Supplement of Atmos. Chem. Phys., 14, 8813–8839, 2014
http://www.atmos-chem-phys.net/14/8813/2014/
doi:10.5194/acp-14-8813-2014-supplement
© Author(s) 2014. CC Attribution 3.0 License.

Atmospheric
Chemistry
and Physics
Open Access

*Supplement of*

# Sources and geographical origins of fine aerosols in Paris (France)

**M. Bressi et al.**

*Correspondence to:* M. Bressi (michael.bressi@ensiacet.fr)

Supplement of Atmos. Chem. Phys., 14, 8813–8839, 2014
http://www.atmos-chem-phys.net/14/8813/2014/
doi:10.5194/acp-14-8813-2014-supplement
© Author(s) 2014. CC Attribution 3.0 License.

Atmospheric Chemistry and Physics
Open Access

*Supplement of*

# Sources and geographical origins of fine aerosols in Paris (France)

**M. Bressi et al.**

*Correspondence to:* M. Bressi (michael.bressi@ensiacet.fr)

# Supplementary Material

## S1. Complement to Sect. 2.2 "Identification and contribution of the major sources of $PM_{2.5}$"

To identify the major sources of $PM_{2.5}$ and estimate their contribution to fine aerosol masses, source apportionment (SA) models have been extensively developed in the last three decades (Cooper and Watson, 1980; Gordon, 1980; Hopke, 1981, 1985; Watson et al., 2002). Three main groups of SA methods can be distinguished according to Viana et al. (2008): (1) methods based on the evaluation of monitoring data using basic numerical data treatment (e.g. Lenschow et al., 2001), (2) methods based on emission inventories and/or dispersion models to simulate aerosol emission, formation, transport and deposition (e.g. Visser et al., 2001) and (3) methods based on the statistical evaluation of PM chemical data acquired at receptor sites (so-called receptor models). The latter class of techniques was here chosen because of their advanced mathematical approach, their robustness (e.g. Hopke et al., 2006) and their widespread use in the literature (Belis et al., 2013; Zhang et al., 2011), hence allowing comparisons between methods and results achieved (e.g. Poirot et al., 2001; Zheng et al., 2002).

A receptor model was used to identify the major sources of $PM_{2.5}$ and estimate their contribution to fine aerosol masses. Receptor models assume mass conservation and use a mass balance analysis to identify and apportion sources of airborne PM (Hopke et al., 2006). Equation S1 resumes this principle for a dataset consisting of n samples made of m chemical species emitted by p independent sources:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} * f_{kj} \qquad \text{(S1)}$$

where $x_{ij}$ is the measured concentration of the $j^{th}$ species in the $i^{th}$ sample, $g_{ik}$ is the contribution of the $k^{th}$ source to the $i^{th}$ sample and $f_{kj}$ is the concentration of the $j^{th}$ chemical species in the material emitted by the $k^{th}$ source. Different models can be used to solve Eq. (S1), including Principal Component Analysis (Blifford and Meeker, 1967), Chemical Mass Balance (Friedlander, 1973; Miller et al., 1972; Winchester and Nifong, 1971), UNMIX (Henry and Kim, 1990; Kim and Henry, 1999, 2000) and Positive Matrix Factorization (Paatero and Tapper, 1994; Paatero, 1997), to name a few.

For the present study, we decided to use the PMF model because: i) it does not require a priori knowledge about source profiles i.e. source chemical composition, ii) it accounts for measurement uncertainties allowing weighting of individual samples and iii) it forces every source contribution and source profile to be non-negative (Reff et al., 2007). The mathematical model in its matrix form is:

$$X = G * F + E \qquad \text{(S2)}$$

where X is the chemical dataset matrix, G is the source contribution matrix, F the source profile matrix and E the so-called residual matrix. In index notation Eq. (S2) can be written as:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} * f_{kj} + e_{ij} \qquad \text{(S3)}$$

where $e_{ij}$ represents the residual element, or the PMF model error, for the species j measured in the sample i (see Eq. (S1) for the explanation of the other parameters). The PMF model aims at resolving Eq. (S3) by minimising a Q function defined as:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} (\frac{e_{ij}}{\sigma_{ij}})^2 \qquad \text{(S4)}$$

where $\sigma_{ij}$ is the uncertainty associated to the j$^{th}$ species in the i$^{th}$ sample. Different Q functions can be defined: $Q_{true}$ calculated including all data and $Q_{robust}$ calculated excluding outliers i.e. data for which the scaled residual ($e_{ij}/\sigma_{ij}$) is greater than 4. (Note that $Q_{theoretical}$ will not be studied here as explained in Sect. S2.) Mathematically speaking, PMF is thus a constrained weighted least square method attempting to find G and F matrices that best reproduce X. Different programs have been developed to solve Eq. (S3) by minimising the Q function, which includes PMF2 (Paatero and Tapper, 1994; Ulbrich et al., 2009), PMF3 (Paatero, 1997) and Multilinear Engine (Paatero, 1999). A second version of the ME program was used here (ME-2; Paatero, 2000; Norris et al., 2009; Canonaco et al., 2013). It was integrated in the standalone PMF version of the United States Environmental Protection Agency (US-EPA) known as EPA PMF3.0 (Norris et al., 2008) and can be downloaded at http://www.epa.gov/heasd/products/pmf/pmf.html. Comparisons between the different solving programs can be found in the literature (e.g. Amato et al., 2009; Dutton et al., 2010; Ramadan et al., 2003) and attest the validity and the robustness of the ME-2 version.

## S2. Complement to Sect. 2.2.2 “Data preparation”

The construction of the concentration dataset (X matrix) demands a precise and critical analysis of the chemical dataset. We initially used the whole chemical database described in Bressi et al. (2013) and Poulakis et al. (2012), in addition with the species mentioned in the Sect.1.1.2. This comprises $PM_{2.5}$ mass and 29 chemical components: $PM_{grav}$, EC, OM, $Cl^-$, $NO_3^-$, $SO_4^{2-}$, $Na^+$, $NH_4^+$, $K^+$, $Mg^{2+}$, $Ca^{2+}$, Al, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Cd, Pb, levoglucosan, mannosan, arabitol and mannitol. Unfortunately, for specific compounds, atmospheric concentrations were mostly below the Method Quantification Limit (MQL) and therefore difficult to interpret. After testing different threshold values, we decided to exclude every compound exhibiting less than 40% of their values above the MQL, which concerns Al, Cr, As, arabitol and mannitol. In addition, to avoid redundant species, Ca was excluded from the dataset but $Ca^{2+}$ was taken into account, and levoglucosan and mannosan concentrations were subtracted from organic matter levels.

We also had to deal with missing species measurements due to analytical problems for specific samples. EPA PMF3.0 does not allow missing data to be implemented in the X matrix. These data are generally replaced by “virtual” values having larger uncertainties in order to lower their influences in the PMF modelling (Huang et al., 1999; Polissar et al., 1998; Reff et al., 2007). Every missing data was here replaced by the median of the corresponding species, and associated an uncertainty of four times the species-specific median as suggested in Norris et al. (2008). The detailed list of every processed missing data can be found in Table S2.

