BOREAL ENVIRONMENT RESEARCH 13: 347–358 © 2008
ISSN 1239-6095 (print) ISSN 1797-2469 (online) Helsinki 28 August 2008

Comparison of multivariate source apportionment of urban $PM_{2.5}$ with chemical mass closure

Marko Vallius[1], Juhani Ruuskanen[2] and Juha Pekkanen[3][4]

[1] *National Public Health Institute, Department of Environmental Health, P.O. Box 95, FI-70701 Kuopio, Finland*
[2] *University of Kuopio, Department of Environmental Sciences, P.O. Box 1627, FI-70211 Kuopio, Finland*
[3] *National Public Health Institute, Environmental Epidemiology Unit, P.O. Box 95, FI-70701 Kuopio, Finland*
[4] *University of Kuopio, School of Public Health and Clinical Nutrition, P.O. Box 1627, FI-70211 Kuopio, Finland*

Received 2 Feb. 2007, accepted 6 Aug. 2007 (Editor in charge of this article: Veli-Matti Kerminen)

Vallius, M., Ruuskanen, J. & Pekkanen, J. 2008: Comparison of multivariate source apportionment of urban $PM_{2.5}$ with chemical mass closure. *Boreal Env. Res.* 13: 347–358.

In the lack of comparisons of different source apportionment methods, we resolved daily contributions of five source categories to $PM_{2.5}$ by applying principal component analysis and multiple linear regression to air pollution data from Helsinki, Finland. From the same data we estimated mass concentrations of four major components of $PM_{2.5}$ using a chemical mass closure model. Multiple linear regression analysis suggested that secondary and other long-range transported particles contributed 58% to total $PM_{2.5}$, on the average, while traffic and mixed local combustion sources accounted for 19%, oil combustion for 14%, crustal source for 4.9% and salt for 2.4%. Mass closure suggested average contributions of 50%, 34%, 4.5% and 1.2% from ammonium sulphate, combustion-related particles, crustal material and sea salt, respectively. The crustal source and salt were apportioned similar amounts of $PM_{2.5}$ whereas results from the two methods were less comparable for the long-range transported and secondary particles, and the combustion-related source components.

Introduction

Promulgation of national and international strategies to control ambient air pollution requires knowledge of the contribution of major sources to particulate matter (PM; $PM_{2.5}$ stands for particulate matter less than 2.5 μm in aerodynamic diameter) levels and the relative importance of local and distant sources in different parts of a country or continent. In addition, both qualitative and quantitative information of source-specific PM are needed for the assessment of human exposure to particles from different sources.

The contribution of different sources to ambient particulate matter can be estimated using source apportionment methods, which exploit information of the composition of emissions at their source (e.g. Henry 1977, Henry *et al*. 1984, Gordon 1988, Paatero *et al*. 2002, Hopke 2003). All source apportionment methods require information of the properties of PM both at the site of emissions (source) and the site where PM are measured (receptor). In statistical multivariate source apportionment models, such as principal component analysis and multiple linear regression (PCA-MLR), information of the chemical composition of PM at a given receptor site can be utilised as such to identify potential sources which contribute to the PM concentrations at that site. In chemical mass balance models (Watson *et al.* 1991), on the other hand, knowledge of the chemical composition of PM at its sources is used to model the contribution of these sources to the measured PM at a receptor site. Chemical mass closure is a deterministic modelling approach that relies on mass balance and direct back-calculation of PM constituents from the measured elemental composition of a PM sample (Harrison *et al.* 2003).

Each source apportionment technique is based on a set of assumptions regarding the analysed data and these assumptions need to be substantially satisfied if the models are to work properly. In real life measurement data, however, this is rarely the case. Therefore, different methods yield different results when they are applied to measurement data about air pollution. Thus a comparison is warranted of source apportionment results obtained from the same data but by using different analytical tools.

The main objective of this paper is to compare the results of source apportionment by two fundamentally different methods: a multivariate method (PCA-MLR) and a deterministic method (mass closure, MC). The elemental composition of $PM_{2.5}$ and other air pollution data from winter 1998–1999 in Helsinki were used to estimate daily concentrations of source-specific $PM_{2.5}$. We shall describe and discuss the differences between the two methods both in terms of the average concentrations and the daily variation in source-specific $PM_{2.5}$ fractions.

Material and methods

Measurements

$PM_{2.5}$ mass concentration, absorption coefficient of $PM_{2.5}$ filters (denoted $Abs_{2.5}$), number concentrations of ultrafine particles (diameter 0.01–0.1 μm; denoted $NC_{0.01–0.1}$ in tables and figures) and accumulation mode particles (diameter 0.1–1.0 μm; denoted $NC_{0.1–1.0}$), nitrogen oxides (NO_x), sulphur dioxide (SO_2), and meteorological parameters (temperature, wind speed and wind direction) were monitored in Helsinki, Finland, from November 1998 to April 1999. The measurement of $PM_{2.5}$ and the gaseous pollutans took place at an urban monitoring site located in a small park about two kilometres northeast of the city centre. The distance of the PM monitors and gaseous pollutant monitors from the only nearby major street (about 14 000 vehicles day^{-1} on weekdays) were 60 and 10 meters, respectively. The meteorological parameters were measured 50 m above the ground level at a meteorological station located 1.5 km southwest from the air pollution monitoring site. Detailed descriptions of the sampling and analytical methods, equipment and quality control are provided elsewhere (Koistinen *et al.* 1999, Pekkanen *et al.* 2000, Vallius *et al.* 2003).

