Elsevier Editorial System(tm) for Science of the Total Environment
Manuscript Draft

Manuscript Number: STOTEN-D-15-02718R1

Title: Implementing constrained multi-time approach with bootstrap analysis in ME-2: an application to PM2.5 data from Florence (Italy)

Article Type: Research Paper

Keywords: advanced receptor models; source apportionment; constrained ME-2; bootstrap

Corresponding Author: Dr. Roberta Vecchi,

Corresponding Author's Institution: University of Milan

First Author: Alice Crespi

Order of Authors: Alice Crespi; Vera Bernardoni; Giulia Calzolai; Franco Lucarelli; Silvia Nava; Gianluigi Valli; Roberta Vecchi

Abstract: Advanced receptor models have been recently developed and tested in order to improve the resolution of apportionment problems reducing rotational ambiguity of results and aiming at identifying a larger number of sources. In particular, multi-time model is a factor analysis method able to compute source profiles and contributions using aerosol compositional data with different time resolutions. Unlike traditional factor analysis, each measured value can be inserted into multi-time model with its original time schedule, thus all temporal information can be effectively used in the modeling process. In this work, multi-time model was expanded in order to impose constraints on modeled factors aiming at improving the source identification. Moreover, as far as we know for the first time, a suitable bootstrap technique was implemented in the multi-time scheme to estimate the uncertainty of the final constrained solutions. These implemented approaches were tested on a PM2.5 (particulate matter with aerodynamic diameter < 2.5 μm) dataset composed of 24-hours samples collected during one year and hourly data sampled in parallel for two shorter periods in Florence (Italy). The daily samples were chemically characterized for elements, ions and carbonaceous components while elemental concentrations only were available for high-time resolved samples. The application of the advanced model revealed the major contribution from traffic (accounting for 37% of PM2.5 as annual average) and allowed an accurate characterization of involved emission processes. In particular, exhaust and non-exhaust emissions were identified. The constraints imposed in the continuation run led to a better description of the factor associated to nitrates and also of biomass burning profile and the bootstrap results gave useful information to assess the reliability of source apportionment solutions. Finally, the comparison with the results computed by ME-2 base model applied to daily and hourly compositional data separately demonstrated the advantages provided by the multi-time approach.

Implementing constrained multi-time approach with bootstrap analysis in ME-2: an application to PM2.5 data from Florence (Italy)

A. Crespi[1], V. Bernardoni[1], G. Calzolai[2,3], F. Lucarelli[2,3], S. Nava[3], G. Valli[1], R. Vecchi[1]

[1] *Dept. of Physics, Università degli Studi di Milano & INFN, Via Celoria 16, 20133, Milano, Italy*

[2] *Dept. of Physics, University of Florence, Via Sansone 1, 50019, Sesto Fiorentino (FI), Italy*

[3] *INFN – Section of Florence, Via Sansone 1, 50019, Sesto Fiorentino (FI), Italy*

Corresponding author: Roberta Vecchi (roberta.vecchi@unimi.it)

Abstract

Advanced receptor models have been recently developed and tested in order to improve the resolution of apportionment problems reducing rotational ambiguity of results and aiming at identifying a larger number of sources. In particular, multi-time model is a factor analysis method able to compute source profiles and contributions using aerosol compositional data with different time resolutions. Unlike traditional factor analysis, each measured value can be inserted into multi-time model with its original time schedule, thus all temporal information can be effectively used in the modeling process. In this work, multi-time model was expanded in order to impose constraints on modeled factors aiming at improving the source identification. Moreover, as far as we know for the first time, a suitable bootstrap technique was implemented in the multi-time scheme to estimate the uncertainty of the final constrained solutions. These implemented approaches were tested on a PM2.5 (particulate matter with aerodynamic diameter < 2.5 µm) dataset composed of 24-hours samples collected during one year and hourly data sampled in parallel for two shorter periods in Florence (Italy). The daily samples were chemically characterized for elements, ions and carbonaceous components while elemental concentrations only were available for high-time resolved samples. The application of the advanced model revealed the major contribution from traffic (accounting for 37% of PM2.5 as annual average) and allowed an accurate characterization of involved emission processes. In particular, exhaust and non-exhaust emissions were identified. The constraints imposed in the continuation run led to a better description of the factor associated to nitrates and also of biomass burning profile and the bootstrap results gave useful information to assess the reliability of source apportionment solutions. Finally, the comparison with the results computed by ME-2 base model applied to daily and hourly compositional data separately demonstrated the advantages provided by the multi-time approach.

Keywords*: advanced receptor models, source apportionment, constrained ME-2, bootstrap*

1. Introduction

Negative effects induced by atmospheric aerosol on the environment (IPCC, 2013) and human health (Pope and Dockery, 2006; WHO, 2013) have been assessed so far. In order to formulate strategies to reach air quality objectives in a specific area, air quality modeling could be useful to assess impacting sources and their contributions (Viana et al., 2008).

The most recent developments in monitoring techniques allow aerosol components measurement with higher and higher time resolution (typically from 24-h down to minutes) thus giving detailed information that can be used also in source apportionment studies. Besides 24-h resolution data carried out in compliance with regulatory purposes, hourly samples of different PM fractions can be collected by higher resolution instruments, such as streaker samplers and rotating drum impactors (RDI). IBA (Ion Beam Analysis) and SR-XRF (Synchrotron Radiation-induced X-Ray Fluorescence Spectrometry) techniques have been widely applied to streaker and RDI samples, respectively, providing elemental concentrations on hourly basis (Lucarelli et al., 2014; Visser et al., 2015). These data are usually combined with compositional information extracted from lower resolution data to gain more and more information producing a dataset of mixed time resolution. Traditional multivariate techniques, such as PCA (Principal Component Analysis) or PMF-2 (Positive Matrix Factorization), need to reduce all data to a single time schedule; either the modeler has to average high-time resolution data over the longest sampling interval - typically 24-h - or to interpolate the low-resolution data to the shortest sampling period (Zhou et al., 2004). Therefore, the source apportionment modeling is typically limited to the separate analysis of high-time or daily samples thus either important source markers given by the complete chemical speciation often available on daily samples or the high-time resolution information get lost.

In order to properly take into account all the temporal details provided by experimental data without averaging or interpolating, an advanced receptor model was developed (Zhou et al., 2004; Ogulei et al., 2005) using Multilinear Engine (ME-2) which is a flexible algorithm program implementing a scripting language able to solve multivariate receptor problems (Paatero, 1999). By the multi-time approach, apportionment problems can be solved inserting each data value in its original time schedule.

By means of the flexibility of ME-2 program, in this work the multi-time model has been expanded 1) to perform a continuation run; and 2) to assess uncertainties on the solutions. More precisely, during the continuation run physically meaningful constraints can be imposed on the resolved sources in order to reduce rotational ambiguity and enhance the problem solution. Usually the constraints derive from a-priori knowledge on the features of sources influencing the monitoring site. An important step in receptor modeling is also the evaluation of the uncertainty on the final

solution. As highlighted in Paatero et al. (2014), the uncertainty in receptor analysis arises from three main causes: (1) random errors in data values; (2) rotational ambiguity and (3) modeling errors. In order to capture these different sources of uncertainty, three estimation methods were developed and are now available in the latest version of EPA PMF 5.0 (EPA, 2014): bootstrapping, dQ-controlled displacement of factor elements and bootstrapping enhanced with displacement. In this work, a method based on bootstrap technique for estimating uncertainty on model results has been implemented in multi-time scheme in order to assess the reliability of an advanced factor analysis solution. Apart the very recent work by Liao et al. (2015), this is one of the very few examples of multi-time resolution ME-2 analysis with constraints reported in the literature. Moreover, as far as we know, this is the first application of bootstrap analysis on multiple time resolution data.