The second dataset, dealing with uncertainties associated with every species measurement, was constructed following the procedure described by Norris et al. (2008) and adapted from Polissar et al. (1998). This is an equation-based method that requires each species Method Detection Limit (MDL) -in the same unit as the species concentration (here in µg/m$^3$)- and its analytical uncertainty (u) -in percentage-. The PMF uncertainty (σ) is calculated as:

$$if\ x_{ij} \leq MDL_i\ , \forall j, then\ \sigma_{ij} = \frac{5}{6} * MDL_i \tag{S5a}$$

$$if\ x_{ij} > MDL_i\ , \forall j, then\ \sigma_{ij} = \sqrt{\left(u_i * x_{ij}\right)^2 + (MDL_i)^2} \tag{S5b}$$

MDLs and analytical uncertainties of our dataset were initially taken from Bressi et al. (2013), Poulakis et al. (2012), and Ghersi et al. (2010), but lead to meaningless and non-robust results. An adjustment was empirically made for every chemical species to have i) normally distributed scaled residuals, ii) relevant MDL and u regarding the analytical method used and iii) mathematically satisfactory and robust PMF results (see Sect. 2). Several empirical tests showed that when the theoretical MDL of a given compound is not appropriate, taking its 5th percentile as its MDL often leads to suitable results. In fact, this method lowers the statistical weight of the smallest concentrations of a given chemical species. However, it will prevent us from using the $Q_{theoretical}$ function to assess the quality of our uncertainty matrix, since the latter was optimized as described above. MDLs and analytical uncertainties chosen are reported in Table S3.

With the EPA PMF3.0 software, additional uncertainties can be optionally attributed to single species and/or to the overall dataset. Single species additional uncertainties were suggested by Paatero and Hopke (2003) with a calculation based on the signal-to-noise ratio (S/N), which is defined in EPA PMF3.0 as:

$$(\frac{S}{N})_j = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij}-\sigma_{ij})^2}{\sum_{i=1}^{n}\sigma_{ij}{}^2}} \tag{S2}$$

According to S/N values, species are categorised as "bad" (S/N<0.2), "weak" (0.2≤S/N≤2) or "strong" (S/N>2), although the authors mention the relatively arbitrary nature of these figures. The bad categorisation excludes the species from the dataset, the weak triples the PMF uncertainty whereas the strong does not add supplemental uncertainties. According to these criteria, Cd, Zn and Cu were defined weak whereas all other species were categorised strong. PM mass was defined as a "Total Variable" i.e. was regarded as being the sum of all PM chemical components; all available mass was thus apportioned across the factors although part of this mass might still be unexplained. PM mass variable was categorised weak to lower its influence in the final PMF results. An optional "Extra Modeling Uncertainty" can be added in the EPA PMF3.0 version but was not applied here.

## S3. Complement to Sect. 2.2.3 "Robustness of PMF results"

Bootstrap is a statistical inference method introduced by Efron (1979) to "estimate the sampling distribution of some prespecified random variable on the basis of the observed dataset". It gives information on the accuracy of an estimate or a statistic. The key idea is to "resample from the original data to create replicate datasets from which the variability of the quantities of interest will be assessed" (Davison and Hinkley, 1997). Resampling can either be performed from an empirical data distribution (nonparametric bootstrap) or via a fitted model describing a specific distribution (parametric bootstrap). This led to different bootstrap versions reported in Wehrens et al. (2000). One major assumption required for having reliable bootstrap results, is that data are independent and identically distributed (Singh, 1981), which is questionable for real atmospheric samples. A way to solve this dependence issue is to use the so-called "blocked bootstrap" method, which is randomly selecting blocks of successive data, instead of selecting individual elements (see Lahiri, 2003 for more details). The relevant block length to be chosen is discussed in Politis and White (2004). More information concerning the bootstrap theory can be found in Efron and Tibshirani (1993).

We used the bootstrap method suggested in the EPA PMF3.0 software which is a nonparametric "blocked bootstrap" version. Non-overlapping blocks of consecutive samples are randomly selected with replacement to create a new "bootstrap matrix" that has the same dimensions as the original X matrix (bootstrap matrices will be noted with an "*" later on). PMF is then run on the X* matrix and bootstrap factors are assigned to base run factors by comparing their contribution to PM mass (i.e. by comparing G* and G columns, respectively). Bootstrap factors are "mapped" with base run factors if their G* versus G column correlation is higher than a user defined threshold (see Sect. 1.2.4), and "unmapped" otherwise. Further details concerning the bootstrap method used here can be found in Norris et al. (2008).

Table S1. Days discarded from the PMF chemical dataset.

Note: days were discarded due to power failure during sampling (61%), problems with Leckel sampler (14%), chemical analysis (14%), sampler calibration (4%), local contamination (4%) and other reasons (4%).

| **Month/Year** | **Day of month** |
|---|---|
| 09/09 | 15, 16 |
| 10/09 | 7 |
| 11/09 | 7 |
| 12/09 | 16, 17, 18 |
| 01/10 | 7, 11, 12 |
| 02/10 | 19, 25, 28 |
| 03/10 | 1, 27, 28, 29 |
| 04/10 | 3, 4, 5, 6, 11, 12 |
| 05/10 | - |
| 06/10 | - |
| 07/10 | 16, 17, 18, 19 |
| 08/10 | 7 |
| 09/10 | - |
| **Total (days)** | **28** |

Table S2. List of the missing data replaced by the median of the corresponding species concentrations.

