above 100ngm^{-3} are associated with bushfire events.

Richmond

At Richmond between 1 January and 31 December 2014 there were 1323 sampling days with *Error Code* =0. The average fine mass was (7.10±5.9)µgm^{-3} with a median of 5.58µgm^{-3} and a 24-hour maximum value of 61µgm^{-3}. The average $PM_{2.5}$ chemical composition over the 15-year study period was 8% salt, 22% ammonium sulfate, 5% soil, 29% organics and 12% BC. The remaining percentage mass was made up of water vapour and nitrates which we do not measure on our Teflon filters.

Table 7-2. The average, median, standard deviation (SD) and maximum of the daily measured values of the common chemical species at Richmond, 2000–14.

Richmond	Species	Average (µgm^{-3})	Median (µgm^{-3})	SD (µgm^{-3})	Max (µgm^{-3})
1323 days	Mass	7.10	5.58	5.9	61
$PM_{2.5}$	RCM	5.51	4.42	4.3	43
	RCM%	80.1	78.4	13.5	146
	Salt	0.56	0.34	0.7	5
	Ammonium sulfate	1.59	1.20	1.3	11
	Soil	0.33	0.21	0.5	9
	Organics	2.08	0.99	3.3	34
	Black carbon	0.86	0.72	0.5	4
	K_{non} (ngm^{-3})	47.9	25.2	65.3	555

The average value of K_{non} was only 48ngm^{-3} however **Figure 7-2** shows that at this site K_{non} varied from –100 to 600ngm^{-3} during the sampling period 2000–14. The negative K_{non} values were associated with dust storm events.

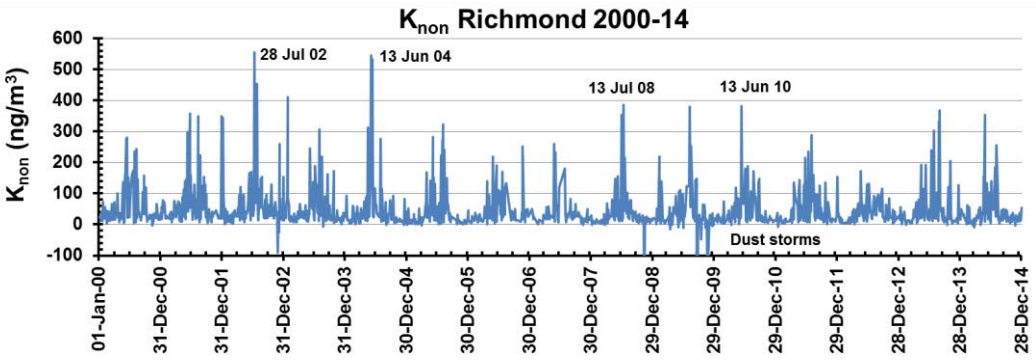

Figure 7-2. Time series plot of the smoke estimator, K_{non}, 2000–14.

There is a well-defined summer–winter seasonal variation for K_{non} between 0ngm^{-3} and 200ngm^{-3} for Richmond showing this site is significantly impacted by the burning of wood for domestic heating. K_{non} values above 200ngm^{-3} are associated with bushfire and controlled burning events.

Mascot

At Mascot between 1 January and 31 December 2014 there were 1469 sampling days with *Error Code* =0. The average fine mass was (8.21±9.5)µgm^{-3} with a median value of 6.59µgm^{-3} and a 24-hour maximum value of 311µgm^{-3}. The average $PM_{2.5}$ chemical composition over the 15-year study period was 14% salt, 20% ammonium sulfate, 7% soil, 23% organics and 18%

BC. The remaining percentage mass was made up of water vapour and nitrates which we do not measure on our Teflon filters.

Table 7-3. The average, median, standard deviation (SD) and maximum of the daily measured values of the common chemical species at Mascot, 2000–14.

Mascot	Species	Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
1469 days	Mass	8.21	6.59	9.5	311
$PM_{2.5}$	RCM	6.68	5.41	7.3	236
	RCM%	82.6	82.2	12.1	149
	Salt	1.16	0.88	1.1	9
	Ammonium sulfate	1.63	1.32	1.1	8
	Soil	0.55	0.32	2.3	84
	Organics	1.87	0.91	5.0	164
	Black carbon	1.48	1.13	1.1	8
	K_{non} (ngm^{-3})	5.82	9.24	129	458

The average value of K_{non} was only $6ngm^{-3}$ however **Figure 7-3** shows that at this site K_{non} varied from $-100ngm^{-3}$ to $500ngm^{-3}$ during the sampling period 2000–14. The negative K_{non} values were associated with dust storm events.

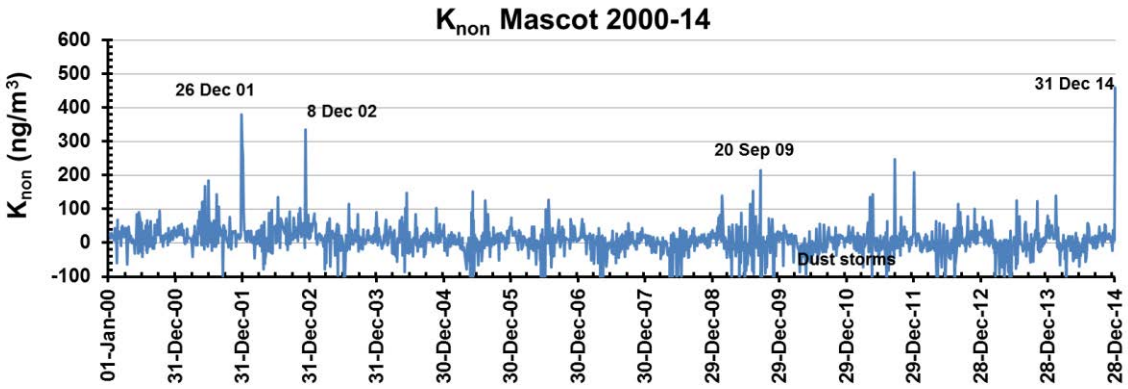

Figure 7-3. Time series plot of the smoke estimator, K_{non}, 2000–14.

As with the Lucas Heights site there was little seasonal variation in K_{non} at this site, with K_{non} mostly between $0ngm^{-3}$ and $50ngm^{-3}$. K_{non} values above $150ngm^{-3}$ are associated with bushfire and controlled burning events.

Liverpool

At Liverpool between 1 January and 31 December 2014 there were 1510 sampling days with *Error Code* =0. The average fine mass was $(8.76\pm9.9)\mu gm^{-3}$ with a median of $6.94\mu gm^{-3}$ and a 24-hour maximum value of $308\mu gm^{-3}$. The average $PM_{2.5}$ chemical composition over the 15-year study period was 9% salt, 19% ammonium sulfate, 7% soil, 27% organics and 17% BC. The remaining percentage mass was made up of water vapour and nitrates which we do not measure on our Teflon filters.

Table 7-4. The average, median, standard deviation (SD) and maximum of the daily measured values of the common chemical species at Liverpool, 2000–14.

Liverpool	Species	Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
1510 days	Mass	8.76	6.94	9.9	308
$PM_{2.5}$	RCM	6.98	5.74	5.9	150
	RCM%	82.0	81.3	10.8	133
	Salt	0.83	0.59	0.9	6
	Ammonium sulfate	1.70	1.36	1.3	11
	Soil	0.57	0.33	2.3	83
	Organics	2.32	1.18	3.7	77
	Black carbon	1.52	1.21	1.1	9
	K_{non} (ngm^{-3})	29.6	18.2	130.5	980

The average value of K_{non} was only $30ngm^{-3}$ however **Figure 7-4** shows that at this site K_{non} varied from –100 to $600ngm^{-3}$ during the sampling period 2000–14. The negative K_{non} values were associated with dust storm events.

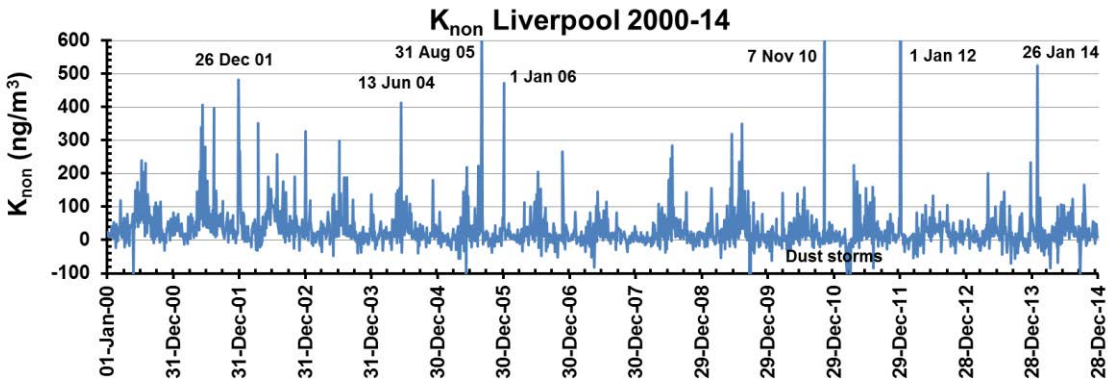

Figure 7-4. Time series plot of the smoke estimator, K_{non}, 2000–14.

There is a well-defined summer–winter seasonal variation for K_{non} between $0ngm^{-3}$ and $200ngm^{-3}$ for Liverpool, showing this site is significantly impacted by the burning of wood for domestic heating. K_{non} values above $200ngm^{-3}$ are associated with bushfire and controlled burning events.

All sites summary

The average $PM_{2.5}$ chemical composition across all four sites does not vary a lot. The sampling sites are different, some being rural and others heavily urbanised so the total fine mass varies but the chemical composition is surprisingly similar. This is not too unexpected as all four sites are generally within a common airshed of the major Sydney metropolitan region and so fine $PM_{2.5}$ ambient air composition is expected to vary little.

The average chemical composition across all four sites for the sampling period 2000–14 is shown in **Figure 7-5**. It shows that the average fine mass of $7.392\mu gm^{-3}$ is 12% sea salt, 22% ammonium sulfate, 6% soil, 24% organics and 15% BC. The reconstructed mass (RCM) =79%; the 21% missing mass was probably water (8–10%) and nitrates (11–13%) not currently measured by us.

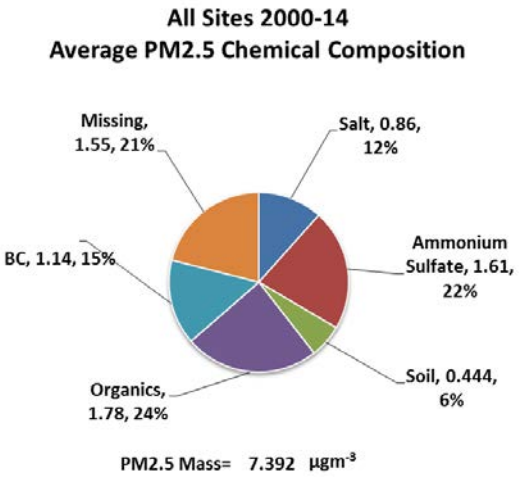

Figure 7-5. Average chemical composition for all four sites, 2000–14.

Figure 7-5 shows the average soil content across all four sites was 6%, Equation (6) shows this estimate was composed of the five oxides of Al, Si, Ca, Ti and Fe. If soil, as defined, was the primary source of these oxides then they should be highly correlated.

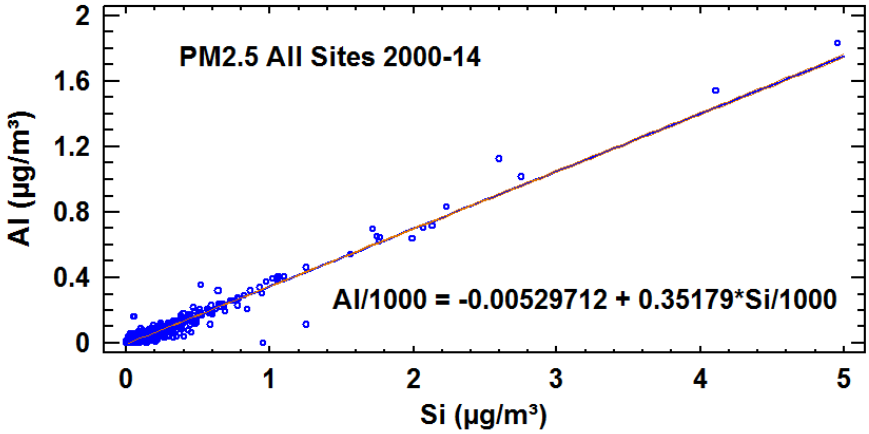

Figure 7-6. Plot of Al versus Si concentrations for all four sites, 2000–14.

Figure 7-6 is a plot of Al versus Si for each of the 5776 sampling days in the study period between 2000 and 2014. These two elements are highly correlated with the ratio (Al/Si) =0.35 showing that they do indeed originate from the one source type. Correlation plots of this type are typical for elements that originate from the same source.

The pie chart of **Figure 7-5** shows the average ammonium sulfate content of the $PM_{2.5}$ mass was 22%. Ammonium sulfate estimates were defined in Equation (4) above. As the ammonium sulfate is $(NH_4)_2(SO_4)$ we would expect the ratio of (H/S) =0.25.

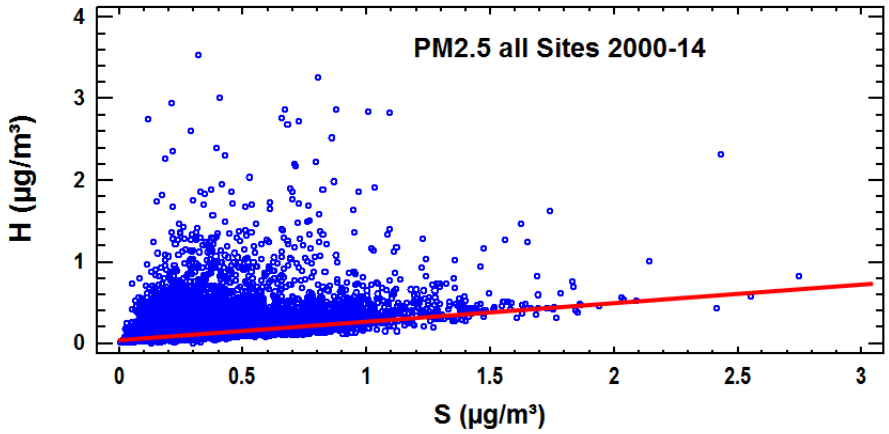

Figure 7-7. Plot of H versus S concentrations for all four sites, 2000–14.

Figure 7-7 is a correlation plot for the total H versus total S for all four sites during the study period 2000–14. The line drawn beneath the points is the (H/S) =0.25 line. It shows that ammonium sulfate is present for many of these points but that there are also many point above the line which correspond to excess hydrogen (H) not associated with sulfur. This would be hydrogen in organics as defined in Equation (5) for example. Note also that there are very few points below the (H/S) line in this plot; this demonstrates that fine sulfate (SO_4^{2-}) particles on the filters are fully neutralised and do not occur as bisulfate (NH_4HSO_4) or sulfuric acid (H_2SO_4) particles.

Scatter plots similar to that shown in **Figure 7-7** are typical for elements that have multiple sources. In NSW sulfur is typically sourced from coal combustion for power generation where 25 million tons of coal with a sulfur content of around 0.5% is burnt annually, with minor contributions from fossil fuel combustion in motor vehicles, from oil refining and from biomass burning of vegetation. It is also present in small amounts in sea spray where the ratio (S/Na) =0.083.

8. Daily mass exceedances

As the daily gravimetric mass was measured at each sampling site on each Sunday and Wednesday during the study period from 2000–14 the number of 24-hour exceedances can be obtained. An exceedance is based on the National Environmental Protection (Ambient Air Quality) Measure (Air NEPM) value of 25µgm^{-3} in a 24-hour sampling period. As sampling only occurred twice a week on Wednesdays and Sundays only 104 filters were collected each year.

The daily mass exceedances for the four sites for the study period from 2000–14 are summarised in **Table 8-1** to **Table 8-4**. Each table lists the date, the measured daily $PM_{2.5}$ mass for that day together with the measured soil, organics, ammonium sulfate, BC and the non-soil potassium (K_{non}) concentration which is a biomass smoke indicator. This helps identify the possible reason for the exceedance in many of the cases. Here we have tried to identify the known events due to windblown soil or dust and smoke from biomass burning and added this in the comments column of the tables. Dust from windblown soil was identified from the daily log sheet comments, and for days where soil >3µgm^{-3} and greater than 10% of the fine mass. Smoke and organic exceedances were also taken from the daily log sheet comments, and events where K_{non} >0.3µgm^{-3} and organics >25µgm^{-3} and greater than 50% of the total fine mass respectively.

Lucas Heights

Between 1 January 2000 and 31 December 2014 there were 11 measured $PM_{2.5}$ 24-hour mass measurements exceeding 25µgm^{-3} at this site. All these exceedances are listed in **Table 8-1**.

Table 8-1. A summary of the number of Air NEPM maximum daily mass exceedances for the Lucas Heights site, 2000–14.

Site	Date	Mass (µgm^{-3})	Soil (µgm^{-3})	Organics (µgm^{-3})	Sulfate (µgm^{-3})	Salt (µgm^{-3})	BC (µgm^{-3})	K_{non} (µgm^{-3})	Comment
Lucas Heights $PM_{2.5}$	20-Sep-00	41.9	0.18	29.9	0.49	0.00	3.4	0.30	Smoke
	26-Dec-01	54.6	1.71	31.8	0.88	0.00	3.2	0.50	Smoke
	30-Dec-01	53.2	0.28	29.7	2.76	0.00	1.6	0.31	Smoke
	02-Jan-02	42.8	0.74	25.3	0.89	0.00	1.9	0.26	Smoke
	15-Sep-02	36.8	0.51	17.8	3.99	0.03	1.8	0.17	
	03-Nov-02	28.9	2.44	9.5	3.26	2.10	1.6	0.18	
	08-Dec-02	70.0	0.19	37.9	1.33	0.00	1.8	0.24	Smoke?
	26-Jan-03	35.4	0.64	14.4	3.18	1.81	2.2	0.13	
	22-Nov-06	40.3	4.47	16.4	3.16	2.80	2.4	0.35	Smoke/ Dust
	23-Sep-09	235.1	76.13	22.4	0.20	0.52	0.0	–3.48	Dust
	03-Aug-14	26.6	0.12	13.2	0.58	0.81	1.4	0.21	

Richmond

Between 1 January 2000 and 31 December 2014 there were 28 measured $PM_{2.5}$ 24-hour mass measurements exceeding 25µgm^{-3} at this site. All these exceedances are listed in **Table 8-2**.

Table 8-2. A summary of the number of Air NEPM maximum daily mass exceedances for the Richmond site, 2000–14.

Site	Date	Mass (µgm^{-3})	Soil (µgm^{-3})	Organics (µgm^{-3})	Sulfate (µgm^{-3})	Salt (µgm^{-3})	BC (µgm^{-3})	K_{non} (µgm^{-3})	Comment
Richmond $PM_{2.5}$	24-Jun-01	36.4	0.26	17.8	0.89	0.00	2.6	0.36	Smoke
	12-Aug-01	35.0	1.30	17.0	2.11	0.00	2.3	0.35	Smoke
	30-Dec-01	61.5	0.62	33.6	3.32	0.00	2.1	0.35	Smoke
	06-Jan-02	50.3	0.71	18.7	10.03	3.54	2.1	0.34	Smoke
	14-Jul-02	42.6	2.06	28.0	2.99	0.86	3.4	0.55	Smoke
	28-Jul-02	35.6	1.18	19.7	1.53	0.00	2.9	0.45	Smoke
	23-Oct-02	55.9	6.27	28.0	4.52	1.26	1.3	0.13	Dust
	08-Dec-02	29.9	0.34	18.7	0.64	0.00	2.1	0.26	
	26-Jan-03	59.7	3.09	28.5	4.16	2.27	3.9	0.41	Smoke/ Dust
	29-Jan-03	25.5	0.66	2.3	4.25	2.17	0.8	0.08	
	03-Aug-03	28.0	2.09	13.8	1.22	0.00	2.8	0.31	Smoke
	23-May-04	36.2	2.33	20.3	1.72	0.00	2.7	0.23	
	02-Jun-04	25.9	1.93	15.0	1.08	0.00	3.7	0.21	
	06-Jun-04	40.3	0.98	24.1	1.77	1.47	3.2	0.55	Smoke
	01-Aug-04	30.0	0.86	12.3	3.94	0.00	2.4	0.28	
	01-May-05	30.2	0.52	19.5	0.72	0.00	1.6	0.17	
	05-Jun-05	26.0	1.31	11.4	1.69	0.55	2.2	0.28	
	22-Nov-06	30.2	2.66	11.8	2.82	2.84	1.5	0.25	
	13-Jul-08	33.2	0.85	16.2	1.56	0.00	3.1	0.39	Smoke
	23-Sep-09	334	91.2	483	0.24	0.53	0.0	–5.46	Dust
	10-Jun-12	25.4	1.18	12.2	1.57	0.24	2.4	0.06	

Site	Date	Mass (μgm^{-3})	Soil (μgm^{-3})	Organics (μgm^{-3})	Sulfate (μgm^{-3})	Salt (μgm^{-3})	BC (μgm^{-3})	K_{non} (μgm^{-3})	Comment
	28-Jul-13	26.5	0.69	12.3	1.48	0.92	2.6	0.30	
	25-Aug-13	34.1	0.70	19.5	1.36	1.00	2.8	0.33	
	01-Sep-13	34.3	0.95	19.2	1.44	1.31	2.0	0.37	
	27-Oct-13	25.8	0.38	13.7	1.64	1.06	1.4	0.12	
	03-Nov-13	38.8	3.67	17.6	1.90	2.91	2.4	0.20	
	25-May-14	29.8	1.65	15.5	1.01	0.51	2.8	0.35	
	06-Aug-14	26.0	0.57	14.4	0.89	0.40	1.9	0.19	

Mascot

Between 1 January 2000 and 31 December 2014 there were 23 measured $PM_{2.5}$ 24-hour mass measurements exceeding 25μgm^{-3} at this site. All these exceedances are listed in **Table 8-3**.

Table 8-3. A summary of the number of Air NEPM maximum daily mass exceedances for the Mascot site, 2000–14.

Site	Date	Mass (μgm^{-3})	Soil (μgm^{-3})	Organics (μgm^{-3})	Sulfate (μgm^{-3})	Salt (μgm^{-3})	BC (μgm^{-3})	K_{non} (μgm^{-3})	Comment
Mascot $PM_{2.5}$	25-Jun-00	29.0	1.49	13.6	2.71	0.00	5.5	0.06	BC
	02-Aug-00	31.2	1.47	10.7	5.07	0.00	6.5	–0.03	BC
	06-Jun-01	36.5	2.19	9.1	6.80	0.00	8.4	–0.02	BC
	01-Jul-01	38.3	1.28	19.4	3.58	0.00	6.5	0.18	BC
	12-Aug-01	26.5	1.16	9.5	3.36	1.51	2.9	0.14	
	26-Dec-01	39.4	1.53	24.4	0.77	0.56	2.8	0.38	Organics/ Smoke
	30-Dec-01	54.2	0.43	29.1	3.62	0.00	2.5	0.30	Organics/ Smoke
	02-Jan-02	50.8	2.08	28.6	2.71	0.00	2.7	0.25	Organics/ Smoke
	08-Dec-02	52.0	0.92	25.3	3.56	2.89	3.1	0.33	Organics/ Smoke
	01-Aug-04	32.9	1.16	11.6	4.48	0.40	4.4	0.08	
	05-Jun-05	27.0	1.73	11.1	3.07	1.60	3.5	0.15	
	23-Jul-06	28.3	1.90	10.7	2.26	2.80	4.3	0.13	
	25-May-08	26.2	1.16	8.8	2.95	0.58	2.7	–0.01	
	14-Jun-09	31.5	1.66	14.2	1.02	0.00	4.8	–0.02	
	26-Jul-09	29.1	1.39	12.9	1.96	0.00	4.9	0.11	
	09-Aug-09	32.1	2.13	12.1	3.01	2.40	3.7	0.15	
	20-Sep-09	42.4	1.37	22.2	2.92	0.00	3.2	0.22	
	23-Sep-09	311	84.2	164	0.32	0.70	0.0	–4.60	Dust
	22-Nov-09	25.3	8.10	4.4	5.30	1.70	1.2	0.07	Dust
	22-May-11	37.1	1.90	18.9	2.85	0.69	3.5	0.14	Organics
	26-Jun-11	25.2	3.13	10.9	2.93	0.00	3.3	–0.06	Dust
	14-Jul-13	31.5	1.89	13.8	1.99	1.42	4.0	0.13	
	03-Nov-13	30.0	1.73	12.0	2.50	4.01	3.1	0.12	

Liverpool

Between 1 January 2000 and 31 December 2014 there were 43 measured $PM_{2.5}$ 24-hour mass measurements exceeding 25μgm^{-3} at this site. All these exceedances are listed in **Table 8-4**.

Table 8-4. A summary of the number of Air NEPM maximum daily mass exceedances for the Liverpool site, 2000–14.

Site	Date	Mass (μgm^{-3})	Soil (μgm^{-3})	Organics (μgm^{-3})	Sulfate (μgm^{-3})	Salt (μgm^{-3})	BC (μgm^{-3})	K_{non} (μgm^{-3})	Comment
Liverpool $PM_{2.5}$	24-May-00	31.6	4.90	11.6	2.25	0.00	9.0	–0.10	Dust
	09-Jul-00	30.8	0.86	15.2	1.77	0.00	4.8	0.24	
	23-Jul-00	32.0	1.60	14.6	2.28	2.04	5.5	0.20	
	30-Jul-00	32.3	1.18	16.8	3.00	0.00	5.6	0.23	
	18-Apr-01	25.0	1.53	5.2	8.84	0.72	3.5	0.02	
	03-Jun-01	42.9	1.03	18.2	4.27	0.00	5.8	0.34	Smoke
	06-Jun-01	41.3	2.16	13.0	7.18	0.00	8.0	0.09	
	24-Jun-01	43.5	1.38	17.2	2.52	0.00	7.0	0.28	
	08-Jul-01	27.9	0.42	13.3	1.26	0.00	3.7	0.18	
	12-Aug-01	34.1	1.15	12.3	4.51	0.00	3.4	0.40	Smoke
	26-Dec-01	62.8	1.90	32.0	1.68	0.00	2.8	0.48	Smoke
	30-Dec-01	51.3	0.46	27.6	2.81	0.00	2.1	0.27	
	02-Jan-02	48.7	1.97	25.2	1.63	0.00	2.9	0.26	
	14-Apr-02	42.2	1.76	22.0	2.95	0.00	5.2	0.35	Smoke
	24-Jul-02	25.4	1.56	12.2	1.46	0.73	7.2	0.06	
	28-Jul-02	25.7	0.75	13.8	2.12	0.66	3.8	0.26	
	15-Sep-02	30.9	0.72	15.2	3.33	0.32	2.3	0.14	
	30-Oct-02	28.1	2.92	9.7	4.25	0.62	4.4	–0.03	
	03-Nov-02	31.6	3.28	11.4	3.34	2.97	2.4	0.19	
	26-Jan-03	39.3	0.88	14.3	2.97	2.42	2.7	0.12	
	06-Jul-03	31.1	0.62	27.8	1.20	0.00	4.3	0.30	Smoke
	03-Aug-03	27.0	1.67	16.4	2.52	0.37	3.7	0.19	
	02-Jun-04	29.3	2.66	11.7	1.17	0.00	7.2	0.03	
	13-Jun-04	26.9	0.79	13.1	1.55	0.00	4.0	0.41	Smoke
	16-Jun-04	29.0	3.17	11.0	3.46	0.38	3.9	–0.04	
	08-Jun-05	36.4	3.66	11.6	6.70	0.47	6.1	–0.11	
	31-Aug-05	25.6	1.61	11.7	1.05	0.90	2.4	0.68	Smoke
	09-Jul-06	27.3	1.45	12.2	1.39	0.57	3.5	0.20	
	22-Nov-06	42.8	3.19	18.6	2.88	2.47	3.3	0.27	
	29-Jun-08	29.4	1.49	13.2	1.62	0.39	3.4	0.18	
	06-Jul-08	25.8	0.57	11.4	1.52	0.48	3.9	0.19	
	13-Jul-08	31.0	1.20	13.6	1.84	0.61	3.7	0.25	
	20-Jul-08	30.3	1.35	14.2	1.01	0.31	5.0	0.28	
	14-Jun-09	34.8	1.26	18.5	1.23	0.34	5.1	0.32	Smoke
	26-Jul-09	29.9	0.65	13.4	1.35	0.00	3.2	0.24	
	09-Aug-09	27.7	1.28	11.1	2.09	1.28	4.3	0.35	Smoke
	19-Aug-09	26.3	3.08	10.0	2.14	1.30	2.8	0.09	
	23-Sep-09	308	83.5	77.5	0.20	0.78	0.0	–4.29	Dust
	22-May-11	31.8	1.40	17.2	2.21	0.00	3.3	0.14	
	26-Jun-11	26.6	1.98	12.4	2.01	0.00	3.8	0.16	
	31-Jul-11	25.5	1.66	11.5	1.21	0.00	3.6	0.16	
	29-Aug-12	29.3	2.86	12.3	2.46	0.61	4.6	0.04	
	28-Apr-13	35.6	1.60	19.2	1.88	0.99	4.4	0.20	

9. Positive matrix factorisation

In **Figure 7-6** and **Figure 7-7** above we showed how two-dimensional (2D) elemental correlation plots can be used in a crude way to determine soil and sulfate sources. This method is simple but only works when chemical species are clearly associated with just one source. Once we get into the real life situations where one chemical species can have multiple sources we need to go to multi-dimensional statistical processes to unfold the inter-element correlations and to associate an element with a given factor or source term. Positive matrix factorisation (PMF) (Paatero & Tapper 1994; Paatero 2004) is such a multi-dimensional process.

Here we apply this PMF receptor modelling technique to ANSTO's fifteen year elemental dataset related to the Richmond, Mascot, Liverpool and Lucas Heights sampling sites. The PMF technique is a one step-process that statistically partitions the elemental data into correlated elements representing source fingerprints and their contributions to the total $PM_{2.5}$ mass at each site.

It should be noted that the original DOS version of the PMF analysis codes developed by Paatero (Paatero & Tapper 1994; Paatero 2004) was used in this work and not the modified US EPA PMF version 5 codes based on multi-linear engine (ME) available on the internet. We utilise the PMF-DOS version as it is a more flexible code designed for researchers and we have invested significant time and energy over the past 20 years into developing local software scripts for running this code, as well as standardising and plotting many of its output files. A recent collaborative study for the Upper Hunter Valley region of NSW (Hibberd et al. 2013) involving ANSTO and CSIRO showed good agreement in PMF solutions obtained independently using the two different PMF versions. This is important as the solutions derived from the PMF technique are not unique, but are produced by a least squares iterative process based on the following matrix equation:

$$X = F * G + E \qquad (10)$$

where $X(n,m)$ is a measurement matrix of n sampling dates (i.e. elemental database dates in rows) and m chemical species (i.e. elemental database in columns), $F(p,m)$ is a factor matrix of p source fingerprints each with m elements, $G(n, p)$ is a contribution matrix of the p source fingerprints for each of the n sampling dates, and $E(n,m)$ is an error matrix which is minimised during the PMF process. The PMF codes work statistically towards minimising the unexplained part of the $X(n,m)$ matrix. What differentiates PMF from other statistical source apportionment methods such as chemical mass balance (CMB) or principal components analysis (PCA) is that the G and F matrices are constrained to have positive values only. The error term matrix, $E(n,m)$ is reduced in the PMF process by minimising the quantity, Q where:

$$Q = \sum_i \sum_j \left(\frac{e_{ij}}{s_{ij}}\right)^2 \qquad (11)$$

and e_{ij} are the elements in the error matrix term E and s_{ij} are the estimated errors of the experimental measurements assessed by PMF itself. In our case, we used:

$$s_{ij} = Err_{ij} + MDL_{ij} \qquad (12)$$

where Err and MDL have been defined earlier in Equations (1) and (2) respectively.

Note that one of the powers of this PMF technique is that each *(i,j)* in the error matrix Err_{ij} and the MDL_{ij} matrix has its own unique error and MDL which are used in the PMF codes to weight the significance of that single measurement on that individual day.

Equation (10) can be rewritten as:

$$X = (F/c) * (cG) + E \quad (13)$$

where *c* is any constant. For our solutions presented below we use a constant *c* that normalises the fingerprint matrix *F* to having its maximum element (*m*) with a fraction equal to unity. This makes it easier to identify the nature of most fingerprints. For example, a fingerprint with Si=1.0 and Al=0.33 is probably a soil fingerprint and a fingerprint with S=1.0 and H=0.25 is probably a secondary sulfate fingerprint driven by ammonium sulfate. As the *F* matrix has been renormalised (by constant *c*) the *G* matrix has also been adjusted by the same constant *c* as shown in Equation (13); so to make the *G* matrix meaningful we perform a multi-linear least squares regression analysis on the *G* matrix contributions and the total $PM_{2.5}$ gravimetric mass, so the final source fingerprint masses on average add up to the total $PM_{2.5}$ gravimetric mass.

The optimal PMF solution

As the PMF solutions are not unique it is important to define a process that selects the optimal solutions for each site and time series used. This is done by performing the PMF analysis for multiple factors (p) in order to determine and compare the best statistical fit while maintaining a solution that meets the following six criteria:

i. All factors (p) are recognised as possible key source fingerprints for that site and can be given a meaningful name.
ii. All factors *F(p,m)* are driven by more than one element (*m*>1).
iii. All factors have positive mass contributions when their masses are multi-least squared fitted to the total gravimetric mass.
iv. The average factor mass contributions, after multi-linear regression, are generally greater than 1%.
v. All multi-linear regression *p*-values are less than 0.05.
vi. χ^2 is close to unity, where:

$$\chi^2 = Q / Q_{theory} \quad (14)$$

and

$$Q_{theory} = mn - p(m+n) \quad (15)$$

and *Q* is calculated by the PMF codes and is given by Equation (11) above.

Inspection of Equations (11), (14) and (15) shows that if $\chi^2 >> 1$ then either there are too few fingerprints being used in the fit (*p* is too small) or the experimental error estimates s_{ij} used are too small. For $\chi^2 << 1$ the reverse applies; that is, *p* is too large or the error estimates s_{ij} are too large. Generally experience has shown that if the six criteria mentioned are met then solutions with $0.7 < \chi^2 < 1.1$ are optimal.

Clearly the PMF analysis does not assign a source name to each resulting factor *F(p,m)*. It only provides the factors and their contributing elements, which are then given a name by the data analyst. This is done based on sampling site knowledge and experience in identifying the most

likely source associated with the elemental fingerprint for each factor. For example, a factor dominated by Al, Si, Ca, Ti and Fe in the elemental fingerprint is typically associated with windblown soil, or H and S or Na and Cl present in a factor in the correct ratios can be used to identify a secondary sulfate and sea spray factor, respectively.

Additional information and examples related to ANSTO's application of IBA techniques and statistical PMF source apportionment to air pollution studies in Australia and Asia can be found in a number of peer-reviewed publications (Chan et al. 2005, 2008a, 2008b, 2011; Cohen et al. 1996, 2000, 2004a, 2004b, 2005, 2010a, 2010b, 2011, 2012 and 2014).

PMF errors

The PMF codes of Paatero 1994 provide error estimates on each element within a given source fingerprint *F(p,m)*. When we plot these fingerprints we include the ±95% confidence interval error bars on the elemental fractions for each element. The errors on the *G(n,p)* matrix contributions are also estimated by the PMF codes and these are again included as ±95% confidence intervals in the tabulated values of the fingerprint contributions provided below.

10. Site fingerprints and fingerprint contributions

The PMF source apportionment codes were run with 7, 8 and 9 factors to start with to test the optimal solutions that would fit the six criteria defined above. At all four sites seven factors (p=7) were found to best fit these criteria. Solutions with eight or nine factors invariably had $\chi^2 << 1$, such as factors that were driven by only one element (m=1) or a combination of elements in fractional ratios that did not fit expectations, for example, incorrect (H/S) ratios in ammonium sulfate and poor (Al/Si) ratios in soil fingerprints. Consequently, seven factors or fingerprints were used for PMF fits to all four sites. This produced $0.78 < \chi^2 < 1.07$ for the site fits using the same 23 elements, namely, H, Na, Al, Si, P, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Se, Br, Pb, BC and TotN. Only data with *Error Code* =0 were used for each site fit.

It should be pointed out that typical environmental data as collected here over a 24-hour period for 15 years will not be normally distributed. Some days will have very high mass concentrations more than six standard deviations above the mean, particularly during severe dust storms or major bushfire events for example. Environmental data of this type are usually log-normally distributed. This means that distribution averages and standard deviations can be misleading if extreme outliers are included. The aim of these analyses was to determine average source fingerprints and their contributions for the average days during the study period and not to include the extreme outlier events, as these could unduly bias the fits and hence the results. In order to eliminate these extreme events from the data several approaches were taken. Firstly, before the PMF analysis was performed all daily events with reconstructed mass (RCM) concentrations below 40% and above 150% of the gravimetric mass were removed, as these were considered to be outliers. Secondly, after the first PMF iteration the fitted PMF mass (FitMass), obtained by summing the mass of each of the fingerprints, was plotted against the gravimetric mass for each day in the study period. Any points on this plot lying outside the tramlines defined by 4–6 standard deviations about the mean were rejected and another PMF iteration performed again using the seven factors fits. These PMF mass versus observed or gravimetric mass plots are reproduced with their tramlines for each of the sites in the sections below.

A further data reduction followed by another PMF iteration was also performed by removing outliers in the plots of the PMF concentrations for each element against the original IBA elemental concentrations. These graphs are reproduced for each site in Appendix A.

Lucas Heights

The 15-year dataset at Lucas Heights was optimally fitted with seven factors or fingerprints using 1383 sampling days with a χ^2 =0.84. The seven factors were, Soil, Sea, mixed secondary sulfate (Mixed-2ndryS), mixed industrial calcium and aged sulfate (Mixed-Ind-Ca-Saged), mixed smoke and automobiles (Mixed-Smoke-Auto) and two automobile sources (Auto1 and Auto2). Each of these is described in detail in the graphs and tables below. The first plots are the seven source fingerprints with the maximum driving element fraction normalised to unity in each plot and the second plots are the percentage contribution each element makes to each given fingerprint shown in the first plots. As indicated previously several of these source fingerprints are mixed sources being driven by several different emission sources. For example, the smoke fingerprint is primarily driven by fine potassium from biomass burning but also contains components from automobiles (Fe, Pb, Br); hence it is named Mixed-Smoke-Auto to reflect this.

For this site sulfur primarily occurs in three fingerprints: 80% in Mixed-2ndryS, 9.5% in Mixed-Smoke-Auto and 6.5% in Auto2; the remainder is essentially in Sea spray. In the Mixed-2ndryS fingerprint sulfur clearly occurs as ammonium sulfate as the (H/S) ratio is approximately correct at 5.9 and the presence of total nitrogen (TotN) confirms there is ammonium present. The presence of V and Ni also confirms sulfate from heavy oil combustion. There is a known small oil fired power station a few kilometres from the Lucas Heights sampling site. This secondary sulfate fingerprint also contains 15% of the measured BC and 12% of the Na, which may originate from the municipal landfill waste tip which is also adjacent to the sampling site. As there are clearly several contributing sources in this fingerprint we call it a mixed fingerprint.

The Mixed Ind-Ca-Saged fingerprint is driven by the H, Na, Si, S, Ca, Fe and BC; it contains little or no total nitrogen (TotN). It contains 59% of the total fine Ca, 49% of the Na with 5.9% BC and very little TotN suggesting there are very few ammonium or nitrate compounds present and the Na and Ca probably occur as oxides of carbonates. There is a cement works a few kilometres north of the sampling site and there were significant building and construction works around the site after 2011 which would account for this source.

The two Auto fingerprints together (Auto1 + Auto2) contain nearly 100% of the total Zn and 85% of the total Cu, 70% of the total Pb, 77% of the BC and 28% of the TotN between them, which points to motor vehicular emissions.

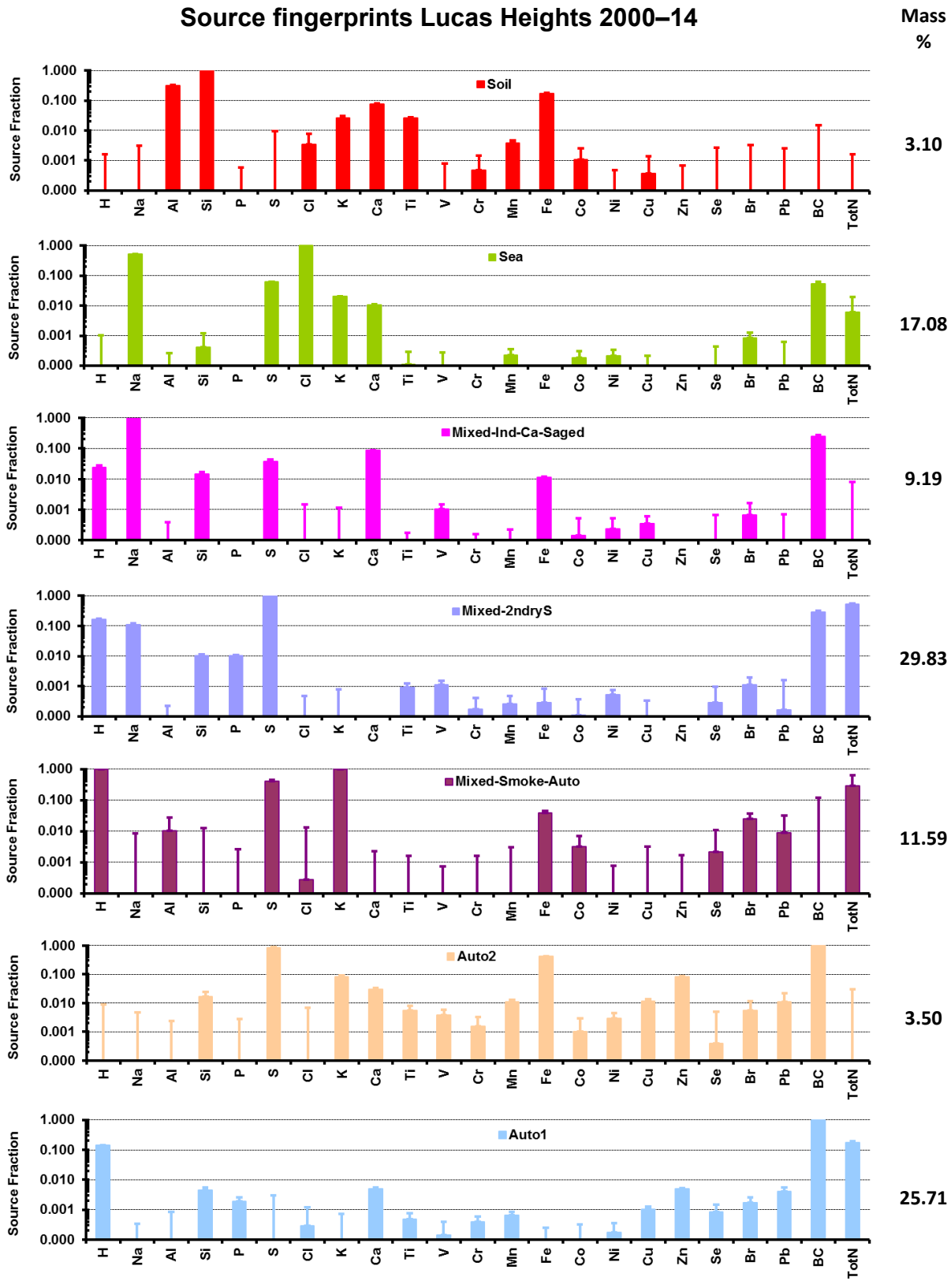

Figure 10-1. Plot of the PMF elemental fractions in the fingerprints for the Lucas Heights site normalised to unity for maximum element.

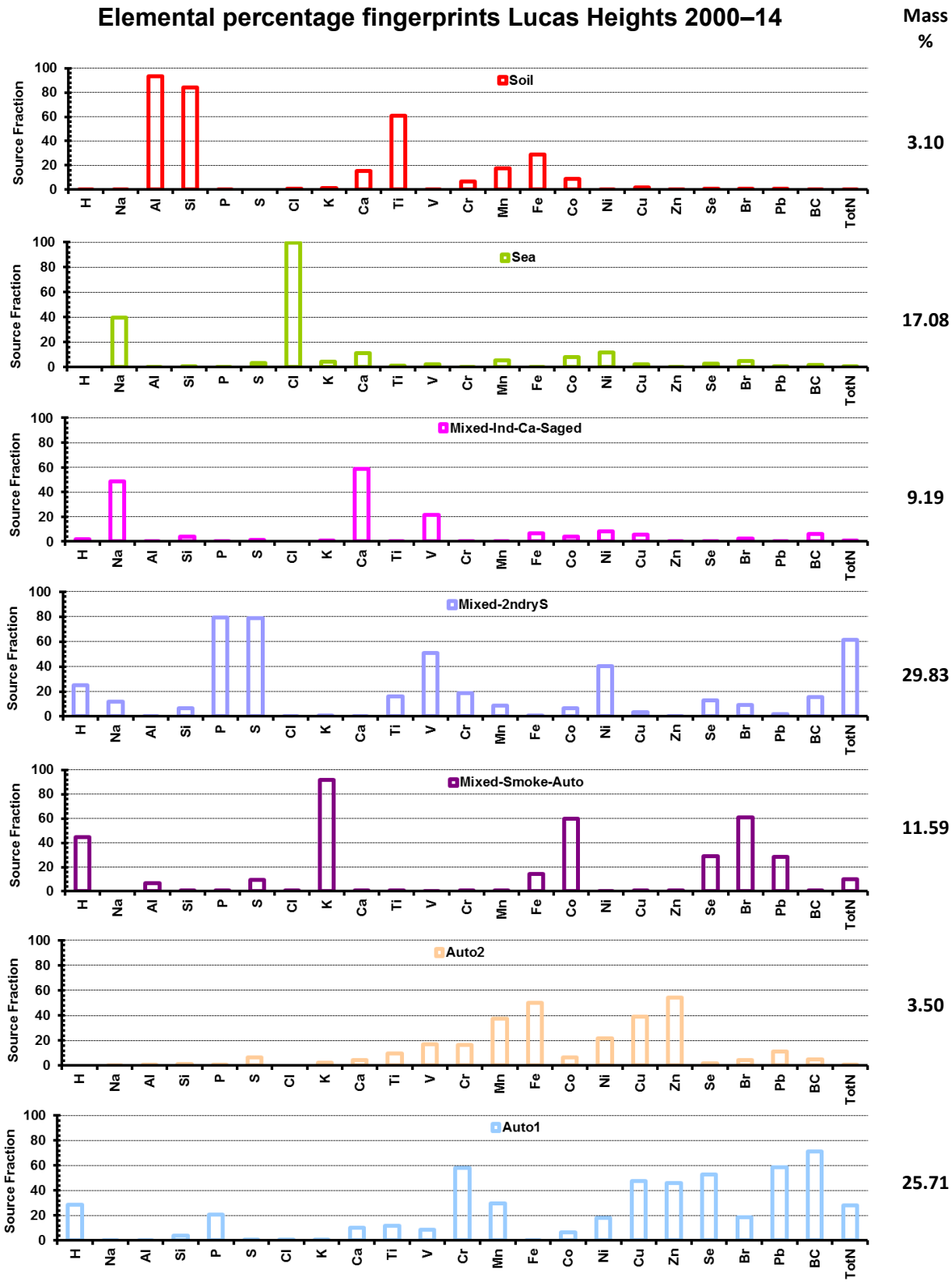

Figure 10-2. Plot of the percentage elemental concentrations in each fingerprint at Lucas Heights.

Figure 10-3 shows the plots of the sum of the calculated PMF fingerprint masses (FitMass) versus the observed or gravimetric $PM_{2.5}$ mass. Note the fit has a gradient of 0.984 which means that the PMF analysis has fitted the total $PM_{2.5}$ mass to better than 1.6% over the mass range 0–20µgm^{-3}; also, the correlation coefficient was 0.92 which is excellent.

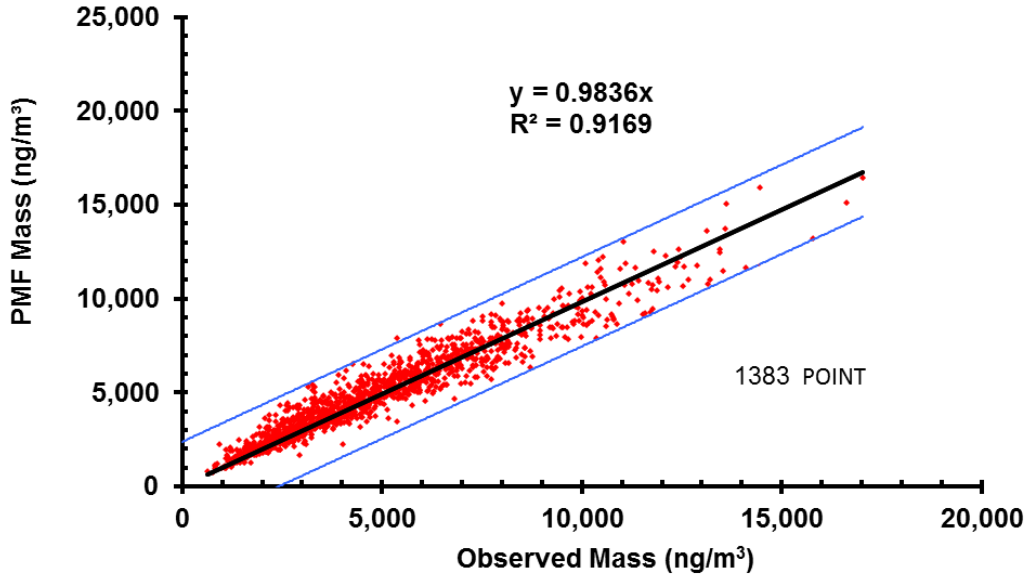

Figure 10-3. Lucas Heights PMF mass versus gravimetric mass, 2000–14. The dashed tramlines represent five standard deviations either side of the linear least squares fitted solid line.

The two tramlines either side of the line of best fit represent ±5 standard deviations from this line and it is obvious that most points outside these tramlines have been removed.

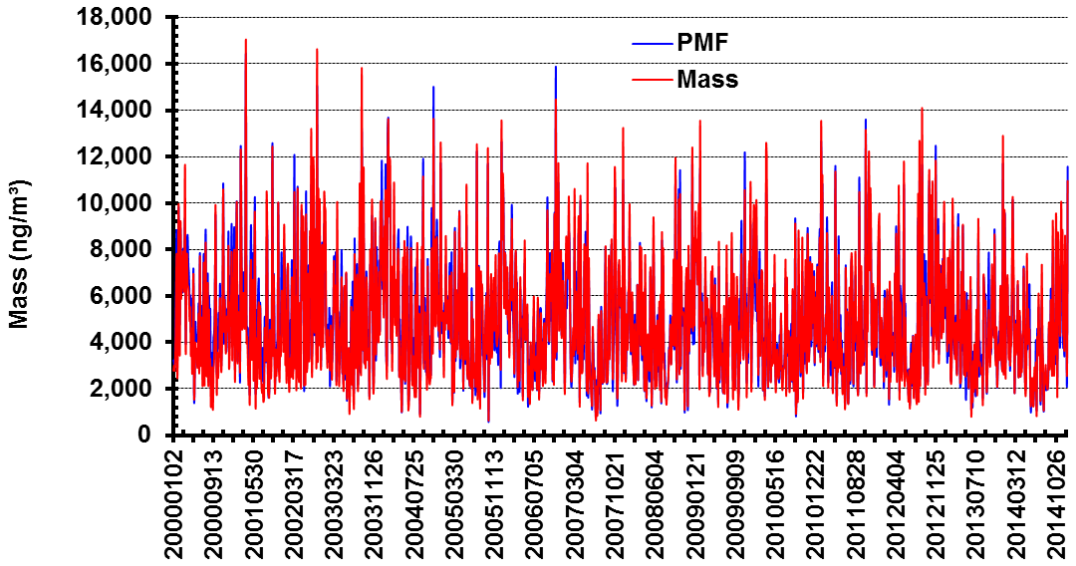

Figure 10-4. Lucas Heights time series PMF mass and gravimetric mass, 2000–14.