The ME-2 model implemented in this work was applied to PM2.5 samples collected at a traffic site in Florence (Italy) both with daily and hourly time resolution in the framework of the regional project PATOS2 - Particolato Atmosferico in TOScana 2 – (PATOS2, 2014) as described in detail below.

2.Material and methods

2.1 Aerosol measurements

The PM2.5 aerosol samples analyzed by multi-time model were collected from April 2009 to March 2010 at a traffic site in Florence (Italy). The monitoring site was located at one of the major roads in the city beltway, close to a traffic light and a bus stop. In addition, there is a mechanic workshop behind the site and many car parks are located all along the road sides.

PM2.5 daily data were collected in parallel on PTFE and quartz-fibre filters. Samples on Teflon filters were collected every other day while samples on quartz-fibre filters everyday. PM2.5 mass concentrations were determined by gravimetric technique on Teflon filters. Filters were weighed by an analytical balance after a 24-hours conditioning period in controlled temperature (20 ± 1° C) and relative humidity (50 ± 5 %). Moreover, a ionization gun was used to avoid electrostatic effects. They were then analyzed by PIXE (Particle Induced X-Ray Emission) to obtain the elemental composition (Lucarelli et al., 2010) and by Ion Chromatography (IC) to determine water-soluble inorganic ions (Morganti et al., 2007). Concentrations of Total Carbon (TC), Organic Carbon (OC) and Elemental Carbon (EC) were retrieved by thermo-optical-transmission analysis on 1.5 cm^2 quartz-fibre filter portions using the Sunset Lab analyzer with NIOSH-like protocol (CEN/TC 264, 2011). The accuracy of PIXE elemental concentrations was determined by summing up independent uncertainties on standard samples thickness (5%) and X-rays counting statistics (from 2% to 20% or

higher when values approach detection limits); uncertainties on carbonaceous components and ions are in the range 5-10%. Irrespective of the analytical technique, elemental concentrations are affected by additional uncertainties (~ 5%) due to other sampling parameters (e.g. flow rate variations).

During shorter periods in early autumn (23 September - 7 October 2009) and in winter (14 - 28 January 2010) hourly PM2.5 samples were collected by a two-stage streaker sampler. Briefly, this device separates particles on different stages using a pre-impactor and an impactor. PM with aerodynamic diameter D_{ae} >10 µm is removed by pre-impactor while the impactor deposits aerosol coarse fraction (2.5 µm < D_{ae} < 10 µm) on a Kapton foil and the fine fraction (D_{ae} < 2.5 µm) is collected on a Nuclepore polycarbonate membrane. The two collecting supports are kept in rotation at constant speed producing a circular continuous deposition on both stages (D'Alessandro et al., 2003).

PIXE analysis of these samples by a properly collimated proton beam, scanning the deposit in steps corresponding to 1-hour of aerosol sampling, provided the elemental concentrations with hourly time resolution (Lucarelli et al., 2010 and 2014).

2.2 Model description

Multivariate receptor models represent a valuable tool to solve source apportionment problems of airborne particulate matter. From the input matrix containing concentration data of mass and chemical species collected at the monitoring site, receptor models estimate the number of sources influencing the area and decompose the initial matrix into two matrices representing source time contributions and chemical compositions.

In receptor modeling, the principle of mass conservation is assumed and a mass balance analysis is performed to apportion PM sources. Using the index *p* for factors, *i* for samples and *j* for species, the base equation in a factor analysis model can be written as below:

$$x_{ij} = \sum_{p=1}^{P} g_{ip} f_{pj} + e_{ij} \qquad (1)$$

where x_{ij} is the concentration of the *j-th* species in the *i-th* sample, f_{pj} is the mass fraction of the *j-th* species from the *p-th* source, g_{ip} is the contribution of the *p-th* source in the *i-th* sample and e_{ij} are the residuals. Besides factor chemical profiles and time series of source contributions, model outputs include also the Explained Variations (EV), i.e. the fraction of a species variation explained by each source. Indicating with s the uncertainties of input data, EV values are computed by the model as follows (Paatero, 2010):

$$EV_{pj} = \frac{\sum_{i=1}^{N}\left(\frac{|g_{ip}f_{pj}|}{s_{ij}}\right)}{\sum_{i=1}^{N}\left(\frac{\sum_{p=1}^{P}|g_{ip}f_{pj}| + |e_{ij}|}{s_{ij}}\right)} \quad (2)$$

The multi-time model was developed in order to insert every measured concentration data with its original time schedule. Each concentration is influenced by emissions from several sources; therefore, in the modeling source contributions need to be averaged in order to refer to the same sampling period of the corresponding concentration value. To this aim, the main equation (Eq.1) is modified as below (Zhou et al., 2004):

$$x_{sj} = \frac{1}{t_{s2} - t_{s1} + 1}\sum_{p=1}^{P} f_{jp}\sum_{i=t_{s1}}^{t_{s2}} g_{ip}\eta_{jm} + e_{sj} \quad (3)$$

where s is the sample number, j represents the species, t_{s1} and t_{s2} are starting and ending time, respectively, for the *s-th* sample described by the number of time units. The time unit corresponds to the shortest sampling interval of available data (e.g. 1 hour for the dataset used in this work). As in Eq.1, x_{sj} is the concentration of the *j-th* species in the *s-th* sample, f_{jp} is the mass fraction of the *j-th* species from the *p-th* source, g_{ip} is the contribution of the *p-th* source during the sampling period for the *s-th* sample and e_{sj} are the residuals. η_{jm} are adjustment factors for replicated species measured by more than one analytical method with different time resolutions. Adjustment factors for non-replicated species can be set to unity by default while an adjustment factor next to unity indicates a good agreement between the different measurements (Ogulei et al., 2005). In this work, the adjustment factors were always set to unity as the same analytical technique (i.e. PIXE) was used to measure both daily and hourly elemental concentrations.

Since some sources could contain no or few species measured with high time resolution, a regularization equation (Eq.4) is used to smooth the time series of their contributions:

$$g_{i+1p} - g_{ip} = 0 + \varepsilon_i \quad (4)$$

where g_{ip} is the source contribution of the *p-th* source during the *i-th* time unit and ε_i is the residual.

Eqs.3 and 4 were solved using Multilinear Engine algorithm (ME) (Paatero, 1999) which provides a weighted least-squares solution to the problem minimizing the sum of squares, i.e. the so-called object function Q:

$$Q = Q_{main} + Q_{aux} = \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{e_{ij}}{\sigma_{ij}} + \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{e'_{ij}}{\sigma'_{ij}} \quad (5)$$

As shown in Eq.5, both the main equation (Eq.3) and the auxiliary equations (e.g. regularization equations, pulling equations and constraints) are taken into account in the object function Q; σ are the uncertainties of input data and σ' are the uncertainties related to the auxiliary equations.

In this work, the model was expanded in order to perform the continuation run. Indeed, starting from the base-case solution obtained by the multi-time model, suitable constraints to the identified factors were inserted in the model in the form of auxiliary equations. As examples, specific marker ratios can be constrained to stoichiometric values and ambiguous chemical contributions in a source profile can be pulled to zero or to a specific value (Amato and Hopke, 2012). The strength of each constraint is controlled by the Q_{main} increase limit set by the user, the greater is the allowed increase, the stronger is the effect of the constraining (Paatero and Hopke, 2009). Due to the imposed conditions, resolved factors can be more clearly identified as the rotational ambiguity on the solutions is reduced.

