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

M. Bressi¹,², J. Sciare¹, V. Ghersi³, N. Mihalopoulos⁴, J.-E. Petit¹,⁵, J. B. Nicolas¹,², S. Moukhtar³, A. Rosso³, A. Féron¹, N. Bonnaire¹, E. Poulakis⁴ and C. Theodosi⁴.

¹Laboratoire des Sciences du Climat et de l’Environnement, LSCE, UMR8212, CNRS-CEA-UVSQ, Gif-sur-Yvette, 91191, France.
²French Environment and Energy Management Agency, ADEME, 20 avenue du Grésillé, BP90406 49004, Angers Cedex 01, France.
³AIRPARIF, Surveillance de la Qualité de l’Air en Ile de-France, Paris, 75004, France.
⁴Environmental Chemical Processes Laboratory, ECPL, Heraklion, Voutes, Greece.

Correspondence to: michael.bressi@ensiacet.fr

Abstract

The present study aims at identifying and apportioning the major sources of fine aerosols in Paris (France) - the second most populated “Larger Urban Zone” in Europe-, and determining their geographical origins. It is based on the daily chemical composition of PM$_{2.5}$ characterised during one year at an urban background site of Paris (Bressi et al., 2013). Positive Matrix Factorization (EPA PMF3.0) was used to identify and apportion the sources of fine aerosols; bootstrapping was performed to determine the adequate number of PMF factors, and statistics (root mean square error, coefficient of determination, etc.) were examined to better model PM$_{2.5}$ mass and chemical components. Potential Source Contribution Function (PSCF) and Conditional Probability Function (CPF) allowed the geographical origins of the sources to be assessed; special attention was paid to implement suitable weighting functions. Seven factors namely ammonium sulfate (A.S.) rich factor, ammonium nitrate (A.N.) rich factor, heavy oil combustion, road traffic, biomass burning, marine aerosols and metal industry were identified; a detailed discussion of their chemical characteristics is reported. They contribute 27, 24, 17, 14, 12, 6 and 1% of PM$_{2.5}$ mass (14.7 µg/m$^3$) respectively on the annual average; their seasonal variability is discussed. The A.S. and A.N. rich factors have undergone north-eastward mid- or long-range transport from Continental Europe, heavy oil combustion mainly stems from northern France and the English Channel, whereas road traffic and biomass burning are primarily locally emitted. Therefore, on average more than half of PM$_{2.5}$ mass measured in the city of
Paris is due to mid- or long-range transport of secondary aerosols stemming from continental Europe, whereas local sources only contribute a quarter of the annual averaged mass. These results imply that fine aerosols abatement policies conducted at the local scale may not be sufficient to notably reduce PM$_{2.5}$ levels at urban background sites in Paris, suggesting instead more coordinated strategies amongst neighbouring countries. Similar conclusions might be drawn in other continental urban background sites given the transboundary nature of PM$_{2.5}$ pollution.

1. Introduction

Aerosols are airborne solid or liquid particles arising from various natural and anthropogenic sources (IPCC, 2007). They are directly emitted in the atmosphere as particles (primary aerosols) or result from gas to particle conversions (secondary aerosols, Raes et al., 2000). Their chemical characteristics are miscellaneous given the diversity of their sources as well as their formation and transformation processes. Aerosols are subjects of concern for sanitary (Bernstein et al., 2004; Pope and Dockery, 2006), climatic (Forster et al., 2007; Isaksen et al., 2009) and economic reasons (Aphekom, 2012; US-EPA, 2011a), to name a few (see US-EPA, 2011b for further details). Due to their enhanced adverse health effects in particular, fine particles (PM$_{2.5}$ i.e. particles with an aerodynamic diameter less than or equal to 2.5 µm) have been subject to a stringent legislative framework during the last years.

The city of Paris (France) is concerned by the aforementioned issues. First, about 11 million inhabitants (ca. 18% of the French population) are exposed to PM$_{2.5}$ pollution in this “Larger Urban Zone” (LUZ), which is the second most populated in Europe (Eurostat, 2012). Aphekom (2011) estimated that reducing PM$_{2.5}$ levels in Paris to the recommended World Health Organisation (WHO) value of 10 µg/m$^3$ would lead to a gain in life expectancy of ca. half a year in this city. Second, because different megacities in the world have been reported to impact their regional climates (Molina and Molina, 2004 and references therein), the anthropogenic emissions of air pollutants in Paris could lead to the same consequences. Third, substantial economic benefits should result from a reduction of PM$_{2.5}$ levels in Paris, due to the decrease of hospital admissions and corresponding work losses. For instance, Aphekom (2011) estimated that a reduction of PM$_{2.5}$ levels in Paris to the WHO guidelines would lead to more than 4 billion euros benefits. Therefore, there is a need to lower fine aerosol levels in Paris, which requires effective PM$_{2.5}$ abatement strategies. It should be mentioned that in a broader context, PM$_{2.5}$ levels measured in Paris are generally lower than in other European urban environments: Paris (17.8 µg.m$^{-3}$), Zurich (19.0 µg.m$^{-3}$), Prague (19.8 µg.m$^{-3}$), Vienna (21.8 µg.m$^{-3}$), Barcelona (28.2 µg.m$^{-3}$) (Bressi, 2012; Putaud et al., 2010). Implementing effective PM$_{2.5}$ abatement strategies is thus not only necessary in Paris but also in most European urban environments.
At the present times, such strategies seem to be rather insufficient in this city. Despite the abatement policies implemented (e.g. prefectoral order n° 2011-00832 of the 27 October 2011 targeting sources such as wood burning, agricultural fertilizers, industrial emissions, etc.), PM$_{2.5}$ annual levels in Paris have remained rather stable during the last ten years (AIRPARIF, 2012). The lack of knowledge of the sources and the geographical origins of fine aerosols in this city may explain the ineffectiveness of such policies. In fact, until now the major sources of PM$_{2.5}$ have only been estimated through emission inventories (EI), a methodology that leads to significant uncertainties. As an illustration, the French Interprofessional Technical Centre for Studies on Air Pollution (CITEPA) estimated uncertainties of 48% for PM$_{2.5}$ emissions in France in 2008 (CITEPA, 2010). Comparisons with the EI implemented by AIRPARIF (which is the regional air quality network of Paris) lead to substantial differences (Bressi, 2012); discrepancies between AIRPARIF EI and the European Monitoring and Evaluation Program (EMEP) are also considerable (Hodzic et al., 2005). Furthermore, such approaches do not take into account the secondary fraction of fine aerosols, which are however predominant in Europe (Putaud et al., 2010) and in Paris in particular (Bressi et al., 2013). By contrast, source apportionment (SA) techniques -such as Positive Matrix Factorization (PMF)- would allow considering this secondary aerosol fraction, and would thus appear more suitable to identify and apportion PM sources (Viana et al., 2008; Belis et al., 2013). Nonetheless, this type of studies has not been conducted on aerosols on the annual scale in Paris yet and is rare in France (Karagulian and Belis, 2011).

In addition, the geographical origins of PM$_{2.5}$ are poorly documented in this city. To the best of our knowledge, only one study conducted by Sciare et al. (2010) has addressed this issue for PM$_{2.5}$. They reported that eastward long-range transport can significantly affect PM$_{2.5}$ levels in the region of Paris by bringing high levels of secondary aerosols mainly composed of ammonium sulfate (A.S.) and ammonium nitrate (A.N.). Freutel et al. (2013) reach the same conclusion, reporting the highest PM$_{1}$ levels in the region of Paris when air masses are advected from continental Europe. Interestingly, modelling studies conducted by Vautard et al. (2003) and Bessagnet et al. (2005) have also reported a noticeable influence of eastward long-range transport on ozone and PM$_{10}$ levels, respectively, observed in the region of Paris. Nevertheless, the results reported by Sciare et al. (2010) and Freutel et al. (2013) on fine aerosols were based on few weeks periods (19 and 30 days, respectively) occurring during late spring/summer and thus suffer from a lack of representativeness on a longer time scale. The determination of the geographical origins of PM$_{2.5}$ in Paris thus requires longer observations to reach more robust conclusions, which could ask for the use of statistical tools such as Potential Source Contribution Function (PSCF) and Conditional Probability Function (CPF).

In this context, the "Particles" research project involving the regional air quality network (AIRPARIF) and the Climate and Environmental Sciences Laboratory (LSCE) was implemented. It aims...
at documenting the chemistry, the sources and the geographical origins of fine aerosols in the region of Paris, during one-year, on a daily basis. A full description of the project can be found in AIRPARIF and LSCE (2012) and Ghersi et al. (2010, 2012). The daily chemical composition of PM$_{2.5}$ in the region of Paris obtained within the "Particles" project has been discussed in detail in Bressi et al. (2013). Based on this work, the present paper aims at:

1. identifying the sources of fine aerosols at an urban site in Paris (Sect. 4.1),
2. identifying the geographical origins of these sources (Sect. 4.2),
3. determining the contribution of each source to PM$_{2.5}$ mass on yearly and seasonal bases (Sect. 4.3).

Section 2 will briefly describe i) the sampling procedure and the chemical analyses conducted and ii) the statistical tools used to fulfil these objectives (PMF, PSCF and CPF). Section 3 will show how the appropriate number of PMF factors can be chosen through the bootstrap technique. Technical results regarding the ability of PMF to model PM$_{2.5}$ mass and chemical components will be presented, and the methodology used to determine the suitable PSCF and CPF weighting functions discussed. The identification of PMF factors to real physical sources will be reported in Sect. 4.1, after having compared their chemical profiles to the literature. Section 4.2 will focus on the geographical origins of PM$_{2.5}$ sources discussing PSCF and CPF results. Finally, the yearly and seasonal contributions of each source will be compared to other European studies, chosen according to their presumable geographical origins (Sect. 4.3).

2. Material and methods

2.1. Sampling and chemical analyses

A full description of the sampling site and the analytical methods used can be found in Bressi et al. (2013) and Poulakis et al. (2012); only the essential information will be reported here.

2.1.1. Sampling

The sampling site is located in the city centre of Paris (4$^{th}$ district, 48°50’56”N, 02°21’55”E, 20 m above ground level, a.g.l.) and is representative of an urban background (Bressi et al., 2013; Ghersi et al., 2010, 2012). It is worthwhile noting that PM$_{2.5}$ levels and chemical composition are very homogeneous in the Paris LUZ (Bressi et al., 2013). For instance, urban and suburban sites (distant by 10 km) typically exhibit PM$_{2.5}$ levels that are not statistically significantly different, whereas levels

...
measured at rural locations (distant by 50 km) are ca. 25% lower than at the urban site. This urban sampling site is thus regarded as being representative of Paris metropolitan area. It should however be highlighted that this site is at 20m a.g.l. which might prevent near ground sources (e.g. road dust) to be considered in our study. Fine aerosols (PM$_{2.5}$) were collected every day from 00:00 to 23:59 LT, during one year from 11 September 2009 to 10 September 2010. Two collocated Leckel low volume samplers (SEQ47/50) running at 2.3 m$^3$/h were used for filter sampling. One Leckel sampler was equipped with Quartz filters (QMA, Whatman, 47 mm diameter) for carbon analyses, the other with Teflon filters (PTFE, Pall, 47 mm diameter, 2.0 µm porosity) for gravimetric, ion and metal measurements. 28 samples (i.e. 8% of the dataset) were discarded because of power failures, analytical problems, etc. (see Table S1 for the detailed list).