Figure 10-4 is a daily time series plot of the PMF mass and gravimetric mass of the data shown in **Figure 10-3** above. The x-axis is the date using the convention 20040725 to represent 25 July 2004. It again demonstrates the excellent fit of the PMF calculates to the measured data. All the main daily and seasonal variations over the 15-year study period are well reproduced.

Table 10-1 summarises the seven fingerprint descriptions for the Lucas Heights site. The PMF calculations use these seven factors or source fingerprints for the whole of the study period. They are unique to this site and under this analysis their elemental composition does not change with season, only their relative contributions to the total measured fine mass.

The table lists the fingerprint names, the key driver elements for that fingerprint which helped define the name, the key percentage elements contributing to that fingerprint, which elements in the fingerprint were poorly fitted by the PMF process and finally comments as to why this factor was so named. Elements that were poorly fitted by the PMF process were generally of low concentration (<20ngm^{-3}), had high MDLs and large errors associated with the measurement. Consequently their weighting was such that they did not drive the PMF fits and their contributions to the least squares fitting process were minimal.

Table 10-1. Summary of the seven fingerprint descriptions for the Lucas Heights site.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
Lucas Heights ASP1 $PM_{2.5}$	F1	Soil	3.10	0.84	1,383	Al, Si, K, Ca, Ti, Fe	Al, Si, Ti, Fe	P,V, Cr, Co, Ni, Se, Br, Pb, TotN	None	Clearly a soil factor with (Al/Si) ratio =0.32, contains 95%of the Al, 85% of the Si, 60% of the Ti and 30% of the Fe.
	F2	Sea	17.08			Na, S, Cl, Br, BC, TotN	Na, Cl		Summer high	A sea salt factor dominated by Na and Cl with (Cl/Na)=2.0 and (S/Na)=0.15, a little high but consistent with NaCl and sea salt. Contains 40% of the total Na, 100% of the total Cl, with traces of BC and TotN but no H.
	F3	Mixed-2ndryS	29.83			H, Na, S, BC, TotN	H, Na, P, S, V, Ni, TotN		Summer high	A (S/H) ratio =5.9 consistent with neutralised ammonium sulfate. Excess S probably associated with the 15% Na as sodium sulfate. Also contains 50% of the V and 40% of the Ni which indicates a heavy oil combustion component as well. Contains 80% of the total S, 25% of the total H and 60% of the TotN associated with ammonia.
	F4	Mixed-Ind-Ca-Saged	9.19			H, Na, Si, S, Ca, Fe, BC	Na, Ca, V		Summer high	Contains 49% of total Na and 59% of the total Ca could be a local cement factory and/ or local construction/ earthworks which the time series plot shows increased after 2011. Contains only 5.9% of the total BC and no TotN, suggesting it is low in nitrates and ammonium.
	F5	Mixed-Smoke-Auto	11.59			H, S, K, Fe, Br, Pb, TotN	H, K, Co, Se, Br, Pb		Winter high	Smoke can originate from diesel vehicles, bushfires and domestic wood heating. Wood heating tends to show seasonal trends being higher in the winter months. This factor is driven by H (organics) and fine potassium from biomass burning. Contains 45% of the total H and 95% of the total K, hence labelled smoke. Contains 60% of the total Co and Br and 10% of the total S and TotN, so it is probably ammonium sulfate.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
	F6	Auto1	25.71			H, Si, P, Ca, Zn, Pb, BC, TotN	Cr, Mn, Cu, Zn, Se, Pb, BC, TotN		Winter high	This factor is similar to F7 but with extra H (25%). This factor represents 26% of the fine mass and is consistent with automobile source emissions. Factors F6 and F7 contain nearly 100% of the total Zn, 85% of the total Cu and 70% of the total Pb between them, which points to vehicular emissions.
	F7	Auto2	3.50			Si, S, K, Ca, Fe, Zn, BC	Mn, Fe, Cu, Zn		Slight winter high	This is driven by S, Fe, Zn and BC with a range of heavy metals which can be associated with motor vehicles. Contains 50% of the total Fe and Zn. Fine Zn can be associated with vehicular tyre wear.

Richmond

The 15-year dataset at Richmond was optimally fitted with seven factors or fingerprints using 1248 sampling days with a χ^2 =0.78. The seven factors were, Soil, mixed industrial aged sulfate (Mixed-Ind-Saged), Sea, mixed secondary sulfate (Mixed-2ndryS), mixed smoke and automobiles (Mixed-Smoke-Auto) and two automobile sources (Auto1 and Auto2). Each of these is described in detail in the graphs and tables below. The first plots are the seven source fingerprints with the maximum driving element fraction normalised to unity in each plot and the second plots are the percentage contribution each element makes to each given fingerprint shown in the first plots. As indicated previously several of these source fingerprints are mixed sources being driven by several different emission sources. For example, the smoke fingerprint is primarily driven by fine potassium from biomass burning but also contains components from automobiles (Fe, Pb, Br); hence it is named Mixed-Smoke-Auto to reflect this.

The largest contributors to the total $PM_{2.5}$ mass at Richmond are the Mixed-Smoke-Auto (30%) the Mixed 2ndryS (26%) and the Mixed-Ind-Saged (15%) fingerprints. Sulfur appears as 80% in the Mixed-2ndryS and 20% in the Mixed-Ind-Saged fingerprints. The ratio of (H/S) =0.25 in the Mixed-2ndryS fingerprint, which is consistent with fully neutralised ammonium sulfate. It also contains 6% of the BC and 48% of the total nitrogen (TotN).

The Mixed-Smoke-Auto fingerprint is the largest single contributor (30%) to the total fine mass at this site and is driven by H, S, K, BC and TotN with some re-entrained soil (Al, Si, Ti) and traces of automobile components (S, Cu, Zn, Br and Pb). It contains 60% of the total H, 74% of the total K, 38% of the total Zn, 37% of the total BC and 52% of the total TotN, probably as ammonium nitrate; hence it is labelled smoke.

The Mixed-Ind-Saged fingerprint is driven by Na, S and BC with no Cl and no TotN, so it cannot be NaCl or $NaNO_3$; it is probably a mixed aged industrial sulfate source. The (Na/S) ratio =3.0 is higher than expected in sodium sulfate [(Na/S)=1.4] so the excess Na must be in another chemical form such as carbonate or oxide (not nitrate as TotN is small). It does not appear to be calcium sulfate either as we have excess Ca, as either oxide or carbonate. This is primarily an industrial fingerprint but it may also be a partial coal combustion fingerprint from power stations as it does contain 46% of the total Ca, 18% of the total S, 13% of the total selenium (Se) and 8% of the total BC. It also contains 31% of the total V and 20% of the total Ni, which is indicative of heavy oil combustion or processing; hence the mixed industrial aged sulfur naming of this fingerprint.

The two Auto fingerprints together represent 20% of the total fine mass and contain 61% of the total Zn and 89% of the total Cu, 88% of the total Cr, 33% of the total Pb and 50% of the BC between, them which points to motor vehicular emissions.

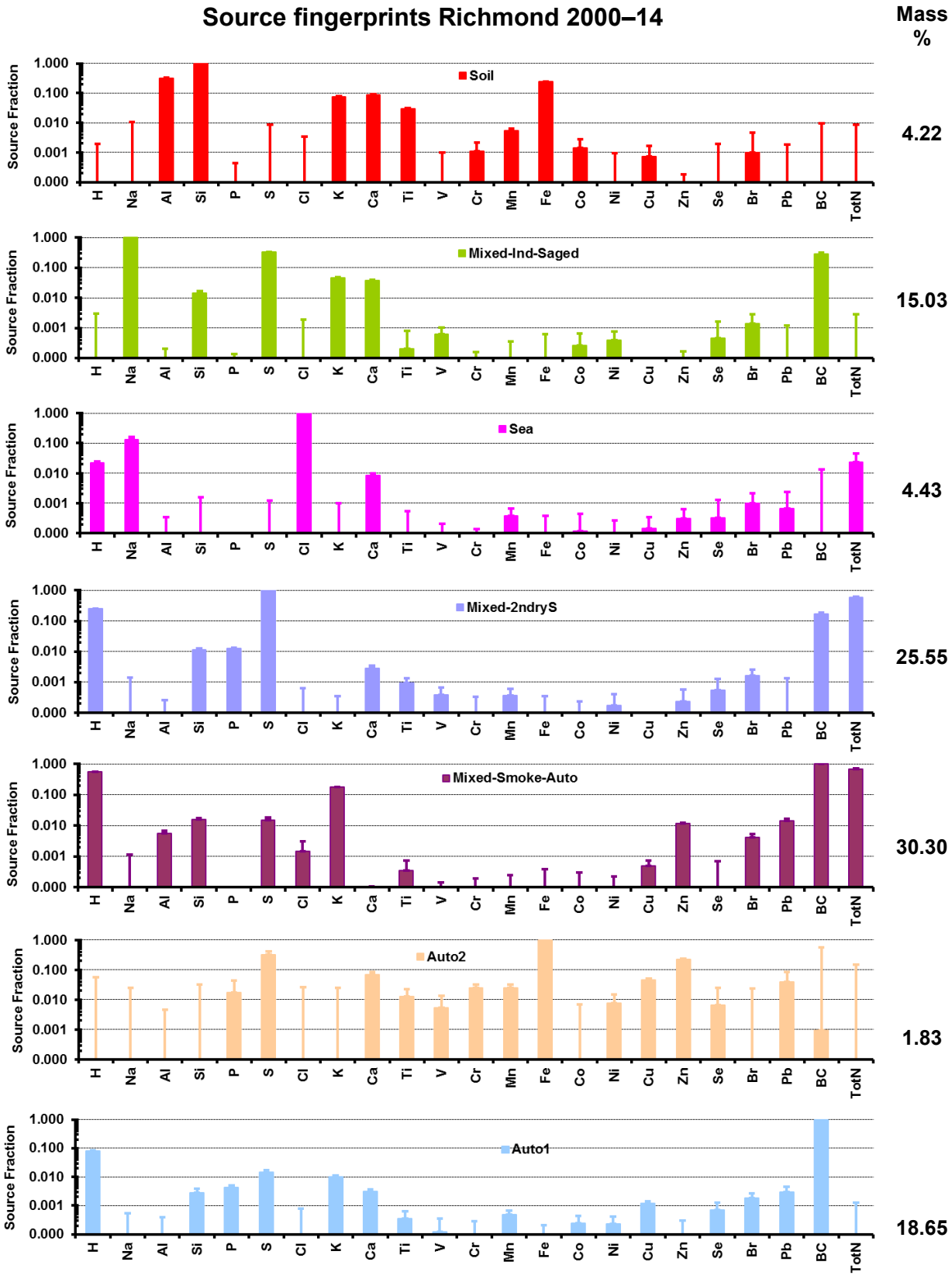

Figure 10-5. Plot of the PMF elemental fractions in the fingerprints for the Richmond site normalised to unity for maximum element.

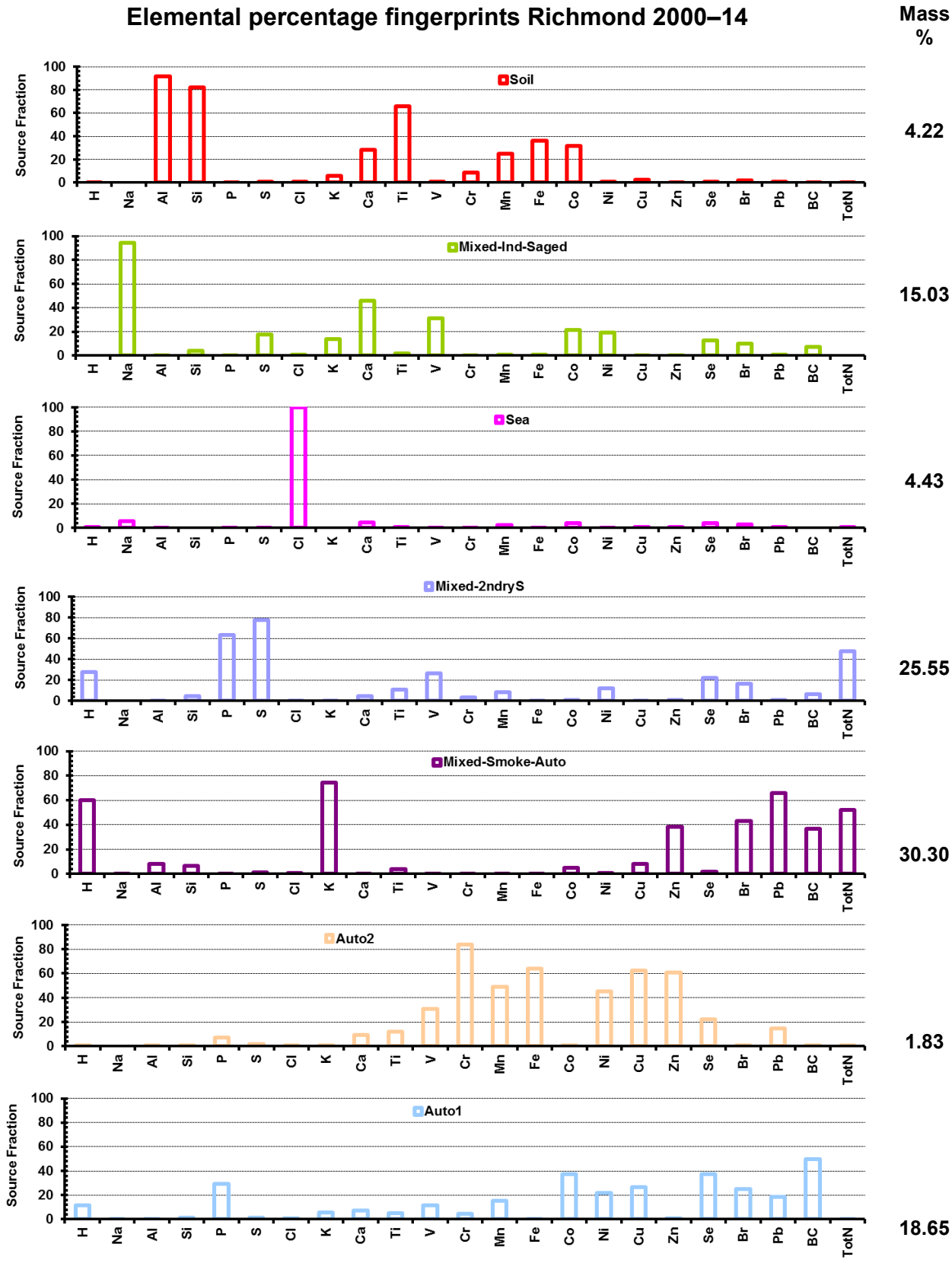

Figure 10-6. Plot of the percentage elemental concentrations in each fingerprint at Richmond.

Figure 10-7 shows the plots of the sum of the calculated PMF fingerprint masses (FitMass) versus the observed or gravimetric $PM_{2.5}$ mass. Note the fit has a gradient of 0.986 which means that the PMF analysis has fitted the total $PM_{2.5}$ mass to better than 1.4% over the mass range 0–40µgm^{-3}; also, the correlation coefficient was 0.95 which is excellent.

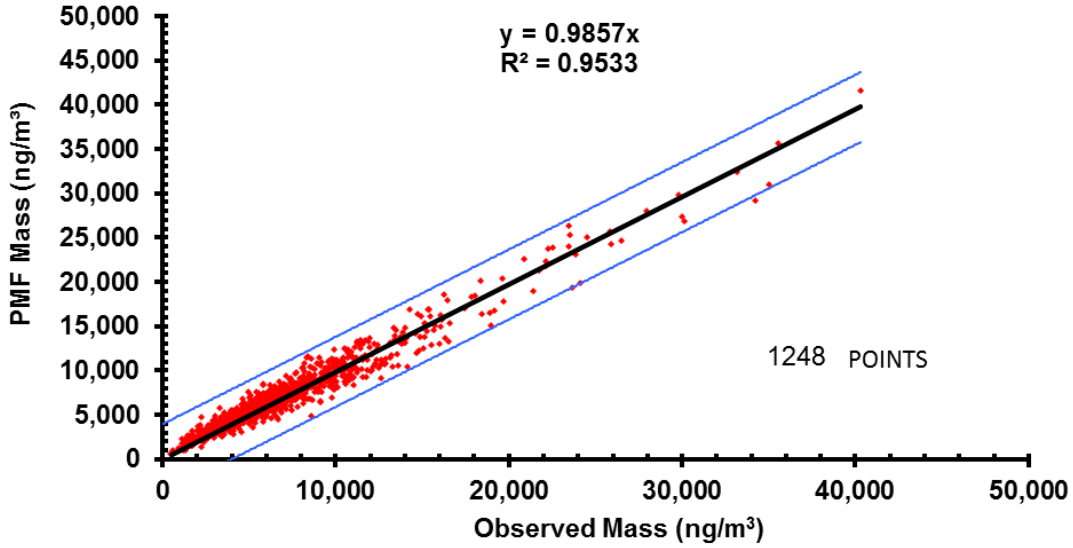

Figure 10-7. Richmond PMF mass versus gravimetric mass, 2000–14. The tramlines represent six standard deviations either side of the linear least squares fitted solid line.

The two tramlines either side of the line of best fit represent ±6 standard deviations from this line and it is obvious that most points outside these tramlines have been removed.

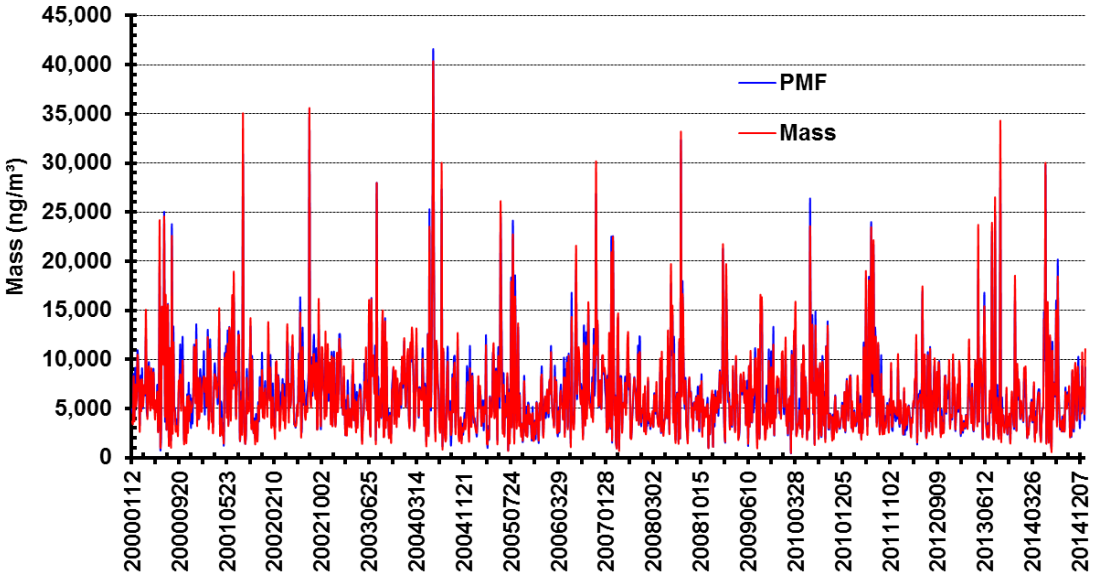

Figure 10-8. Richmond time series PMF mass and gravimetric mass, 2000–14.

Figure 10-8 is a daily time series plot of the PMF mass and gravimetric mass of the data shown in **Figure 10-7** above. The x-axis is the date using the convention 20040725 to represent 25 July 2004. It again demonstrates the excellent fit of the PMF calculates to the measured data. All the main daily and seasonal variations over the15-year study period are well reproduced.

Table 10-2 summarises the seven fingerprint descriptions for the Richmond site. The PMF calculations use these seven factors or source fingerprints for the whole of the study period. They are unique to this site and under this analysis their elemental composition does not change with season, only their relative contributions to the total measured fine mass.

The table lists the fingerprint names, the key driver elements for that fingerprint which helped define the name, the key percentage elements contributing to that fingerprint, which elements in the fingerprint were poorly fitted by the PMF process and finally comments as to why this factor was so named. Elements that were poorly fitted by the PMF process were generally of low concentration (< 20ngm^{-3}), had high MDLs and large errors associated with the measurement. Consequently their weighting was such that they did not drive the PMF fits and their contributions to the least squares fitting process were minimal.

Table 10-2. Summary of the seven fingerprint descriptions for the Richmond site.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
Richmond ASP18 $PM_{2.5}$	F1	Soil	4.22	0.78	1,248	Al, Si, K, Ca, Ti, Fe	Al, Si, Ca, Ti, Fe	P, V, Cr, Co, Ni, Se	None	Clearly a soil factor with (Al/Si) ratio =0.32, contains 90%of the total Al, 80% of the total Si, 65% of the Ti and 35% of the Fe and Co.
	F2	Mixed-2ndryS	25.55			H, S, BC, TotN	H, P, S, V, Ni, Se, TotN		Summer high	A (S/H) ratio =4.0 consistent with fully neutralised ammonium sulfate. Contains 30% of the total H, 60% of the total P, 80% of the total S, 27% of the total V and 12% of the total Ni which indicates a heavy oil combustion component as well and 48% of the total TotN indicating ammonium sulfate. The fact that we have a significant separate secondary sulfate source suggests that a large fraction of this fingerprint is related to coal combustion for power generation as coal combustion for power produces 98% of the emitted sulfur dioxide in NSW.
	F3	Mixed-Ind-Saged	15.03			Na, Si, S, K, Ca, V, Ni, BC	Na, S, Ca, V, Ni		Summer high	This factor is driven by Na, S and BC with no Cl and no TotN, so it cannot be NaCl or $NaNO_3$; it is probably a mixed aged industrial sulfate source. The (Na/S) ratio =3.0 is higher than expected in sodium sulfate [(Na/S)=1.4] so the excess Na must be in another chemical form such as carbonate or oxide (not nitrate). It does not appear to be calcium sulfate either as we have excess Ca maybe as either oxide or carbonate. This may also contain some components from coal combustion fingerprint from power stations as it does contain 18% of the total S, 13% of the total Se and 8% of the total BC.
	F4	Sea	4.43			H, Na, Cl, Ca, TotN	Na, Cl		Summer high	Clearly a sea salt factor dominated by Na and Cl with some Ca. Contains 5% of the total Na, 100% of the total Cl, so it is a weak source at 4.4% of the total fine mass. Richmond is an inland site. There is significant excess Cl not as NaCl in sea salt; this is probably occurring as $CaCl_2$ and ammonium chloride as this best fits the missing mass for this factor and TotN also drives this factor.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
	F5	Mixed-Smoke-Auto	30.30			H, Al, Si, S, Cl, K, Zn, Br, Pb, BC, TotN	H, K, Zn, Br, Pb, BC, TotN		Winter high	This factor is driven by H (organics), K, Zn and BC from biomass burning with significant soil component (Al, Si) and Br and Pb. Contains 60% of the total H, 75% of the total K, 38% of the total Zn, 37% of the total BC and 52% of the total TotN, probably as ammonium nitrate; hence labelled smoke. It represents 30% of the total fine mass at Richmond and has a strong seasonal variation, being higher in the winter months.
	F6	Auto2	1.83			S, Ca, Fe, Zn	V, Cr, Mn, Fe, Ni, Cu, Zn		Slight winter high	This factor is driven by a range of metals P, Ca, Fe, Zn and Pb related to combustion engines and oils. Contains 30% of the total V, 80% of the total Cr, 50% of the Mn, 60% of the Fe, 45% of the total Ni, 60% of the total Cu and total Zn. It is a relatively small factor representing only 2% of the total fine mass.
	F7	Auto1	18.65			H, BC	P, Co, Ni, Cu, Se, BC		Winter high	This factor is driven mainly by H(organics) and BC with minor heavy metal tracers. It contains only 10% of the total H, no TotN but 50% of the total BC. It represents 17% of the total fine mass and has strong seasonal variations being higher in the winter months.

Mascot

The 15-year dataset at Mascot was optimally fitted with seven factors or fingerprints using 1402 sampling days with a χ^2 =1.07. The seven factors were, Soil, mixed industrial aged sulfate (Mixed-Ind-Saged), Sea, mixed secondary sulfate (Mixed-2ndryS), mixed smoke and automobiles (Mixed-Smoke-Auto) and two automobile sources (Auto1 and Auto2). Each of these is described in detail in the graphs and tables below. The first plots are the seven source fingerprints with the maximum driving element fraction normalised to unity in each plot and the second plots are the percentage contribution each element makes to each given fingerprint shown in the first plots. As indicated previously several of these source fingerprints are mixed sources being driven by several different emission sources. For example, the smoke fingerprint is primarily driven by fine potassium from biomass burning but also contains components from automobiles and retrained soils (Al, Si, S, Fe, Cu, Zn, Pb, Br, BC); hence it is named Mixed-Smoke-Auto to reflect this.

At Mascot sulfur appears primarily in Mixed-2ndryS (74%) and Mixed-Ind-Saged (20%). For the Mixed-2ndryS fingerprint the ratio (H/S) =0.16, a little lower than expected for ammonium sulfate. This fingerprint contains 17% of the total H, 88% of the total P, 74% of the total S, 48% of the total V, 29% of the total Ni, 1% of the total BC and 43% of the total TotN, clearly a secondary sulfate factor with a significant heavy oil combustion component as well.

The Mixed-Ind-Saged fingerprint is driven mainly by Na, S, Ca and BC but contains 55% of the total Ca, around 25% of the total V, Co and Ni. It contains 95% of the total Na but no TotN. The (Na/S) ratio =5.2, much higher than expected for sodium sulfate [(Na/S)=1.4] so sodium must be in another form, oxide or carbonate for example. It contains no Cl and 20% of the total sulfur so aged sea salt with Cl loss can only be a small fraction of this fingerprint, which is 23% of the total fine mass. It is similar in nature to the Mixed-Ind-Saged fingerprint at Richmond and may be a fossil fuel combustion source mixed with other industrial sources at Mascot.

The Mixed-Smoke-Auto fingerprint is driven by total H (organics), S, K, BC and TotN from biomass burning with contributions from soil (Al, Si) and motor vehicles (Fe, Cu, Zn, Br, Pb). It appears to be a mixed smoke source. It contains 49% of the total H, 68% of the total K, around 40% of the total Zn and Se, 57% of the total Br, 70% of the total Pb, 29% of the total BC and 55% of the total TotN, probably as ammonium nitrate. It represents 23% of the total fine mass at Mascot and has a strong seasonal variation, being higher in the winter months.

The two Auto fingerprints together represent 23% of the total fine mass and contain 61% of the total Zn and 89% of the total Cu, 76% of the total Cr, 30% of the total Pb and 46% of the BC between them, which points to motor vehicular emissions.

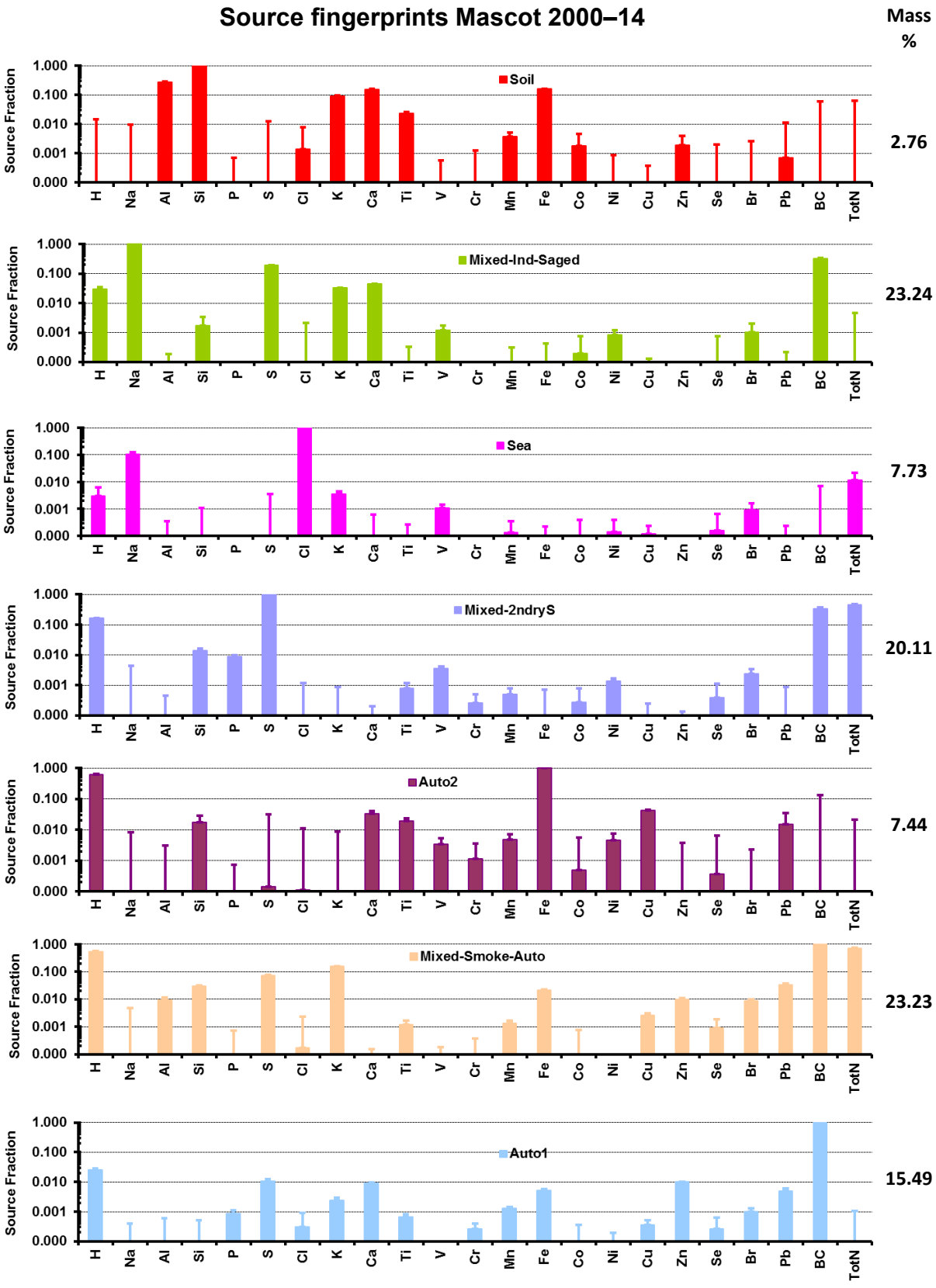

Figure 10-9. Plot of the PMF elemental fractions in the fingerprints for the Mascot site normalised to unity for maximum element.

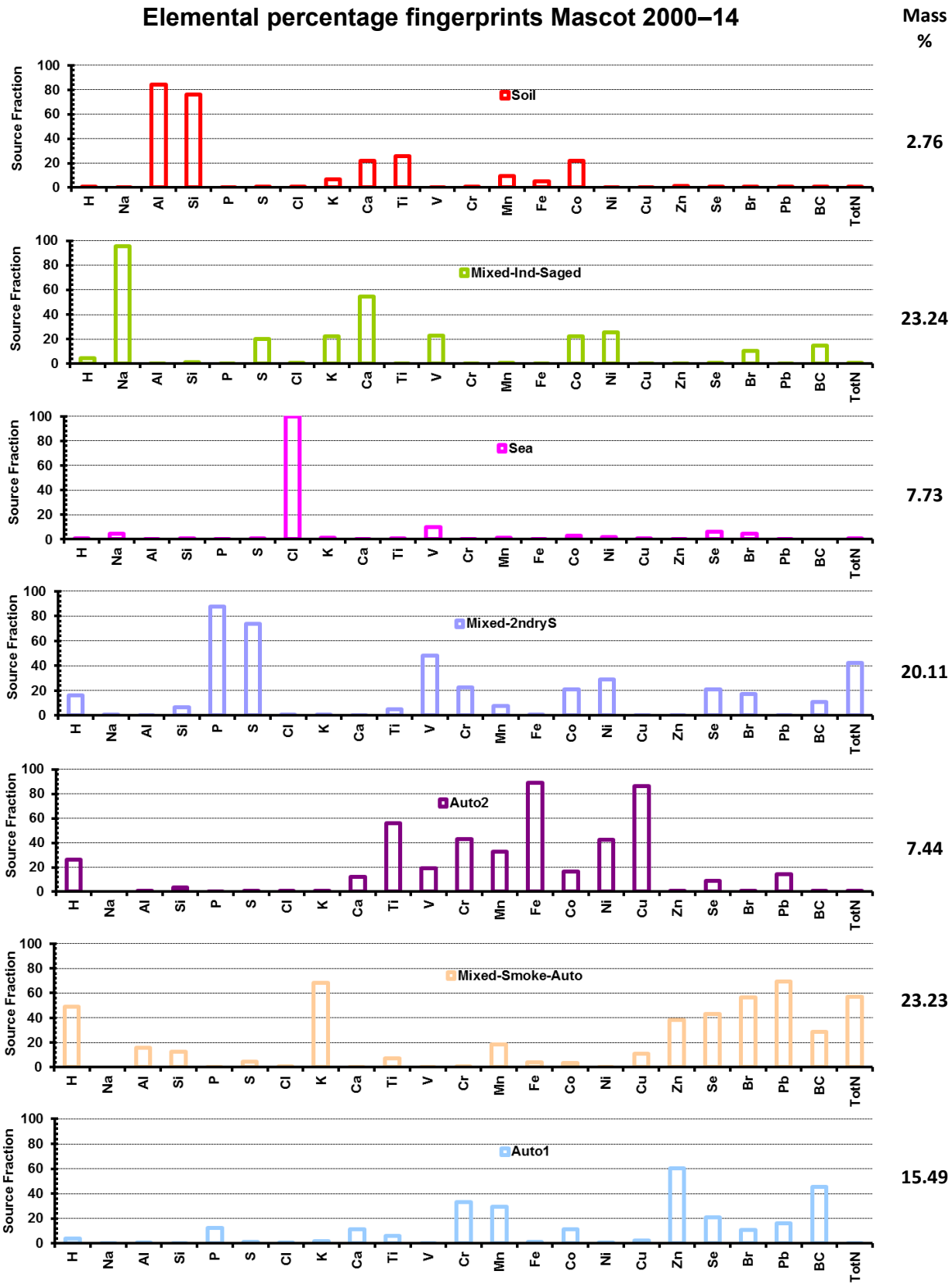

Figure 10-10. Plot of the percentage elemental concentrations in each fingerprint at Mascot.

Figure 10-11 shows the plots of the sum of the calculated PMF fingerprint masses (FitMass) versus the observed or gravimetric $PM_{2.5}$ mass. Note the fit has a gradient of 0.987 which means that the PMF analysis has fitted the total $PM_{2.5}$ mass to better than 1.3% over the mass range 0–40µgm^{-3}; also, the correlation coefficient was 0.95 which is excellent.

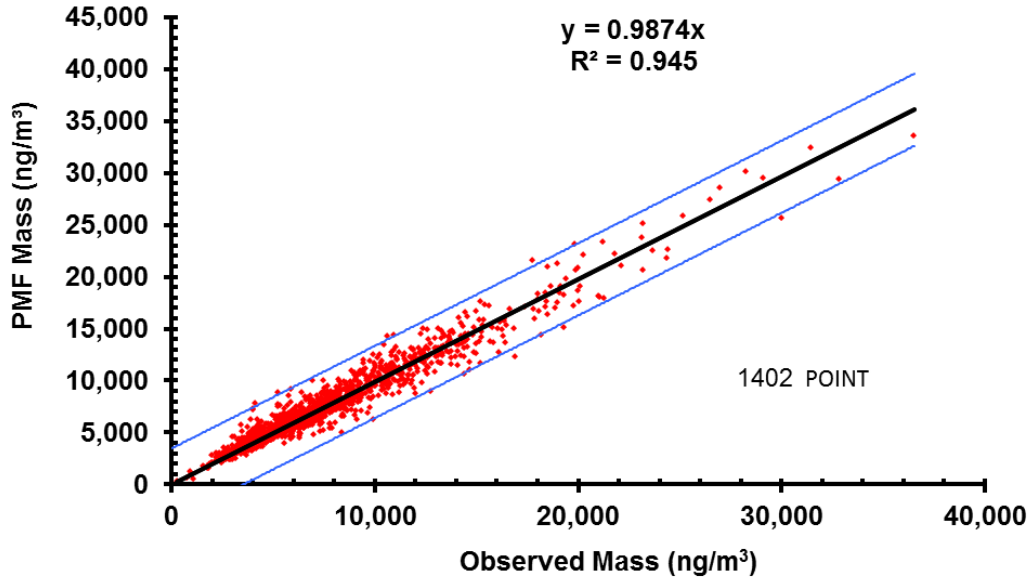

Figure 10-11. Mascot PMF mass versus gravimetric mass, 2000–14. The tramlines represent five standard deviations either side of the linear least squares fitted solid line.

The two tramlines either side of the line of best fit represent ±5 standard deviations from this line and it is obvious that most points outside these tramlines have been removed.

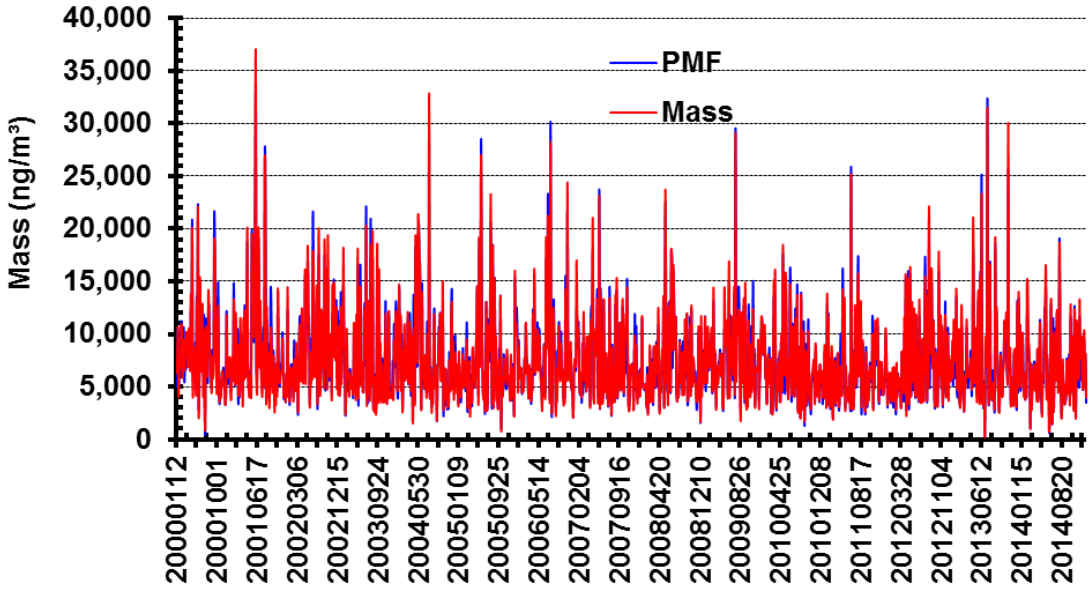

Figure 10-12. Mascot daily time series PMF mass and gravimetric mass, 2000–14.

Figure 10-12 is a daily time series plot of the PMF mass and gravimetric mass of the data shown in **Figure 10-11** above. The x-axis is the date using the convention 20040725 to represent 25 July 2004. It again demonstrates the excellent fit of the PMF calculates to the

measured data. All the main daily and seasonal variations over the15-year study period are well reproduced.

Table 10-3 summarises the seven fingerprint descriptions for the Mascot site. The PMF calculations use these seven factors or source fingerprints for the whole of the study period. They are unique to this site and under this analysis their elemental composition does not change with season, only their relative contributions to the total measured fine mass.

The table lists the fingerprint names, the key driver elements for that fingerprint which helped define the name, the key percentage elements contributing to that fingerprint, which elements in the fingerprint were poorly fitted by the PMF process and finally comments as to why this factor was so named. Elements that were poorly fitted by the PMF process were generally of low concentration (<20ngm^{-3}), had high MDLs and large errors associated with the measurement. Consequently their weighting was such that they did not drive the PMF fits and their contributions to the least squares fitting process were minimal.

Table 10-3. Summary the seven fingerprint descriptions for the Mascot site.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
Mascot ASP23 $PM_{2.5}$	F1	Soil	2.76	1.07	1402	Al, Si, K, Ca, Ti, Fe	Al, Si, Ca, Ti, Co	V, Cr, Co, Ni, Se, Br	None	Clearly a soil factor with (Al/Si) =0.28, contains 85% of the total Al, 80% of the total Si but only 5% of the total Fe, but does have 20% of the total Co.
	F2	Mixed-Ind-Saged	23.24			H, Na, S, K, Ca, BC	Na, S, K, Ca, V, Co, Ni, BC		Summer high	This factor is driven mainly by Na, S, Ca and BC but contains 55% of the total Ca, around 25% of the total V, Co and Ni. It contains 95% of the total Na but no TotN. The (Na/S) ratio =5.2, much higher than expected for sodium sulfate [(Na/S)=1.4] so sodium must be in another form, oxide or carbonate for example. It contains no Cl and 20% of the total S so aged sea salt with Cl loss can only be a small fraction of this factor which is 23% of the total fine mass. It is similar in nature to Mixed-Ind-Saged factor at Richmond and may be a fossil fuel combustion source mixed with other industrial sources at Mascot.
	F3	Mixed-2ndryS	20.11			H, S, V, Ni, BC, TotN	P, S, V, Ni, TotN		Summer high	A (S/H) ratio =6.1, high but consistent with neutralised ammonium sulfate. Contains 17% of the total H, 88% of the total P, 74% of the total S, 48% of the total V, 29% of the total Ni, 1% of the total BC and 43% of the total TotN, clearly a secondary sulfate factor with a significant heavy oil combustion component as well.
	F4	Sea	7.73			Na, S, Cl	Cl		Summer high	A sea salt factor dominated by Na and Cl with (Cl/Na) ratio =9.7 which is high for sea salt, indicating excess Cl or reduced Na. Contains only 5% of the total Na, 100% of the total Cl and no TotN, so the excess Cl cannot be ammonium chloride, similar again to the Sea factor at Richmond. Its contribution to the total fine mass of 8% is twice that at Richmond as Mascot is closer to the ocean.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
	F5	Mixed-Smoke-Auto	23.23			H, Al, Si, S, K, Fe, Zn, Br, Pb, BC, TotN	H, K, Zn, Se, Br, Pb, BC, TotN		Winter high	This factor is driven by H (organics), S, K, BC and TotN from biomass burning with contributions from soil (Al, Si) and motor vehicles (Fe, Cu, Zn, Br, Pb). It appears to be a mixed smoke source. It contains 49% of the total H, 68% of the total K, around 40% of the total Zn and Se, 57% of the total Br, 70% of the total Pb, 29% of the total BC and 55% of the total TotN, probably as ammonium nitrate. It represents 23% of the total fine at Mascot and has a strong seasonal variation, being higher in the winter months.
	F6	Auto2	7.44			H, Fe	H, Ti, V, Cr, Mn, Fe, Co, Ni, Cu		Winter high	This factor is driven by H and Fe and a range of heavy metals. Contains 55% of the total Ti, 20% of the total V, 40% of the total Cr, 35% of the Mn, 90% of the Fe, 40% of the total Ni, 85% of the total Cu but no BC and no TotN. It represents 7% of the total fine mass and is similar to the Auto2 factor at Richmond.
	F7	Auto1	15.49			H, S, Ca, Fe, Zn, Pb, BC	Cr, Mn, Zn, BC		Winter high	This factor is driven mainly by H(organics) and BC with Ca, Fe, Zn and Pb. It contains less than 5% of the total H but 60% of the total Zn and 45% of the total BC and no TotN. It represents 15% of the total fine mass and has strong seasonal variations being higher in the winter months.

Liverpool

The 15-year dataset at Liverpool was optimally fitted with seven factors or fingerprints using 1453 sampling days with a χ^2 =0.97. The seven factors were, Soil, mixed industrial aged sulfate (Mixed-Ind-Saged), Sea, mixed secondary sulfate (Mixed-2ndryS), mixed smoke and automobiles (Mixed-Smoke-Auto) and two automobile sources (Auto1 and Auto2). Each of these is described in detail in the graphs and tables below. The first plots are the seven source fingerprints with the maximum driving element fraction normalised to unity in each plot and the second plots are the percentage contribution each element makes to each given fingerprint shown in the first plots. As indicated previously several of these source fingerprints are mixed sources being driven by several different emission sources. For example, the smoke fingerprint is primarily driven by fine potassium from biomass burning but also contains components from automobiles (Fe, Pb, Br); hence it is named Mixed-Smoke-Auto to reflect this.

As with the Mascot site, total sulfur at the Liverpool site is dominated by the two sources Mixed-2ndryS (77%) and Mixed-Ind-Saged (22%). The Mixed-2ndryS fingerprint has a ratio (H/S) =0.23, consistent with neutralised ammonium sulfate. It contains 25% of the total H, 96% of the total P, 77% of the total S, 41% of the total V, 26% of the total Ni, 7% of the BC and 44% of the TotN. This is clearly a secondary sulfate factor with a heavy oil combustion component as well.

The Mixed-Ind-Saged fingerprint is driven mainly by Na, S, Ca, V, and Ni. It contains 93% of the total Na, 37% of the total Ca, 40% of the total V and Ni. It has no BC or TotN. The (Na/S) ratio =3.2, much higher than expected for sodium sulfate [(Na/S)=1.4] so sodium must be in another form, oxide or carbonate for example. It contains no Cl and 22% of the total sulfur so aged sea salt with Cl loss can only be a small fraction of this factor, which is 14% of the total fine mass. The excess Na in this factor cannot be sodium nitrate as TotN is zero. It is similar in nature to Mixed-Ind-Saged factors at the Richmond and Mascot sites and may be a partial fossil fuel, heavy oil burning or coal combustion source mixed with other industrial sulfate sources at Liverpool.

The two Auto fingerprints together represent 20% of the total fine mass and contain 75% of the total Zn, 29% of the total Cu, 52% of the total Cr, 51% of the total Fe, 8% of the total Pb and 47% of the BC between them, which points to motor vehicular emissions.

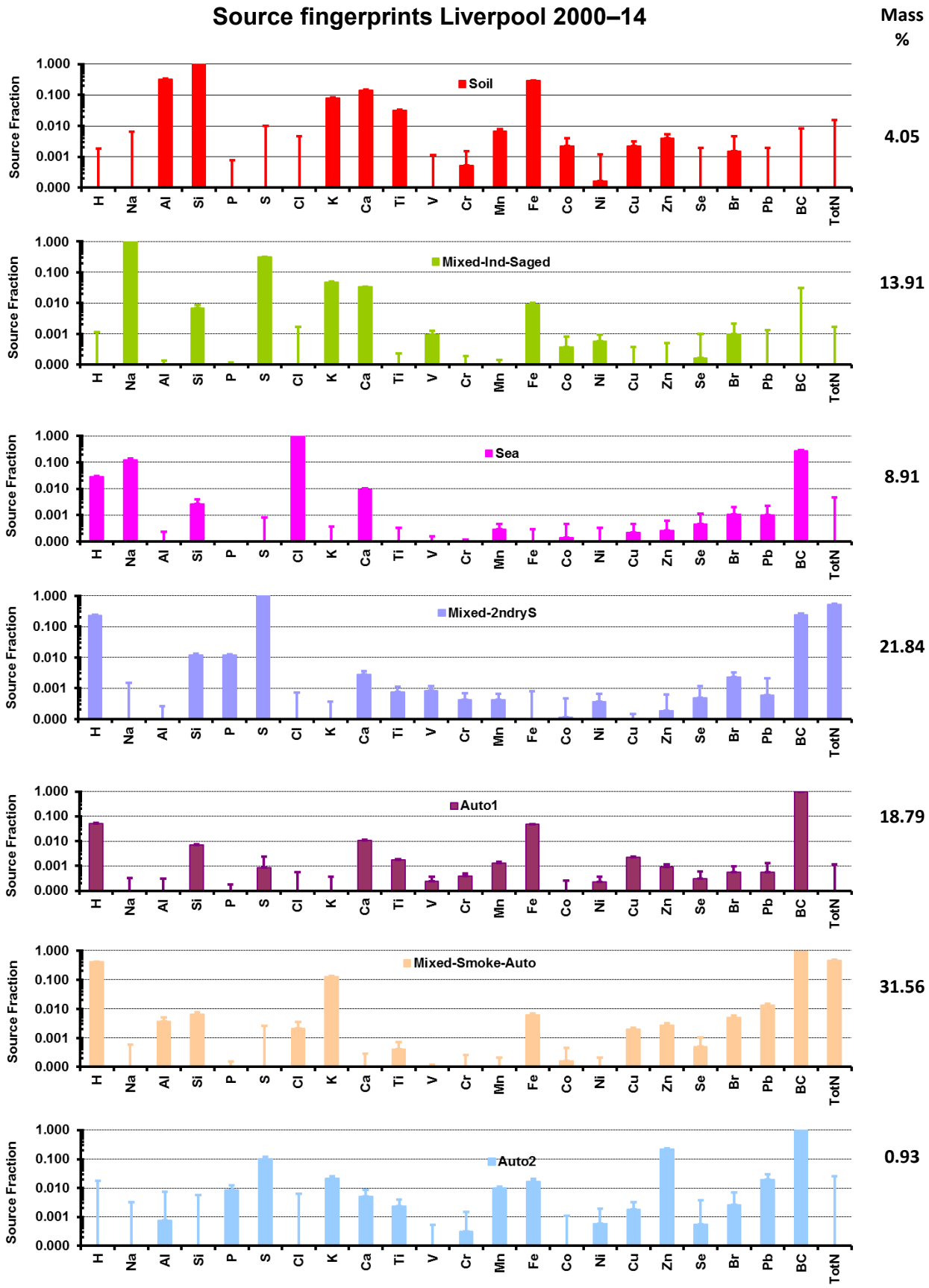

Figure 10-13. Plot of the PMF elemental fractions in the fingerprints for the Liverpool site normalised to unity for maximum element.

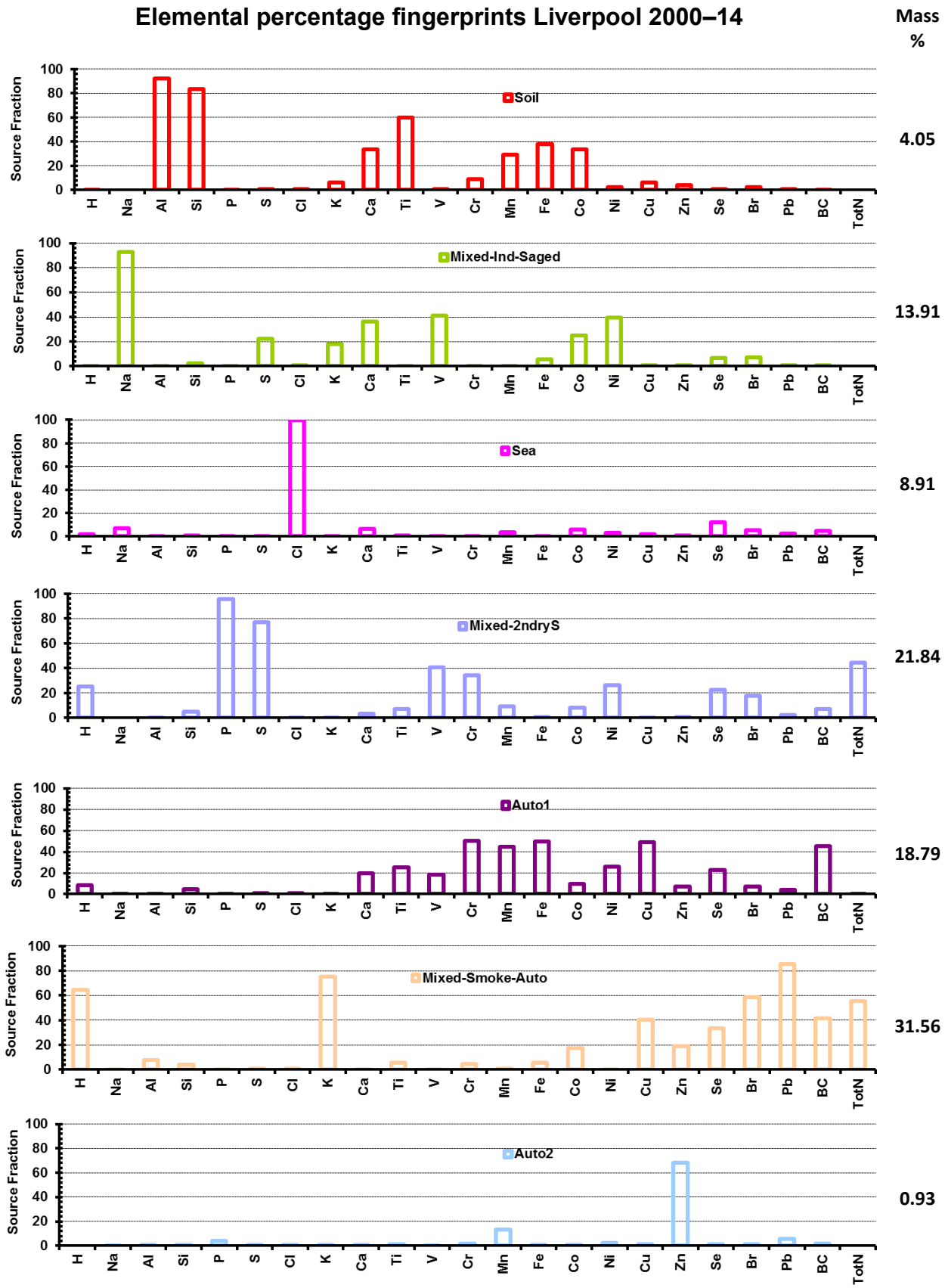

Figure 10-14. Plot of the percentage elemental concentrations in each fingerprint at Liverpool.