*Legend: numbers (unless bolded) refer to days of month.*

| | Month/Year | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **09/09** | **10/09** | **11/09** | **12/09** | **01/10** | **02/10** | **03/10** | **04/10** | **05/10** | **06/10** | **07/10** | **08/10** | **09/10** | **Total (days)** |
| **PM** | 15 | - | - | 17 | 28 | - | 28 | 4, 5 | - | 9 | 16-19 | 7 | - | **12** |
| **OM** | 15, 16 | 7, 13, 14 | 7, 15 | 16-18 | 7, 11, 12 | 19, 25, 28 | 13, 27, 28 | 3-5, 11 | 17, 19, 20 | 9, 17, 18 | 16-19 | 7 | - | **34** |
| **EC** | - | 7 | - | 16-18 | 7 | 19, 25 | 14, 28 | 4, 5, 29 | - | 5, 9 | 17, 18 | - | - | **16** |
| **NO3** | 28 | 28 | - | - | - | - | 2, 28 | 4, 5 | - | - | 17, 18 | - | - | **8** |
| **SO4** | 17, 18 | 28 | - | - | - | - | 2, 28 | 4, 5 | - | - | 17, 18 | - | - | **9** |
| **NH4** | 28 | - | - | - | - | - | 28 | 4, 5 | - | - | 17, 18 | - | - | **6** |
| **Na** | - | - | - | - | - | - | 28 | 4, 5 | - | - | 17, 18 | - | - | **5** |
| **Cl** | - | 28 | - | 12 | 13 | - | 2, 28 | 4, 5 | - | - | 17, 18 | - | - | **9** |
| **Mg** | - | - | 11 | - | 1 | - | 28 | 4, 5 | - | - | 17, 18 | - | - | **7** |
| **K** | - | - | 8-10 | - | 1 | - | 28 | 4, 5 | - | - | 17, 18 | - | - | **9** |
| **Lev** | - | 13, 14 | 15 | 17, 18 | - | - | 13, 27, 28 | 4, 5 | 17 | - | 17, 18 | - | - | **13** |
| **Man** | - | 13, 14 | 15 | 17, 18 | - | - | 13, 27, 28 | 4, 5 | 17 | - | 17, 18 | - | - | **13** |
| **V** | - | 19, 20 | - | 20, 30 | 16 | - | 9, 27, 28 | 4, 5, 19 | 21 | 15 | 17, 18, 21 | - | - | **16** |
| **Ni** | 16, 25 | 3, 20 | - | 30 | 16 | 16 | 9, 27, 28 | 4, 5, 18, 28-30 | 1-20, 22, 27-31 | 1, 3, 4, 8-13, 15-16 | 2, 17, 18, 21, 25 | 23 | - | **59** |
| **Fe** | - | - | 20 | 30 | - | 4, 14 | 9, 27, 28 | 4, 5 | - | 15 | 17, 18, 21 | - | - | **13** |
| **Mn** | - | - | - | 27, 30 | - | - | 9, 28-31 | 4, 5 | - | 14, 15 | 14, 17, 18, 21, 24 | 17 | - | **17** |
| **Cu** | - | - | 20 | 17, 25, 30 | 1 | 14 | 9, 27, 28 | 4, 5, 27 | - | 15 | 17, 18, 21 | - | - | **16** |
| **Cd** | 23, 29 | 29 | 25 | 17, 30 | 25-27 | - | 9, 28-31 | 4, 5 | 13 | 15, 23-30 | 1-5, 17, 18, 21 | - | - | **34** |
| **Pb** | - | - | - | 17, 25, 30 | 26 | - | 9, 28-31 | 4, 5, 15 | - | 14, 15 | 17, 18, 21 | - | - | **17** |

Table S3. Method Detection Limits (MDL) and Analytical Uncertainties (u) chosen for PMF runs.

| | **MDL** | **u** |
|---|---|---|
| | µg.m$^{-3}$ | % |
| **PM** | 2.0E-01 | 5 |
| **OM** | 1.0E-01 | 15 |
| **EC** | 3.0E-01 | 15 |
| **NO3** | 1.3E-01 | 5 |
| **SO4** | 4.3E-01 | 5 |
| **NH4** | 3.0E-01 | 5 |
| **Na** | 2.6E-01 | 10 |
| **Cl** | 6.9E-02 | 10 |
| **Mg** | 5.5E-03 | 15 |
| **K** | 2.0E-02 | 10 |
| **Lev** | 5.2E-03 | 10 |
| **Man** | 5.1E-04 | 10 |
| **V** | 2.0E-04 | 10 |
| **Ni** | 2.9E-04 | 20 |
| **Fe** | 3.6E-02 | 20 |
| **Mn** | 5.6E-04 | 15 |
| **Cu** | 1.0E-03 | 20 |
| **Cd** | 2.1E-05 | 15 |
| **Pb** | 1.1E-03 | 15 |

Table S4. Using bootstrapping to determine the adequate number of factors in PMF.

*Legend: r-value: minimum coefficient of determination used to assign a boot factor to a base factor, rows: bootstrap factors, columns: base factors, Un: unmapped. The less robust bootstrap factor is boldfaced and coloured in red.*

Note: Bootstrapping was not performed on the same base run for the 0.6 and 0.7 r-value configurations. In fact, for each configuration, 20 base runs were conducted prior to bootstrapping and the base run showing the lowest Q-value was retained. Therefore, according to the r-value configuration, the different factor numbers will not correspond to the same physical source.

r-value=0.6

| | **1** | **2** | **3** | **4** | **5** | **6** | **Un** |
|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| **3** | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| **4** | 0 | 0 | 1 | 99 | 0 | 0 | 0 |
| **5** | 0 | 0 | 1 | 0 | **94** | 0 | 5 |
| **6** | 0 | 0 | 0 | 0 | 0 | 100 | 0 |

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Un** |
|---|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | **96** | 1 | 0 | 0 | 0 | 0 | 3 |
| **3** | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| **4** | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| **5** | 0 | 0 | 0 | 0 | 98 | 0 | 0 | 2 |
| **6** | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **Un** |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 99 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **3** | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| **4** | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| **5** | 11 | 1 | 1 | 0 | **78** | 0 | 0 | 3 | 6 |
| **6** | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| **8** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 5 |

r-value=0.7

| | **1** | **2** | **3** | **4** | **5** | **6** | **Un** |
|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 99 | 0 | 0 | 0 | 0 | 1 |
| **3** | 0 | 0 | **83** | 1 | 0 | 0 | 16 |
| **4** | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| **5** | 0 | 0 | 0 | 0 | 99 | 0 | 1 |
| **6** | 0 | 0 | 0 | 0 | 0 | 100 | 0 |

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Un** |
|---|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 96 | 0 | 0 | 0 | 0 | 0 | 4 |
| **3** | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| **4** | 1 | 0 | 0 | **95** | 0 | 0 | 1 | 3 |
| **5** | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| **6** | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **Un** |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 93 | 0 | 0 | 1 | 0 | 0 | 0 | 6 |
| **3** | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| **4** | 0 | 0 | 0 | 99 | 0 | 0 | 0 | 0 | 1 |
| **5** | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| **6** | 3 | 3 | 0 | 0 | 0 | **75** | 0 | 2 | 17 |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| **8** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |

Table S5. Sum of the squared difference between the scaled residuals (d-values) calculated for each pair of base runs.
Note: Cf. Norris et al. (2008) and Sect. 2.1 for more information on this variable.