2.1.2 Chemical analyses

Chemical analyses of the major PM$_{2.5}$ components are thoroughly described in Bressi et al. (2013). Briefly i) gravimetric mass (PM$_{\text{grav}}$) was determined with a microbalance (Sartorius, MC21S), ii) elemental and organic carbon (EC and OC, respectively) were analysed by a thermal-optical method (Sunset Lab., OR, USA) using the EUSAAR_2 protocol (Cavalli et al., 2010) and iii) water-soluble ions (Cl$^-$, NO$_3^-$, SO$_4^{2-}$, Na$^+$, NH$_4^+$, K$^+$, Mg$^{2+}$, Ca$^{2+}$) were quantified with Ion Chromatographs (IC). Note that the gravimetric procedure used underestimates PM$_{2.5}$ mass compared to EU reference methods (EN 14907) by ca. 20% on average (see Bressi et al., 2013). Organic matter (OM) was inferred from OC measurements using an OC to OM conversion factor of 1.95 (Bressi et al., 2013). Metals including Al, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Cd and Pb were analysed after acid microwave digestion by Inductively Coupled Plasma and Mass Spectrometry as reported in Poulakis et al. (2012) and Theodosi et al. (2010). Note that some minerals (e.g. Al, Ti, etc.) might be underestimated due to the acid microwave digestion procedure used here (with HNO$_3$), which might not be able to dissolve entirely these compounds (see e.g. Robache et al., 2000).

Monosaccharides and sugar alcohols, comprising levoglucosan, mannosan, arabitol and mannitol were also analysed. They were determined following the technique reported in Iinuma et al. (2009), using a high performance anion exchange chromatograph (HPAEC, DIONEX, model ICS 3000) with pulsed amperometric detection (PAD). Separation was performed with a Dionex CarboPac MA1 4-mm diameter column (see Sciare et al., 2011 for further information).

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; Hopke et al., 2006; Watson et al., 2002). An extensive description of SA methods and receptor models can be found in the supplementary material (Sect. S1).

2.2.1. Positive Matrix Factorization (PMF)

The PMF model (Paatero and Tapper, 1994; Paatero, 1997; Ulbrich et al., 2009; Zhang et al., 2011) is used here (see Sect. S1). PMF is a receptor model that assumes mass conservation and uses a mass balance analysis to identify and apportion sources of PM; it aims at resolving the following equation:

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

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, \( f_{kj} \) is the concentration of the \( j^{th} \) chemical species in the material emitted by the \( k^{th} \) source and \( e_{ij} \) represents the residual element, or the PMF model error, for the species \( j \) measured in the sample \( i \). Equation 1 is solved by minimising a Q function defined as:

\[ Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( \frac{e_{ij}}{\sigma_{ij}} \right)^2 \]  

where \( \sigma_{ij} \) is the uncertainty associated to the \( j^{th} \) species in the \( i^{th} \) sample. Different Q functions can be defined: \( Q_{\text{true}} \) calculated including all data and \( Q_{\text{robust}} \) calculated excluding outliers i.e. data for which the scaled residual \( (e_{ij}/\sigma_{ij}) \) is greater than 4. (Note that \( Q_{\text{theoretical}} \) will not be studied here as explained in Sect. S2.) A standalone version of PMF using the second version of the multi-linear engine algorithm (ME-2; Paatero, 2000; Norris et al., 2009; Canonaco et al., 2013) has been developed by the United States Environmental Protection Agency (US-EPA) and is used in our study. This version will be named EPA PMF3.0 in the following and can be downloaded at http://www.epa.gov/heasd/products/pmf/pmf.html.

2.2.2 Data preparation

Two input datasets are required by the EPA PMF3.0 model: one with the chemical species atmospheric concentrations of every sample, the other with their associated uncertainties. Both datasets were here constructed following the advices given by Reff et al. (2007) in his review on PMF existing methods and Norris et al. (2008) in the EPA PMF3.0 user guide. A detailed description of both datasets can be found in Sect. S2. It is worthwhile noting that Al, Cr, As, arabitol and mannitol...
have not been taken into account for PMF analysis since their atmospheric concentrations were mostly below their Method Quantification Limit (see Sect. S2).

2.2.3 Robustness of PMF results

Robustness of PMF results can be assessed by different methods that will be discussed in Sect. 3, including Q function analysis, residual analysis, predicted versus observed concentrations interpretation, etc. In addition, the bootstrap method (Davison and Hinkley, 1997; Efron, 1979; Efron and Tibshirani, 1993; Singh, 1981; Wehrens et al, 2000) implemented in the EPA PMF3.0 software has been performed to estimate the stability and the uncertainty of the PMF solution, with a focus on the F matrix. It will be shown in Sect. 3.1 that it will also help in better determining the adequate number of factors to choose. Further information on the bootstrap theory and its application to our study can be found in Sect. S3. Note that bootstrap matrices will be noted with an “*” in the following.

2.2.4 PMF technical parameters

Concerning base model runs - i.e. runs without performing bootstrapping - (1) twenty runs were conducted, (2) the initial F and G matrices (so-called "seed") were randomly selected and (3) different numbers of factors ranging from 3 to 10 were tested (a detailed discussion of the number of factor chosen will be made in Sect. 3.1). The run exhibiting the lowest Q_{robust} value was retained for further analysis. Bootstrapping was then carried out, performing 100 bootstrap runs, using a random seed (initial F* and G* matrices), a block size of 52 - determined by the methodology of Politis and White (2004) - and a minimum correlation coefficient (R-value) of 0.6 (unless otherwise stated later on). Results will be discussed in Sect. 3.1.

2.3. Geographical origins of PM$_{2.5}$

Geographical origins of PM$_{2.5}$ chemical compounds and sources were assessed by two different methods that are the Conditional Probability Function (CPF) and the Potential Source Contribution Function (PSCF).

2.3.1 Conditional Probability Function
The Conditional Probability Function was applied to PMF results. It estimates the probability that a source contribution, from a given wind direction, exceeds a predetermined threshold criterion (Ashbaugh et al., 1985; Kim and Hopke, 2004; Kim et al., 2003). It is defined as:

\[ \text{CPF}_\theta = \frac{m_\theta}{n_\theta} \]  

(3)

where \( m_\theta \) is the number of occurrence that a source contribution, coming from the wind direction \( \theta \), exceeds a pre-determined threshold criterion, and \( n_\theta \) is the total number of times the wind came from that same \( \theta \) direction. 48-hours air mass back-trajectories with an altitude endpoint of 500 meters were calculated every 6 hours from 11 September 2009 06:00:00 LT to 11 September 2010 00:00:00 LT using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Rolph, 2011). Backtrajectories were then defined according to their overall path in one of the sixteen \( \theta \) directions separated by 22.5° (i.e. N, NNE, NE, etc.). This procedure allows curved backtrajectories to be binned in the appropriate direction, but is laborious and prone to user-approximations. Calm winds (i.e. wind speed below 1 m.s\(^{-1}\)) were excluded from the dataset, which represents 3% of wind data. A total of 1,417 air mass back-trajectories were taken into account for CPF calculations. Different threshold criteria were tested and the 75\(^{th}\) percentile was retained as it better illustrates source locations. This threshold is in line with what is reported elsewhere (e.g. Amato and Hopke, 2012; Jeong et al., 2011a; Kim et al., 2004). Furthermore, a weighting function was empirically implemented to lower uncertainties associated with low \( n_\theta \) values (thus resulting in high \( \text{CPF}_\theta \) values, see Sect. 3.4). This function was defined as:

\[
W_{\text{CPF}}(n_\theta) = \begin{cases} 
1.00 & \text{for } n_\theta \geq 0.75 \times \max(n_\theta) \\
0.75 & \text{for } 0.75 \times \max(n_\theta) > n_\theta \geq 0.50 \times \max(n_\theta) \\
0.50 & \text{for } 0.50 \times \max(n_\theta) > n_\theta \geq 0.25 \times \max(n_\theta) \\
0.25 & \text{for } 0.25 \times \max(n_\theta) > n_\theta 
\end{cases}
\]  

(4)

where \( \max(n_\theta) = 131 \) in this study (for the SW direction).

2.3.2 Potential Source Contribution Function

Potential Source Contribution Function (PSCF) was introduced by Ashbaugh et al. (1985) and can be defined as "a conditional probability describing the spatial distribution of probable geographical source locations inferred by using trajectories arriving at the sampling site" (Polissar et al., 1999).

48-hour back-trajectories, with an altitude endpoint of 500m, were calculated every six hours from 11 September 2009 06:00:00 LT to 11 September 2010 00:00:00 LT, using a PC-based version of HYSPLIT (version 4.9; Draxler and Hess, 1997). Meteorological parameters comprising ambient temperature, relative humidity and precipitation were determined along each trajectory. Wet deposition was estimated by assuming that precipitation (\( \geq 0.1 \) mm) will clean up the air parcel.
(PSCF=0). PSCF was set to 0 for all air parcels determined before (in terms of time, but after in terms of back-trajectory calculation) precipitation occurred.

The PSCF calculation method (Polissar et al., 1999, 2001a) can be resumed as:

\[ PSCF_{ij} = \frac{m_{ij}}{n_{ij}} \]  

(5)

where \( n_{ij} \) is the total number of endpoints falling in the air parcel of address \((i, j)\), and \( m_{ij} \) is the number of endpoints of that parcel for which measured concentrations exceed a user-determined threshold criterion. The threshold chosen is the 75th percentile, which will allow a comparison with CPF results, and which is in agreement with the literature (e.g. Begum et al., 2010; Hsu et al., 2003; Sunder Raman and Hopke, 2007).

To remove high PSCF uncertainties associated with small \( n_{ij} \) values, a weighting function - \( W_{PSCF}(n_{ij}) \) - is generally implemented (e.g. Hwang and Hopke, 2007; Jeong et al., 2011b; Polissar et al., 2001a, 2001b; Zeng and Hopke, 1989). Weighting factors were empirically determined and the resulting weighting function is defined as:

\[
W_{PSCF}(n_{ij}) = \begin{cases} 
1.00 & \text{for } n_{ij} \geq 0.85 \times \max[\log(n_{ij} + 1)] \\
0.73 & \text{for } 0.85 \times \max[\log(n_{ij} + 1)] > n_{ij} \geq 0.60 \times \max[\log(n_{ij} + 1)] \\
0.48 & \text{for } 0.60 \times \max[\log(n_{ij} + 1)] > n_{ij} \geq 0.35 \times \max[\log(n_{ij} + 1)] \\
0.18 & \text{for } 0.35 \times \max[\log(n_{ij} + 1)] > n_{ij}
\end{cases}
\]  

(6)

where \( \max[\log(n_{ij}+1)] = 3.6 \) or \( \max(n_{ij}) = 3980 \) in our study. The latter value corresponds to the maximum number of trajectories going through a sole cell. A binomial smoothing (i.e. a Gaussian filter) implemented in the IGOR Pro 6 software (http://www.wavemetrics.com/) was then applied to PSCF results.

3. Results

3.1. Factors and chemical species to retain

3.1.1 Number of factors

Choosing the accurate number of factors (p-values) in models has always been a challenging question (Cattell, 1966; Henry, 2002; Henry et al., 1999; Malinowski, 1977). Too few factors will result in a mixing of different sources in the same factor as well as high residuals, whereas too many factors will lead to meaningless sources made up of a sole chemical species. Different parameters are used to determine the appropriate p-value, including the examination of Q-values, scaled residuals, or post-PMF regression, to name a few (Norris et al., 2008; Reff et al., 2007). All these parameters are
here investigated, but a special focus on Q-values, bootstrap results and the physical meaning of factor profiles has been made to determine the adequate number of factors to choose.

