

SoFi, an Igor based interface

F. Canonaco et al.

SoFi, an Igor based interface for the efficient use of the generalized multilinear engine (ME-2) for source apportionment: application to aerosol mass spectrometer data

F. Canonaco, M. Crippa, J. G. Slowik, U. Baltensperger, and A. S. H. Prévôt

Paul Scherrer Institute, Laboratory of Atmospheric Chemistry, 5232 Villigen PSI, Switzerland

Received: 24 April 2013 – Accepted: 26 June 2013 – Published: 15 July 2013

Correspondence to: A. S. H. Prévôt (andre.prevot@psi.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Source apportionment using the bilinear model through the multilinear engine (ME-2) was successfully applied to non-refractory organic aerosol (OA) mass spectra collected during winter 2011 and 2012 in Zurich, Switzerland using the aerosol chemical speciation monitor ACSM. Five factors were identified: low-volatility oxygenated OA (LV-OOA), semivolatile oxygenated OA (SV-OOA), hydrocarbon-like OA (HOA), cooking OA (COA) and biomass burning OA (BBOA). A graphical user interface SoFi (Source Finder) was developed at PSI in order to facilitate the testing of different rotational techniques available within the ME-2 engine by providing a priori factor profiles for some or all of the expected factors. ME-2 was used to test the positive matrix factorization (PMF) model, the fully constrained chemical mass balance (CMB) model, and partially constrained models utilizing *a* values and pulling equations. Within the set of model solutions determined to be environmentally reasonable, BBOA and SV-OOA factor mass spectra and time series showed the greatest variability. This variability represents uncertainty in the model solution and indicates that analysis of model rotations provides a useful approach for assessing the uncertainty of bilinear source apportionment models.

1 Introduction

Atmospheric aerosols are of scientific and political interest due to their highly uncertain direct and indirect effects on the solar radiation balance of the Earth's atmosphere (IPCC, 2007). Moreover, aerosols have a strong negative effect on human health (Peng et al., 2005), visibility (Watson, 2002), ecosystems and agricultural areas via acidification and eutrophication (Matson et al., 2002). Therefore, reliable source identification and quantification is essential for the development of effective political abatement strategies. Atmospheric aerosols can be roughly separated based on their formation process into primary and secondary aerosols, i.e. directly emitted and formed from gas-phase reactions of emitted precursor gases, respectively. However, detailed aerosol

AMTD

6, 6409–6443, 2013

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



formation processes are still poorly understood, in particular the submicron organic fraction of particulate matter (PM₁) (Hallquist et al., 2009), which comprises 20–70 % of the total submicron aerosol mass depending on the measurement location (Jimenez et al., 2009), is poorly characterized.

5 The Aerodyne aerosol mass spectrometer (AMS) provides on-line quantitative mass spectra of the non-refractory (inorganic and organic) components of the non-refractory submicron aerosol fraction with high time resolution, i.e., seconds to minutes (Canagaratna et al., 2007). Through the knowledge of the typical mass spectral fragmentation patterns, these spectra can be assigned to several inorganic components and to the organic fraction (Allan et al., 2004). However, interpretation of the organic fraction is challenging due to the enormous number of possible compounds. Over 10 the past years numerous ambient studies have successfully exploited the positive matrix factorization (PMF) algorithm apportioning the measured organic mass spectra in terms of source/process-related components, see Zhang et al. (2011) for a review. The 15 statistical tool PMF (Paatero and Tapper, 1994; Paatero, 1997) is a bilinear model that represents the time series of measured organic mass spectra as a linear combination of static factor profiles (i.e. mass spectra) and their respective time series. However, if all measured variables, i.e. the mass to charge ratios (m/z) for mass spectrometer exhibit temporal covariation, e.g. due to meteorological events such as rainfall or 20 boundary layer evolution or if the model solution has high rotational ambiguity, then the apportionment with PMF can yield non-meaningful or mixed factors. Under such conditions, the bilinear model can be directed towards an optimal solution by utilizing a priori information in the form of the factor profiles and/or time series. If all factor profiles are predetermined, the approach is called chemical mass balance (CMB). At the other extreme, in PMF the factor profiles are calculated entirely by the algorithm. The multilinear 25 engine algorithm (ME-2) is capable of solving both these extremes and all intermediate cases in accordance with the constraints provided by the user (Paatero, 1999; Paatero and Hopke, 2009). Several PM source apportionment studies in which PMF did not properly represent the measured data have utilized ME-2 to find acceptable solutions,

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



e.g., (Lanz et al., 2008; Amato and Hopke, 2012; Reche et al., 2012). However, such studies are scarce, possibly due to the need for manual configuration and analysis of the results of the powerful ME-2 package. Therefore, to facilitate the choice of the initial conditions for the ME-2 engine and the analysis of the results, the authors have written the graphical user interface SoFi (Source Finder) within the software package Igor Pro Wavemetrics, Inc., Portland, OR, USA. This package can be provided to all interested ME-2/PMF users.

In this study the ME-2 engine was successfully applied to organic mass spectra obtained with the recently developed aerosol chemical speciation monitor (ACSM) (Ng et al., 2011b), an instrument based on AMS technology and optimized for long-term sampling. The ACSM was deployed in downtown Zurich, Switzerland, from January 2011 to February 2012.