Figure 10-15 shows the plots of the sum of the calculated PMF fingerprint masses (FitMass) versus the observed or gravimetric $PM_{2.5}$ mass. Note the fit has a gradient of 0.990 which means that the PMF analysis has fitted the total $PM_{2.5}$ mass to better than 1.0% over the mass range 0–40µgm^{-3}; also, the correlation coefficient was 0.96 which is excellent.

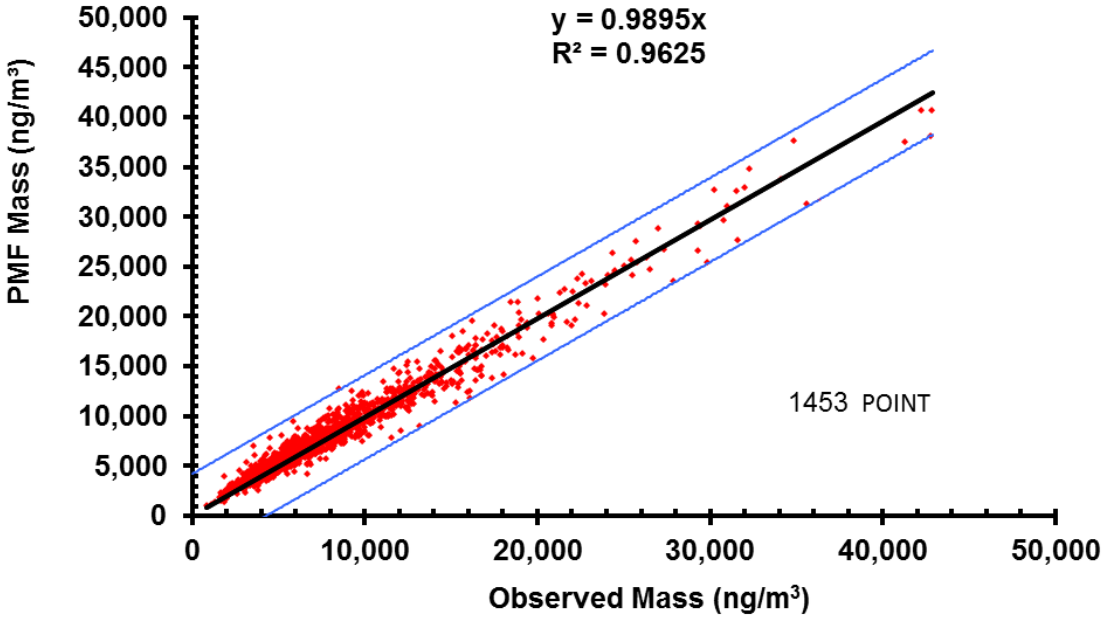

Figure 10-15. Liverpool PMF mass versus gravimetric mass, 2000–14. The tramlines represent five standard deviations either side of the linear least squares fitted solid line.

The two tramlines either side of the line of best fit represent ±5 standard deviations from this line and it is obvious that most points outside these tramlines have been removed.

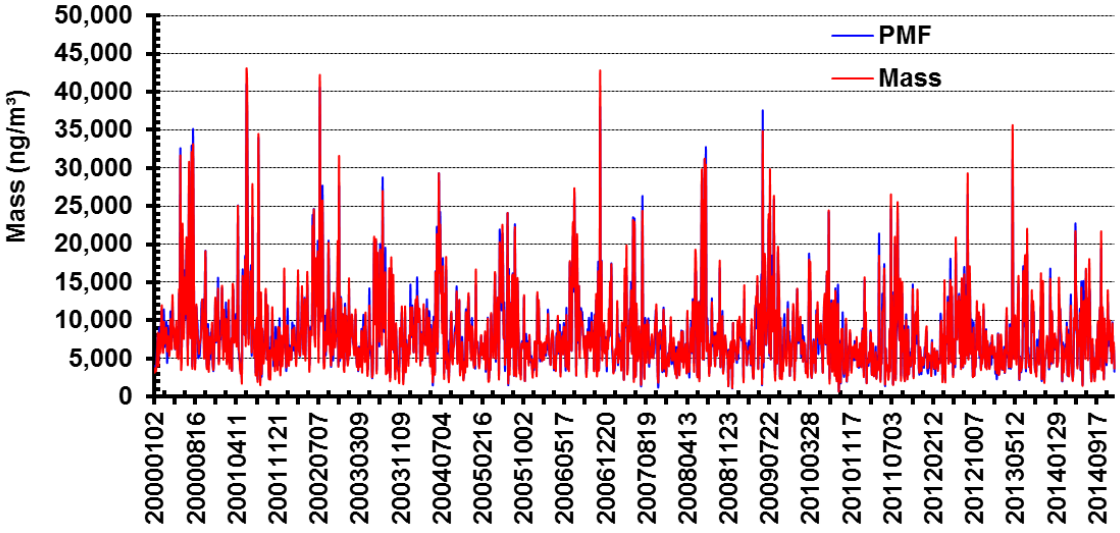

Figure 10-16. Liverpool daily time series PMF mass and gravimetric mass, 2000–14.

Figure 10-16 is a daily time series plot of the PMF mass and gravimetric mass of the data shown in **Figure 10-15** above. The x-axis is the date using the convention 20040725 to represent 25 July 2004. It again demonstrates the excellent fit of the PMF calculates to the

measured data. All the main daily and seasonal variations over the15-year study period are well reproduced.

Table 10-4 summarises the seven fingerprint descriptions for the Liverpool site. The PMF calculations use these seven factors or source fingerprints for the whole of the study period. They are unique to this site and under this analysis their elemental composition does not change with season, only their relative contributions to the total measured fine mass.

The table lists the fingerprint names, the key driver elements for that fingerprint which helped define the name, the key percentage elements contributing to that fingerprint, which elements in the fingerprint were poorly fitted by the PMF process and finally comments as to why this factor was so named. Elements that were poorly fitted by the PMF process were generally of low concentration (<20ngm^{-3}), had high MDLs and large errors associated with the measurement. Consequently their weighting was such that they did not drive the PMF fits and their contributions to the least squares fitting process were minimal.

Table 10-4. Summary of the seven fingerprint descriptions for the Liverpool site.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
Liverpool ASP31 $PM_{2.5}$	F1	Soil	4.05	0.97	1,453	Al, Si, K, Ca, Ti, Fe	Al, Si, Ca, Ti, Mn, Fe, Co	V, Cr, Co, Ni, Se	None	Clearly a soil factor with (Al/Si) =0.33, contains 90% of the total Al, 80% of the total Si, 35% of the total Ca, 60% of the total Ti, 30% of the total Mn, 40% of the total Fe and 35% of the total Co.
	F2	Mixed-Ind-Saged	13.91			Na, S, K, Ca, Fe	Na, Ca, V, Co, Ni		Summer high	This factor is driven mainly by Na, S, Ca, V, and Ni. It contains 93% of the total Na, 37% of the total Ca, 40% of the total V and Ni. It has no BC or TotN. The (Na/S) ratio =3.2, much higher than expected for sodium sulfate [(Na/S)=1.4] so sodium must be in another form, oxide or carbonate for example. It contains no Cl and 22% of the total S so aged sea salt with Cl loss can only be a small fraction of this factor, which is 14% of the total fine mass. The excess Na in this factor cannot be sodium nitrate as TotN is zero. It is similar in nature to Mixed-Ind-Saged factors at the Richmond and Mascot sites and may partially be a fossil fuel, heavy oil or coal combustion source mixed with other industrial sources at Liverpool. The fact that we have a significant separate secondary sulfate source suggests that a large fraction of this fingerprint is related to coal combustion for power generation, as coal combustion for power produces 98% of the emitted sulfur dioxide in NSW.
	F3	Mixed-2ndryS	21.84			H, S, BC, V, Ni, TotN	H, P, S, V, Cr, Ni, Se, TotN		Summer high	A (S/H) ratio =4.3, consistent with neutralised ammonium sulfate. Contains 25% of the total H, 96% of the total P, 77% of the total S, 41% of the total V, 26% of the total Ni, 7% of the BC and 44% of the TotN. This is clearly a secondary sulfate factor with a heavy oil combustion component as well. The fact that we have a significant separate secondary sulfate source suggests that a large fraction of this fingerprint is related to coal combustion for power generation, as coal combustion for power produces 98% of the emitted sulfur dioxide in NSW.

Site	Factor	Notional name	%	χ^2	Days	Driver elements	Drivers percent	Elements not well fitted by PMF	Seasonality	Comments
	F4	Sea	8.91			H, Na, Cl, Ca, BC	Na, Cl		Summer high	A sea salt factor dominated by Na and Cl with (Cl/Na) ratio =8.1, which is high for sea salt, indicating excess Cl or reduced Na. It contains only 7% of the total Na, 100% of the total Cl, similar again to Sea factor at Richmond with the excess Cl. Its contribution to the total fine mass of 9% is similar to that at Mascot, which shows the effects of the afternoon sea breezes on Liverpool.
	F5	Mixed-Smoke-Auto	31.56			H, K, BC, TotN	H, K, Cu, Zn, Se, Br, Pb, BC, TotN		Winter high	This factor is driven by H(organics), K, BC and TotN from biomass burning with contributions from soil (Al, Si) and motor vehicles (Fe, Cu, Zn, Br, Pb). It appears to be a mixed smoke source. It contains 60% of the total H, 75% of the total K, 40% of the total Cu and Se, 60% of the total Br, 85% of the total Pb, 40% of the total BC and 55% of the total TotN. It represents 32% of the total fine mass at Liverpool and has a strong seasonal variation, being higher in the winter months.
	F6	Auto1	18.79			H, Si, Ca, Fe, BC	Mn, Zn		Winter high	This factor is driven by H(organics), Fe and BC with Si, Ca and a range of heavy metals. It contains less than 10% of the total H, 50% of the total Cr, 45% of the total Mn, 50% of the total Fe and total Cu and 45% of the total BC with no TotN. It represents 19% of the total fine mass and has strong seasonal variations, being higher in the winter months.
	F7	Auto2	0.93			S, K, Mn, Fe, Zn, Pb, BC	Mn, Zn		Winter high	This factor is driven by S, K, Fe, Pb, Zn and BC. Contains 15% of the total Mn and 70% of the total Zn. It represents only 0.9% of the total fine mass and is similar to the Auto2 factor at Richmond.

11. Fingerprint 15-year summary

In this section we summarise the PMF source fingerprint contributions to the total measured $PM_{2.5}$ mass at each of the four sites. We include in the tables the mass concentrations in μgm^{-3} and the percentage contributions (%) of each fingerprint to the total fine mass.

Lucas Heights

At the Lucas Heights site the $PM_{2.5}$ mass of $(4.93 \pm 2.64)\mu gm^{-3}$ was composed of 3% Soil, 17% Sea spray, 9% Mixed industrial and aged sulfate, 30% secondary sulfates, 12% smoke from biomass burning and diesel motor vehicles and 29% motor vehicles from two different sources.

Table 11-1. Average 15-year fingerprint contributions to the $PM_{2.5}$ total mass at Lucas Heights.

		Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
Lucas Heights	Mass	4.93	4.41	2.64	17.0
1383 days	Soil	0.161	0.076	0.303	4.02
$PM_{2.5}$	Sea	0.739	0.366	0.979	6.85
	Mixed-Ind-Ca-Saged	0.463	0.323	0.490	4.28
	Mixed-2ndryS	1.65	1.14	1.57	12.8
	Mixed-Smoke-Auto	0.611	0.434	0.583	5.50
	Auto1	1.19	0.902	0.921	5.08
	Auto2	0.156	0.127	0.125	0.915
	FitMass	4.97	4.50	2.46	16.4
	Soil%	3.10	1.73	4.44	63.7
	Sea%	17.1	9.55	19.23	78.9
	Mixed-Ind-Ca-Saged%	9.19	8.42	7.40	54.9
	Mixed-2ndryS%	29.8	26.7	17.6	84.8
	Mixed-Smoke-Auto%	11.6	9.87	7.96	69.1
	Auto1%	25.7	23.5	15.4	74.9
	Auto2%	3.50	2.77	3.09	29.5

Richmond

At the Richmond site the $PM_{2.5}$ mass of $(6.50 \pm 4.55)\mu gm^{-3}$ was composed of 4% Soil, 4% Sea spray, 15% Mixed industrial and aged sulfate, 26% secondary sulfates, 30% smoke from biomass burning and diesel motor vehicles and 20% motor vehicles from two different sources.

Table 11-2. Average 15-year fingerprint contributions to the $PM_{2.5}$ total mass at Richmond.

		Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
Richmond	Mass	6.50	5.40	4.55	40.3
1248 days	Soil	0.263	0.144	0.392	5.63
$PM_{2.5}$	Sea	0.234	0.073	0.385	3.42
	Mixed-Ind-Saged	0.815	0.539	0.917	5.64
	Mixed-2ndryS	1.62	1.13	1.50	8.61

		Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
	Mixed-Smoke-Auto	2.52	1.20	3.71	35.4
	Auto1	1.01	0.874	0.634	6.55
	Auto2	0.109	0.090	0.091	0.885
	FitMass	6.57	5.54	4.35	41.6
	Soil%	4.22	2.66	5.37	76.5
	Sea%	4.43	1.39	6.90	42.9
	Mixed-Ind-Saged%	15.0	10.9	14.2	54.7
	Mixed-2ndryS%	25.5	22.5	18.1	80.8
	Mixed-Smoke-Auto%	30.3	25.4	22.1	89.8
	Auto1%	18.6	16.9	10.7	56.4
	Auto2%	1.83	1.54	1.35	9.25

Mascot

At the Mascot site the $PM_{2.5}$ mass of (7.59±4.27)μgm^{-3} was composed of 3% Soil, 8% Sea spray, 23% Mixed industrial and aged sulfate, 20% secondary sulfates, 23% smoke from biomass burning and diesel motor vehicles and 23% motor vehicles from two different sources.

Table 11-3. Average 15-year fingerprint contributions to the $PM_{2.5}$ total mass at Mascot.

		Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
Mascot	Mass	7.59	6.48	4.27	36.5
1402 days	Soil	0.208	0.113	0.275	2.80
$PM_{2.5}$	Sea	0.466	0.273	0.543	4.32
	Mixed-Ind-Saged	1.56	1.28	1.32	8.11
	Mixed-2ndryS	1.52	1.12	1.31	10.5
	Mixed-Smoke-Auto	2.10	1.12	2.65	21.9
	Auto1	1.15	0.775	1.16	8.89
	Auto2	0.584	0.401	0.592	4.85
	FitMass	7.59	6.54	4.14	33.5
	Soil%	2.76	1.74	3.47	46.8
	Sea%	7.73	4.26	8.63	38.5
	Mixed-Ind-Saged%	23.2	22.0	17.1	64.6
	Mixed-2ndryS%	20.1	17.8	13.0	69.0
	Mixed-Smoke-Auto%	23.2	18.8	17.6	75.2
	Auto1%	15.5	13.3	11.4	61.2
	Auto2%	7.44	6.48	5.13	32.6

Liverpool

At the Liverpool site the $PM_{2.5}$ mass of (8.17±5.36)μgm^{-3} was composed of 4% Soil, 9% Sea spray, 14% Mixed industrial and aged sulfate, 22% secondary sulfates, 32% smoke from biomass burning and diesel motor vehicles and 20% motor vehicles from two different sources.

Table 11-4. Average 15-year fingerprint contributions to the $PM_{2.5}$ total mass at Liverpool.

		Average (μgm^{-3})	Median (μgm^{-3})	SD (μgm^{-3})	Max. (μgm^{-3})
Liverpool	Mass	8.17	6.82	5.36	42.9
1453 days	Soil	0.345	0.181	0.495	5.24
$PM_{2.5}$	Sea	0.560	0.310	0.695	5.85
	Mixed-Ind-Saged	0.921	0.704	0.873	7.05
	Mixed-2ndryS	1.75	1.27	1.59	12.7
	Mixed-Smoke-Auto	3.08	1.73	3.78	30.5
	Auto1	1.50	1.15	1.28	11.9
	Auto2	0.080	0.046	0.102	1.04
	FitMass	8.24	6.96	5.15	40.7
	Soil%	4.05	2.58	5.02	69.0
	Sea%	8.91	4.44	11.2	57.2
	Mixed-Ind-Saged%	13.9	11.7	12.0	51.8
	Mixed-2ndryS%	21.8	18.4	15.5	72.3
	Mixed-Smoke-Auto%	31.6	28.5	21.4	85.6
	Auto1%	18.8	16.8	10.5	58.5
	Auto2%	0.93	0.66	1.03	13.9

The seven PMF fingerprints used to fit each of the four sites are remarkably similar and show that the $PM_{2.5}$ mass at each site can be broken down into soil, secondary sulfates, industrial, smoke and automobile contributions whose relative contributions vary at each of the four sites.

12. Time series analysis

As mentioned, PMF analysis is a one-step process that produces both source fingerprints from the inter-element correlations as well as the contribution of each of these fingerprints to the measured fine mass on a daily basis. This allows each fingerprint to be plotted as a function of sampling day over the whole 15-year sampling period. In the sections below we reproduce these daily time series plots for each site for both the daily concentrations (in ngm^{-3}) and the percentage contributions to the total daily $PM_{2.5}$ mass.

Lucas Heights – daily variations

Figure 12-1 shows the daily source contributions during the 15-year study period for the seven source fingerprints at the Lucas Heights site. The bottom plot shows the daily time series for the total $PM_{2.5}$ mass for comparison purposes. The horizontal axis is the same for all plots and uses the date format 20040711 to represent 11 July 2004. Note also that the vertical axis is in (ngm^{-3}) with some plots having different scales so daily variations are more obvious.

Clear summer–winter seasonal variations can be seen in most fingerprints except for the soil fingerprint. Averaging these seasonal variations over a single month or year makes for a more meaningful time series comparison and we have done this in the next section.

The mixed smoke automobile fingerprint represents between 20% and 30% of the fine $PM_{2.5}$ mass in the winter months falling to between 5% and 10% in the summer months. It shows a significant drop from 2000 to 2014, the annual median value dropping from $584ngm^{-3}$ to

312ngm^{-3} respectively, a 47% fall over the 15-year sampling period. Daily spikes in this smoke fingerprint were due to local short-term bushfires or controlled burning events.

The Auto1 plus Auto2 wintertime absolute concentration peaks appear to trend downwards, with the annual median value dropping from 1595ngm^{-3} to 774ngm^{-3} respectively, a 51% fall over the 15-year sampling period. The annual median $PM_{2.5}$ mass dropped from 4.34µgm^{-3} in 2000 to 4.14µgm^{-3} in 2014 or only 5%. So the percentage drop in the Auto1+2 fingerprint from 2000 to 2014 was essentially 49% at Lucas Heights. This fingerprint contributes between 40% and 60% to the total $PM_{2.5}$ mass in the winter months at this site.

The mixed secondary sulfate fingerprint also dropped from 2000 to 2014, the annual median value dropping from 1257ngm^{-3} to 785ngm^{-3} respectively, a 38% fall over the 15-year sampling period.

Like the mixed secondary sulfate fingerprint the mixed industrial, calcium, aged sulfur fingerprint at Lucas Heights has a strong seasonal variation being higher in the summer months. However it had an interesting feature between 2012 and 2014 where it increased by 80% on its average value before 2012. As mentioned in the previous section this was due to building and construction work (hence the high Ca component) in and around the site during this period. Without this step function increase after 2012 this fingerprint would look more like the Mixed-Ind-Saged fingerprint of the other sites.

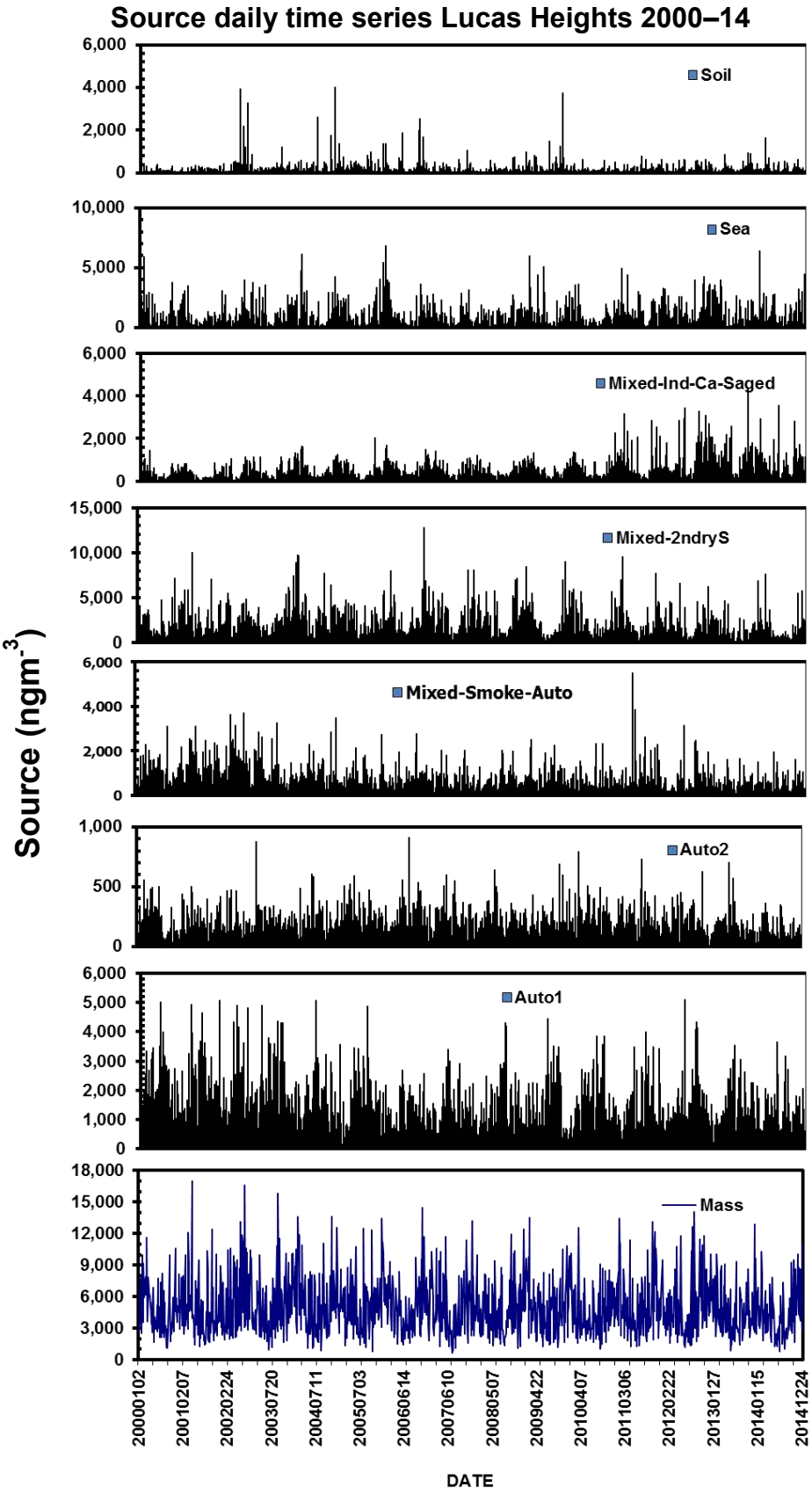

Figure 12-1. Lucas Heights PMF seven factor source contributions to total mass, 2000–14.

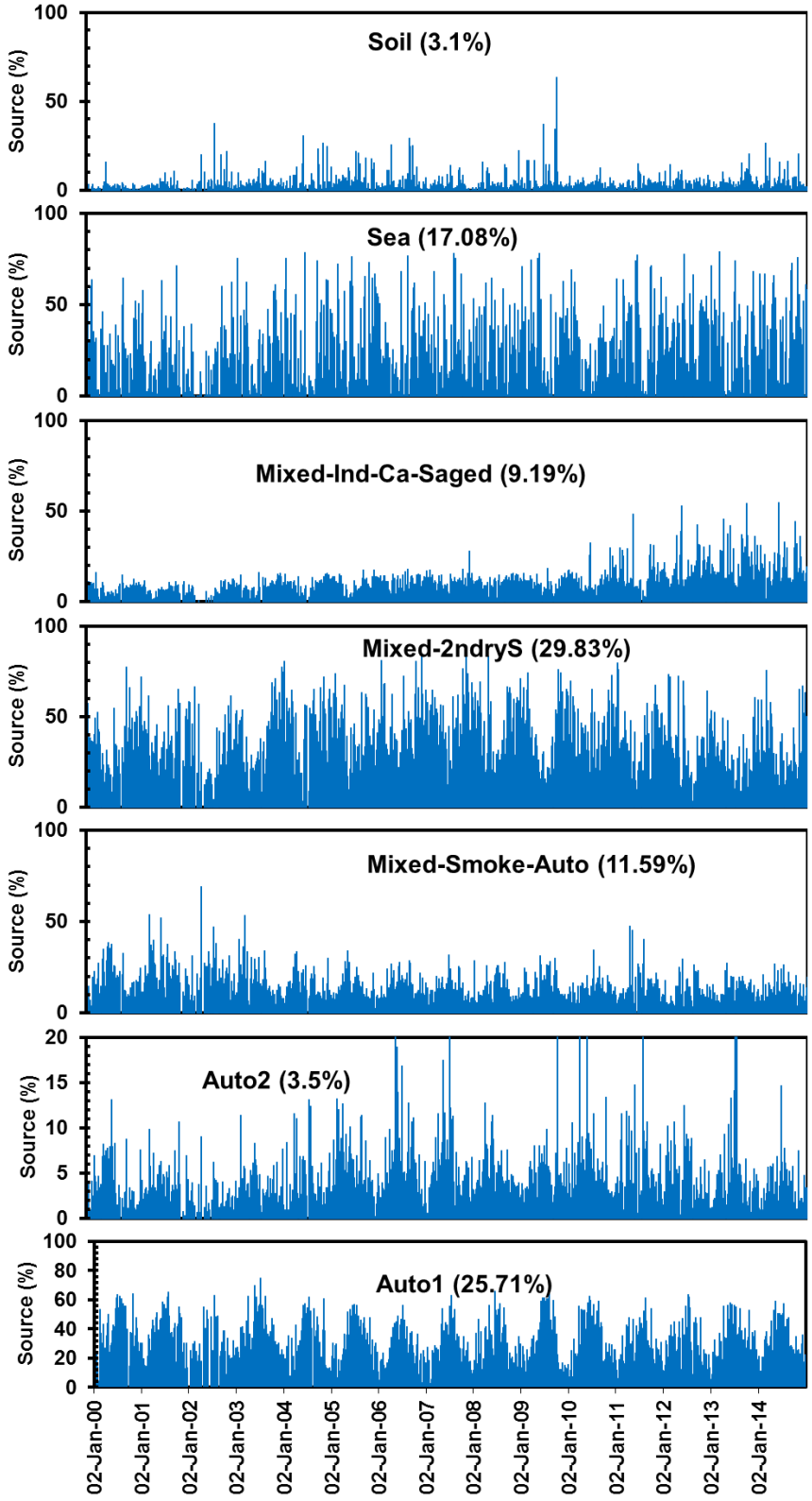

Figure 12-2. Lucas Heights percentage PMF source contributions, 2000–14.

Richmond – daily variations

Figure 12-3 shows the daily source contributions during the 15-year study period for the seven source fingerprints at the Richmond site. The bottom plot shows the daily time series for the total $PM_{2.5}$ mass for comparison purposes. The horizontal axis is the same for all plots and uses the date format 20040711 to represent 11 July 2004. Note also that the vertical axis is in (ngm^{-3}) with some plots having different scales so daily variations are more obvious.

Clear summer–winter seasonal variations can be seen in most fingerprints except for the soil fingerprint. Averaging or looking at the median values for these seasonal variations over a single month or year makes for a more meaningful time series comparison and we have done this in the next section.

The mixed smoke automobile fingerprint at Richmond was more than twice the concentrations of the same fingerprint at the Lucas Heights site. This fingerprint represents between 70% and 90% of the fine $PM_{2.5}$ mass in the winter months, falling to between 10% and 20% in the summer months. It shows a significant drop from 2000 to 2014, the annual median value dropping from 1283ngm^{-3} to 961ngm^{-3} respectively, a 25% fall over the 15-year sampling period. This was probably driven by the varying temperature differences between the two winter periods in these years. Daily spikes in this smoke fingerprint were due to local short-term bushfires or controlled burning events.

The mixed secondary sulfate fingerprint was responsible for between 60% and 70% of the fine $PM_{2.5}$ total mass in the summer months. Its concentrations also dropped from 2000 to 2014, the annual median value dropping from 1501ngm^{-3} to 651ngm^{-3} respectively, a 57% fall over the 15-year sampling period.

The mixed industrial aged sulfur fingerprint at Richmond has a strong seasonal variation, being higher in the summer months. It accounted for between 30% and 50% in the summer months. It varies by up to a factor of two from year to year with no clear long-term trend either up or down.

Again we add the two Auto fingerprints together and see a slight but significant drop in the annual median value, dropping from 1311ngm^{-3} to 822ngm^{-3} (37%) between 2000 and 2014. The annual median $PM_{2.5}$ mass dropped from 5.85μgm^{-3} in 2000 to 5.31μgm^{-3} in 2014 or only 9%. So the percentage drop in the Auto1+2 fingerprint from 2000 to 2014 was essentially 34% at Richmond. This fingerprint contributed between 30% and 50% to the total $PM_{2.5}$ mass in the winter months at this site.

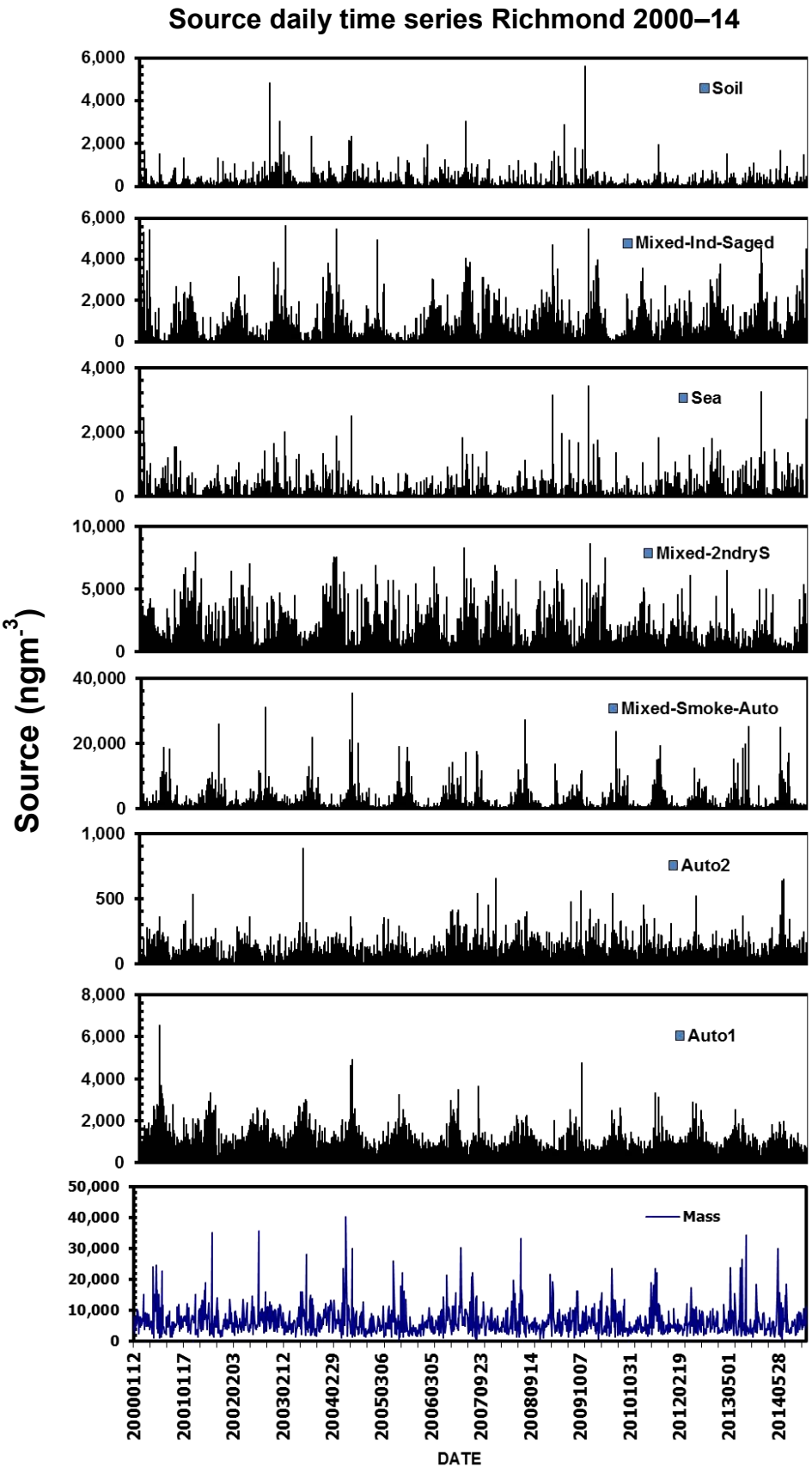

Figure 12-3. Richmond PMF seven factor source contributions to total mass, 2000–14.

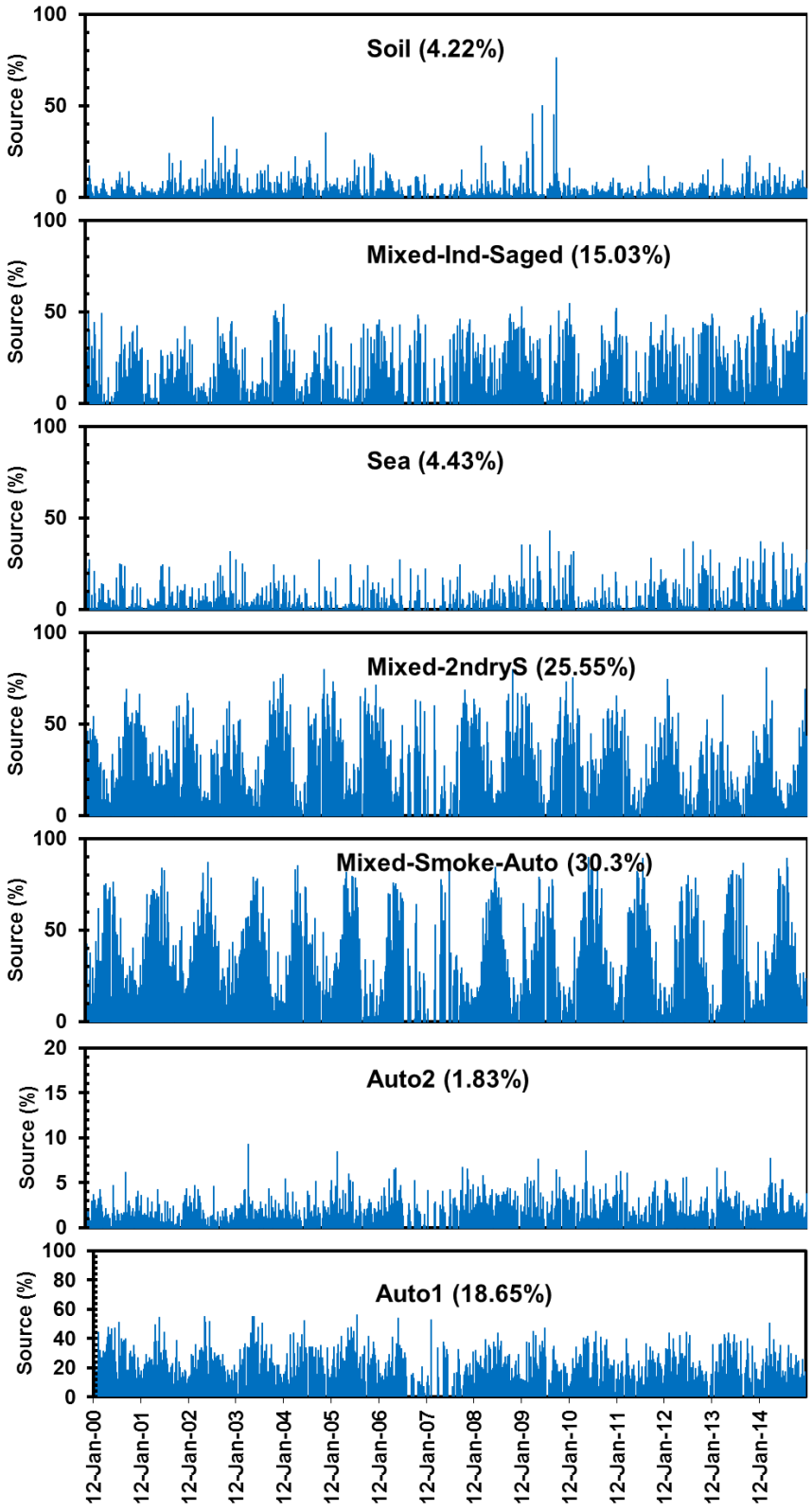

Figure 12-4. Richmond percentage PMF source contributions, 2000–14.

Mascot – daily variations

Figure 12-5 shows the daily source contributions during the 15-year study period for the seven source fingerprints at the Mascot site. The bottom plot shows the daily time series for the total $PM_{2.5}$ mass for comparison purposes. The horizontal axis is the same for all plots and uses the date format 20040711 to represent 11 July 2004. Note also that the vertical axis is in (ngm^{-3}) with some plots having different scales so daily variations are more obvious.

Clear summer–winter seasonal variations can be seen in most fingerprints except for the soil fingerprint. Averaging or looking at the median values for these seasonal variations over a single month or year makes for a more meaningful time series comparison and we have done this in the next section.

The mixed secondary sulfate fingerprint contributed between 40% and 50% to the total $PM_{2.5}$ mass in the summer months. It dropped from 2000 to 2014, the annual median value dropping from 1399ngm^{-3} to 821ngm^{-3} respectively, a 41% fall over the 15-year sampling period.

The mixed industrial aged sulfur fingerprint at Mascot contributed between 40% and 50% to the total $PM_{2.5}$ mass in the summer months. It has a strong seasonal variation, being higher in the summer months. It varies by over 100% over some years however over the 15-year study period the annual median value has increased from 1071ngm^{-3} in 2000 to 1659ngm^{-3} in 2014.

Again we add the two Auto fingerprints together and see a slight but significant drop in the annual median value, dropping from 1575ngm^{-3} to 847ngm^{-3} (46%) between 2000 and 2014. The annual median $PM_{2.5}$ mass dropped from 7.02μgm^{-3} in 2000 to 6.15μgm^{-3} in 2014 or 12%. So the percentage drop in the Auto1+2 fingerprint from 2000 to 2014 was essentially 40% at Mascot. This fingerprint contributed between 40% and 60% to the total $PM_{2.5}$ mass in the winter months at this site.

The mixed smoke automobile fingerprint at Mascot is comparable with the same fingerprint at the Richmond site. This fingerprint represents between 60% and 80% of the fine $PM_{2.5}$ mass in the winter months, falling to between 10% and 20% in the summer months. It shows a significant drop from 2000 to 2014, the annual median value dropping from 1616ngm^{-3} to 751ngm^{-3} respectively, a 54% fall over the 15-year sampling period. As mention previously this smoke-automobile fingerprint is probably more influenced by smoke from polluting diesel vehicles at this site than by smoke from domestic wood heaters, as is Richmond. Since the Auto fingerprints show significant drops between 2000 and 2014 as the fleet gets cleaner, it is not surprising that this smoke-auto fingerprint at Mascot also shows a similar long-term decrease with time.

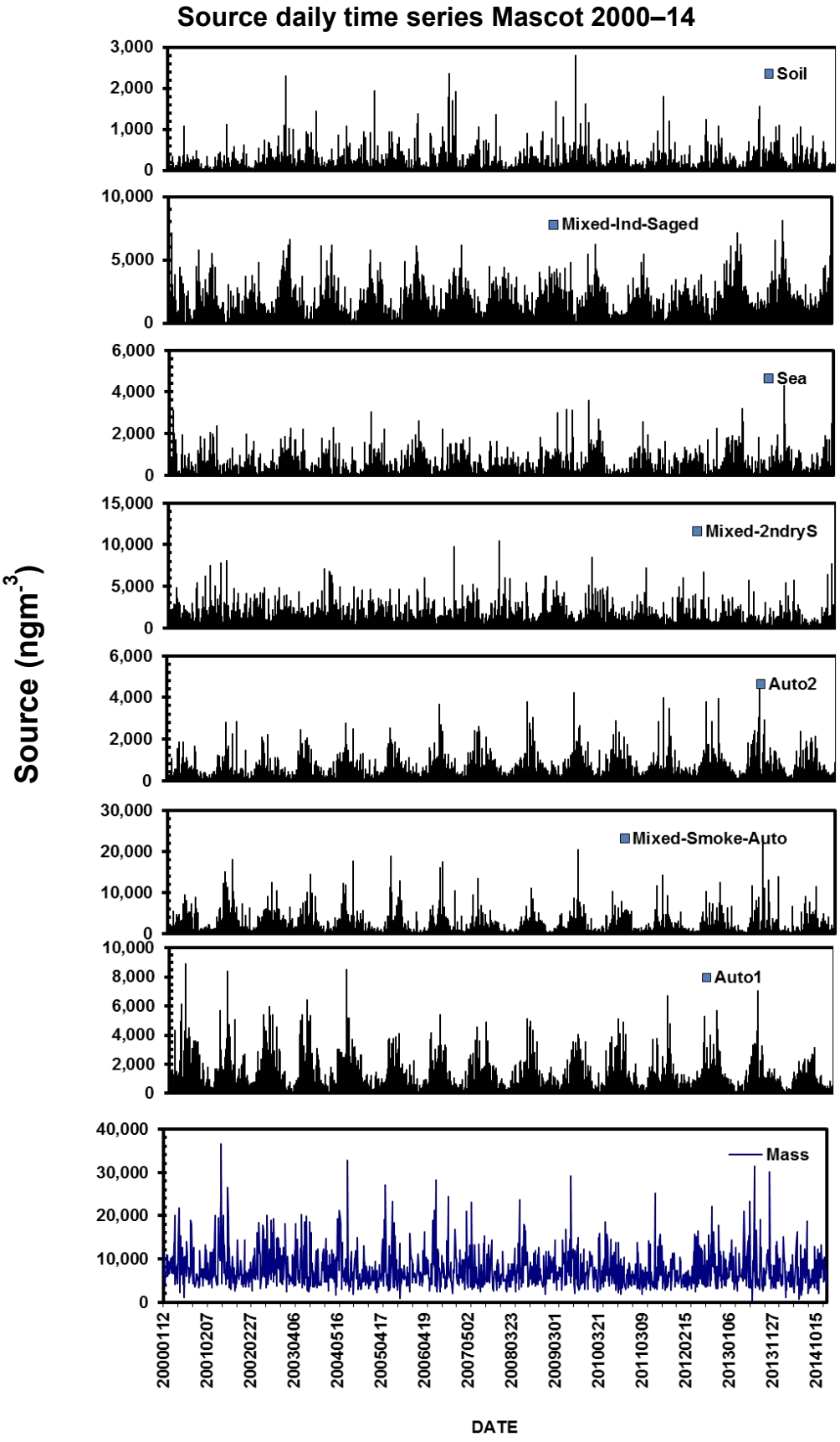

Figure 12-5. Mascot PMF 7 factor source contributions to total mass, 2000–14.

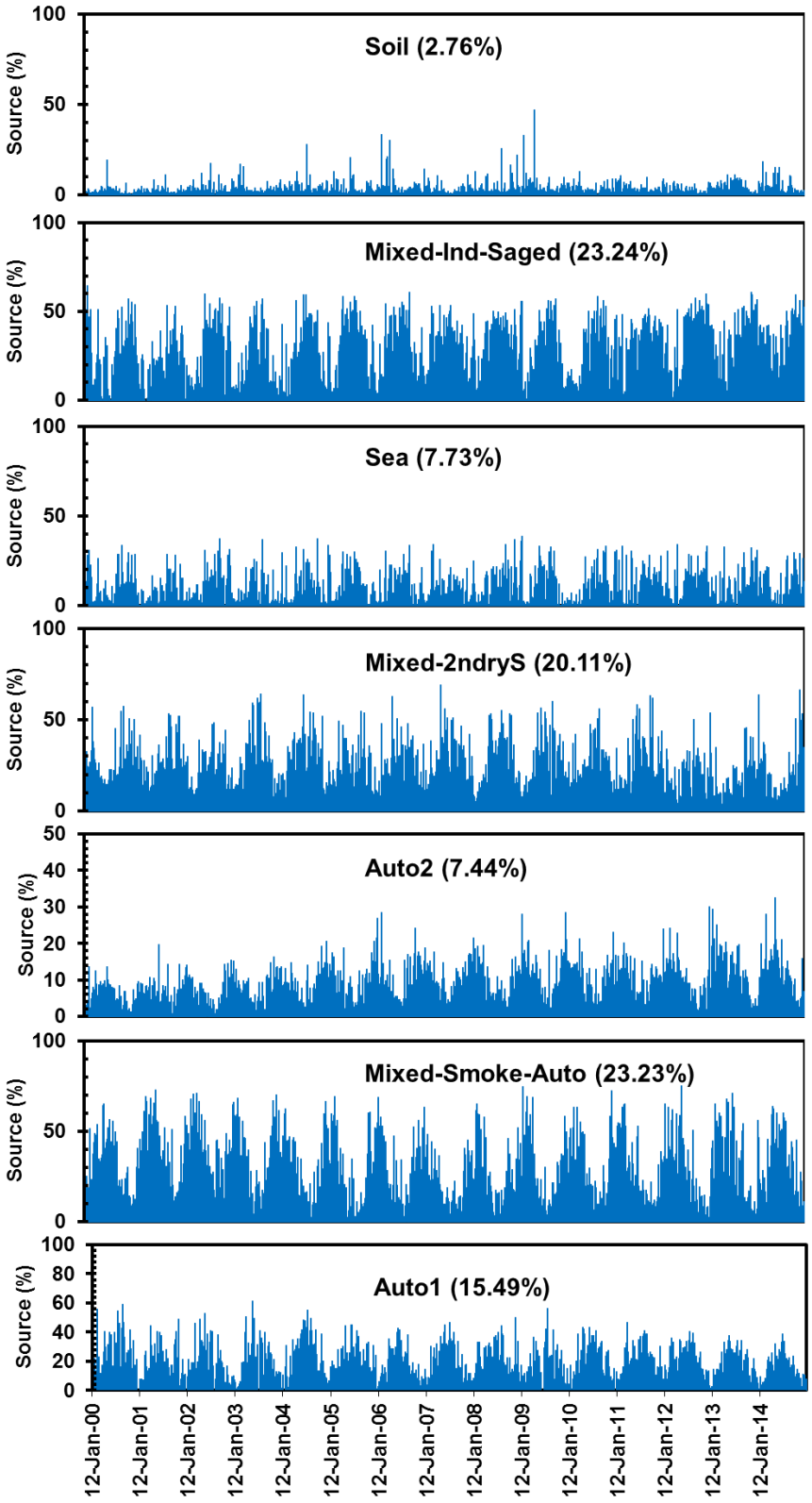

Figure 12-6. Mascot percentage PMF source contributions, 2000–14.

Liverpool – daily variations

Figure 12-7 shows the daily source contributions during the 15-year study period for the seven source fingerprints at the Liverpool site. The bottom plot shows the daily time series for the total $PM_{2.5}$ mass for comparison purposes. The horizontal axis is the same for all plots and uses the date format 20040711 to represent 11 July 2004. Note also that the vertical axis is in (ngm^{-3}) with some plots having different scales so daily variations are more obvious.

Clear summer–winter seasonal variations can be seen in most fingerprints except for the soil fingerprint. Averaging or looking at the median values for these seasonal variations over a single month or year makes for a more meaningful time series comparison and we have done this in the next section.

The mixed secondary sulfate fingerprint contributed between 40% and 60% of the total $PM_{2.5}$ mass in the summer months at Liverpool. The annual median value dropping from 1648ngm^{-3} to 870ngm^{-3} when comparing the year 2000 and 2014 respectively, a 47% fall over the 15-year sampling period.

The mixed industrial aged sulfur fingerprint at Liverpool has a strong seasonal variation, being higher in the summer months. It was responsible for 20–30% of the total fine $PM_{2.5}$ mass in the summer months. It varied significantly from year to year however over the 15-year study period the annual median value showed a slight increased from 703ngm^{-3} in 2000 to 875ngm^{-3} in 2014, a 24% increase.

Again we add the two Auto fingerprints together and see a slight but significant drop in the annual median value, dropping from 1324ngm^{-3} to 1059ngm^{-3} (20%) between 2000 and 2014. The annual median $PM_{2.5}$ mass dropped from 7.790μgm^{-3} in 2000 to 6.866μgm^{-3} in 2014 or 12%. So the percentage drop in the Auto1+2 fingerprint from 2000 to 2014 was essentially 18% at Liverpool. This fingerprint contributed between 30% and 40% to the total $PM_{2.5}$ mass in the winter months at this site.

The mixed smoke automobile fingerprint at Liverpool is comparable with but larger than the same fingerprint at the Richmond site. This fingerprint represents between 70% and 80% of the fine $PM_{2.5}$ mass in the winter months, falling to between 10% and 20% in the summer months. It showed a significant drop from 2000 to 2014, the annual median value dropping from 2865ngm^{-3} to 1226ngm^{-3} respectively, a 57% fall over the 15-year sampling period. As mentioned previously this smoke-automobile fingerprint at Liverpool is probably more influenced by smoke from domestic wood heating in the winter months than by smoke from diesel vehicles as with the Mascot site. The significant fall in this fingerprint with time is probably due to the increasing winter temperatures in western Sydney over the past decade and hence the reduced need for domestic wood heaters.