The figures mentioned in the following refer to simulations run with the optimal number of chemical species (discussed below). Eight different configurations are tested, with p-values ranging from 3 to 10, each configuration being run 20 times as mentioned in Sect. 2.2.4. Configurations with 3, 4, 5, 9 and 10 factors are not suitable because:

i) a high base run variability is noticeable (unless for 3 factors) when examining the sum of the squared difference between the scaled residuals for each pair of base runs (d-values) and

ii) they lead to questionable factor profiles with a clear combinations of multiple sources in an individual factor for 3, 4 and 5 factor configurations, and factors with a single chemical species for 9 and 10 factor configurations.

Configurations with 6, 7 and 8 factors give on the other hand fairly good results with i) stable base runs and ii) meaningful factor profiles. To discriminate between these three simulations, bootstrap results are inspected in more detail (Table S4). Regarding the 6- and 7-factor configurations, each boot factor is assigned (r≥0.6) to base factors for at least 94 and 96% of the runs (n=100), respectively, hence highlighting their robustness. On the other hand, the 8-factor solution shows less satisfactory results, with a boot factor being assigned to the corresponding base factor for 78% of the runs. This 8-factors configuration was consequently rejected. Since no significant bootstrap discrepancies are observed for 6 and 7 factors, further tests are conducted by increasing the r-value of the bootstrap mapping. With r≥0.7, the 6-factor configuration shows a less robust factor (83%), than the 7-factor one (95%); the latter assumption will therefore be retained in the following.

Although bootstrapping is usually not used for this purpose, it consequently appears to be a valuable statistical tool to choose the adequate number of factors in PMF simulations.

The physical meaning of factor profiles will be discussed in detail in Sect. 4.1.

3.1.2 Appropriate chemical species

Chemical species were primarily retained in or excluded from simulations according to the coefficient of determination of their observed versus predicted concentrations. We decided to categorise as bad (i.e. exclude from the dataset) every species exhibiting an r-value lower than 0.5, which concerns Ca$^{2+}$, Zn and Ti (r=0.08, 0.13 and 0.17, respectively). The only exception was made for Ni showing a coefficient of determination equal to 0.47, partly due to a lack of data during the months of April and May, nevertheless bringing valuable information for source identification (Sect. 4.1).
3.2. Technical results

Further technical results concerning the 7-factor configuration will now be reported to discuss the robustness and the quality of our PMF results. First, no significant base run variability is observed as it is attested by $Q_{\text{robust}}$ values (5569.0±0.1 on average, n=20) and d-values (Table S5). The $Q_{\text{true}}$ to $Q_{\text{robust}}$ ratio is equal to one (1.00±0.00 on average, n=20) indicating that no peak events are substantially influencing the model.

Table 1 reports statistics based on the annual comparison between observed (i.e. measured) and predicted (i.e. modelled) concentrations for each chemical species and for PM mass. PM is very well reproduced by PMF, showing a coefficient of determination and a slope close to one ($r^2$=0.97, $y=1.01\pm0.01x - 0.25\pm0.18 \, \mu g/m^3$, n=337). Most chemical species also exhibit very good coefficient of determination ($r^2$ higher than 0.8 for 11 compounds, and between 0.7 and 0.8 for 4 compounds), with the exception of EC, Cd and Ni showing reasonably good coefficients (between 0.4 and 0.6). Slopes are close to one for most species (higher than 0.7 for 14 compounds), except for Ni (0.4). The limitations regarding the ability of the model to simulate Ni concentrations should be borne in mind when discussing its results.

The seasonal variability of statistics describing the ability of PMF to simulate PM$_{2.5}$ mass is reported in Table 2. Three variables were studied: the coefficient of determination ($r^2$), the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE). The two latter are defined as:

\begin{align}
RMSE &= \sqrt{\frac{1}{n} \sum_{i} (PM_{\text{modeled}} - PM_{\text{measured}})^2} \\
MAPE &= \frac{100}{n} \sum_{i} \left| \frac{PM_{\text{modeled}} - PM_{\text{measured}}}{0.5 \times (PM_{\text{modeled}} + PM_{\text{measured}})} \right|
\end{align}

where PM$_{\text{modeled}}$ is the sum of the seven source contributions for a given day, and n is the number of samples. These statistics are widely used for the evaluation of models (e.g. Stern et al., 2008). Very good coefficients of determination are found throughout the year, ranging from 0.89 to 0.98. Good RMSE are also observed and range from 1.4 to 2.0 $\mu g/m^3$, whereas MAPE values vary between 6 and 10 %. The summer season is the less well simulated. This can be due to the lower PM levels observed during this season, resulting in lower $r^2$ and MAPE, but comparable RMSE values compared with other seasons. It could also be related to the absence of clearly identified biogenic and mineral dust sources in our study (see Sect. 4.1) for which emissions are prevalent during summer. Although those statistics give valuable information on the ability of PMF to model PM mass, they are generally not reported in PMF studies thus making impossible comparisons with our results.

3.3. PMF factors
The 7-factor profiles are reported in Fig. 1 and Fig. 2. Figure 1 allows factor identification, by highlighting the relative contribution of every chemical species in a given factor. Fig. 2 gives the contribution (in µg/m³) of chemical species to each source, i.e. gives the influence of each chemical compound on source contributions to PM mass. Interpreting bootstrap profiles, instead of factor profiles of the optimal base run, is preferred here as it allows uncertainties to be estimated. These uncertainties are displayed by different percentiles of bootstrap runs (Fig. 1). Figures 1 and 2 will be discussed in detail in Sect. 4.1.

Regarding factor contributions, to the best of our knowledge bootstrap results are not documented for the G matrix in EPA PMF3.0 output files. This thus does not allow the uncertainties associated with this G matrix to be estimated. The results given here correspond to the base run that gave the smallest Q_{robust}. Figures 3 and S1 report the daily contribution (in µg/m³) of each source to PM mass during the whole campaign; it should be recalled that some days were excluded from the dataset (Table S1). Correlations between factor time series and their presumable tracers are reported in Table S6. Figure 4 shows the relative contribution (in %) of each source to every chemical species, giving valuable information on the apportionment of compounds emitted by different sources (e.g. OM), and on the real ability of chemical constituents to be source-tracers (e.g. levoglucosan). The contribution of the unaccounted fractions (i.e. proportion of a chemical species that is not attributed to any factor) is below 5% for most species, with the exception of nitrate, K, Cu, Pb and Cd (6, 7, 10, 13 and 17%, respectively). Figure 3, 4 and S1 will be discussed in Sect. 4.3.

3.4. Geographical origins

The geographical distribution of the 48-hour air mass back-trajectories observed during the entire project is reported in Fig. S2. On the left hand side, a logarithmic scale was implemented to better illustrate the number of trajectories going through each cell (n_{ij}, ranging from 0 to 3980). This figure was constructed by plotting the logarithm of (n_{ij}+1) for each cell of address (i, j). Note that n_{ij} values will be used for PSCF calculations (see Sect. 2.3.2). On the right hand side, the number of trajectories per wind direction (n_θ) is plotted and is used for CPF calculations (Sect. 2.3.1). Regarding this last method, relatively high numbers of air mass trajectories are observed from SSW to NE sectors (ranging from 90 to 131 trajectories according to wind directions) and lower numbers are reported from ENE to S sectors (from 26 to 78). Applying the weighting function defined in Eq. 4 allows CPF values to be lowered for ENE to S sectors. Comparable results are found with the PSCF methodology, exhibiting a high number of trajectories per air parcel all around the region of Paris (>500) but in the S to ENE directions. Contrarily to the previous figure, this illustration gives further information on the
distances travelled by air masses with respect to Paris. The number of trajectories per cell is generally
i) higher than 500 from west of France to Benelux, ii) between 50 and 500 from southwest of France,
through England until Denmark and eastern Germany and iii) lower than 20 for further geographical
regions. The PSCF weighting function (Eq. 6) will again allow PSCF values to be reduced in the cells
exhibiting low $n_{ij}$ values. Hence, the assessment of the influence of emissions from southern or
eastern Europe on the city of Paris will not be possible in our study, due to the low number of
trajectories per cell found in these areas, leading to a lack of statistical robustness of CPF and PSCF
results.

4. Discussion

4.1 Source identification

Each PMF factor was interpreted by studying its chemical profile (F matrix). The interpretation of the
7 factors will be discussed from the easiest to the most complicated PMF factor to interpret. A
comparison with other European source apportionment studies will be given at the end of this Sect.

4.1.1. Biomass burning

The physical and chemical characteristics of biomass burning aerosols have extensively been studied
(Crutzen and Goldammer, 1993; Reid et al., 2005). Submicron particles of biomass burning origin are
typically made up of OC (80%), EC (5-9%) and trace inorganic compounds (12-15%) such as
potassium, sulfate, chloride and nitrate (Reid et al., 2005). Carbonaceous material (EC and a
proportion of OC), potassium and chloride are likely in the particle core (Posfai et al., 2003), whereas
sulfate, nitrate, organic acids and semi-volatile organic species are condensed on pre-existing
particles (Reid et al., 2005). It should be noted that fuel types and combustion efficiencies will lead to
a wide variety of specific chemical compositions (Fine et al., 2001, 2002, 2004). Good tracers of this
source are monosaccharide derivatives from the pyrolysis of cellulose and hemicellulose, such as
levoglucosan, mannosan and galactosan (Locker, 1988; Puxbaum et al., 2007; Simoneit, 2002;
Simoneit et al., 1999).

In this study, a biomass burning (BB) source is identified through the strong presence of
levoglucosan and mannosan in a single factor (84 and 80% of their mass, respectively, Fig. 1; unless
otherwise stated median values will be reported when referring to Fig. 1). In addition, noticeable
proportions of potassium, OM, chloride, EC, nitrate and ammonium are present (26, 19, 15, 12, 9 and
4%, respectively). Trace metal elements such as Pb and Ni are also observed (11 and 7%,
respectively) and may result from the absorption of heavy metals present in soil and water by biomass (Sharma and Dubey, 2005). Both compounds have been found in PM$_{2.5}$ resulting from wood combustion in Europe (Alves et al., 2011).

Figure 2 reports the mass contribution (in $\mu$g/m$^3$) of every chemical compound in this BB source. The major contributors are OM, nitrate, EC and levoglucosan (61, 13, 9 and 7% of the source mass, respectively; unless otherwise stated average values will be reported when referring to Fig. 2), the other compounds accounting for less than 5% by weight of this source. Hence, the wood burning contribution to PM$_{2.5}$ mass is mainly governed by carbonaceous materials, and especially organic matter. Interestingly, the relatively high proportion (by weight) of nitrate suggests that this biomass burning source has undergone atmospheric ageing, implying that BB aerosols freshly emitted by the region of Paris may not be the main contributor to this source, which is in agreement with its geographical origin (see later in Sect. 4.2) and the literature (Crippa et al., 2013a, 2013b).

The OC/EC, OC/Levoglucosan, K$^+$/Levoglucosan ratios are 3.4, 4.7, and 0.24, respectively (with an OC to OM conversion factor of 1.95, Bressi et al., 2013). Only insights into the nature of this biomass source can be given through these ratios, as they are highly variable according to the type of biomass combusted (softwood, hardwood, leaves, straws, etc.), the combustion conditions, the type of locations and the measurement techniques used (especially for EC and OC concentrations). Our OC to EC ratio of 3.4 is on the same order of magnitude as the ratios reported by Schmidl (2005 cited by Puxbaum et al., 2007) for beech and spruce (2.7 and 2.6, respectively) that are widespread trees in France and neighbouring countries (Simpson et al., 1999). Our OC to Levoglucosan ratio of 4.7 is close to the ratios reported by Schauer et al. (2001) of 3.9 and 4.3 for pine and oak, respectively, and by Schmidl (2005 cited by Puxbaum et al., 2007) of 5.0 for spruce. It is however lower than the recommended average US ratio of 7.35 (Fine et al., 2002), and Austria ratio of 7.1 (Schmidl, 2005 cited by Puxbaum et al., 2007). Interestingly, our corresponding OM to levoglucosan ratio of 9.2 is close to the values of 10.3 and 10.8 estimated for fine wood burning aerosols in the region of Paris by Sciare et al. (2011) and in the French Alpine region (Grenoble) by Favez et al. (2010), respectively. Finally, our K$^+$ to Levoglucosan ratio of 0.24 is in the 0.03 to 0.90 range of the different types of biomass combustion ratios compiled by Puxbaum et al. (2007), and appeared to be close to the 0.20 value reported by Schauer et al. (2001) for pine, or 0.16 value reported by Fine et al. (2001) for softwood.