2 Materials and methods

2.1 Measurements

From January 2011 to February 2012, an ACSM (Aerodyne Research, Inc., Billerica, MA, USA) was deployed at Zurich Kaserne (Switzerland), an urban background station in the centre of a metropolitan area with about one million inhabitants. The ACSM is a compact, low-maintenance aerosol mass spectrometer designed for long-term measurements of non-refractory particulate matter with vacuum aerodynamic diameters smaller than $1 \mu\text{m}$ (NR-PM₁). The instrument is described in detail by Ng et al. (2011b) whereas the reader is referred to Jayne et al. (2000), Jimenez et al. (2003), Allan et al. (2003b), Allan et al. (2004), and Canagaratna et al. (2007) for a detailed description of the AMS technique.

Briefly, at the Kaserne station in Zurich, ambient aerosol entered the temperature controlled room and was subsequently drawn to a cyclone (model SCC 1.829 Cyclon von BGI, INC.) with a size cut-off of $2.5 \mu\text{m}$ using a flow of 5 L min^{-1} for removing coarse

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



mode particles. The resulting aerosol flow passed through a Nafion drier (MD-110-48S-4, PermaPure LLC, Toms River, NJ, USA) and a subsequent ~ 2 m long stainless steel sampling tube (6 mm o.d.) before reaching the ACSM inlet. In the ACSM, the dried aerosol particles are sampled continuously (averaging time 30 min) through a $100\ \mu\text{m}$ aperture ($\sim 90\ \text{cm}^3\ \text{min}^{-1}$), to pass through an aerodynamic lens (~ 2 torr) where they are focused into a narrow beam. The particle beam impacts a resistively heated surface at $\sim 600^\circ\text{C}$ where the non-refractory fraction is flash vaporized. The resulting gas is ionized by electron impact (70 eV) and analyzed with a quadrupole mass spectrometer. The final aerosol signal is retrieved by subtracting filtered air representing the background signal, under the same sampling conditions.

To obtain quantitative mass concentrations for the ACSM, a collection efficiency parameter (CE) needs to be applied to account for the incomplete detection of the aerosol species (Middlebrook et al., 2012). The CE is a function of the lens system, of the shape and of the bouncing of the aerosol particles on the vaporizer. The latter term was found to be influenced by several parameters, such as the mass fraction of ammonium nitrate, particle acidity, and water content (Matthew et al., 2008). Water content does not affect the present study because the particles are dried. The effects of the nitrate mass fraction and particle acidity on CE have recently been parameterized for ambient data (Middlebrook et al., 2012). However, for the present study, this parameterization underestimates the CE, as demonstrated by higher CE-corrected mass concentrations for the ACSM compared to simultaneous PM_{10} measurements by a tapered element oscillating microbalance (TEOM, FDMS 8500 Thermo Scientific). The CE will be investigated in detail in a future publication; here we assume $\text{CE} = 1$, which provides lower limit for ACSM-measured mass concentrations. Note that since the CE is applied to all measured species, changes in the CE do not affect the relative intensity of m/z within a mass spectrum and hence the ME-2 results reported in this manuscript.

The meteorological parameters and trace gases were measured with conventional instruments by the Swiss National Air Pollution Monitoring Network, NABEL (Empa, 2011). The time resolution of all these measurements was ten minutes. NO_x was

measured by the chemiluminescence technique, and carbon monoxide was monitored by non-dispersive Fourier transform infrared spectroscopy (APNA 360, Horiba, Kyoto, Japan), UV absorption spectroscopy was employed to determine the temporal variation of ozone (TEI 49C, Thermo Electron Corp., Waltham, MA) and black carbon was estimated utilizing an aethalometer AE 31 (Magee Scientific Inc.) based on measured light absorption coefficients at different wavelengths. In addition, the measured absorption coefficients at the wavelengths 470 and 880 nm were used in order to retrieve the related black carbon contributions from the traffic (BC_{traffic}) and wood burning (BC_{wb}) source (Sandradewi et al., 2008; Herich et al., 2011).

2.2 The multilinear engine (ME-2)

The organic mass spectra measured by the ACSM can be represented as a matrix \mathbf{X} where the columns j are the m/z 's and each row i represents a single mass spectrum. A frequently used method is to group the variables into distinct factors based on certain criteria. The simplest and most commonly used approach is to group the variables into two constant matrices, the so-called bilinear model, e.g., principal component analysis PCA (Wold et al., 1987) or positive matrix factorization PMF (Paatero and Tapper, 1994). The bilinear factor analytic model in matrix notation is defined as:

$$\mathbf{X} = \mathbf{GF} + \mathbf{E} \quad (1)$$

where the measured matrix \mathbf{X} is approximated by the product of \mathbf{G} and \mathbf{F} and \mathbf{E} is the model residual. p is then defined as the number of factors of the chosen model solution, i.e. the number of columns of \mathbf{G} and at the same time the number of rows of \mathbf{F} . Each column of the matrix \mathbf{G} represents the time series of a factor, whereas each row of \mathbf{F} represents the profile (e.g. mass spectrum) of this factor. The differences between the bilinear models PCA and PMF are only due to the restrictions of the models. PCA imposes orthogonality of the factors, i.e., the scalar of two different rows of \mathbf{F} is zero and does not require non-negative entries. In contrast, PMF requires non-negative entries

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



throughout **G** and **F**. This constraint makes the PMF algorithm particularly suitable, especially for chemometrics or environmental studies, where mass concentrations must be non-negative. For the ME-2 and the PMF engine the entries in **G** and **F** are fit using a least squares algorithm that minimizes iteratively the quantity Q^m , defined as the sum of the squared residuals weighted by their respective uncertainties, where the uncertainty may contain the measurement and model uncertainty:

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

Here e_{ij} are the elements of the residual matrix **E** and σ_{ij} are the measurement uncertainties for the input points ij . Data points where $\sigma_{ij} \ll e_{ij}$ constitute a large fraction of Q^m and these points will have a high impact during the model iteration. Normally this ensures that data with high signal-to-noise has a higher impact than measurements near detection limit. However, $\sigma_{ij} \ll e_{ij}$ may also occur due to dominant and rare local events or electronic noise within the measurement equipment, where neither of them should be considered by the model. To prevent the solution to be driven by few strong outliers, the model is generally run in the “robust” mode, in which pulling of the solution by outliers is reduced. At each step of the solution process, outliers are defined based on the ratio of residuals to uncertainties:

$$\text{outlier} = \left| \frac{e_{ij}}{\sigma_{ij}} \right| > \alpha \quad (3)$$

where α is the user-defined threshold value. A value of four is recommended as a defining criterion for outliers within the robust mode (Paatero, 1997). The residuals are reweighted dynamically to reduce and ideally to remove the dependence of the rate of change of Q^m with respect to the rate of change of the residuals of the outliers:

$$\frac{dQ^m}{dE_{\text{outliers}}} \cong 0 \quad (4)$$

Thus, it is highly recommended to always run the model in the “robust” mode.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



2.2.1 Normalizing Q by the expected value of Q (Q_{exp})

Normally, monitoring the total Q is not meaningful because the expected value for a “good” solution depends on the size of the data matrix and on the number of chosen factors. One therefore normalizes Q^m by the degree of freedom of the model solution which is both, a function of the size of the data matrix and of the number of factors, called Q_{exp} .

$$Q_{\text{exp}} \cong n \cdot m - p \cdot (m + n) \quad (5)$$

In the past, AMS studies reported typical values for the ratio of Q/Q_{exp} between 1 and 5. However, this is purely empirical and the absolute value cannot be used as a metric for judging model results. Instead, one should investigate the relative change of this ratio across different model runs (large changes indicate significantly decreased residuals and suggest an improved solution), to assist in choosing reasonable model solutions.

2.2.2 Rotational ambiguity of the model solutions

Solutions of the PMF algorithm may have a high degree of rotational ambiguity (Paatero et al., 2002). There are two different kinds of rotations that are allowed, namely the pure and the approximate rotations. For pure rotations, the object function Q^m does not change after the rotation:

$$\bar{G} = G \mathbf{T} \text{ and } \bar{F} = \mathbf{T}^{-1} F \quad (6)$$

where \mathbf{T} is a nonsingular matrix of dimension $p \times p$, \mathbf{T}^{-1} is its inverse, and \bar{G} and \bar{F} are the rotated matrices. The matrix multiplication of \bar{G} and \bar{F} leads again to the same product as for G and F and therefore Q^m remains unchanged. If the transformation matrix \mathbf{T} does not fulfill Eq. (6), the rotation is called an approximate rotation and Q^m changes. For the ME-2 and the PMF engine there is a user-specific parameter called

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SoFi, an Igor based interfaceF. Canonaco et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

fpeak, denoted by φ for the global control of such approximate rotations. For positive φ , elementary rotations or a series of elementary rotations are performed that increase columns of the matrix \mathbf{G} and decrease rows of the matrix \mathbf{F} while conserving mass. The opposite occurs for negative φ . However, the fpeak tool explores only rotations in one dimension of the multidimensional space and if the entries of \mathbf{G} and \mathbf{F} are positive and more than one factor is chosen then the rotational space is multidimensional and the corresponding ambiguity can be very large (e.g. for three factors, the rotational space is nine-dimensional). An advantage of the ME-2 engine compared to the PMF engine is improved rotational control, e.g. selected factors can be summed/subtracted together rather than transforming the entire matrix. Thus, the rotations can be studied in a more controlled environment. Normally, the user should explore the rotational space, on one hand since it is rare to find the environmentally optimized solution with the unrotated case, on the other hand in order to evaluate the stability of the chosen solution in the rotational space. Alternatively, to reduce the rotational ambiguity, a priori information in form of known rows of \mathbf{F} (factor profiles) or of known columns of \mathbf{G} (factor time series) can be added to the model (Paatero and Hopke, 2009). This a priori information prevents the model to rotate and provides a nearly unique model solution. Three main approaches can be exploited with the ME-2 engine, i.e. the chemical mass balance (CMB), the a value and the pulling technique (described below).

The use of a priori information at the stage of the calculation of the model solution provides a more efficient and sensitive exploration of the model space than possible with e.g. the fpeak tool in PMF (Paatero and Hopke, 2009). For this reason, we developed a user-friendly interface (Fig. S.1), SoFi (Source Finder), to facilitate the testing of the different rotational techniques available within the ME-2 model. Three different approaches were exploited, i.e., the chemical mass balance (CMB), the a value and the pulling technique, using the bilinear model based on the criterion of positive entries in \mathbf{G} and \mathbf{F} . The application of these techniques is described in detail in Paatero and Hopke, (2009) and only a brief description is presented here. In addition this interface allows running the PMF algorithm with/without the abovementioned techniques

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



for combined data sets, e.g. particle and gas-phase data, in the robust mode. This technique was first tested using a pseudo robust mode by Slowik et al. (2010). Crippa et al. (2013a) exploited this interface to perform a combined source apportionment on ambient AMS and PTR-MS (Proton Transfer Reaction Mass Spectrometry) data from the Paris campaign 2009/2010 entirely in the robust mode.