Data processing

A total of 168 $PM_{2.5}$ samples were available for energy dispersive X-ray fluorescence analysis (ED-XRF). We calculated the daily concentrations for each element by subtracting the median field blank concentration from the total concentration. The negative concentrations resulting from field blank subtraction were set to zero. We substituted any elemental concentrations below the limit of detection with a value derived from the sample-specific detection limit for the element in question. During the course of data checking and validation, we excluded three outliers from the elemental concentration data thus resulting in 165 valid 24-hour samples. All of the excluded concentration values were > 4 times S.D. above the corresponding average elemental concentration. We substituted missing 24-hour average values for sulphur dioxide (SO_2) and ultrafine ($NC_{0.01–0.1}$) and accumulation mode ($NC_{0.1–1.0}$) particle counts with values obtained from multivariate regression analyses, where the dependent variable having the missing value(s) was regressed on the independent variables which resulted in the highest correlation coefficient (R^2). One missing value (0.6%) for SO_2 and four values (2.4%) for particle number counts ($NC_{0.01–0.1}$ and $NC_{0.1–1.0}$) were substituted with values obtained from the regression analyses. Concentrations of 12 elements (Al, Si, S, Cl, K, Ca, V, Mn, Fe, Ni, Cu, Zn) out of the 18 determined were detectable in more than 60% of the samples and also had satisfactory results for the duplicate samples. We excluded the remaining six elements (Br, Cr, Rb, Sr, Ti, Pb) from further analyses.

Source identification and source apportionment

Principal component analysis and multiple linear regression

In order to identify the main source categories of $PM_{2.5}$, we analysed the elemental composition of PM samples and other air pollution data from 165 days using principal component analysis (PCA). We identified source categories by examining the elemental profiles of rotated (varimax) principal components (Thurston and Spengler 1985). We then assigned each principal component to a source category based on prior knowledge of the chemical composition of particulate matter from different sources. Since PCA alone does not provide estimates of the daily source-specific $PM_{2.5}$ concentrations at the measurement site, we obtained these estimates by regressing the measured $PM_{2.5}$ on absolute principal component scores (APCS) (Thurston and Spengler 1985).

Several criteria were used in selecting the optimal models: in terms of source identification (PCA), we required identification of major sources with physically reasonable principal components whose eigenvalues were larger than 1 after varimax rotation. In terms of source apportionment (multiple linear regression) we required positive regression coefficients for all sources, a positive and moderate model intercept, and a high model R^2. We used SAS/STAT® statistical software version 8.02 of the SAS System for Windows (SAS Institute Inc. 1999) for all statistical analyses.

Chemical mass closure

Mass closure (MC) refers to a simple mass balance modelling procedure that can be used for apportioning particulate matter to its components (or sources). The basis of mass closure is the analysis of particulate matter sample(s) for a number of chemical components which can then be used as tracers of major aerosol constituents (Andrews *et al*. 2000, Harrison *et al*. 2003). The sources of particulate matter have to be either assumed based on prior information, or identified by using for example PCA before their contributions can be modelled in mass closure. In this work, we selected the major source components to be modelled based on the results of PCA of the current data and previously published information on the sources of fine particulate matter in Helsinki area (Ojanen *et al*. 1998, Pakkanen *et al*. 2001).

Unlike in the PCA-MLR where also gaseous pollutants and particle number concentrations were utilised, the mass closure of $PM_{2.5}$ was done using only the concentrations of elements which were available from the XRF analysis of the $PM_{2.5}$ filters. Consequently, we were able to model the mass concentrations of three $PM_{2.5}$ components using chemical mass closure: ammonium sulphate (denoted $(NH_4)_2SO_{4mc}$; subscript mc stands for mass closure), crustal material ($Crustal_{mc}$) and sea salt ($Salt_{mc}$). The residual $PM_{2.5}$ mass is by default related to an undefined number of sources and chemical constituents of PM which are not accounted for by the three components specified above. In practice, however, majority of the residual $PM_{2.5}$ can be assumed to be combustion related particulate matter from traffic, oil combustion, industrial sources and small-scale combustion (e.g., residential heating). Therefore, although this will be somewhat of a simplification, we refer to the residual $PM_{2.5}$ component as $Combustion_{mc}$ later in this paper. Based on previous data on the composition of fine particulate matter in urban Helsinki (Ojanen *et al*. 1998), the $Combustion_{mc}$ fraction comprises mostly elemental and organic carbon and nitrate. The division of PM components in mass closure is analogous to that used in a previous study in Helsinki (Koistinen *et al*. 2004).

Ammonium sulphate

Secondary inorganic particles comprise mostly sulphate, nitrate and ammonium. However, since these components were not measured routinely during the study, we had to restrict ourselves to estimation of the ammonium sulphate concentration from the sulphur concentration obtained from the XRF analysis:

$$(NH_4)_2SO_{4mc} = 4.125[S_{xrf}], \quad (1)$$

where $[S_{xrf}]$ is the sulphur concentration from the XRF analysis, and the factor 4.125 converts the atomic weight of S into the molecular weight of $(NH_4)_2SO_4$. This equation assumes that all S on the $PM_{2.5}$ filter is bound to ammonium sulphate.

Crustal material

A formula that has previously been used by Eldred *et al*. (1997) and Shiraki and Holmen (2002) was selected for mass closure of the crustal component:

$$\text{Crustal}_{mc} = 2.49[Si] + 2.20[Al] + 2.42[Fe] + 1.63[Ca] + 1.94[Ti], \quad (2)$$

where the factor 2.42 for Fe takes into account the usual fraction of K relative to Fe in soil.

Sea salt

The equation for calculating the sea salt component was

$$\text{Salt}_{mc} = 0.89 \times 1.816[Cl], \quad (3)$$

where 0.89 is the previously reported concentration ratio of Cl from sea salt and total Cl in the Helsinki area (Ojanen *et al*. 1998), and the factor 1.816 takes into account other major constituents in sea salt (Seinfeld and Pandis 1998). This formula was used previously (Koistinen *et al*. 2004) to reconstruct a salt source component in $PM_{2.5}$ in Helsinki. The total factor of 1.62 in Eq. 3 is very close to the factor of 1.65 used earlier by Harrison *et al*. (2003) for mass closure calculation of NaCl in marine aerosol.

