Atmospheric Environment 61 (2012) 316–326

Contents lists available at SciVerse ScienceDirect

Atmospheric Environment

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

Comparison of three techniques for analysis of data from an Aerosol Time-of-Flight Mass Spectrometer

Chiara Giorio [a,*], Andrea Tapparo [a], Manuel Dall'Osto [b], Roy M. Harrison [c,d], David C.S. Beddows [c], Chiara Di Marco [e], Eiko Nemitz [e]

[a] Dipartimento di Scienze Chimiche, Università degli studi di Padova, Via Marzolo 1, 35131 Padova, Italy
[b] Institute of Environmental Assessment and Water Research (IDÆA), Consejo Superior de Investigaciones Científicas (CSIC), C/ LLuis Solé i Sabarís S/N, 08028 Barcelona, Spain
[c] National Centre for Atmospheric Science, School of Geography, Earth and Environmental Sciences, Division of Environmental Health and Risk Management, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom
[d] Department of Environmental Sciences, Center of Excellence in Environmental Studies, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[e] Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian EH26 0QB, United Kingdom

HIGHLIGHTS

- Comparison of three techniques for ATOFMS data analysis: PMF, K-means and Art-2a.
- PMF is applied, for the first time, to individual particle mass spectra.
- Efficient mass spectra deconvolution by single particle PMF analysis.
- Extraction of different organic families and two EC contributions.
- Results have been compared with independent measurements.

ARTICLE INFO

Article history:
Received 28 March 2012
Received in revised form
20 July 2012
Accepted 22 July 2012

Keywords:
Aerosol
ATOFMS
PMF
Single particle analysis
K-means
ART-2a

ABSTRACT

The Aerosol Time-of-Flight Mass Spectrometer (ATOFMS) is one of few instruments able to measure the size and mass spectra of individual airborne particles with high temporal resolution. Data analysis is challenging and in the present study, we apply three different techniques (PMF, ART-2a and K-means) to a regional ATOFMS dataset collected at Harwell, UK. For the first time, Positive Matrix Factorization (PMF) was directly applied to single particle mass spectra as opposed to clusters already generated by the other methods. The analysis was performed on a total of 56,898 single particle mass spectra allowing the extraction of 10 factors, their temporal trends and size distributions, named CNO−COOH (cyanide, oxidized organic nitrogen and carboxylic acids), SUL (sulphate), NH4-OOA (ammonium and oxidized organic aerosol), NaCl, EC+ (elemental carbon positive fragments), OC-Arom (aromatic organic carbon), EC− (elemental carbon negative fragments), K (potassium), NIT (nitrate) and OC−CHNO (organic nitrogen). The 10 factor solution from single particle PMF analysis explained 45% of variance of the total dataset, but the factors are well defined from a chemical point of view. Different EC and OC components were separated: fresh EC (factor EC−) from aged EC (factor EC+) and different organic families (factors NH4-OOA, OC-Arom, OC−CHNO and CNO−COOH). A comparison was conducted between PMF, K-means cluster analysis and the ART-2a artificial neural network. K-means and ART-2a give broadly overlapping results (with 9 clusters, each describing the full composition of a particle type), while PMF, by effecting spectral deconvolution, was able to extract and separate the different chemical species contributing to particles, but loses some information on internal mixing. Relationships were also examined between the estimated volumes of ATOFMS PMF factors and species concentrations measured independently by GRAEGOR and AMS instruments, showing generally moderate to strong correlations.

1. Introduction

In the last decade numerous epidemiological studies have revealed a significant correlation between environmental particulate matter concentrations and adverse health effects. However,

* Corresponding author. Tel.: +39 049 8275180; fax: +39 049 8275271.
E-mail address: chiara.giorio@unipd.it (C. Giorio).

1352-2310/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.atmosenv.2012.07.054

since most studies have used PM_{10} or $PM_{2.5}$ mass concentrations to investigate correlations with human health outcomes it is likely that the health impacts of PM have been in most cases underestimated (Harrison et al., 2010). Atmospheric aerosol is especially problematic to characterize because of its complex and variable composition, wide size range and a broad spectrum of both natural and anthropogenic sources. In this connection, on-line measurements deploying Mass Spectrometric techniques are very promising in order to characterize both aerosol size and chemical composition for a wide range of substances (Pratt and Prather, 2011). Aerosol Time-of-Flight Mass Spectrometry (ATOFMS) is particularly attractive as it allows size and chemical characterisation by measuring the aerodynamic diameter and positive and negative ion mass spectra of individual particles in real time within the diameter range of 0.1–3 μm (Rebotier and Prather, 2007; Gard et al., 1997; Dall'Osto et al., 2004; Drewnick et al., 2008). The ATOFMS can measure in a single campaign hundreds of thousands of single particle mass spectra which present a considerable data analysis challenge.

Successful analysis of ATOFMS data requires fast and reliable processing and interpretation of the huge amount of data generated. In order to reduce the time of analysis and the predeterministic nature of the manual classification, statistical methods can be used. The general aim of classification is to find a structure, i.e. groups of similar or related objects in the available dataset (Hinds, 1999). The main difference between a clustering method and manual classification is that the clustering method has the ability to perform analysis over the whole spectrum, rather than as individual peaks. By applying a statistical algorithm to the ATOFMS dataset, the user bias of determining which chemical information is more important in the spectra is minimised. Therefore single particle data are usually treated with a clustering algorithm, such as K-means or ART-2a, in order to group particles of similar size range and chemical composition (Rebotier and Prather, 2007; Gross et al., 2010; Healy et al., 2009; Pekney et al., 2006).

In environmental studies, factor analysis techniques (PCA, PCFA, PMF) are widely used to perform source apportionment from data taken at receptor sites. PMF analysis has been successfully applied to 24 h averaged data from analysis of particles collected on filters (Stortini et al., 2009; Jia et al., 2010; Doğan et al., 2008; Bari et al., 2009; Alleman et al., 2010) whose principal limitation is the possibility of losing the point source contributions as the characteristic time of plumes from local sources is short. Thus the results obtained are usually limited to the extraction of the 3 or 4 main sources like crustal, marine, combustion sources and secondary particulate matter, while other sources can be extracted only with a wide range of chemical analyses, size segregation and more frequent measurements (Pekney et al., 2006; Wexler and Johnston, 2008). On the other hand, PMF applied to high-resolution data (only obtainable for long periods with an on-line technique) can be a useful tool for this purpose. For example PMF analysis was successfully applied to 1 h semi-continuous characterization data of both particulate and gas phase composition leading to the extraction of 6 main sources, while by combining ATOFMS and AMS (aerosol mass spectrometry) data to the original dataset the PMF was able to identify 16 factors during a field campaign in Riverside, CA (Eatough et al., 2008).

PMF has previously been applied to ATOFMS data *after* clustering by another technique (e.g. McGuire et al., 2011), but not to data *before* clustering. In the present study, for the first time, PMF analysis is directly applied to single particle mass spectra in order to deconvolve the different chemical species which contribute to ambient particulate matter in a rural background location in Harwell (UK). A comparison among three different data treatment techniques (PMF, K-means, ART-2a) is also conducted. Hourly temporal trends of the factors extracted from single particle analysis are compared to each other in order to highlight possible correlations and to study the mixing state of ambient particles. Moreover, temporal trends of factors and clusters are compared with independent ion (and non refractory organic carbon) measurements to evaluate the performance of the data analysis.

2. Methodology

2.1. Measurement site and instrumentation

The sampling campaign was conducted in Harwell (51°34′32″N, 1°18′49″W), a rural background site in Oxfordshire (UK) from the 4th October to the 17th October 2008 deploying two on-line mass spectrometric instruments, an Aerosol Time-of-Flight Mass Spectrometer (ATOFMS TSI Model 3800-100) and an Aerosol Mass Spectrometer (Aerodyne high-resolution-ToF-AMS) (Drewnick et al., 2005; DeCarlo et al., 2006; Canagaratna et al., 2007; Jimenez et al., 2003), and a GRAEGOR (Thomas et al., 2009), which performs semi-continuous measurements of water-soluble trace gas species (NH_3, HNO_3, HONO, HCl and SO_2) collected by two wet-annular rotating denuders and their related particulate compounds (NH_4^+, NO_3^-, Cl^-, SO_4^{2-}) collected in series by two steam-jet aerosol collectors (SJAC). Sample solutions are analysed on-line by ion chromatography for anions and flow injection analysis for ammonia and ammonium (Thomas et al., 2009). During the campaign, the two inlets of GRAEGOR were placed at the same height (roughly 2 m above ground) collecting TSP and $PM_{2.5}$ simultaneously.