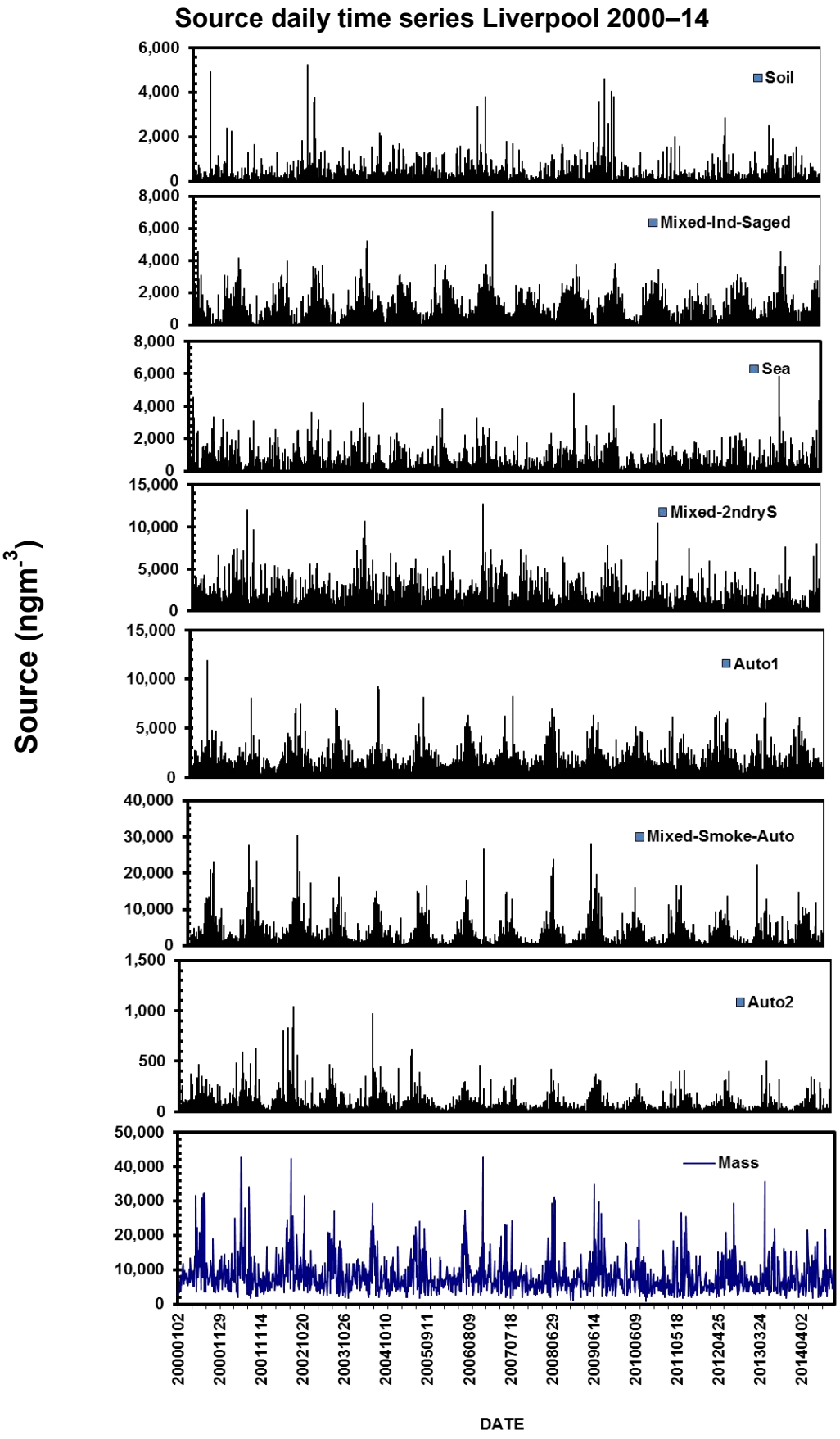

Figure 12-7. Liverpool PMF seven factor source contributions to total mass, 2000–14.

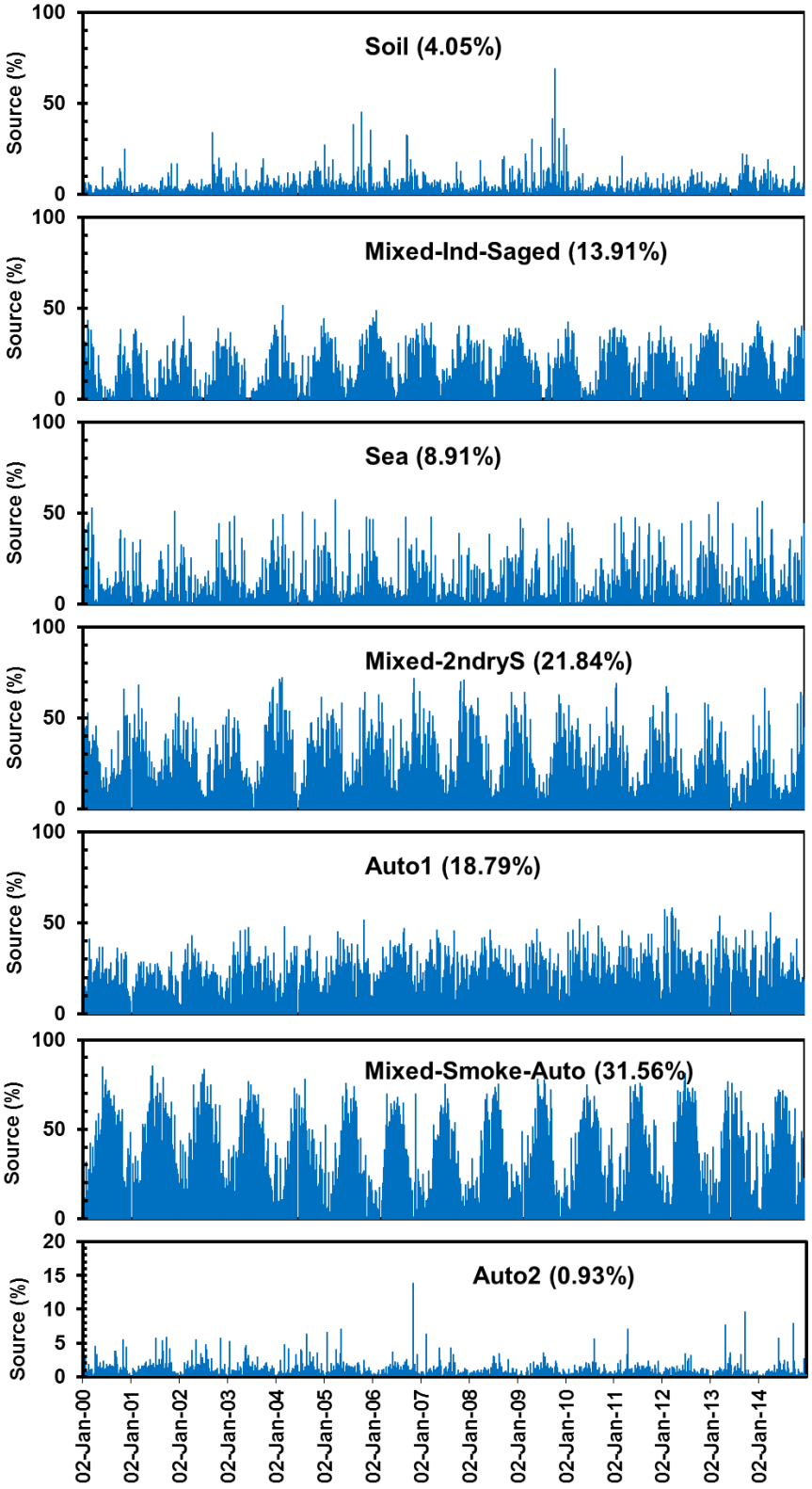

Figure 12-8. Liverpool percentage PMF source contributions, 2000–14.

Domestic wood burning

Domestic wood heaters have a significant effect on ambient fine particle air quality particular in western parts of Sydney in the winter months. An obvious feature of the time series plots shown in **Figure 12-4** and **Figure 12-8** is the increased Mixed-Smoke-Auto fingerprint contributions for Richmond and Liverpool compared with the Lucas Heights and Mascot sites. **Figure 12-9** plots the daily time series for Mixed-Smoke-Auto factors for Liverpool and Mascot from 2000–14 on the same graph. The Liverpool factor is consistently higher than the Mascot factor and the seasonal trends follow each other closely.

Figure 12-9 has two distinct features. Firstly, the seasonal, summer–winter troughs and peaks extending over several weeks each year, with the winter peaks clearly visible against the summer minima.

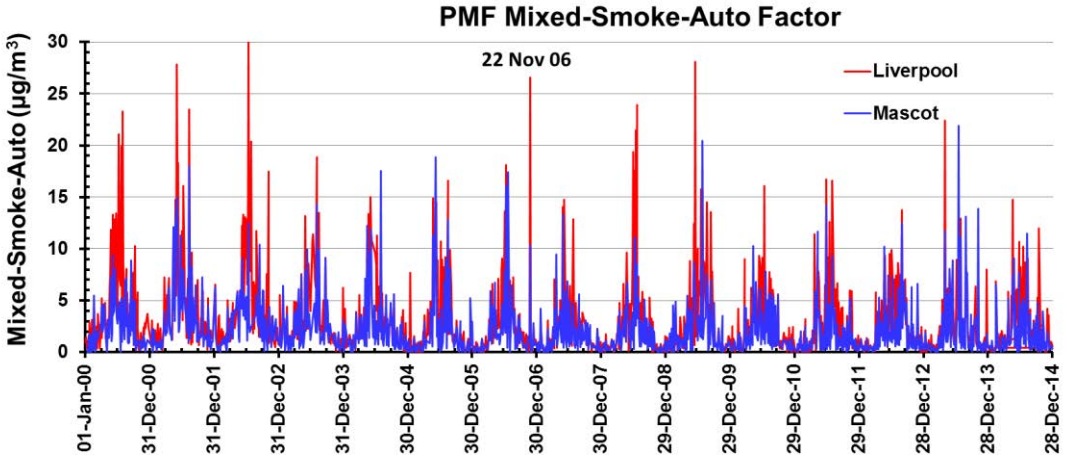

Figure 12-9. Daily time series plot of the Mixed-Smoke-Auto factors for Liverpool and Mascot, 2000–14.

These wintertime peaks (not the spikes) reach 60–80% of the total measured gravimetric mass. In the summertime they fall to less than 10% of the total $PM_{2.5}$ mass as shown in **Figure 12-10** where the Mixed-Smoke-Auto factor is plotted as a percentage of the total $PM_{2.5}$ gravimetric mass at Liverpool. Clearly this factor is a major component of the ambient $PM_{2.5}$ mass in the winter months at Liverpool. Similar data are obtained for the Richmond site, showing that western Sydney is impacted significantly by smoke in the winter months.

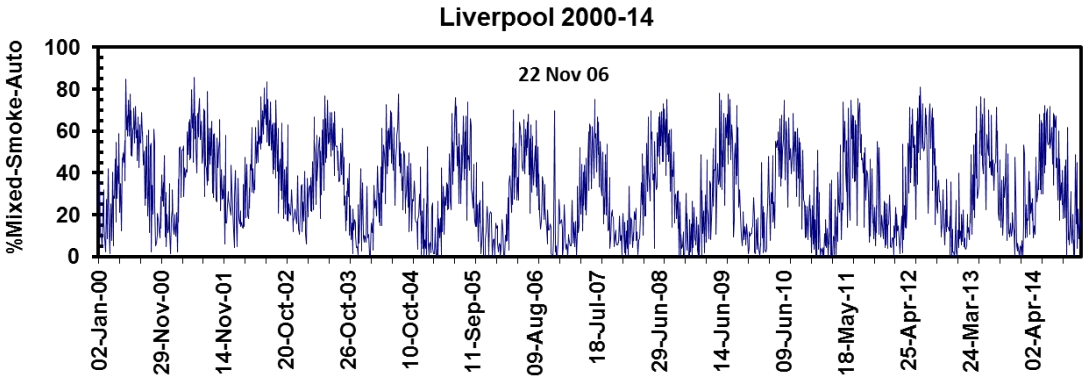

Figure 12-10. Plot of the percentage Mixed-Smoke-Auto factor for Liverpool, 2000–14.

Secondly, clear short-term daily spikes reaching several tens of (µgm^{-3}) correspond to short-term controlled burning events or major bushfire events in and around the greater Sydney metropolitan region. The spike labelled 22 November 2006 in **Figure 12-11** is a clear example of a bushfire event, with Liverpool showing 27µgm^{-3} and Mascot 10µgm^{-3} of Mixed–Smoke-Auto on this day. These daily events are clearly distinguishable in the plot in **Figure 12-11** from the wintertime domestic burning events that extend over several weeks.

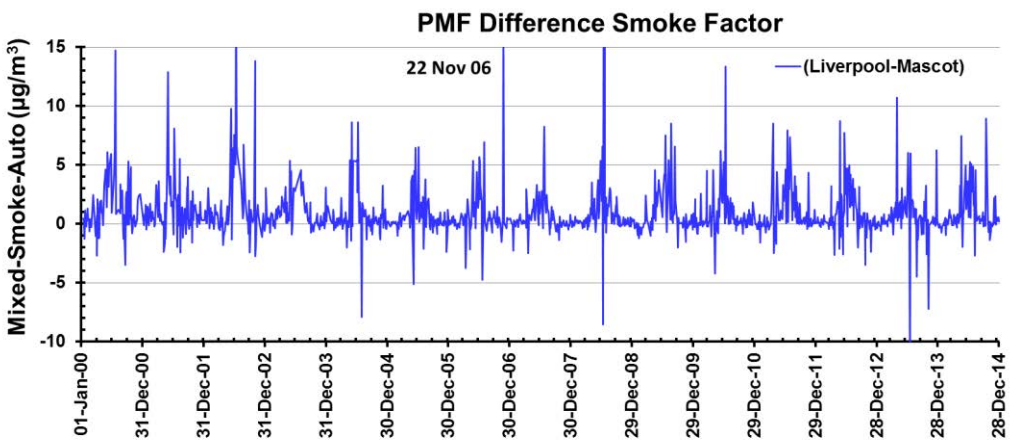

Figure 12-11. A plot of the daily difference between the Liverpool and Mascot Mixed-Smoke-Auto factors, 2000–14.

We chose Liverpool and Mascot as we expect they are impacted by similar types of motor vehicles but have different domestic wood fire contributions over any given year. If we subtract the Mascot Mixed-Smoke-Auto factor from the Liverpool factor and look at the seasonal variations (weekly not daily variations) we would obtain an estimate of domestic wood burning contributions in the Liverpool area. **Figure 12-11** is such a plot. Ignoring daily spikes, we see that the wintertime seasonal peaks in this difference plot average around 5µgm^{-3} for several weeks at a time each year over the 15-year sampling period. This clearly shows the effect of domestic wood burning on the western parts of Sydney in the winter months. For the summer periods we see that the difference plots show Liverpool and Mascot are essentially the same. The negative spikes show days where Mascot was more impacted by controlled burning or bushfire events than Liverpool.

Lucas Heights – PMF monthly, yearly variations

In this section we plot the same daily data presented above but as box and whisker plots for each month of the year and for each of the 15 individual years in this study period of 2000–14. The left hand side of each figure is by month and the right hand side by year.

A box and whisker plot is used here because it contains a wealth of information about the distribution of the data points plotted. The (+) sign represents the average value, the horizontal line in the box represents the median value and the vertical box size represents 25% to 75% of all data points. The vertical whiskers represent the 95% confidence intervals and the dots are the extreme outlier points lying beyond the whiskers.

Figure 12-12 shows the box and whisker plots for the Lucas Heights site by month and by year for each of the PMF source fingerprint concentrations (in µgm^{-3}). Seasonal variations are clear in the monthly plots with sulfate based fingerprints high in the summer months and Auto fingerprints high in the winter months.

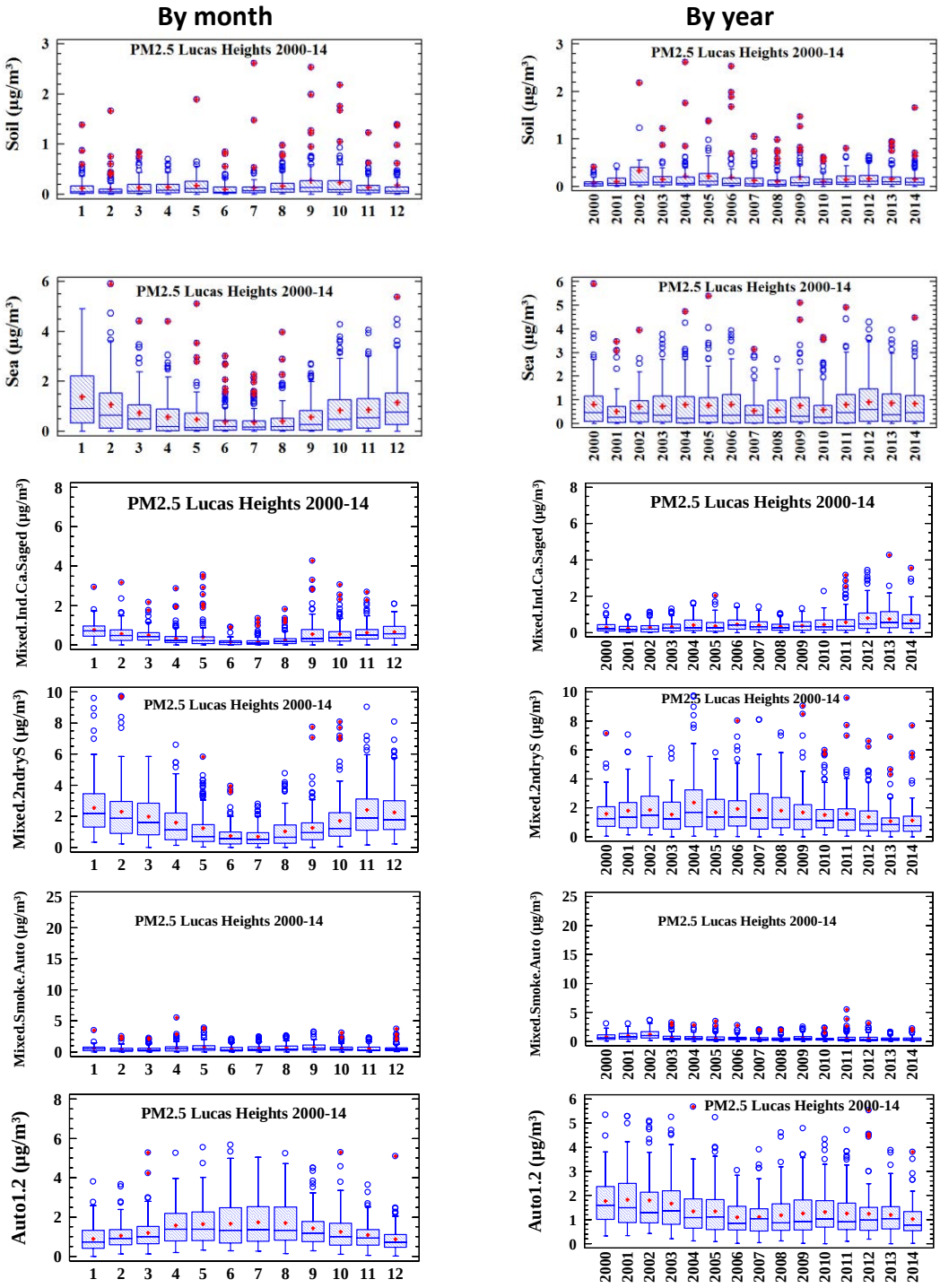

Figure 12-12. Box and whisker plots of source fingerprint monthly and yearly variations for the Lucas Heights site, 2000–14.

Richmond – PMF monthly, yearly variations

Figure 12-13 shows the box and whisker plots for the Richmond site by month and by year for each of the PMF source fingerprint concentrations (in µgm^{-3}).

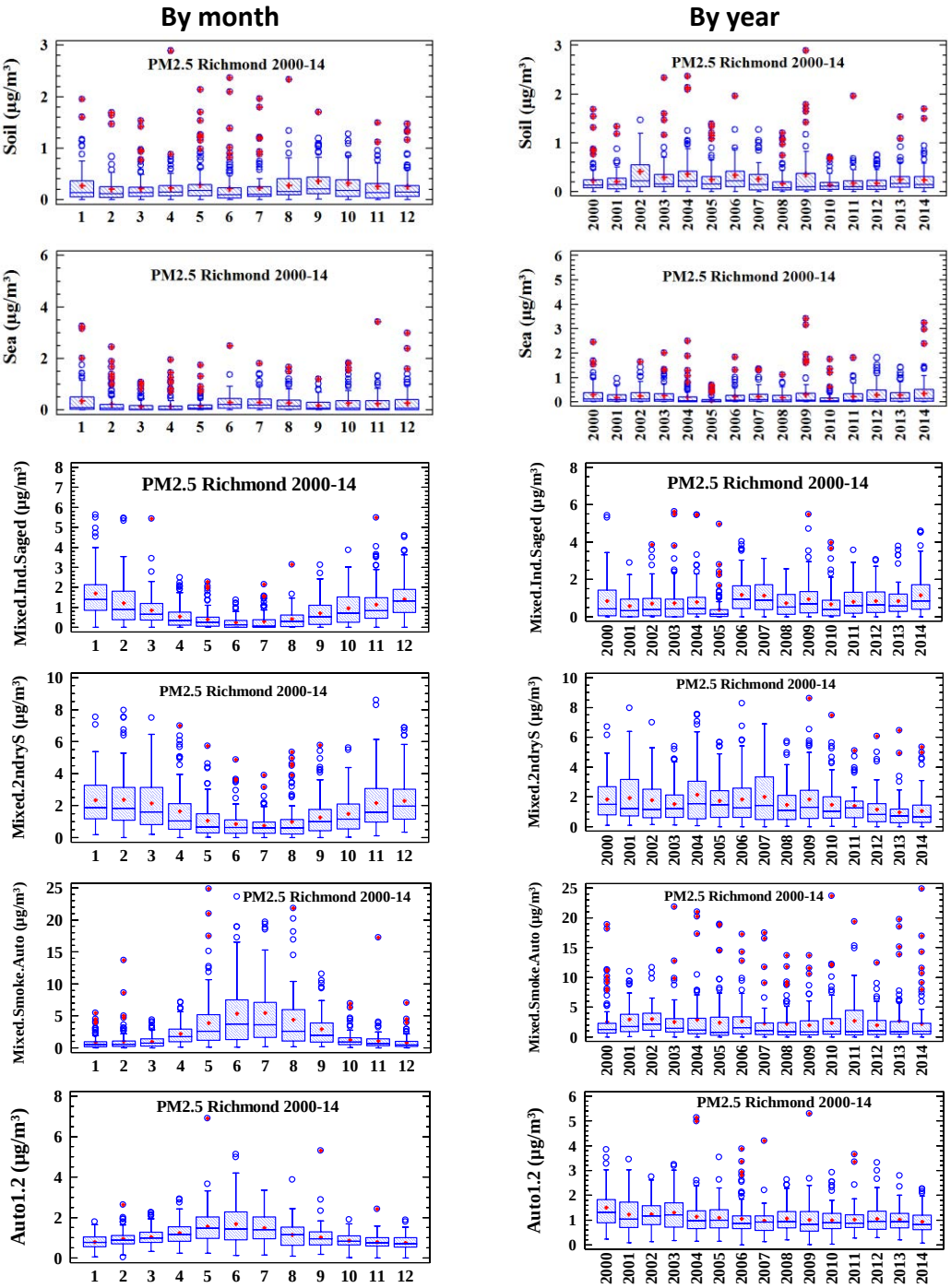

Figure 12-13. Box and whisker plots of source fingerprint monthly and yearly variations for the Richmond site, 2000–14.

Seasonal variations are clear in the monthly plots with sulfate based fingerprints high in the summer months and Smoke and Auto fingerprints high in the winter months.

Mascot – PMF monthly, yearly variations

Figure 12-14 shows the box and whisker plots for the Mascot site by month and by year for each of the PMF source fingerprint concentrations (in µgm^{-3}).

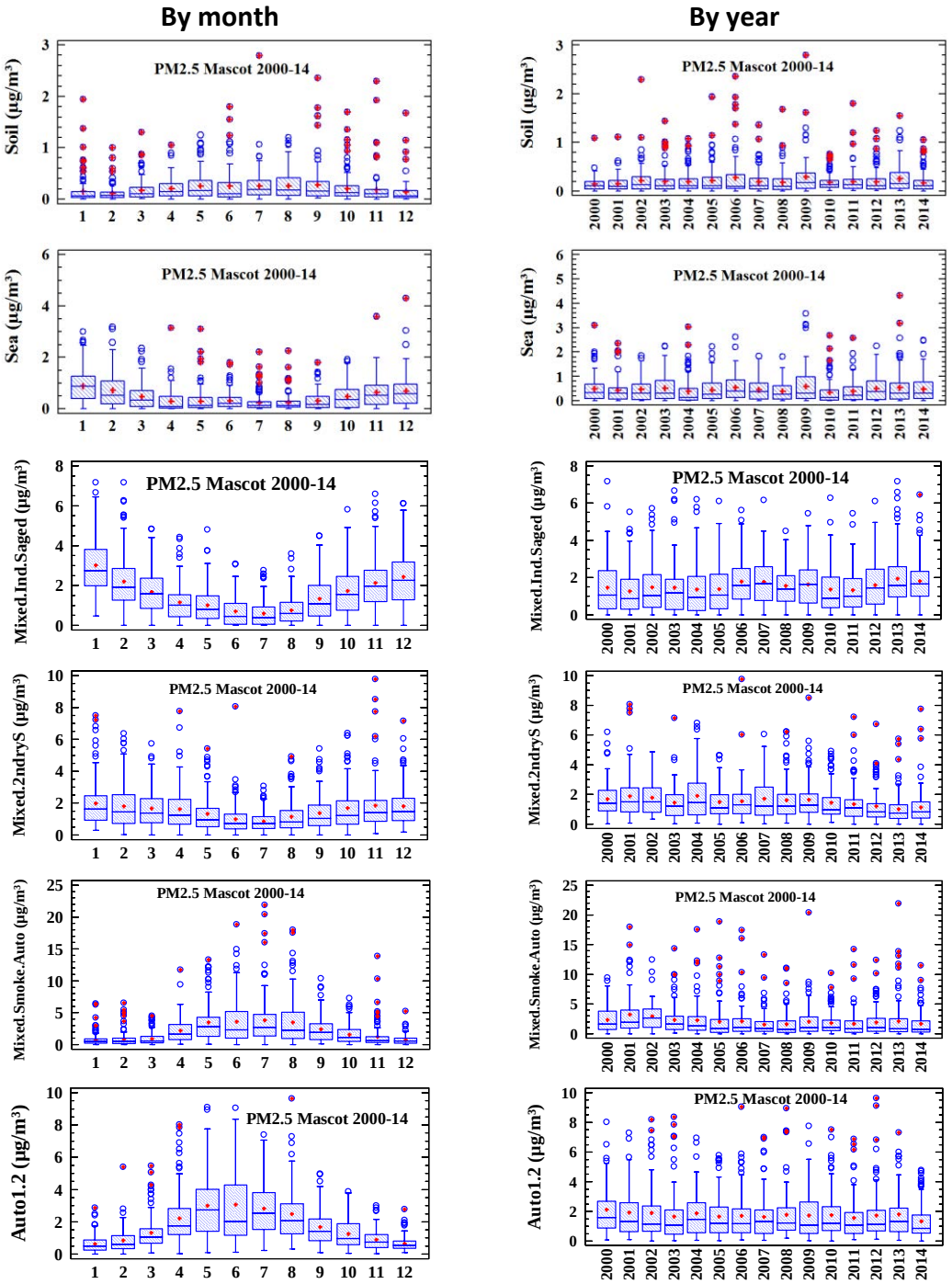

Figure 12-14. Box and whisker plots of source fingerprint monthly and yearly variations for the Mascot site, 2000–14.

Seasonal variations are clear in the monthly plots with sulfate based fingerprints high in the summer months and Smoke and Auto fingerprints high in the winter months.

As discussed previously the annual Auto plot shows a 46% decrease in concentration from 2000 to 2014 at Mascot.

Liverpool – PMF monthly, yearly variations

Figure 12-15 shows the box and whisker plots for the Liverpool site by month and by year for each of the PMF source fingerprint concentrations (in µgm^{-3}). Seasonal variations are clear in the monthly plots with sulfate based fingerprints high in the summer months and Smoke and Auto fingerprints high in the winter months.

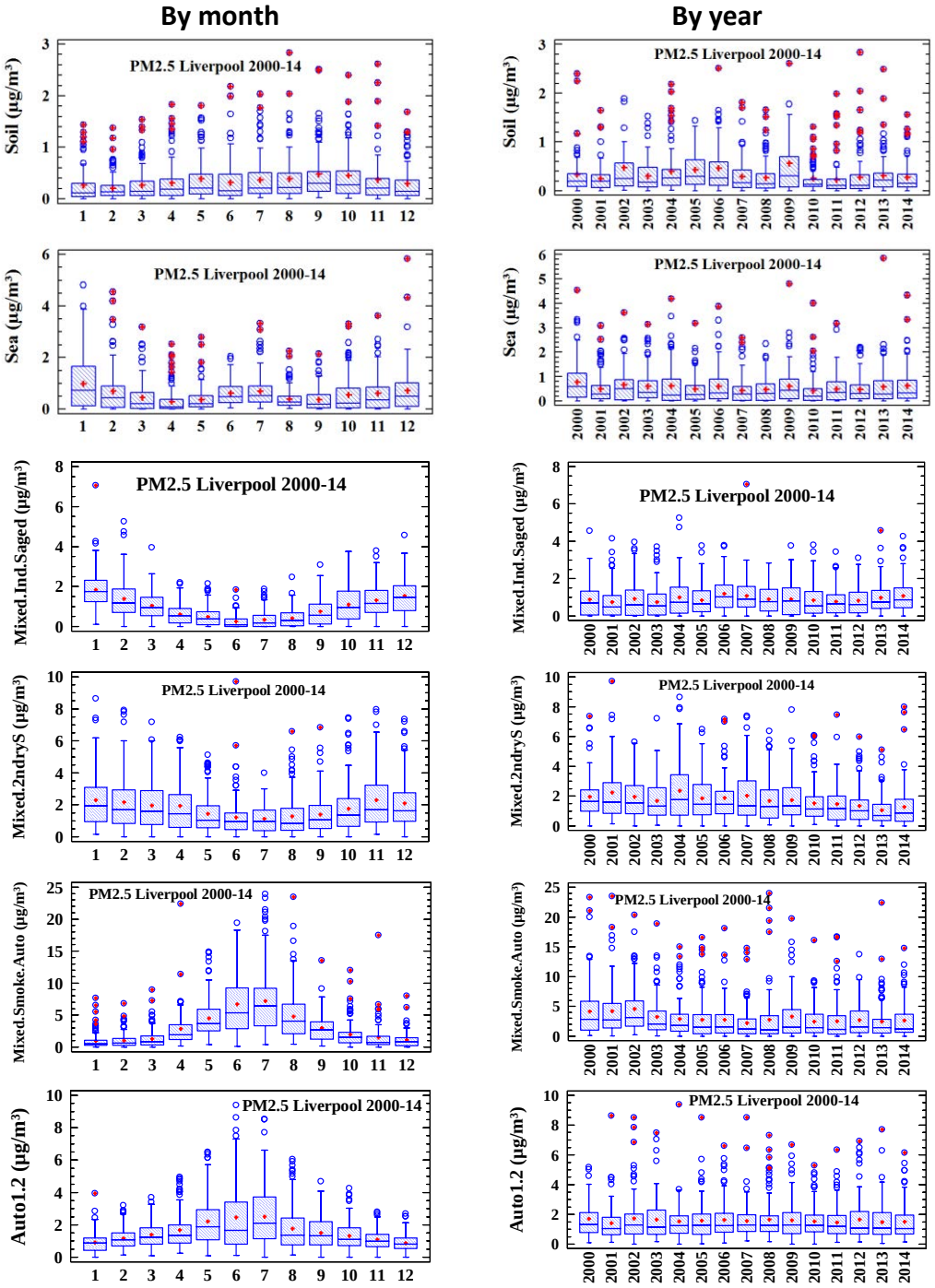

Figure 12-15. Box and whisker plots of source fingerprint monthly and yearly variations for the Liverpool site, 2000–14.

The median mixed smoke-auto fingerprint shows a 57% decrease from 2000 to 2014 at Liverpool, while the mixed secondary sulfate decreases by 47% over the same study period. For the Liverpool site the combined Auto1+2 fingerprint drops by only 20% between 2000 and 2014.

Automobiles

The PMF seven fingerprint fits to the total $PM_{2.5}$ measured mass all gave two distinct automobile fingerprints: Auto1 and Auto2. These consistently occurred in approximately the same proportions whether 6, 7 or 8 source factors were used in the PMF fits. The Auto1 fingerprint was always much higher in concentration than the Auto2 fingerprint. At Mascot where the number of motor vehicles was higher than at the other sites both Auto fingerprints were driven by between 15% and 20% of the total measured lead (Pb). Leaded petrol was used in NSW up until January 2001 so we believe the Auto2 fingerprint being much smaller than the Auto1 fingerprint is probably directly related to the use of leaded petrol in motor vehicles, whereas the lead associated with the Auto1 fingerprint is probably part of the retrained road dust kicked up by motor vehicle movements. It is convenient to add these two fingerprints together to produce the Auto1+2 fingerprint as the Auto2 contributions are generally small and the seasonal variations are similar.

The number of registered motor vehicles in NSW between 2001 and 2014 is provided in **Table 12-1**. It shows an increase of 40% in the number of registered vehicles between 2001 and 2014.

Table 12-1. The number of registered motor vehicles in NSW, 2001–14. Sourced from the Australian Bureau of Statistics (ABS).

Year	2001	2005	2010	2014
Number of registered motor vehicles in NSW	3.745M	4.170M	4.681M	5.247M
% increase over 2001		11.3%	25.0%	40.1%

In **Figure 12-16** we show the annual median Auto1+2 fingerprint concentrations for the four sites for each of the sampling years between 2000 and 2014. The average across all four sites for the annual median Auto1+2 fingerprint concentration fell from 1.45µgm^{-3} in 2000 to 0.875µgm^{-3} in 2014. This was a decrease of 40% even though the average number of registered motor vehicles in NSW rose by more than 40% over this same time period. This can only mean the motor vehicles in NSW have become cleaner and less polluting over the past 15 years.

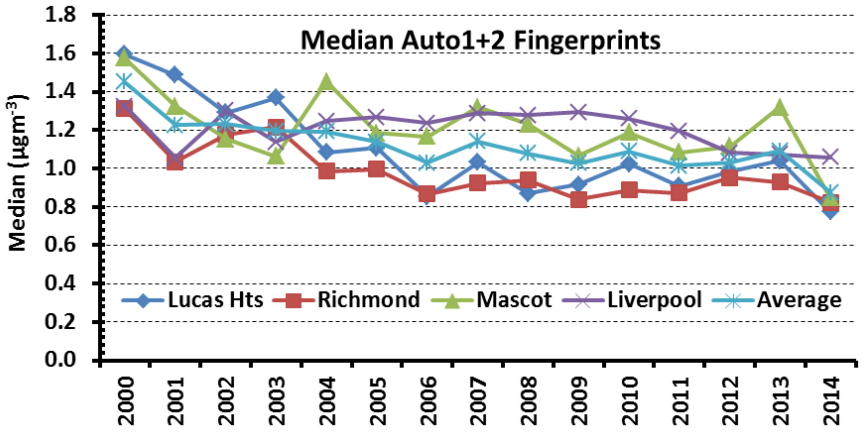

Figure 12-16. The annual median Auto1+2 fingerprint concentration by site, 2000–14.

Modern cars may be cleaner than their older counterparts but the motor vehicle fleet in NSW is still increasing at around 3% per annum. **Figure 12-17** is a box and whisker plot showing the percentage contribution of the Auto1+2 fingerprint to the total $PM_{2.5}$ mass for each of the four sampling sites. Over the 15-year study period Liverpool, Lucas Heights, Mascot and Richmond each contributed 10–25%, 15–40%, 10–30% and 10–25%, respectively.

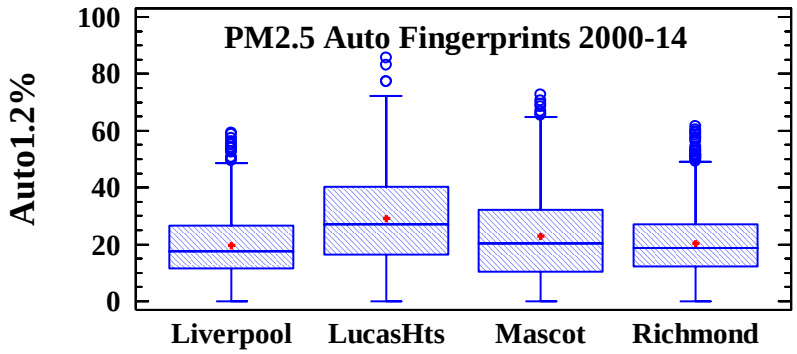

Figure 12-17. The percentage contribution of the Auto1+2 fingerprint to the total $PM_{2.5}$ mass for each of the four sampling sites, 2000–14.

It should be pointed out these Auto1+2 percentages are probably an underestimate of the total contributions from motor vehicles since there are further automobile components in the Mixed-Smoke-Auto and the Mixed-2ndryS fingerprints which also have contributions from motor vehicle emissions. Automobiles are still a significant source of ambient air pollution in the Sydney region. Light, medium and heavy diesel motor vehicles are major contributors to this type of fine particle air pollution.

Secondary sulfates

Analysis showed that, in Sydney, on average about 22% of the $PM_{2.5}$ mass is ammonium sulfate. The PMF analysis across all the Sydney sites showed that the total sulfur (S) mass was mostly associated with two of the seven fingerprints: Mixed-2ndryS (~80% of total S) and Mixed-Ind-Saged (~20% of total S). Secondary sulfate ions originate from the conversion of sulfur dioxide gas in the presence of water vapour and sunlight. Typical conversion rates are about 1–2% per hour in the atmosphere. These ions are neutralised to ammonium sulfate if enough ammonium is present.

Figure 12-18 shows the percentage contribution of the Mixed-2ndryS and the Mixed-Ind-Saged fingerprints to the total $PM_{2.5}$ mass for each of the four sampling sites for the study period of 2000–14.

The contributions to the Mixed-2ndryS fingerprints are between 10% and 40% across all sites but are higher at the inland sites of Liverpool, Lucas Heights and Richmond than at the coastal Mascot site, while for the Mixed-Ind-Saged fingerprints contributions lie between 5% and 40% across all sites but are higher at Mascot, which is impacted more by industry around the international airport and Port Botany shipping container depot. Significant secondary sulfate sources in the study area include coal fired power stations, oil refineries at Rosehill and Kurnell, metals industry in Port Kembla to the south of Sydney, shipping in ports, diesel vehicles and multiple light to medium industries.

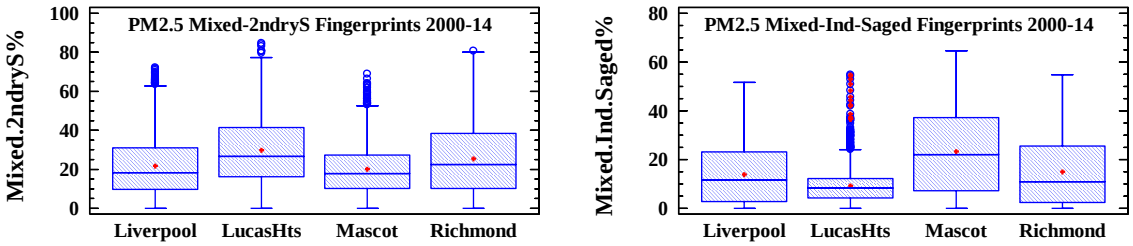

Figure 12-18. The percentage contribution of the Mixed-2ndryS and the Mixed-Ind-Saged fingerprints to the total $PM_{2.5}$ mass for each of the four sampling sites, 2000–14.

The NSW EPA air emission inventory for 2008 showed that 98% of the total 246kT of sulfur dioxide emissions in NSW are associated with the coal fired power stations to the north and west of Sydney, burning 25 million tonnes of coal annually. This coal typically contains at least 0.5% sulfur by weight. We have shown in previous publications (Cohen et al. 2014 and Appendix C) this secondary sulfate from coal combustion is transported into the Sydney basin to sites like Liverpool and Richmond, so we expect much of this Mixed-2ndryS and part of the Mixed-Ind-Saged to be associated with coal burning for power generation. Key signature elements for coal burning are sulfur (S), arsenic (As), not measured here, and selenium (Se). We see around 20% of the total measured selenium occurs in the Mixed-2ndryS at each of the sampling sites, further evidence that coal burning is a significant contributor to secondary sulfates in the Sydney basin.

13. 15-year site comparisons

Figure 13-1 shows the box and whisker plots for each of the seven common fingerprint average concentrations for the 15-year study period at each of the sites, so inter-site comparisons can be made for a given fingerprint.

For most fingerprints the distributions at each site have similar concentrations. The obvious differences being:

- the reduced sea spray at Richmond, which is not unexpected for this inland site
- the increased mixed industry and aged sulfur concentrations at the Mascot site
- the reduced levels of mixed smoke-auto at the Lucas Heights site, due to the site location being on top of the escarpment and less affected by winds in the Sydney basin itself
- the increase Auto2 and hence total automobile contributions at the Mascot site, again totally expected because of the inner urban nature of the site and its proximity to heavy vehicle movements around the international airport and Port Botany.

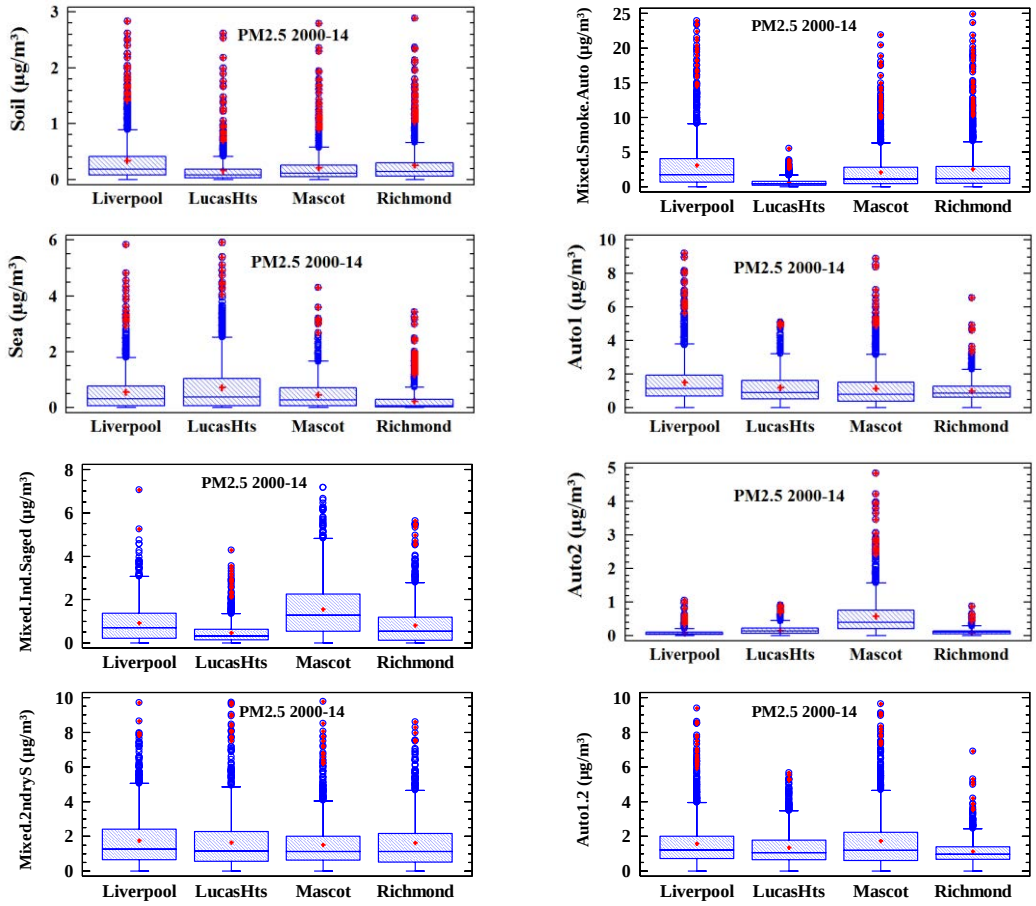

Figure 13-1. Average site fingerprint comparisons for each of the seven fingerprints, 2000–14.

14. Five-year PMF fingerprint analysis

To investigate if there are any long-term trends over the 15-year study period from 2000 to 2014 inclusive we have split the period into three five-year blocks and redone the PMF analysis just for the Liverpool site with the same number of fingerprints. The results are summarised below in tabular form and the full PMF data analyses can be accessed through out PMF database macro discussed in detail in Appendix B.

It should be pointed out that doing shorter timeframes like this reduces the total dataset by a factor of three, which is generally not desirable for effective PMF analysis. You can produce different elemental fingerprints as it can change the inter-element correlations. Furthermore the PMF analysis relies on the variability in multi-dimensional space to determine these fingerprint correlations and thereby separate out possible source fingerprints. Reducing the size of the dataset reduces the variance across the dataset and hence makes this separation more difficult. The more summer–winter cycles and outlier or extreme events you have the better PMF analysis will work. So we do expect some differences between the five-year block analysis and the preferred single 15-year block analysis. Nevertheless it can be a useful exercise to investigate any long-term trends that may occur and look for differences between this method and the full daily analysis of the 15-year dataset.

Liverpool

Table 14-1 summarises the average percentage fingerprint contributions to the total fine mass for the Liverpool site. The PMF analysis was done in three five-year blocks using seven fingerprints and similar techniques to the full 15-year analysis described above.

Table 14-1. The average percentage fingerprint contributions to the total $PM_{2.5}$ mass for the Liverpool site analysed in five-year blocks over the 15-year study period.

Liverpool $PM_{2.5}$	2000–04	2005–09	2010–14	2000–14
Soil%	3.73±0.57	4.09±0.37	2.41±0.50	4.05±0.25
Sea%	10.6±0.89	5.92±0.62	8.79±0.86	8.91±0.47
Mixed-Ind-Saged%	10.6±1.0	19.1±0.87	16.2±1.1	13.9±0.59
Mixed-2ndryS%	23.8±0.75	19.9±0.65	21.7±0.61	21.8±0.40
Mixed-Smoke-Auto%	27.2±0.66	27.0±0.42	35.1±0.54	31.6±0.29
Auto1%	17.9±1.2	19.2±0.58	13.7±0.75	18.8±0.51
Auto2%	6.19±1.1	4.87±0.45	2.07±0.55	0.93±0.34
$PM_{2.5}$ mass (μgm^{-3})	8.86±5.6	7.99±5.1	7.04±4.1	8.17±5.4
Days	477	468	471	1453
χ^2	0.85	0.77	0.89	0.97

The soil fingerprint, which contributes only a few percent to the total fine mass, may have reduced a little in 2010–14 compared with 2000–04. This fingerprint is entirely driven by meteorological conditions and how many significant dust storms impact the sampling site in a given year, so differences of the size shown in **Table 14-1** are not unexpected.

The sea spray fingerprint probably has not changed significantly over the 15-year sampling period; again the amount of sea spray reaching the inland Liverpool site is dependent on wind conditions and the number of coastal storms we have during the sampling period that will transport sea spray and other coastal pollution inland to Liverpool.

The mixed industrial aged sulfur fingerprint is primarily an industrial source. It appears to be quite variable during the study, changing by 30–40% over the 5 year period. This could be due to meteorological conditions such as long dry spells, lack of sea breezes changing atmospheric chemistry or different wind conditions impacting the Liverpool site over the time period as well as variations in this source itself. We have also linked parts of this source to possible coal fire power station and heavy oil burning emissions which if true we do not expect to vary much over this time scale.

The mixed secondary sulfate fingerprint has varied little over the 15-year study period, sitting between 19% and 24%. Secondary sulfates originate from coal fired power stations, motor vehicles (particularly diesel lorries and buses) and heavy oil combustion. It is driven by the conversion of sulfur dioxide gas emissions to sulfate particles. With coal fired power stations emitting 98% of the sulfur dioxide in NSW we therefore expect this fingerprint to be mainly representative of power station emissions.

The mixed smoke and automobile fingerprint is consistently high between 27% and 32% but probably does not show any real increase over time. As the major contributions come from smoke from domestic wood burning in the winter months at this site the absolute value of this fingerprint is strongly dependent on the temperatures during winter months and how much domestic heating is required.

The sum of the two Auto fingerprints was 24%, 24% and 16% of the total fine mass for each of the three five-year blocks respectively. As the $PM_{2.5}$ mass consistently fell 10% between 2000–04 and 2005–09 and 11% between 2005–09 and 2010–14 this means in absolute terms

the contribution from automobiles fell between 2000 and 2014 by 48% [from (1.28±1.3)µgm^{-3} to (0.723±0.67)µgm^{-3}]. Since the number of automobiles on the road clearly increased during this period we can only conclude that the motor vehicle fleet is becoming cleaner and less polluting with time.

Auto fingerprint differences over 15 years

Despite the number of registered motor vehicles in NSW increasing by 40% over 15 years, their contribution to the PMF fine particle Auto fingerprint pollution has decreased by 48%. By doing the PMF analysis in five-year blocks we can look at the differences between Auto1 fingerprints across each five-year period between 2000 and 2014.

Figure 14-1 is such a plot for the Auto1 fingerprint for the Liverpool site for 2000–04, 2005–09 and 2010–2014. The fingerprints are similar for each of the five-year blocks, the obvious differences being the increased Br and Pb in the 2000–04 block compared with the 2010–14 block, reflecting the use of leaded petrol in NSW up until 2001. Also, the total hydrogen (H), which is a signature for organics, is 65% higher in the 2000–04 block than the other two later five-year blocks, again demonstrating that the automobile emissions are getting cleaner with time.

The technique of analysing data over five years is basically no different to taking five-year subsets of the daily data analysed over 15 years and discussed earlier under the *Time series analysis* section, so there is no need to performed similar analyses for the other three sites.

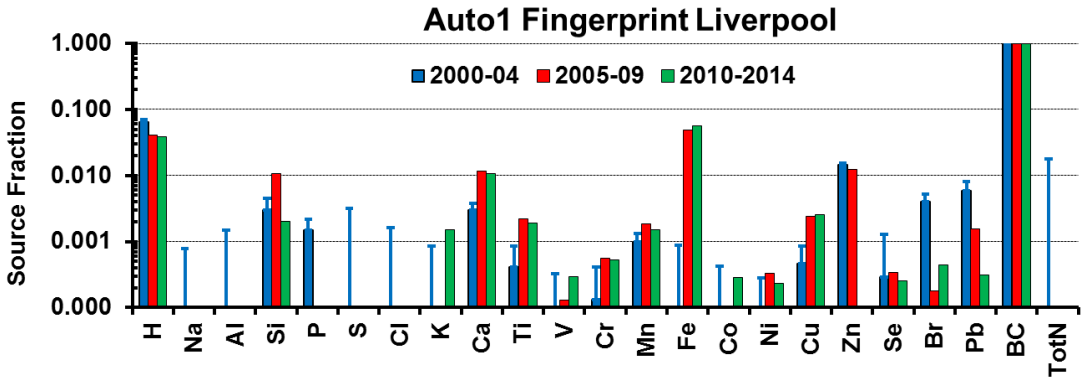

Figure 14-1. Plot of the Auto1 fingerprint for the Liverpool site for 2000–04, 2005–09 and 2010–14.

The detailed fingerprints and their contributions to the total fine mass for the five-year blocks analysed over the 15-year study period for the Liverpool site are included in the PMF dataset described in Appendix B and are provided as part of this report.

15. Summary

This study has used the existing ANSTO database covering daily $PM_{2.5}$ sampling at four sites in the greater Sydney region during the period 1 January 2000 to 31 December 2014. The database consists of over 6000 sampling days from these sites. Each daily filter was weighed and analysed for 23 different elemental and chemical species using the ion beam analysis (IBA) techniques at ANSTO. The results of these analyses have been put through a PMF source

apportionment code to determine seven different source fingerprints and their contributions to the total measured $PM_{2.5}$ mass at each site for each sampling day.

All the aims of this study have been met, namely:

- to convert the existing 15-year $PM_{2.5}$ mass and elemental datasets for four given sites in the Sydney basin into identifiable source fingerprints
- to quantify the absolute and the percentage contribution of each of these fingerprints to the total fine $PM_{2.5}$ mass
- to provide seasonal and annual variations for each of the source fingerprints
- to provide a readily accessible database containing the daily source fingerprints and their contributions covering the 15-year period from 2000 to 2014 for four given sites in the Sydney basin
- if possible, to identify and quantify the major contributors of fine particle pollution to the ambient air quality in Sydney.