Combustion PM

The residual part of $PM_{2.5}$ comprises all other components except those estimated with Eqs. 1, 2 and 3. Therefore, several sources of particulate matter contribute to this fraction. However since majority of this residual $PM_{2.5}$ is related combustion sources either directly or indirectly we named this component Combustion_{mc} and calculated it from

$$\text{Combustion}_{mc} = \text{measured } PM_{2.5} - [(NH_4)_2SO_{4mc}] - [\text{Crustal}_{mc}] - [\text{Salt}_{mc}]. \quad (4)$$

Results and discussion

We obtained PCA-MLR results using a modification of our previously published source apportionment model (Vallius *et al*. 2003). The five sources that we identified in this re-analyses were long-range transported air pollution (denoted LRT_{mlr}), traffic and local combustion (Traffic_{mlr}), fuel oil combustion (Oil_{mlr}), crustal source (Crustal_{mlr}) and sea salt (Salt_{mlr}) (Table 1). The subscript mlr stands for multiple linear regression that we used for calculating the mass contributions from the identified source categories (Table 2). In the original model the sea salt source was mixed with the contribution from an unrecognised source of Pb (Vallius *et al*. 2003). The new principal component analyses without Pb were performed in order to extract a distinct sea salt source. This facilitates a more straightforward comparison of this principal (or source) component with the sea salt component calculated using mass closure. In the PCA-MLR re-analyses, we again identified five source categories with minute differences with regard to interpretation of PCA results (except for the sea salt component) and only moderate differences in terms of average contributions of the identified sources to ambient $PM_{2.5}$ as compared with our earlier results (Vallius *et al*. 2003). The largest changes in average $PM_{2.5}$ mass contributions from the previous to current analyses were the increased long-range transported particulate matter component (+8%), and respective decreases in the sea salt (–5%) and traffic and local combustion (–4%) source contributions.

Long-range transported and secondary particulate matter

The modelled concentrations of LRT_{mlr} and $(NH_4)_2SO_{4mc}$ were close to each other (Table 2). However, we did expect the mass of ammonium sulphate to be lower than that of LRT_{mlr}, because long-range transported particulate matter comprises also other secondary inorganic compounds (mainly ammonium nitrate) and various carbonaceous compounds (Viidanoja *et al*. 2002). Ojanen *et al*. (1998) estimated that about 30% of $PM_{2.3}$ in urban Helsinki 1996–1997 is attributable to long-range transported nitrate, organic carbon and elemental carbon. In our study this 30% would correspond to a 3.8 μg m^{-3} $PM_{2.5}$ concentration, which is considerably higher than the 1.0 μg m^{-3} difference we observed between the LRT_{mlr} and $(NH_4)_2SO_{4mc}$ concentrations.

Our estimate for $(NH_4)_2SO_4$ from mass closure (50% of $PM_{2.5}$) is high in comparison with the reported aggregate amount of less than 30%

Table 1. Principal components of the air pollution data (varimax rotated). Number of cases (days analysed) = 165.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Al	0.11	0.89	0.13	0.07	–0.06
Si	0.13	0.94	0.08	0.03	–0.14
S	–0.04	–0.11	0.24	0.86	–0.18
Cl	–0.02	–0.07	–0.06	–0.10	0.96
K	0.23	0.39	0.20	0.78	0.04
Ca	0.12	0.84	0.14	0.05	0.27
V	0.04	0.18	0.85	0.32	0.04
Mn	0.47	0.51	0.17	0.35	–0.18
Fe	0.47	0.76	0.08	0.21	–0.16
Ni	0.19	0.19	0.90	0.20	–0.01
Cu	0.80	0.28	0.03	0.24	–0.01
Zn	0.55	0.25	0.35	0.55	–0.02
$Abs_{2.5}$	0.77	0.16	0.22	0.49	0.01
NO_x	0.94	0.10	0.09	–0.03	0.02
SO_2	0.32	0.03	0.81	0.09	–0.12
$NC_{0.01–0.1}$	0.87	0.13	0.27	–0.16	–0.05
$NC_{0.1–1.0}$	0.66	0.20	0.26	0.54	0.03
Eigenvalue	4.29	3.70	2.69	2.59	1.14
Source category	Traffic, local combustion	Crustal, resuspension	Oil combustion	LRT, secondary particles	Salt

Table 2. Descriptive statistics for the mass concentration of $PM_{2.5}$ components (µg m^{-3}) calculated from multiple linear regression (mlr) and mass closure (mc).

	N	Average (µg m^{-3})	% of $PM_{2.5}$	5th percentile (µg m^{-3})	Median (µg m^{-3})	95th percentile (µg m^{-3})
$PM_{2.5}$	165	12.8		5.17	10.6	25.7
LRT_{mlr}	165	7.38	58	0.49	6.05	18.4
$Traffic_{mlr}$	165	2.48	19	0.64	2.21	5.71
Oil_{mlr}	165	1.74	14	0.02	1.39	4.41
$Crustal_{mlr}$	165	0.63	4.9	–0.32	0.44	2.11
$Salt_{mlr}$	165	0.31	2.4	–0.01	0.20	0.91
$(NH_4)_2SO_{4mc}$	165	6.40	50	1.68	5.73	14.1
$Combustion_{mc}$	165	5.63	44	1.31	4.55	14.3
$Crustal_{mc}$	165	0.58	4.5	0.13	0.42	1.34
$Salt_{mc}$	165	0.16	1.2	0.01	0.06	0.62
SO_4^{2-}	68	5.00		2.12	4.62	10.0
NO_3^-	68	1.28		0.11	0.92	3.84