To summarize, a biomass burning source was identified with the help of specific tracers, and could possibly originate from the wood combustion of trees such as beech, spruce, pine and oak (that are widespread in France and surrounding countries), although the contribution of agricultural and garden waste burning cannot be excluded. This source has undergone atmospheric ageing, suggesting that a proportion is imported from outside Paris.
4.1.2. Road traffic

Road traffic aerosols are of high complexity due to the diversity of emission processes (exhaust versus non-exhaust), and their primary and secondary natures. Tailpipe aerosols are primarily composed of OC and EC, although significant amounts of inorganic species such as ammonium nitrate can rapidly be formed by gas-to-particle conversion (Fraser et al., 1998). Non-exhaust aerosols typically arise from break wear, tyre wear, road wear, and road dust abrasion, and can be distinguished from exhaust aerosols by their high contents of heavy metals (e.g. Fe, Cu, Mn, Sb, etc.).

However, the finding of chemical tracers related to each abrasion process still constitutes an active field of research (Thorpe and Harrison, 2008).

In our study, the road traffic source was identified through the presence of characteristic metals and carbonaceous materials. Figure 1 shows that 87, 75, 58, 41, 25, 12 and 8% of Mn, Fe, Cu, EC, OM, Ni and Mg$^{2+}$, respectively, contribute to this source. Mn, Fe, Cu, Ni and Mg$^{2+}$ certainly stem from non-exhaust processes, and have all been detected from brake wear (Garg et al., 2000; Hildemann et al., 1991; Kennedy and Gadd, 2003), tire wear (Adachi and Tainosho, 2004), road wear (Kennedy and Gadd, 2003) and road dust (Schauer et al., 2006) emissions. As already mentioned, it remains complicated - if not impossible - to discriminate the contribution of each abrasion process to non-exhaust road particles; Thorpe and Harrison (2008) state that only brake dust particles may be identified from copper, but the wide range of proportions found in the literature do not allow a single Cu-to-brake dust particle conversion factor to be used. OM and EC arise from exhaust and non-exhaust emissions and will be discussed in more detail later on. Interestingly, no significant amounts of secondary inorganic species (ammonium, sulfate and nitrate) are found here, suggesting that this source is most plausibly freshly emitted and of local origin. Hence, it can be inferred that OM and EC are also likely of primary origin. Finally, given that road salt are exclusively made of NaCl (99% of its mass) in Paris (Le Priol et al., 2013), the absence of sodium and chloride in this factor indicates that road salting does not influence this traffic-related source on a year-basis, which gives further confidence on the abrasion nature of magnesium here. Note that the lack of mineral tracers mentioned in Sect. 2.1.2 might prevent us from identifying a road dust fraction in this factor.

As shown in Fig. 2, road traffic source mass is essentially composed of OM and EC (63 and 28%, respectively) and to a much smaller extent of Fe (6%). Both OM and EC are thought to stem from exhaust and non-exhaust processes in comparable proportions. In fact, in different European cities the contributions of exhaust and non-exhaust processes to traffic-related emissions of PM are approximately equal (Querol et al., 2004). In addition, the importance of non-exhaust particles emitted in the region of Paris has been reported in an emission inventory study (Jaecker-Voirol and
Pelt, 2000). Since carbonaceous materials represent more than 90% of our road traffic source mass, an equal contribution of both processes to OM and EC can be assumed. The low OC to EC ratio of 1.2 found in this source can be explained by the large proportion of diesel vehicles in the region of Paris, the low influence of secondary organic aerosols in this factor and the analytical method used to quantify both chemical compounds (EUSAAR_2 protocol). As a comparison, Ruellan and Cachier (2001) reported a 2.4 OC to Black Carbon ratio near a high flow road in Paris, Giugliano et al. (2005) a 1.3 OC to EC ratio at a tunnel site in Milan (Italy) and El Haddad et al. (2009) a 0.6 value for primary vehicular exhaust emissions in France. The secondary nature of road traffic related aerosols will be found in other factors (see Sect. 4.1.6 for instance).

In a few words, a factor was interpreted as a road traffic source mainly composed of primary carbonaceous and metallic particles that are likely freshly emitted and result from exhaust and non-exhaust processes.

4.1.3. Marine aerosols

A marine aerosol source was identified by the high proportion of sodium, chloride and magnesium in a single factor (79, 77 and 68%, respectively, Fig. 1). These chemical compounds are related to primary sea-salt aerosols produced by the mechanical disruption of the ocean surface (O’Dowd et al., 1997). The Cl/Na+ and Mg²⁺/Na+ ionic ratios of 0.96 and 0.13, respectively, are on the same order of magnitude as the standard sea water composition of 1.17 and 0.11, respectively (Sverdrup et al., 1942; Tang et al., 1997). The lower proportion of chloride with respect to sodium can be due to acid-base reactions between sea salt particles and sulphuric and/or nitric acids, which would lead to the evaporation of gaseous HCl in the atmosphere (Eriksson, 1959 in McInnes et al., 1994). The high sulfate to sodium ratio of 0.096 compared to 0.060 in sea water is in agreement with this assumption; the very high nitrate to sodium ratio of 1.08 likely implies another source for this latter compound. In fact, the amount of nitrate plus twice the sulfate formed should not exceed the chloride lost, on a molar basis.

To a lower extent, a small proportion of EC, Cu, K⁺, nitrate, OM and Ni is found in this marine source (10, 7, 6, 6, 5 and 4%, respectively). As mentioned above, nitrate and a fraction of OM might originate from gas-to-particle conversion of NOₓ and organic compounds, respectively, onto pre-existing sea-salt particles (Fitzgerald, 1991). EC, Cu, K⁺ and Ni are unlikely associated with natural marine processes as these chemical compounds are mainly of anthropogenic origin (with the exception of potassium). Shipping transport is a possible source of EC and Ni, because it emits large amounts of particles made of carbonaceous material and heavy metals in marine areas (Lack et al., 2009; Murphy et al., 2009), onto which nitrate could condense. However, the presence of sulfate for
both interpretations would be expected. Sea salt particles could also be enriched by anthropogenic compounds during their transport from marine regions to Paris, due to inland emissions (e.g. of EC, Ni, K, Cu) from combustion processes. Finally, uncertainties related to PMF simulations should not be excluded as well (e.g. the slope of the linear regression between observed and predicted concentrations for chloride and EC are 0.62 and 0.68, respectively, Table 1).

The resulting mass contributions to this source are 0.2±0.1, 0.2±0.1 and 0.1±0.0 µg/m³ for OM, nitrate and EC, respectively, 0.2±0.0.0 and 0.2±0.02 µg/m³ for Na⁺ and Cl⁻, respectively, and minor for the other compounds (Fig. 2). The primary sea-salt fraction of this source (Na⁺, Cl⁻ and Mg²⁺) hence accounts for ca. 37% of its mass and the likely anthropogenic fraction (EC, OM and nitrate) for the other 63%.

In conclusion, a marine aerosol source comprising sea salt particles and a large fraction of anthropogenic aerosols - that could possibly originate from combustion processes - has been identified.

4.1.4. Heavy oil combustion

A strong proportion of V, Ni and SO₄²⁻ (87, 64 and 33%, respectively) is found in a single factor. Vanadium and nickel are primarily emitted by heavy oil combustion, whose sources are industrial boilers (e.g. used in refineries), electricity generation boilers (e.g. oil power stations), large shipping ports, etc. (Jang et al., 2007; Moreno et al., 2010; Pacyna et al., 2007). It is difficult to distinguish between these sources, and "heavy oil combustion" seems to be the most suitable label for this factor. The presence of a significant proportion of sulfate is in agreement with most source apportionment studies having identified this type of source (e.g. Vallius et al., 2005; Viana et al., 2008). A part of ammonium, OM, EC, Mg²⁺ and Fe is also noticeable (17, 16, 15, 13 and 9%, respectively). Typical fuel oils naturally contain carbonaceous material, but also magnesium and iron (Miller et al., 1998), whereas ammonium is a secondary compound resulting here from the reaction with acidic sulfate to form ammonium sulfate. Larger uncertainties are associated with the other chemical elements (e.g. 25th-75th percentiles of 1-32, 2-25 and 0-12% for Cd, Pb and Cu, respectively), which will therefore not be regarded as part of this factor.

The main contributors to the mass of this heavy oil combustion source are OM, sulfate, ammonium, EC and nitrate (0.8±0.3, 0.7±0.2, 0.2±0.1, 0.2±0.1 and 0.2±0.2 µg/m³ on average, respectively). Hence, this source is at least for 45% of its mass of secondary nature, if OM and EC are assumed to be of primary origin only. This probably implies an aged and imported, instead of freshly emitted and local source. The V/Ni ratio might give insights on the sources associated with oil combustion as suggested by Pandolfi et al. (2010) and Moreno et al. (2010). Pandolfi et al. (2010)
managed to discriminate between shipping and industrial emissions in a study conducted in the vicinity of a port in southern Spain (Algeciras), and showed that the former source exhibit higher vanadium to nickel ratio (ca. 3.0, range 2.1–3.1) than the later (range 0.9–1.9 for a stainless steel plant). The same conclusions are reached by Moreno et al. (2010). In our study, the V/Ni ratio in the heavy oil combustion factor is 1.4 on average, suggesting that industrial emissions (e.g. oil power station, petrochemical complex, boilers and furnaces) are prevalent. However, the geographical origin of this factor (Sect. 4.2) indicates that shipping emissions cannot be neglected either.

4.1.5. Metal industry

As shown in Fig. 1, strong proportions of Cd, Pb and Cu are found in the same factor (47, 32 and 16% of their mass, respectively), although high interquartile ranges are observed (25th-75th percentiles of 29-55, 21-45, and 9-30%, respectively). High uncertainties are thus associated with this source, which is partly due to the difficulty for PMF to model cadmium (coefficient of determination of 0.58 for observed versus predicted concentrations, see Table 1). Cadmium and lead emission inventories have been reported for Europe by Pacyna et al. (2007) for the year 2000. The major sources of heavy metals have been taken into account, including combustion of coal/oil in industrial, residential, and commercial boilers, iron and steel production, waste incineration, gasoline combustion, etc. Although substantial uncertainties are associated with each emission category (e.g. ±20% for stationary fossil fuel combustion, ±25% for iron and steel production, etc.), they conclude that the main source of cadmium is fuel combustion to produce heat and electricity (62% by weight), whereas Pb is first emitted by gasoline combustion (51%).

The Pb/Cd ratio can be further investigated to discriminate between these types of sources. In our study, the Pb/Cd ratio is 27 on average (weight/weight ratio), which is far lower than the expected value for gasoline combustion aerosols (2,300), but closer to the mean ratio of anthropogenic European emissions (46), and to the low range of values (5-15) reported for non-ferrous metal production (Dulac et al., 1987; Pacyna, 1983). This is in agreement with the geographical origins of this source (see later in Sect. 4.2). The highest mass contributions to this source are attributed to OM, nitrate, sulfate and EC (0.03, 0.03, 0.02 and 0.01 µg/m³, respectively, Fig. 2). Very high uncertainties are associated with these concentrations that are close to, or lower than, method quantification limits. The overall contribution to PM mass is negligible (0.10 µg/m³).