Hourly data for gaseous pollutant concentrations measured as part of the UK national air quality network and local weather were obtained from the UK national air quality archive (www.airquality.co.uk). Five day air mass back-trajectories arriving at Harwell at three different altitudes (100, 500 and 1000 m) were obtained using HYSPLIT (Hybrid Single Particle Lagrangian Integrated Trajectory Model) (Draxler and Rolph, 2003). Details of Harwell aerosol characterization and air mass trajectories have been provided in supplementary material.

2.2. ATOFMS technique

The ATOFMS (TSI 3800-100) collects, in real-time, bipolar mass spectra of individual aerosol particles. The instrument is constituted by an aerosol inlet, a sizing region and a mass spectrometer detector. In the aerosol inlet, particles are introduced into a vacuum system region through a converging nozzle, then focused through aerodynamic lenses into a narrow particle beam, which travels through the sizing region. The aerodynamic diameter of individual particles is determined from the time of flight between two continuous-wave laser beams ($\lambda = 532$ nm). After that, particles enter into the mass spectrometer region where a pulsed high power desorption/ionization laser ($\lambda = 266$ nm) is triggered on the basis of the transit time of the particle measured in the sizing region. Mass analysis is then provided by a bipolar time of flight reflectron mass spectrometer (Gard et al., 1997; Dall'Osto et al., 2004; Drewnick et al., 2008).

During the campaign, the ATOFMS sampled aerosol through a 3/4 inch diameter copper pipe mounted vertically and in-line with the Aerodynamic Focussing Lens (AFL). The inlet of the copper pipe (roughly 4 m above the ground) was protected using a simple hockey stick rain cap. The ATOFMS itself was fitted with a TSI 3800-100 AFL which admitted the aerosol at nominal volumetric flow rate of 0.1 L min^{-1} operating at a pressure of 2 torr. The device has a quoted size range of 100–3000 nm (Su et al., 2004) although in practice during the sampling campaign our system was capable of

hitting 56,898 particles with a measured aerodynamic diameter up to 3019 nm.

Before data analysis, single particles mass spectra were exported using the TSI MS-Analyze software. The peak-list were constructed using the following parameters: minimum peak height of 20 units above the baseline, minimum area of 20 units and representing at least the 0.005% of the total area in the particle mass spectrum. The data obtained were analysed using positive matrix factorization (PMF), K-means cluster analysis and artificial neural network (ART-2a) analysis.

2.3. Positive matrix factorization (PMF) analysis

The PMF analysis was performed using the program PMF2 (Paatero and Tapper, 1994; Paatero, 1998). Briefly, the positive matrix factorization model (whose principles are detailed elsewhere (Paatero and Tapper, 1994; Paatero, 1998)) solves the following equation $\boldsymbol{X} = \boldsymbol{GF} + \boldsymbol{E}$ where $\boldsymbol{X}$ is the original $n \times m$ data matrix, $\boldsymbol{G}$ is the $n \times p$ scores matrix (factors weight) and $\boldsymbol{F}$ is the $p \times m$ loadings matrix (factors profile), $\boldsymbol{E}$ represents the $n \times m$ residuals matrix. In the present case n is the number of particles, m is the number of m/z signals of the spectra and p is the number of factors. The exact number of factors to use was determined by monitoring the parameters suggested by Lee et al. (1999) and the chemical interpretation of the factors profile.

Data matrices. Before the PMF analysis the dataset was reduced to 106 major m/z values (−146, −144, −124, −121, −119, −104, −101, −99, −98, −97, −96, −95, −89, −88, −85, −84, −81, −80, −79, −76, −73, −72, −71, −64, −63, −62, −61, −60, −59, −49, −48, −46, −45, −44, −43, −42, −37, −36, −35, −27, −26, −25, −24, −17, −16, −15, −14, −13, −12, 7, 12, 15, 18, 23, 24, 27, 36, 37, 39, 41, 43, 46, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 69, 70, 71, 72, 73, 74, 75, 77, 81, 83, 84, 85, 86, 87, 88, 91, 94, 96, 108, 115, 118, 120, 128, 132, 138, 139, 207) and 55,357 particles by eliminating the bad variables (the ones that have more than 55,000 zero point values on a total number of particles of 56,898) and the particles with a diameter below the calibration range. Absolute area of peaks was considered for the analysis, which was directly applied to single particle mass spectra.

2.3.1. Data uncertainties

Positive Matrix Factorization relies on the accuracy of error estimates to produce reliable non-negative results and uses the estimates of the error in the data to provide both variable and sample weighting. This is particularly important when less robust datasets have to be used because of the presence of many missing or below detection limit values, as in the case of mass spectra, that could have the ability to define real sources or even be source markers (Owega et al., 2004; Paatero and Tapper, 1994; Paatero, 1998; Zhang et al., 2008). The original noise of the data ($\bar{x}_b = 4$, $\sigma_b = 4$), evaluated in zones of particle mass spectra without peaks, was added to the input matrix by simulating it with random numbers between 0 and 8, to avoid multiple zero entries. In fact, circa 70% of data in the input matrix are null values. The detection limit was evaluated as the blank value plus three times its standard deviation by integrating the mass spectra signals in several regions without peaks. The uncertainty of the data was evaluated in a laboratory experiment in which equimolar solutions of various salts were nebulised and analysed with the ATOFMS. The data reproducibility was about 50% and 80% on the average signals for positive and negative ions respectively. Moreover, the particle diameter does not influence the signal intensity. These high uncertainties reflect the principal limits of the ATOFMS analyzer which reside in the size-dependent transmission losses (Allen et al., 2000; Wenzel et al., 2003), laser intensity shot-to-shot variations (Bhave et al., 2002), ionization matrix effects (Reilly et al., 2000), different sensitivities among chemical species that make a semi-quantitative analysis possible to achieve only beside independent sampling measurements (Bhave et al., 2002; Gross et al., 2000; McGuire et al., 2011).

The data uncertainties used for the PMF analysis were then calculated as follow $s_{ij} = t + v \cdot x_{ij}$, where $t = 4DL = 64$ and $v = 0.4$ in order to give the same weight to both low and high intensity signals and to avoid the effect of background noise upon the analysis. The data uncertainty of 40% was chosen because there were no further improvements by using a higher uncertainty or different uncertainties for positive and negative ions in terms of quality of the fit and explained variations. Although Q/Qexp = 0.43 could indicate a slight overestimation of real data uncertainty, the optimized value seems to be a good compromise considering laboratory experimental data.

2.3.2. PMF solution

The robustness of factor solutions was inspected by comparing the temporal trends of factors through the different PMF solutions. The global minimum of the factor solution was achieved by starting from 50 seeds (pseudorandom starting points). The rotational ambiguity was also tested by modifying the *F*peak parameter from −2.5 to 2.5. The effect of this variation was not significant with values in the range −0.5–0.5 while PMF analysis did not converge with larger *F*peak values. Thus the PMF solution obtained could be considered unique and *F*peak = 0 was used for the final analysis. After the PMF analysis, factor loadings ($\boldsymbol{F}$) and scores ($\boldsymbol{G}$) obtained were respectively normalized and weighted as follows: each factor loading vector was normalized by dividing it by a scalar value $b_h = \sum_{j=1}^{m} f_{hj}$ and the corresponding score vector was weighted by multiplying it by the same scalar b_h.

2.4. Cluster analysis

2.4.1. K-means

ATOFMS particle mass spectra were directly imported into ENCHILADA, an open source single particle mass spectra software package (Gross et al., 2010), and 56,898 single particle mass spectra were clustered using the K-means/Euclidean square algorithm (MacQueen, 1967). K-means, which is a non hierarchical clustering technique, starts with the random subdivision of objects (in this case single particles) into a number of clusters previously defined by the operator. The algorithm computes the total heterogeneity of the system $E_T = \sum_{c=1}^{C} \sum_{i=1}^{I_c} \sum_{v=1}^{V_c} (x_{ivc} - \bar{x}_{vc})^2$, which is related to the Euclidean distance of every object to the centroid of the cluster to which the object belongs to, and moves objects from a cluster to another until it finds the minimum of system heterogeneity (MacQueen, 1967; Gross et al., 2010). In the current study, data analysis was repeated several times with increasing numbers of clusters. The exact number of clusters to use was chosen by monitoring E_T and the chemical interpretation of the cluster centroid mass spectra.

2.4.2. ART-2a

The ATOFMS dataset was imported into YAADA (Yet Another ATOFMS Data Analyzer) and single particle mass spectra were grouped with Adaptive Resonance Theory neural network, ART-2a (Song et al., 1999). The parameters used for ART-2a in this experiment were: learning rate 0.05, vigilance factor 0.85 and iterations 20. These are standard setting used in the ART-2a procedure on ATOFMS data and further details of the parameters can be found elsewhere (Song et al., 1999; Dall'Osto and Harrison, 2006; Rebotier and Prather, 2007). An ART-2a area matrix (AM) of a particle cluster represents the average intensity for each m/z for all particles within

a group. An ART-2a AM therefore reflects the typical mass spectrum of the particles within a group.