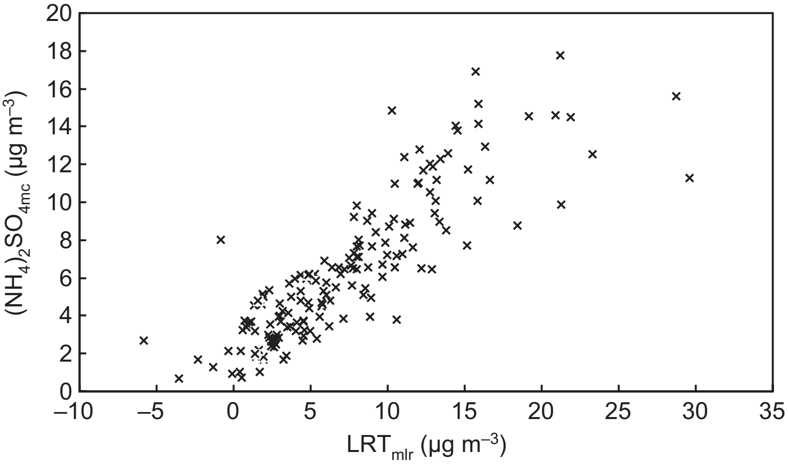

Fig. 1. Relationship between the estimated concentrations of long-range transported particulate matter from PCA-MLR (LRT_{mlr}) and ammonium sulphate from mass closure ($(NH_4)_2SO_{4mc}$).

of sulphate and ammonium in fine particulate matter ($PM_{2.3}$) in urban Helsinki in 1996–1997 (Ojanen *et al*. 1998). However, this difference can be fully explained by the lower concentration of sulphate aerosol over Finland during 1996–1997 compared to spring 1999 of this study (Co-operative programme for monitoring and evaluation of the long-range transmission of air pollutants in Europe: Chemical Coordinating Centre (EMEP-CCC). Available at: http://www.nilu.no/projects/ccc/network/index.html), assuming that local and regional sources had a similar effect on the ammonium sulphate concentrations in this study and the study by Ojanen *et al*. (1998).

The average sulphate concentrations observed during January–April in this study were higher than those observed during the same time period at two background air quality monitoring stations (EMEP-CCC 2007) which are operated by the Finnish Meteorological Institute and are part of the EMEP network (station codes FI37 and FI17). This suggests that not only the long-range sources but also local and regional sources, e.g. heating and energy production, could have affected the average sulphate concentrations in this study.

Correlation between LRT_{mlr} and $(NH_4)_2SO_{4mc}$ was quite high (R = 0.88; Fig. 1 and Table 3) and, by default, very close to the loading of S in the fourth principal component (Table 1) since $(NH_4)_2SO_{4mc}$ (Eq. 1) is a linear transformation of the S_{xrf} concentration used in the PCA-MLR analysis.

Sixty-eight $PM_{2.5}$ samples collected in January–April 1999 were analysed for sulphate and nitrate (Table 2) using ion chromatography (IC), making it possible to compare the measured sulphate (IC) and estimated sulphate (from S_{xrf}) concentrations. There was a good correlation of SO_4^{2-} and S_{xrf} (Spearman $R = 0.96$; Fig. 2). However, the average ratio of SO_4^{2-} and S_{xrf} was 3.2, which suggests that the factor of 3.0 that we used to convert S_{xrf} to SO_4^{2-} (Eq. 1) may have resulted in a slight underestimation of the true concentration of sulphate species. On the other hand, a factor of 4.125 in Eq. 1 would overestimate the

Table 3. Spearman correlation coefficients of specific $PM_{2.5}$ components from multiple linear regression (mlr) with mass closure components (mc).

	$(NH_4)_2SO_{4mc}$	$Combustion_{mc}$	$Crustal_{mc}$	$Salt_{mc}$
LRT_{mlr}	0.88	0.46	0.02	–0.09
$Traffic_{mlr}$	–0.02	0.45	0.31	0.02
Oil_{mlr}	0.35	0.25	0.13	–0.09
$Traffic_{mlr} + Oil_{mlr}$	0.17	0.53	0.34	0.01
$Crustal_{mlr}$	–0.14	0.06	0.90	–0.05
$Salt_{mlr}$	–0.15	0.41	–0.08	0.83

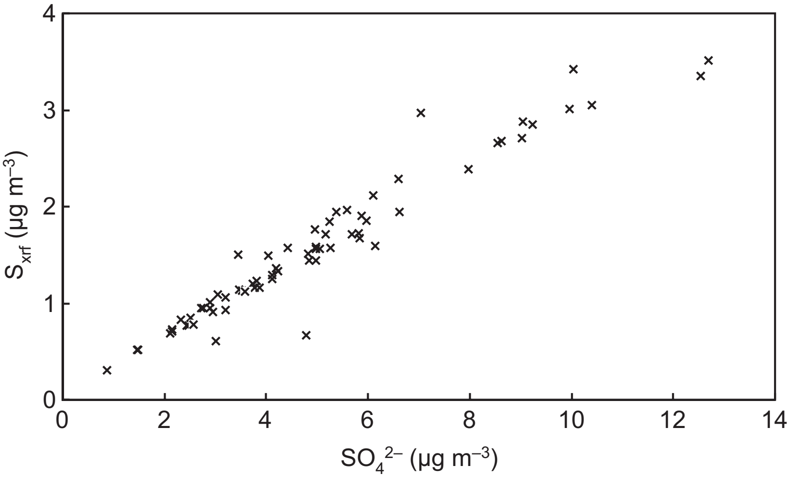

Fig. 2. Relationship between sulphate (SO_4^{2-}; ion chromatography) and sulphur (S_{xrf}; X-ray fluorescence) concentrations on the $PM_{2.5}$ filter samples (*N* = 68).