In summary, this study identified four main pollution sources or source factors in the Sydney airshed between 2000 and 2014 which significantly contributed to the $PM_{2.5}$ mass affecting all four sites. These were:

- smoke from domestic wood heaters which contribute significantly to fine particle levels particularly in western Sydney in winter months, peaking between 60% and 80% in the wintertime at some sites
- secondary sulfates from the burning of large amounts of coal for power generation and fine particles from industry and motor vehicles mainly generated by light, medium and heavy diesel vehicles, peaking between 50% and 70% in the summer months at some sites
- significant contributions to the $PM_{2.5}$ mass loading from industrial sources represented by the Mixed–Ind-Saged source fingerprints at all of the four sites studied during the 15 years from 2000 to 2014, peaking between 30% and 50% in the summer months at some sites
- vehicle emissions which contribute to the Auto1, Auto2 and Mixed-Smoke-Auto factors; reductions in vehicle contributions occurred during the period despite increases in the vehicle fleet, but such reductions seem to have tailed off more recently with Auto1+2 contributions of about $1\mu gm^{-3}$ or between 10% and 15% of the annual $PM_{2.5}$ mass in 2014.

The approach adopted was focused on identifying sources which contribute significantly to daily, seasonal and annual average concentrations. During the analysis, smoke from biomass burning was identified as a significant intermittent source contributing to short-term spikes in fine particle ambient mass concentrations.

A database of each of seven source fingerprints for each sampling day for each of the four sites has been produced and has been made readily available. An instruction manual to do this has been reproduced in full in Appendix B. All the data in this database is fully accessible and extractable by the user.

16. Recommendations

For expanding and making future similar fine particle source apportionment studies more effective and better able to resolve and identify individual source types we make the following recommendations:

- The number of elemental and chemical species should be expanded from the basic 23 currently used. The present study did not directly measure magnesium (Mg), arsenic (As), ammonium ions, nitrate ions, elemental carbon (EC) or organic carbon (OC), which are key signatures for several sources.
- The measurement of chemical species like levoglucosan and mannosan tracers for biomass burning, methanosulfonate (MSA^-) and oxalate ($C_2O_4^{2-}$) ions would help to better differentiate biomass burning, fossil fuel combustion, automobile and individual industrial sources. The measurement of these species together with the missing species mentioned above could bring the total elemental and chemical species list to over 35, making PMF analysis much more effective in identifying and quantifying more source types.
- The combination of PMF sources with wind and back trajectory modelling for every hour of every sampling day over long timeframes and tied to major known point source locations (as done in a recent paper by Cohen et al. 2012 and 2014, Appendix C), should be undertaken to identify and quantify $PM_{2.5}$ windblown soil and secondary sulfate sources at other existing long-term sampling sites in NSW. Chemical transport modelling could also be used to predict these sources. The PMF particle characterisation data provided by this study should be used in the development and validation of chemical transport models which can then be used to provide a refined projection of source contributions to both primary and secondary particles.

It should be pointed out that the IBA analysis techniques used at ANSTO are non-destructive, whereas the extra analyses suggested above such as ion chromatography (IC) and (EC/OC) analyses destroy the filters.

Also, historical filters from this study are not available for any further destructive analysis and therefore the recommendations above relate to the collection of future samples.

17. Acknowledgements

The authors would like to acknowledge funding from the NSW Environment Protection Authority (NSW EPA) to carry out this study. We would also like to acknowledge Botany Council for financial support of the Mascot site, staff at the University of Western Sydney for support in operating the Richmond site, and the NSW EPA and the NSW Office of Environment and Heritage for support with the Liverpool site throughout this study. We would also like to thank the ANSTO accelerator support staff for running the accelerators and the ASP aerosol labs throughout this study.

18. References

Chan YC, Simpson R, Cohen D, Hawas O, Stelcer E, Denison L, Wong N, Golding G, Christensen E, Gore W, Hodge M, Comino E and Carswell S 2005, A preliminary analysis of sources of emissions of particles in four major Australian cities by Positive Matrix Factorisation, in *Proceedings of The 17th International Clean Air and Environment Conference, Hobart 3–6 May 2005*.

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

Chan YC, Cohen DD, Hawas O, Stelcer E, Simpson R, Denison L, Wong N, Hodge M, Comino E and Carswell S 2008b, Apportionment of sources of fine and coarse particles in four major Australian cities by positive matrix factorisatiion, *Atmospheric Environment*, vol.42, pp.374–389.

Chan YC, Hawas O, Hawker D, Vowles P, Cohen DD, Stelcer E, Simpson R, Golding G and Christensen E 2011, Using multiple type composition data and wind data in PMF analysis to apportion and locate sources of air pollutants, *Atmospheric Environment*, vol.45, pp.439–449.

Cohen DD, Bailey GM, and Kondepudi R 1996, Elemental analysis by PIXE and other IBA techniques and their application to source fingerprinting of atmospheric fine particle pollution, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, vol.109, pp.218–226.

Cohen DD, Garton D and Stelcer E 2000, Multi-elemental methods for fine particle source apportionment at the global baseline station at Cape Grim, Tasmania, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, vol.161, pp.775–779.

Cohen DD, Stelcer E, Hawas O and Garton D 2004a, IBA methods for characterisation of fine particulate atmospheric pollution: a local, regional and global research problem, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, vol.219, pp.145–152.

Cohen DD, Kahn R, Anderson J, Anderson TL, Bates T, Brechtel F, Carrico, CM, Clarke A, Doherty SJ, Dutton E, Flagan R, Frouin R, Fukushima H, Holben B, Howell S, Huebert B, Jefferson A, Jonsson H, Kalashnikova O, Kim J, Kim S-W, Kus P, Li W-H, Livingston JM, McNaughton C, Merrill J, Mukai S, Murayama T, Nakajima T, Quinn P, Redemann J, Rood M, Russell P, Sano I, Schmid B, Seinfeld J, Sugimoto N, Wang J, Welton EJ, Won J-G and Yoon S-C 2004b, Multielemental analysis and characterization of fine aerosols at several key ACE-Asia sites, *Journal of Geophysical Research–Atmospheres*, vol.109(D19).

Cohen DD, Gulson BL, Davis JM, Stelcer E, Garton D, Hawas O and Taylor A 2005, Fine-particle Mn and other metals linked to the introduction of MMT into gasoline in Sydney, Australia: Results of a natural experiment, *Atmospheric Environment*, vol.39(36), pp.6885–6896.

Cohen DD, Crawford J, Stelcer E and Bac VT 2010a, Characterisation and source apportionment of fine particulate sources at Hanoi from 2001 to 2008, *Atmospheric Environment*, vol.44(3), pp.320–328.

Cohen DD, Crawford J, Stelcer E and Bac VT 2010b, Long range transport of fine particle windblown soils and coal fired power station emissions into Hanoi between 2001 to 2008, *Atmospheric Environment*, vol.44(31), pp.3761–3769.

Cohen DD, Stelcer E, Garton D and Crawford J 2011, Fine particle characterisation, source apportionment and long-range dust transport into the Sydney Basin: a long term study between 1998 and 2009, *Atmospheric Pollution Research*, vol.2(2), pp.182–189.

Cohen DD, Crawford J, Stelcer E and Atanacio AJ 2012, Application of positive matrix factorization, multi-linear engine and back trajectory techniques to the quantification of coal-fired power station pollution in metropolitan Sydney, *Atmospheric Environment*, vol.61, pp.204–211.

Cohen DD, Stelcer E, Atanacio AJ and Crawford J 2014,The application of IBA techniques to air pollution source fingerprinting and source apportionment, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, vol.318, pp.113–118.

Hibberd MF, Selleck PW, Keywood MD, Cohen DD, Stelcer E and Atanacio AJ 2013, *Upper Hunter Valley Characterisation Study*, CSIRO, Australia.

Malm WC, Sisler JF, Huffman D, Eldred RA and Cahill TA 1994, Spatial and seasonal trends in particle concentration and optical extinction in the United-States, *Journal of Geophysical Research–Atmospheres*, vol.99(D1), pp.1347–1370.

Paatero P and Tapper U 1994, Positive matrix factorisation: a non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, vol.5, pp.111–126.

Paatero P 2004, *Users guide for positive matrix factorisation programs PMF2 and PMF3*, University of Helsinki, Helsinki, Finland.

Taha G, Box GP, Cohen DA and Stelcer E 2007, Black carbon measurement using laser integrating plate method, *Aerosol Science and Technology*, vol.41(3), pp.266–276.

19. Figure captions

20. Table captions

21. Appendix A- PMF Fits by Element

This Appendix shows the individual plots for each of the 23 elemental species used for each of the sites. Each plot is the PMF fitted mass for that element against the measured IBA mass for that element. The least squared fitted line to data is shown as the solid red line and the gradient and the coefficient of regression R^2 for each plot are shown on the plot. For a perfect fit we would expect the gradient to be 1 and the coefficient of regression R^2=1. For well determined elements with low errors and MDLs the gradients and the R^2 are close to unity. For trace elements like selenium (Se) with low concentrations the errors and the MDL's are large and the PMF fits correspondingly poor.

Lucas Heights

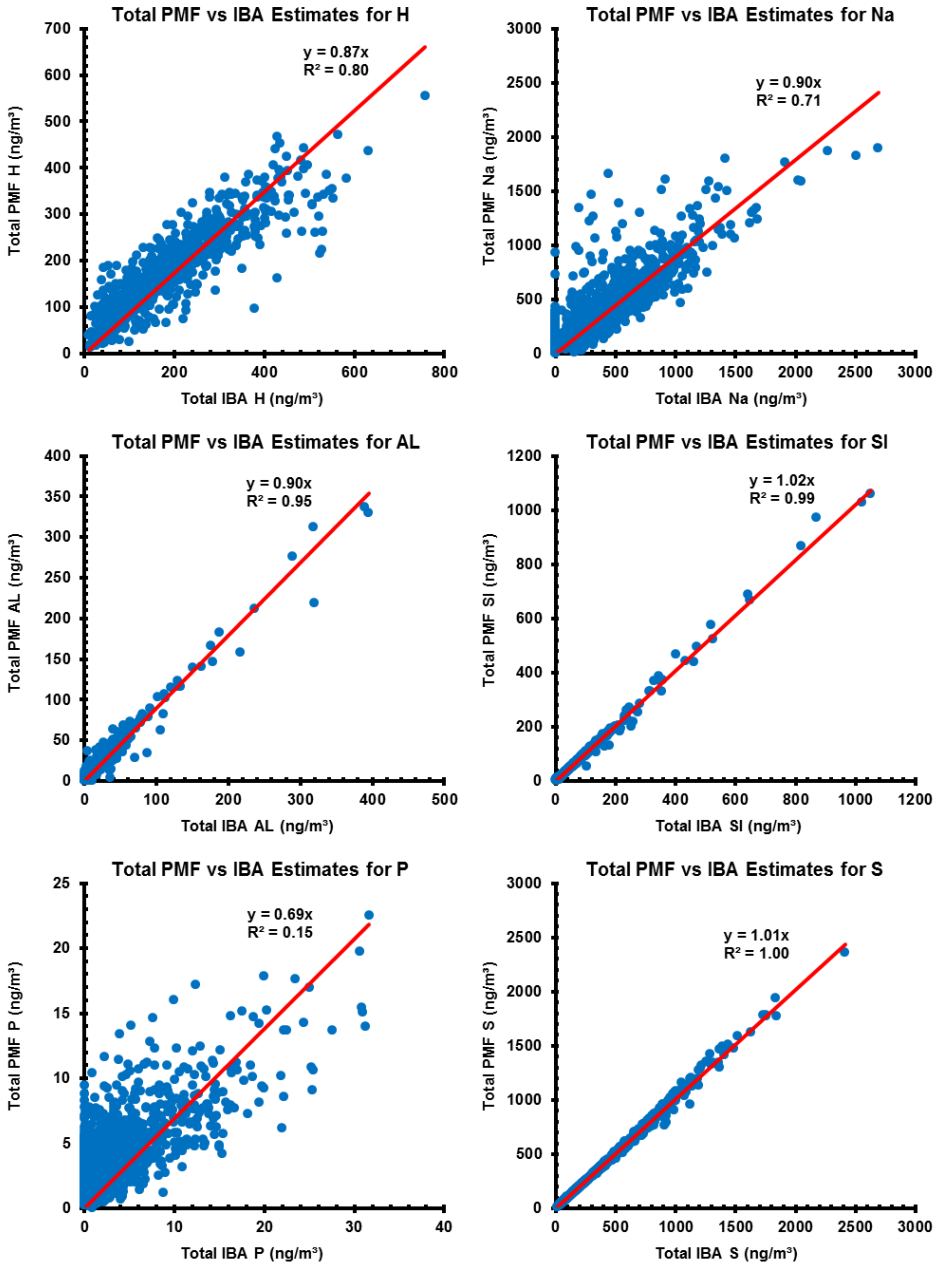

Figure 21-1. Lucas Heights elemental correlation plots of PMF vs. IBA

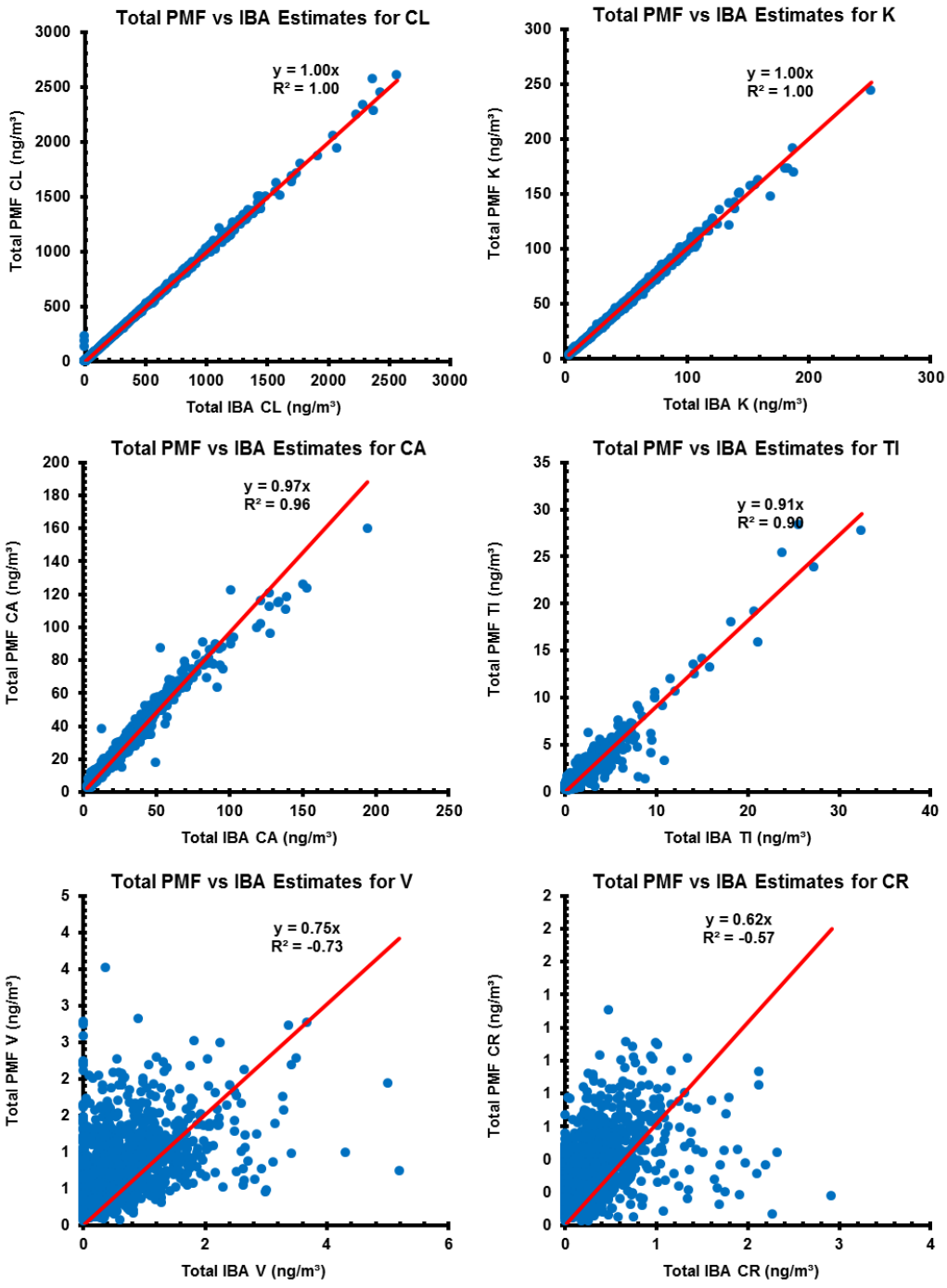

Figure 21-2. Lucas Heights elemental correlation plots of PMF vs. IBA (cont'd)

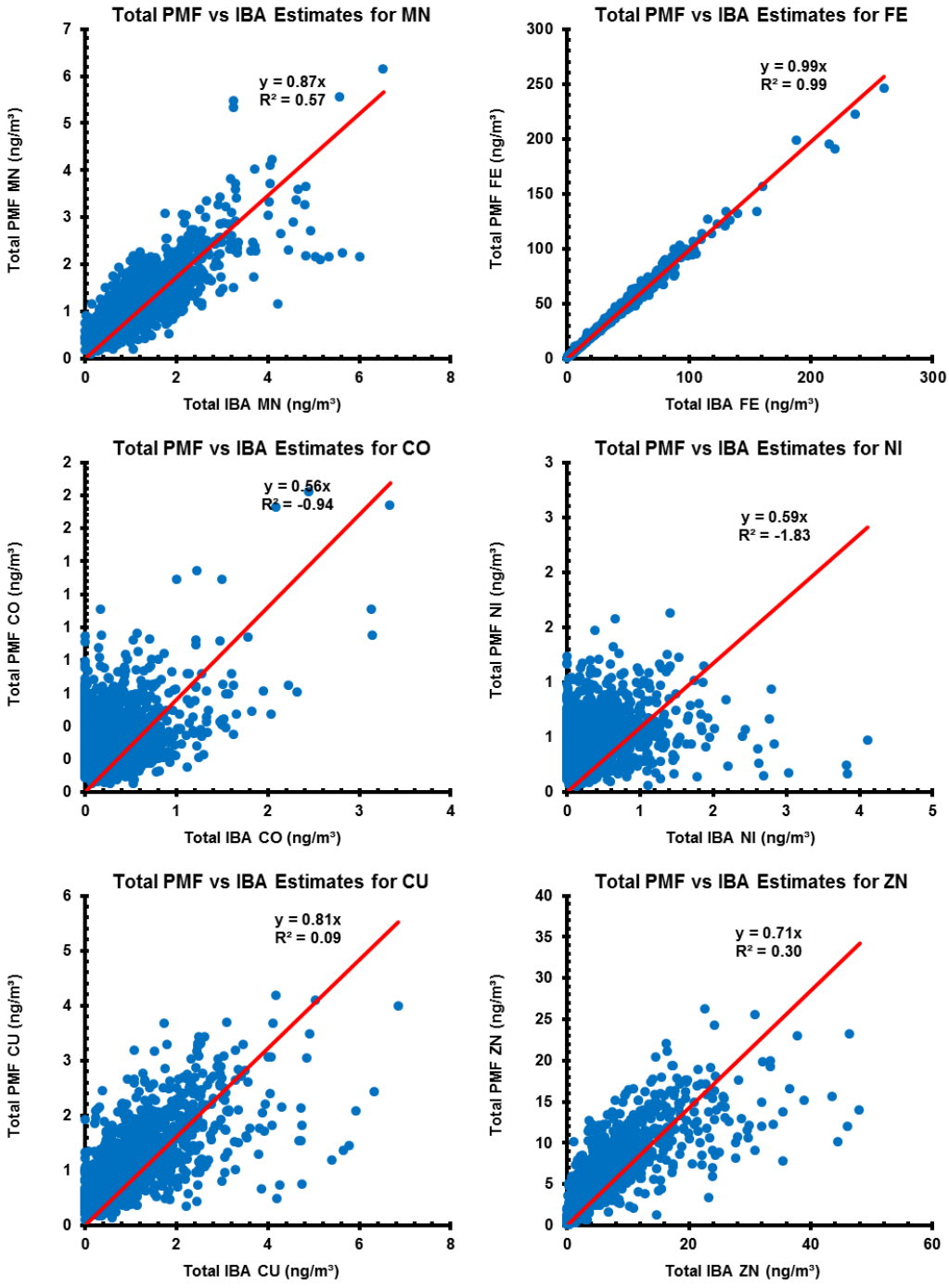

Figure 21-3. Lucas Heights elemental correlation plots of PMF vs. IBA (cont'd)

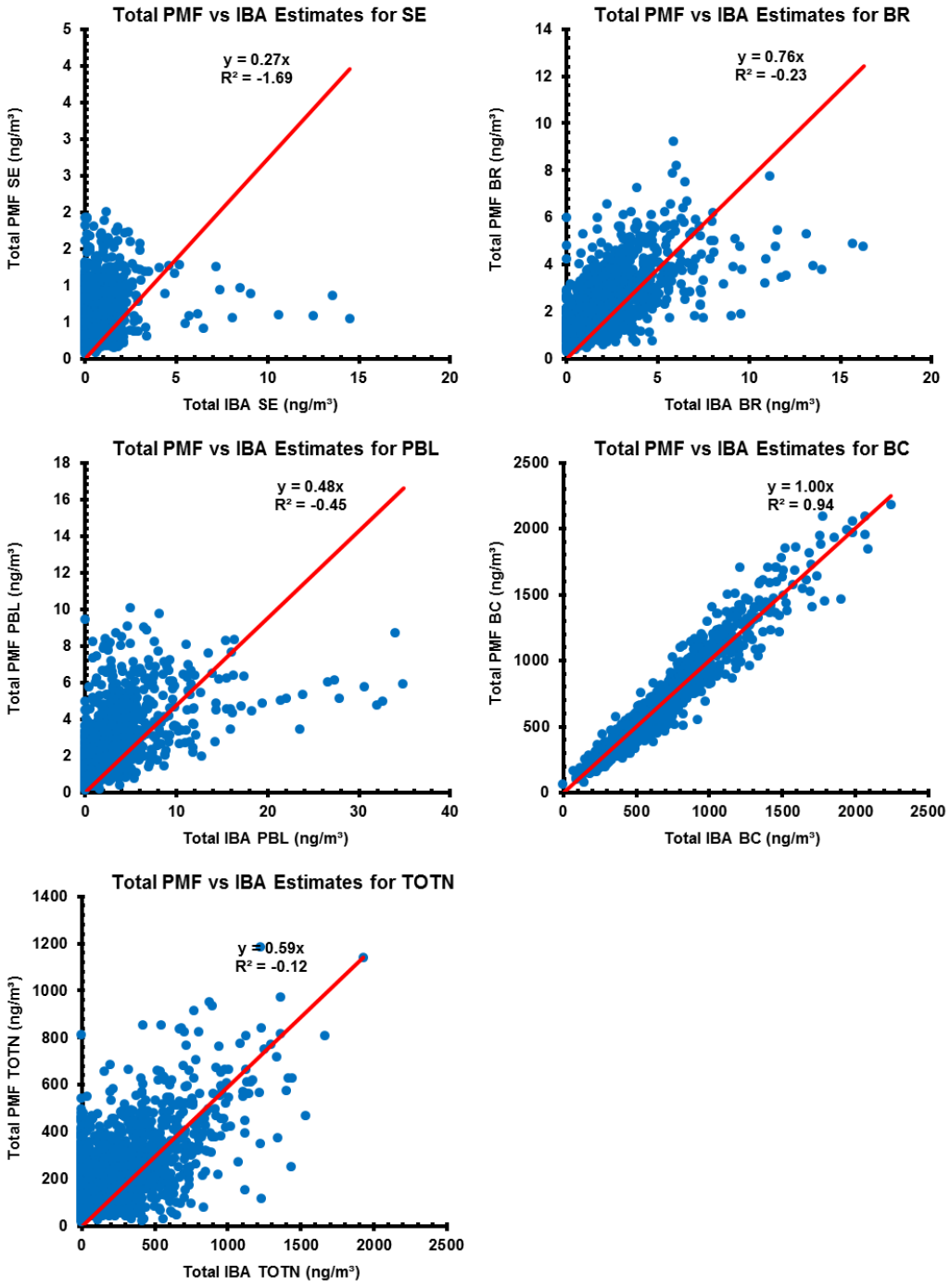

Figure 21-4. Lucas Heights elemental correlation plots of PMF vs. IBA (cont'd)

Richmond

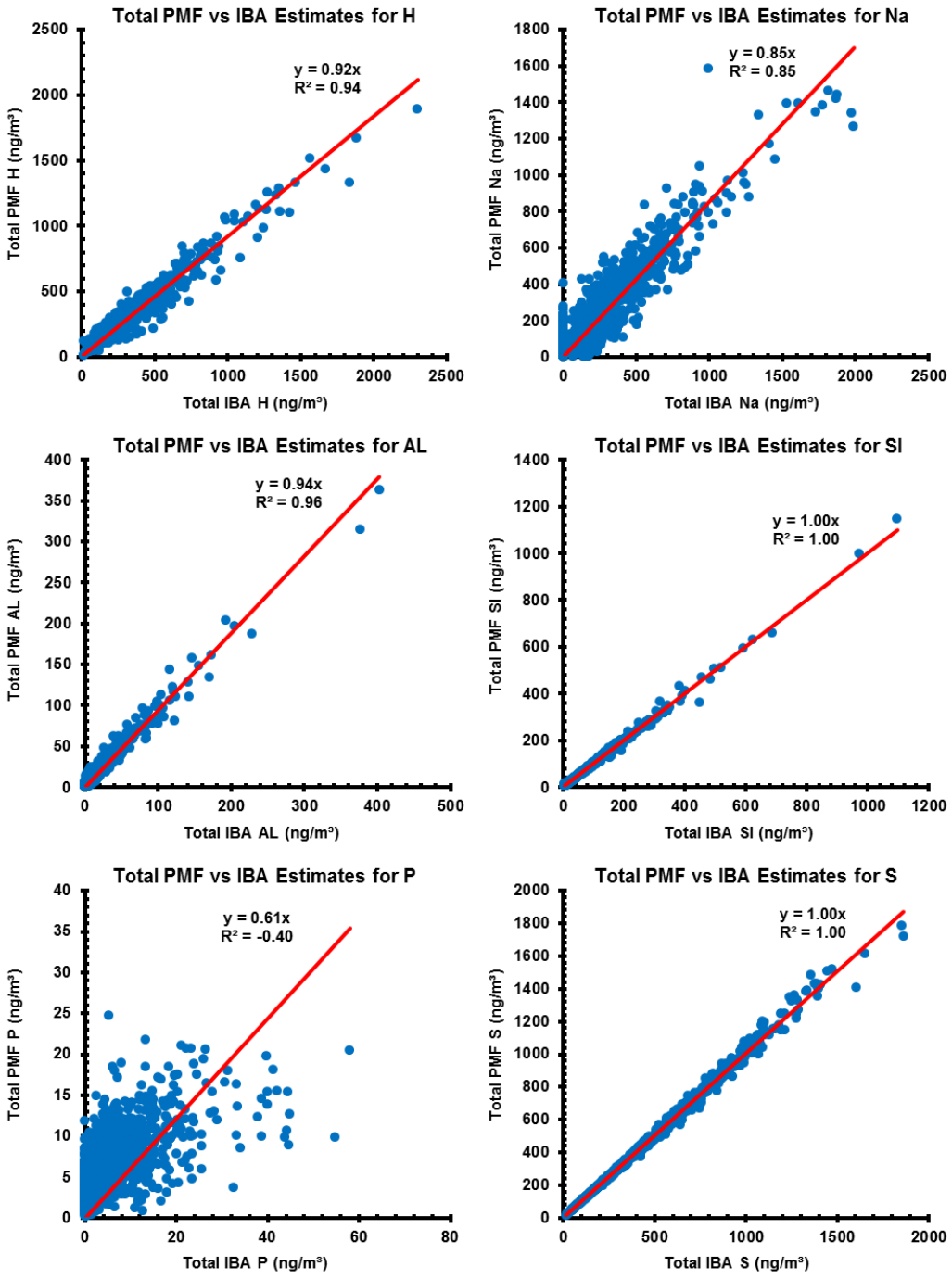

Figure 21-5. Richmond elemental correlation plots of PMF vs. IBA

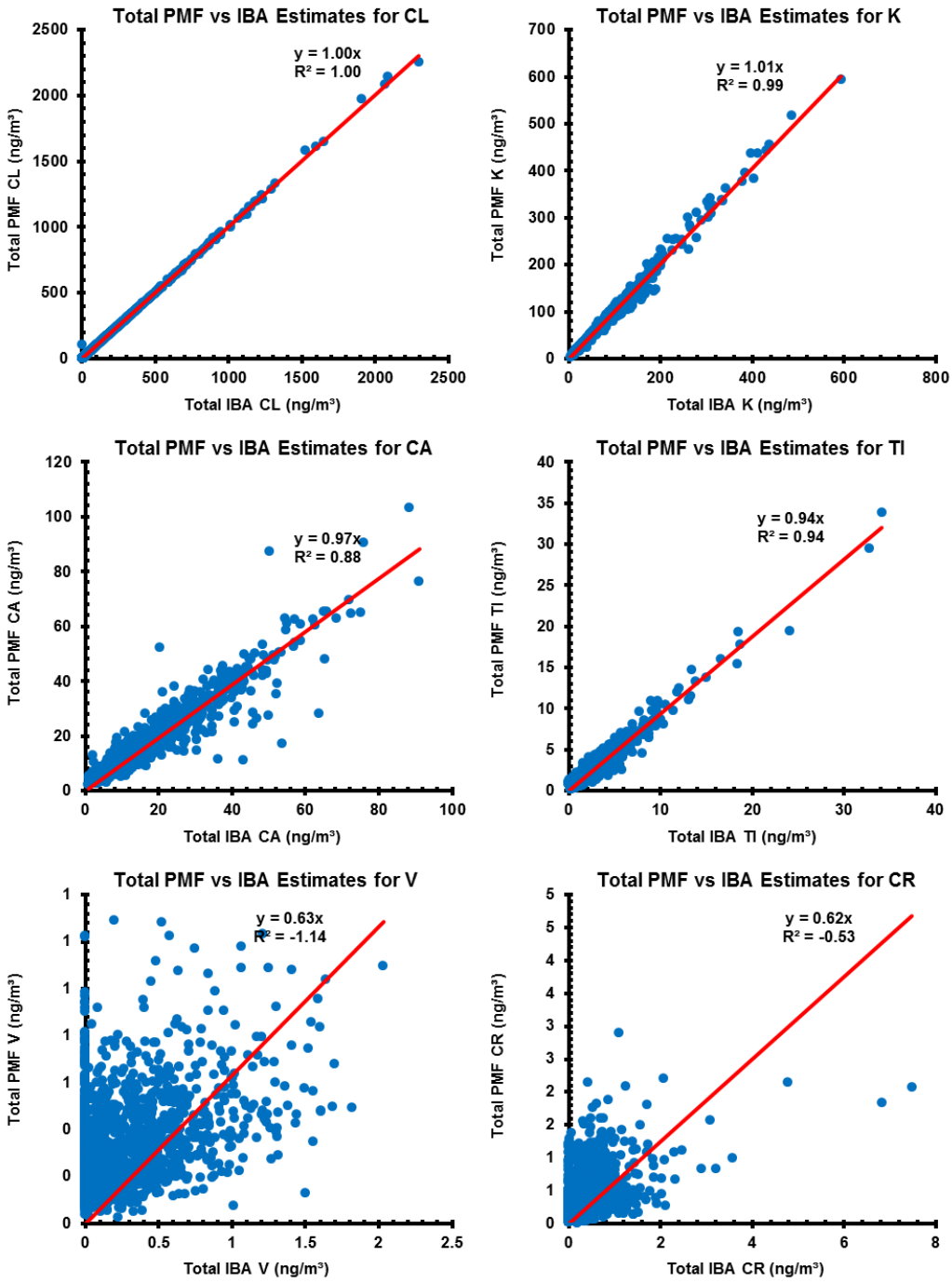

Figure 21-6. Richmond elemental correlation plots of PMF vs. IBA

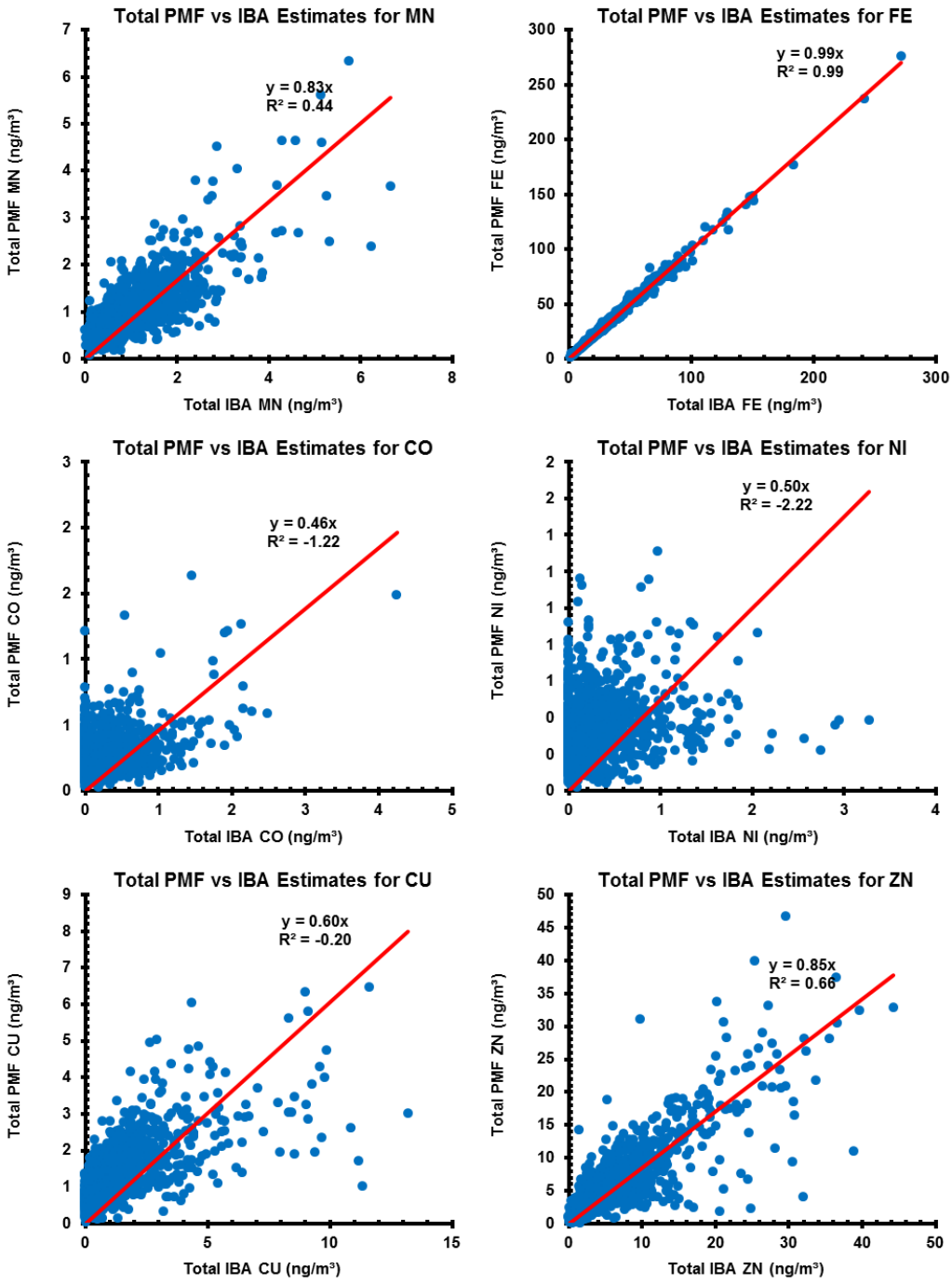

Figure 21-7. Richmond elemental correlation plots of PMF vs. IBA (cont'd)

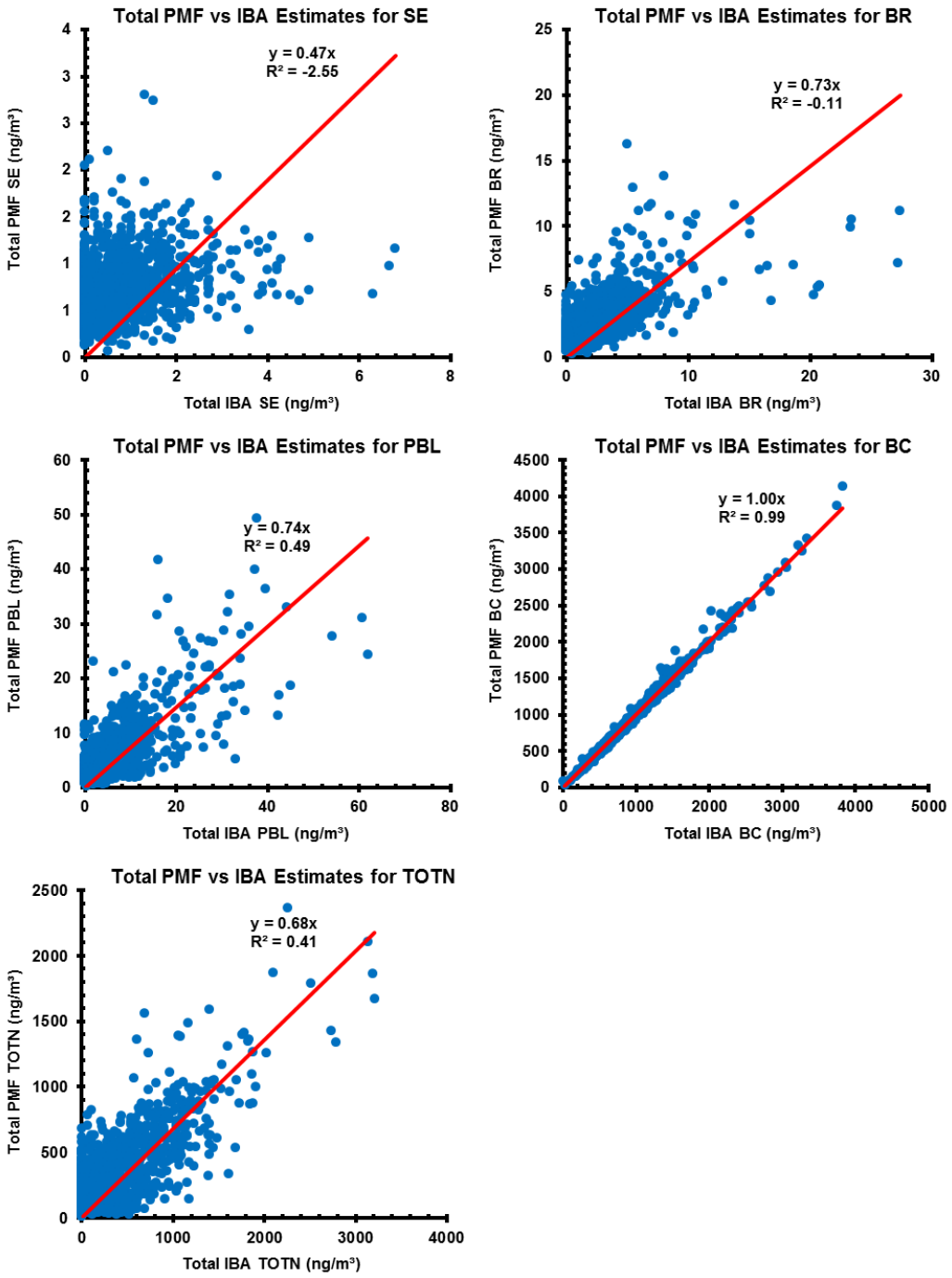

Figure 21-8. Richmond elemental correlation plots of PMF vs. IBA (cont'd)

Mascot

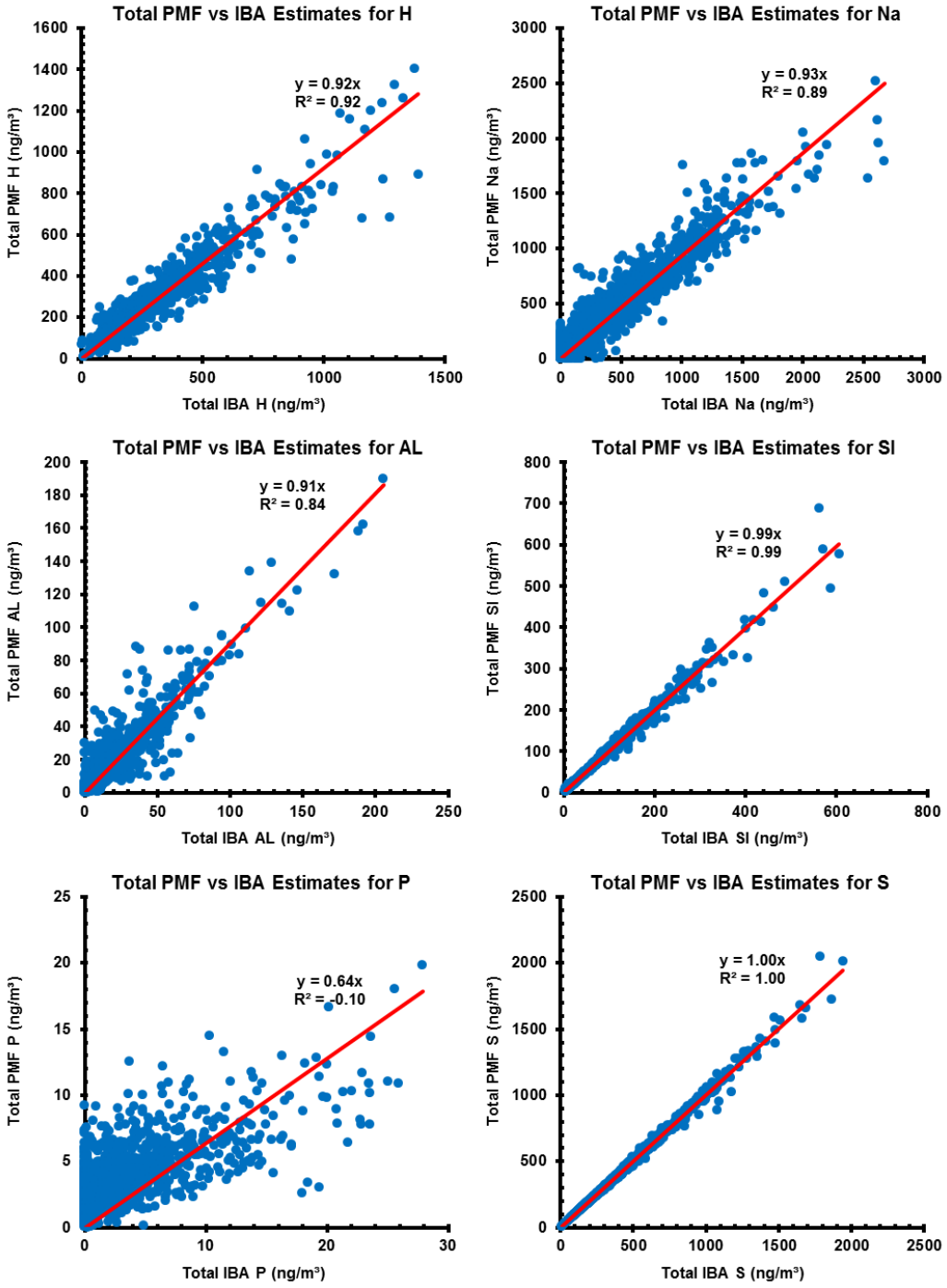

Figure 21-9. Mascot elemental correlation plots of PMF vs. IBA

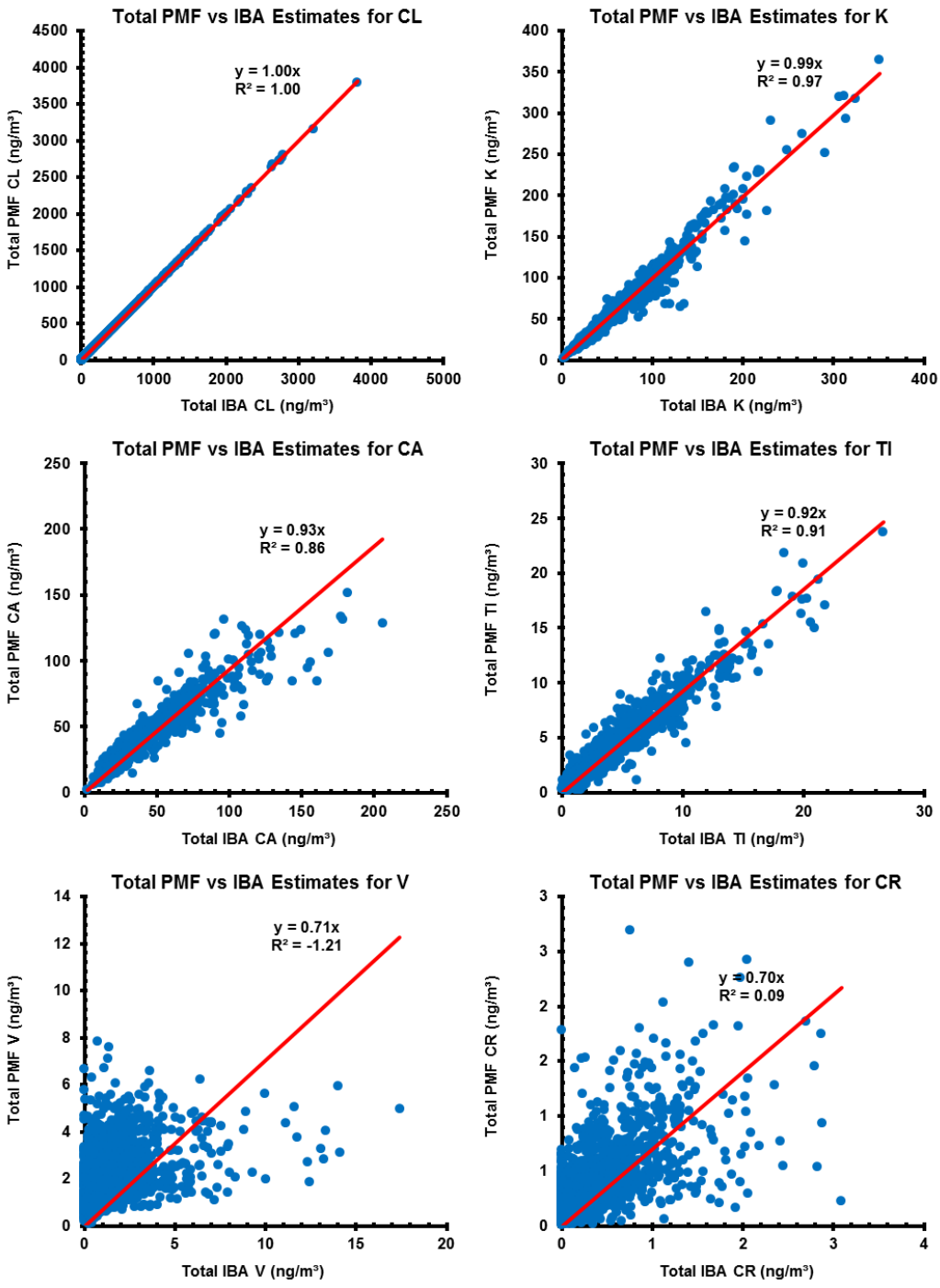

Figure 21-10. Mascot elemental correlation plots of PMF vs. IBA (cont'd)

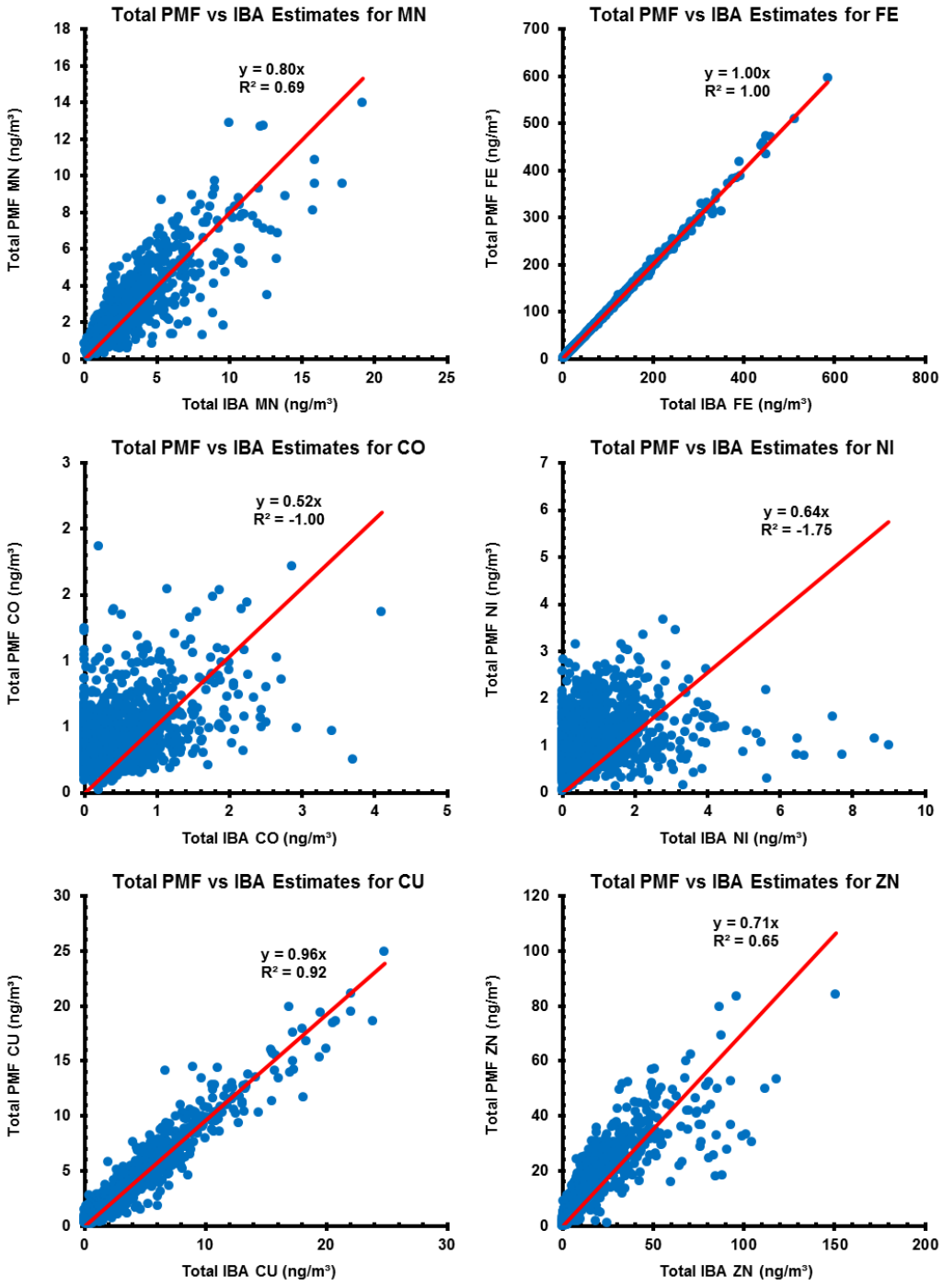

Figure 21-11. Mascot elemental correlation plots of PMF vs. IBA (cont'd)