2.5. Positive matrix factorization of AMS data

Standard unit mass resolution PMF analysis was carried out on the organic matrix of the AMS dataset (Ulbrich et al., 2009). Two general factors were found: LV-OOA (low-volatile oxidized organic aerosol) and a SV-OOA (semi-volatile oxidized organic aerosol). Whilst the mass spectrum of LV-OOA factor was found to be equivalent to previous standard factor (Ulbrich et al., 2009), the factor SV-OOA contains the standard aliphatic series together with a high m/z 44 and m/z 60 signals, indicating a contribution from biomass burning (Lanz et al., 2007).

3. Results and discussion

3.1. PMF analysis on individual particle mass spectra

Single particle mass spectra were subjected to Positive Matrix Factorization analysis with solutions varying from 3 to 15 factors. According to both mathematical parameters and chemical interpretation of factor profiles, the 10 factor solution was selected. The factors extracted are:

- F1 "CNO−COOH", explaining 2% of variance, presents peaks of (CN^-) (m/z −26) and oxidized species (CNO^-) (m/z −42), ($CHOO^-$) (m/z −45) and (CH_3COO^-) (m/z −59), i.e. carboxylic acids and organic nitrogen species (Angelino et al., 2001; Dall'Osto and Harrison, 2006; Moffet et al., 2008);
- F2 "SUL" explaining 2% of variance, is characterized by the main peak of sulphate (m/z −97);
- F3 "NH4-OOA" with an explained variation of 4%, is characterized by peaks of (NH_4^+) (m/z 18) and secondary organic species ($C_2H_3^+$) (m/z 27) and ($C_2H_3O^+$) (m/z 43);
- F4 "NaCl" explaining 6% of variance, is characterized by peaks of (Na^+) (m/z 23), (Na_2^+) (m/z 46), (Na_2O^+) (m/z 62), (Na_2OH^+) (m/z 63) and (Na_2Cl^+) (m/z 81/83);
- F5 "EC+" explaining 7% of data variation, contains the elemental carbon positive ions (C^+, C_2^+, C_3^+ at $m/z = 12, 24, 36$);
- F6 "OC-Arom" explaining 5% of variance, contains signals related to organic carbon and the benzene fragment (m/z 27, 41, 43, 51, 53, 55, 57, 63, 69, 77, 87, 91, 115) (McLafferty, 1983);
- F7 "EC−" explaining 3%, is characterized by elemental carbon signals in the negative mass spectrum (C^-, C_2^-, C_3^- at $m/z = -12, -24, -36$);
- F8 "K" explaining 7%, contains the potassium signals (m/z 39/41);
- F9 "NIT" explaining 4%, is characterized by the nitrate peaks (m/z −46/−62);
- F10 "OC−CHNO" with an explained variation of 5%, is characterized by organic carbon and organic carbon related to nitrogen signals (m/z −26, 27, 37, 49−52, 60−63, 84−87).

The 10 factors obtained can explain only 45% of the total data variance but they are characterized by clear and well defined chemical patterns (Fig. 1). Despite the low explained variance, the main signals constituting the factors are well represented and they account for up to 89% of the variance of potassium for example. Sulphate is explained at 84%, while the majority of the bad variables (m/z values with low signal/noise ratio, i.e. $m/z = -146, -144, -124, -121, -119, -104, -101$) are not explained at all.

From inspection of residuals (Fig. S1) it appears that the PMF analysis failed to extract a few components: this includes chloride signals ($m/z = -35, -37$), which are not present in the NaCl factor, water signals and some other signals probably related to m/z miscalibration problems (Dall'Osto and Harrison, 2006); however, these signals do not influence the interpretation of factors. It should be noted that despite the limited explained variance, which could be a problem in relation to quantification, the factors' chemical profiles obtained are clear and well-defined and thus of qualitative value with the only exceptions of chloride and water signals.

The results obtained demonstrate that Positive Matrix Factorization analysis applied to individual particle mass spectra allows the deconvolution of the mass spectra into the contributing specific chemical species (factors K, NIT, SUL, NaCl) or their related classes (factors EC+, EC−, OC-Arom, OC−CHNO, CNO−COOH, NH4-OOA) as well as the extraction of their temporal trends and size distributions (Fig. S2). Positive and negative m/z signals are split into different factors (EC+ and EC−, K, NIT, SUL for example) due to different temporal trends either representing changing source contributions or varying relative ionization efficiencies (Bhave et al., 2002; Dall'Osto et al., 2006; Gross et al., 2000). Unlike K-means or ART-2a, PMF does not cluster whole spectra, but disaggregates them into chemical constituents, or groups of constituents. The factors are used to reconstitute actual particle mass spectra as shown in Fig. S3. From Fig. S3, it may be seen that more than one factor is necessary to reconstruct each particle mass spectrum, demonstrating the mass spectral deconvolution made by PMF analysis on single particles. For example, in Fig. S3a, the particle mass spectrum is reconstructed by 10% CNO−COOH, 3% SUL, 6% NH4-OOA, 26% OC-Arom, 10% EC−, 8% K, 8% NIT, and 29% remains unexplained.

Factor time-series were calculated as hourly sum of factor scores (not shown) and in equivalent numbers of particle (Fig. 2a). Factor time series in equivalent number of particles are calculated as sum over each hour of particle fractions attributable to each factor by first calculating the fraction of particle i attributable to factor h as

$$fF_{ih} = \frac{\sum_{j=1}^{m} g_{ih} f_{hj}}{\sum_{j=1}^{m} \left(g_{ih} f_{hj} + e_{ij} \right)} \quad (1)$$

and then summing over each hour of particle fractions attributable to factor h:

$$NfF_{h,\mathrm{hour}} = \sum_{\mathrm{hour}} fF_{ih} \quad (2)$$

The number size distributions were calculated by summing the factor scores of particles within the same size bin (size bin width of 0.01 μm). The factor size distributions are very similar to each other and all are dominated by the accumulation mode. The only exceptions are F4-NaCl, which presents a coarse distribution because of its origin from sea spray, and F9-NIT which presents both an accumulation and a coarse mode (Fig. S2). Moreover, EC− and OC−CHNO factors clearly show a distribution that is shifted towards smaller particles with a tail in the direction of the Aitken mode particles. Despite not being corrected for size-dependent inlet efficiencies, these distributions show predictable differences.

The analysis of the correlations between temporal trends of the factors, obtained by summing the score values of each factor within an hour, may give deeper insight into particle components and their sources. Correlations between factors were studied through the correlation coefficients (Fig. 3 and Table S1) in the Pearson correlation test. Almost every correlation is statistically significant (p-value <0.05) but to different degrees. The NaCl seems to be an independent factor because it has no strong