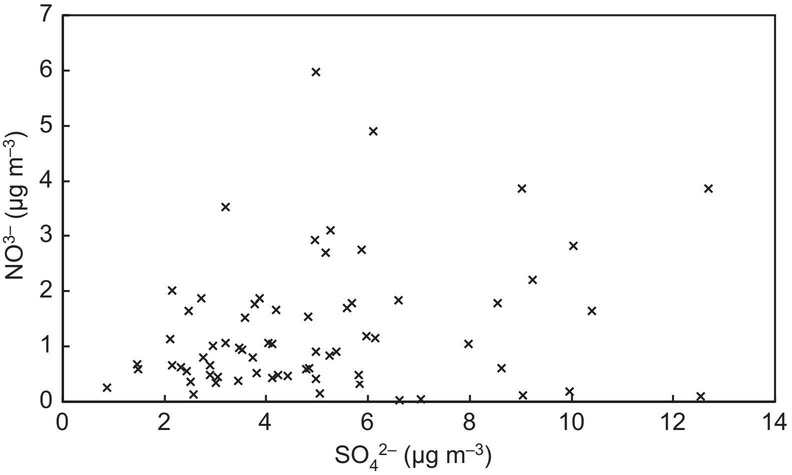

Fig. 3. Relationship between the measured (IC) sulphate (SO_4^{2-}) and nitrate (NO_3^-) concentrations on the $PM_{2.5}$ filter samples (*N* = 68).

mass concentration of the sulphate species if the sulphate were in fact not completely neutralised to ammonium sulphate but present partly as sulphuric acid (H_2SO_4), ammonium bisulphate (NH_4HSO_4), or other sulphates.

The low correlation between SO_4^{2-} and NO_3^- (R = 0.17, Fig. 3) demonstrates the partially different origin of these two components and, on the other hand, the fact that variable amounts of nitrate are lost during conventional filter sampling of $PM_{2.5}$. Pakkanen *et al*. (2001) reported that about 2/3 of nitrate evaporates from teflon filters during the sampling of urban $PM_{2.3}$, and that the percentage of nitrate lost was highly variable in time due to variable meteorological conditions. Thus, in this work, it was not feasible to estimate the nitrate concentration from the average sulphate/nitrate ratio or by using regression analysis.

NO_3^- correlated with Cl (R = 0.64) and the $Salt_{mlr}$ component (R = 0.67). A reduced PCA-MLR where SO_4^{2-} and NO_3^- were included (N = 68) resulted in a Cl–NO_3^- component which contributed 12% to the average $PM_{2.5}$ during that measurement period. The highest concentrations attributable to the Cl–NO_3^- component were associated with winds from south and southwest blowing in from the Baltic Sea and downtown Helsinki. The mass contribution of the Cl–NO_3^- component is consistent with another study where nitrate made up 12% of the urban $PM_{2.3}$ in Helsinki (Pakkanen *et al*. 2001). About 40% of NO_3^- in urban Helsinki is of local origin according to an earlier study (Ojanen *et al*. 1998).

Particulate matter from combustion sources

The share of Combustion$_{mc}$ in the measured PM$_{2.5}$ was 44%, which is higher than the summed contribution of 33% from Traffic$_{mlr}$ and Oil$_{mlr}$. This difference in the average mass contributions by these component was expected, since in the mass closure model all carbonaceous material and inorganics (except ammonium sulphate) that were not estimated by the specific Eqs. 1–3 were included in the Combustion$_{mc}$ component, regardless whether they were in fact combustion related or not. In PCA-MLR however, the carbonaceous particles and inorganics are not assigned to any source component *a priori* but distributed between different source categories.

The correlation coefficients (Table 3) of Combustion$_{mc}$ with Traffic$_{mlr}$ (Fig. 4a) and Oil$_{mlr}$ (Fig. 4b) were rather low, suggesting that these components obtained from MC and PCA-MLR represented different fractions of PM$_{2.5}$. The correlation of Combustion$_{mc}$ with the sum of Traffic$_{mlr}$ and Oil$_{mlr}$ was only somewhat higher (Table 3). There were six specific cases (encircled in Fig. 4a) that were associated with relatively high concentrations of PM$_{2.5}$, Abs$_{2.5}$ and NO_3^-. We could find no definite explanation for the discrepancy between the masses from PCA-MLR and mass closure during these six days. It seems, however, that the difference was due to both an underestimation of the total mass of long-range transported particulate matter by the mass closure model, and the inability of the PCA-MLR to model the occasional high concentrations of locally produced particulate matter.

The moderate correlation (Table 4) of Combustion$_{mc}$ with both NO_3^- (R = 0.79) and Abs$_{2.5}$ (R = 0.73) suggests that local sources which emit carbonaceous PM also contributed to the production of NO_3^-. Correlation of NO_3^- was lower with Traffic$_{mlr}$ (R = 0.48) than with Salt$_{mlr}$ (R = 0.67). This could be due to the similar association of Cl and the long-range transported fraction of NO_3^- with wind direction, since both were strongly associated with south-westerly winds blowing in from the Baltic Sea and sweeping over harbours and densely trafficked downtown Helsinki.

Crustal particles

The median contributions to PM$_{2.5}$ from the crustal source component were similar, although estimates from the two methods were quite different both for very high and very low concentrations of crustal particulate matter (Fig. 4c). One possible reason for these differences is that PCA-MLR associated material other than "pure" crustal particles with this source component, especially if their concentrations correlated with PM emissions from the soil and streets. This could well be the case for traffic-induced resuspension and exhaust emissions, which can contain considerable amounts of typical crustal elements (Wang *et al*. 2003). On the other hand,

Table 4. Spearman correlation coefficients of the specific PM$_{2.5}$ components from multiple linear regression (mlr) and mass closure (mc) with the measured pollutants.