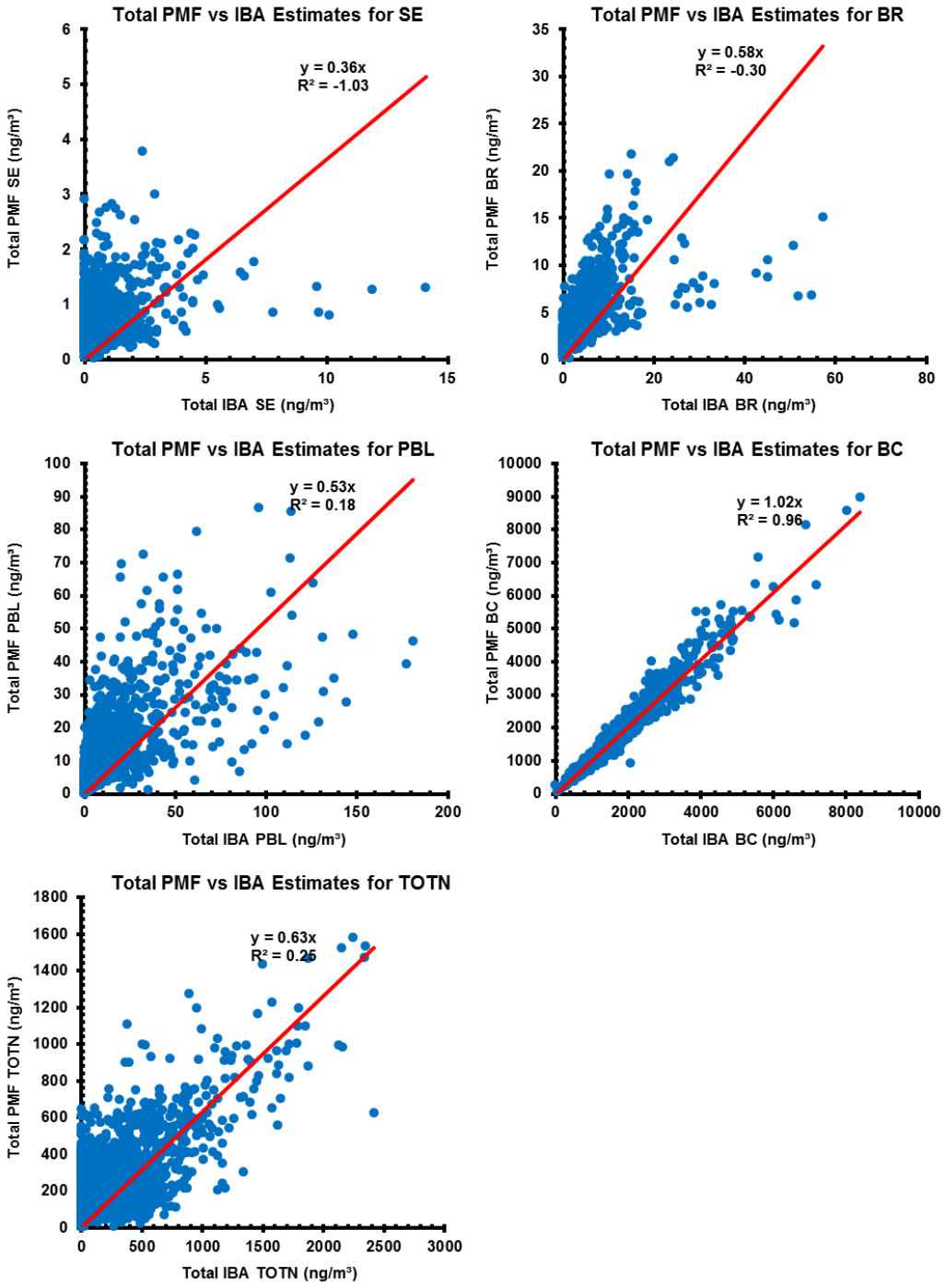

Figure 21-12. Mascot elemental correlation plots of PMF vs. IBA (cont'd)

Liverpool

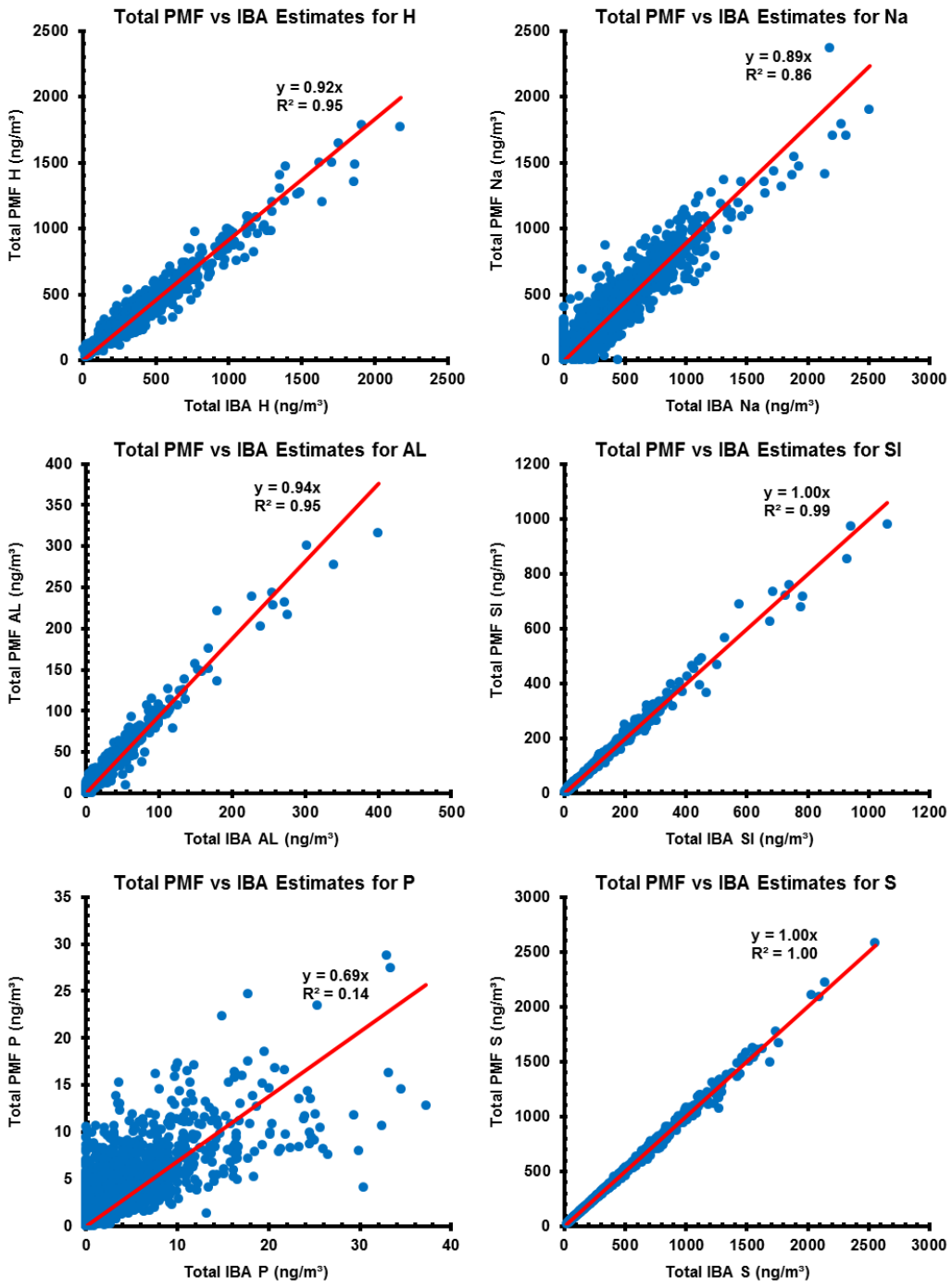

Figure 21-13. Liverpool elemental correlation plots of PMF vs. IBA

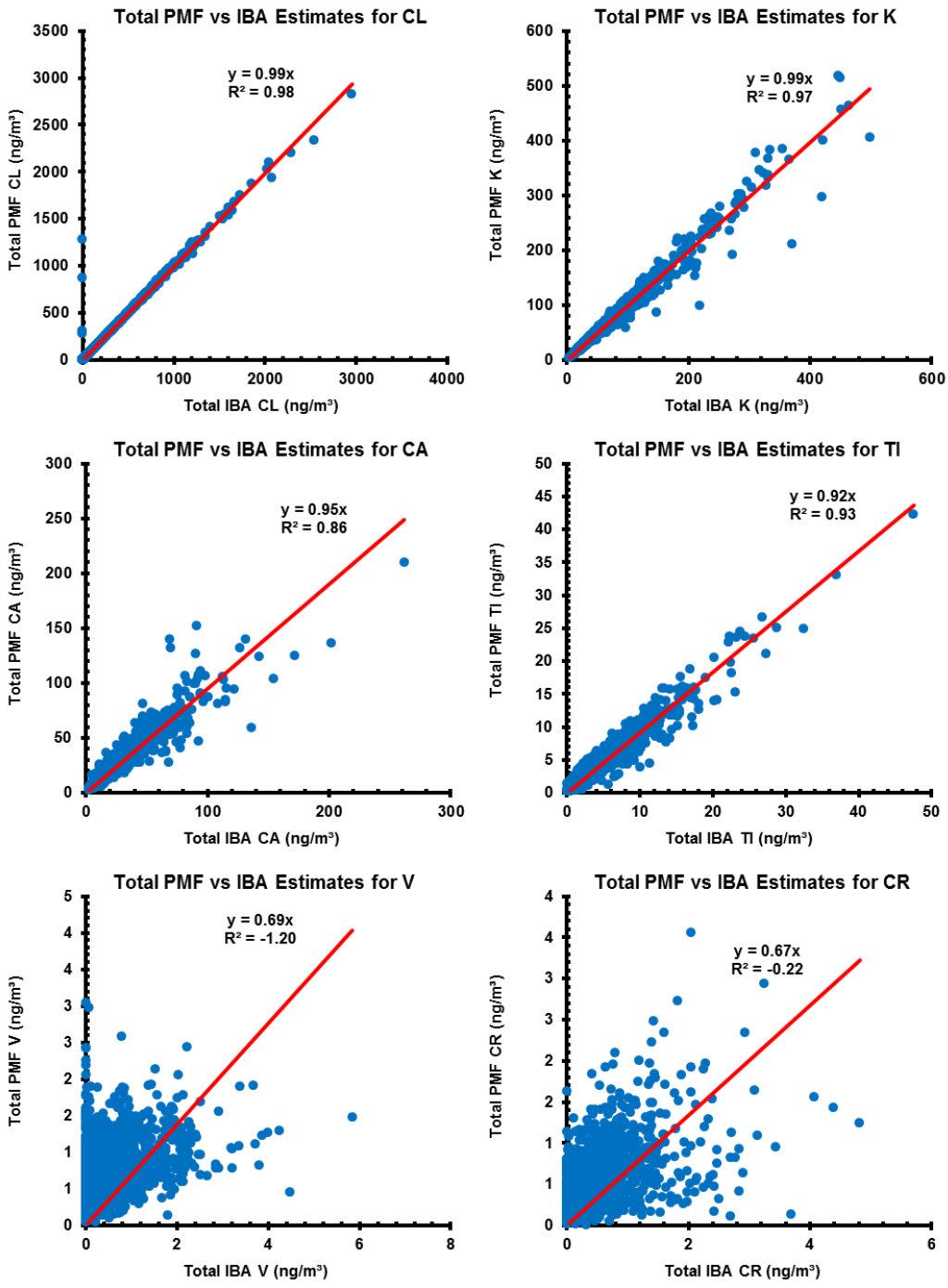

Figure 21-14. Liverpool elemental correlation plots of PMF vs. IBA (cont'd)

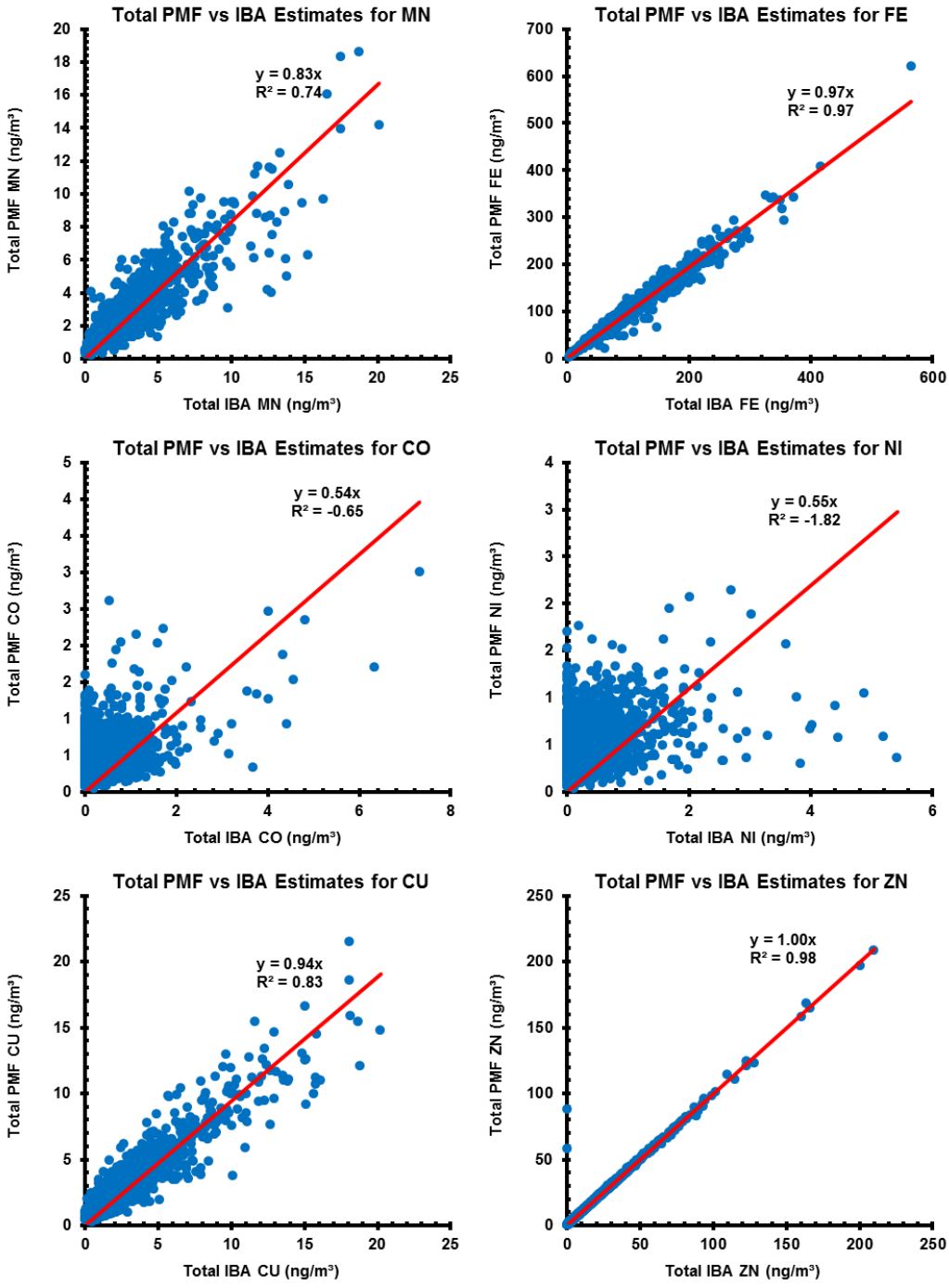

Figure 21-15. Liverpool elemental correlation plots of PMF vs. IBA (cont'd)

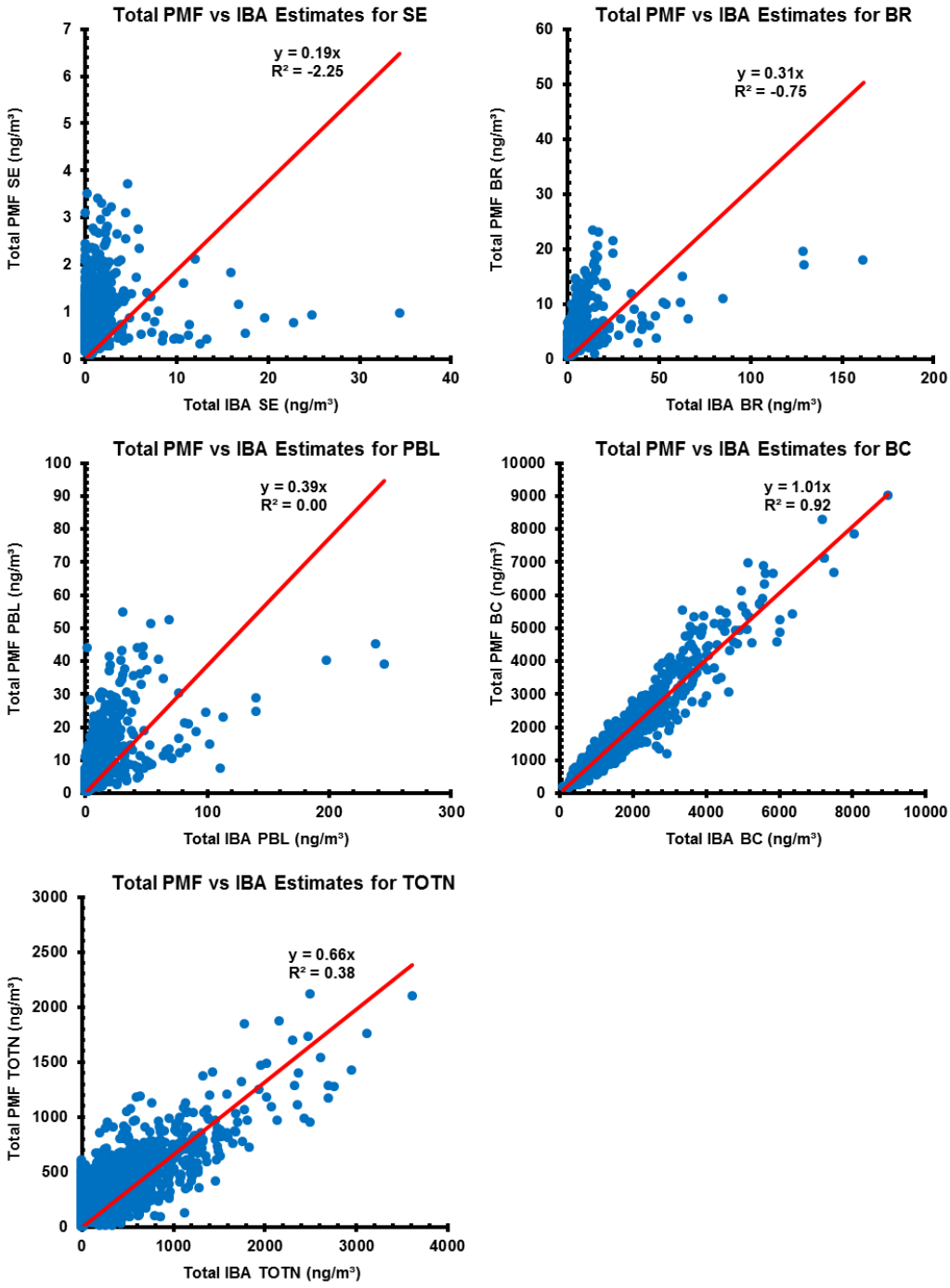

Figure 21-16. Liverpool elemental correlation plots of PMF vs. IBA (cont'd)

22. Appendix B - ANSTO $PM_{2.5}$ PMF Receptor Source Database

Sydney Particle Characterisation Study

$PM_{2.5}$ PMF Receptor Source Database for Greater Sydney Metropolitan Area Sites

(2000-2014)

Instruction Manual

Prepared by

Armand J. Atanacio

Australian Nuclear Science and Technology Organisation (ANSTO)

April 2016

Introduction

This document accompanies the Sydney Particle Characterisation Study - Receptor Source Positive Matrix Factorisation (PMF) Database File.

The aim of this document is to provide instructional steps and related information necessary for navigating and utilising the PMF database. It is important to note that interpretation of the receptor source PMF fingerprints and apportionment contained in the database is beyond the scope of this document and can be found in the full study report.

The Receptor Source PMF Database Macro

The receptor source PMF fingerprint database is provided as the following zipped file: *SydneyParticleStudy2000-14.zip*. This zipped file contains the following three files: (1) the main VBA macro-enabled excel file: *SydneyParticleStudy2000-14.xlsm*, (2) *SamplingSiteMap.html*, and (3) an electronic copy of this instruction manual. Please note, for the remainder of this document, the Excel macro file will be referred to as the *database*. In an effort to minimise the size of this database macro, it only contains the PMF data (without the associated PMF plots) from each of the sites (15 years of PMF data for one site alone is over 2Mb!). The Excel visual basic for applications (VBA) macro functions are then utilised to automatically generate the associated PMF plots from the data for each site as required.

The macro options are accessed from the **Menu** worksheet and their functions described in the following sections of this document. Please note: these macro functions have been written and tested to operate correctly in Excel version 2007-2010.

Important: Depending on your Microsoft Excel settings, you may encounter a "Security Warning Macros have been disabled" alert when you open the PMF Master database files (see **Figure 22-1**). You will need to press the "Enable Content" button for the macros to function correctly.

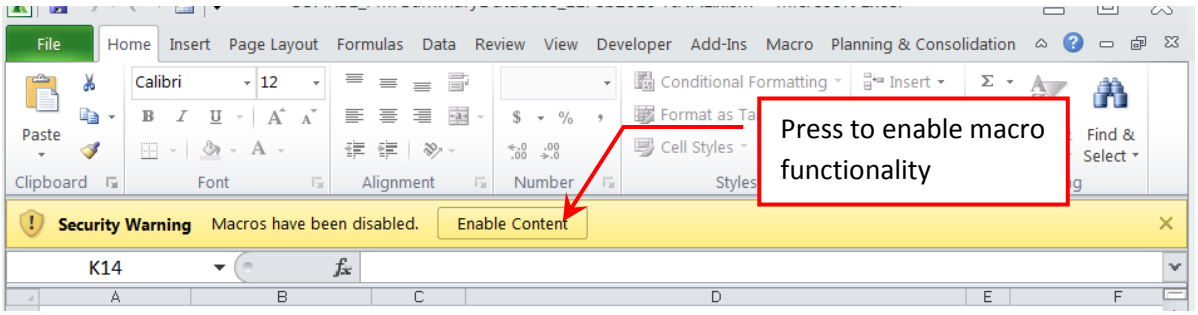

Figure 22-1. Microsoft Excel macro security warning and enable content button

Initialising the Program

When you open the database for the first time, you will land on the "Menu" worksheet (See **Figure 22-2**). For convenience, a brief version of instructions is also available at the top of this page (**Figure 22-2**). To begin using the database, the program and data must first be initialised. This process clears any previous data and plots and also prepares the required macro functionality. Initialisation is performed by pressing the button labelled **"(1) INITIALISE DATABASE"** (**Figure 22-2**). A progress bar will be displayed in bottom left of the database window.

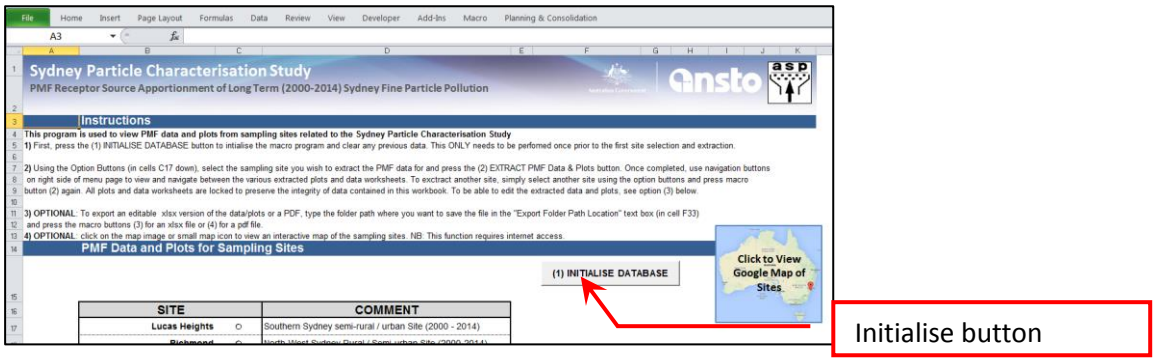

Figure 22-2. Section of the Menu page showing the location of the initialise database button

Extract PMF Data and Plots

Now select the site to extract the PMF data. This is done by clicking on one of the site selection buttons (**Figure 22-3**). Please note, only one site can be selected for extraction at a time.

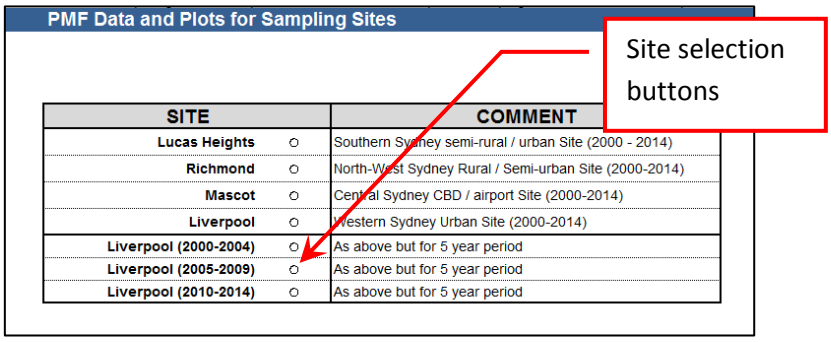

Figure 22-3. Section of the Menu page showing the SITE name, site selection buttons and the related comments

Once a site has been selected, press the **"(2) EXTRACT PMF Data & PLOTS**" button (**Figure 22-4**). As there is a significant amount of data for each site, the extraction and plotting may take up to a minute to complete depending on the processing speed of your computer. The progress bar will be shown in the bottom left corner of the excel window during the extraction process.

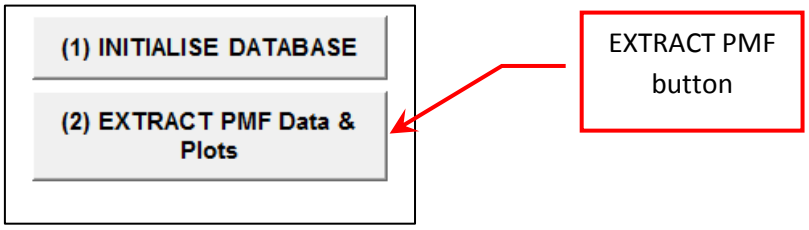

Figure 22-4. Extract PMF Data and plots button

Once the extraction is complete, the Menu worksheet will be updated with additional SUMMARY INFORMATION, worksheet navigation buttons and file export buttons (**Figure 22-5**).

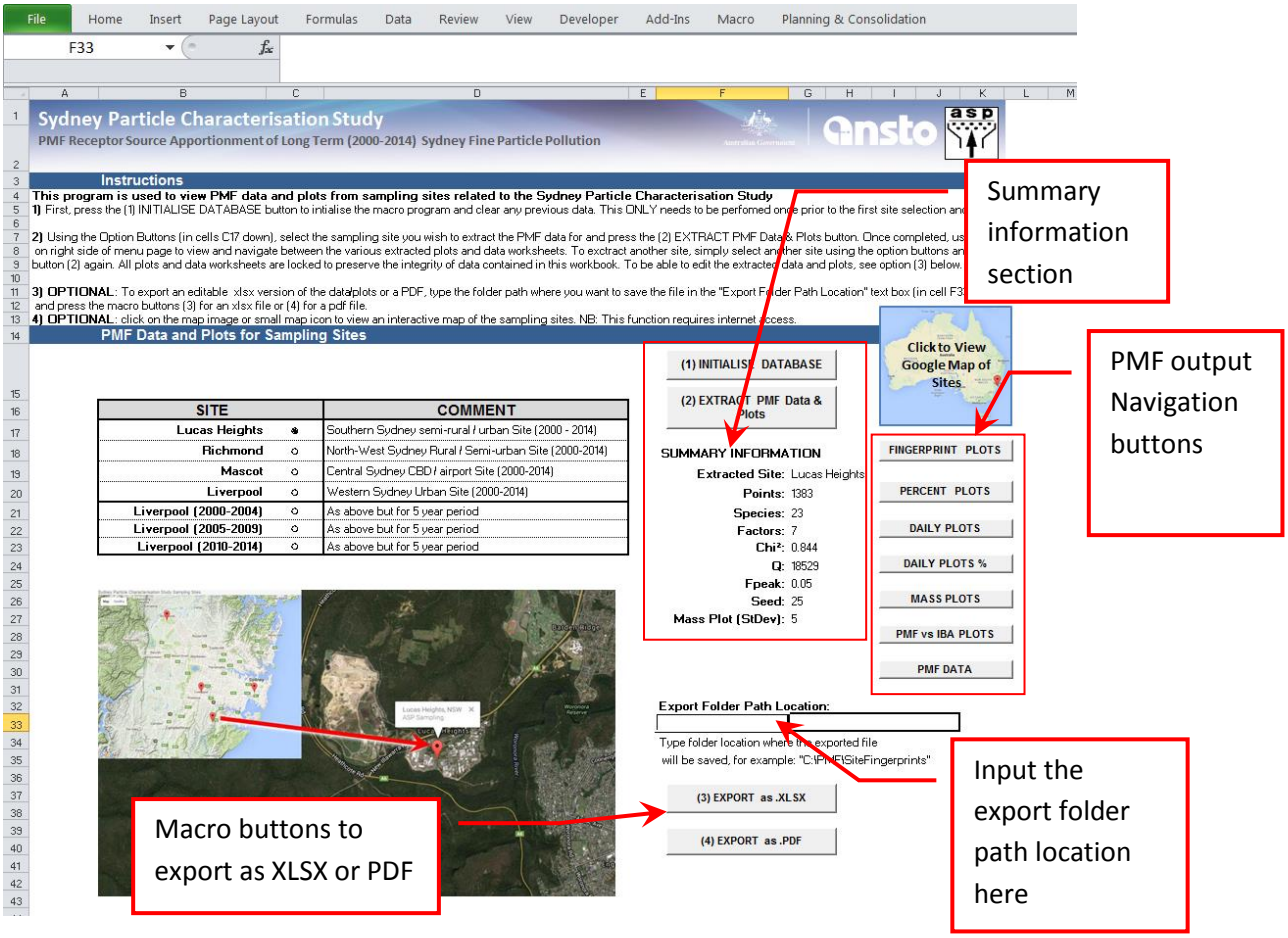

Figure 22-5. Location of summary information, worksheet navigation buttons, xlsx and pdf file export text box and buttons on Menu page

This summary information provides the following details: the extracted site, number of points (i.e. sampling days) used in the PMF analysis, number of species (or elements) used in the PMF analysis, number of fingerprints obtained with PMF analysis, chi-squared (χ^2) value, the Q-value, F-peak value, Seed value and the number of standard deviations used for the tramlines on the Mass Plot (**Figure 22-5**).

The Menu worksheet will also be updated with additional navigational buttons which can be used to go between the sites PMF data and plots (**Figure 22-5**). These navigation buttons are also available on each of the generated worksheet.

Export Data and plots as xlsx or pdf file

XLSX

The various data and plots can be viewed within the database using the navigational button. However, all of the database worksheets are **locked** and cannot be edited. This is to maintain the integrity of the data contained in the database program itself. If you would like to edit the extracted data or plots (for example, to re-plot the data in another program), you will need to export the extracted data as an **unlocked** xlsx file. To do this, type an existing folder path location where you would like this .xlsx file to be exported to in cell F33, for example, "C:\PMF\Australia\fingerprints". Now press the macro button "**(3) EXPORT as .XLSX**" to export the file (see **Figure 22-5**). The filename of the exported xlsx file will be automatically generated as the selected site name. This xlsx file will contain the following: a summary information cover page (CoverPage) (see **Figure 22-7)**, worksheet containing the PMF data

(named as the site that was extracted), worksheet of fingerprint plots (FingerprintPlots), worksheet of percent plots (PercentPlots), worksheet of daily plots (DailyPlots), worksheet of daily plot percent (DailyPlots%), worksheet of PMF Mass vs gravimetric Mass plots (MassPlots) and a worksheet of PMF vs IBA plots for each specie.

PDF

To export a pdf file, follow the same process described above for the xlsx file but press the macro button "**(4)EXPORT as .PDF**". The generated pdf file contains the same cover page and plots as the xlsx file, however, it does not include the PMF data worksheet – as converting this worksheet to pdf can fill more than 20 pages depending on the amount of data associated with the selected site. The filename of the pdf file will be automatically generated to be the same as the selected site.

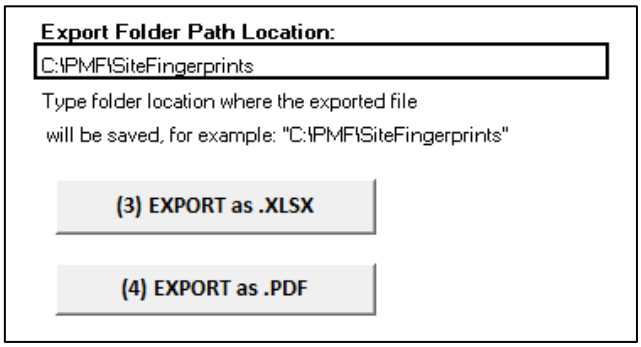

Figure 22-6. Section of Menu page related to xlsx or pdf file export

Cover Page

A cover page is automatically generated and included when you export the data as either an xlsx or pdf file (**Figure 22-7**). This cover page provides the date the data was extracted, site, comment (with space to input additional comments if required), PMF analysis summary information, and a map showing the immediate vicinity of the site (**Figure 22-7**).

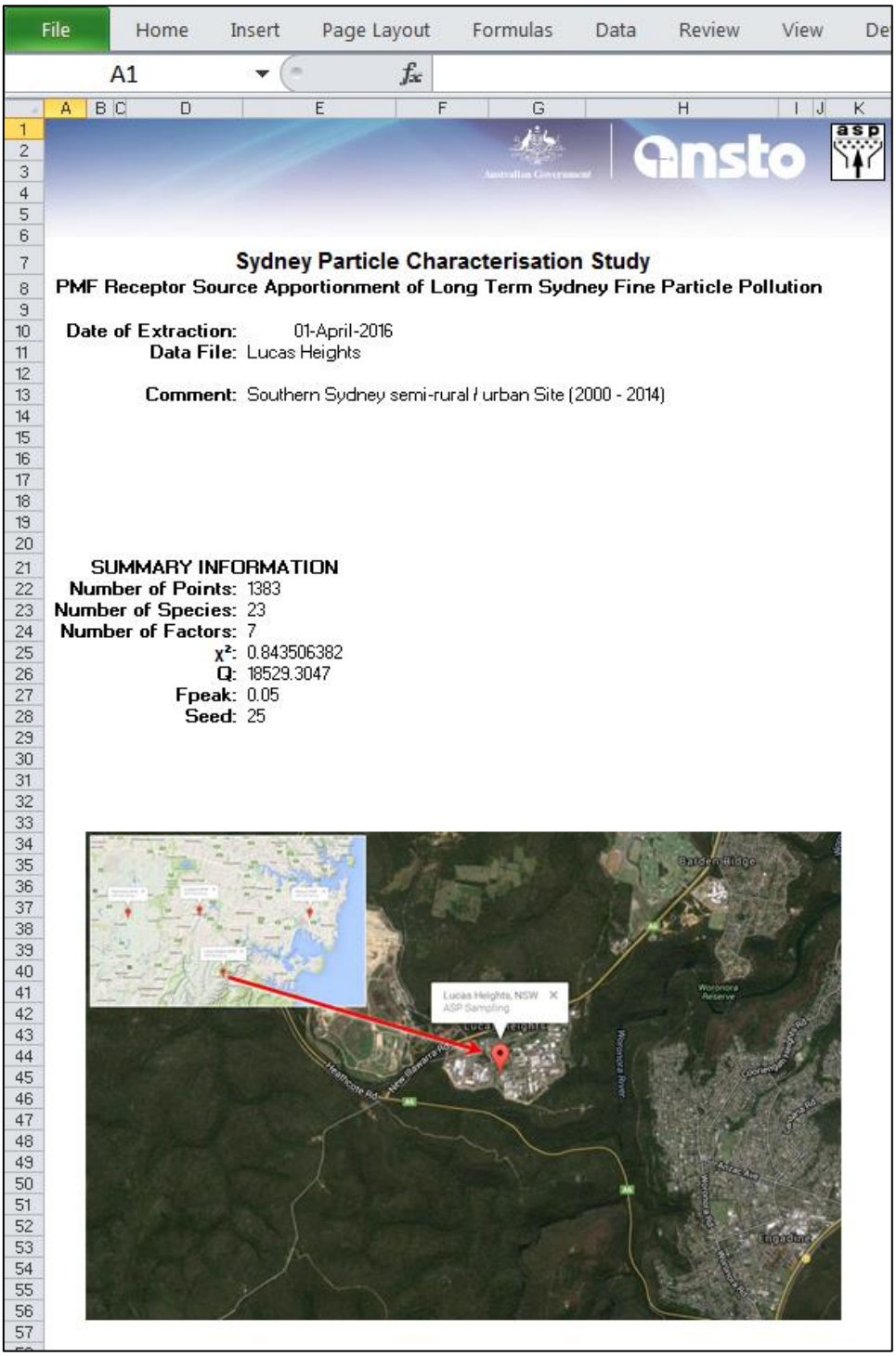

Figure 22-7. Example of the cover page worksheet included in exported xlsx or pdf files

Description of Extracted Worksheets

The following sections describe the various worksheets generated from the PMF data extraction process described earlier**.**

Fingerint Plots

The PMF fingerprint plots are displayed on the FingerprintPlots worksheet (**Figure 22-8**). Each fingerprint is comprised of the fractional ratios for each correlating element. The elemental ratios in each fingerprint have been normalised to have the maximum element with a value of 1. This clearly identifies the main driving element/s for a particular fingerprint and assist the data analyst in assigning possible fingerprint names. The Y-axis of each fingerpint is a 4-decade log plot which allows trace elements to also be easily identified. The error bars relate to 3 standard deviations calculated by the PMF analysis codes. The percentage value in brackets next to the title of each fingerprint denotes the percentage of that fingerprint to the total 'fitted' mass, not the total gravimertic mass.

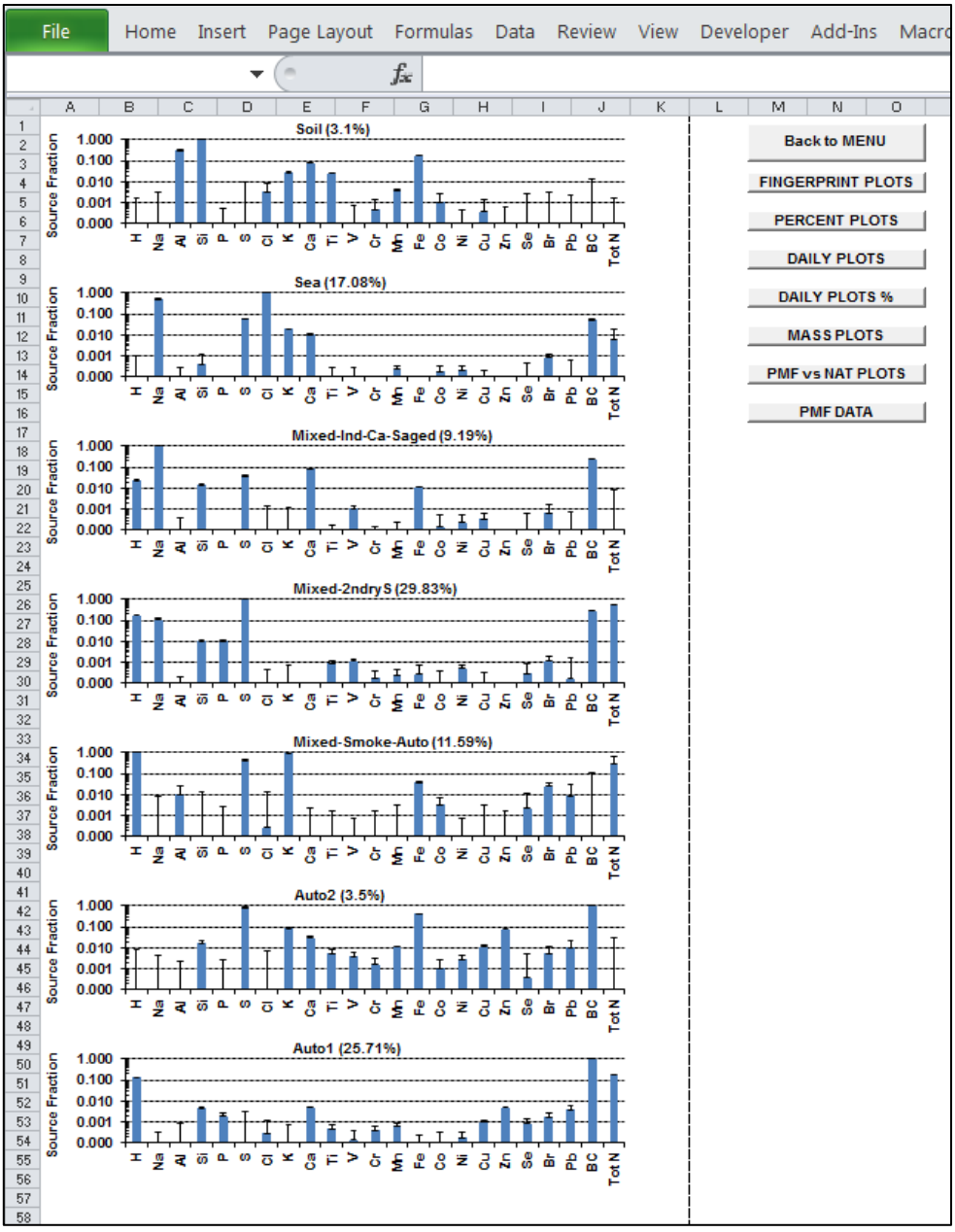

Figure 22-8. Example of a set of PMF fingerprints displayed on the FingerprintPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.

Percentage Plots

The plots on the PercentPlots worksheet represent the percentage distribution of each of the elements across all identified fingerprints. For example, adding the Na percentage shown in each of the 7 fingerprints (see **Figure 22-9**) will account for 100% of the Na fitted mass. These plots are NOT source fingerprints and should not be used as such. They are useful to assess where the major contributions of each element reside in terms of fingerprints. Correct PMF fingerprint interpretation requires the use of both the fingerprints plots (**Figure 22-8**) and the percentage plots (**Figure 22-9**) in conjunction to best to identify meaningful receptor sources.

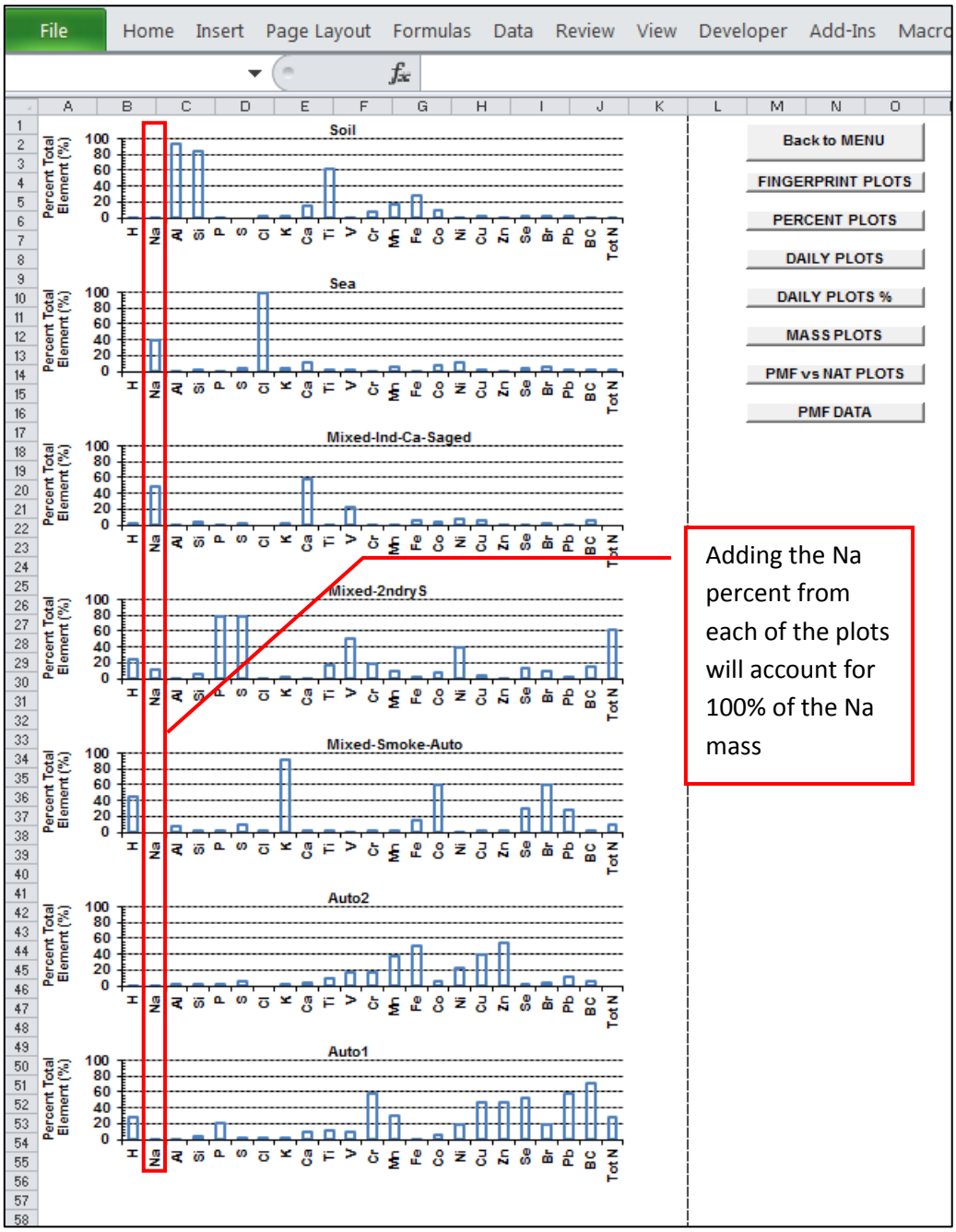

Figure 22-9. Example of a set of PMF percentage plots on the PercentPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.

Daily Plots

The plots on the DailyPlots worksheet (**Figure 22-10**) represent the daily contribution of each fingerprint in nanograms (ng/m³). The date on the x-axis has the format DD-MMM-YY, for example 02-Jan-14 denotes Day =02, Month =January and Year =2014.

These plots are useful to observe daily, monthly and yearly trends for each identified fingerprint. Such trends may include: gradual yearly reduction in receptor sources as a result of the implementation of pollution reduction policies, regular monthly seasonal variations of particular receptor sources (e.g. summer-winter), or extreme daily events such as dust storms or bushfires.

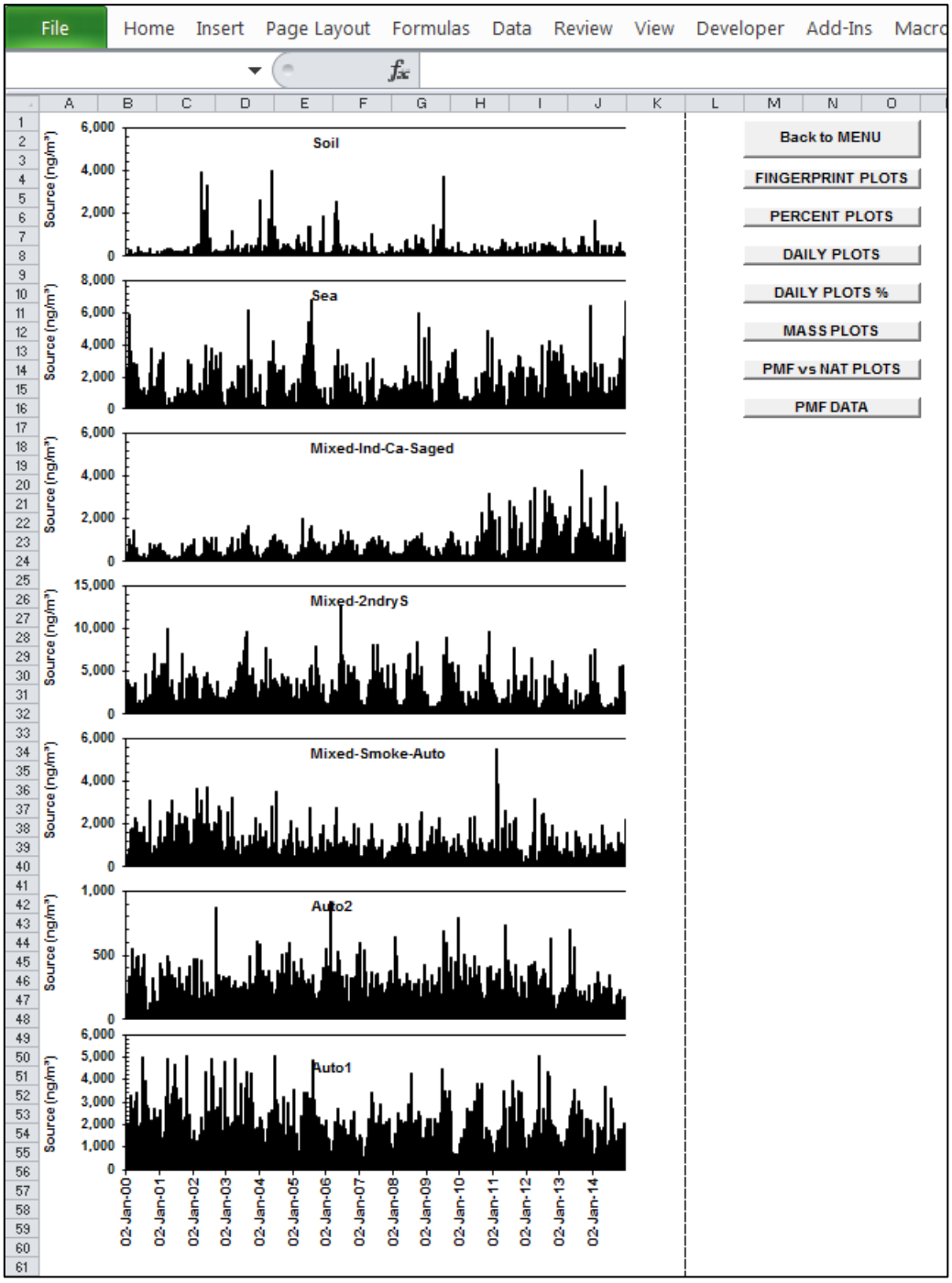

Figure 22-10. Example of plots on the DailyPlots worksheet showing the daily time series contribution of each PMF fingerprints in ng/m³. The worksheet navigation button panel is also seen on the right side of the image.

Daily Percentage Plots

The plots on the DailyPlots% worksheet (**Figure 22-11**) are similar to the daily plots (**Figure 22-10**) but instead of daily concentration in ng/m³, these plots represent the percentage (%) contribution of each identified receptor source to total pollution as a daily time series. The date on the x-axis has the format DD-MMM-YY, for example 02-Jan-14 denotes Day =02, Month =January and Year =2014.

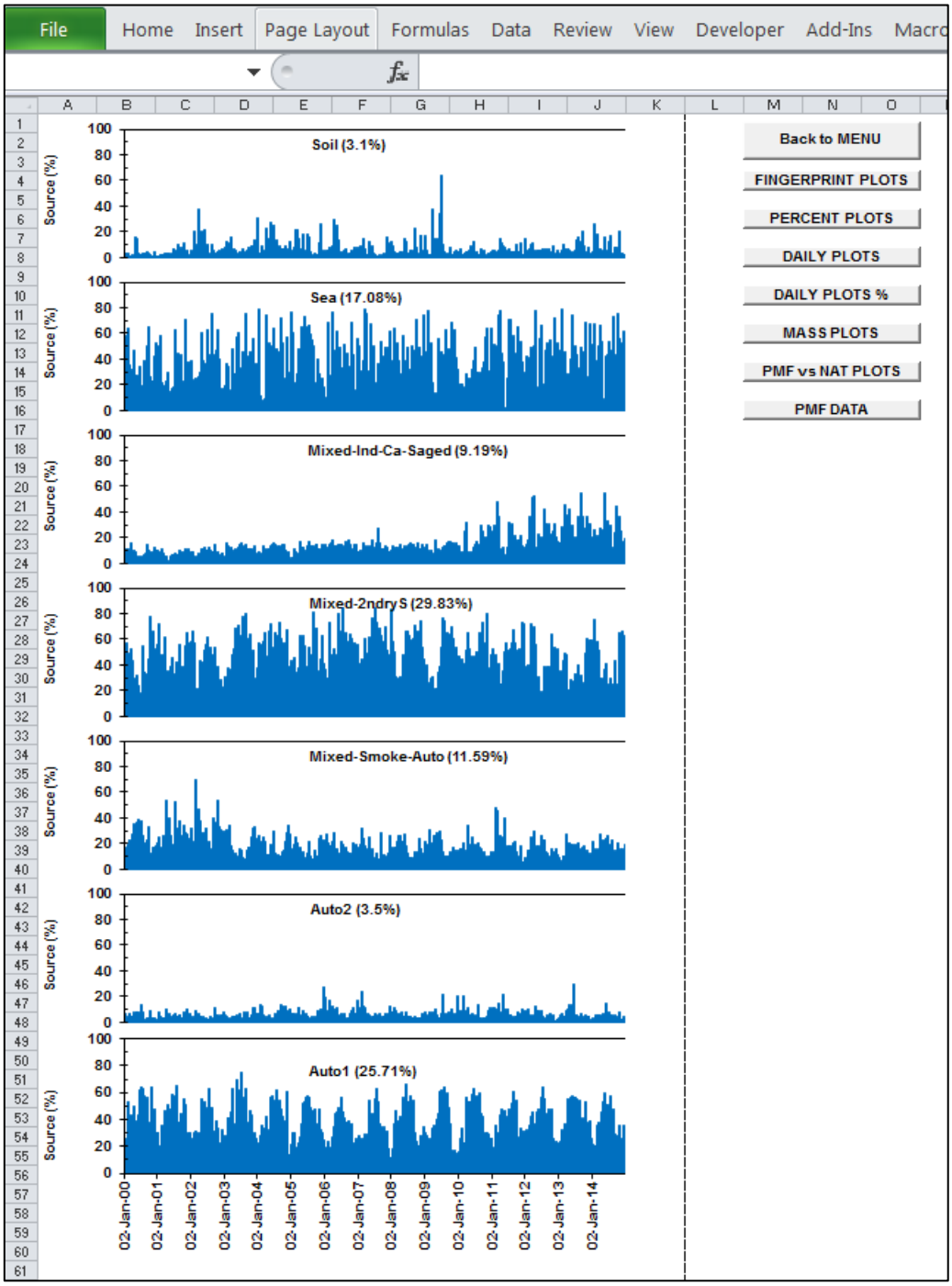

Figure 22-11. Example of plots on the DailyPlots% worksheet showing the each PMF fingerprints daily percentage (%) contribution. The worksheet navigation button panel is also seen on the right side of the image.