2.2.3 Fully unconstrained matrices **G** and **F**: positive matrix factorization (PMF)

For a completely unconstrained PMF run, the algorithm models the entries of **G** and **F** autonomously.

2.2.4 Fully constrained matrix **F**: chemical mass balance (CMB)

Within the spirit of the chemical mass balance, all elements of the **F** matrix, i.e. factor profiles, are set to non-negative values by the user. The entries of the matrix **G** remain variable and are evaluated by the model.

2.2.5 Constrained matrices **F/G**: *a* value approach (*a* value)

Here the elements of the **F** matrix (factor profiles) and/or of the **G** matrix (factor time series) can be constrained by the user. The user inputs one or more factor profiles (rows of **F**)/factor time series (columns of **G**) and a constraint defined by the scalar *a* that can be applied to the entire profile/time series or to individual elements of the profile/time series only. The *a* value determines the extent to which the output **F/G** is allowed to vary from the input **F / G**, according to:

$$f_{j,\text{solution}} = f_j \pm a \cdot f_j \quad (7)$$

$$g_{i,\text{solution}} = g_i \pm a \cdot g_i \quad (8)$$

where f and g represent a row and the column of the matrices \mathbf{F} and \mathbf{G} , respectively. The index j varies between 0 and the number of variables and i varies between 0 and the number of measured points.

5 The situation of the chemical mass balance described in Sect. 2.2.4 is achieved by using the scalar a set to zero for all factor profiles.

2.2.6 Constrained matrices F/G: pulling approach (pulling)

The user has the possibility to introduce pulling equations into the model that pull factor elements towards predefined anchor values (here shown for a row of the matrix \mathbf{F} only):

$$10 \quad \bar{a}_j = \bar{f}_j + r_j \quad (9)$$

In Eq. (9), a_j represents the anchor to which the model pulls the iterative value \bar{f}_j and r_j represents the residual. The anchor is a known value introduced as a priori information by the user. The pulling equations create an additional auxiliary term Q^{aux} that is added to Q^m . Thus, if pulling equations are introduced, the model will minimize the argument of Q

$$15 \quad \arg \min_{G,F}(Q) = \arg \min_{G,F}(Q^m + Q^{\text{aux}}) \quad (10)$$

The term of Q^{aux} has a conceptually similar aspect to Q^m :

$$20 \quad Q^{\text{aux}} = \sum_{k=1}^K \left(\frac{r_k}{s_k} \right)^2 \quad (11)$$

The index j from Eq. (11) has been replaced by k , since k denotes the index of the pulling equations added to the model (over many factor profiles/time series). The pulling parameter s_k specifies the softness of the pull. The smaller s_k becomes, the higher the

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



impact of Q^{aux} of the k -th pull within the iterative process. The pulling approach is a sensitive technique in that if the pulling equation is not compatible with the specific data matrix, the decrease of Q^{aux} (Eq. 11) obtained as \bar{f}_j reaches its anchor value a_j (Eq. 9) is negligible compared to a larger increase of Q^m (Eq. 2), then the pull falls off.

5 Adding known factor profiles/time series and using the pulling technique might be seen as a “soft” and self-regulating constraining technique. Generally, the user provides the total acceptable limits of Q^m , denoted as dQ . Changing these limits and the pulling parameter s_k allows to monitor the change in Q^m and its acceptability can be judged.

2.2.7 The correct solution and number of factors

10 Generally, increasing the number of factors decreases Q and the ratio of Q to Q_{exp} , due to the additional degrees of freedom provided, allowing a better fit to the measured matrix. However, these additional factors may not be physically meaningful. As a first metric in judging the correct number of factors Paatero and Tapper (1993) recommended to consider the size of the decrease of Q or Q/Q_{exp} as a function of added

15 factors, rather than its absolute value. Changes in Q or Q/Q_{exp} over different model runs of a few percentages are acceptable, if the model solution is enhanced. If the difference is however, of tens of percentages, further investigation is required.

In addition to the Q analysis, Paatero (2007) introduced another metric based on the estimation of the measurement variation explained by the factors. The explained variation EV is a dimensionless quantity that indicates how much variation in time or for each variable is explained by each factor. As an example the equation for the explained variation for the i th point in time for the factor k is given by:

$$EV_{ik} = \frac{\sum_{j=1}^m (|g_{ik} \cdot f_{kj}| / \sigma_{ij})}{\sum_{j=1}^m \left(\left(\sum_{h=1}^P |g_{ih} \cdot f_{hj}| + |e_{ij}| \right) / \sigma_{ij} \right)} \quad \text{for } k = 1, \dots, P \quad (12)$$

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Similar equations can be formulated for the unexplained variation UEV, by replacing the product $g_{ik} \cdot f_{kj}$ in the numerator with e_{ij} . Expressing the explained and the unexplained variation for a variable j as EV_{jk} results by simply replacing the sum over j in the ratio by the sum over i . If all variation is explained by the model, then $EV = 1$. According to Paatero (2007) a variable should be regarded as explained, only if the UEV for this variable is less than 25 %.

Besides these mathematical instruments, it is crucial to compare the model output with measurements or reference values that were not included in the model solution. This aids in the selection and verification of the factor solutions.