In addition, a method for estimating uncertainty in the multi-time model solutions was here implemented in order to evaluate the robustness of our results. It is based on the bootstrap analysis, inspired by one of the available techniques from the EPA PMF 5.0 script. Factors are initialized by the final solution values and the analysis is repeated many times (typically more than 100 times). As in the EPA PMF 5.0 script, the implemented bootstrap analysis is based on three different steps: *resampling*, *reweighting* and *random rotational pulling*. In particular, during the last phase of the procedure some random-selected elements of time contribution matrix are pulled up or down toward a specific value in order to assess the solution uncertainty due to rotational ambiguity. In this work, the bootstrap method was slightly modified in order to properly take into account the different dimensions of input and output matrices used in multi-time approach.

The variables used for the model were classified in strong, weak or bad according to the noise to signal ratio (S/N) criterion (Paatero, 2013; Norris and Duvall, 2014) and for measured, missing and below MDL (minimum detection limit specific for each component) data the procedure described in Polissar et al. (1998) was followed. Finally, for multi-time analysis 22 daily variables (PM2.5 mass, Na, Mg, Al, Si, Cl, K, Ca, Ti, V, Mn, Fe, Ni, Cu, Zn, Br, Pb, SO_4^{2-}, NO_3^-, NH_4^+, OC and EC) and 14 hourly species (Na, Mg, Al, Cl, K, Ca, Ti, V, Mn, Fe, Ni, Cu, Zn and Pb) were chosen. Input matrix shows a sub-block structure where the concentrations of the chemical species measured by the same method are grouped together and the row number corresponds to the number of available samples. For each row the start, the duration and the end of the sampling have to be expressed as time units. In addition, each column containing concentration values must be followed by a column reporting uncertainty on concentrations. The input matrix for this work was composed by two sub-

blocks with 776 samples and the total number of time units was 4104. In Figure 1 an example of the input data used for multi-time ME model is shown.

3. Results and discussion

In this section, the base-case and constrained solutions are discussed separately. The base-case results were obtained without imposing any constraint and were useful to get a preliminary identification of the factor profiles and contributions. In order to enhance the chemical profiles obtained and also to improve the interpretation of factors, some constraints were imposed on the base-case solutions and the continuation run was performed. Major improvements obtained from the continuation run analysis are then described and compared to the base case results and the uncertainty analysis results on the final constrained solution are reported too. Finally, the multi-time model solutions were compared to those obtained from the application of ME-2 base model to PM2.5 hourly and daily data separately, in order to highlight the benefits to insert multi-time resolution details in the same modeling process.

3.1 Base run results

Before the modeling phase, a mass closure analysis was performed on daily concentration data. In order to take into account contributions from oxidized forms of some aerosol chemical species and other undetected components, preferential oxides for metals and crustal elements were calculated and Organic Carbon (OC) was converted in Organic Matter (OM). The OC-to-OM conversion coefficient, whose value is typically in the range 1.2-2.3, was set to 1.4 which can be considered representative for an urban area, as reported for example in Turpin and Lim (2001). The reconstructed mass explained about 95% of the measured PM2.5 mass and OM represented the major contribution (about 45%).

In addition, the ion balance was calculated considering SO_4^{2-}, NO_3^-, NH_4^+, Ca^{2+}, K^+, Mg^{2+}, Na^+ and Cl^-. The balance showed an excess of anions (r^2 =0.97 and slope m=0.85, see Figure S1) likely due to the presence of ammonium bisulphate (NH_4HSO_4).

In Table S1 mean, median, standard deviation, quartiles, maximum and minimum concentration values of the measured species are reported.

After carrying out 30 runs with ME multi-time analysis, a 7-factor solution was selected corresponding to the minimum Q value (Q=5026.44). It is noteworthy that reducing or increasing the number of factors, the solutions showed ambiguous chemical profiles and inconsistent mass closures (i.e. the sum of species contributions exceeded the 100% of total mass in some factor profiles). For the selected solution, most of the scaled residual were between ±3 with a random

distribution of positive and negative values; only for Cl, the residuals showed a longer tail on positive values. Since the mass is reconstructed on hourly basis by the multi-time model, it is not directly comparable with the mass measured with 24-h resolution and therefore a regression analysis cannot be performed. However, the computed hourly mass concentrations were averaged and the daily mean values were compared with the 24-h mass measurements obtaining a good agreement with a regression coefficient of 0.94 and r^2=0.95 (Figure S2). Finally, scaled source profiles and contributions were obtained by dividing each profile by the corresponding mass fraction and multiplying each factor score by the same value.

The seven factors resolved by ME-2 were tentatively assigned to specific aerosol sources on the basis of their chemical profiles and EV. Two of them are likely related to traffic (exhaust and non-exhaust emissions) while the others were associated to biomass burning, soil dust, sulphate, nitrate and aged sea salt. In Figure 2 contributions of each species to the chemical profile (μg/μg, grey bars) as well as EV values (black points) are represented. In Table 1 average source contributions to PM2.5 - both absolute (μg/m^3) and relative (%) concentrations - are reported for summer and winter time. The source contribution temporal patterns computed by multi-time approach are represented in Figure S3.

Two factors were assigned to traffic emissions: the one which mainly contributed to total PM2.5 mass concentration showed high EV and relevant concentration values for EC, OC, Zn and Pb while the other was characterized by high values for Ca, Cu, Fe, but also for Ti, Al and other crustal elements. Referring to previous literature works (Pant and Harrison, 2013; Amato et al., 2009) these factors were identified as traffic exhaust and non-exhaust emissions, respectively, due to the presence of specific markers as outlined in the following.

Traffic exhaust emissions account for particles which are typically in the fine fractions and are generated by the vehicle combustion of fossil fuels and directly emitted from the tailpipe. In particular, carbonaceous compounds represent the most useful markers for this source (EC and OC) and their relative contributions depend strongly on engine type and age of vehicles. Indeed, gasoline engines are known to release higher OC concentrations while diesel engines emit more EC (Ntziachristos et al., 2007). This factor represents the main contribution to total PM2.5 mass in both seasons and its profile is composed mainly by EC and OC which account for 57% and 41%, respectively. The obtained OC/EC ratio results 0.72, comparable with literature values for primary emissions, e.g. Pio et al. (2011). The presence of Si was previously found in the PM1 profile of a factor associated to a traffic source (Vecchi et al., 2008) and negligible contributions in the profile derive from heavy metals, particularly from Zn and Pb. On a yearly basis, this factor associated to exhaust emissions contributed for about 37% to PM2.5 mass.

The traffic non-exhaust emissions are typically due to tyre and brake wear, road surface abrasion, wear and corrosion of other vehicle components such as the clutch and resuspension of road surface dust (Pant and Harrison, 2013). They are typically characterized by trace metals (e.g. Cu, Fe, Zn, Ca, Mn) even if their composition is site-dependent. The identification of traffic non-exhaust emissions is relevant because of their contribution to heavy metals and carcinogenic components in atmosphere but also in relationship with the exceedances of air quality guidelines and limits (Amato et al., 2011). The chemical profile of this factor is composed mainly by Fe, Ca and Cu, with contributions from other species, such as Al, Ti and Mg, generally attributed to crustal components. The presence of EC confirms the anthropogenic nature of this factor while the significant contribution of SO_4^{2-} needs to be more investigated even if a sulphate contribution to road material resuspended by traffic has been already observed by Amato et al. (2009). The contribution from this non-exhaust emissions factor is comparable in both seasons and accounts for about 7% as annual average.