| **RUN#** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **2** | 0.028 | | | | | | | | | | | | | | | | | | |
| **3** | 0.019 | 0.003 | | | | | | | | | | | | | | | | | |
| **4** | 0.020 | 0.005 | 0.005 | | | | | | | | | | | | | | | | |
| **5** | 0.019 | 0.005 | 0.001 | 0.006 | | | | | | | | | | | | | | | |
| **6** | 0.007 | 0.020 | 0.011 | 0.012 | 0.010 | | | | | | | | | | | | | | |
| **7** | 0.025 | 0.001 | 0.003 | 0.004 | 0.005 | 0.017 | | | | | | | | | | | | | |
| **8** | 0.013 | 0.019 | 0.012 | 0.010 | 0.009 | 0.003 | 0.018 | | | | | | | | | | | | |
| **9** | 0.032 | 0.007 | 0.009 | 0.004 | 0.010 | 0.020 | 0.007 | 0.017 | | | | | | | | | | | |
| **10** | 0.025 | 0.012 | 0.010 | 0.007 | 0.008 | 0.011 | 0.012 | 0.006 | 0.009 | | | | | | | | | | |
| **11** | 0.005 | 0.028 | 0.016 | 0.022 | 0.014 | 0.005 | 0.026 | 0.010 | 0.032 | 0.023 | | | | | | | | | |
| **12** | 0.019 | 0.008 | 0.005 | 0.004 | 0.004 | 0.008 | 0.007 | 0.005 | 0.008 | 0.002 | 0.017 | | | | | | | | |
| **13** | 0.007 | 0.022 | 0.013 | 0.014 | 0.011 | 0.001 | 0.020 | 0.003 | 0.022 | 0.012 | 0.006 | 0.009 | | | | | | | |
| **14** | 0.041 | 0.088 | 0.071 | 0.087 | 0.069 | 0.063 | 0.085 | 0.075 | 0.106 | 0.099 | 0.037 | 0.084 | 0.062 | | | | | | |
| **15** | 0.008 | 0.022 | 0.013 | 0.013 | 0.011 | 0.001 | 0.019 | 0.003 | 0.021 | 0.011 | 0.006 | 0.008 | 0.001 | 0.064 | | | | | |
| **16** | 0.017 | 0.005 | 0.002 | 0.004 | 0.001 | 0.007 | 0.005 | 0.007 | 0.009 | 0.006 | 0.013 | 0.002 | 0.009 | 0.073 | 0.008 | | | | |
| **17** | 0.048 | 0.106 | 0.094 | 0.101 | 0.096 | 0.082 | 0.102 | 0.098 | 0.122 | 0.122 | 0.057 | 0.107 | 0.081 | 0.016 | 0.084 | 0.097 | | | |
| **18** | 0.018 | 0.016 | 0.011 | 0.009 | 0.009 | 0.006 | 0.015 | 0.003 | 0.012 | 0.002 | 0.016 | 0.003 | 0.007 | 0.087 | 0.006 | 0.006 | 0.109 | | |
| **19** | 0.035 | 0.005 | 0.010 | 0.004 | 0.012 | 0.025 | 0.006 | 0.023 | 0.002 | 0.013 | 0.038 | 0.011 | 0.028 | 0.110 | 0.027 | 0.011 | 0.121 | 0.017 | |
| **20** | 0.031 | 0.006 | 0.009 | 0.003 | 0.010 | 0.019 | 0.006 | 0.015 | 0.003 | 0.006 | 0.033 | 0.006 | 0.021 | 0.108 | 0.020 | 0.008 | 0.122 | 0.010 | 0.003 |

Table S6. Correlation (Pearson coefficient, R) between PMF factor time series and their presumable tracers.

*Legend: BCff: fossil fuel black carbon, Lev: levoglucosan, Mann: mannosan, BCwb: wood burning black carbon. Pearson coefficients higher than 0.5 are indicated in bold.*

Note: More information on this table can be found in Bressi et al. (2014).

| | $SO_4$ | $NO_3$ | V | NOx | BCff | Lev | Mann | BCwb | Na | Cl | Mg | Pb | Cd |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A.S. rich factor | **0.85** | **0.58** | 0.07 | 0.25 | 0.25 | 0.21 | 0.34 | 0.20 | -0.30 | -0.17 | -0.32 | **0.51** | 0.40 |
| A.N. rich factor | **0.73** | **0.99** | 0.24 | 0.27 | 0.19 | 0.29 | 0.39 | 0.31 | -0.24 | -0.03 | -0.26 | 0.43 | **0.52** |
| Heavy oil combustion | 0.35 | 0.11 | **0.75** | 0.09 | 0.22 | -0.06 | -0.05 | -0.01 | -0.25 | -0.26 | -0.16 | 0.29 | 0.27 |
| Road traffic | -0.09 | -0.15 | 0.08 | **0.50** | **0.50** | -0.01 | 0.01 | 0.14 | -0.25 | -0.23 | -0.19 | 0.23 | 0.20 |
| Biomass burning | 0.21 | 0.44 | 0.12 | **0.61** | 0.22 | **0.99** | **0.98** | **0.89** | -0.12 | 0.18 | -0.17 | 0.42 | 0.31 |
| Marine aerosols | -0.45 | -0.25 | -0.25 | -0.20 | -0.22 | -0.11 | -0.15 | -0.14 | **0.91** | **0.82** | **0.88** | -0.35 | -0.37 |
| Metal industry | 0.33 | 0.33 | 0.42 | 0.36 | 0.32 | 0.21 | 0.28 | 0.27 | -0.19 | -0.07 | -0.20 | **0.85** | **0.92** |

Table S7. Description of the different receptor model studies compared in Fig. 8 and discussed in Sect. 3.3.1.
Note: Sites are indicated as: “City (Country)-Type of sites”. Urb: urban, Rur: rural, Kerb: kerbside.