For testing the aforementioned rotational approaches, we used a data matrix containing the winter data from both 2011 and 2012 from downtown Zurich measured with the ACSM. The measurement error matrix was calculated according to the method of Allan et al. (2003a) and Allan et al. (2004), the m/z 44 related m/z 's and weak and bad m/z were downweighted as in Ulbrich et al. (2009).

3 Results

3.1 Unconstrained matrices G and F (PMF)

The first step in the source apportionment analysis was to perform the bilinear model without any a priori information in the modeled matrices (PMF) for different numbers of factors, e.g. one to ten factors to estimate an environmentally reasonable number of factors. PMF analysis of aerosol mass spectra has previously been described in detail (e.g. Lanz et al., 2007, 2010; Ulbrich et al., 2009), and similar metrics for determining the appropriate number of factors were employed in this study. Specifically, the solution was chosen based on an analysis of the dependence of Q/Q_{exp} and explained variation on the number of factors, as well as the correlation of the retrieved factor profiles and time series with reference spectra and collocated measurements. A five-factor solution was selected for further analysis. This solution is summarized below and additional

details are provided in the Supplement (Sect. 6.2). PMF solutions with a higher number of factors are not considered, due to purely mathematical splits of the factor profiles which did not represent additional sources.

The 5-factor solution consists of 3 primary factors and 2 secondary factors. The primary factors are hydrocarbon-like organic aerosol (HOA), cooking organic aerosol (COA) and biomass burning organic aerosol (BBOA), while the secondary factors are semi-volatile oxygenated organic aerosol (SV-OOA) and low-volatility oxygenated organic aerosol (LV-OOA). These factors have been identified in many previous studies and only a brief description of their most important characteristics is given here. Factor mass spectra are shown in the supplementary Fig. S2, the time series are shown in Fig. S3 and their diurnal patterns are shown in Fig. S4. The HOA spectrum shows characteristic high signal at m/z typical of aliphatic hydrocarbons (Canagaratna et al., 2004; Zhang et al., 2005). The time series and diurnal pattern of HOA are correlated with traffic-related species like NO_x , CO and $\text{BC}_{\text{traffic}}$. The COA profile is qualitatively similar to HOA but has higher m/z 55 and less m/z 57, similar to previous results (Allan et al., 2010; He et al., 2010; Slowik et al., 2010; Sun et al., 2011; Mohr et al., 2012; Crippa et al., 2013b). The diurnal cycle shows the characteristic lunch peak at noon. The BBOA profile has significantly higher contributions at m/z 60 and 73. These fragments are characteristic of sugars such as levoglucosan (Alfarra et al., 2007) which are released during wood combustion. The BBOA diurnal pattern has higher contributions at night, consistent with domestic heating activities in winter. SV-OOA and LV-OOA have significantly higher contributions at m/z 44, which is typically dominated by the CO_2^+ ion. This ion results from the thermal decomposition and fragmentation of highly oxygenated species such as organic acids (Ng et al., 2010). Compared to SV-OOA, LV-OOA typically has a higher mass fraction at m/z 44, suggesting a more aged and less volatile aerosol. Their time series correlate with the time series of secondary species like sulfate, nitrate and ammonium aerosol.

Note that while features of the factors described above can be identified from the PMF analysis, there is no unequivocal apportionment of each factor to one specific

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



source. Hence, the labeled factors in the supplementary Sect. 6.2 are only indicative. For example, the characteristic COA peak at noon is visible but rather broad between 8 a.m. and 12 a.m. The primary factors HOA and COA both contain signal from m/z 44 and m/z 60, suggesting that some biomass burning aerosol may be apportioned to these factors. These features reveal a mixed situation for the PMF factor solution. In order to retrieve an environmentally satisfactory model solution, further investigation of the multidimensional solution space is needed. One possible method is to make use of the global rotational parameter f_{peak} . Nonetheless, the outcome might not always be satisfactory, as was the case for this study. The ME-2 solver provides three alternative options for exploration:

1. Application of user-specific rotations to search for solutions that better describe the measured data matrix.
2. Adding specific pulling equations on e.g. retrieved factor profiles and/or time series from earlier unconstrained PMF solutions.
3. Utilization of a priori information, thus strongly reducing the rotational ambiguity.

This study investigates only the third approach, although the user-specific rotations and specific pulling equations are potentially valuable techniques and should be further investigated in future source apportionment studies.

3.2 Comparison of solutions constraining matrix **F**

3.2.1 Overview

Besides the PMF run using unconstrained matrices **G** and **F** described in the last section, the subsequent model runs constraining the matrix **F** or parts of it have been tested. The following runs are summarized in Table 1.

- CMB, with all five factors fixed (see Sect. 2.2.4)

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- *a* value approach (see Sect. 2.2.5) where the primary factors HOA, COA and BBOA were constrained and the other factors left free. Different *a* values were tested, i.e. with an *a* value of zero to 0.3 applied simultaneously to all constrained profiles. Note that an *a* value of zero yields a “partial CMB” model where the primary factors are fully constrained and the secondary factors are fully free.
- pulling approach (see Sect. 2.2.6) where the primary factors HOA, COA and BBOA were constrained and the other factors left free. The parameters tested were $dQ = 100$ and softness *s* between 0.01 and 0.05. Since dQ stayed invariant, the only value reported for the pulling runs in the following graphs is the softness *s*.

The primary factors (HOA, COA, BBOA) employed have been taken from Crippa et al. (2013b), a PMF analysis where the primary sources have successfully been separated, whereas the secondary factors (SV-OOA, LV-OOA) were the mean mass spectra reported by (Ng et al., 2011a).