	LRT$_{mlr}$	$(NH_4)_2SO_{4mc}$	Traffic$_{mlr}$	Oil$_{mlr}$	Combustion$_{mc}$	Crustal$_{mlr}$	Crustal$_{mc}$	Salt$_{mlr}$	Salt$_{mc}$
PM$_{2.5}$	0.84	0.85	0.22	0.36	0.74	–0.02	0.22	0.13	–0.04
PM$_1$	0.76	0.80	0.34	0.41	0.72	0.02	0.29	0.05	–0.13
PM$_{10}$	0.60	0.60	0.08	0.29	0.58	0.30	0.48	0.16	–0.04
SO_4^{2-}	0.81	0.96	0.20	0.16	0.29	–0.17	–0.02	–0.01	–0.10
NO_3^-	0.35	0.18	0.48	0.16	0.79	–0.14	–0.04	0.67	0.64
Abs$_{2.5}$	0.49	0.45	0.71	0.25	0.73	0.14	0.43	0.12	0.02
NO_x	–0.10	–0.04	0.87	0.15	0.47	0.14	0.38	0.03	0.04
CO	0.28	0.23	0.40	–0.01	0.45	–0.14	0.02	0.18	0.17
SO_2	0.20	0.39	0.29	0.76	0.30	0.03	0.26	–0.10	–0.17
NC$_{0.01–0.1}$	–0.18	–0.04	0.81	0.20	0.28	0.21	0.43	–0.10	–0.11
NC$_{0.1–1.0}$	0.53	0.51	0.57	0.33	0.84	0.14	0.42	0.18	0.00
Temp	0.03	–0.14	–0.43	–0.26	0.04	0.06	–0.12	0.43	0.31
WS	0.03	0.00	–0.58	–0.04	–0.28	0.01	–0.22	0.20	0.22

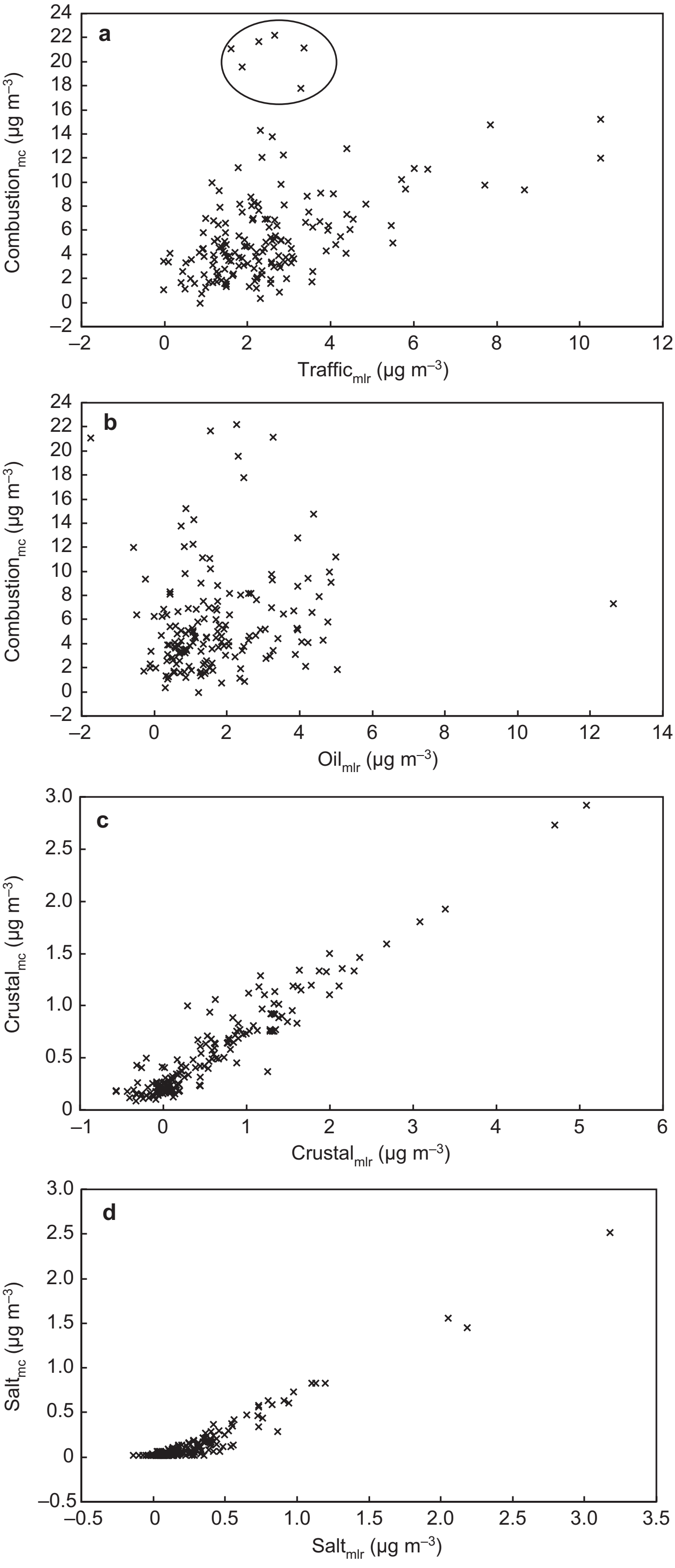

Fig. 4. Relationship between the estimated concentrations of particulate matter related to (**a**) local traffic and combustion from PCA-MLR (Traffic$_{mlr}$) and residual PM$_{2.5}$ from mass closure (Combustion$_{mc}$); (**b**) oil combustion from PCA-MLR (Oil$_{mlr}$) and residual PM$_{2.5}$ from mass closure (Combustion$_{mc}$); (**c**) crustal particulate matter from PCA-MLR (Crustal$_{mlr}$) and mass closure (Crustal$_{mc}$); (**d**) sea salt from PCA-MLR (Salt$_{mlr}$) and mass closure (Salt$_{mc}$).

while the mass closure model yields a more "pure" crustal component, it possibly underestimates the total amount of $PM_{2.5}$ that is associated with the crustal source component, for example, by neglecting the mass of resuspended organic material. Irrespective of the potentially different interpretation of the components obtained from the two methods, the correlation between the crustal $PM_{2.5}$ from PCA-MLR and MC was high ($R = 0.90$).