Biomass burning factor is often identified by levoglucosan and its isomers (mannosan and galactosan) generated by cellulose pyrolysis. As levoglucosan concentrations were not measured in this study, biomass burning emissions - which include contributions from pellet and wood heating, agricultural activities and forest fires - were characterized by other inorganic elements/compounds, mainly potassium salts (KCl, K_2SO_4 and KNO_3), metals (Zn, Pb) and refractory compounds including Ca, Mg and Si. Actually, the use of K could be disadvantageous because it has multiple source emissions (Pachon et al., 2013) and the availability of levoglucosan could have been more helpful to better resolve the contributions from biomass burning reducing the interference from other sources. Since K emitted from biomass burning is water-soluble while K from mineral dust is not, the K^+ ion water-soluble fraction measured by ion chromatography (IC) could represent a suitable marker for biomass burning emissions especially in fine aerosol fractions (Reche et al., 2012). K^+ was measured during this campaign and, due to the high correlation between K and K^+ in our dataset (Figure S4), in this work K was used as a tracer for the biomass burning source (Nava et al., 2015). Indeed, K is an element detected by PIXE on both daily and hourly samples thus allowing the use of unitary adjustment factors in the multi-time model (as described in paragraph 2.2). The modeled chemical profile (Figure 2b) was dominated mainly by OC and K. As already mentioned above the presence of Cl and SO_4^{2-} could be ascribed to potassium salts. Although already observed in previous literature works (Reche et al., 2012; Gildemeister et al., 2007; Lee and Hopke, 2006 as examples), the total lack of an EC contribution in the chemical profile of biomass burning was considered a bit suspicious and it needed further investigations (see paragraph 3.2 for more details). The typical seasonal pattern with relevant contribution values (up to 30 μg/m^3 as

hourly concentration) during winter time was observed (Figure S3b). A strong impact of biomass burning for Central Italy has been already shown by Nava et al. (2015) during a PM10 sampling campaign carried out at different sites in Tuscany (Italy).

The ME-2 factor assigned to sulphate is traced by SO_4^{2-} (which accounts for about 78% of the PM2.5 mass in the factor profile) and NH_4^+. Secondary sulphate is produced in atmosphere by photochemical oxidation of sulphur oxides emitted as gases from local sources or long range transports (Seinfeld and Pandis, 1998). This factor is characterized by a clear seasonality with major contributions during summer months when a more intense insolation drives photochemical processes with the accumulation of reaction products in atmosphere. Moreover, the atmospheric circulation conditions in summer foster the transport of SO_2 and sulphate which are produced on regional scale. The SO_4^{2-} to NH_4^+ ratio for this factor is 3.3, i.e. it is intermediate between the stoichiometric ratio for $(NH_4)_2 SO_4$ (2.7) and that relative to ammonium bisulphate (5.3).

The nitrate factor is characterized by NO_3^- and NH_4^+ as major contributors in the chemical profile accounting for 45% and 9% of PM2.5 mass, respectively. Nitrates represent a secondary component due to oxidation of the nitrogen oxides produced by local combustion processes, typically traffic and heating systems in urban areas. The temporal pattern shows a strong seasonality with high values during winter months and very low values during summer. This effect could be explained both by the residential heating emissions during colder months and by the atmospheric gas-to-particle equilibrium shifted towards the gaseous phase during summertime as well as by negative sampling artifacts (Vecchi et al., 2009). NO_3^-/NH_4^+ ratio is 4.8, i.e. higher than the expected stoichiometric value for ammonium nitrate (3.4). This higher ratio might be due to the influence from other sources (such as biomass burning processes which produce KNO_3), or gaseous pollutants reactions with mineral particles (such as Ca bearing particles forming $Ca(NO_3)_2$) during long-range transport episodes as shown in Perrone et al. (2013). There is also a relevant contribution of EC in the factor profile. As EC has a primary origin, its presence in this factor associated to secondary nitrate is suspicious and will be investigated in the following paragraph.

The soil dust factor represents particulate matter deposited on the ground and resuspended by natural (e.g. wind) or anthropogenic agents (e.g. traffic) as well as due to medium- and long-range transport events. Principal tracers for this factor are crustal elements, such as Al, Si and Ti (Viana et al., 2008; Bernardoni et al., 2011 and therein in cited literature). The Al/Si ratio in the ME-2 chemical profile factor is 0.30, which is comparable to the value expected for average crustal matter (Mason, 1966). Both relative and absolute contributions of the soil dust factor to PM2.5 mass are higher during summertime (6%, 1.1 μg/m^3) than wintertime (3%, 0.7 μg/m^3). Indeed, stronger

atmospheric turbulence and drier soil conditions foster particles resuspension during summertime, as often observed in Mediterranean countries (Vecchi et al., 2008).

The factor assigned to aged sea salt represents the contribution of marine air masses transported to receptor site from the coast, about 80 km far from Florence. Tracers for this source are Na, Mg and Cl; EV for Na e Mg show the highest values while their contributions in source profiles are respectively 41% and 6% of the emitted PM2.5 mass. Na/Mg ratio is 0.13, in agreement with that expected for bulk-sea water (0.12). Cl/Na ratio is 0.06, significantly lower than the typical value for fresh sea salt particles (1.8); this is likely a consequence of Cl deplection effects due to heterogeneous reactions involving NaCl during transport to the monitoring site (D'Alessandro et al., 2003). This source represents the lowest contribution to PM2.5 mass (i.e. 4%) and its average contribution is about 0.6 $\mu g/m^3$.

3.2 Continuation run: results

Although the factors were quite clearly resolved in the base-case modeling reported so far, the obtained solution can be improved imposing proper constraints. In particular, two constraints have been imposed on the nitrate profile in order to enhance the characterization of this source. First of all, the NO_3^-/NH_4^+ ratio in the chemical profile was constrained to the stoichiometric value for ammonium nitrate (NO_3^-/NH_4^+ = 3.4) as in the base case it resulted a bit higher (4.8) suggesting nitrate and ammonium could have been over- and under-estimated, respectively. The strength associated to this constraining equation was defined by setting the Q_{main} increase limit to 100 units. Secondly, the presence of EC in a factor which was attributed mainly to a secondary PM component was not expected therefore a pulling equation was applied to EC in the profile in order to force to zero its contribution to the secondary nitrate factor. For this constraint the Q_{main} allowed increase was set to 10 units implying a weaker effect associated to this equation than the first one.

The small increase in the object function (i.e. 20 units or 0.4% more than the base case) obtained in the continuation run represented a useful diagnostic instrument to evaluate the reliability of the imposed constraints and the overall goodness of the analysis. Although a definitive criterion does not exist yet and tolerance depends on the strength and type of constraints, increments of a few tens are generally considered acceptable (Paatero and Hopke, 2009).

Both the targets imposed on the secondary nitrate factor were achieved. In Figure 3c, the resulting profile is compared to that provided by the base run. The content of EC is zeroed and the contributions from ionic compounds were also modified as NH_4^+ increased and NO_3^- slightly decreased. Nevertheless, a relevant amount of SO_4^{2-} appeared in the factor profile and its origin is

not clear yet (see below); in addition, a fraction of SO_4^{2-} was still contained in every chemical profile, thus SO_4^{2-} presence has to be investigated further.