To summarize, this PMF "metal industry" source presumably reflects a mesoscale background aerosol, composed of a high proportion of heavy metals that likely originate from industrial activities (non-ferrous metal production, industrial boilers, etc.).
4.1.6. Ammonium nitrate (A.N.) rich factor

The majority of nitrate and ammonium is found in a single factor (75 and 52%, respectively) while an important proportion of sulfate is also present (17%). Smaller contributions of Cd, Mn, Cl, K+ and OM are also observable (9, 9, 7, 6 and 5%, respectively). Figure 2 shows that nitrate, ammonium, sulfate and OM account for 2.0±0.2, 0.7±0.1, 0.4±0.1, 0.3±0.2 µg/m^3, respectively. This source thus represents secondary inorganic aerosols, with a stronger proportion of ammonium nitrate than ammonium sulfate, the latter being discussed in detail in the following section (Sect. 4.1.7). Ammonium nitrate stems from chemical reactions between ammonia and nitric acid, the latter compound resulting from the oxidation of NOx (NO and NO2), (Schaap et al., 2004). It therefore appears necessary to identify the major sources of NOx and ammonia to know the sources of this factor.

In Europe, atmospheric ammonia is predominantly emitted by agricultural activities -such as volatilization from animal waste and synthetic fertilizers- which have been estimated to contribute 94% of their mass emissions in 2004 for example (Pay et al., 2012). In France, emission inventories also reach the same conclusion, with agricultural activities accounting for 97% of total emissions during the same year, but also during the years 2009 and 2010 corresponding to this study (CITEPA, 2012). Other sources of ammonia such as biomass burning, fossil fuel combustion, natural emissions, etc. (Krupa, 2003; Simpson et al., 1999) will thus not be regarded as contributing to ammonia emissions here.

NOx on the other hand is produced by a variety of sources, including the combustion of fossil fuel, biomass burning, lightning, microbiological emissions from soils, etc. (Lee et al., 1997; Logan, 1983). In Europe, based on emissions of 2004 reported by Pay et al. (2012), the major anthropogenic sources of NOx are road and non-road transport (33 and 31%, respectively), followed by energy transformation and industrial combustion (17 and 11%, respectively), using the Selected Nomenclature for Air Pollution. In France, using a slightly different nomenclature (so-called SECTEN), CITEPA (2012) reported ,for the selected years 2004, 2009 and 2010, that NOx emissions primarily stem from road transport (55%), manufacturing industry (13-15% according to years), agriculture (9-10%), residential and service sectors (7-10%), and energy transformation (8-9%). The heavy metals present in this factor presumably come from some of the aforementioned activities such as road transport, manufacturing industry, energy transformation, etc. In addition, although they are not referred to in these emission inventories, the possible contribution of biomass burning in this factor should not be excluded, as suggested by the presence of potassium, chloride and OM. In that case, the unexpected absence of levoglucosan and mannosan could be explained by the imported nature
of this source (see Sect. 4.2), which could lead to the degradation of these tracers during their
transport (Hoffmann et al., 2009, see Sect. 4.1.7 for further details).

To summarize, the univocal identification of this PMF factor is rendered difficult by its
secondary nature and the diversity of the sources of its precursor gases. It can only be inferred from
emission inventories that this factor stems from a large variety of sources, likely mainly being road
and non-road transport, industrial activity, agriculture, and biomass burning.

4.1.7. Ammonium sulfate (A.S.) rich factor

This last factor is certainly the most complicated to interpret given the high proportions of
miscellaneous chemical compounds (Fig. 1), implying the contribution of a wide variety of sources. A
strong proportion of K⁺, SO₄²⁻, OM, NH₄⁺, Pb, EC and Cd (54, 46, 29, 26, 24, 17 and 17%, respectively)
and a smaller fraction of Mg²⁺, Na⁺, Ni and Fe (17, 8, 6, 5 and 5%, respectively) are observed. Mass
contributions to this source are dominated by OM, SO₄²⁻, NH₄⁺ and EC (1.6±0.3, 1.0±0.2, 0.4±0.1 and
0.3±0.2 µg/m³, respectively, Fig. 2). Based on these data, we will try to associate chemical
compounds likely to result from the same source.

Sulfate is certainly primarily bound with ammonium - (NH₄)₂SO₄ - as aerosols sampled in Paris
are neutral, and as ammonium neutralizes most of nitrate and sulfate (Bressi et al., 2013).
Ammonium sulfate aerosols come from the chemical reaction between ammonia and sulfuric acid,
the latter compound resulting from the oxidation of sulfur dioxide. Ammonia is almost exclusively
emitted by agricultural activities as mentioned in the previous section, whereas sulfur dioxide is
principally emitted by energy transformation (56%), non-road transport (17%) and industrial
combustion (13%), according to the aforementioned study of Pay et al. (2012). In France, CITEPA
(2012) states that energy transformation (54%) and manufacturing industry (30%) are the main
sources of SO₂ (in 2009), without taking into account maritime transport. These industrial activities
could explain the presence of metals such as Ni, Cd, Fe and Pb, as well as a fraction of carbonaceous
matter in this factor. Ni, Cd, Fe and Pb might also come from coal burning emissions (Junninen et al.,
2009) which could have been transported from Central/Eastern Europe to Paris (see Sect. 4.2.).

The substantial presence of potassium is presumably related to biomass burning emissions.
The absence of levoglucosan and mannosan is unexpected but could be explained by their
degradation during transport due to oxidative reactions with OH radicals (Hoffmann et al., 2009;
Kundu et al., 2010), as this source is thought to be mainly imported (Sect. 4.2). For instance, the
seasonal average levoglucosan concentration of our dataset (13.5 and 411.8 ng/m³ in summer and
winter, respectively) could be degraded in less than 2 hours in summer, and less than two days and a
half in winter (57h), following the degradation rates given in Hoffmann et al. (2009) of 7.2 ng.m⁻³.h⁻¹.
and 4.7 ng.m\(^{-3}\).h\(^{-1}\) in summer and winter, respectively. The aged property of biomass burning particles contributing to this source is in line with the absence of chloride in this factor, which could be due to the chemical conversion of KCl particles to K\(_2\)SO\(_4\) (or to a lesser extent KNO\(_3\)), after having undergone similar heterogeneous reactions mentioned for marine aerosol particles in Sect. 4.1.3 (Li, 2003). The aforementioned literature study reported that more than 90% of KCl particles coming from biomass burning were converted to potassium sulfate or nitrate after only 24 minutes in southern Africa. Nevertheless, given the geographical origins of this source, we do not exclude the potential contribution of potassium industries (e.g. fertilizer industries) in this source as well, which could produce potassium sulfate and potassium nitrate compounds.

Finally, because of the high proportions of sulfate and ammonium, this source is essentially secondary in nature. Therefore, OM can here be assumed to principally refer to secondary organic aerosols (SOA), as it is supported by the high OM to EC ratio of 5.6. The complexity and the multiplicity of the chemical processes leading to the formation of SOA do not allow us to determine its precise sources. Beekmann et al. (2012) reported that SOA could be of mixed anthropogenic (fossil fuel) and biogenic origins in the region of Paris (see also Crippa et al., 2013a, 2013b on this subject).

To summarize, this factor is primarily made of secondary aerosols, which stem from a variety of sources including agriculture, industrial activities, non-road transport and biomass burning, to name a few.

### 4.1.8. Comparison with other source apportionment (SA) studies

A comparison with source apportionment studies conducted throughout Europe, based on the review of Viana et al. (2008), will now be presented here. As reported in our work, most studies identify a vehicular (with carbon, Fe, Cu), a sea salt (Na\(^+\), Mg\(^{2+}\), Cl), a mixed industrial/fuel-oil combustion (V, Ni, SO\(_4^{2-}\)) and a secondary aerosol (SO\(_4^{2-}\), NO\(_3^-\), NH\(_4^+\)) source (Viana et al., 2008) - although the secondary aerosol source has been apportioned in two distinct factors in our case-. Biomass burning sources have been reported worldwide in more recent SA studies (Gu et al., 2011; Larsen et al., 2012; Thurston et al., 2011). The metal industry source found in our work is less common, which could be related to its very low contribution to PM\(_{2.5}\) mass. Finally, it can be noted that a crustal or mineral dust source has not been identified in our work, contrary to what is ordinarily reported elsewhere. This type of source is generally characterised by high contents of aluminium, silicon, calcium and iron. Calcium in particular, has already been used to trace mineral dusts in the city of Paris (Guinot et al., 2007); discarding this element from PMF simulations thus appears problematic. The difficulty encountered by PMF to model this compound is certainly related to a local source contamination of calcium (renovation of building façades) near the sampling site.
(Bressi et al., 2013). Nonetheless, it has been estimated to only contribute 3% of PM$_{2.5}$ mass on average during this one-year project (Bressi et al., 2013), and hence does not represent a major source of fine aerosols in Paris. Further research should however be conducted to better characterize mineral dust contribution to fine aerosols in the region of Paris.

### 4.2 Source geographical origins

The geographical origins of each PM$_{2.5}$ source determined by PSCF and CPF are reported in Fig. 5. This figure aims at providing insights on source localization but does not claim to be accurate at the pixel or the degree level. PSCF and CPF results will first be compared and similar results only will be further interpreted for each source. Note that the values of the probabilities given by PSCF and CPF are not directly comparable as weighting functions and smoothing procedures differ from one methodology to the other.

Regarding the A.S. rich factor, its high probability to come from geographical regions located northeast of Paris is highlighted by both methodologies. In fact, the probabilities for this factor to exceed the 75th percentile in CPF are clearly higher for air masses coming from NNE to ENE than from other directions (40 versus 12% on average, respectively). Similarly, a hot spot is observable in this NE direction with PSCF, with probabilities to exceed the aforementioned criterion being higher than 80% from northeast of France to Benelux and southwest Germany. Interestingly, these geographical regions are amongst the major emitters of sulfur dioxide in Europe (Pay et al., 2012), which is - with ammonia - a precursor of ammonium sulfate. High probabilities (ca. 55%) are however observed with PSCF for almost all France and southeast England contrarily to CPF. Given the long lifetime compounds present in this factor, it is possible that its high contributions result from anticyclonic conditions, involving stagnating air masses that could come from any regions around Paris. In addition, because such results are not observed in CPF, bias related to the binomial smoothing used in PSCF may not be excluded.

The A.N. rich factor is likely coming from regions located NNE of Paris. CPF values are significantly higher for NNE and NE than for the other directions (49 versus 11% on average, respectively). Similarly, PSCF probabilities are the highest in this direction (generally above 60%), against probabilities generally below 20% in the other directions. This is also in line with the European map depicted by Pay et al (2012) for total nitrate (HNO$_3$+NO$_3^-$) concentrations, which appear higher in this geographical area.

The heavy oil combustion source presumably comes from north of France although a local influence is not excluded. CPF suggests this source originates from NNW to NNE directions (mean of 42% against 12% for the other directions), and PSCF shows its highest probabilities in the NNW
direction (higher than 80% in northern France and the English Channel). Northern France is a highly industrialized region (e.g. the so-called Nord-Pas-de-Calais region, located near Belgium and the English Channel is the fourth industrialized French region), comprising some of the largest harbours of the country (e.g. Le Havre, Dunkirk, Calais, etc.). These activities are in line with the industrial feature of this source mentioned in Sect. 4.1.4, and will be further discussed in Sect. 4.3.1. On the other hand, the high PSCF values observed in the English Channel suggest that maritime transport clearly affects the contribution of this factor. The low V to Ni ratio reported in our study (Sect. 4.1.4) might thus not be the best proxy to distinguish between industrial and maritime heavy oil combustion. Finally, influences of local sources cannot be excluded as well, given the high number of industrial activities in the region of Paris. As PSCF and CPF only focus on the highest contributions of sources, local emissions could be omitted by both methodologies, because they would constantly increase the concentrations of this factor without however triggering pollution events.