| | Authors | Site | Receptor model | Year | PM fraction | PM mass (µg/m$^3$) | Source contribution (µg/m$^3$) | Source contribution (%) |
|---|---|---|---|---|---|---|---|---|
| Ammonium nitrate + ammonium sulfate rich sources | Vallius et al., 2005 | Amsterdam (Neth) - Urb | PCA | 1998-1999 | $PM_{2.5}$ | 20.0 | 6.8 | 34 |
| | Andersen et al., 2007 | Copenhagen (Den) - Urb | COPREM | 1999-2004 | $PM_{10}$ | 23.3 | 6.9 | 29 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 7.4 | 51 |
| | Mooibroek et al., 2011 | Schiedam (Neth) - Urb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 13 | 8.6 | 66 |
| | Mooibroek et al., 2011 | Hellendoorn (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 12.5 | 9.1 | 73 |
| | Mooibroek et al., 2011 | Rotterdam (Neth) - Kerb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 16.4 | 10.0 | 61 |
| | Mooibroek et al., 2011 | Vredepeel (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 14.5 | 10.7 | 74 |
| | Mooibroek et al., 2011 | Cabauw (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 17.5 | 12.6 | 72 |
| | Quass et al., 2004 | Duisburg (Ger) - Urb | PMF | 2003-2004 | $PM_{2.5}$ | 22.8 | 13 | 57 |
| Ammonium nitrate rich source | Andersen et al., 2007 | Copenhagen (Den) - Urb | COPREM | 1999-2004 | $PM_{10}$ | 23.3 | 3.3 | 14 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 3.5 | 24 |
| | Mooibroek et al., 2011 | Schiedam (Neth) - Urb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 13.0 | 5.6 | 43 |
| | Mooibroek et al., 2011 | Hellendoorn (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 12.5 | 6.0 | 48 |
| | Mooibroek et al., 2011 | Vredepeel (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 14.5 | 6.4 | 44 |
| | Mooibroek et al., 2011 | Rotterdam (Neth) - Kerb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 16.4 | 6.7 | 41 |
| | Mooibroek et al., 2011 | Cabauw (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 17.5 | 7.7 | 44 |
| Ammonium sulfate rich source | Mooibroek et al., 2011 | Schiedam (Neth) - Urb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 13.0 | 3.0 | 23 |
| | Mooibroek et al., 2011 | Hellendoorn (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 12.5 | 3.1 | 25 |
| | Mooibroek et al., 2011 | Rotterdam (Neth) - Kerb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 16.4 | 3.3 | 20 |
| | Andersen et al., 2007 | Copenhagen (Den) - Urb | COPREM | 1999-2004 | $PM_{10}$ | 23.3 | 3.5 | 15 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 3.9 | 27 |
| | Mooibroek et al., 2011 | Vredepeel (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 14.5 | 4.4 | 30 |
| | Mooibroek et al., 2011 | Cabauw (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 17.5 | 4.9 | 28 |
| Road traffic | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 2.1 | 14 |
| | Lee et al., 2003 | Toronto (Ca) - Urb | PMF | 2000-2001 | $PM_{2.5}$ | 12.8 | 2.3 | 18 |
| | Maykut et al., 2003 | Seattle (USA) - Urb | PMF and UNMIX | 1996-1999 | $PM_{2.5}$ | 9.2 | 2.3 | 25 |
| | Minguillón et al., 2012 | Zurich (Swi) - Urb | EPA PMF3.0 | Summer and Winter 2009 | $PM_{1.0}$ | 10.2 | 3.8 | 37 |
| | Perronne et al., 2012 | Milan (It) - Urb | CMB | 2006-2009 | $PM_{2.5}$ | 35.5 | 7.8 | 22 |
| Biomass combustion | Karanasiou et al., 2009 | Athens (Gr) - Urb | PMF3 | 2002 | $PM_{2.0}$ | 5.3 | 0.8 | 15 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 1.8 | 12 |
| | Perronne et al., 2012 | Milan (It) - Urb | CMB | 2006-2009 | $PM_{2.5}$ | 35.5 | 7.1 | 16 |
| | Andersen et al., 2007 | Copenhagen (Den) - Urb | COPREM | 1999-2004 | $PM_{10}$ | 23.3 | 7.3 | 15 |
| Heavy oil combustion | Andersen et al., 2007 | Copenhagen (Den) - Urb | COPREM | 1999-2004 | $PM_{10}$ | 23.3 | 2.2 | 9 |
| | Alleman et al., 2010 | Dunkirk (Fr) - Urb | PMF2 | 2003-2005 | $PM_{10}$ | 25.0 | 2.3 | 9 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 2.4 | 16 |
| | Vallius et al., 2005 | Amsterdam (Neth) - Urb | PCA | 1998-1999 | $PM_{2.5}$ | 20.0 | 3.5 | 9 |
| Marine aerosols | Mooibroek et al., 2011 | Hellendoorn (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 12.5 | 0.8 | 6 |
| | This study | Paris (Fr) - Urb | EPA PMF3.0 | 2009-2010 | $PM_{2.5}$ | 14.7 | 0.8 | 6 |
| | Mooibroek et al., 2011 | Rotterdam (Neth) - Kerb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 16.4 | 0.8 | 5 |
| | Mooibroek et al., 2011 | Vredepeel (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 14.5 | 0.9 | 6 |
| | Vallius et al., 2005 | Helsinki (Fin) - Urb | PCA | 1998-1999 | $PM_{2.5}$ | 12.2 | 0.9 | 7 |
| | Karanasiou et al., 2009 | Athens (Gr) - Urb | PMF3 | 2002 | $PM_{2.0}$ | 5.3 | 1.1 | 21 |
| | Mooibroek et al., 2011 | Schiedam (Neth) - Urb | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 13.0 | 1.2 | 9 |
| | Mooibroek et al., 2011 | Cabauw (Neth) - Rur | EPA PMF3.0 | 2007-2008 | $PM_{2.5}$ | 17.5 | 1.6 | 9 |

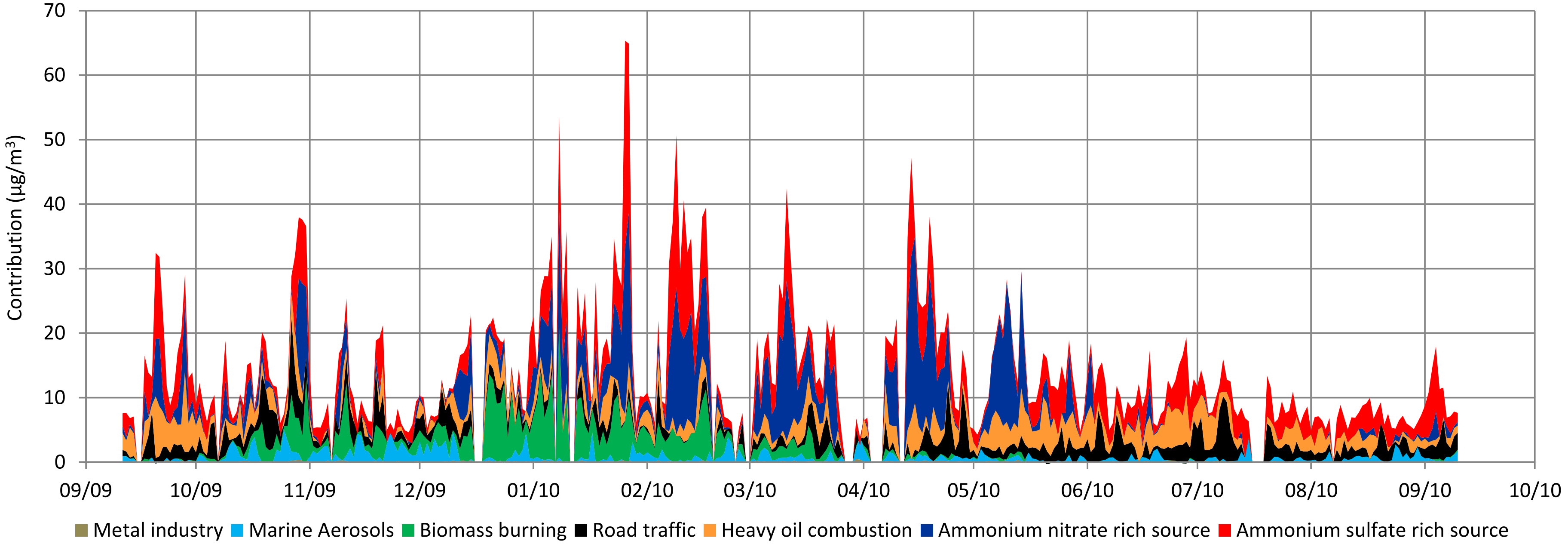

Fig. S1. Daily contribution (µg/m$^3$) of the seven sources to PM$_{2.5}$ mass from 11 September 2009 to 10 September 2010.
Note: results were taken from the base run exhibiting the lowest Q$_{robust}$.