Besides the secondary nitrate profile, it is necessary to study the effects of the imposed constraints on the other source profiles to check if their chemical composition significantly changed. Traffic exhaust emissions, traffic non-exhaust emissions, sulphate, soil dust and aged sea salt profiles did not change while a significant amount of NO_3^- appeared in biomass burning profile. This is due to the excess of NO_3^- in secondary nitrate which - after the continuation run - was associated to biomass burning factor as a consequence of the species redistribution generated by the constraints. It is noteworthy that the nitrate presence in the biomass burning chemical profile is physically meaningful as both NO_3^- and SO_4^{2-} contributions may derive from potassium salts (KNO_3 and K_2SO_4 particularly) produced by biomass combustion processes (Nava et al., 2015).

Even if absolute and relative source apportionment values for the factors are almost unchanged in comparison with base run results (Table S2), the continuation run helped in understanding more accurately the modeled source profiles and appeared to be a suitable tool for improving results also for the multi-time model.

No improvements concerning possible appearance of EC in biomass burning source was registered after the continuation run; thus, we tested the addition of an extra constraint on the chemical profile of this factor. In particular, a pulling-to-maximum equation (Paatero, 2007) was added to the constraints mentioned above in the continuation run analysis. By this equation, the content of EC in the profile was pulled up as far as possible keeping the change in Q_{main} below a fixed value (set to 5 units in this analysis). The imposed condition was satisfied: a certain amount of EC was associated to the biomass burning chemical profile with an increase of Q_{main} of only 4 units in comparison to the continuation run analysis described above. This result suggested the presence of EC in the profile is reliable, as – at least in our experience - constraints which have not physical meaning are not achieved by the model. In addition, the suspicious amount of SO_4^{2-} in the chemical profile of nitrate was pulled to zero and this constraint was reached suggesting that the contribution of this secondary species in the profile is not relevant. Nevertheless, we chose to keep the number of imposed constraints on the final solution as small as possible and the results of these tests are shown separately in Figures S5-S6, respectively.

3.3 Continuation run: uncertainty evaluation

The bootstrap analysis provided the uncertainty associated to the constrained source profiles. The technique evaluates mostly the uncertainty due to random errors in input data. As mentioned above, also a *random rotational pulling* phase was performed, so that a fraction of rotational ambiguity can

be accounted for. During this phase, a certain number of entries of the G matrix is random-chosen and pulled up or down of a fixed amount. The criteria used for selecting entries to be pulled was here slightly modified because of the different dimensions of matrices used in multi-time approach respect to the traditional scheme. 100 repetitions were carried out and for each run the chemical profiles were computed. The interquartile range (25th-75th percentile) for each species contribution was chosen as the variability interval which should contain the best solution value. In Figure 4 the comparison between bootstrap results and modeled solutions for each source profile is shown. In this case, source profiles computed by bootstrap are not normalized and the values are expressed in $\mu g/m^3$ on a logarithmic scale. The species tracers for each source show generally small confidence ranges which include the modeled solution values while, as expected, for the species which are not specific markers in that source profile the agreement is weaker and interquartile ranges result wider. However, the agreement for species tracers could be considered as index of the goodness of the model results.

3.4 Comparison with ME-2 base model solutions

The advantages obtained by the multi-time approach here proposed can be seen comparing the base-case solutions to the results obtained from ME-2 base model performed with hourly and daily data separately.

As far as hourly data obtained by streaker sampler are concerned, it is noteworthy that a complete source apportionment cannot be computed as mass concentration is typically not available from streaker samples. In addition, on streaker samples only elemental concentrations can be detected, thus the lack of any information about the carbonaceous fraction and ionic species prevents the identification of significant contributions impacting on the investigated area (e.g. nitrates). Evaluating the elemental composition of chemical profiles obtained by the application of ME-2 base model to PM2.5 hourly data, five sources were identified: sulphate, traffic, soil dust, aged sea salt and biomass burning. The analysis with six factors was rejected because it led to a split of biomass burning profile in two factors dominated almost exclusively by K and Pb, respectively, which cannot be further identified without other additional information. The results provided from hourly data analysis are shown in Figure S7.

As regards the application of ME-2 base model to daily data only, six factors were identified: traffic, biomass burning, nitrate, sulphate, soil dust and aged sea salt. It is noteworthy that the base model was not effective in resolving the different traffic contributions revealed by multi-time approach (exhaust and non exhaust emissions). This is likely due to the absence of the high temporal details contained in the hourly data, which permit to take into account in the modeling

process the short-time variations of elemental tracers occurring in real cases. The chemical profiles of the other factors were similar to those computed by multi-time approach and also stechiometric values for nitrate and sulphate were fully comparable. Performing the same analysis on the daily dataset at seven factors, one factor had significantly high EV only for Cl and, in general, the results showed more ambiguous chemical profiles; therefore, the six-factor solution was preferred. The results and apportionment values for the six factors solution of ME-2 base model on daily data are reported in Figure S8 and Table S3.

4. Conclusions

The multi-time model allowed the joint analysis of both daily and hourly aerosol compositional data taking advantage of the detailed chemical composition related to the daily dataset and the high-time resolution patterns given by hourly elemental data. The high temporal details improved the modeling, resolved more accurately the factors and together with the use of source markers on a daily basis improved the factor-to-source assignments. In particular, the multi-time approach led to the identification of two well-distinct contributions from traffic emissions (exhaust and non-exhaust emissions) which also represent relevant PM2.5 sources, as expected considering the location of the receptor site. The factor assigned to traffic exhaust emissions gives the most relevant contribution to PM2.5 mass (37% and 8 $\mu g/m^3$ as annual average). The characterization of traffic emissions is very useful to improve the actual knowledge about the different impacts of this source on urban areas. The comparison with the results obtained from the application of ME-2 base model to daily and hourly compositional data separately demonstrated that the integration of higher resolution data can be advantageous to better distinguish these traffic emissions. In addition, the present work allowed testing the imposition of constraints on factor profiles computed by the multi-time ME-2 model to reduce rotational ambiguity and identify more accurately the sources. Two constraints were imposed on the nitrate profile leading to a better definition of this factor dominated by secondary components and – as a result of species redistribution - a more detailed profile of biomass burning factor was also obtained. Finally, a bootstrap method was implemented to evaluate the uncertainty of final constrained solutions. This method allows accounting for the uncertainty due to random errors of input data and to rotational ambiguity. Confidence intervals on the seven source profiles were computed and the general agreement to solution values suggests the goodness of the final results.

In conclusion, this application of the enhanced ME-2 revealed the importance of developing advanced receptor models which could lead to a better understanding of the sources impacting on a studied area. In particular, the multi-time model can be considered an advantageous tool to resolve

apportionment problems as high-time details allow to take into account fine scale variations of specific tracers improving identification of factors and resolving in certain situations a greater number of sources. For future studies, the integration of ancillary data (e.g. mass and/or other chemical species with hourly resolution) or longer periods with high resolution sampling could be helpful to obtain more reliable and robust solutions.

Acknowledgments

The PATOS2 project was supported by the Regional Government of Tuscany.

References

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

Amato, F., Pandolfi, M., Moreno, T., Furger, M., Pey, J., Alastuey, A., Bukowiecki, N., Prevot, A.S.H., Baltensperger, U., Querol, X., 2011. Sources and variability of inhalable road dust particles in three European cities. Atmospheric Environment, 45, 6777-6787.

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

Bernardoni, V., Vecchi, R., Valli, G., Piazzalunga, A., Fermo, P. (2011). PM10 source apportionment in Milan (Italy) using time-resolved data. Science of the Total Environment, 409, 4788-4795.