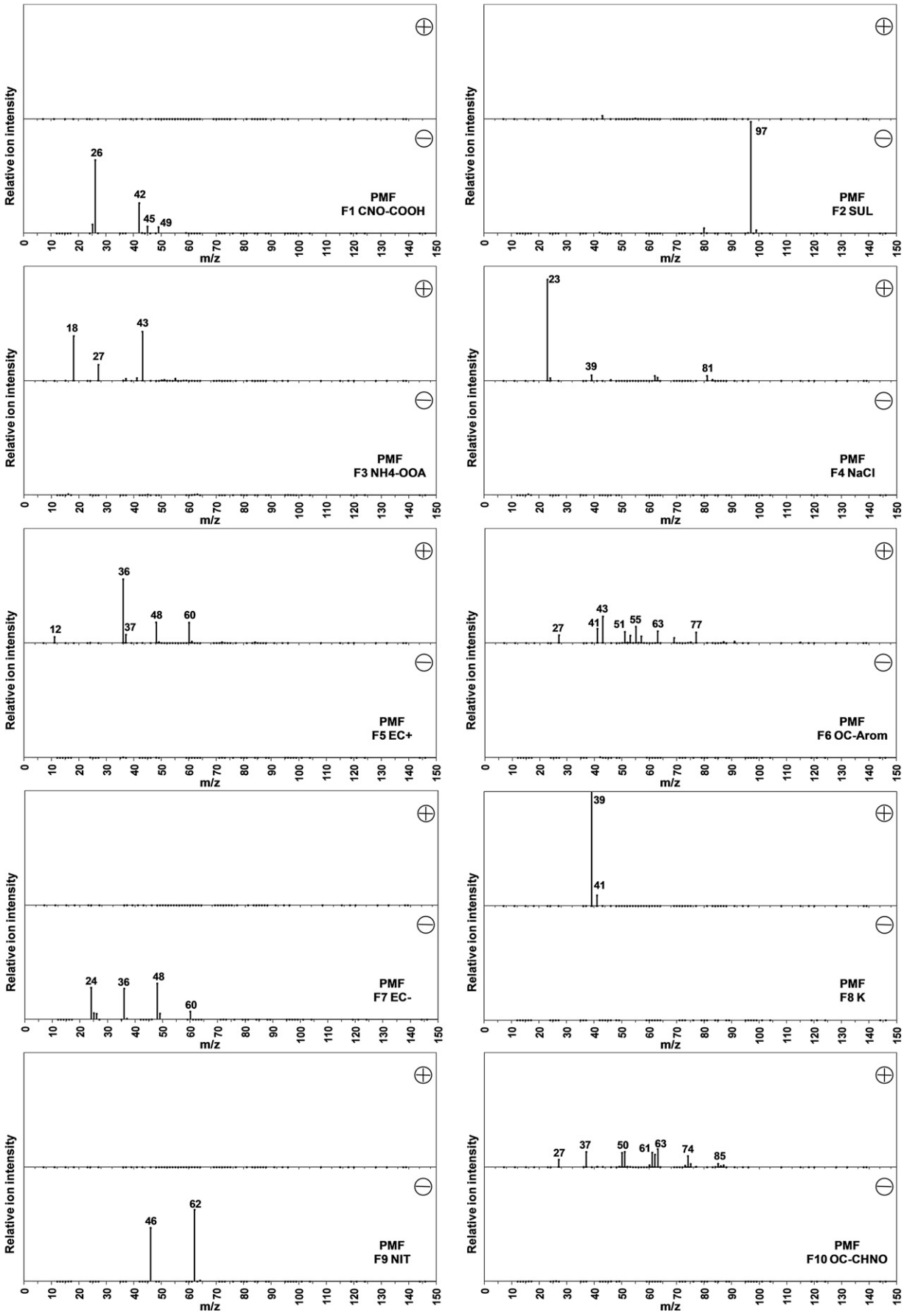

Fig. 1. Mass Spectra of the 10 PMF factors.

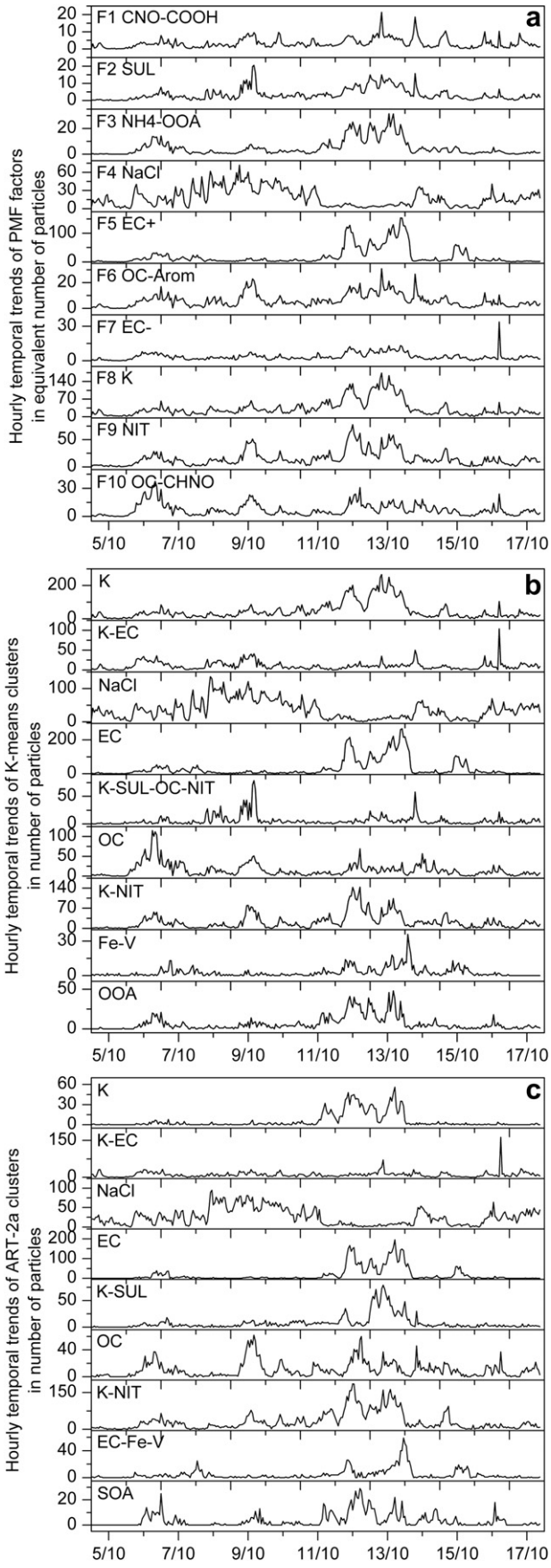

Fig. 2. Temporal trends of (a) PMF factors expressed in equivalent number of particles, (b) K-means clusters and (c) ART-2a clusters in number of particles.

correlations with the other factors, according to the Cohen classification (Cohen, 1988) and it is not correlated to sulphate, EC−, potassium, nitrate and OC−CHNO (p-value >0.05). Sulphate is strongly correlated with potassium ($r = 0.64$), nitrate ($r = 0.76$) and the organic carbon factors, OC-Arom (0.82) in particular. Potassium and nitrate are strongly correlated with almost every factor and are the dominant species, present in the majority of the particles collected. This reflects, at least in part, the very high sensitivity of the ATOFMS to these species (Gross et al., 2000).

Hourly temporal trends of EC− and EC+ present only a correlation of medium strength ($r = 0.41$). EC− is strongly correlated with OC−CHNO ($r = 0.64$) while EC+ is correlated more with secondary species (r coefficients for NH4-OOA, 0.87 and NIT, 0.61). This result suggests that the splitting of elemental carbon signals into two factors may not only reflect different ionization and detection efficiencies between positive and negative ions. It seems that the ionization pattern is influenced by the matrix composition (Reilly et al., 2000) distinguishing two different elemental carbon components: one probably freshly emitted (EC−) and one more aged (EC+), modified by oxidation reactions, and internally mixed with secondary species. In fact, as proposed by Reinard and Johnston (2008) secondary species like nitrate and sulphate could limit the electron availability, leading to a suppression of elemental carbon fragments negatively charged, while potassium and sodium, on the contrary, could lead to an enhancement of them. Moreover, the temporal trend of the EC− is characterized by a peak event on 16/10/2008 probably due to a combustion event near the sampling site.

3.2. Cluster analyses

3.2.1. K-means

The K-means analysis separated 13 clusters. Clusters obtained from miscalibrated mass spectra were eliminated and clusters with similar profiles and temporal trends were recombined to generate a total of 9 clusters (mass spectra are reported in Fig. S4a and their temporal trends, expressed as the number of particles are reported in Fig. 2b). The clusters are:

- K (14,140 particles, 25%), which presents high potassium signals and some signals of low intensity due to Na^+, cyanide, nitrate and sulphate;
- K−EC (3252 particles, 6%), which presents negative ions signals related to elemental carbon, and to a lesser extent nitrate and sulphate signals, while in the positive mass spectrum it presents signals of a low intensity, related to oxidized organic carbon, potassium and sodium;
- NaCl (10,872, 19%), which mainly presents signals of sodium, chloride, potassium and nitrate;
- EC (9436 particles, 17%), which presents both positive and negative signals related to elemental carbon and signals of nitrate and sulphate;
- K−SUL−OC−NIT (1832 particles, 3%) presents CN^-, NO^-, NO_2^-, SO_3^-, HSO_3^-, HSO_4^- signals in the negative mass spectrum and potassium and OC aromatic signals in the positive mass spectrum;
- OC (4625 particles, 8%) presents both aromatic, amine and oxygenated carbon signals and traces of ammonium, nitrate, sulphate and cyanide;
- K−NIT (6829 particles, 12%) is mainly characterized by potassium and nitrate signals along with the presence of cyanide, sulphate, ammonium and oxidized organic aerosol fragments ($m/z = +27/+43$);
- OOA (2006 particles, 4%), composed of signals corresponding to $C_2H_3^+$, $C_2H_3O^+$ and carboxylic acids along with ammonium, potassium, nitrate and sulphate;

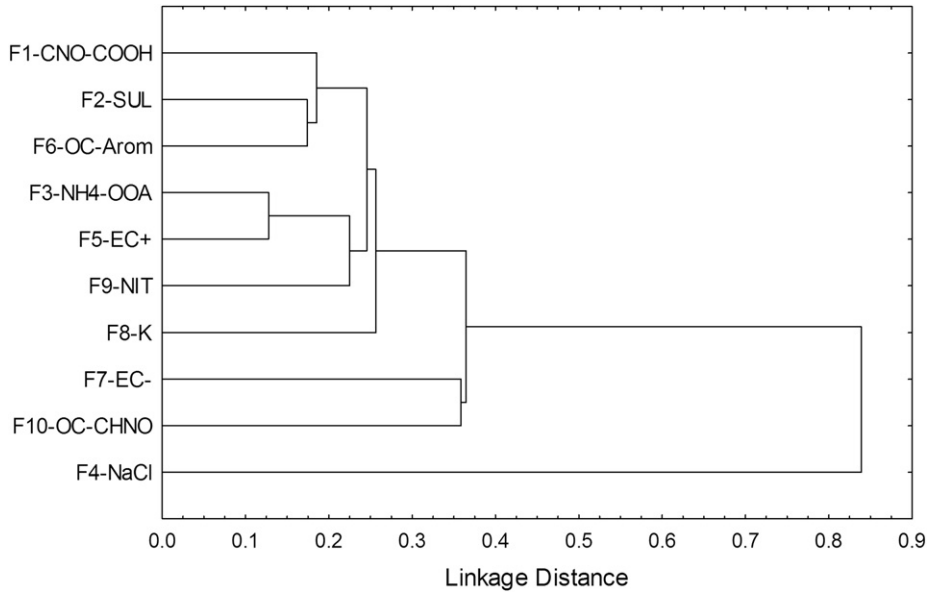

Fig. 3. Dendogram obtained from the hierarchical cluster analysis of the temporal trends of PMF factors (single linkage method, *r*-Pearson correlation coefficient distance measure).