The road traffic source is primarily of local origin. Nevertheless, CPF and PCSF also indicate the influence of central France, which is unlikely and could be related to an artefact discussed below. High probabilities are observed with CPF for S to SSW (42% on average) and E directions (33%) compared with the other air mass origins (16% on average). PSCF probabilities are also higher for S to SW directions (above 80%), but contrarily to CPF the eastern direction is not highlighted here. Instead, moderate probabilities are rather uniformly distributed all around the region of Paris (ranging from 50 to 70%) that could be related to a local origin for this source. The eastern influence shown by CPF will not be regarded as meaningful given its divergence with PSCF values. Differences between both methodologies could also be related to the local feature of this source. In addition, it is very unlikely that primary particles with road transport characteristics measured in Paris were imported from central France given the high number of vehicles present in the former megacity. Furthermore, a comparison between our EC concentrations (45% of EC is found in this factor; Fig. 4) and those measured at a rural site located 60 km southward does not show any correlation ($r^2=0.03$, slope=0.27, n=335, Bressi et al., 2013). Instead, air masses originating from south of Paris could be related to low boundary layer heights that would enhance local road traffic aerosol concentrations. We attempted to quantify this phenomenon and found that 40% of southward back-trajectories (n=123) displays BLH below 600m (corresponding to 26th percentile of BLH values measured during the campaign, see Bressi et al., 2013 for further information on BLH measurements). Other meteorological parameters (e.g. atmospheric pressure) should be taken into account to fully understand the characteristics of these southward air masses.

The biomass burning source is likely both locally emitted and imported from south of Paris. CPF shows fairly homogeneous probabilities from WSW to SSE (ranging from 9 to 20%) and higher values from S to SW directions (22-30%). Note that the absolute values of CPF probabilities are the
lowest for this source, signifying that its geographical origins are less marked. PSCF also shows relatively homogeneous probabilities all around Paris (ca. 60%) with however significantly higher values south to southwest of this megacity (higher than 80%). Two assumptions could explain such results. First, this BB source could be locally emitted as suggested by relatively isentropic results for both approaches with the exception of S to SW directions. In that case, the hot spot highlighted S to SW of Paris would be due to the same feature described previously for the road traffic source (specific meteorological conditions related to southward air masses such as low BLH). This assumption is in line with previous studies stating that BB aerosols are locally emitted in the region of Paris (Favez et al., 2009; Sciare et al., 2011). Second, a proportion of this source could actually be imported from south of Paris. This is supported by a comparison conducted between atmospheric concentrations of levoglucosan measured at our urban site and at the aforementioned rural site (located 60 km southward our sampling site). Very good correlations are observed between both datasets on the entire duration of the project ($r^2=0.84$, slope=0.84, $n=331$; Beekmann et al., 2012), suggesting that a noticeable proportion of biomass burning aerosols could be imported from south of Paris. Further research should be conducted on biomass burning sources in Paris to fully explain this surprisingly influence of southward geographical areas.

Marine aerosols are mostly coming from the Atlantic Ocean and to a lower extent from the North Sea, although anthropogenic contributions from inland emissions are noticeable. CPF exhibits high probabilities from SSW to W (38% on average), intermediates from NNW to N (24%) and low values from NE to S (4%). PSCF results are in agreement showing high probabilities from the Atlantic Ocean to western France (above 80%), intermediates in the North Sea (ca. 60%) and low values from NE to S (typically below 20%). Interestingly, the hot spot highlighted in the Atlantic Ocean corresponds to a geographical area where the biggest salt ponds of the country lie (e.g. Guérande, Noirmoutier, etc.). As suggested by high PSCF probabilities in western France, the anthropogenic fraction of this source most plausibly stem from inland anthropogenic emissions that could be (internally or externally) mixed with sea-salt particles, or could affect their chemical composition.

Lastly, the metal industry source seems to reflect a regional haze, although the influence of areas located northeast of Paris is underlined. CPF displays higher probabilities from NNW to NE than for the other directions (31 versus 14% on average, respectively). PSCF also points to high probabilities in the NE direction with values higher than 80% in northeastern France. Contrarily to CPF, Paris and Central France also exhibit high PSCF values (above 80%). Discrepancies observed between CPF and PSCF results might reflect the presumable regional background properties of this factor, characterizing a mesoscale haze of metal industry emissions. They could also be due to the very low atmospheric concentrations of this source (representing 1% of PM$_{2.5}$ mass on average) leading to large uncertainties.
4.3 Source contribution

4.3.1 Annual average

The annual average contribution of the seven sources to PM$_{2.5}$ mass is reported in Fig. 6. The two predominant factors are the ammonium sulfate and the ammonium nitrate rich factors accounting for ca. half of PM$_{2.5}$ mass (51%). Heavy oil combustion, road traffic and biomass burning also contribute significantly to fine aerosol mass (17, 14 and 12%, respectively), whereas marine aerosols and metal industry sources have a far lower contribution (6 and 1%, respectively). These contributions were compared with source apportionment studies (see Fig. 7 and Table S7), chosen according to the following criteria:

i) SA is performed on PM mass (PM$_{2.5}$ in most of the cases);

ii) each SA study is representative of one year minimum;

iii) when possible, SA studies have been chosen according to their presumable geographical origins (e.g. continental Europe for A.S. and A.N. rich sources);

iv) similar source categories (i.e. factor identifications) are reported.

The prevalence of an ammonium sulfate rich factor in European SA studies is widely reported (Viana et al., 2008). It is for instance illustrated in a study conducted by Mooibroek et al. (2011) on PM$_{2.5}$ sampled during one year (2007-2008), at five sites in the Netherlands (one urban, one kerbside and three rural sites). An A.S. rich factor was identified by EPA PMF3.0, and contributes from 20 to 30% of PM$_{2.5}$ mass, with a PM$_{2.5}$ annual average concentration ranging from 12.5 to 17.5 µg/m$^3$ (i.e. on the same order of magnitude as our mean PM$_{2.5}$ level of 14.7 µg/m$^3$). The absolute contributions of this source are 4.4 and 4.9 µg/m$^3$ at two rural sites (Vredepeel and Cabauw sites, respectively, values calculated from concentrations given in Weijs et al., 2011), which is higher than the contribution of 3.9 µg/m$^3$ reported in our study. Interesting results are also reported in a SA study conducted at an urban background site in Copenhagen (Denmark) by Andersen et al. (2007). The comparison with our results is much more limited here, as this study was conducted on PM$_{10}$ for a 6-year period (1999-2004), and as a hybrid receptor model combining CMB and PMF approaches (COPREM model) was used. Nevertheless, most of the compounds found in our A.S. rich factor (ammonium sulfate and SOA) are assumed to be in the fine mode, and the sources identified with COPREM are very similar to ours. The resulting contribution of their A.S. rich factor is 3.5 µg/m$^3$, which is again close to the value of 3.9 µg/m$^3$ reported in our study. The contribution of the ammonium sulfate rich factor to PM$_{2.5}$ mass found in our work is hence in the range of values reported in other European SA studies, and the presumable influence of countries located northeast
of France appears relevant, regarding the high contributions of this A.S. factor in this geographical area.

The A.N. rich factor is also a predominant contributor to PM$_{2.5}$ in European SA studies (Viana et al., 2008). Mooibroek et al. (2011) report a very high contribution of this source in the Netherlands, ranging from 5.6 to 7.7 µg/m$^3$ according to sites, compared to a contribution of 3.5 µg/m$^3$ in our study. Andersen et al. (2007) report a contribution of 3.3 µg/m$^3$ on average in Copenhagen, which is in line with our value.

Considering both A.S. and A.N. factors as a single source would allow more comparisons with other SA studies. Combining both factors is acceptable as they mainly stem from common sources of precursor gases and are imported from the same geographical area in our study (see Sect. 4.1.6, 4.1.7 and 4.2). A secondary aerosol source was identified by Quass et al. (2004) with PMF in Duisburg (Germany), based on one-year measurements of PM$_{2.5}$ (2003-2004). Its annual contribution to PM$_{2.5}$ mass is higher than the value reported in our study (57% versus 51%, respectively), so are its absolute concentrations (13.0 µg/m$^3$ versus 7.4 µg/m$^3$, respectively). On the other hand, Vallius et al. (2005) reported at an urban site in Amsterdam (study conducted from November 1998 to June 1999) a contribution of a PM$_{2.5}$ secondary aerosol source of 6.8 µg/m$^3$ that is comparable to ours. Finally the summed contribution of A.S. and A.N. factors reaches 6.8 µg/m$^3$ in the study of Andersen et al. (2007), and ranges from 8.6 to 12.6 µg/m$^3$ according to sites in Mooibroek et al. (2011). Therefore, the predominant contribution of secondary aerosol sources to fine aerosol mass estimated in Paris is in line with most SA studies conducted in Europe. The high proportions of such sources in countries located northeast of France support the idea that this region significantly affects secondary aerosol concentration levels measured in Paris.

Regarding the heavy oil combustion source, its important contribution to PM$_{2.5}$ mass of 17% (2.4 µg/m$^3$) is relevant with its imported feature from northern France (Cf. Sect. 4.2), where there is a high density of industries and strong emissions from maritime transport in the English Channel. The influence of industrial activities on aerosol levels in this geographical area has been reported by Alleman et al. (2010) in a study conducted in the highly industrialised harbour of Dunkirk, which is one of the largest French commercial harbours (freight transport: 58 million tons in 2008). These authors applied various SA techniques including PMF to identify and apportion the sources of PM$_{10}$ sampled at an urban background site during almost two years (June 2003 – March 2005). A source labelled "petrochemistry" was identified because of its high contents of vanadium and nickel, and includes emissions from fuel refineries, fossil fuel power plants but also boat transport. Note that only small proportions of V and Ni are found in the coarse fraction at our site (Poulakis et al., 2012) thus making the comparison between our heavy oil combustion factor and this petrochemistry factor pertinent. This source shows a mean contribution of 2.3 µg/m$^3$ in Dunkirk (value calculated from a
mean PM$_{10}$ concentration of 25 µg/m$^3$ estimated from www.atmo-npdc.fr/home.htm), which is very close to the contribution of our heavy oil combustion source of 2.4 µg/m$^3$ in Paris. In addition, if all industry-related sources identified in Dunkirk (such as metallurgical sintering plant, metallurgical coke plant, etc.) are taken into account, their average contribution reaches 9.3 µg/m$^3$, which represents 37% of PM$_{10}$ mass. The different types of plants located in Dunkirk may all contribute to our oil combustion factor by increasing its carbonaceous and secondary inorganic content, but may not be distinguished as specific tracers analysed in the Dunkirk study were not quantified in Paris (e.g. Rb, Cs, Bi, Th, etc.). The levels measured for this oil combustion source in Paris (2.4 µg/m$^3$) are comparable to what has been reported in Amsterdam and Copenhagen, which cities are located in the vicinity of petrochemical activities and maritime transport. In the former city, a mean contribution of 2.2 µg/m$^3$ was estimated by Vallius et al. (2005) whereas in the latter, this contribution reaches 3.5 µg/m$^3$ (Andersen et al., 2007). Further research investigating the contribution of heavy oil combustion sources to fine aerosols should be conducted in the region of Paris, given the surprisingly high levels found in our study.