CEN/TC 264, 2011. Ambient air quality — Guide for the measurement of elemental carbon (EC) and organic carbon (OC) deposited on filters.

D'Alessandro, A., Lucarelli, F., Mandò, P.A., Marcazzan, G., Nava, S., Prati, P., Valli, G., Vecchi, R., Zucchiatti, A., 2003. Hourly elemental composition and sources identification of fine and coarse PM10 particulate matter in four Italian towns. Journal of Aerosol Science, 34, 243-259.

Gildemeister, A.E., Hopke, P.K., Kim, E., 2007. Sources of fine urban particulate matter in Detroit, MI. Chemosphere, 69, 1064-1074.

IPCC, 2013: Technical summary. In Climate Change 2013: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Website: www.ipcc.ch.

Lee, J.H., and Hopke, P.K., 2006. Apportioning sources of PM2.5 in St. Louis, MO using speciation trends network data. Atmospheric Environment, 40, S360-S377.

Liao, H.-T., Chou, C.C.-K., Chow, J.C., Watson, J.G., Hopke, P.K., Wu, C.-F., 2015. Source and risk apportionment of selected VOCs and PM2.5 species using partially constrained receptor models with multiple time resolution data. Environmental Pollution, 205, 121-130.

Lucarelli, F., Nava S., Calzolai, G., Chiari, M., Udisti, R., Marino, F., 2010. Is PIXE still a useful technique for the analysis of atmospheric aerosols? The LABEC experience. X-Ray Spectrometry, 40, 162–167.

Lucarelli, F., Calzolai, G., Chiari, M., Giannoni, M., Mochi, D., Nava, S., Carraresi, L., 2014. The upgraded external-beam PIXE/PIGE set-up at LABEC for very fast measurements on aerosol samples. Nuclear Instruments and Methods in Physics Research B, 318, 55-59.

EPA, 2014. http://www.epa.gov/heasd/research/pmf.html

Mason, B., 1966. Principles of Geochemistry. Wiley and Sons, New York.

Morganti, A., Becagli, S., Castellano, E., Severi, M., Traversi, R., and Udisti, R., 2007. An improved flow analysis-ion chromatography method for determination of cationic and anionic species at trace levels in Antarctic ice cores. Analytica Chimica Acta, 603, 190-198.

Nava, S., Lucarelli, F., Amato, F., Becagli, S., Calzolai, G., Chiari, M., Giannoni, M., Traversi, R., Udisti, R., 2015. Biomass burning contributions estimated by synergistic coupling of daily and hourly aerosol composition records. Science of Total Environment, 511, 11-20.

Norris, G., and Duvall, R., 2014. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide.

Ntziachristos, L., Ning, Z., Geller, M.D., Sheesley, R.J., Schauer, J.J., Sioutas, C., 2007. Fine, ultrafine and nanoparticle trace element compositions near a major freeway with a high heavy-duty diesel fraction. Atmospheric Environment, 41, 5684–5696.

Ogulei, D., Hopke, P.K., Zhou, L., Paatero, P., Park, S.S., Ondov, J.M., 2005. Receptor modeling for multiple time resolved species: The Baltimore supersite. Atmospheric Environment, 39, 3751-3762.

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., 2007. The Multilinear Engine (ME-2) script language. Available with the program ME-2 (me2scrip.txt).

Paatero, P., and Hopke, P.K., 2009. Rotational tools for factor analytic models. Journal of Chemometrics, 23, 91-100.

Paatero, P., 2010. User's guide for Positive Matrix Factorization programs PMF2 and PMF3, Part 2: reference.

Paatero, P., 2013. User's guide for Positive Matrix Factorization programs PMF2 and PMF3, Part 1: tutorial.

Paatero, P., Eberly, S., Brown, S.G., Norris, G.A., 2014. Methods for estimating uncertainty in factor analytic solutions. Atmospheric Measurement Techniques, 7, 781-797.

Pachon, J.E., Weber, R.J., Zhang, X., Mulholland, J.A., Russell, A.G., 2013. Revising the use of potassium (K) in the source apportionment of PM2.5. Atmospheric Pollution Research, 4, 14–21.

Pant, P., Harrison, R.M., 2013. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: A review. Atmospheric Environment, 77, 78-97.

PATOS2, 2014. Regional report (in Italian). http://servizi2.regione.toscana.it/aria/img/getfile_img1.php?id=23513

Perrone, M.R., Becagli, S., Garcia Orza, J.A., Vecchi, R., Dinoi, A., Udisti, U., Cabello, M., 2013. The impact of long-range-transport on PM1 and PM2.5 at a Central Mediterranean site. Atmospheric Environment, 71, 176-186.

Pio, C., Cerqueira, M., Harrison, R.M., Nunes, T., Mirante, F., Alves, C., Oliveira, C., de la Campa, A.S., Artiñano, B., Matos, M., 2011. OC/EC ratio observations in Europe: Re-thinking the approach for apportionment between primary and secondary organic carbon. Atmospheric Environment, 45, 6121-6132.

Polissar, A., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: elemental composition and sources. Journal of Geophysical Research, 103, 19045-19057.

Pope, C.A., and Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. Journal of the Air and Waste Management Association 56, 709-742.

Reche, C., Viana, M., Amato, F., Alastuey, A., Moreno, T., Hillamo, R., Teinila, K., Saarnio, K., Seco, R., Penuelas, J., Mohr, C., Prévot, A.S.H., Querol, X., 2012. Biomass burning contributions to urban aerosols in a coastal Mediterranean City. Science of the Total Environment, 427-428, 175–190.

Seinfeld, J.H., and Pandis, S.N., 1998. Atmospheric Chemistry and Physics. New York: John Wiley.

Turpin, B.J., and Lim, H.J, 2001. Species contributions to pm2.5 mass concentrations: Revisiting common assumptions for estimating organic mass. Aerosol Science and Technology, 35, 602-610.

Vecchi, R., Chiari, M., D'Alessandro, A., Fermo, P., Lucarelli, F., Mazzei, F., Nava, S., Piazzalunga, A., Prati, P., Silvani, F., Valli, G., 2008. A mass closure and PMF source

apportionment study on the sub-micron sized aerosol fraction at urban sites in Italy. Atmospheric Environment, 42, 2240-2253.

Vecchi, R., Valli, G:, Fermo, P., D'Alessandro, A:, Piazzalunga, A., Bernardoni, V., 2009. Organic and inorganic sampling artefacts assessment. Atmospheric Environment, 43, 1713-1720.

Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, M., Prevot, A.S.H., Szidat, S., Hueglin, C., Bloemen, H., Wahlin, P., Vecchi, R., Miranda, A.I., Kasper- Giebl, A., Maenhaut, W., Hitzenberger, R., 2008. Source apportionment of particulate matter in Europe: A review of methods and results. Journal of Aerosol Science, 39, 827-849.

Visser, S., Slowik, J. G., Furger, M., Zotter, P., Bukowiecki, N., Dressler, R., Flechsig, U., Appel, K., Green, D. C., Tremper, A. H., Young, D. E., Williams, P. I., Allan, J. D., Herndon, S. C., Williams, L. R., Mohr, C., Xu, L., Ng, N. L., Detournay, A., Barlow, J. F., Halios, C. H., Fleming, Z. L., Baltensperger, U., and Prévôt, A. S. H., 2015. Kerb and urban increment of highly time-resolved trace elements in PM10, PM2.5 and PM1.0 winter aerosol in London during ClearfLo 2012, Atmosheric Chemistry and Physics, 15, 2367-2386.