Figure 1 shows Q/Q_{exp} for the mentioned runs. This graph and the successive ones are structured such that model runs with weaker boundaries, i.e. larger *a* values or softer pulling parameters are on the outside, while runs with stronger constraints are inside. Therefore, PMF represents the outer edge and CMB is in the center. Note that for the *a* value approach, the value in the graph indicates the lower and upper limit. For the pulling runs, the value reported stands for the softness *s* of the pull (dQ is constant at 100).

The CMB result has the worst compatibility to the measured data matrix, as shown by the higher Q/Q_{exp} ratio. This is also reflected in the plot of explained variation (EV) (Fig. 2), where the CMB run shows the highest unexplained variation (UEV). In general there is a considerable change in the distribution of EV between the different model configurations. In particular the EV for the secondary species SV-OOA and LV-OOA varies significantly. The EV for the primary species COA, HOA and BBOA stays

approximately constant for the a values runs between zero and 0.2 and the pulling runs between 0.01 and 0.02.

The mean mass concentration of all factors as a function of all model runs is shown in Fig. 3. The black rectangles in Fig. 3 (and Fig. 4) denote environmentally reasonable solutions, as discussed later. The figure shows that the CMB approach lacks in representing the measured data, due to the very inferior explained mass compared to the other models as already mentioned for Figs. 1 and 2. In addition, the continuous redistribution of the mass contributions to the five factors as the tightness of constraint changes is also apparent.

As mentioned in Sect. 2.2.7, an important criterion along with the Q/Q_{exp} and the explained variation for judging acceptable source apportionment solutions is the comparison with external information. Figure 4 lists R^2 (Pearson) for the correlations between the time series of HOA with the traffic species NO_x and $\text{BC}_{\text{traffic}}$ as well as between BBOA and $\text{BC}_{\text{wood burning}}$ and between LV-OOA and NR- PM_1 sulfate as a function of the different ME-2 configurations. Acceptable correlation values fall within the black rectangle. Attention is also drawn to the fact that, although R^2 (Pearson) for BBOA with BC_{wb} is highest for the pulling model run with $s = 0.05$, this run is still rejected due to the other degraded correlations, in particular that for the traffic factor HOA. Further support for identifying the model solutions within the rectangles as environmentally reasonable is provided by the analysis of the diurnal cycle of HOA and COA (Sect. 2.2.3 and in the Supplement Sect. 6.4), where the expected diurnal patterns for the traffic and cooking factors are retrieved. The mean absolute and relative mass concentrations for all selected solutions are shown in Table 2. The high standard deviation for BBOA indicates that the apportionment of this species is more uncertain, while COA and HOA show very little variation.

3.2.2 Comparison of factor profiles

Figure 5 shows the factor profiles of all environmentally reasonable model solutions. Models based on a values and pulling equations are shown in the left and right column,

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



respectively. Different constraint levels are shown with different symbols. As noted in the previous section, the selected solutions lie in a relatively small range of a values (0–0.2) and strong pulling strengths (0.01–0.02).

As seen from Fig. 5 there is no significant variation of the primary factor profiles HOA, COA and BBOA as a function of the different model runs, due to the imposed tight constraint. In contrary, the unconstrained factors, especially SV-OOA, show more model-dependent variation. In particular, the high variation of m/z 43 of SV-OOA highlights the high uncertainty in apportioning this variable. Figure 3 highlights the fact that moving from a constrained run to a less constrained situation apportions less mass to LV-OOA and more to SV-OOA as well as to the three primary factors HOA, COA and in particular BBOA. This is evidenced in the factor profile with the increase of m/z 43 in SV-OOA for less constrained model runs.

3.2.3 Comparison of factor time series

Diurnal cycles for the environmentally reasonable model solutions are shown in Fig. 6. In addition, NO_x and $\text{BC}_{\text{traffic}}$ are plotted together with HOA, while BC_{wb} is plotted with BBOA.

The diurnal trends of HOA, NO_x , and $\text{BC}_{\text{traffic}}$ are highly correlated. The diurnal cycle of the cooking factor COA manifests a strong peak during meal activities, similar to COA factors found in other source apportionment studies conducted on NR-PM_1 data in other cities such as Barcelona (Mohr et al., 2012) or Paris (Crippa et al., 2013b). BBOA is correlated with $\text{BC}_{\text{wood burning}}$ and is highest at night, consistent with domestic heating activities and previous measurements in Zurich (Lanz et al., 2008). The diurnal cycle of SV-OOA is anticorrelated with temperature, suggesting that the factor represents semivolatile material which is influenced by temperature-driven partitioning whereas the diurnal cycle of LV-OOA does not show strong diurnal trends.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4 Discussion

4.1 Uncertainty in bilinear model results

As discussed in the previous section, the bilinear model analysis using constraints in **F** solved by the ME-2 engine yield a set of environmentally reasonable solutions which definitely improve the source apportionment. Note that the fully unconstrained PMF run did not even fall into the range of environmentally acceptable solutions. While the constrained ME-2 solutions have many features in common, the reported profiles and mass concentrations differ (see Table 2, Figs. 5 and 6). This variation reflects the model uncertainty for the bilinear system. Rotational techniques, such as the *a* value, pulling equations or individual rotations as well as the frequently used global rotational tool *f*peak are valuable tools for the quantitative assessment of the bilinear model uncertainty.