The road traffic source contributes 14% of PM$_{2.5}$ mass which represents 2.1 µg/m$^3$. This contribution is noticeable, but was expected to be more important given the high density of vehicles in Paris. It is for instance markedly lower than the 3.8 µg/m$^3$ estimated by PMF for PM$_{1.0}$ though, at an urban background site in Zurich (Switzerland) by Minguillón et al. (2012) from a winter and summer campaign. It is also significantly lower than the 7.8 µg/m$^3$ estimated by CMB for PM$_{2.5}$, at an urban background site in Milan (Italy) by Perrone et al. (2012) for a 3-year period (2006-2009). As mentioned in Sect. 4.1.2, the absence of a road dust fraction might partly explain the relatively low contribution of our road traffic source. Nevertheless the level estimated in our study is comparable with other highly populated urban areas in the world. For instance, at an urban site in Toronto (Canada, ca. 5.6 million inhabitants in the metropolitan area), from PM$_{2.5}$ sampled during one year (2000-2001) and apportioned by PMF, Lee et al. (2003) estimated a contribution of 2.3 µg/m$^3$ to a road transport source. Their resulting contribution to PM$_{2.5}$ mass is slightly higher than ours (18 versus 14%, respectively). Similar levels were also reported in Seattle (U.S.A., ca. 3.5 million inhabitants in the metropolitan area) by Maykut et al. (2003) from multiannual measurements (1996-1999) conducted at an urban site. The PMF and UNMIX approaches lead to a contribution of 2.0 and 2.5 µg/m$^3$ for this source, i.e. 22 and 28% of PM$_{2.5}$ mass respectively.

The biomass burning source is the last considerable contributor to PM$_{2.5}$ in Paris (12%, 1.8 µg/m$^3$). To the best of our knowledge, the contribution of such source to particulate matter mass in Paris is estimated for the first time on the annual scale in our study. However, two studies attempted to estimate it from campaigns of few days or weeks (Favez et al., 2009; Sciare et al., 2011) that do not allow suitable comparisons to be performed on the annual scale. (Note that comparisons will
nonetheless be conducted in the next section on the seasonal scale only.) In Europe as well, few studies report the contribution of BB to PM mass. Andersen et al. (2007) estimated a very large contribution of 7.3 µg/m³ for this BB source, representing 15% of their PM₁₀ samples in Copenhagen. Note that biomass burning sources are presumably entirely found in the fine fraction (e.g. Karanasiou et al., 2009), making the previous comparison relevant. Perrone et al. (2012) also report a substantial contribution of 7.1 µg/m³ representing 16% of their PM₂.₅ samples in Milan (Italy). Finally, Karanasiou et al. (2009) estimated by PMF this contribution to be 0.8 µg/m³ in Athens (Greece), representing 15% of their PM₂.₅ samples.

The contribution of the marine aerosol source is fairly low (6%, 0.8 µg/m³), likely because its mass size distribution is mainly located in the coarse mode. It is comparable to values reported in the Netherlands (e.g. 0.8 µg/m³ at a rural site, Mooibroek et al., 2011), in Finland (0.9 µg/m³ at an urban site of Helsinki, Vallius et al., 2003) or in Greece (1.1 µg/m³ at an urban site in Athens, Karanasiou et al., 2009). This comparison however presents some limitations since the distance from the coast is substantially higher for our sampling site (ca. 300-500 km depending on the directions) compared to the aforementioned sites (around or below 100km). Finally, metal industry contributes to very low levels of PM₂.₅ in our study (1%, 0.1 µg/m³) that certainly reflects a haze due to large scale pollutions, as it is reported in Poulakis et al. (2012).

4.3.2 Seasonal variability

The seasonal variability of the sources of PM₂.₅ is reported in Table 3, Figs. 8, 9, S3 and S4. As expected, each source presents singular patterns due to variations of source emissions, thermodynamic conditions or meteorological parameters in general according to seasons. First, the A.S. rich factor exhibits high contributions all along the year, ranging from 2.7 to 5.6 µg/m³ i.e. from 21 to 39% of PM₂.₅ mass on average according to seasons (Table 3). Its highest contribution in absolute concentrations is observed during winter and is more than 40% higher than its annual average value. This can be explained by the numerous pollution events occurring during January and February (e.g. 26 January 2010: 31.9 µg/m³ or 9 February 2010: 23.5 µg/m³; Fig. 3) that are related to specific meteorological conditions (anticyclonic conditions, eastward imported air masses and low boundary layer heights; Bressi et al., 2013). Note that in terms of relative proportion to PM₂.₅ mass, the highest contribution of this A.S. rich factor is on the other hand observed during summer (39%, Table 3, Fig. 9), which can mathematically be explained by its continuously high absolute concentrations along the year whereas PM₂.₅ levels are notably lower during summer. Photochemistry could also play a role in the high contribution observed for this secondary source during summer.
The A.N. rich factor shows a very clear seasonal pattern with significantly higher concentrations during winter and spring than autumn and summer seasons (6.8, 5.1, 1.9 and 0.5 µg/m³, respectively). This was expected due to thermodynamic conditions (especially low temperatures) observed during winter and spring in Paris (see Bressi et al., 2013), thus favouring the condensation of ammonium nitrate (Clegg et al., 1998). This source contributes approximately one third of PM_{2.5} mass on average during winter and spring (32 and 34%, respectively), compared to a lower contribution during autumn and especially summer (15 and 5%, respectively).

The heavy oil combustion source presents fairly stable seasonal concentrations ranging from 2.0 to 3.0 µg/m³ (Table 3). Higher concentrations are however observed during spring and summer (Fig. S4), which could have different explanations. First, this could be an artefact due to the high number of Ni values replaced by the median of its concentrations during May and June months (61%, n=61), as it is suggested by the increased baseline during these months on Fig. 3. However, the methodology detailed in Sects. 2.2.2 and S2 was implemented to lower the influence of median-replaced concentrations and this artefact should be minimal. Second, it could be due to enhanced marine vessel activities during spring and summer, in addition to non-dispersive meteorological conditions enhancing the influence of industrial activities. Such phenomenon has been reported by Mooibroek et al. (2011) who also found a clear seasonal pattern for their oil combustion source in the Netherlands, exhibiting an increased contribution during summer (summer median more than twice higher than the annual one). They partly explain this pattern by the significant height of the flue gas stacks of petrochemical industry: during winter, the flue gases and particles can be exhausted above the boundary layer height whereas during summer they are exhausted below, which results on greater impacts on ground level atmospheric concentrations during the latter season. This interpretation appears suitable to our heavy oil combustion source as well, given the presumably high contributions of industrial activities. Third, during summer photochemistry could favour the formation of secondary compounds such as ammonium sulfate (representing 39% of this factor’s mass).

The road traffic source exhibits rather stable concentrations all along the year (annual average of 2.1±2.1 µg/m³), with however a smaller contribution during winter (1.3±1.4 µg.m⁻³). An overestimation of EC content in the biomass burning factor and an underestimation in the road traffic one could explain this observation. However, fairly good correlations between fossil fuel black carbon and the road traffic source are observed throughout the year (r=0.50, n=327, Table S6 and Bressi, 2012), suggesting that this pattern is real although it is not fully explained by the authors. From a mathematical standpoint, this could be explained by the absence of clear pollution events for this (primary) road traffic source during winter, in contrary to what is observed during the other seasons. In fact, 8, 1, 5 and 8 days show contributions higher than 6 µg/m³ during autumn, winter,
spring and summer, respectively (Fig. 3). These pollution events are mostly driven by low boundary
layer height conditions instead of increased emissions from road traffic.

As expected, the biomass burning source exhibits significantly higher concentrations during
autumn and winter than during spring and summer seasons. The maximum contribution of 4.7±3.7
µg/m$^3$ is observed during winter and represents 22% of PM$_{2.5}$ levels on average during this season
(Table 3). A day by day calculation leads to an averaged relative contribution of the BB source to
PM$_{2.5}$ mass of 24±14%. This estimation is on the same order of magnitude as, but slightly higher than,
the previous estimations made for PM$_{2.5}$ at urban sites of Paris during shorter time periods in winter.
Based on light absorption measurements, the averaged contribution of a biomass burning source to
PM$_{2.5}$ mass has been estimated to be 20±10% in Favez et al. (2009), and 15±11% in Sciare et al.
(2011), after 40- and 10-days measurements, respectively. A similar result has been reported by
Perrone et al. (2012) in Milan (Italy), where a biomass burning source has shown to contribute to
25% of PM$_{2.5}$ on average during winter, with however an absolute concentration more than 3 times
higher than in Paris (14.6±6.5 against 4.7±3.7 µg/m$^3$, respectively).

The seasonal variations of marine aerosols and metal industry sources are illustrated in more
detail in Fig. S4. Marine aerosols display higher contributions during autumn than the rest of the year
(1.3 against 0.6-0.7 µg/m$^3$, respectively) due to higher occurrences of air masses originating from the
Atlantic Ocean or the North Sea. The metal industry source does not show any seasonal pattern, with
seasonal averaged concentrations ranging from 0.10 to 0.12 µg/m$^3$, which is in line with its regional
background characteristic.

5. Conclusions and perspectives

Based on one-year PM$_{2.5}$ sampling at an urban site located in Paris (France), and on the use of
statistical tools (EPA PMF3.0, CPF, PSCF), this paper allowed i) the identification of seven PMF factors
that were related to real aerosol sources, ii) the identification of the geographical origins of each
factor, and iii) the apportionment of each factor to PM$_{2.5}$ mass discussed on yearly and seasonal
bases. The main results can be summarized as follow:

(1) the ammonium sulfate and ammonium nitrate rich factors contribute ca. half of PM$_{2.5}$ mass
on average during the whole study (27 and 24%, or 3.9 and 3.5 µg/m$^3$, respectively). These
factors are made of secondary organic and inorganic aerosols, originating from various
sources (including road traffic, industry, agriculture and biomass burning) that are difficult to
distinguish. Both factors have primarily undergone north-eastward mid- or long-range
transport.
(2) a heavy oil combustion source exhibits a noticeable contribution to PM$_{2.5}$ mass (17%, 2.4 µg/m$^3$ on average). It has been identified through a strong signature of specific tracers (V and Ni), and mainly stems from industrial activities (e.g. oil power station, petrochemical complex, etc.) and shipping emissions. It likely originates from northern France and the English Channel where a high density of industries, large harbours and shipping lies, although a local influence may not be excluded.

(3) a road traffic source accounts for 14% of PM$_{2.5}$ mass on average (2.1 µg/m$^3$), which is relatively low regarding the expected high contribution of the numerous vehicles of Paris. This source includes exhaust and non-exhaust particles that are almost solely composed of carbonaceous materials. It is a local source which contributions could be enhanced by the meteorological conditions associated with southward air masses (e.g. low BLHs).