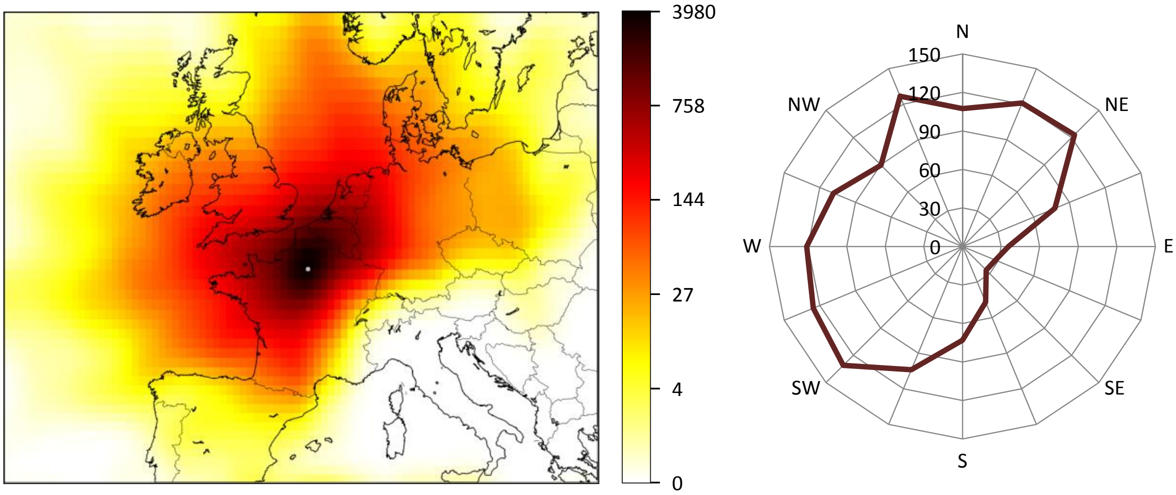

Fig. S2. Geographical distribution of 48-hour air mass back-trajectories from 11 September 2009 to 10 September 2010. Left: number of back-trajectories per cell used for PSCF (logarithmic scale); right: number of back-trajectories per wind direction used for CPF (linear scale).
*Note: The city of Paris is indicated by a grey dot on the left figure.*

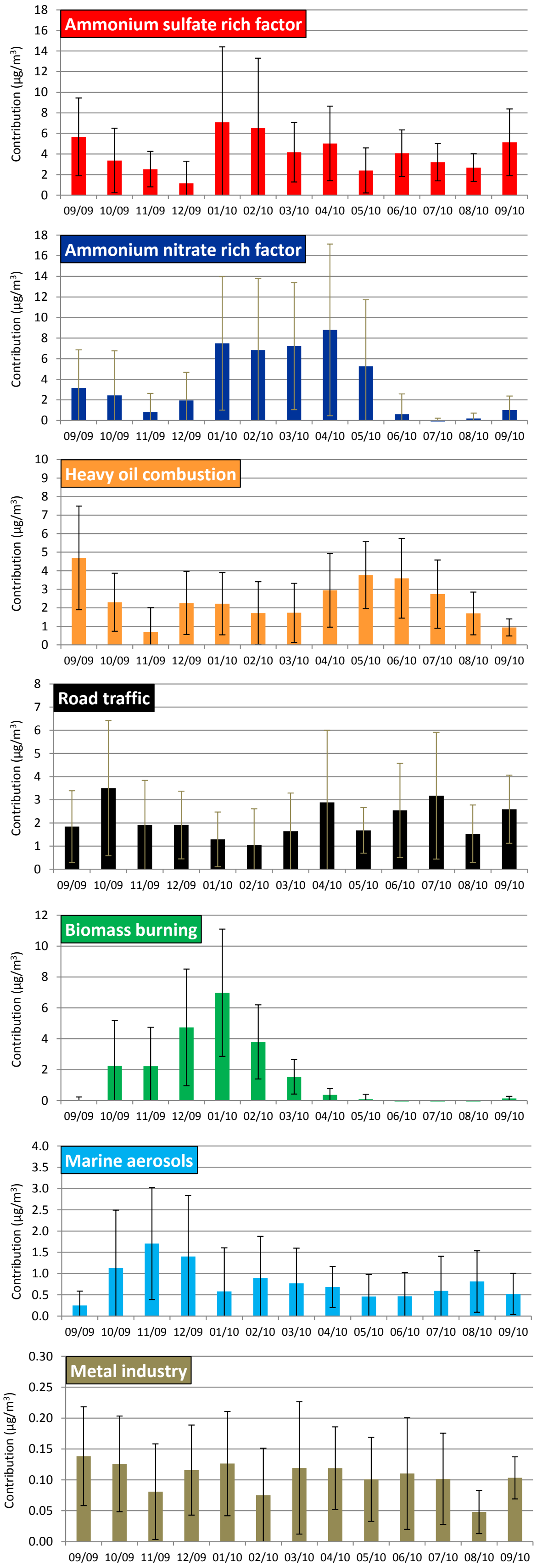

Fig. S3. Monthly mean contribution (µg/m$^3$) of PM$_{2.5}$ sources from 11 September 2009 to 10 September 2010. Constructed from the base run exhibiting the lowest Q$_{robust}$. Error bars represent ±1σ.