- Fe–V (840 particles, 1%), characterized by signals at $m/z = +51/+56/+67$ that could be assigned respectively to V^+, Fe^+ and VO^+ and by signals at $m/z = +58/+60$ that could be attributed to nickel and, to a lesser extent, by sodium, potassium, elemental carbon and nitrate.

3.2.2. ART-2a

The ART-2a algorithm generated 389 clusters used to describe the dataset (total particles 56,898). The 50 most populated clusters represent more than 63% of the mass spectra from the study and thus were used for the results presented in this paper. The remaining clusters were mostly made up of a majority of mis-calibrated mass spectra. By manually merging similar clusters according to their chemical and temporal profiles with the standard procedure elsewhere described (Dall'Osto and Harrison, 2006), the total number of clusters describing the whole database was reduced to 9, representing about 63% of the total number of particles sampled (Fig. S4b). The rest of the particles presented low signal to noise ratios and therefore were not classified. The 9 clusters are:

- K–NIT (9613 particles, 17%) composed by potassium, nitrate, cyanide and sulphate;
- NaCl (7852 particles, 14%) characterized by a mass spectrum identical to K-means NaCl;
- OC (3172 particles, 6%) composed mainly by oxidized organic aerosol and aromatic compounds, and potassium, cyanide, nitrate and sulphate signals;
- K–SUL (2355 particles, 4%) with high potassium and sulphate signals, along with ammonium, nitrate and aromatic organic compounds;
- EC (5416 particles, 10%) which present a mass spectrum identical to the K-means EC;
- K (1656 particles, 3%) with a high potassium signal and Na^+, C_3^+, nitrate and sulphate signals;
- EC–Fe–V (1337 particles, 2%) composed by high signals of elemental carbon, and V^+, Fe^+ and VO^+ in the positive mass spectrum while it does not present significant signals in the negative mass spectrum;
- SOA (1066 particles, 2%) composed by ammonium, $C_2H_3^+$, $C_2H_3O^+$, potassium, low elemental and organic carbon signals, nitrate and sulphate;
- K–EC (3390 particles, 6%) elemental carbon signals, potassium, sodium, cyanide, nitrate and sulphate.

The mass spectra of the 9 clusters are shown in Fig. S4b and their time-series, expressed as number of particles are reported in Fig. 2c. Results obtained from ART-2a analysis are very similar to the K-means results. The two NaCl and EC clusters present the same mass spectrum. The two K-EC clusters are similar but the ART-2a cluster is characterized by higher cyanide, nitrate and sulphate signals. The OC ART-2a cluster presents higher aromatic signals than OC K-means cluster. The K–NIT K-means cluster presents aromatic organic carbon signals which are indeed not present in the K–NIT ART-2a cluster. The OOA K-means cluster has a mass spectrum similar to the SOA ART-2a cluster, with a high contribution of NH_4^+, $C_2H_3^+$ and $C_2H_3O^+$, but the former presents carboxylic acids signals while the latter presents more aromatic organic compounds signals. The main differences reside in the abundance of the K cluster, which is probably overestimated by K-means (25% of particles in the K-means clustering and 3% of particles in the ART-2a clustering), the EC–Fe–V ART-2a cluster which exhibits strong elemental carbon signals that are not present in the Fe–V K-means cluster, and K–SUL ART-2a cluster which has a different positive mass spectrum, dominated by the potassium signal, while the K–SUL–OC–NIT K-means cluster has OC aromatic signals in the positive mass spectrum.

The differences between the two techniques could reside in the different approach to clustering the data. In K-means cluster analysis, all particles are assigned to the clusters by dividing them into groups of similarity. The number of clusters is chosen by the operator who proceeds with a trial-and-error approach by incrementing the number of clusters until the division into more clusters is chemically meaningless (13 clusters in this case). On the contrary, ART-2a (running with standard parameters) usually produces a huge number of clusters (389 in this case). After that, clusters made by only few particles are eliminated and only the main contributing clusters are considered, and clusters of similar composition and size distribution are merged manually. Thus, ART-

2a may give more clear and well defined clusters than K-means which considers more particles than the former in the final solution.

3.3. Comparison between results of PMF analysis on single particles, K-means cluster analysis and ART-2a artificial neural network analysis

PMF and cluster analysis can be viewed as complementary techniques. While K-means and Art-2 give a rapid classification of whole particles by dividing them into classes of similarity the PMF analysis on single particle mass spectra permits the extraction of the chemical species constituting the particles. Much of the information on internal mixing is lost.

The results of the correlation analysis (Pearson correlation test) among cluster and factor temporal trends (in equivalent number of particles) which have a similar chemical profile show a good agreement (Table 1). For instance, taking into account K-means results, PMF F8-K is correlated to the K cluster with $r^2 = 0.99$ and p-value = <0.001 (Pearson correlation test); F4-NaCl is correlated to the NaCl cluster ($r^2 = 0.91$, p-value = <0.001) and F3 NH4-OOA is correlated to the OOA cluster ($r^2 = 0.83$, p-value = <0.001). The cross correlations between factors and cluster temporal trends confirm the conclusions obtained from the cross correlations between PMF factors (Fig. 3 and Table S1). The NaCl cluster presents a strong correlation only with the NaCl factor. In fact, NaCl is an independent particle type which is directly associated with the sea spray source. The EC+ factor is strongly correlated with clusters characterized by secondary aerosol (OOA $r^2 = 0.53$, K $r^2 = 0.59$, K–NIT $r^2 = 0.36$) while EC− is not strongly correlated with any cluster, confirming the two different elemental carbon contributions to aged (EC+) and fresh (EC−) particles. The cluster Fe–V is strongly correlated with the EC+ factor ($r^2 = 0.49$) probably because of a common origin from oil based fuel combustion (Korn et al., 2007) or transported from coal-fired power plants in Central Europe. In fact, EC+ abundance increased during long-range transport of air masses from Central Europe (see SI). Moreover, the K–EC cluster which is moderately correlated to organic factors as well as EC−, could represent a biomass burning signature (Bi et al., 2011; Healy et al., 2012). CNO–COOH, SUL and NH4-OOA PMF factors, as expected, are present in multiple clusters as they are highly oxidized aerosol components produced during ageing processes.

3.4. Comparison of PMF analysis results with independent measurements

Alongside the ATOFMS, inorganic water soluble components in the TSP and $PM_{2.5}$ size fractions were measured by GRAEGOR, and in non-refractory PM_1 (NR-PM_1) by the AMS, defined as those components within PM_1 that volatilise rapidly at the vaporiser temperature of 600 °C. In order to validate the PMF factor temporal trends, a correlation analysis (r-Pearson test) was made between them and these independent measurements. In Fig. 4 the sulphate, nitrate, chloride, ammonium and organic concentrations are reported compared to the corresponding PMF ATOFMS factors.

For this purpose, factor temporal trends were calculated under the simplifying assumptions that all particles are homogenous, spherical and a constant mass of material is ionized from each particle, irrespective of their size (Dall'Osto et al., 2006). Particle volume was multiplied by the percentage contribution of each factor to it. The hourly time-series (in volume) of the factors were then calculated by summing the partial volume of each particle attributable to each factor (Fig. 4). For comparison with AMS PM_1 concentrations, PMF factor partial volumes were integrated for particles of <1 μm diameter. It is important to note that ATOFMS time-series were not corrected for size-dependent inlet efficiencies (Dall'Osto et al., 2006).