An additional source of uncertainty in the model results derives from the selection of the anchoring factor profiles and the magnitude of their constraints. The effects of the latter are evident from the CMB result (see Sect. 2.2.1). In particular, the mass contribution of the SV-OOA factor was almost negligible. This occurred because the profiles were completely fixed, leaving the model with no chance to adapt the SV-OOA profile to the measured data. However, a semi-volatile fraction has been modeled for the *a* value, pulling and even in the PMF approaches. This underlines the fact that if an anchor profile is too tight, a legitimate factor may be excluded from the model result.

The effect of choosing various factor profiles is highlighted by the comparison with the results of Lanz et al. (2008), in which winter AMS data from Zurich was analyzed using a constrained HOA factor profile with an *a* value of 60 %. This value is considerably higher than the maximum *a* value of 20 % selected in this study. The difference in the required *a* value for these two studies is likely due to the choice of the HOA factor profile. In the present study, the employed HOA anchor profile has a non-zero contribution in *m/z* 44 (1.4 %). Most probably, a large *a* value would lead to a mixing situation based on the variable *m/z* 44 which is avoided by using only smaller *a* values. This

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



was not the case in the constrained factor used in Lanz et al. (2008) as no m/z 44 was present at all (0%).

However, testing of the influence of different anchoring profiles and the tightness of their constraints, before a solution lacks in environmental interpretation, is ongoing and will be methodically discussed in a future study.

4.2 Recommendations for ME-2 analysis of aerosol mass spectra

Currently, there is very little experience regarding the inclusion of a priori information in the source apportionment for aerosol mass spectrometer data. The scope of this work is to facilitate the source apportionment in this respect by testing in a semi-automatic way different rotational tools of ME-2 with the graphical user interface SoFi.

Generally, the user can anchor factor information (profiles or time series) and easily vary the tightness of the constraint, while monitoring the various criteria for evaluation of a solution (see Sect. 2.2.7). Based on the experience gained in this study, for NR-PM₁ source apportionments it is recommended to constrain the primary factors (HOA, COA, BBOA), whenever the PMF run reveals indications for such sources in the PMF model result and or in the corresponding residuals. The secondary species can in first runs stay unconstrained, since they do not represent specific emissions rather span the range of aging processes in a specific location during the measurement time. Thus, it is difficult to match this evolution with auxiliary data. However, this topic is under study and more information will be provided in future.

The user should in any case perform sensitivity tests on the tightness of the constraint (a value or pulling parameters), to assess the environmentally reasonable solutions and present this range rather than only a single solution. As stated in Sect. 3.1, these parameters highly depend on the anchoring profiles employed and thus no limits for these values can be suggested at this stage. However, in general the increase of Q/Q_{exp} should not be larger than a few percent points as stated already in Paatero and Tapper (1993).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Crippa et al. (2013) are performing the source apportionment with the ME-2 engine on the AMS EUCAARI data measured in 2008/2009, with the aim to standardize the method of source apportionment on NR-PM₁ data with ME-2 to some extend.

5 Conclusions

Source apportionment using the bilinear model as implemented through the multilinear engine (ME-2) was successfully applied to non-refractory organic aerosol (OA) mass spectra measured during winter 2011 and 2012 in Zurich, Switzerland using the aerosol chemical speciation monitor ACSM. The solutions were analyzed exploiting the newly developed source finder SoFi. The selected solution consists of two secondary factors and three primary factors. The secondary factors are a semi-volatile oxidized OA (SV-OOA) and a low-volatility oxidized OA (LV-OOA). The three primary factors are traffic-related hydrocarbon-like OA (HOA), cooking OA (COA) and biomass burning OA (BBOA).

Different rotational approaches were investigated employing the ME-2 engine. The tested implementations consisted of (unconstrained matrix **F** and **G**) positive matrix factorization (PMF), (fully constrained) chemical mass balance (CMB), and partially constrained models using the *a* value parameter and pulling equations for the matrix **F**. In addition, for the *a* value and pulling model runs a sensitivity test on the constrained profiles was performed. This allowed to identify the set of environmentally reasonable solutions.

Moreover, such analysis provides insight into the robustness and uncertainty of the bilinear model solution, e.g. the primary factor BBOA and the secondary semivolatile factor SV-OOA show the highest variability across models (implying the highest model uncertainty), while COA and HOA have the least variability (smallest model uncertainty).

Finally, some recommendations for future NR-PM₁ source apportionments exploiting ME-2 are reported.

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Supplementary material related to this article is available online at:
[http://www.atmos-meas-tech-discuss.net/6/6409/2013/
amtd-6-6409-2013-supplement.pdf](http://www.atmos-meas-tech-discuss.net/6/6409/2013/amtd-6-6409-2013-supplement.pdf).

Acknowledgements. The ACSM measurements were supported by the Swiss Federal Office for the Environment (FOEN). J. Slowik acknowledges support from the Swiss National Science Foundation (SNF) through the “Ambizione” program. The authors would like to thank P. Paatero for the valuable comments during the analysis of the data and especially to the Environmental group of the Swiss Federal Laboratories for Materials and Testing (EMPA) for their support. Special thanks are also owed to Michele Canonaco for a critical reading of this manuscript.

References

Alfarra, M. R., Prevot, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D., Mohr, M., and Baltensperger, U.: Identification of the mass spectral signature of organic aerosols from wood burning emissions, *Environ. Sci. Technol.*, 41, 5770–5777, 2007.