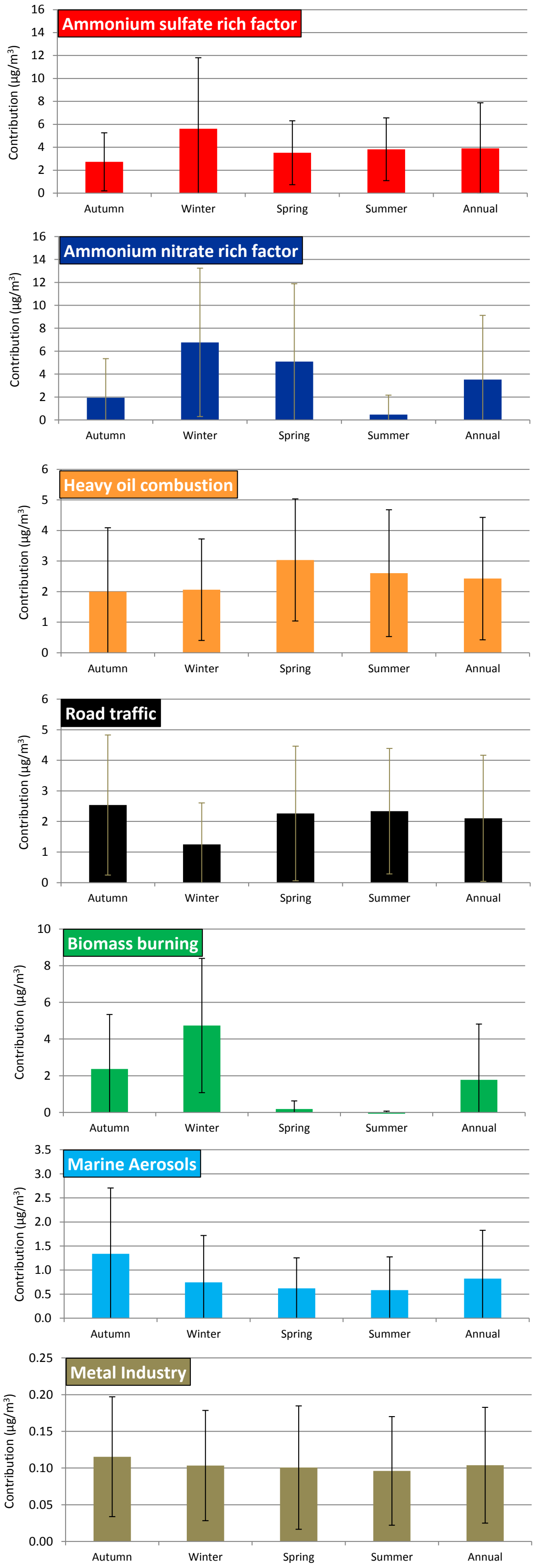

Fig. S4. Seasonal mean contribution (µg/m$^3$) of PM$_{2.5}$ sources from 11 September 2009 to 10 September 2010. Constructed from the base run exhibiting the lowest Q$_{robust}$. Error bars represent ±1σ.

## References

Alleman, L. Y., Lamaison, L., Perdrix, E., Robache, A. and Galloo, J.-C.: PM10 metal concentrations and source identification using positive matrix factorization and wind sectoring in a French industrial zone, Atmospheric Research, 96(4), 612–625, doi:10.1016/j.atmosres.2010.02.008, 2010.

Amato, F., Pandolfi, M., Escrig, A., Querol, X., Alastuey, A., Pey, J., Perez, N. and Hopke, P. K.: Quantifying road dust resuspension in urban environment by Multilinear Engine: A comparison with PMF2, Atmospheric Environment, 43(17), 2770–2780, doi:10.1016/j.atmosenv.2009. 02.039, 2009.

Andersen, Z. J., Wahlin, P., Raaschou-Nielsen, O., Scheike, T. and Loft, S.: Ambient particle source apportionment and daily hospital admissions among children and elderly in Copenhagen, Journal of Exposure Science and Environmental Epidemiology, 17(7), 625–636, doi:10.1038/sj.jes.7500546, 2007.

Belis, C.A., Karagulian, F., Larsen, B.R., Hopke, P.K.: Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe, Atmos. Environ., 69, 94–108, 2013.

Blifford, I. H. and Meeker, G. O.: A factor analysis model of large scale pollution, Atmospheric Environment (1967), 1(2), 147–157, doi:10.1016/0004-6981(67)90042-X, 1967.

Bressi, M., Sciare, J., Ghersi, V., Mihalopoulos, N., Petit, J.-E., Nicolas, J.B., Moukhtar, S., Rosso, A., Féron, A., Bonnaire, N., Poulakis, E. and Theodosi, C.: Response to Referee 3 related to: "Sources and geographical origins of fine aerosols in Paris (France), Atmosph. Chem. Phys. Discuss., 13, 33237–33309. doi:10.5194/acpd-13-33237-2013", Available from: http://www.atmos-chem-phys-discuss.net/13/C13512/2014/acpd-13-C13512-2014.pdf (accessed 27 May 2014), 2014.

Bressi, M., Sciare, J., Ghersi, V., Bonnaire, N., Nicolas, J. B., Petit, J.-E., Moukhtar, S., Rosso, A., Mihalopoulos, N. and Féron, A.: A one-year comprehensive chemical characterisation of fine aerosols (PM2.5) at urban, suburban and rural background sites in the region of Paris (France), Atmos. Chem. Phys., 13, 7825-7844, doi:10.5194/acp-13-7825-2013, 2013.

Canonaco, F., Crippa, M., Slowik, J.G., Baltensperger, U. and Prévôt, A.S.H.: SoFi, an Igor based interface for the efficient use of the generalized multilinear engine (ME-2) for source apportionment: application to aerosol mass spectrometer data, Atmos. Meas. Tech., 6, 3649–3661. doi:10.5194/amt-6-3649-2013, 2013.

Davison, A. C. and Hinkley, D. V.: Bootstrap Methods and Their Application, Cambridge University Press, 1997.

Dutton, S. J., Vedal, S., Piedrahita, R., Milford, J. B., Miller, S. L. and Hannigan, M. P.: Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated PM2.5, Atmospheric Environment, 44(23), 2731–2741, doi:10.1016/ j.atmosenv.2010.04.038, 2010.

Efron, B.: 1977 Rietz Lecture - Bootstrap Methods - Another Look at the Jackknife, Ann. Stat., 7(1), 1–26, doi:10.1214/aos/1176344552, 1979.

Efron, B. and Tibshirani, R.: An Introduction to the Bootstrap, Chapman & Hall, 1993.

Friedlander, S. K.: Chemical element balances and identification of air pollution sources, Environ. Sci. Technol., 7(3), 235–240, doi:10.1021/es60075a005, 1973.

Ghersi, V., Rosso, A., Moukhtar, S., Lameloise, P., Sciare, J., Bressi, M., Nicolas, J. B., Féron, A. and Bonnaire, N.: Etude de contribution des sources de particules fines (PM2,5) en Ile-de-France, Pollution Atmosphérique, 63–72, 2010.

Henry, R. C. and Kim, B. M.: Extension of self-modeling curve resolution to mixtures of more than three components: Part 1. Finding the basic feasible region, Chemometrics and Intelligent Laboratory Systems, 8(2), 205–216, doi:10.1016/0169-7439(90)80136-T, 1990.

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., and Thurston, G.D.: PM source apportionment and health effects: 1. Intercomparison of source apportionment results, J. Expo. Sci. Env. Epid., 16, 275–286, 2006.

Huang, S., Rahn, K. A. and Arimoto, R.: Testing and optimizing two factor-analysis techniques on aerosol at Narragansett, Rhode Island, Atmospheric Environment, 33(14), 2169–2185, doi:10.1016/S1352-2310(98)00324-0, 1999.

Karanasiou, A. A., Siskos, P. A. and Eleftheriadis, K.: Assessment of source apportionment by Positive Matrix Factorization analysis on fine and coarse urban aerosol size fractions, Atmospheric Environment, 43(21), 3385–3395, doi:10.1016/j.atmosenv.2009.03.051, 2009.

Kim, B. M. and Henry, R. C.: Extension of self-modeling curve resolution to mixtures of more than three components: Part 2. Finding the complete solution, Chemometrics and Intelligent Laboratory Systems, 49(1), 67–77, doi:10.1016/S0169-7439(99)00029-5, 1999.