The SUL factor (expressed in volume of particles) is significantly correlated with sulphate concentrations in $PM_{2.5}$ ($r^2 = 0.34$, p-value = <0.001) and in AMS PM_1 ($r^2 = 0.41$, p-value = <0.001). In the case of nitrate, the NIT PMF factor temporal trend is weakly correlated with nitrate concentration in $PM_{2.5}$ ($r^2 = 0.07$, p-value = <0.001), but is strongly correlated with nitrate in NR-PM_1 ($r^2 = 0.54$, p-value = <0.001). The difference in the correlations may reflect different instrumental inlet characteristics leading to different large particle contributions to the temporal patterns. In fact, while NR-PM_1 is fairly specific to NH_4NO_3, $PM_{2.5}$ can contain also significant amount of $NaNO_3$, produced by sea salt processing through HNO_3. NIT PMF factor presents both an accumulation and a coarse mode, and the latter could be measured with higher efficiency than the former, and would also contain contributions that are not included in the NR-PM_1. The high correlations seen for the NR-PM_1 fraction are however reassuring.

The NaCl factor is weakly but significantly correlated to the chloride measurements in $PM_{2.5}$ ($r^2 = 0.11$, p-value = <0.001). On the contrary AMS chloride is not significantly correlated with the GRAEGOR chloride measurements (p-value = 0.32 for TSP and 0.46 for $PM_{2.5}$), which shows much larger concentrations, because the AMS only detects the non-refractory fraction which is thought to be dominated by NH_4Cl. The NH4-OOA factor, which contains both OOA and ammonium signals, is correlated with the ammonium concentration in $PM_{2.5}$ ($r^2 = 0.61$, p-value = <0.001) and in NR-PM_1 ($r^2 = 0.59$, p-value = <0.001) and to the organic component measured by the AMS ($r^2 = 0.60$, p-value = <0.001). The non-refractory organic concentration measured by AMS is strongly correlated with ammonium concentration and presents the highest

Table 1
Coefficient of determination (r^2) values of the linear regressions between hourly temporal trends of PMF factors (equivalent number of particles) and K-means clusters or ART-2a clusters.[a]

PMF factors	r^2 (PMF factors vs K-means clusters)									r^2 (PMF factors vs ART-2a clusters)								
	K	K–EC	NaCl	EC	K–SUL–OC–NIT	OC	K–NIT	Fe–V	OOA	K–NIT	NaCl	OC	K–SUL	EC	K	EC–Fe–V	SOA	K–EC
CNO–COOH	0.43	0.34	0.00	0.09	0.28	0.09	0.25	0.02	0.12	0.29	0.00	0.19	0.27	0.10	0.09	0.01	0.04	0.16
SUL	**0.53**	0.20	0.00	0.28	**0.51**	0.12	0.45	0.10	0.37	0.47	0.00	0.33	0.42	0.31	0.30	0.08	0.22	0.07
NH4-OOA	**0.76**	0.01	0.11	**0.68**	0.01	0.17	**0.64**	0.27	**0.83**	**0.78**	0.12	0.19	0.46	**0.76**	**0.73**	0.23	0.48	0.03
NaCl	0.06	0.06	**0.91**	0.12	0.06	0.00	0.03	0.06	0.08	0.07	**0.82**	0.00	0.05	0.12	0.12	0.05	0.07	0.01
EC+	**0.59**	0.00	0.14	**0.99**	0.00	0.02	0.36	0.49	**0.53**	**0.53**	0.16	0.04	0.38	**0.87**	**0.60**	**0.66**	0.24	0.01
OC-Arom	**0.54**	0.31	0.00	0.25	0.36	0.29	**0.59**	0.08	0.42	**0.53**	0.00	0.45	0.34	0.30	0.29	0.05	0.26	0.11
EC−	**0.53**	0.35	0.02	0.42	0.06	0.17	0.42	0.16	0.39	0.43	0.02	0.17	0.26	0.42	0.32	0.18	0.23	0.08
K	**0.99**	0.03	0.04	**0.56**	0.03	0.04	0.47	0.17	**0.61**	**0.78**	0.05	0.11	**0.64**	**0.58**	**0.60**	0.21	0.25	0.07
NIT	**0.60**	0.12	0.00	0.37	0.12	0.19	**0.92**	0.12	**0.64**	**0.77**	0.00	**0.50**	0.21	**0.50**	**0.54**	0.08	0.42	0.06
OC–CHNO	0.14	0.40	0.00	0.08	0.11	**0.83**	0.49	0.03	0.30	0.24	0.00	**0.51**	0.03	0.14	0.12	0.01	0.36	0.11

[a] Strong correlated results ($r^2 > 0.5$) are presented in bold.

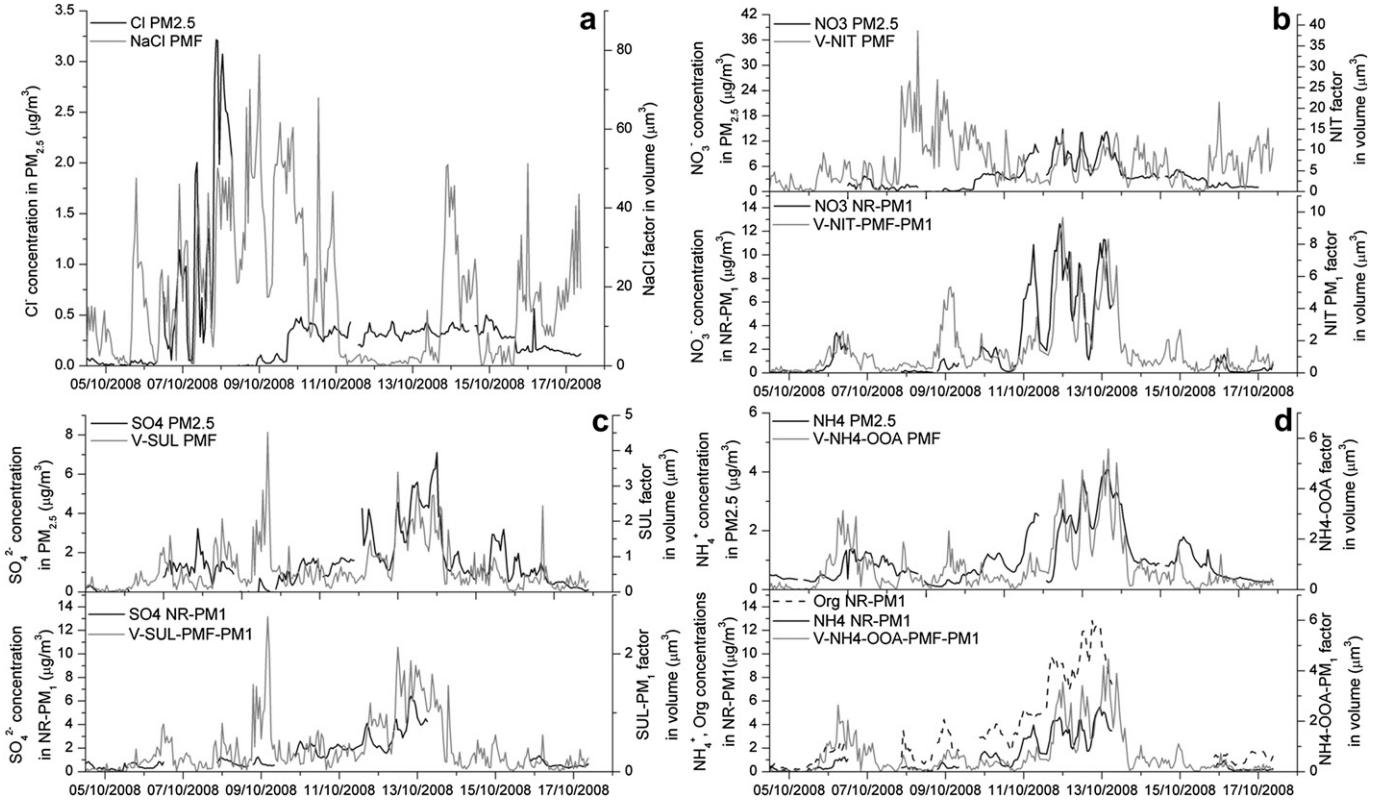

Fig. 4. Hourly time-series of (a) chloride concentrations and NaCl factor, (b) nitrate concentrations and NIT factor, (c) sulphate concentrations and SUL factor, (d) ammonium, organic concentrations and NH4-OOA factor.

correlation with the NH4-OOA PMF factor rather than with other organic factors. ATOFMS factors were also compared to TSP ion measurements, but because of different inlet characteristics, the correlations are in general weak or not significant and the results are not reported.