World Health Organization (WHO), 2013. Review of evidence on health aspects of air pollution – REVIHAAP Project.

Zhou, L., Hopke, P.K., Paatero, P., Ondov, J.M., Pancras, J.P., Pekney, N.J., Davidson, C.I., 2004. Advanced factor analysis for multiple time resolution aerosol composition data. Atmospheric Environment, 38, 4909-4920.

FIGURE AND TABLE CAPTIONS

Figure 1: Example of the input data for multi-time ME where *unc* indicates the uncertainty column.

Figure 2: Base-case solution: source profiles (bars) and EV values (dots).

Figure 3: Comparison between source profiles provided by the base-case solution and the continuation run solution.

Figure 4: Bootstrap results for the constrained source profiles. Interquartile intervals are represented by blue bars, the median values by black lines and the whiskers extend to the data extremes. The constrained solution values are indicated by red circles.

Table 1: Absolute and relative average source contributions to PM2.5 mass during the whole sampling period, summer and winter.

Table 1

Table 1) Absolute and relative average source contributions to PM2.5 mass during the whole sampling period, summer and winter.

Factor	Summer [µg/m^3]	Winter [µg/m^3]	Annual average [µg/m^3]
Traffic exhaust emissions	8.4 (45%)	7.9 (31%)	8.0 (37%)
Biomass burning	2.5 (13%)	7.8 (25%)	5.6 (20%)
Nitrate	0.7 (4%)	7.0 (21%)	4.4 (15%)
Sulphate	4.0 (21%)	3.5 (11%)	3.7 (14%)
Traffic non-exhaust emissions	1.3 (7%)	1.5 (6%)	1.4 (7%)
Soil dust	1.1 (6%)	0.7 (3%)	0.9 (4%)
Aged sea salt	0.6 (4%)	0.6 (3%)	0.6 (4%)

Figure 1
Click here to download Figure: fig_1.pdf

Start time dd/mm/yy hh:mm	End time dd/mm/yy hh:mm	Start of sampling Unit (1h)	Sampling duration Unit (1h)	End of sampling Unit (1h)	Mass $\mu g/m^3$	Mass unc $\mu g/m^3$	EC $\mu g/m^3$	EC unc $\mu g/m^3$	Al $\mu g/m^3$	Al unc $\mu g/m^3$
24/09/09 0:00	24/09/09 1:00	1729	1	1729					0.12	0.04
24/09/09 1:00	24/09/09 2:00	1730	1	1730					0.08	0.04
24/09/09 2:00	24/09/09 3:00	1731	1	1731					0.08	0.04
24/09/09 3:00	24/09/09 4:00	1732	1	1732					0.08	0.03
...	...	...	...	...	...	...	...	...	...	...
11/04/09 0:00	12/04/09 0:00	1	24	24	24.45	97.83	6.48	1.02	0.09	0.01
13/04/09 0:00	14/04/09 0:00	25	24	48	27.38	109.51	6.12	0.98	0.07	0.01
15/04/09 0:00	16/04/09 0:00	49	24	72	28.20	112.78	6.55	1.03	0.10	0.02
...	...	...	...	...	...	...	...	...	...	...