Mass Plots

The MassPlots worksheet contains two mass related plots. The first plot compares the calculated PMF mass with the measured gravimetric mass, both in ng/m³. Ideally, both the linear fit and R^2 value should be close to 1. However, this linear fit can be skewed significantly by outlying points causing the fitted data to be representative of neither the majority of data nor the outlying point, but somewhere in-between. To avoid this, the data analyst generally tries to remove most of the outlying or extreme points outside of the "tramlines" represented by the dotted lines in the top plot of **Figure 22-12**. These tramlines represent a certain number of standard deviations (SD) from the fitted line, generally between 3 to 6 standard deviations. The tramline SD value which was applied in each site's PMF analysis is provided in the SUMMARY INFORMATION section (**Figure 22-5**).

The second plot on the MassPlots worksheet is a time series comparison of both the PMF mass and gravimetric mass. It is used to check that the daily PMF fit obtained from statistical analysis matches the measured daily gravimetric mass.

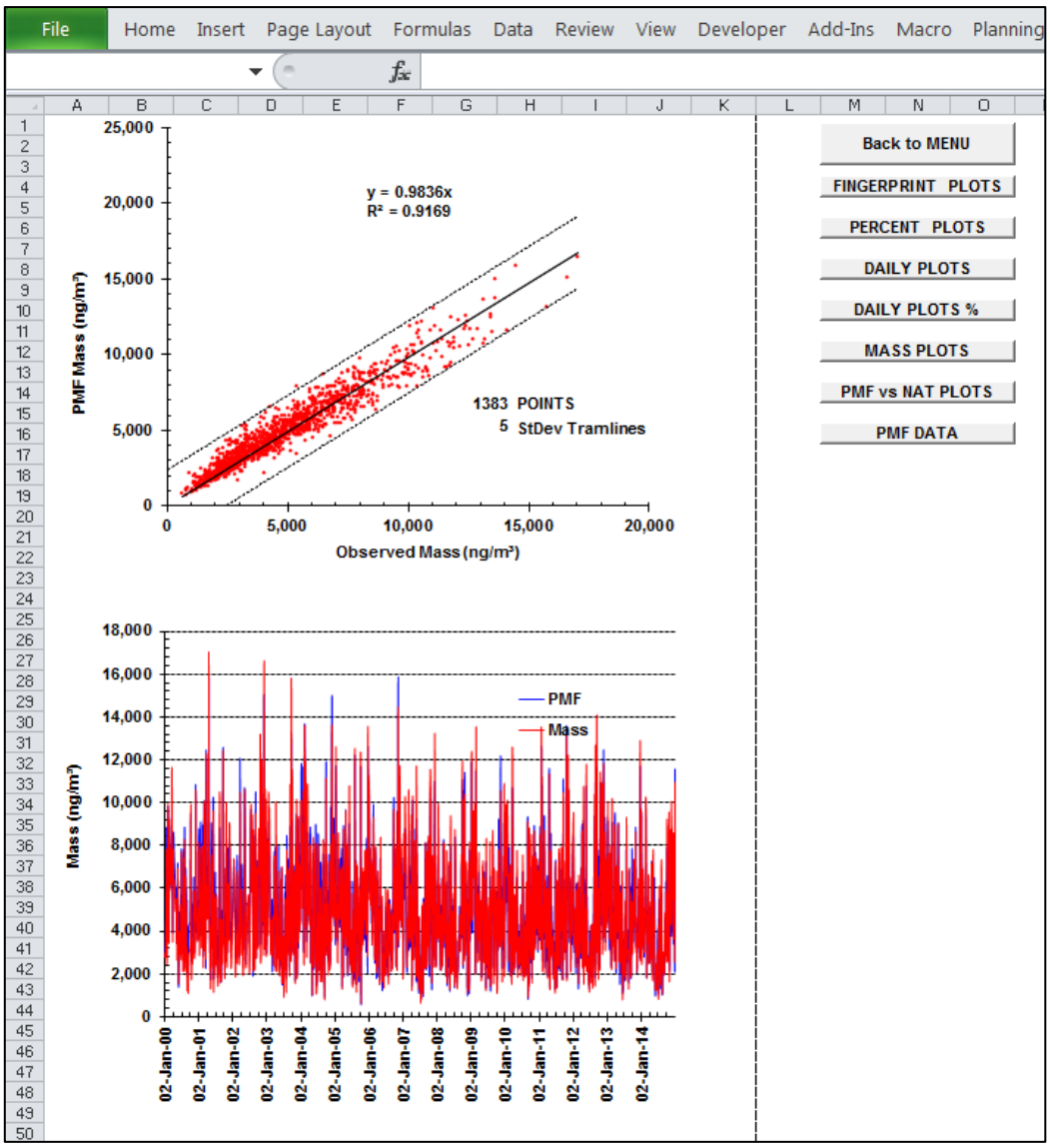

Figure 22-12. Example of plots on the MassPlots worksheet comparing PMF mass with gravimetric mass. The worksheet navigation button panel is also seen on the right side of the image.

Elemental PMF versus Measured Plots

The plots on the PMFvsIBAplots worksheet (**Figure 22-13**) represent the correlation of each elements PMF fitted value against its measured concentration value obtained with accelerator ion beam analysis (IBA). These plots provide a clear visualisation, element by element, of which elements have been fitted well by the PMF process. Ideally, the obtained gradient and R^2 for each element would be 1 if the PMF fit was perfect. This linear fit can be skewed significantly by outlier points resulting in a solution that is representative of neither the majority of data for that element nor the outlying point, but somewhere in-between. Therefore, during the PMF analysis process, the data analyst is able to systematically remove outliers from these plots in order to get closer to this ideal condition. However, for trace elements or elements with very large associated errors/MDLs, this is not often possible nor even advantageous as the PMF process will always be driven by elements with more accurate and precise measurements. These plots are also userful during the PMF analysis process in determining which elements can be removed from the PMF analysis as they have minimal impact on the final fits.

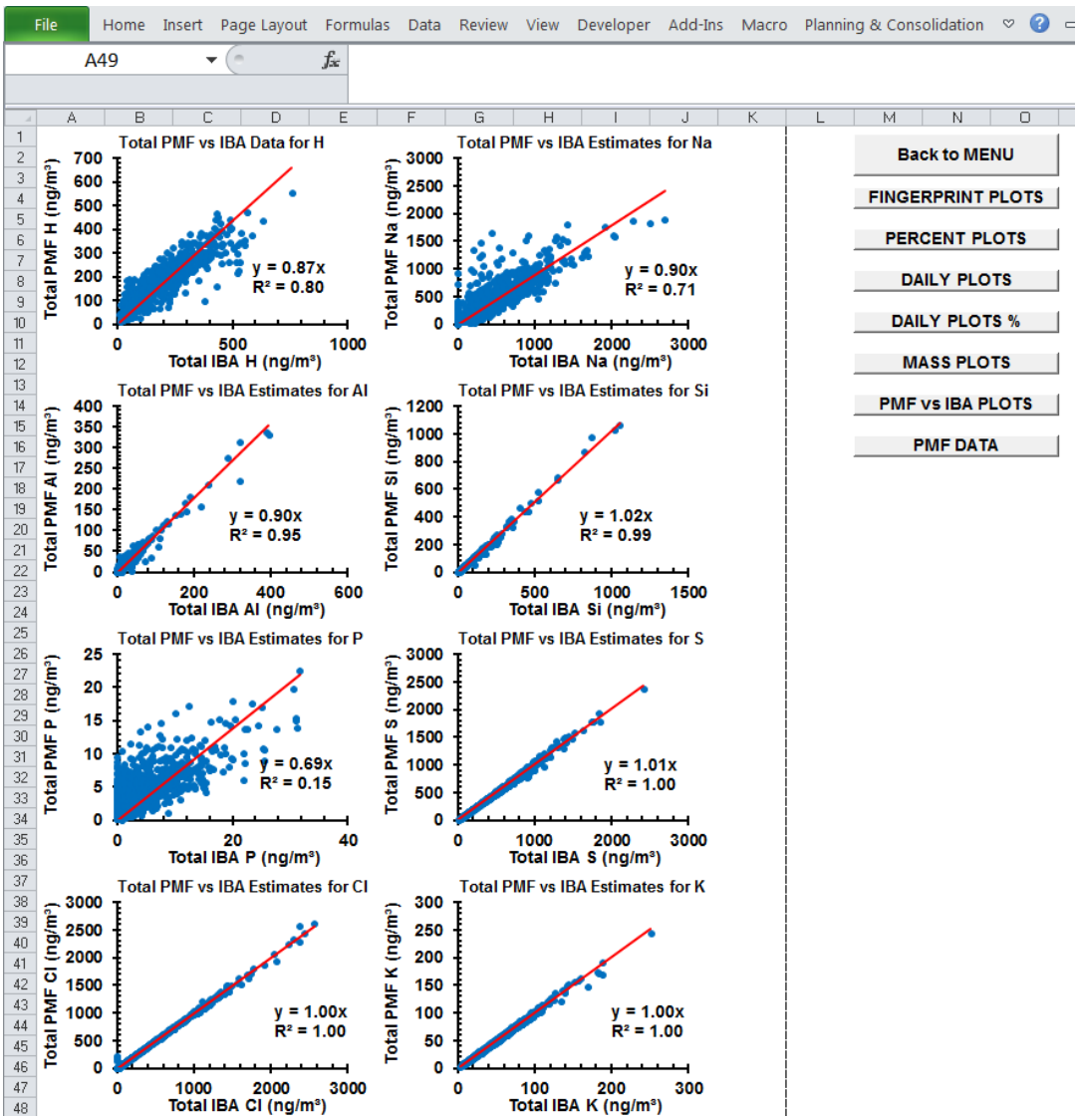

Figure 22-13. Example of plots on the PMF versus IBA plots worksheet comparing PMF fitted concentration (ng/m³) against the ion beam analysis (IBA) measured concentration (ng/m³) for each element. The worksheet navigation button panel is also seen on the right side of the image.

PMF Master Data

The data contained on the worksheet with the selected sampling site name (e.g. Lucas Heights) represents the PMF Master Data related to that site and from which all of the previously described plots are generated. The following section of the present document will describe the various sections of data located on the PMF Master Data worksheet using the excel cell reference method of letter and column, e.g. “E10” refers to column E and row 10 (see **Figure 22-14**).

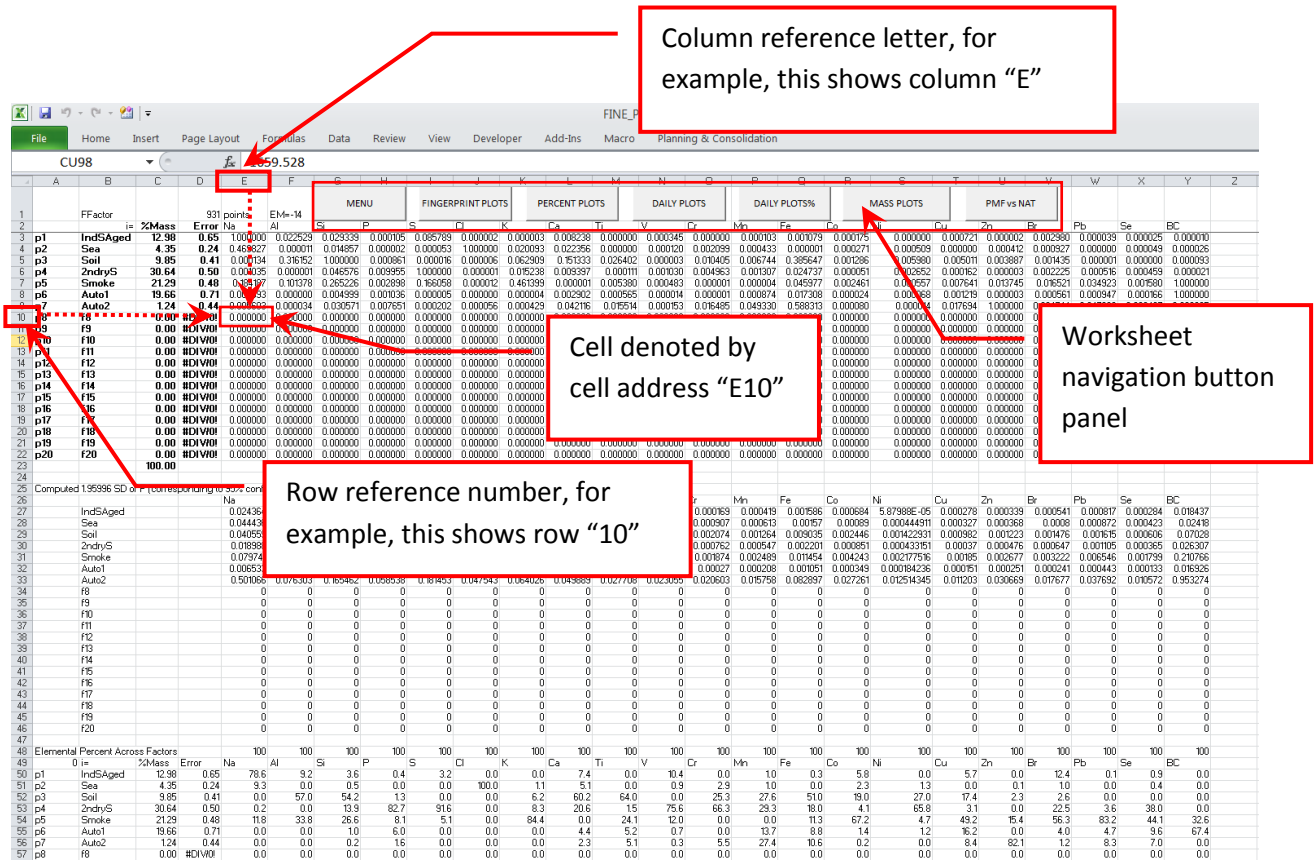

Figure 22-14. Example section of the PMF Data worksheet showing the cell reference method, for example cell E10, as show in the image. Please note: the worksheet navigation button panel is in a slightly different format/location to previous worksheets and is seen along the top of the worksheet.

Important Note

When viewing the PMF Master Data worksheet for each member state, it should be noted that the PMF anlaysis program which generates this Master Data worksheet is capable of processing up to 20 different PMF fingerprints, but in reality we would rarely use that many fingerprints. Most sites to date are adequately accomodated by between 5-10 fingerprints. Therefore, you will notice many cells containing either zero or #DIV/0! values as these are related to unused factors of the available 20.

F-matrix

Rows 3 to 22 (across): Data related to fractional ratio contribution of each element in each fingerprint. It represents the F-matrix. In this data, the element with the largest contribution in each fingerprint has been normalised to 1. The fingerprint names are listed in cells B3 to B22, with their corresponding percentage mass contributions and associated error listed in cells C3 to C22 and D3 to D22,

respectively. The elements included in the PMF analysis are listed in row 2 from column E and onwards across the worksheet (depending on how many elements were included in the analysis). This data is used to generate the fingerprint plots.

Cell D1: Number of points (i.e. days) included in the PMF analysis.

Rows 26 to 46 (across): Data related to calculated 95% confidence interval of the F-matrix data from column B and onwards across the worksheet depending on how many elements included in the analysis. This data is used to generate the error bars on each element in the figerprint plots.

Rows 48 to 69 (across): Data related to the percentage distribution of each element across all of the fingerprints from column A and onwards across the worksheet depending on how many elements included in the analysis. Therefore, the sum of each element's percentage contribution in each fingerprint, e.g. cells E50 to E69, results in a value of 100 percent as show in row 48 from columns E onwards (depending on how many elements were included in the analysis)

Rows 70 to 90 (across): Data related to manual Fkey pull-down strength used during the PMF analysis from column A and onwards across the worksheet depending on how many elements included in the analysis. Column B70 to B90 lists the fingerprint names related to that row of data. Columns E (for rows 70-90) onwards relate to the elements in each fingerprint (as listed in row 2). For example, cell E71 would refer to the Fkey value applied to the first element in the first fingerprint.

The optional Fkey option available during PMF analysis refers to the use of known external information to impose additional control on the rotation of the PMF analysis. For example, if specific elements in specific values are **known** to be zero, such as black carbon (BC) in a pure sea spray fingerprint, then assigning an Fkey value to that element can be used to force the PMF solution toward zero for BC values in that sea spray fingerprint. A range of Fkey pulldown strength may have been used between one, which has minimal pull-down strength, to nine which has the strongest Fkey pulldown strength. A value of zero, denotes no Fkey pull-down has been used.

Cell A96: Shows the year range of the dataset used in the PMF analysis

G-matrix

Cells A96 to AA97 (down): Data related to daily contribution of each fingerprint to total mass. This data represent the G-matrix. The names of the fingerprints are listed horizontally from cell F96 to Y96 with their data listed in the corresponding column below each name. Cells A97 to D97 (downward depending on the number of analysed days) relate to the corresponding Site, Day, Month and Year of each G-matrix data row. Column E97 and Z97 relate to the graviemetric mass concentration (Cmass) and fitted PMF mass concentration (FitCmass), respectively. Column AA97 relates to the date (DD-MMM-YY) used in several of the plot worksheets described earlier.

Rows E92 and E92 (across): Data related to the average (mean) and standard deviation values, respectively, for the columns in "**Cells A96 to AA97 (down):**" described above.

Cell E94: Q-value obtained from analysis

Cell G94: Fpeak value used for analysis

Cell J94: Chi-squared value obtained from analysis

Cell M94: Seed value used for analysis

Row E95 to AA95 (across): Least squares linear regression coefficient values corresponding to each of the G-matrix columns from F98 to Y98 (down).

Rows AB96 AU96 (down): Daily percentage of each fingerprint to the total PMF fitted mass concentration (i.e. FitCmass). Summation of each row, e.g. AB98 to AU98 which represents the individual percentage contribution of each fingerprint for that day, should equal a value of 100%.

Tramlines

Columns AW97 to AZ97 (down): Data used to create the standard deviation tramlines on the PMF mass vs. Observed Mass plot (see MassPlots worksheet, **Figure 22-12**).

Cell BA97: Standard deviation tolerance for the upper and lower tramlines

Cell BB97: Standard deviation of the values listed in the "Distances" column A97 (down)

Additional Functions - Interactive Sampling Site Map

*NB: this function requires **internet connection** and the "SamplingSiteMap.html" file must be located in the same folder as the Database file.*

Clicking the interactive map button on the Menu worksheet accesses generates a map of the sampling sites which it displays in a new internet browser window using Google Maps.

NB: If this function fails to open correctly from within the database, please try opening the SamplingSiteMap.html file directly using an internet browser other than Internet Explorer (.e.g. Chrome).

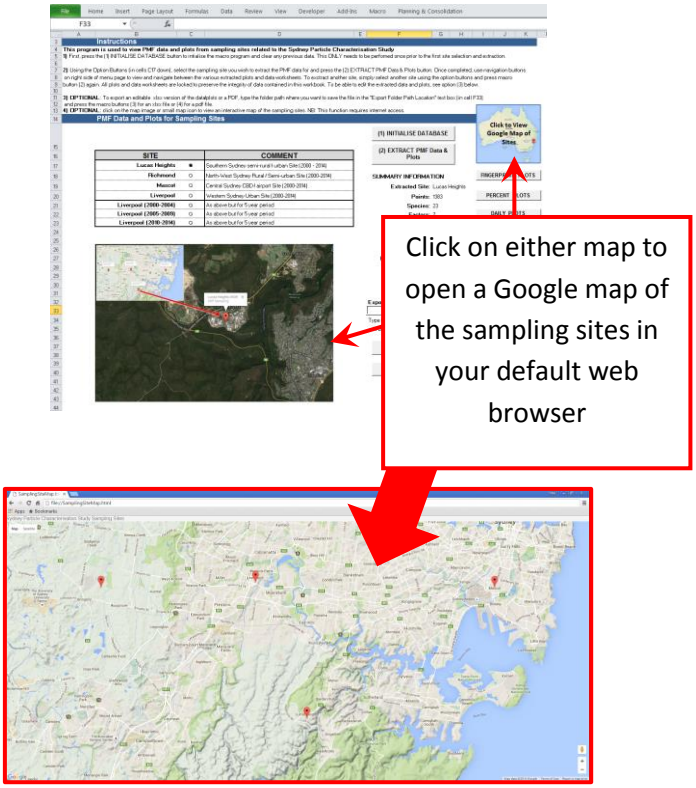

Figure 22-15. Location of hyperlinks that will open an interactive Google map showing sampling site locations in your internet browser. NB: This function requires an active internet connection.

23. Appendix C-

Cohen et al. Atmos. Environ. 61 (2012) 204-211.

Atmospheric Environment 61 (2012) 204–211

Contents lists available at SciVerse ScienceDirect

Atmospheric Environment

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

Application of positive matrix factorization, multi-linear engine and back trajectory techniques to the quantification of coal-fired power station pollution in metropolitan Sydney

David D. Cohen*, Jagoda Crawford, Eduard Stelcer, Armand J. Atanacio

Australian Nuclear Science and Technology Organisation, Locked Bag 2001, Kirrawee DC, NSW, 2232, Australia

HIGHLIGHTS

- ► IBA methods have characterized an 11 year fine particle pollution dataset.
- ► PMF, ME methods combined with back trajectory data identified coal-fired fingerprints.
- ► Quantitative fingerprints and their contributions have been determined.
- ► Coal-fired power stations contribute between 30 and 50% to the total fine sulfate mass.

ARTICLE INFO

Article history:
Received 14 April 2012
Received in revised form
16 July 2012
Accepted 16 July 2012

Keywords:
Ion beam analysis
Fine particles
PIXE
Factor analysis
HYSPLIT

ABSTRACT

Over 900 fine particle Teflon filters were collected within the Sydney Basin between 1 January 2001 and 31 December 2011 and analyzed using simultaneous PIXE, PIGE, RBS and PESA techniques to determine 21 different elements between hydrogen and lead. These elements were used in positive matrix factorization (PMF) and multi-linear engine (ME) techniques together with HYSPLIT wind back trajectory techniques to quantitatively determine source fingerprints and their contributions from coal-fired power stations. The power stations were many kilometers outside the greater Sydney metropolitan area but still had a significant impact on the fine particle mass loadings measured at the sampling site within this metropolitan area. The PM2.5 eleven year average mass at the sampling site was 6.48 μg m^{-3}. The corresponding ammonium sulfate estimate was 1.65 μg m^{-3} or 26% of the PM2.5 mass. By applying back trajectory data and (ME) analysis methods, two power related fingerprints, secondary sulfate (*2ndrySPower*) and aged industrial sulfur (*IndSagedPower*) were determined. These two power related fingerprints were responsible for between 14 and 18% of the total PM2.5 mass and 34–47% of the total sulfate measured at the sampling site. That is on average somewhere between a third and a half of all the sulfate measured in the greater Sydney region could be attributed to coal-fired power station emissions.

1. Introduction

Globally atmospheric fine particle pollution has significant impacts on human health, visibility, climate change and can be transported many hundreds of kilometers every day. Over the last decade studies of fine particle pollution have moved from just chemical characterization to using statistical techniques such as positive matrix factorization (PMF) and multi-linear engine (ME) methods to not just determine source elemental fingerprints but also quantify source contributions to the total measured fine particle mass (Paatero and Tapper, 1994, 1997, 1999, 2004, 2009). More recently these PMF and ME source apportionment techniques have been combined with wind back trajectory data to tie known emitters more directly to measured sources at the receptor sites (Yli-Tuomi et al., 2003; Chan et al., 2011; Cohen et al., 2010a, 2010b, 2011, 2012 and the references therein).

Coal-fired power stations are known emitters of fine particles (PM2.5) and pollutant gases such as SO_x and NO_x. It is therefore important to better understand their contributions to atmospheric pollution especially in large urban areas. In this paper we report, for the first time, on the contributions of eight coal-fired power stations, burning over 25 MT yr^{-1} of low grade sulfur coal, to the PM2.5 mass loading in the greater Sydney metropolitan area. We apply and compare PMF and ME techniques for quantitative source

* Corresponding author.
E-mail address: dcz@ansto.gov.au (D.D. Cohen).

1352-2310/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.atmosenv.2012.07.037

apportionment and elemental fingerprint identification and use ME methods that include wind back trajectory data to uniquely tie these known sources, external to the Sydney metropolitan region, to fine particle measurements made within this region.

The new techniques described here can be applied generally to fingerprint any known point sources and quantify their contributions to the total measured fine particle mass at a receptor site hundreds of kilometers away.

2. Study site and local conditions

New South Wales is one of six Australian States on the east coast of Australia, see Fig. 1. Sydney is on the coast, it is the largest city with over 4.5 million people and nearly 3 million motor vehicles. The locations of the eight coal-fired power stations supplying Sydney's electricity are shown in Fig. 1, their longitude, latitude and distance to the sampling site at Richmond are given in Table 1. These stations generally lie well outside the greater Sydney metropolitan region which is typically within 40 km of the central business district (CBD). The boxes used to represent these power stations in Fig. 1 are ±0.1°, or ±10 km, about each power station's longitude and latitude location and have been used to define back trajectory intersections above power stations.

The average sulfur content of coal burnt for power generation in NSW is around 0.5% by weight. The eight power stations listed in Table 1 represent 243 kT yr^{-1} of sulfur dioxide emissions across NSW. Power stations are by far the largest sulfur dioxide emitters in NSW, other significant point source emitters include aluminum production (11 kT yr^{-1}), steel production (9 kT yr^{-1}), petrol refining (5 kT yr^{-1}) and diffuse sources such as motor vehicles with around 20 kT yr^{-1}. This suggests that, in the State of NSW, of the total 290 kT yr^{-1} of sulfur dioxide emissions produced each year over 80% are from coal-fired power stations.

The sampling site is located at Richmond (−33.618°S, 150.748°E) in an open grassed area in the grounds of the University of Western Sydney. It was selected to be on the outskirts of the urban-residential areas and bordering on the rural zone west of Sydney. The site is influenced by both sources from within the Sydney CBD region and sources external to the greater metropolitan area such as the power stations. In the summertime afternoon on-shore sea breezes transport inner Sydney pollution to the Richmond site. In the winter time westerly breezes bring pollution from the Lithgow

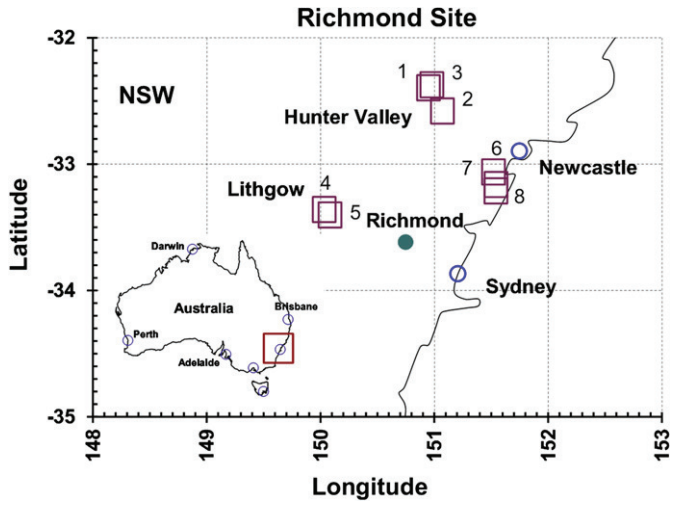

Fig. 1. Location of Richmond site (●) in NSW and the eight associated power stations (boxes) whose locations are given in Table 1. The box sizes are ±1°, or ±10 km, about each location.

Table 1
Location of the power stations, their distance from the Richmond sampling site and the total tonnage of coal burnt annually together with estimated SO_2 emissions each year.

	Name	Lat (°S)	Long (°E)	Distance (km)	Coal burnt (kT yr^{-1})	SO_2 emissions (kT yr^{-1})
1	Bayswater	−32.3953	150.9491	138	7100	69.6
2	Redbank	−32.5784	151.0345	121	200	1.7
3	Liddell	−32.3719	150.9783	142	4900	47.9
4	Mount piper	−33.3934	149.9705	67	4000	39.0
5	Wallerawang	−33.4040	150.0845	73	2100	21.0
6	Eraring	−33.0623	151.5214	96	3800	37.0
7	Vales point B	−33.1596	151.5428	91	2000	20.0
8	Munmorah	−33.2077	151.5393	87	700	6.8

power station sites (4 and 5 on Fig. 1) into the Richmond site. We consider this site to be a suburban site.

A standard Aerosol Sampling Program (ASP) cyclone unit based on the USA IMPROVE system was operated at 22 l min^{-1} for 24 h (mid-night to mid-night) every Wednesday and Sunday at the Richmond site to collect PM2.5 particles on a thin stretched Teflon filters. Nine hundred and twelve filters were collected during the 11 year study period from 1 January 2001 to 31 December 2011. The site was 53 km northwest of the Sydney CBD. The average daytime summer and winter temperatures were 16 °C and 26 °C respectively. The long hours of sunlight, particularly in the summer, together with the high humidity were ideal for the conversion of SO_2 gas to sulfate particles.

3. IBA chemical composition

Ion beam analysis (IBA) techniques have been used for many decades to determine chemical composition of fine particulate matter. They are ideally suited to this task as they have the sensitivity to non-destructively measure elements from hydrogen to lead with concentration down to nanograms per cubic meter of air sampled. Four simultaneous techniques of PIXE, PIGE, RBS and PESA (Cohen et al., 1996, 1998, 2004a,b) together with laser absorption methods for black carbon analysis (Taha et al., 2007) were employed to determine 21 different element species between hydrogen (H) and lead (Pb), presented in Table 2.

Table 2 below gives the PM2.5 average, standard deviation (SD), median, maximum, minimum detectable limits (MDL) and typical errors for our IBA analysis of the Teflon filters collected at the Richmond site during the study period.

The standard deviations on the data are much larger than the typical experimental measurement errors because they have large seasonal variations. The source of typical IBA experimental errors given in Table 2 have been published previously (Cohen et al., 2002). The reconstructed mass (RCM: estimated according to Malm et al., 1994) represents the sum of all the masses of the analyzed major components (79%) and is generally less than the measured gravimetric mass (100%) because we did not measure nitrates and the water vapor content of the filters which made up this missing mass (21%). Organics, ammonium sulfate, soil and sea salt were also estimated according to the formalism of Malm et al. (1994).

4. PMF fingerprints and source contributions

For a dataset of sufficient size, the statistical approach of positive matrix factorization (PMF) can be applied to both identify source elemental fingerprints as well as quantify the contributions of these sources to the total PM2.5 concentration. PMF is a 2-way, bilinear

Table 2
Elemental concentrations (μg m^{-3}) from IBA results for 912 days for the Richmond site between 1 January 2001 and 31 December 2011.

Chemical species	Av. (μg m^{-3})	SD (μg m^{-3})	Median (μg m^{-3})	Max (μg m^{-3})	MDL (μg m^{-3})	Error (μg m^{-3})
H	0.254	0.214	0.194	2.301	0.008	0.012
Na	0.182	0.252	0.102	1.877	0.164	0.082
Al	0.019	0.029	0.009	0.276	0.006	0.0027
Si	0.067	0.084	0.041	0.779	0.003	0.0028
P	0.007	0.008	0.004	0.078	0.003	0.0014
S	0.401	0.315	0.304	2.750	0.003	0.018
Cl	0.131	0.217	0.039	2.066	0.003	0.0040
K	0.059	0.058	0.040	0.594	0.002	0.0025
Ca	0.016	0.012	0.013	0.096	0.002	0.0011
Ti	0.003	0.003	0.002	0.024	0.001	0.0005
V	0.0003	0.0003	0.0002	0.002	0.0005	0.0004
Cr	0.0005	0.001	0.0004	0.023	0.0005	0.0004
Mn	0.001	0.001	0.001	0.007	0.0007	0.0003
Fe	0.027	0.023	0.021	0.256	0.001	0.0015
Co	0.0003	0.0004	0.0002	0.007	0.002	0.0007
Ni	0.0004	0.001	0.0002	0.030	0.0005	0.0004
Cu	0.002	0.002	0.001	0.024	0.0005	0.0005
Zn	0.006	0.007	0.004	0.086	0.0005	0.0008
Br	0.002	0.002	0.002	0.013	0.004	0.0018
Pb	0.005	0.005	0.003	0.038	0.007	0.0035
BC	0.824	0.454	0.710	3.345	0.029	0.055
Mass	6.484	4.267	5.482	40.312	0.16	0.20
Ammonium sulfate	1.654	1.300	1.254	11.344	0.012	0.074
Soil	0.304	0.334	0.203	3.090	0.015	0.014
Organics	1.715	2.234	0.969	24.137	0.064	0.083
Sea salt	0.462	0.640	0.259	4.768	0.41	0.208
RCM	5.040	3.164	4.295	32.321	0.030	0.25
RCM%	79	12	78	146		

factor analysis problem and, following Paatero and Tapper (1994), can be represented as:

$$X = GF + E \quad (1)$$

or

$$x_{i,j} = \sum_{k=1}^{p} g_{i,k} f_{k,j} + e_{i,j} \quad (2)$$

where the matrix X contains the measured quantities, i.e. $x_{i,j}$ represents the concentration of chemical species j in the ith sample. Matrices G and F are factor matrices to be determined and E is the error matrix of residuals. If n observations are available, each containing m chemical species and if a p-factor model is being considered, G is an n x p matrix of source contributions, describing the temporal variation of the source strengths. The matrix F is a p by m matrix of source chemical compositions, or source fingerprints.

The objective of PMF is to minimize the penalty function Q under the constraints that the factor elements remain non-negative:

$$Q_{\text{main}} = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{e_{i,j}^2}{s_{i,j}^2} \quad (3)$$

$$Q_{\text{expt}} = nm - p(n+m) \quad (4)$$

where $e_{i,j}$ are the error terms in equ. 2 and $s_{i,j}$ is a specified experimental error for each data value (i,j) in the X matrix. The following form of $s_{i,j}$ is used:

$$s_{i,j} = \text{MDL}_{i,j} + \text{Error}_{i,j} \max\left(|x_{i,j}|, |y_{i,j}|\right) \quad (5)$$

where, $\text{MDL}_{i,j}$ is the specified Minimum Detectible Limit and $\text{Error}_{i,j}$ is the specified error to account for experimental error, peak area determination, counting and statistical error and calibration error, and $y_{i,j}$ is the fitted value for the (i,j) element of the Y matrix given by G*F. The analysis of the solution includes techniques reported by Paatero and Tapper (1994), Paatero (1997, 2004), including: analysis of the penalty function Q as well as examining the G matrix for correlations between factors.

In the current analysis there were 912 sampling days ($n = 912$) and 21 different elemental concentrations ($m = 21$). Thus, with seven factors or sources ($p = 7$) an expected value of $Q_{\text{expt}} = 12{,}621$, according to equ. 4 (Paatero and Tapper, 1994), the final PMF solution had a. $Q_{\text{main}} = 6608$. The fit to the 912 days was excellent with the PMF(Mass) = (0.986 ± 0.12)*Gravimetric(Mass) with an $R^2 = 0.951$ and the 7 factor fit to the total mass having a P-value, $p < 0.01$ for each of the seven factors. That is the total PM2.5 gravimetric mass could be explained to within 2% by these 7 fingerprints and each elemental fingerprint was significantly determined at the 99% confidence interval in this PMF model. We also considered 6 and 8 factor solutions, but considered the 7 factor solution as optimal for this dataset in terms of standard PMF criteria (as mentioned above) and our ability to associate factors with known sources in the region.

The 7 factors or fingerprints are discussed below and plotted in Fig. 2.

Factor 1-2ndryS: This was the secondary sulfate fingerprint, produced by the conversion of SO_2 gas to the particulate sulfate phase in the presence of sunlight and water vapor (Seinfeld and Pandis, 1998). It had the correct hydrogen (H) to sulfur (S) ratio [H/S] = 4 for fully neutralized ammonium sulfate $(NH_4)_2SO_4$. SO_2

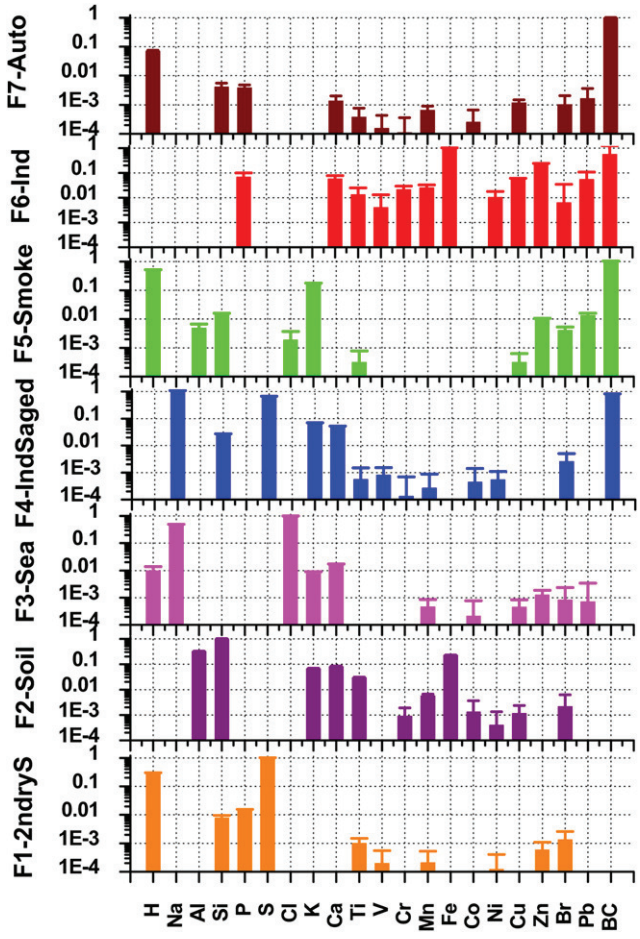

Fig. 2. The seven factors (fingerprints) obtained from the standard PMF analysis and the F matrix of equ. 1.

originates from motor vehicles, coal-fired power stations and other industrial activities. This factor shows strong seasonal variations with higher concentrations in summer (Fig. 3) as expected.

Factor 2-Soil: This factor contains the five key elements (Al, Si, Ca, Ti and Fe) commonly associated with windblown soils (Wu et al., 2009; Amato and Hopke, 2012). It also has typical [Al/Si] ratios for common alumina-silicates (Cohen et al., 2009, 2010) between (0.25–0.35).

Factor 3-Sea: This factor is dominated by Na and Cl, with small amounts of Br which is indicative of sea spray (e.g. Qin et al., 1997; Wu et al., 2009). The [Cl/Na] ratio was higher than the expected ratio for fresh sea salt (1.54) but this could be due to an underestimate of Na or an excess of Cl in this source. The latter is more probable as the Richmond site is at least 50 km from the coast and sea air passes over the Sydney CBD before reaching this inland site picking up other Cl sources such as motor vehicle exhaust.

Factor 4-IndSaged: This factor is dominated by Na, S and BC. The Na to S ratio was indicative of Na_2SO_4 probably formed by chemical reactions of sea spray particles with sulfate particles as seen by others (Qin et al., 1997; Wu et al., 2009) hence we call it aged industrial sulfur.

Factor 5-Smoke: This factor was dominated by high H from organic sources, K and BC from biomass burning (Yli-Tuomi et al., 2003; Lee et al., 2008, Cohen et al., 2010a,b), with traces of soil which represented biomass burning and vegetation burning common in domestic heating in the winter time and bush fire smoke in the summer months. This factor shows a strong seasonal component with high concentrations in winter and correlating with known domestic wood heater emissions during the colder winter months.

Factors 6-Ind: This factor was dominated by BC, Fe and Zn with the time series having sharp peaks, indicative of possible industrial emissions like metal smelting or processing. The concentration to the total mass from this factor was only 1.8% consistent with it being a suburban site influenced very little by industrial sources.

Factor7-Auto: This factor was dominated by H, BC and trace elements associated with motor vehicles such as Zn from tyre wear, P and Ca from engine oils and small amounts of Pb and Br associated with historic leaded petrol use. For this study we label this factor as Auto. The contribution to the total mass from this fingerprint (11.2%) was also consistent with the motor vehicle use in the vicinity of the Richmond site.

The time series plots for the PM2.5 mass (in μg m^{-3}) together with the percentage mass contributions for each of the 7 fingerprints are given in Fig. 3. The average percentage contributions to the total PM2.5 mass for each of these standard PMF sources during the study period is given in Table 3. The two major sulfur sources, *2ndryS* and *IndSaged* accounted for 40% of the mass followed closely by *Smoke* from biomass burning with 37%. Table 4 shows that for the standard PMF analysis 73% of the measured total sulfur was associated with the *2ndryS* fingerprint and 27% with the *IndSaged* fingerprint. The remaining 4 fingerprints account for 23% of the fine mass but are not really the focus of the current study as they are generally associated with diffuse or distributed sources and not with the eight coal-fired power stations point sources which we want to quantify here.

It should be emphasized that this 7 factor standard PMF analysis has not used any back trajectory information at this stage. This aspect is considered in detail below.

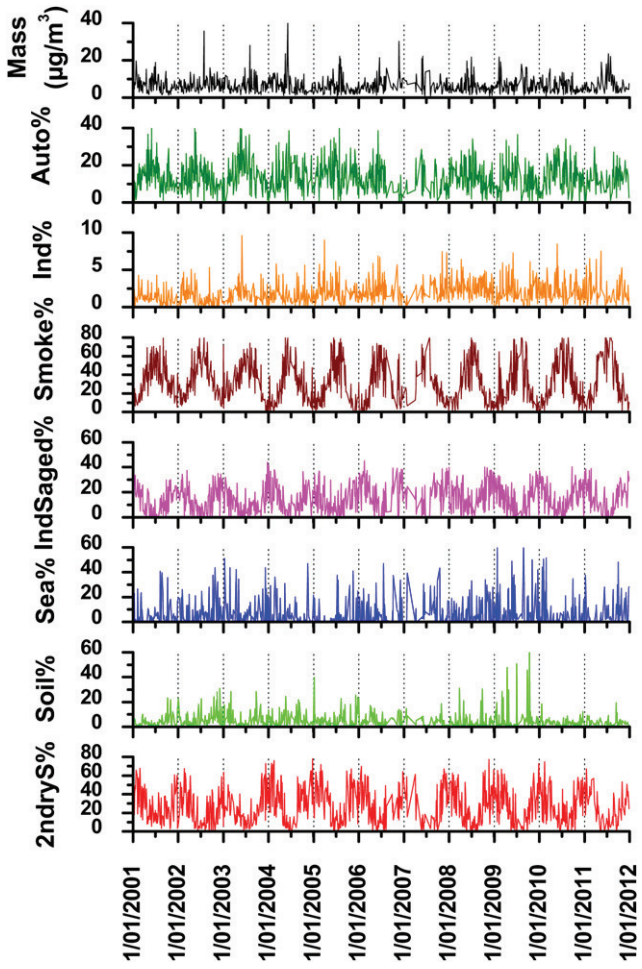

Fig. 3. Time series plots for the total PM2.5 mass and the percentage mass contributions of the 7 standard PMF factors or fingerprints.

5. Back trajectories

In recent publications (Cohen et al., 2009, 2010, 2011 and 2012) we described a new approach to the combination of IBA multi-elemental data and wind back trajectory data to tie known sources to the receptor site. We were able to quantify source contributions to long range transport of fine desert dusts from China into Vietnam and from Australian desert regions into urban areas of Sydney. We applied these same techniques, of linking source and receptor sites through back trajectories, here to look at the

Table 3
Percentage contributions to the total PM2.5 mass for each of the fingerprints in the standard PMF analyses and for *Scenarios 1* and *2* described below for the ME analyses.

%Fingerprint masses for different scenarios			
Fingerprint	Standard PMF	Scenario ME 9 factors	
	7 factors	1	2
2ndryS (total)	27.3 ± 0.6	25.3 ± 0.8	25.5 ± 0.8
2ndrySPower		12.7 ± 0.4	10.1 ± 0.3
2ndrySnoPower		12.6 ± 0.4	15.5 ± 0.5
IndSaged (total)	12.4 ± 0.8	14.3 ± 0.7	13.9 ± 0.8
IndSagedPower		5.60 ± 0.3	4.54 ± 0.3
IndSagednoPower		8.70 ± 0.4	9.39 ± 0.5
Soil	4.76 ± 0.4	4.90	4.77
Sea	5.54 ± 0.4	5.40	5.55
Smoke	37.1 ± 0.7	33.8	33.7
Industry	1.75 ± 0.7	1.50	1.71
Auto	11.2 ± 0.8	14.8	14.8
Total%	100 ± 5	100 ± 5	100 ± 5

Table 4
Percentage contributions to the total measured sulfur content in the fine PM2.5 fraction for each of the fingerprints in the standard PMF analyses and for *Scenarios 1* and *2* described below for the ME analyses.

Fingerprint	Sulfur contributions (%)		
	Standard PMF	Scenario ME 9 factors	
	7 factors	1	2
2ndryS (total)	72.9 ± 4	74.8 ± 4	74.8 ± 4
2ndrySPower		36.5 ± 2	26.0 ± 2
2ndrySnoPower		38.3 ± 2	48.8 ± 2
IndSaged (total)	27.1 ± 2	25.2 ± 2	25.2 ± 2
IndSagedPower		10.6 ± 0.8	8.1 ± 0.6
IndSagednoPower		14.6 ± 1	17.1 ± 1

transport of emissions from the eight coal-fired power stations listed in Table 1 into the Richmond sampling site.

As before, hourly back trajectory data for each 24 h sampling day (from mid-night to mid-night) was taken from the HYSPLIT model of Draxler (1991) and Draxler and Rolph, 2003. For the study period from 1 January 2001 to 31 December 2011 there were 912 sampling days corresponding to a total of 21,888 hourly back trajectories. The back trajectories were defined by their spatial position ($long_i$, lat_i) at 30 min intervals for five days back from the receptor site. That is 240 separate spatial locations determined each of the 24 back trajectories for each sampling day. For a 300 m starting height above the Richmond site there were 3898 back trajectories (or 18%) on 490 days for which at least one back trajectory passed over a power station. The technique has been previously detailed in Cohen et al., (2010a,b, 2011 and 2012), in summary, a rectangular region, centered on the power station ($long_0$, lat_0), was defined, contributions from the power station were then considered possible for a specified hour of the day, if the corresponding back trajectory passed over the spatial region defined by the rectangle, i.e. if $(long_0 - 0.1^\circ) \leq long_i \leq (long_0 + 0.1^\circ)$ and $(lat_0 - 0.1^\circ) \leq lat_i \leq (lat_0 + 0.1^\circ)$, for at least one i, $i = 1,\ldots,240$. In this region 0.1° of both longitude and latitude corresponds to approximately 10 km.

Table 5 shows the percentage breakdown of hourly back trajectories passing over each of the eight power stations shown as boxes in Fig. 1 for a sampling site back trajectory starting height of 300 m.

For events with sulfur greater that zero, the western power stations near Lithgow have the highest percentage of back trajectories originating from the Richmond sampling site, followed closely by the three coastal power stations with a larger gap back to the northern Hunter Valley power stations.

For the three Northern and three Coastal power stations the percentage of intersecting back trajectories increases (by around a factor of 2) with increasing daily sulfur concentrations while for the closer Western power stations the percentage of intersecting back trajectories remains approximately constant at between 5% and 7%.

The wind speed above the power stations relates critically to the size of the boxes used to define the influence of a power station (Fig. 1). If the wind speed is too high and the box is too small the analysis code may not provide an intersection of the trajectory with the power station's defining box. For the starting height of 300 m at Richmond Fig. 4 shows a histogram of the wind speeds above each of the eight power stations.

Fig. 4 shows that 92% of the wind speeds above all eight power stations were below 40 km h^{-1}. This corresponds to power station box sizes of at least 20 × 20 km for the 30 min time intervals used here along each back trajectory. That is the boxes should be no less than (±0.1°) around each longitude and latitude position of each power station as each 1° longitude and latitude corresponds to approximately 100 km. These arguments explain why we used ±10 km boxes to represent each power station.

It was also important to consider the height at which each back trajectory passed over each power station. If this was too high and well outside the mixing layer then the power station was less likely to impact the Richmond site. From the HYSPLIT code we estimated that 92% of the trajectories pass over power stations at heights below 700 m well within the average mixing layer depths provided by HYSPLIT and consistent with typical stack heights of around 100 m for emissions from these power plants. Given this, all the back trajectories were used in the analysis and none were excluded.

Our analytical method (Cohen et al., 2012) relates not just the back trajectories passing over source receptor sites but also the source contributions on that day at the receptor site to determine the relative contributions of a given source to the total fine particle mass. Table 5 shows that the number of back trajectory intersections from each of the power stations generally increases when the sulfur concentration increases. Only the two western power stations show a small decrease. That is the higher the daily sulfur concentration at the Richmond site the more influence the northern and coastal power stations had while the western power station influence on the Richmond site was relatively constant.

The 7 factor standard PMF analysis produced two sulfur containing fingerprints, *2ndryS* and *IndSaged* both of which were a mixture of several possible sources. Linking the 7 fingerprints with HYSPLIT back trajectory data provided an estimate of between

Table 5
Hourly back trajectories from the Richmond sampling site with daily sulfur concentrations $S > 0$, 0.4 (mean), 0.7 (mean + one standard deviation) μg m^{-3} and a starting height at the sampling site of 300 m passing over the eight power stations during the study period, expressed as a percentage of the total number of back trajectories corresponding to sulfur concentrations satisfying the conditions.

Power stations	Name	%Trajectory intersections	%Trajectory intersections	%Trajectory intersections
		$S > 0$ μg m^{-3}	$S > 0.4$ μg m^{-3} (mean)	$S > 0.7$ μg m^{-3} (mean + 1 standard dev.)
Northern	1 Bayswater	1.8	3.5	4.2
	2 Redbank	2.5	4.9	5.7
	3 Liddell	1.8	3.4	4.2
Western	4 Mount Piper	6.3	5.5	5.5
	5 Wallerawang	7.0	6.5	6.6
Coastal	6 Eraring	5.4	9.3	10.7
	7 Vales Point B	5.4	9.3	10.1
	8 Munmorah	5.4	9.3	10.3

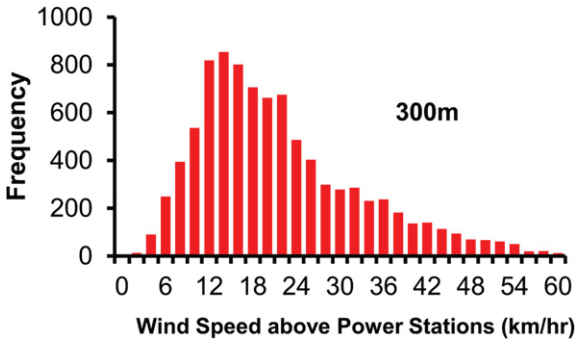

Fig. 4. Histogram of wind speeds for each hour back trajectory that passes over a power station during the study period.

3% and 9% of the days with sulfur above the mean concentration for which the eight power stations contributed to these two fingerprints. It did not provide a power station fingerprint or a breakdown of how much of *2ndryS* or *IndSaged* sources were power station related. To achieve this we apply more sophisticated ME analysis to our dataset.