Kim, B. M. and Henry, R. C.: Extension of self-modeling curve resolution to mixtures of more than three components: Part 3. Atmospheric aerosol data simulation studies, Chemometrics and Intelligent Laboratory Systems, 52(2), 145–154, doi:10.1016/S0169-7439(00)00077-0, 2000.

Lahiri, S. N.: Resampling Methods for Dependent Data, Springer, 2003.

Lee, P. K. H., Brook, J. R., Dabek-Zlotorzynska, E. and Mabury, S. A.: Identification of the Major Sources Contributing to PM2.5 Observed in Toronto, Environ. Sci. Technol., 37(21), 4831–4840, doi:10.1021/es026473i, 2003.

Lenschow, P., Abraham, H. J., Kutzner, K., Lutz, M., Preuss, J. D., and Reichenbacher, W.: Some ideas about the sources of PM10, Atmos. Environ., 35, S23–S33, 2001.

Maykut, N. N., Lewtas, J., Kim, E. and Larson, T. V.: Source Apportionment of PM2.5 at an Urban IMPROVE Site in Seattle, Washington, Environ. Sci. Technol., 37(22), 5135–5142, doi:10.1021/es030370y, 2003.

Miller, M., Friedlander, S. and Hidy, G.: A chemical element balance for the Pasadena aerosol, Journal of Colloid and Interface Science, 39(1), 165–176, doi:10.1016/0021-9797(72)90152-X, 1972.

Minguillón, M. C., Querol, X., Baltensperger, U. and Prévôt, A. S. H.: Fine and coarse PM composition and sources in rural and urban sites in Switzerland: Local or regional pollution?, Science of The Total Environment, 427–428(0), 191–202, doi:10.1016/j.scitotenv.2012.04.030, 2012.

Mooibroek, D., Schaap, M., Weijers, E. P. and Hoogerbrugge, R.: Source apportionment and spatial variability of PM2.5 using measurements at five sites in the Netherlands, Atmospheric Environment, 45(25), 4180–4191, doi:10.1016/j.atmosenv.2011.05.017, 2011.

Norris, G. A., Vedantham, R., Wade, K., Zhan, P., Brown, S., Pentti, P., Eberly, S. I. and Foley, C.: Guidance document for PMF applications with the Multilinear Engine, U.S. Environmental Protection agency, Washington DC. Available from: cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=488084 (Accessed 28 June 2012), 2009.

Norris, G., Vedantham, R., Wade, K., Brown, S., Prouty, J. and Foley, C.: EPA Positive Matrix Factorization (PMF) 3.0: fundamentals & user guide, U.S. Environmental Protection Agency, 2008.

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

Paatero, P.: The multilinear engine - A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model, J. Comput. Graph. Stat., 8(4), 854–888, doi:10.2307/1390831, 1999.

Paatero, P.: User's guide for the multilinear engine program "ME2" for fitting multilinear and quasimultilinear models, University of Helsinki, Finland, 2000.

Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, Analytica Chimica Acta, 490(1-2), 277–289, 2003.

Paatero, P. and Tapper, U.: Positive Matrix Factorization - a Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values, Environmetrics, 5(2), 111–126, doi:10.1002/env.3170050203, 1994.

Perrone, M. G., Larsen, B. R., Ferrero, L., Sangiorgi, G., De Gennaro, G., Udisti, R., Zangrando, R., Gambaro, A. and Bolzacchini, E.: Sources of high PM2.5 concentrations in Milan, Northern Italy: Molecular marker data and CMB modelling, Science of The Total Environment, 414(0), 343–355, doi:10.1016/j.scitotenv.2011.11.026, 2012.

Poirot, R. L., Wishinski, P. R., Hopke, P. K. and Polissar, A. V.: Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont, Environmental science & technology, 35(23), 4622–4636, 2001.

Polissar, A., Hopke, P. and Paatero, P.: Atmospheric aerosol over Alaska - 2. Elemental composition and sources, J. Geophys. Res.-Atmos., 103(D15), 19045–19057, doi:10.1029/98JD01212, 1998.

Politis, D. N. and White, H.: Automatic Block-Length Selection for the Dependent Bootstrap, Econometric Reviews, 23(1), 53–70, doi:10.1081/ETC-120028836, 2004.

Poulakis, E., Theodosi, C., Sciare, J., Bressi, M., Ghersi, V. and Mihalopoulos, N.: Airborne mineral components and trace metals in Paris region: Spatial and temporal variability, manuscript in preparation, 2012.

Quass, U., Kuhlbusch, T. and Koch, M.: Identification of source groups of fine dust, Public report to the Environment Ministry of North Rhine Westphalia, Germany. Available from: http://www.lanuv.nrw.de/luft/berichte/FeinstaubNRW_2004_Summary.pdf (Accessed 24 July 2012), 2004.

Ramadan, Z., Eickhout, B., Song, X. H., Buydens, L. and Hopke, P. K.: Comparison of positive matrix factorization and multilinear engine for the source apportionment of particulate pollutants, Chemometrics and Intelligent Laboratory Systems, 66(1), 15–28, 2003.

Reff, A., Eberly, S. I. and Bhave, P. V.: Receptor modeling of ambient particulate matter data using positive matrix factorization: review of existing methods, Journal of the Air & Waste Management Association, 57(2), 146, 2007.

Singh, K.: On the Asymptotic Accuracy of Efron's Bootstrap, Ann. Statist., 9(6), 1187–1195, doi:10.1214/aos/1176345636, 1981.

Ulbrich, I.M., Canagaratna, M.R., Zhang, Q., Worsnop, D.R. and Jimenez, J.L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891–2918, 2009.

Vallius, M., Janssen, N. A. H., Heinrich, J., Hoek, G., Ruuskanen, J., Cyrys, J., Van Grieken, R., de Hartog, J. J., Kreyling, W. G. and Pekkanen, J.: Sources and elemental composition of ambient PM2.5 in three European cities, Science of The Total Environment, 337(1–3), 147–162, doi:10.1016/j.scitotenv.2004.06.018, 2005.

Wehrens, R., Putter, H. and Buydens, L.: The bootstrap: a tutorial, Chemometrics and Intelligent Laboratory Systems, 54(1), 35–52, 2000.

Winchester, J. W. and Nifong, G. D.: Water pollution in Lake Michigan by trace elements from pollution aerosol fallout, Water, Air, & Soil Pollution, 1(1), 50–64, 1971.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R. and Sun, Y.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, Anal. Bioanal. Chem., 401, 3045–3067, doi:10.1007/s00216-011-5355-y, 2011.

Zheng, M., Cass, G. R., Schauer, J. J. and Edgerton, E. S.: Source Apportionment of PM2.5 in the Southeastern United States Using Solvent-Extractable Organic Compounds as Tracers, Environmental Science & Technology, 36(11), 2361–2371, doi:10.1021/es011275x, 2002.