Figure 2

Click here to download Figure: fig2_rev.pdf

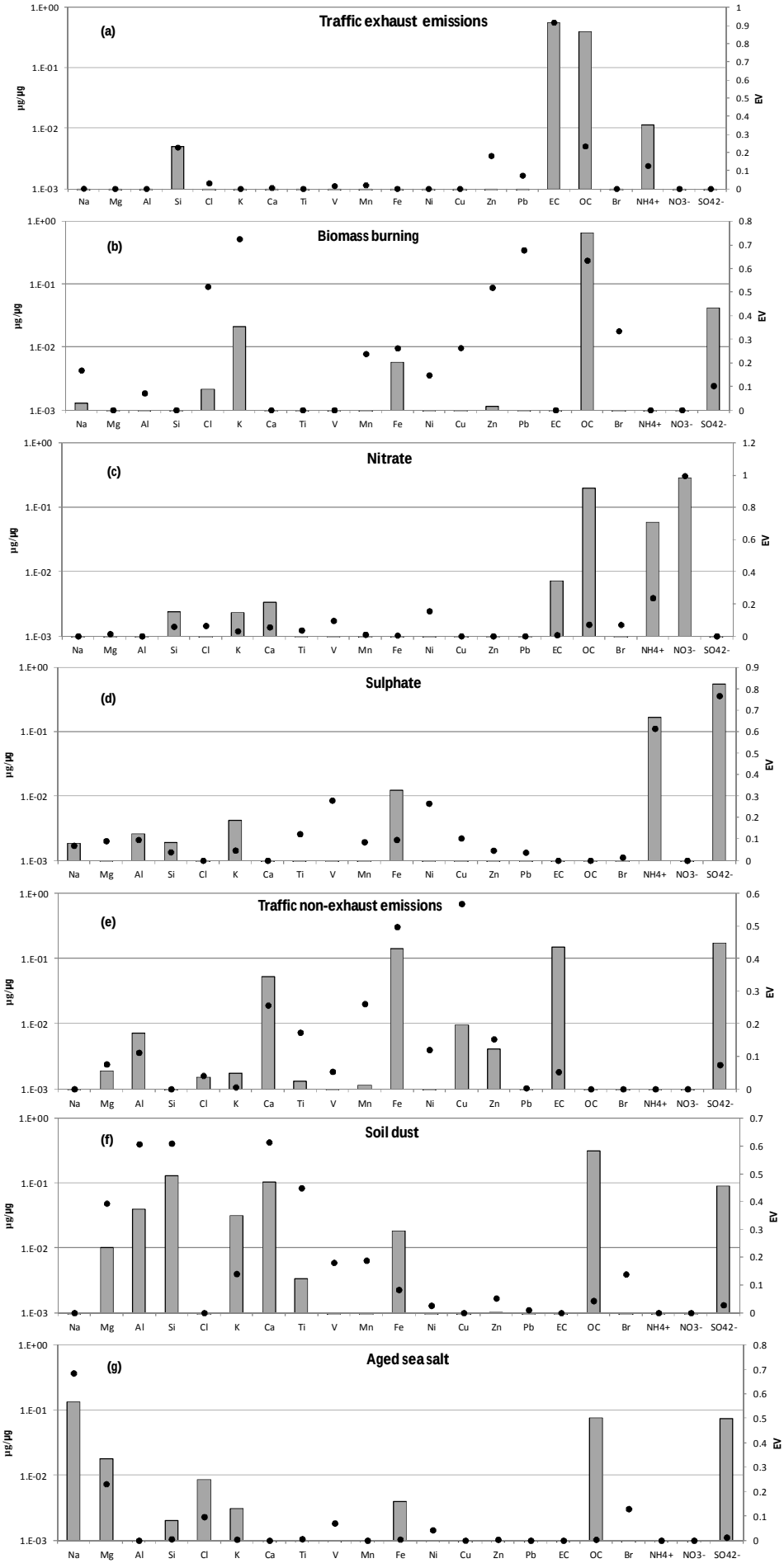


Figure 3

Click here to download Figure: fig3.pdf

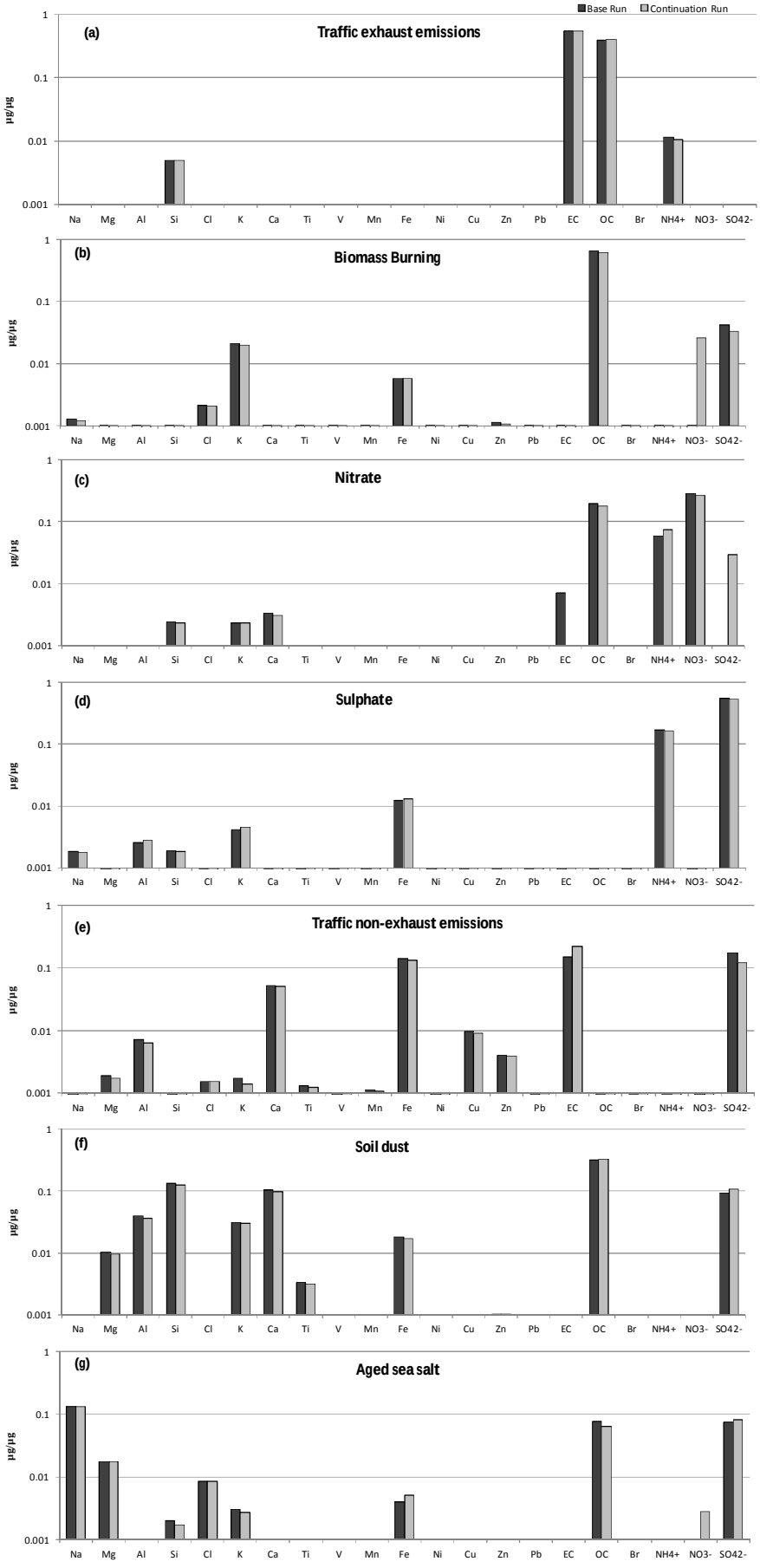
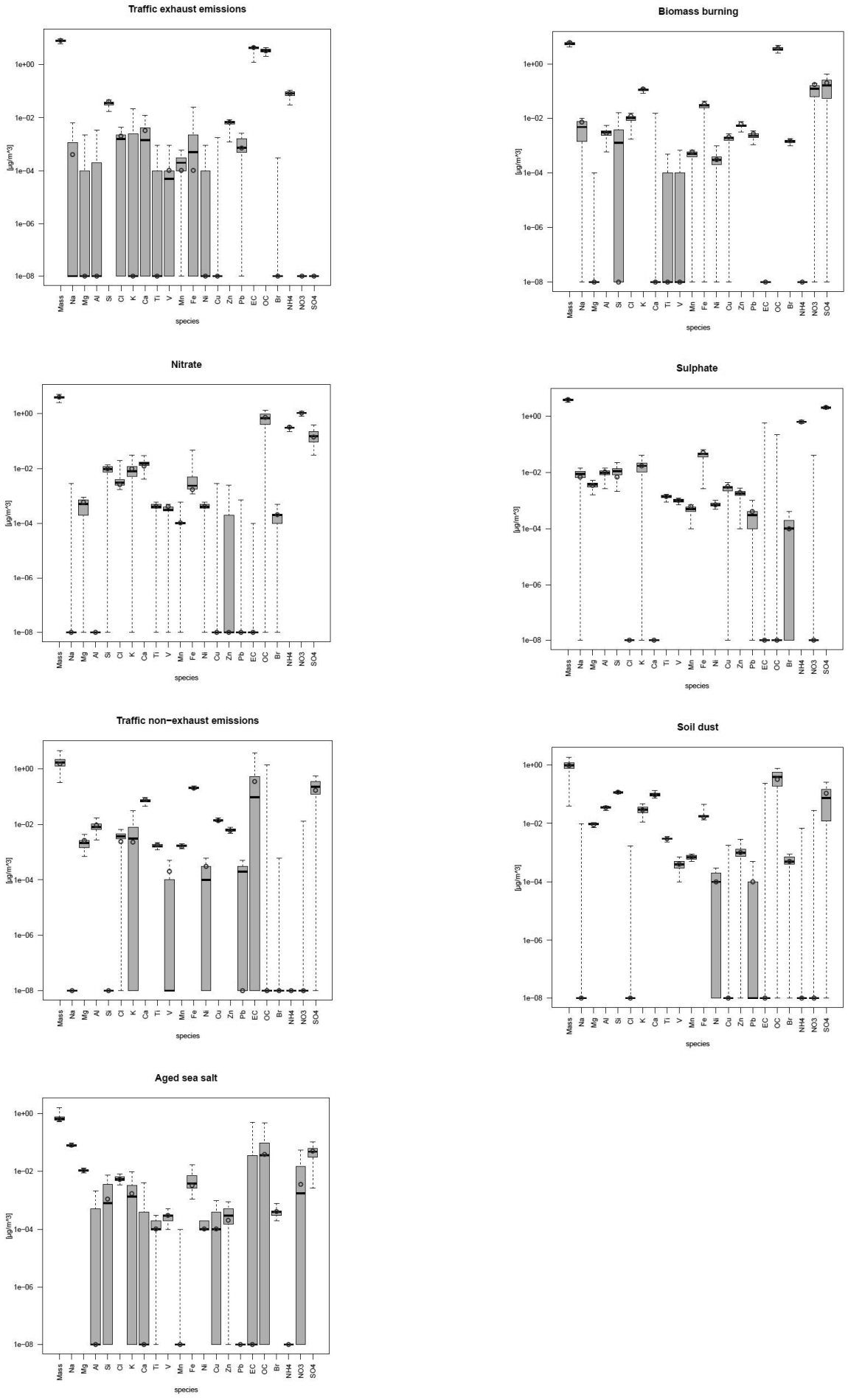


Figure 4

Click here to download Figure: fig4_revBW.pdf



Supplementary material for on-line publication only

Click here to download Supplementary material for on-line publication only: Supplementary Material_rev.pdf