The analysis shows clearly that PMF factors are highly significantly correlated with the corresponding chemical species mass concentrations, with a better agreement with NR-PM_1 (if ATOFMS PMF factors are integrated for particles <1 μm). On the contrary, clustering analytical techniques such as K-means and ART-2a cannot disaggregate the contribution of the different chemical species present in the particles. For this reason, a direct comparison between the time-series of a cluster and the mass concentration of one of its components is not appropriate. In fact, such correlation would be highly dependent on particle mixing-state. Thus, the disaggregation of species made by the PMF analysis (on single particles) proves very useful for quantification purposes of the principal substances or classes of substances constituting the particles. The determination coefficient, slope and intercept of the linear regressions between ATOFMS factors and the species concentrations measured by AMS in NR-PM1 are reported in Supplementary Material. (Table S2). Moreover, the correlation between PMF factors and the corresponding species concentrations may be even stronger if ATOFMS data are corrected for size-dependent transmission losses (Jeong et al., 2011).

3.5. Comparison between ATOFMS-PMF factors and AMS-PMF factors for secondary organic aerosol

In order to further validate the PMF analysis on single particle ATOFMS spectra, the factors obtained were compared with standard factors (Ulbrich et al., 2009) extracted by PMF analysis on the organic matrix of the AMS measurements (Table S3, Supplementary Materials). The comparison was conducted considering ATOFMS-PMF factor time-series in volume (integrated over particles of <1 μm diameter) because AMS-PMF factors are expressed in mass concentration (μg m^{-3}). The results show that ATOFMS-PMF factors associated with aged aerosol (NH4-OOA and EC+) are better correlated with the most aged LV-OOA AMS-PMF factor ($r^2 = 0.66$ and 0.67 for NH4-OOA and EC+ respectively) rather than with SV-OOA ($r^2 = 0.55$ and 0.43 for NH4-OOA and EC+ respectively). On the contrary, fresh or less aged components (ATOFMS-PMF factors OC-Arom and OC−CHNO) are better correlated with the less aged SV-OOA AMS-PMF factor ($r^2 = 0.54$ and 0.37 for OC-Arom and OC−CHNO respectively) rather than with LV-OOA ($r^2 = 0.36$ and 0.07 for OC-Arom and OC−CHNO respectively).

Unexpectedly, EC− presented a correlation of medium intensity with both AMS-PMF factors ($r^2 = 0.45$ for LV-OOA and $r^2 = 0.43$ for SV-OOA). However, the correlations are stronger, especially with respect to the less aged SV-OOA if the EC− time-series is expressed as the equivalent number of particles ($r^2 = 0.62$). This may be due to the fact that using time-series calculated in volume we may further underestimate the contribution of small particles because of size-dependent transmission losses (Gross et al., 2000; Dall'Osto et al., 2006).

3.6. Harwell aerosol characterization

From the study of the back-trajectories of air masses arriving in Harwell during the sampling campaign (detailed in S.M.), it was clear that the NaCl factor was dominant during the sampling of marine-polar air masses, while during periods of sampling continental air masses (from Central Europe) elemental carbon, potassium, nitrate and sulphate concentrations increased. This is as

expected as Harwell is a rural background site and it should not be influenced substantially by local primary sources. More interesting were two marine-continental periods. The first was characterized by air masses coming from the ocean, crossing Scotland and England before arriving at the Harwell site. It was characterized by high concentrations of NO_x and other primary gaseous pollutants, and a high abundance of the OC–CHNO and EC– factors. The second was characterized by air masses coming from the west coast of France, low concentrations of primary gaseous pollutants and a high amount of CNO–COOH and EC+ PMF factors (Table S4). Thus, the first period was characterized by freshly emitted aerosol while the second period is characterized by aged and chemically oxidized particles.

4. Conclusions

PMF analysis has been applied to single particle ATOFMS mass spectra and allows the extraction and separation of significant contributing chemical components. In general, PMF factor profiles identify well defined chemical species or classes of substances from inorganic (NaCl, K, NIT, SUL) to organic families (EC+, EC–, OC-Arom, OC–CHNO, CNO–COOH, NH4-OOA). There is a partial loss of information on internal mixing of particles.

From the cross correlation analysis among temporal trends of PMF factors it was possible to identify two elemental carbon components: the EC– factor, correlated to OC–CHNO, probably related to anthropogenic primary emissions and the EC+ factor present in aged particles internally mixed with secondary species. Furthermore, this is the first time in which different families of organic carbon have been extracted from ATOFMS data, including aromatic, oxidized organic compound and two different organic nitrogen components: primary (OC–CHNO) and oxidized (CNO–COOH). Oxidized carbon in the form of oxidized organic nitrogen and carboxylic acids is found only in aged aerosol while nitrate and sulphate are found in different proportions: the former in less aged aerosol such as in urban plumes while sulphate arose predominantly from long-range transport from continental sources.

From the comparison of different data treatment techniques it emerges that K-means cluster analysis and ART-2a artificial neural network analysis give similar results, with particles grouped in clusters of similar composition, reflective of aerosol sources, chemical processes and a combination of both, while PMF analysis of single particle mass spectra allows the deconvolution of the mass spectra and the extraction of some constituent components. Moreover, when expressed in volume, the temporal trends of PMF factors are highly significantly correlated to the corresponding chemical species concentration measured by independent instruments, even in the case of highly internally mixed particles, while the correlation between cluster temporal trends and corresponding chemical species concentration is highly dependent upon particle mixing state. Thus PMF analysis may prove useful for the quantification of the main components of PM data collected with the ATOFMS instrument. However, better repeatability of the ionization process and higher efficiency of particle detection would improve its quantification capability.

Acknowledgements

The GRAEGOR and AMS measurements were funded by the UK Department for Environment, Food and Rural Affairs within the UK contribution to the Intensive Measurement Periods of the EMEP Programme of the UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP). We would like to thank Dr Andre Prevot and Mr Francesco Canonaco from the Paul Scherrer Institut (PSI, Switzerland) who conducted the PMF analysis on the organic AMS matrix.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.atmosenv.2012.07.054.

References

Alleman, L.Y., Lamaison, L., Perdrix, E., Robache, A., Galloo, J.C., 2010. PM_{10} metal concentrations and source identification using positive matrix factorization and wind sectoring in a French industrial zone. Atmospheric Research 96, 612–625.

Allen, J.O., Fergenson, D.P., Gard, E.E., Hughes, L.S., Morrical, B.D., Kleeman, M.J., Gross, D.S., Gälli, M.E., Prather, K.A., Cass, G.R., 2000. Particle detection efficiencies of aerosol time of flight mass spectrometers under ambient sampling conditions. Environmental Science & Technology 34, 211–217.

Angelino, S., Suess, D.T., Prather, K.A., 2001. Formation of aerosol particles from reactions of secondary and tertiary alkylamines: characterization by aerosol time-of-flight mass spectrometry. Environmental Science & Technology 35, 3130–3138.

Bari, M.A., Baumbach, G., Kuch, B., Scheffknecht, G., 2009. Wood smoke as a source of particle-phase organic compounds in residential areas. Atmospheric Environment 43, 4722–4732.

Bhave, P.V., Allen, J.O., Morrical, B.D., Fergenson, D.P., Cass, G.R., Prather, K.A., 2002. A field-based approach for determining ATOFMS instrument sensitivities to ammonium and nitrate. Environmental Science & Technology 36, 4868–4879.

Bi, X., Zhang, G., Li, L., Wang, X., Li, M., Sheng, G., Fu, J., Zhou, Z., 2011. Mixing state of biomass burning particles by single particle aerosol mass spectrometer in the urban area of PRD, China. Atmospheric Environment 45, 3447–3453.

Canagaratna, M.R., Jayne, J.T., Jimenez, J.L., Allan, J.D., Alfarra, M.R., Zhang, Q., Onasch, T.B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L.R., Trimborn, A.M., Northway, M.J., DeCarlo, P.F., Kolb, C.E., Davidovits, P., Worsnop, D.R., 2007. Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. Mass Spectrometry Reviews 26, 185–222.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates Inc., New Jersey, USA.

Dall'Osto, M., Harrison, R.M., 2006. Chemical characterisation of single airborne particles in Athens (Greece) by ATOFMS. Atmospheric Environment 40, 7614–7631.

Dall'Osto, M., Beddows, D.C.S., Kinnersley, R.P., Harrison, R.M., 2004. Characterization of individual airborne particles by using aerosol time-of-flight mass spectrometry at Mace Head, Ireland. Journal of Geophysical Research 109, D21302.

Dall'Osto, M., Harrison, R.M., Beddows, D.C.S., 2006. Single-particle efficiencies of aerosol time-of flight mass spectrometry during the North Atlantic marine boundary layer experiment. Environmental Science & Technology 40, 5029–5035.

DeCarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Analytical Chemistry 78, 8281–8289.

Doğan, G., Güllü, G., Tuncel, G., 2008. Sources and source regions effecting the aerosol composition of the Eastern Mediterranean. Microchemical Journal 88, 142–149.

Draxler, R.R., Rolph, G.D., 2003. HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) Model, vol. 4.9. NOAA Air Resource Laboratory, Silver Spring MD. http://ready.arl.noaa.gov/HYSPLIT.php.