Salt

The difference in median concentrations of sea salt (Table 2) is probably due to the different distribution of low salt concentrations in PCA-MLR and MC methods. Similarly to the crustal component, the PCA scores of the sea salt component were possibly affected by particulate matter from other sources. This would result in materials other than salt particles being apportioned to the sea salt component in MLR, thus resulting in higher median source contribution for Salt_{mlr} as compared with that for Salt_{mc}. Estimated daily sea salt concentrations were in good agreement ($R = 0.83$; Fig. 4d).

PCA-MLR *vs.* mass closure

An advantage of mass closure analysis compared to the multivariate techniques is that it can be run even on a single sample of particulate matter. In MC, the samples are characterized one sample at a time making them independent of each other, whereas PCA relies on the analysis of the common variation of chemical constituents in several samples. The least squares fitting in linear regression analysis leads, by default, either to underestimation or overestimation of the source contributions for a single observation (measurement day). The quality of sampling and analyses including field procedures, filter weighing, and XRF analysis, have a like effect on both source apportionment methods.

One major disadvantage of the mass closure is that no other chemical compounds can be attributed to a specific source component except for those which are either measured directly or estimated indirectly by the mass closure equations (in this work, Eqs. 1–3). In addition, many of the chemical constituents of PM originate from two or more sources. Therefore, the total contribution to PM from a given source on one day can be either higher or lower than implied by simple calculations due to emissions of similar elements or chemical species from other sources, and because the true chemical composition profile of the material emitted by that source is not restricted to that implied by the mass closure formula. In PCA, elements which are emitted from two or more source categories can be associated with two or more principal (source) components. Thus the principal component score values of those components and, consequently, the contributions of those source categories to the modelled total PM concentration are affected by the variations of the concentrations of all associated elements. However, the elemental profiles of most PM sources are not entirely stable in time, which may complicate the interpretation of sources derived from PCA.

Chemical mass closure will not yield negative source contributions for any single source. However, the calculated mass of ammonium sulphate, crustal material and sea salt, can add up to more than the measured total $PM_{2.5}$ concentration making the balance in mass closure negative. This happened only once in the 165 days in our study. In PCA-MLR, the contribution of a source category can be negative for certain measurement days (Figs. 1 and 4) due to the occasional negative principal component scores which are used as the independent variables in MLR.

Conclusions

We obtained partially conflicting results in terms of concentrations and daily variations of $PM_{2.5}$ components in urban Helsinki by using two different source apportionment methods. The differences were mainly due to the fact that the two methods are based on fundamentally different approaches to model the contributions from specific sources of particulate matter. While PCA-MLR relies on the daily variation of markers, mass closure relies on the absolute levels of the selected markers.

With regard to mass closure, the lack of data on elemental and organic carbon compounds was the single most important factor that prevented a more complete apportionment of the combustion-related component of urban particulate matter. The importance of carbon speciation in getting a better source resolution for vehicular and other combustion sources is now well recognised (Watson *et al*. 2002, Thurston *et al*. 2005). On the other hand, a comprehensive comparison of different and widely used source apportionment methods showed that, in general, results from various methods are consistent (Thurston *et al*. 2005). The comparison by Thurston *et al*. (2005) also indicated that the variability between results from different source apportionment methods had only a minor effect on the outcome of statistical analyses on associations of source resolved particulate matter and mortality. Some work has been done to compare receptor and emission-oriented source apportionment methods (Held *et al*. 2005, Marmur *et al*. 2006). These comparisons arrived at somewhat different conclusions regarding the level of agreement of source contribution estimates from these conceptually very different modelling approaches.

In our work the agreement between the methods in terms of the daily variations of PM$_{2.5}$ components was good for sea salt and crustal material, the chemical composition of which is well documented for a successful employment in mass closure. Both sea salt and crustal source components have reliable marker elements, which makes interpretation of these components relatively easy in a multivariate source apportionment. However, in terms of average source contributions to PM$_{2.5}$, the relative difference between the two methods was largest for the sea salt component. For the more complex components of PM, the fundamental differences between these two methods rendered it difficult to obtain comparable results from the two methods — both in terms of the average PM$_{2.5}$ mass and the daily variations of modelled components. Our results support those by Hopke *et al*. (2006) who compared various source apportionment methods in two U.S. cities and found that crustal, salt, oil and sulfate sources were most unambiguously identified, whereas they found more between-method variability in contributions from traffic and vegetative burning sources.

Our results suggest that in order to be successful and informative, a mass closure analysis requires a more comprehensive chemical characterization of PM samples than what was done in this study. Similarly, the lack of suitable source markers lowers the power of PCA to identify the sources affecting a receptor site.

Source apportionment is being conducted using various methods depending on the study setting and data that are available for the source analyses. Since the method itself plays a crucial part in defining the final outcome of a source apportionment exercise, we recommend to apply and compare the results of at least two different methods in any particular study.

Acknowledgements: We thank Matti Jantunen, Ph.D., and Tarja Yli-Tuomi, Ph.D., for their useful comments and enlightening discussions during the writing process, and the Environmental Office of the Helsinki Metropolitan Area Council and our colleagues at the Helsinki University Department of Physics for their co-operation during the field work. Financial support for Dr. Vallius from The Finnish Cultural Foundation of Northern Savo and the Maj and Tor Nessling Foundation is warmly acknowledged.

References

Andrews E., Saxena P., Musarra S., Hildemann L.M., Koutrakis P., McMurry P.H., Olmez I. & White W.H. 2000. Concentration and composition of atmospheric aerosols from the 1995 SEAVS experiment and a review of the closure between chemical and gravimetric measurements. *J. Air Waste Manage. Assoc*. 50: 648–664.

Eldred R.A., Cahill T.A. & Flocchini R.G. 1997. Composition of PM$_{2.5}$ and PM$_{10}$ aerosols in the IMPROVE network. *J. Air Waste Manage. Assoc*. 47: 194–203.

Gordon G.E. 1988. Receptor models. *Environ. Sci. Technol*. 22: 1132–1142.

Harrison R.M., Jones A.M. & Lawrence R.G. 2003. A pragmatic mass closure model for airborne particulate matter at urban background and roadside sites. *Atmos. Environ*. 37: 4927–4933.