6. Combining back trajectories and PMF in ME

We now use the multi-linear engine (ME) approach to split the two sulfur fingerprints *2ndryS* and *IndSaged* obtained in the standard PMF analysis into two components each. The first component being contributions with back trajectories from Richmond which passed over at least one of the eight identified power stations (*Power*) and the second component being contributions that did not pass over any power stations (*noPower*). This had the effect of turning the seven standard PMF fingerprints of Fig. 2 into nine fingerprints which were then analyzed by ME techniques under the two different scenarios described below.

Scenario 1. Four sulfur fingerprints *2ndrySPower*, *2ndrySnoPower*, *IndSagedPower* and *IndSagednoPower* were resolved which were forced to have the elemental fractional concentrations close to that of their standard PMF fingerprints above (i.e. *2ndrySPower* and *2ndrySnoPower* were to be close to the *2ndryS* fingerprint resolved by the standard PMF and similarly *IndSagedPower* and *IndSagednoPower* were to be close to *IndSaged)*. The remaining five fingerprints, *Soil*, *Sea*, *Smoke*, *Auto1* and *Auto2* were allowed to vary at will within the nine factor ME analysis.

This was achieved by specifying an auxiliary equation known as a 'pulling equation' (Amato et al., 2009, Paatero and Hopke, 2009; Amato and Hopke, 2012):

$$Q^{k}_{\text{aux}} = \sum_{j=1}^{m} \frac{\left(f_{kj} - a_{kj}\right)^2}{\sigma^2_{kj}} \tag{6}$$

where the elements of the four source factors in the F matrix, f_{kj} are pulled to the target value of a_{kj} $k = 1$, 4, and $j = 1$, m (m being the number of chemical species), and σ_{kj} is the uncertainty, which was set to $0.1a_{kj}$. This allowed only minor changes in the element fractions from their respective PMF sulfur factors. In ME the penalty function Q^{k}_{aux} for the auxiliary equations is then added to the penalty function for the main equations (Q_{main} in equ. 3) for the optimization procedure, i.e.

$$Q_{\text{tot}} = Q_{\text{main}} + \sum_{i=1}^{q} Q^{i}_{\text{aux}} \text{ where } q \text{ is the number of auxiliary equations.} \tag{7}$$

Aside from the pulling equations on the F matrix, the columns in the G matrix of equ. 2 corresponding to the four sulfur fingerprints also had to be modified:

- For the *Power* case, entries in the corresponding G matrix column were pulled to zero if and only if there were <5 hourly back trajectories, out of the possible 24 hourly back trajectories, passing over any of the eight power stations on the day of sample collection.
- For the *noPower* case, entries in the corresponding G matrix column were pulled to zero if and only if there were 5 or more hourly back trajectories, out of the 24 hourly back trajectories, passing over any of the eight power stations on the day of sample collection.

This was implemented by defining a power station vector (PS), where $PS_i = 1$ if 5 or more hourly back trajectories (out of the 24) on the day of sample i passed over a power station, otherwise $PS_i = 0$. The corresponding column entry of the G matrix was pulled to zero, using the pulling equation:

$$Q^{i}_{\text{aux}} = g_{ik} \text{ when } PS_i = 0 \text{ for } k = 1,2 \text{ (columns corresponding to power station fingerprints)} \tag{8}$$

$$Q^{i}_{\text{aux}} = g_{ik} \text{ when } PS_i = 1 \text{ for } k = 3,4 \text{ (columns corresponding to non – power station sulfur containing fingerprints)} \tag{9}$$

For this scenario, this implies that if 5 or more back trajectories (out of 24) on any day passed over a power station then the power station alone contributed all the mass to the sulfur fingerprint otherwise it was a non-power station source.

Cut offs for 5 and 12 hourly back trajectories in a given 24 h day were also tested in this ME analysis scenario and the final elemental fingerprint results varied little. So 5 was selected as 12 or more back trajectories passing over power stations per day (i.e. 50% of the time) gave fewer statistics for ME analysis and hence less reliable results on the power station fingerprint contributions.

Scenario 2. *Scenario 2* was implemented by using the same four sulfur containing fingerprints identified in *Scenario 1* with the same fractional elemental concentrations. The columns in the G matrix of equ. 2 corresponding to the four sulfur fingerprints were again modified:

- For the *Power* case, entries in the corresponding G matrix column were pulled to zero if and only if there were no back trajectories out of the possible 24 passing over any of the eight power stations on the day of sample collection.
- For the *noPower* case, there were no constraints on the G matrix contributions.

This was implemented by defining a power station vector (PS), where $PS_i = 1$ if 1 or more hourly back trajectories (out of the 24) on the day of sample i passed over a power station, otherwise $PS_i = 0$. The corresponding column entry of the G matrix was pulled to zero, using the pulling equation:

$$Q^{i}_{\text{aux}} = g_{ik} \text{ when } PS_i = 0 \text{ for } k = 1,2 \text{ (columns corresponding to power station fingerprints)} \tag{10}$$

These constraints imply that if 1 or more back trajectories (out of the 24) on a day passed over a power station then ME was allowed to determine the contributions to the sulfur containing fingerprints from the power station sources as well as from non-power station sources. The main difference between *Scenario 2* and *1* is that only the corresponding elements in two columns (not four columns as in *Scenario 1*) of the G matrix were pulled to zero on days when no power station was passed. However, no constraints were imposed on the G matrix corresponding to the non-power sulfur containing fingerprints. It was assumed that ME in the factorization process would apportion the sulfur between the power station target factors and the non-power station sulfur containing target factors for those days on which there was a contribution from both power stations and non-power station sources. Again, for *Scenario 2*, the

remaining five fingerprints, *Soil*, *Sea*, *Smoke*, *Ind* and *Auto* were allowed to vary at will within the nine factor ME analysis model.

These two ME scenarios should give us nine source fingerprints, their contributions as well as a range of power and non-power sulfur source fingerprint masses contributing to the total PM2.5 mass loadings at the Richmond site. This is the unique aspect of this approach. They will enable us to not only define *Power* fingerprints but also to put upper and lower limits on the possible contributions these fingerprints (and consequently the power stations) make to the measured PM2.5 mass at the Richmond receptor site.

The four new sulfur containing fingerprints for *Scenario 1* (S1) and *Scenario 2* (S2) are presented in Figs. 5 and 6. In the standard PMF model black carbon (BC) and iron (Fe) were not present in the *2ndryS* fingerprint whereas in the ME model both of these elements were present, at the few percent or lower level relative to sulfur, in *2ndrySPowerS1* and *2ndrySPowerS2* fingerprints but not in the *2ndrySnoPowerS1 and 2ndrySnoPower2* case. In the PMF and ME modes aged industrial sulfur BC was present for both scenarios and for *Power* and *noPower* cases in all fingerprints.

Over a decade of experience has shown that the presence of BC or not in the secondary sulfate fingerprint depends on the sulfur sources being local or regional and whether or not it has had an opportunity to mix with known BC emitters. All power station related sulfur has traveled significant distances (see Table 1 and Fig. 1) and al ong quite different trajectories to reach the Richmond site. In particular, coastal power stations (6,7 and 8 in Fig. 1) in the summertime have trajectories traveling south along the NSW coast picking up sea salt, passing over the Sydney CBD region picking up motor vehicle emissions and then inland to Richmond. This is quite different to the western stations (4 and 5) which in the winter time have direct access via westerly winds to Richmond and only pass over rural areas. We therefore suspect these differing elemental fingerprints for secondary sulfate *Power* and *noPower* cases could actually lead to us being able to distinguish between the different groups of power stations and their seasonal contributions to pollution within the greater Sydney metropolitan region. This is beyond the scope of the present study which is to determine a generic coal-fired power station fingerprint for Sydney and to estimate its average annual contribution. A future publication will look at the more detailed seasonal and regional aspects of the power station contributions to Sydney's pollution at several sites and not just the Richmond site and attempt to distinguish between the contributions of the different coal-fired power stations.

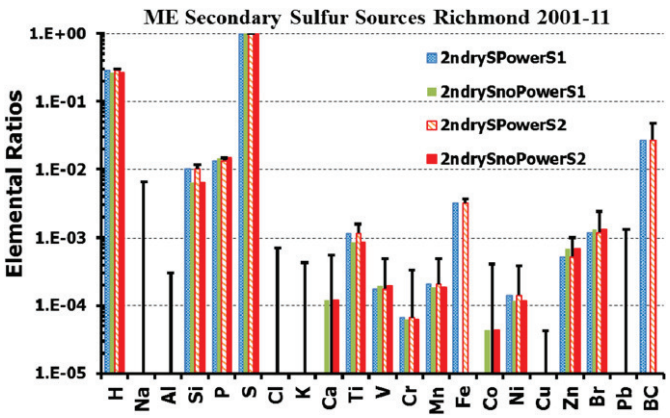

Fig. 5. ME secondary sulfur sources for *Scenario 1* (S1) and *Scenario 2* (S2) for days through power stations (*Power*) and days not through power stations (*noPower*) for all the Richmond PM2.5 data between January 2001 and December 2011. Typical 95% confidence interval error bars (vertical black error bars) are shown for each elemental species.

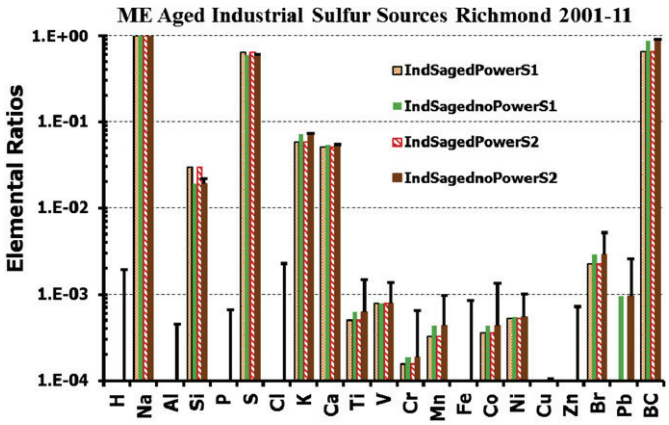

Fig. 6. ME aged industrial sulfur sources for *Scenario1* (S1) and *Scenario 2* (S2) for days through power stations (*Power*) and days not through power stations (*noPower*) for all the Richmond PM2.5 data between January 2001 and December 2011. Typical 95% confidence interval error bars (vertical black error bars) are shown for each elemental species.

Nevertheless, the ME based sulfur fingerprints for *Power* and *noPower* situations differ insignificantly for both the secondary sulfate and the aged industrial sulfate obtained in the standard PMF model. However the ME modeling does allow us to estimate the *Power* and *noPower* contributions to the total sulfate and the total aged industrial sulfate.

The percentage contributions of the four ME sulfur fingerprints in Figs. 5 and 6 and the remaining 5 non-sulfur fingerprints to the total PM2.5 mass and to the total measured sulfur concentration are given in Tables 3 and 4 respectively for both scenarios. These data allow us to quantitatively estimate the effects of the eight coal-fired power stations on the measured PM2.5 mass at the Richmond site and to compare these estimates across the three different modeling approaches namely, PMF and ME *Scenarios 1* and *2*.

In summary, the secondary sulfate fingerprint represented 25–27% of the PM2.5 mass with between 10 and 13% being associated with coal-fired power station emissions. For the aged industrial sulfur fingerprint with 12–14% of the PM2.5 mass between 4 and 6% was power related. Adding up these contributions we see that the average coal-fired power station contribution to the total PM2.5 mass at Richmond was between 14 and 18% during the study period. Furthermore, we now have calculated two coal-fired power station fingerprints, *2ndrySPower* and *IndSaged-Power*. The contributions of the non-sulfur fingerprints, *Soil*, *Sea*, *Smoke*, *Ind and Auto* varied little between the PMF and ME models or scenarios used. Together, they contributed around 60% to the total PM2.5 mass, being dominated by *Smoke* from biomass burning.

7. Summary

Fine particles (PM2.5) collected at Richmond, a rural-urban site on the outskirts of Sydney, over an 11 year period between January 2001 and December 2011were characterized by four simultaneous IBA techniques producing elemental concentrations for 21 different chemical species from hydrogen to lead. PMF techniques were then used to identify and quantify 7 different fingerprints. Two of these fingerprints, secondary sulfate and aged industrial sulfur were each split in two and, through HYSPLIT wind back trajectories, sorted into two groups those passing over coal-fired power stations and those not. ME techniques were then applied using two different scenarios to define fingerprints associated with coal-fired power

stations as well as quantifying their contributions to total PM2.5 mass loadings.

There are eight coal-fired power stations in NSW burning 25 MT yr^{-1} of coal and emitting 243 kT yr^{-1} of SO_2 which is converted by water and sunlight to sulfate particles. These power stations were many kilometers outside the greater Sydney metropolitan area but still had a significant impact on the fine particle mass loadings measured at the sampling site within the metropolitan area. The PM2.5 eleven year average mass at the Richmond sampling site was 6.48 μg m^{-3}. The corresponding ammonium sulfate estimate was 1.65 μg m^{-3} or 26% of the PM2.5 mass. The two PMF fingerprints, secondary sulfate (*2ndryS*) and aged industrial sulfur (*IndSaged*) accounted for 27% and 12% respectively, making a total of 39% of the PM2.5 mass. Using back trajectories to split these two fingerprints into power and non-power contributions and running an ME analysis over them we were able to produce two coal-fired power station fingerprints. These two power related fingerprints were responsible for between 14 and 18% of the total PM2.5 mass and 34–47% of the total sulfate measured at the sampling site. That is on average somewhere between a third and a half of the total sulfate measured in the greater Sydney region can be attributed to coal-fired power station emissions.

Further work is continuing to use these PMF, ME and back trajectory methods described here and long term data measurements from other sites in the Sydney metropolitan region to see if we can better quantify these coal-fired power station contributions and even determine which of the eight stations are the main contributors to sulfate pollution within the urban Sydney region.

Acknowledgments

The NOAA Air Resources Laboratory (ARL) made available the HYSPLIT transport and dispersion model and the relevant input files for generation of back trajectories used in this paper. We would like to acknowledge the help of staff at the University of Western Sydney with fine particle sample collection and the ANSTO accelerator operations staff for running the accelerators and the aerosol labs throughout this study.

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.

Amato, F., Hopke, P.K., 2012. Source apportionment of the ambient PM2.5 across St. Louis using constrained positive matrix factorisation. Atmospheric Environment 46, 329–337.

Chan, Y., Hawas, O., Hawker, D., Vowles, P., Cohen, D.D., Stelcer, E., Simpson, R., Golding, G., Christensen, E., 2011. The use of multivariate type composition data and wind data in the source apportionment of air pollutants. Atmospheric Environment 45, 439–449.

Cohen, D.D., Bailey, G.M., Kondepudi, R., 1996. Elemental analysis by PIXE and other IBA techniques and their application to source fingerprinting of atmospheric fine particle pollution. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 109, 218–226.

Cohen, D.D., 1998. Characterisation of atmospheric fine particles using IBA techniques. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 136, 14–22.

Cohen, D.D., Siegele, R., Orlic, I., Stelcer, E., 2002. Long term accuracy and precision of PIXE and PIGE measurements for thin and thick sample analyses. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 189, 81–85.

Cohen, D.D., Stelcer, E., Hawas, O., Garton, D., 2004a. IBA methods for characterisation of fine particulate atmospheric pollution: a local, regional and global research problem. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 219, 145–152.

Cohen, D.D., Garton, D., Stelcer, E., Hawas, O., 2004b. Accelerator based studies of atmospheric pollution processes. Radiation Physics and Chemistry 71, 759–767.

Cohen, D.D., Stelcer, E., Santos, F.L., Prior, M., Thompson, C., Preciosa, C.B., Pabroa, P.C.B., 2009. Fingerprinting and source apportionment of fine particle pollution in Manila by IBA and PMF Techniques; A 7 year study. X-ray Spectrometry 38, 18–25.

Cohen, D.D., Crawford, J., Stelcer, E., Bac, V.T., 2010a. Characterisation and source apportionment of fine particulate sources at Hanoi from 2001 to 2008. Atmospheric Environment 44, 320–328.

Cohen, D.D., Crawford, J., Stelcer, E., Thu Bac, V., 2010b. Long range transport of fine particle windblown soils and coal-fired power station emissions into Hanoi between 2001 to 2008. Atmospheric Environment 44, 3761–3769.

Cohen, D.D., Stelcer, E., Garton, D., Crawford, J., 2011. Fine particle characterisation, source apportionment and long range dust transport into the Sydney Basin: a long term study between 1998 and 2009. Atmospheric Pollution Research 2, 182–189.

Cohen, D.D., Crawford, J., Stelcer, E., Atanacio, A., 2012. A new approach to the combination of IBA techniques and wind back trajectory data to determine source contributions to long range transport of fine particle air pollution. Nuclear Instruments and Methods in Physics Research B273, 186–188.

Draxler, R.R., 1991. The accuracy of trajectories during ANATEX calculated using dynamic model analysis versus rawinsonde observations. Journal of Applied Meteorology 30, 1466–1467.

Draxler, R.R., Rolph, G.D., 2003. Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT), Model. http://www.arl.noaa.gov/ready/hysplit4.html.

Lee, H.L., Park, S.S., Kim, K.W., Kim, Y.J., 2008. Source identification of PM2.5 particles measured in Gwangju, Korea. Atmospheric Research 88, 199–211.

Malm, W.C., Sisler, J.F., Huffman, D., Eldred, R.A., Cahill, T.A., 1994. Spatial and seasonal trends in particle concentrations and optical estimations in the US. Journal of Geophysical Research 99, 1347–1370.

Paatero, P., Tapper, U., 1994. Positive matrix factorisation: a non-negative factor model with optimal utilisation of error estimates of data values. Environmetrics 5, 111–126.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems 37 (1), 23–35.

Paatero, P., 1999. The multilinear engine – a table-driven least squares program for solving multilinear problems, including the n-way parallel factor analysis model. Journal of Computational and Graphical Statistics 8 (4), 854–888.

Paatero, P., 2004. User's Guide for Positive Matrix Factorization Program PMF2 and PMF3, Part 1 and 2.

Paatero, P., Hopke, P.K., 2009. Rotational tools for factor analysis models. Journal of Chemometrics 23 (2), 91–100.

Qin, Y., Chan, C.K., Chan, L.Y., 1997. Characteristics of chemical compositions of atmospheric aerosols in Hong Kong: spatial and seasonal distribution. The Science of the Total Environment 206, 25–37.

Seinfeld, J.H., Pandis, S.N., 1998. Atmospheric Chemistry and Physics. In: From Air Pollution to Climate Change. John Wiley and Sons, Inc, USA. ISBN 0-471-1-17816-0.

Taha, G., Box, G.P., Cohen, D.D., Stelcer, E., 2007. Black carbon measurement using laser integrating plate method. Aerosol Science and Technology 41, 266–276.

Wu, C., Wu, S., Wu, Y., Cullen, A., Larson, T., Williamson, J., Liu, L.-J., 2009. Cancer risk assessment of selected hazardous air pollutants in Seattle. Environment International 35, 516–522.

Yli-Tuomi, T., Hopke, P.K., Paatero, P., Shamsuzzoha Basunia, M., Landsberger, S., Viisanen, Y., Paatero, J., 2003. Atmospheric aerosol over Finnish Arctic: source analysis by the multilinear engine and the potential source contribution function. Atmospheric Environment 37, 4381–4392.

Cohen et al. NIMB 318 (2014) 113-118.

Nuclear Instruments and Methods in Physics Research B 318 (2014) 113–118

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research B

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

The application of IBA techniques to air pollution source fingerprinting and source apportionment

D.D. Cohen *, E. Stelcer, A. Atanacio, J. Crawford

Australian Nuclear Science and Technology Organisation, Locked Bag 2001, Kirrawee DC, NSW 2232, Australia

ARTICLE INFO

Article history:
Received 22 February 2013
Received in revised form 22 April 2013
Accepted 27 May 2013
Available online 19 July 2013

Keywords:
IBA
PIXE
Air pollution
PMF
Long range transport
HYSPLIT

ABSTRACT

IBA techniques have been used to measure elemental concentrations of more than 20 different elements found in fine particle (PM2.5) air pollution. These data together with their errors and minimum detectable limits were used in Positive Matrix Factorisation (PMF) analyses to quantitatively determine source fingerprints and their contributions to the total measured fine mass. Wind speed and direction back trajectory data from the global HYSPLIT codes were then linked to these PMF fingerprints to quantitatively identify the location of the sources.

1. Introduction

Ion Beam Analysis (IBA) techniques have been used for decades to characterise fine particle air pollution. This is not new, the techniques are well established. Typically 2–3 MeV protons are used to bombard thin filter papers and up to four simultaneous techniques like Particle Induced X-ray Emission (PIXE), Particle Induced Gamma Ray Emission (PIGE), Rutherford Backscattering (RBS) and Elastic Recoil Detection (ERDA) are applied to obtain (μg/g) concentrations for elements from hydrogen to lead [1,2]. Generally low volume samplers are used to sample between 20 and 30 m^3 of air over a 24 h period, this together with IBA's sensitivity means that concentrations down to 1 ng/m^3 of air sampled can be readily achieved with only a few minutes of proton irradiation. With these short irradiation times and high sensitivities for a broad range of elements large numbers of samples can be obtained and analysed very quickly and easily. At the Australian Nuclear Science and Technology Organisation (ANSTO), over the past two decades, we have used IBA methods to acquire a database of over 50,000 filters from over 85 different sites through Australia and Asia, each filter has been analysed for more than 21 different chemical species from hydrogen to lead [1–7].

Large databases extending over many years means that modern statistical techniques like Positive Matrix Factorisation (PMF) [8] can be used to define well characterised source fingerprints and source contributions for a range of different fine particle air pollutants. Here we discuss these PMF techniques and show the power of these methods and how they identify both natural sources like sea spray and windblown soils as well as anthropogenic sources like automobiles, biomass burning, coal-fired power stations and industrial emissions. These data are particularly useful for Governments, environmental protection agencies and managers of pollution to better understand pollution sources and their relative contributions and hence to better manage air pollution.

Current trends are to take these IBA and PMF techniques a step further and to combine them with hourly wind speed and back trajectory data [9] to better pinpoint and identify emission sources. This is now being applied on local, regional and global scales with examples of local industrial pollution in urban areas and with long range industrial air pollution being tracked, for example, from China into Vietnam [3,10].

2. IBA techniques

PIXE is the primary IBA technique used to provide elemental concentrations for elements from Al to Pb [1]. The gamma technique of PIGE is useful for the lighter elements like Na, F, Al, Mg and Li. If aerosol samples are collected on thin filter papers, like 250 μg/cm^2 stretched Teflon filters, then RBS techniques can also be used to determine total C, N, and O concentrations and ERDA to determine total H concentrations [1,2]. Fig. 1 shows a typical

* Corresponding author.
E-mail address: dcz@ansto.gov.au (D.D. Cohen).

0168-583X/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nimb.2013.05.093

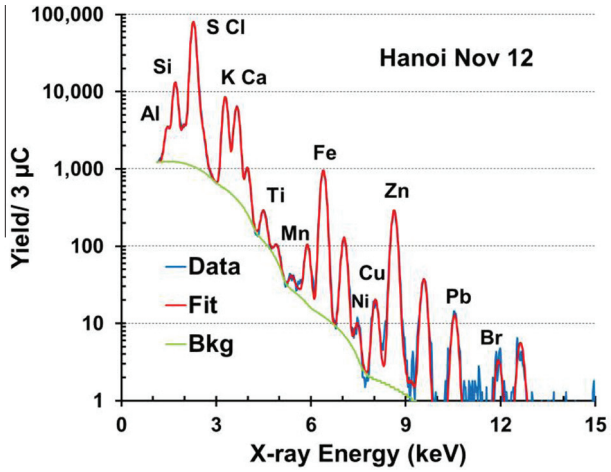

Fig. 1. Typical PIXE spectrum for 3 µC of 2.6 MeV protons on a thin Teflon filter collected at Hanoi, Vietnam over 24 h.

Table 1
Typical elemental signatures for common sources found in the PM2.5 air pollution size fraction [1–7].

Sources	Common signature elements in PM2.5 fraction
Soil	Al, Si, Ti, K, Ca, Fe, Ba, Sr, Rb
Sea spray	Na, Cl, Mg, Br, S
Metals Industry	Al, Ti, V, Cr, Ni, Co, Cu, Zn, Pb, S, As, Mo, Sb
Heavy oil burning (shipping, power generation)	V, Ni, S, Mn, Fe, Cr, As
Coal burning	Al, Si, S, Zn, As, Se, Hg, Cr, Cd, Th, Pb, Sb
Cement industry	Ca
Refuse incineration	K, Zn, Pb, Sb
Biomass burning, wild fires, agricultural burning	K, P, S, Cl, Br
Automobiles, petrol, diesel	S, Zn, K, Cl, Ca, Mn, Fe, Cr, V, Ni, Cu, Br, Pb, Mo, Ba, Sr, Cd, Sb, Ce, La, Pt

PIXE spectrum obtained using 3 µC of 2.6 MeV protons. Modern analysis codes like GUPIX, GeoPIXE and PIXAN are capable of extracting peak areas and hence elemental concentrations on hundreds of spectra in a few seconds [11–13]. With these techniques all being performed on the same sample between 20 and 30 different elements from H to Pb can be determined non-destructively in just a few minutes of proton irradiation.

This is a key advantage of the IBA techniques. In order to perform meaningful source fingerprinting and source apportionment studies it is important to obtain a wide range of elemental concentrations that span all of the possible sources one expects to find in the air sheds being studied.

Table 1 shows key elements associated with several common air pollution sources. Many of these sources are related to combustion of fossil fuels and biomass burning. Table 1 also shows that many elements like Si, S and K appear in many different sources which complicates the identification of separate sources and necessitates the use of PMF techniques.

3. Fingerprinting techniques

In order to properly fingerprint pollution sources two key requirements need to be met; the chemical species measured need to span the sources being studied and the dataset collected must be sufficiently large to include daily, seasonal and yearly variations of the source. Typically this means one needs to collect filters for 2 to 3 times a week for at least 2 years. If there are between 6 and 8 sources contributing to one urban air shed and you require a minimum of 2 or 3 elements to identify one source, this implies 15–30 different elements need to be determined. Obviously the more elements determined and the more filters analysed the better the source fingerprinting and apportionment will be.

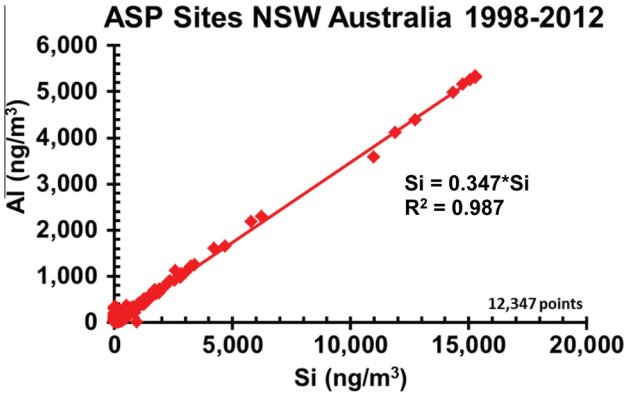

Fig. 2. Typical Al vs Si correlation plot for fine airborne soils in NSW Australia between 1998 and 2012.

Early techniques looked for inter-elemental correlations in the dataset and used these to produce elemental fingerprints. Fig. 2 shows such a correlation for Al vs Si for the soil fingerprint. The data contains 12,347 points covering a 14 year period. The correlation is excellent (R^2 = 0.987) with very few points being off the line of best fit. This demonstrates that Al and Si are indeed associated with the one source soil.

The gradient of the line in Fig. 2 gives the [Al/S] ratio in the soil fingerprint. Similar correlations of Si with Ca, Ti and Fe can be obtained and hence a true soil fingerprint containing the five elements Al, Si, Ti, Ca and Fe may be constructed [1,2].

Similar techniques can be used to obtain a sea spray fingerprint containing Na and Cl (see Fig. 3) at a coastal site at Cape Grim in North western Tasmania, Australia. The line drawn in Fig. 3 shows the expected [Cl/Na] ratio of 1.54 for sodium chloride a major salt component in sea spray. The large spread in the data in Fig. 3 reflects the fact that several chemical reactions can take place to convert NaCl to either sulfate or nitrate salts, hence altering this [Cl/Na] ratio over time.

These correlation techniques work well only if the elements used are associated mainly with one source. Once an element, like S contributes to many different sources then the correlations are not so obvious and the technique becomes more difficult to apply.

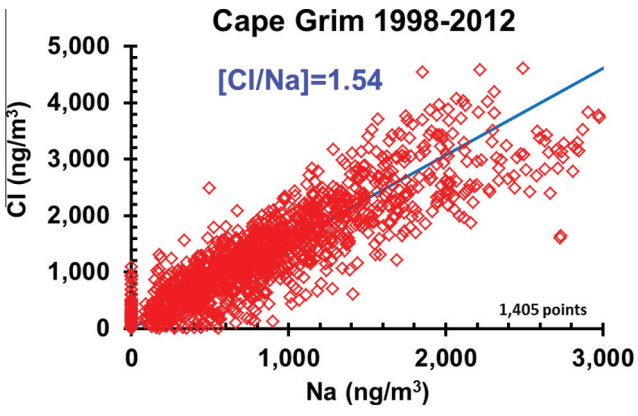

Fig. 3. Na vs Cl correlation at a coastal site at Cape Grim in Tasmania, Australia.

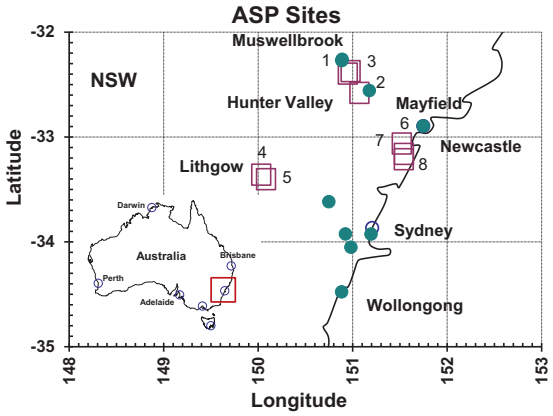

Fig. 4. Fine particle sampling sites (●) along the NSW, Sydney coast. The eight coal fired power stations (□) providing electricity to Sydney are also shown.

4. Positive Matrix Factorisation (PMF)

To overcome the limitations of these earlier source fingerprinting techniques Paatero and Tapper [8] developed the PMF technique. This method is a one step process which calculates both the source fingerprints and their contributions to the total measured gravimetric mass. The application of the method has been described in detail elsewhere [3,4,6,7,13]. Here we will just demonstrate the power of the technique by an example at the industrial/ urban Mayfield site near the coast at Newcastle north of Sydney, see Fig. 4.

PMF solves the matrix equation $M = F^{*}G + Err$. M is the $n^{*}m$ mass matrix, F is the $p^{*}m$ fingerprint matrix, G is the $n^{*}p$ contribution matrix and Err is the error matrix term, where the number of measured elements is m from n individual days and the PMF analysis solves for p different sources. Fig. 5 shows the $p = 8$ fingerprints (Soil, Secondary Sulfur, Sea spray, Smoke, Industrial Mn,

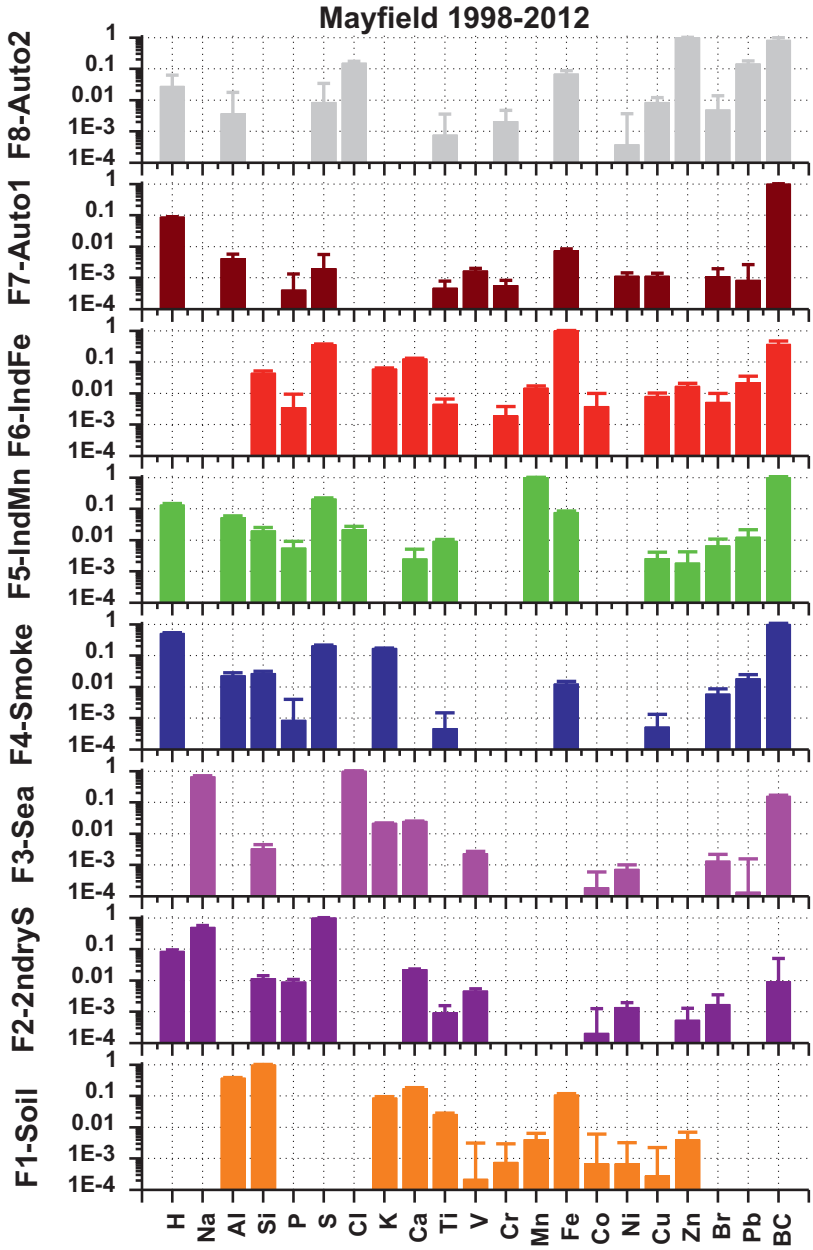

Fig. 5. The F fingerprint matrix obtained from a PMF analysis of the daily Mayfield data from 1998 to 2012 with 21 elements ($m = 21$) and 8 fingerprints ($p = 8$). The vertical axis for each fingerprint has been normalised to the largest element having a fractional contribution of one.

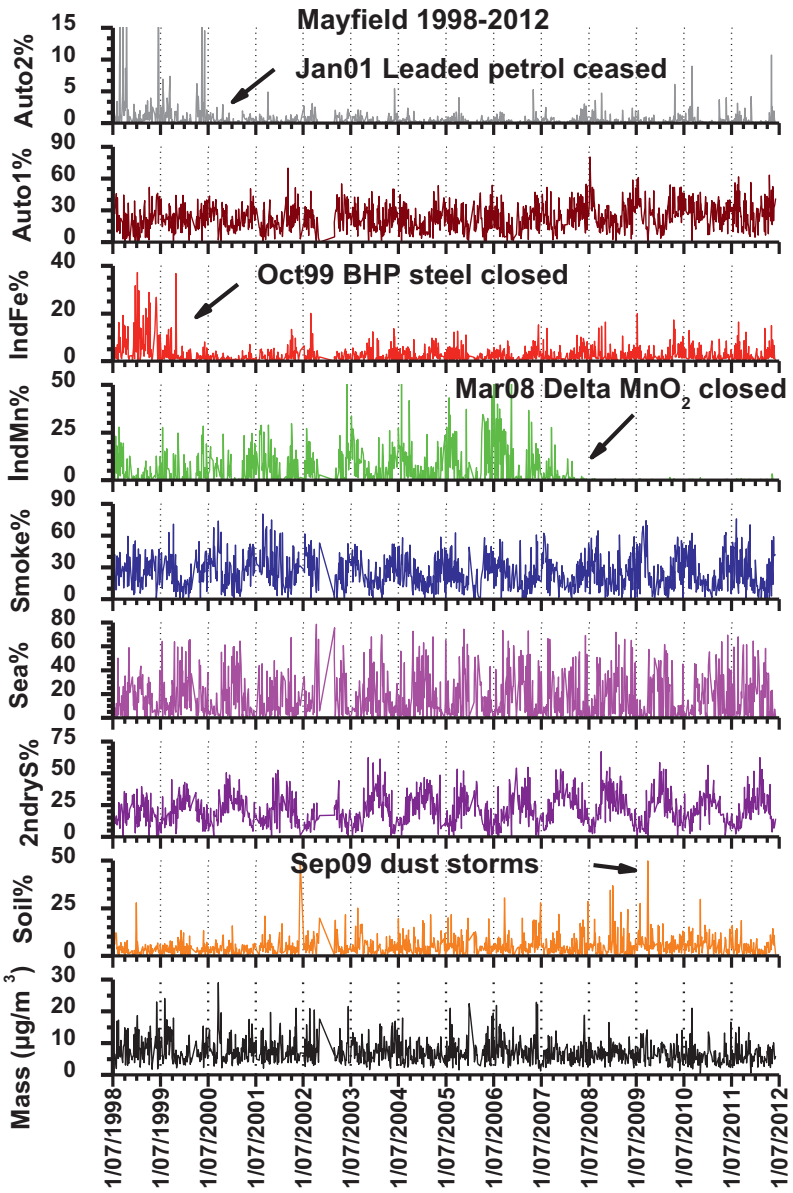

Fig. 6. Daily percentage contributions to the total mass of the 8 fingerprints shown in Fig. 5.

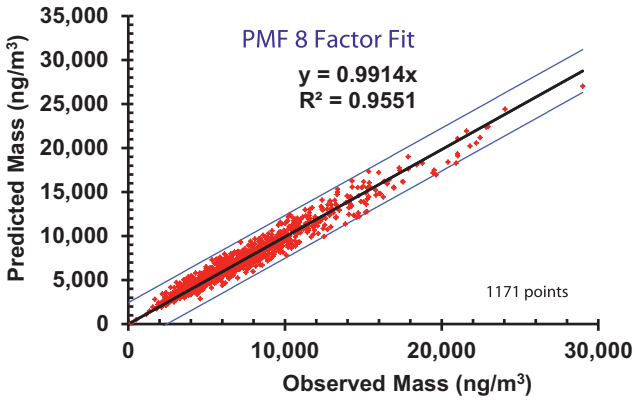

Fig. 7. Plot of the PMF mass against the gravimetric mass for an eight factor PMF fit to the Mayfield data between 1998 and 2012. The tram lines either side of the line of best fit represent the ±4 standard deviations.

Table 2
Average PMF percentage contributions to the fine particle pollution at Mayfield site between 1998 and 2012.

Fingerprint	% Mass total	% Mass when fingerprint present
Soil	5.0 ± 0.3	
2ndryS	21.1 ± 0.4	
Sea	17.1 ± 0.3	
Smoke	25.2 ± 0.3	
IndMn	4.6 ± 0.1	6.9 ± 0.1
IndFe	3.0 ± 0.2	7.8 ± 0.2
Auto1	23.3 ± 0.4	
Auto2	0.7 ± 0.1	1.5 ± 0.1

Industrial Fe and two Autos) obtained using the $m = 21$ elements from H to black carbon (BC), obtained on $n = 1171$ days between 1998 and 2012. The ordinate axis for each of the fingerprints in Fig. 5 shows the fraction of each of the elements present in that fingerprint normalised to the maximum element fraction being unity.

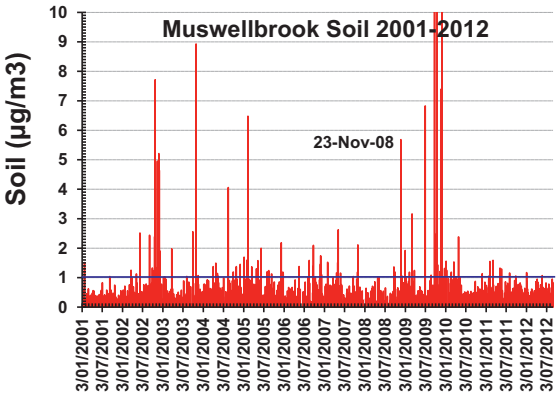

Fig. 8. Plot of the soil fingerprint contributions at Muswellbrook in the Hunter Valley, NSW Australia between 2001 and 2012. The dust storm of 23 November 2008 is shown and the horizontal line is for all dust events above 1 µg/m^3.

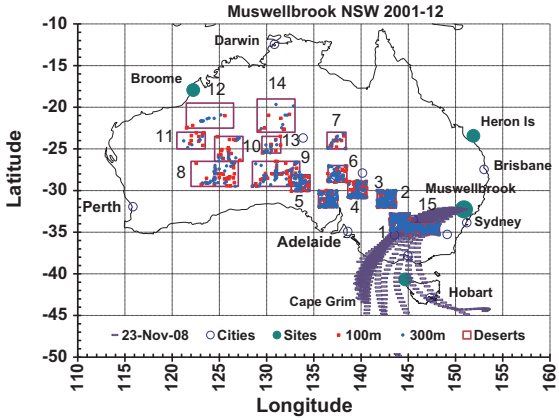

Fig. 9. Map of Australia showing the 14 desert regions as rectangles with the 15th rectangle (in the SE corner) showing the Riverina agricultural area. The 24 hourly back trajectories from Muswellbrook for 23 November 2008 are also shown intersecting with the Riverina rectangle 15. The red dots and blue dots represent HYSPLIT starting heights of 100 m and 300 m above the sampling site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The PMF technique can clearly provide meaningful fingerprints even when several elements like S, K and Zn appear in more than one fingerprint. The daily contribution of each of these fingerprints is given in Fig. 6 as a percentage of the total fine mass, shown at the bottom plot in this figure. Fig. 7 shows the quality of the PMF fit with the sum of each of the masses of each of the 8 fingerprints plotted against the measured gravimetric mass of each filter. More than 99% of the measured mass was explained by these eight factors with a correlation coefficient of $R^2 = 0.955$. A detailed inspection of the daily fingerprint contributions shown in Fig. 6 shows the real power of the PMF method. There are two Auto fingerprints, Auto1 with low Pb and Br and Auto 2 high Pb and Br. Leaded petrol (with added Pb and Br) ceased sales in NSW, Australia in January 2001 and we see the Auto2 contribution drops after January 2001.

Even more impressive are the time series plots of Industrial Fe and Industrial Mn. Mayfield is a significant industrial site with both steel works and metals production. In October 1999 one of the major BHP steelworks closed and in March 2008 the only manganese oxide plant, Delta Manganese, closed down. Both these events are clearly seen in the plots of Fig. 6. The PMF technique not only identifies these events but also provides a quantitative estimate of the change in the fine particle pollution loadings as a consequence of these closures.

Table 3
Number of back trajectory intersections in each of 14 desert regions and the Riverina agricultural area originating from the Muswellbrook sampling site.

Soil >1.0 µg/m^3	100 m		300 m	
Region	Number intersect	% intersect	Number intersect	% intersect
1 Lake Mungo	563	32.9	598	32.0
2 Lake Windaunka	190	11.1	172	9.2
3 EastFlinders	79	4.6	105	5.6
4 Olympic Dam	98	5.7	143	7.7
5 Emu Fields Salt Plains	34	2.0	39	2.1
6 Lake Eyre North	35	2.0	57	3.1
7 Simpson Desert	10	0.6	10	0.5
8 Great Vic West	31	1.8	35	1.9
9 Great Vic East	39	2.3	55	2.9
10 Gibson Desert	18	1.1	13	0.7
11 Little Sandy Desert	9	0.5	6	0.3
12 Great Sandy W Desert	4	0.2	9	0.5
13 Great Sandy E Desert	11	0.6	13	0.7
14 Tanami Desert	7	0.4	7	0.4
15 Riverina	584	34.1	605	32.4
Total hourly intersections	1,712		1,867	
% of all intersections	12.8		14.1	

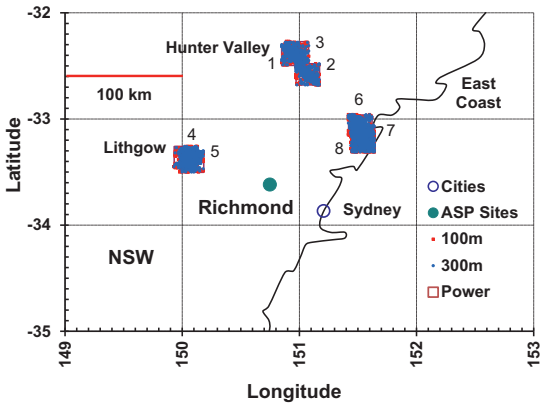

Fig. 10. Map of the eight coal fired power stations outside the Sydney Basin, and the sampling site at Richmond inside the Basin but west of Sydney. The red dots and blue dots represent HYSPLIT starting heights of 100 m and 300 m above the sampling site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The PMF technique also provides basic information for environmental protection agencies and managers of pollution. For example the average PMF source contributions are given in Table 2. The soil and sea spray contribution would be hard to reduce. But significant gains could be made by reducing smoke from biomass burning (25%), cleaning up the automobile fleet (23%) or reducing secondary sulfate (21%) emissions from coal burning, industry and cars.

5. The inclusion of air mass back trajectory data

Having used PMF methods to identify source types the next major recent development is to use hourly wind speed and directional back trajectories to actually determine where these sources are located. Recently we have developed a new approach which ties known point sources to back trajectories from the receptor (or

Table 4
Number of back trajectory intersections at each of the eight power stations (1–8), originating from the Richmond sampling site.

Starting heights	100 m		300 m	
Sulfate source power stations	Number intersect	% intersect	Number intersect	% Intersect
1 Bayswater	204	7.6	193	7.3
2 Redbank	231	8.6	289	11.0
3 Liddell	197	7.3	196	7.4
4 Mount Piper	179	6.6	273	10.4
5 Wallerawang	186	6.9	289	11.0
6 Eraring	592	21.9	481	18.3
7 Vales PointB	552	20.4	459	17.4
8 Munmorah	559	20.7	451	17.1
Total with SO_4 >2 μg/m^3	2700		2631	
% of all intersections		5.1%		5.0%

sampling) site on each sampling day over many years. The technique's strengths and application have been described previously [3,4,6,7,10]. Here we will show the power of this method with two examples. The first for natural long range transport of desert soils from central and outback Australia [3] and the second for the quantification and identification of the contributions of coal fired power stations to pollution in the Sydney basin [7].

Fig. 8 shows the time series plot of the soil fingerprint at the Muswellbrook site (see Fig. 9) between 2001 and 2012. The average soil contribution was (0.67 ± 3) μg/m^3, the large standard deviation reflects the large variability of the data over time and not the experimental error of the result. There were 64 days with soil above 1 μg/m^3. These were considered extreme soil events at Muswellbrook and the aim was to determine the major sources of these events. The Australian continent is 17% desert, in Fig. 9 we represent 14 of these major desert regions with rectangular boxes labelled 1 to 14. The 15th rectangle in the SE of Australia represents a major agricultural region called the Riverina.

A FORTRAN code was written to plot a dot in each rectangle each time an hourly back trajectory from the Muswellbrook site crossed one of these desert regions. This was done for every hour of every sampling day between 2001 and 2012 for which the soil measured at Muswellbrook exceeded 1 μg/m^3. All wind speeds and directions were taken from the HYSPLIT global dataset [9]. Two starting heights at the Muswellbrook site of 100 m and 300 m were used to see if the mixing layer heights affected the results at all.

There were 1712 and 1867 back trajectory intersections starting at 100 m and 300 m respectively. The results are summarised in Fig. 9 and Table 3. The red and blue dots in each rectangle represent a back trajectory intersecting with that rectangle at a starting height (above the sampling site) of 100 m and 300 m respectively. The Riverina agricultural region, shown as rectangle 15, contained 32–34% of all the back trajectory intersections with all 15 rectangles and was the major contributor to all extreme soil events measured at Muswellbrook. This Riverina region is more than 500 km from Muswellbrook and demonstrates the nature of long range fine soil transport within Australia. It should be emphasised that it is not appropriate to equate these percentage intersections directly with percentage contributions from the sources. There may be several different fine particle removal processes, like wet and dry deposition, along the transport path that may change these contributions.

Similar techniques and codes were used to find the sulfate contributions from the eight coal fired power stations outside the Sydney Basin to the Richmond sampling site inside the Basin (see Fig. 10). The average daily sulfate concentration at Richmond between 2001 and 12 was (1.18 ± 1.0) μg/m^3. There were 148 days with values above 2 μg/m^3 considered extreme sulfate days for this exercise.

Table 4 shows the results for HYSPLIT hourly back trajectories from Richmond intersecting with each power station during the study period. The three Hunter Valley power stations (1–3) to the north contributed 26%, the three mid-coast stations (6–8) 53% and the two western Lithgow stations (4–5) 21% of the all back trajectory intersections to the extreme sulfate events measured at Richmond between 2001 and 2012. Again it should be emphasised that it is not appropriate to equate these percentage intersections directly with percentage contributions from the sources. Nevertheless, previous work [7] showed that between 30% and 50% of the total sulfate measured in the Sydney Basin originated from these eight coal-fired power stations shown in Fig. 10. So we now have not only the total sulfate from coal burning entering the Basin, where over 4 million people reside, but also a qualitative measure of the relative contributions from each of these individual power stations.

6. Summary

IBA techniques can efficiently and effectively provide large databases of daily elemental concentrations from the small samples coming from fine particle air pollution studies. When these databases are used as inputs to modern source fingerprinting and source apportionment techniques like PMF they can provide quantitative source characterisation and identification that are extremely useful to environmental pollution agencies and pollution managers. These data can then be further enhanced by linking sources and receptor site measurements to global wind speed and direction datasets like HYSPLIT to produce real verifiable source locations for significant fixed point source emitters like desert regions and coal–fired power stations. We expect the concurrent use of these three methods to increase as we move towards a better understanding of fine particle pollution and its sources.

Acknowledgements

We would like to acknowledge the help of numerous parties in funding sampling sites and in collecting the thousands of fine particle Teflon filters throughout this study. All the IBA analyses were carried out by the ANSTO staff on the 2MV STAR Tandetron accelerator.

References

[1] D.D. Cohen, Nucl. Instrum. Meth. B136–138 (1998) 14.
[2] D.D. Cohen, E. Stelcer, O. Hawas, D. Garton, Nucl. Instrum. Meth. B219–220 (2004) 145–152.
[3] D.D. Cohen, J. Crawford, E. Stelcer, V.T. Bac, Atmos. Environ. 44 (2010) 320–328.
[4] D.D. Cohen, E. Stelcer, D. Garton, J. Crawford, Aerosol Pollut. Res. 2 (2011) 182–189.
[5] M. Radhi, M.A. Box, G.P. Box, M.D. Keywood, D.D. Cohen, E. Stelcer, R.M. Mitchell, Environ. Chem. 8 (2011) 248–262.
[6] D.D. Cohen, J. Crawford, E. Stelcer, A. Atanacio, Nucl. Instrum. Meth. B273 (2012) 186–188.
[7] D.D. Cohen, J. Crawford, E. Stelcer, A. Atanacio, Atmos. Environ. 61 (2012) 204–211.
[8] P. Paatero, U. Tapper, Environmetrics 5 (1994) 111–126.
[9] R.R. Draxler, G.D. Rolph, 2003. HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY Website (http://www.arl.noaa.gov/ready/hysplit4.html). NOAA Air Resources Laboratory, Silver Spring, MD.
[10] D.D. Cohen, J. Crawford, E. Stelcer, V.T. Bac, Atmos. Environ. 44 (2010) 3761–3769.
[11] E. Clayton, D.D. Cohen, P. Duerden, Nucl. Instrum. Meth. 180 (1981) 541–548.
[12] J.A. Maxwell, J.L. Campbell, W.J. Teesdale, Nucl. Instrum. Meth. B43 (1989) 218–230.
[13] C.G. Ryan, D.N. Jamieson, C.L. Churms, J.V. Pilcher, Nucl. Instrum. Meth. 104 (1995) 157–165.

Australian Nuclear Science and Technology Organisation
P: 02 9717 3111
E: enquiries@ansto.gov.au
W: www.ansto.gov.au