Drewnick, F., Hings, S.S., DeCarlo, P., Jayne, J.T., Gonin, M., Fuhrer, K., Weimer, S., Jimenez, J.L., Demerjian, K.L., Borrmann, S., Worsnop, D.R., 2005. A new time-of-flight aerosol mass spectrometer (TOF-AMS)-instrument description and first field deployment. Aerosol Science & Technology 39, 637–658.

Drewnick, F., Dall'Osto, M., Harrison, R.M., 2008. Characterization of aerosol particles from grass mowing by joint deployment of ToF-AMS and ATOFMS instruments. Atmospheric Environment 42, 3006–3017.

Eatough, D.J., Grover, B.D., Woolwine, W.R., Eatough, N.L., Long, R., Farber, R., 2008. Source apportionment of 1h semi-continuous data during the 2005 study of organic aerosols in riverside (SOAR) using positive matrix factorization. Atmospheric Environment 42, 2706–2719.

Gard, E., Mayer, J.E., Morrical, B.D., Dienes, T., Fergenson, D.P., Prather, K.A., 1997. Real-time analysis of individual atmospheric aerosol particles: design and performance of a portable ATOFMS. Analytical Chemistry 69, 4083–4091.

Gross, D.S., Galli, M.E., Silva, P.J., Prather, K.A., 2000. Relative sensitivity factors for alkali metal and ammonium cations in single-particle aerosol time-of-flight mass spectra. Analytical Chemistry 72, 416–422.

Gross, D.S., Atlas, R., Rzeszotarski, J., Turetsky, E., Christensen, J., Benzaid, S., Olson, J., Smith, T., Steinberg, L., Sulman, J., Ritz, A., Anderson, B., Nelson, C., Musicant, D.R.,

Chen, L., Snyder, D.C., Schauer, J.J., 2010. Environmental chemistry through intelligent atmospheric data analysis. Environmental Modelling & Software 25, 760–769.
Harrison, R.M., Giorio, C., Beddows, D.C.S., Dall'Osto, M., 2010. Size distribution of airborne particles controls outcome of epidemiological studies. Science of the Total Environment 409, 289–293.
Healy, R.M., O'Connor, I.P., Hellebust, S., Allanic, A., Sodeau, J.R., Wenger, J.C., 2009. Characterisation of single particles from in-port ship emissions. Atmospheric Environment 43, 6408–6414.
Healy, R.M., Sciare, J., Poulain, L., Kamili, K., Merkel, M., Müller, T., Wiedensohler, A., Eckhardt, S., Stohl, A., Sarda-Estève, R., McGillicuddy, E., O'Connor, I.P., Sodeau, J.R., Wenger, J.C., 2012. Sources and mixing state of size-resolved elemental carbon particles in a European megacity: Paris. Atmospheric Chemistry and Physics 12, 1681–1700.
Hinds, W.C., 1999. Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles. John Wiley & Sons, New York.
Jeong, C.H., McGuire, M.L., Godri, K.J., Slowik, J.G., Rehbein, P.J.G., Evans, G.J., 2011. Quantification of aerosol chemical composition using continuous single particle measurements. Atmospheric Chemistry and Physics 11, 7027–7044.
Jia, Y., Clements, A.L., Fraser, M.P., 2010. Saccharide composition in atmospheric particulate matter in the southwest US and estimates of source contributions. Aerosol Science 41, 62–73.
Jimenez, J.L., Jayne, J.T., Shi, Q., Kolb, E., Worsnop, D.R., Yourshaw, I., Seinfeld, J.H., Flagan, R.C., Zhang, X., Smith, K.A., Morris, J.W., Davidovits, P., 2003. Ambient aerosol sampling using the aerodyne aerosol mass spectrometer. Journal of Geophysical Research 108 (D7), 8425.
Korn, M.G.A., Santos, D.S.S., Welz, B., Rodrigues, M.G., Teixeira, A.P., Lima, D.C., Ferreira, L.C., 2007. Atomic spectrometric methods for the determination of metals and metalloids in automotive fuels – a review. Talanta 73, 1–11.
Lanz, V.A., Alfarra, M.R., Baltensperger, U., Buchmann, B., Hueglin, C., Prevot, A.S.H., 2007. Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra. Atmospheric Chemistry and Physics 7, 1503–1522.
Lee, E., Chan, C.K., Paatero, P., 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmospheric Environment 33, 3201–3212.
MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. University of California Press, Berkeley, pp. 281–297.
McGuire, M.L., Jeong, C.H., Slowik, J.G., Chang, R.Y.W., Corbin, J.C., Lu, G., Mlhele, C., Rehbein, P.J.G., Sills, D.M.L., Abbatt, J.P.D., Brook, J.R., Evans, G.J., 2011. Elucidating determinants of aerosol composition through particle-type-based receptor modelling. Atmospheric Chemistry & Physics 11, 8133–8155.
McLafferty, F.W., 1983. Interpretation of Mass Spectra, third ed. CA University Co. Books, Mill Valley.
Moffet, R.C., de Foy, B., Molina, L.T., Molina, M.J., Prather, K.A., 2008. Measurement of ambient aerosols in northern Mexico City by single particle mass spectrometry. Atmospheric Chemistry & Physics 8, 4499–4516.
Owega, S., Khan, B.U.Z., D'Souza, R., Evans, G.J., Fila, M., Jervis, R.E., 2004. Receptor modeling of Toronto $PM_{2.5}$ characterized by aerosol laser ablation mass spectrometry. Environmental Science & Technology 38, 5712–5720.
Paatero, P., Tapper, U., 1994. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5, 111–126.
Paatero, P., 1998. User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3.
Pekney, N.J., Davidson, C.I., Bein, K.J., Wexler, A.S., Johnston, M.V., 2006. Identification of sources of atmospheric PM at the Pittsburgh supersite, Part I: single particle analysis and filter-based positive matrix factorization. Atmospheric Environment 40, S411–S423.
Pratt, K.A., Prather, K.A., 2011. Mass spectrometry of atmospheric aerosols – recent developments and applications. Part II: on-line mass spectrometry techniques. Mass Spectrometry Reviews. http://dx.doi.org/10.1002/mas.20330.
Rebotier, T.P., Prather, K.A., 2007. Aerosol time-of-flight mass spectrometry data analysis: a benchmark of clustering algorithms. Analytica Chimica Acta 585, 38–54.
Reilly, P.T.A., Lazar, A.C., Gieray, R.A., Whitten, W.B., Ramsey, J.M., 2000. The elucidation of charge-transfer-induced matrix effects in environmental aerosols via real-time aerosol mass spectral analysis of individual airborne particles. Aerosol Science & Technology 33, 135–152.
Reinard, M.S., Johnston, M.V., 2008. Ion formation mechanism in laser desorption ionization of individual particles. Journal of the American Society for Mass Spectrometry 19, 389–399.
Song, X.H., Hopke, P.K., Fergenson, D.P., Prather, K.A., 1999. Classification of single particles analyzed by ATOFMS using an artificial neural network, ART-2A. Analytical Chemistry 71, 860–865.
Stortini, A.M., Freda, A., Cesari, D., Cairns, W.R.L., Contini, D., Barbante, C., Prodi, F., Cescon, P., Gambaro, A., 2009. An evaluation of the $PM_{2.5}$ trace elemental composition in the Venice Lagoon area and an analysis of the possible sources. Atmospheric Environment 43, 6296–6304.
Su, Y., Sipin, M.F., Furutani, H., Prather, K.A., 2004. Development and characterization of an aerosol time-of-flight mass spectrometer with increased detection efficiency. Analytical Chemistry 76, 712–719.
Thomas, R.M., Trebs, I., Otjes, R., Jongejan, P.A.C., Brink, H.T., Phillips, G., Kortner, M., Meixner, F.X., Nemitz, E., 2009. An automated analyzer to measure surface-atmosphere exchange fluxes of water soluble inorganic aerosol compounds and reactive trace gases. Environmental Science & Technology 43, 1412–1418.
Ulbrich, I.M., Canagaratna, M.R., Zhang, Q., Worsnop, D.R., Jimenez, J.L., 2009. Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data. Atmospheric Chemistry and Physics 9, 2891–2918.
Wenzel, R.J., Liu, D.Y., Edgerton, E.S., Prather, K.A., 2003. Aerosol time-of-flight mass spectrometry during the Atlanta supersite experiment: 2. Scaling procedures. Journal of Geophysical Research-Atmosphere 108 (D7), 8427.
Wexler, A.S., Johnston, M.V., 2008. What have we learned from highly time-resolved measurements during EPA's supersites program and related studies? Journal of the Air & Waste Management Association 58, 303–319.
Zhang, T., Claeys, M., Cachier, H., Dong, S., Wang, W., Maenhaut, N., Liu, X., 2008. Identification and estimation of the biomass burning contribution to Beijing aerosol using levoglucosan as a molecular marker. Atmospheric Environment 42, 7013–7021.