Held T., Ying Q., Kleeman M.J., Schauer J.J. & Fraser M.P. 2005. A comparison of the UCD/CIT air quality model and the CMB source–receptor model for primary airborne particulate matter. *Atmos. Environ*. 39: 2281–2297.

Henry R.C. 1977. History and fundamentals of multivariate air quality receptor models. *Chemometrics Intellig. Lab. Syst*. 37: 37–42.

Henry R.C., Lewis C.W., Hopke P.K. & Williamson H.J. 1984. Review of receptor model fundamentals. *Atmos. Environ*. 18: 1507–1515.

Hopke P.K. 2003. Recent developments in receptor modeling. *J. Chemometrics* 17: 255–265.

Hopke P.K., Ito K., Mar T., Christensen W.F., Eatough D.J., Henry R.C., Kim E., Laden F., Lall R., Larson T.V., Liu H., Neas L., Pinto J., Stölzel M., Suh H., Paatero P. & Thurston G.D. 2006. PM source apportionment and health effects: 1. Intercomparison of source apportionment results. *J. Expo. Sci. Environ. Epidemiol.* 16: 275–286.

Koistinen K.J., Edwards R.D., Mathys P., Ruuskanen J., Künzli N. & Jantunen M.J. 2004. Sources of fine particulate matter in personal exposures and residential indoor, residential outdoor and workplace microenvironments in the Helsinki phase of the EXPOLIS study. *Scand. J. Work Environ. Health* 30(S2): 36–46.

Koistinen K.J., Kousa A., Tenhola V., Hänninen O., Jantunen M.J., Oglesby L., Künzli N. & Georgoulis L. 1999. Fine particle (PM2.5) measurement methodology, quality assurance procedures, and pilot results of the EXPOLIS study. *J. Air Waste Manage. Assoc.* 49: 1212–1220.

Marmur A., Park S.-K., Mulholland J.A., Tolbert P.E. & Russell A.G. 2006. Source apportionment of $PM_{2.5}$ in the southeastern United States using receptor and emissions-based models: Conceptual differences and implications for time-series health studies. *Atmos. Environ.* 40: 2533–2551.

Ojanen C., Pakkanen T., Aurela M., Mäkelä T., Meriläinen J., Hillamo R., Aarnio P., Koskentalo T., Hämekoski K., Rantanen L. & Lappi M. 1998. *Size distribution, composition and sources of inhalable particles in the Helsinki metropolitan area*. Helsinki Metropolitan Area Council, Publication Series C1998(7).

Paatero P., Hopke P.K., Song X.H. & Ramadan Z. 2002. Understanding and controlling rotations in factor analytic models. *Chemometrics Intellig. Lab. Syst.* 60: 253–264.

Pakkanen T.A., Loukkola K., Korhonen C.H., Aurela M., Mäkelä T., Hillamo R.E., Aarnio P., Koskentalo T., Kousa A. & Maenhaut W. 2001. Sources and chemical composition of atmospheric fine and coarse particles in the Helsinki area. *Atmos. Environ.* 35: 5381–5391.

Pekkanen J., Timonen K.L., Tiittanen P., Vallius M., Lanki T., Sinkko H., Ruuskanen J., Mirme A., Kulmala M., Vanninen E., Bernard A., Ibald-Mulli A., Wolke G., Stadeler M., Tuch T., Kreyling W., Peters A., Heinrich J., de Hartog J., Oldenwenig M., Kos G., ten Brink H., Khlystov A., van Wijnen J., Brunekreef B. & Hoek G. 2000. *ULTRA Exposure and risk assessment for fine and ultrafine particles in ambient air — study manual and data book*. National Public Health Institute, Helsinki.

SAS Institute Inc. 1999. *SAS/STAT user's guide*. SAS Institute Inc, Cary, NC, U.S.A.

Seinfeld J.H. & Pandis S.N. 1998. *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons, Inc., New York.

Shiraki R. & Holmen B.A. 2002. Airborne respirable silica near a sand and gravel facility in central California: XRD and elemental analysis to distinguish source and background quartz. *Environ. Sci. Technol.* 36: 4956–4961.

Thurston G.D., Ito K., Mar T., Christensen W.F., Eatough D.J., Henry R.C., Kim E., Laden F., Lall R., Larson T.V., Liu H., Neas L., Pinto J., Stölzel M., Suh H. & Hopke P.K. 2005. Workgroup report: workshop on source apportionment of particulate matter health effects - intercomparison of results and implications. *Environ. Health Perspect.* 113: 1768–1774.

Thurston G.D. & Spengler J.D. 1985. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmos. Environ.* 19: 9–25.

Vallius M., Lanki T., Tiittanen P., Koistinen K., Ruuskanen J. & Pekkanen J. 2003. Source apportionment of urban ambient PM2.5 in two successive measurement campaigns in Helsinki, Finland. *Atmos. Environ.* 37: 615–623.

Viidanoja J., Sillanpää M., Laakia J., Kerminen V.M., Hillamo R., Aarnio P. & Koskentalo T. 2002. Organic and black carbon in $PM_{2.5}$ and PM_{10}: 1 year of data from an urban site in Helsinki, Finland. *Atmos. Environ.* 36: 3183–3193.

Wang Y., Huang K., Li C., Mi H., Luo J. & Tsai P. 2003. Emissions of fuel metals content from a diesel vehicle engine. *Atmos. Environ.* 37: 4637–4643.

Watson J.G., Chow J.C. & Pace T.G. 1991. Chemical mass balance. In: Hopke P.K. (ed.), *Receptor modeling for air quality management*, Elsevier Science Publishers, New York, pp. 83–116.

Watson J.G., Zhu T., Chow J.C., Engelbrecht J., Fujita E.M. & Wilson W.E. 2002. Receptor modeling application framework for particle source apportionment. *Chemosphere* 49: 1093–1136.