(4) a biomass burning source contributes 12% of PM$_{2.5}$ mass on average (1.8 µg/m$^3$). It includes both primary and secondary aerosols that mainly come from wood combustion, even though agricultural and garden waste burning contributions may not be excluded. It is likely both locally emitted and imported from southward of Paris. The two last sources named marine aerosols and metal industry only contribute 6 and 1% of PM$_{2.5}$ mass on average, respectively. Based on these source apportionment results, more than half of PM$_{2.5}$ levels in Paris is therefore associated with (mid-) long-range transported pollution of secondary organic and inorganic aerosols. Further work is still required to better characterise their sources. For instance, gas precursors including SO$_2$, NO$_x$, NH$_3$ and Volatile Organic Compounds could be simultaneously measured with aerosol chemical components, before being investigated by PMF. Additional aerosol chemical characteristics such as the isotopic composition of individual elements (e.g. S, N, C, O) would also be valuable for PMF interpretation. The influence of (mid-) long-range transport in Paris suggests that abatement policies implemented at the local, or regional level, may not be sufficient to notably reduce PM$_{2.5}$ concentrations in this city. Instead, a collaborative work should be conducted between surrounding regions or even countries. Similar conclusions may presumably be drawn for studies conducted in the vicinity of France aiming at determining the geographical origins of PM$_{2.5}$, given that French emissions of gaseous precursors of secondary aerosols (NH$_3$, NO$_x$, SO$_2$ and VOCs) are estimated to be of the same order of magnitude as, or higher than, those of neighbouring countries (e.g. Visser et al., 2001). The investigation of forward trajectories from our study would help evaluating the influence of Paris emissions on surrounding geographical areas. It would likely support the idea that a significant part of PM$_{2.5}$ pollution in Europe is transboundary, hence requiring coordinated abatement policies amongst E.U. countries.
Disclaimer. This document has been subjected to Airparif and Ineris reviews and approved for publication. Nevertheless, the conclusions drawn do not necessarily reflect the views of these organizations.

Acknowledgements. The French Environment and Energy Management Agency (ADEME), the French Alternative Energies and Atomic Energy Commission (CEA), the region of Paris (Ile-de-France), the city of Paris, the National Centre for Scientific Research (CNRS) and the University of Versailles Saint-Quentin-en-Yvelines are acknowledged for their support. M. Reynaud and M. Artufel are thanked for their technical and analytical help. The authors are grateful to the anonymous referees and the editor for their valuable comments.

References


Draxler, R. R. and Hess, G. D.: Description of the HYSPLIT_4 modeling system, Air Resources Laboratory, Silver Spring, Maryland, 1997.


Guinot, B., Cachier, H., Oikonomou, K. and others: Geochemical perspectives from a new aerosol chemical mass closure, Atmospheric Chemistry and Physics, 7(6), 1670, 2007.


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


Table 1. Statistics describing measured versus modelled concentrations for each chemical species and for PM$_{2.5}$ mass.

Legend: Interc.: Intercept ($\mu$g/m$^3$), SE: Standard Error.

<table>
<thead>
<tr>
<th></th>
<th>$r^2$</th>
<th>Slope</th>
<th>Slope SE</th>
<th>Interc.</th>
<th>Interc. SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM</td>
<td>0.97</td>
<td>1.01</td>
<td>0.01</td>
<td>-2.5E-01</td>
<td>1.8E-01</td>
</tr>
<tr>
<td>OM</td>
<td>0.85</td>
<td>0.87</td>
<td>0.02</td>
<td>4.9E-01</td>
<td>1.3E-01</td>
</tr>
<tr>
<td>EC</td>
<td>0.56</td>
<td>0.68</td>
<td>0.03</td>
<td>3.6E-01</td>
<td>5.3E-02</td>
</tr>
<tr>
<td>NO3</td>
<td>0.99</td>
<td>1.00</td>
<td>0.01</td>
<td>1.8E-02</td>
<td>2.6E-02</td>
</tr>
<tr>
<td>SO4</td>
<td>0.89</td>
<td>0.89</td>
<td>0.02</td>
<td>2.0E-01</td>
<td>4.4E-02</td>
</tr>
<tr>
<td>NH4</td>
<td>0.95</td>
<td>0.95</td>
<td>0.01</td>
<td>6.5E-02</td>
<td>2.4E-02</td>
</tr>
<tr>
<td>Na</td>
<td>0.82</td>
<td>0.81</td>
<td>0.02</td>
<td>2.5E-02</td>
<td>5.2E-03</td>
</tr>
<tr>
<td>Cl</td>
<td>0.76</td>
<td>0.62</td>
<td>0.02</td>
<td>6.0E-02</td>
<td>5.6E-03</td>
</tr>
<tr>
<td>Mg</td>
<td>0.79</td>
<td>0.82</td>
<td>0.02</td>
<td>3.0E-03</td>
<td>7.9E-04</td>
</tr>
<tr>
<td>K</td>
<td>0.91</td>
<td>0.86</td>
<td>0.01</td>
<td>1.3E-02</td>
<td>2.3E-03</td>
</tr>
<tr>
<td>Lev</td>
<td>0.98</td>
<td>0.91</td>
<td>0.01</td>
<td>8.3E-03</td>
<td>2.1E-03</td>
</tr>
<tr>
<td>Man</td>
<td>0.97</td>
<td>0.96</td>
<td>0.01</td>
<td>2.5E-04</td>
<td>2.0E-05</td>
</tr>
<tr>
<td>V</td>
<td>0.89</td>
<td>0.87</td>
<td>0.02</td>
<td>1.5E-04</td>
<td>2.0E-05</td>
</tr>
<tr>
<td>Ni</td>
<td>0.47</td>
<td>0.42</td>
<td>0.02</td>
<td>6.6E-04</td>
<td>4.0E-05</td>
</tr>
<tr>
<td>Fe</td>
<td>0.84</td>
<td>0.80</td>
<td>0.02</td>
<td>2.4E-02</td>
<td>3.8E-03</td>
</tr>
<tr>
<td>Mn</td>
<td>0.86</td>
<td>0.68</td>
<td>0.01</td>
<td>9.7E-04</td>
<td>9.0E-05</td>
</tr>
<tr>
<td>Cu</td>
<td>0.71</td>
<td>0.72</td>
<td>0.02</td>
<td>1.2E-03</td>
<td>2.0E-04</td>
</tr>
<tr>
<td>Cd</td>
<td>0.58</td>
<td>0.85</td>
<td>0.04</td>
<td>4.0E-05</td>
<td>1.0E-05</td>
</tr>
<tr>
<td>Pb</td>
<td>0.73</td>
<td>0.76</td>
<td>0.03</td>
<td>1.2E-03</td>
<td>1.8E-04</td>
</tr>
</tbody>
</table>
Table 2. Seasonal variability of statistics describing the ability of PMF to model PM$_{2.5}$.

Legend: RMSE: Root Mean Square Error, MAPE: Mean Absolute Percentage Error.

Note: $r^2$ was determined by plotting the modelled (sum of the contributions of the sources) versus the measured PM$_{2.5}$ mass. Calendar seasons were used (see Table 3).

<table>
<thead>
<tr>
<th></th>
<th>Autumn</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of days</td>
<td>85</td>
<td>82</td>
<td>84</td>
<td>86</td>
<td>337</td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.97</td>
<td>0.98</td>
<td>0.95</td>
<td>0.89</td>
<td>0.97</td>
</tr>
<tr>
<td>RMSE $\mu g/m^3$</td>
<td>1.4</td>
<td>1.9</td>
<td>2.0</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>MAPE %</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 3. Seasonal variations of the absolute concentrations of the sources (µg/m$^3$), and their relative proportions to the averaged PM$_{2.5}$ mass (%).


**Note:** Calendar seasons were used i.e. Autumn: from 23 September to 21 December 2009; Winter from 22 December 2009 to 20 March 2010; Spring: from 21 March to 21 June 2010 and Summer: from 11 September to 22 September 2009 plus from 22 June to 10 September 2010. Annual: from 11 September 2009 to 10 September 2010. The meaningless negative contribution of the biomass burning source (marked with an asterisk *) during summer is due to analytical problems with levoglucosan during September 2009.

<table>
<thead>
<tr>
<th>Source</th>
<th>Autumn</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.S. rich factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of days</td>
<td>85</td>
<td>82</td>
<td>84</td>
<td>86</td>
<td>337</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.7</td>
<td>5.6</td>
<td>3.5</td>
<td>3.8</td>
<td>3.9</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.5</td>
<td>6.2</td>
<td>2.8</td>
<td>2.7</td>
<td>4.0</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>21</td>
<td>26</td>
<td>24</td>
<td>39</td>
</tr>
<tr>
<td>A.N. rich factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>1.9</td>
<td>6.8</td>
<td>5.1</td>
<td>0.5</td>
<td>3.5</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>3.4</td>
<td>6.5</td>
<td>6.8</td>
<td>1.7</td>
<td>5.6</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>15</td>
<td>32</td>
<td>34</td>
<td>5</td>
</tr>
<tr>
<td>Heavy oil combustion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.0</td>
<td>2.1</td>
<td>3.0</td>
<td>2.6</td>
<td>2.4</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.1</td>
<td>1.7</td>
<td>2.0</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>15</td>
<td>10</td>
<td>20</td>
<td>27</td>
</tr>
<tr>
<td>Road traffic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.5</td>
<td>1.3</td>
<td>2.3</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.3</td>
<td>1.4</td>
<td>2.2</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>19</td>
<td>6</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>Biomass burning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>2.4</td>
<td>4.7</td>
<td>0.2</td>
<td>-0.1*</td>
<td>1.8</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>3.0</td>
<td>3.7</td>
<td>0.4</td>
<td>0.2</td>
<td>3.0</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>18</td>
<td>22</td>
<td>1</td>
<td>-1*</td>
</tr>
<tr>
<td>Marine aerosols</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>1.3</td>
<td>0.7</td>
<td>0.6</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>1.4</td>
<td>1.0</td>
<td>0.6</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Metal industry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>µg/m$^3$ std</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>%</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Fig. 1. Relative contribution (%) of each chemical species in a given PMF factor. Boxplots are constructed with the 5th, 25th, 50th, 75th and 95th percentiles of bootstrap runs (n=100).

Legend: Lev: Levoglucosan, Man: Mannosan.
Ammonium sulfate rich source

Ammonium nitrate rich source

Heavy oil combustion

Road traffic
Fig. 2. Contribution (µg/m³) of the chemical species to each source (mean ± standard deviation of the bootstrap results, n=100).

Legend: Lev: Levoglucosan, Man: Mannosan.
Fig. 3. Daily contribution (µg/m$^3$) of each source to PM mass from 11 September 2009 to 10 September 2010.

Note: results were taken from the base run exhibiting the lowest $Q_{robust}$. 

Metal industry
Fig. 4. Source contribution (%) to each chemical species (median of the bootstrap results, n=100).

Legend: Lev: Levoglucosan, Man: Mannosan, Unaccounted: proportion of a chemical species that is not attributed to any factor.
Ammonium sulfate rich source

Ammonium nitrate rich source

Heavy oil combustion

Road traffic
Fig. 5. Probability (in %) that the contribution of a source exceeds the 75th percentile of all its contributions, when air masses came from a given air parcel (left, PSCF), or a given wind direction (right, CPF).

Note: The city of Paris is indicated by a grey dot on PSCF figures; for each source, PSCF probabilities have been normalized to 100%.
Fig. 6. Annual average contribution (µg/m$^3$; %) to PM$_{2.5}$ mass (14.7 µg/m$^3$) of the seven sources, from 11 September 2009 to 10 September 2010.
Fig. 7. Comparison of the contribution (in µg/m³) of the major sources of PM determined by receptor model studies at different European locations (see Bressi, 2012, Table S6 and text for more details).

Note: Sites are indicated as: “City (Country)-Type of sites”. Urb: urban, Rur: rural, Kerb: kerbside.
Fig. 8. Variations of the seasonal averaged contributions (µg/m$^3$) of the seven sources of PM$_{2.5}$.

Note: Calendar seasons were used (see Table 3 for more details).
Fig. 9. Averaged seasonal and annual contributions in µg/m$^3$ (left) and in % (right) of the seven sources to PM$_{2.5}$ mass (14.7 µg/m$^3$).

Note: Contributions below 0.2 µg/m$^3$ (left) and 1% of PM$_{2.5}$ mass (right) are not indicated. Calendar seasons were used. Cf. Table 3 for additional information.