Allan, J. D., Alfarra, M. R., Bower, K. N., Williams, P. I., Gallagher, M. W., Jimenez, J. L., McDonald, A. G., Nemitz, E., Canagaratna, M. R., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer: 2. measurements of fine particulate chemical composition in two UK cities, *J. Geophys. Res.-Atmos.*, 108, 4091, doi:10.1029/2002JD002359, 2003a.

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer: 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.-Atmos.*, 108, 4090, doi:10.1029/2002JD002358, 2003b.

Allan, J. D., Delia, A. E., Coe, H., Bower, K. N., Alfarra, M. R., Jimenez, J. L., Middlebrook, A. M., Drewnick, F., Onasch, T. B., Canagaratna, M. R., Jayne, J. T., and Worsnop, D. R.: A generalised method for the extraction of chemically resolved mass spectra from Aerodyne aerosol mass spectrometer data, *J. Aerosol Sci.*, 35, 909–922, 2004.

Allan, J. D., Williams, P. I., Morgan, W. T., Martin, C. L., Flynn, M. J., Lee, J., Nemitz, E., Phillips, G. J., Gallagher, M. W., and Coe, H.: Contributions from transport, solid fuel burning

AMTD

6, 6409–6443, 2013

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and cooking to primary organic aerosols in two UK cities, *Atmos. Chem. Phys.*, 10, 647–668, doi:10.5194/acp-10-647-2010, 2010.

Amato, F. and Hopke, P. K.: Source apportionment of the ambient PM_{2.5} across St. Louis using constrained positive matrix factorization, *Atmos. Environ.*, 46, 329–337, 2012.

5 Canagaratna, M. R., Jayne, J. T., Ghertner, D. A., Herndon, S., Shi, Q., Jimenez, J. L., Silva, P. J., Williams, P., Lanni, T., Drewnick, F., Demerjian, K. L., Kolb, C. E., and Worsnop, D. R.: Chase studies of particulate emissions from in-use New York City vehicles, *Aerosol Sci. Tech.*, 38, 555–573, 2004.

10 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., and Worsnop, D. R.: Chemical and microphysical characterization of ambient aerosols with the Aerodyne aerosol mass spectrometer, *Mass Spectrom. Rev.*, 26, 185–222, doi:10.1002/mas.20115, 2007.

15 Crippa, M., DeCarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., Poulain, L., Freutel, F., Sciare, J., Cozic, J., Di Marco, C. F., Elsasser, M., Nicolas, J. B., Marchand, N., Abidi, E., Wiedensohler, A., Drewnick, F., Schneider, J., Borrmann, S., Nemitz, E., Zimmermann, R., Jaffrezo, J.-L., Prévôt, A. S. H., and Baltensperger, U.: Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris, *Atmos. Chem. Phys.*, 13, 961–981, doi:10.5194/acp-13-961-2013, 2013a.

20 Crippa, M., Canonaco, F., Prevot, A. S. H., and Baltensperger, U.: A compilation of organic aerosol components for AMS datasets across Europe using a newly developed ME-2 based source apportionment strategy, in preparation, 2013b.

Empa: Technischer Bericht zum Nationalen Beobachtungsnetz für Luftfremdstoffe (NABEL), available at: <http://www.empa.ch> (last access: 11 July 2013), 2011.

25 Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, Th. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, *Atmos. Chem. Phys.*, 9, 5155–5236, doi:10.5194/acp-9-5155-2009, 2009.

30 He, L.-Y., Lin, Y., Huang, X.-F., Guo, S., Xue, L., Su, Q., Hu, M., Luan, S.-J., and Zhang, Y.-H.: Characterization of high-resolution aerosol mass spectra of primary organic aerosol emis-

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



sions from Chinese cooking and biomass burning, *Atmos. Chem. Phys.*, 10, 11535–11543, doi:10.5194/acp-10-11535-2010, 2010.

Herich, H., Hueglin, C., and Buchmann, B.: A 2.5 year's source apportionment study of black carbon from wood burning and fossil fuel combustion at urban and rural sites in Switzerland, *Atmos. Meas. Tech.*, 4, 1409–1420, doi:10.5194/amt-4-1409-2011, 2011.

IPCC: IPCC Fourth Assessment Report: The Physical Science Basis, Working Group I, Final Report, Geneva, Switzerland, 2007.

Jayne, J. T., Leard, D. C., Zhang, X. F., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R.: Development of an aerosol mass spectrometer for size and composition analysis of submicron particles, *Aerosol Sci. Tech.*, 33, 49–70, 2000.

Jimenez, J. L., Jayne, J. T., Shi, Q., Kolb, C. E., Worsnop, D. R., Yourshaw, I., Seinfeld, J. H., Flagan, R. C., Zhang, X. F., Smith, K. A., Morris, J. W., and Davidovits, P.: Ambient aerosol sampling using the Aerodyne aerosol mass spectrometer, *J. Geophys. Res.-Atmos.*, 108, 8425, doi:10.1029/2001JD001213, 2003.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimojo, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, *Science*, 326, 1525–1529, 2009.

Khaykin, S. M., Pommereau, J.-P., and Hauchecorne, A.: Impact of land convection on the thermal structure of the lower stratosphere as inferred from COSMIC GPS radio occultations, *Atmos. Chem. Phys. Discuss.*, 13, 1–31, doi:10.5194/acpd-13-1-2013, 2013a.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prévôt, A. S. H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, *Atmos. Chem. Phys.*, 7, 1503–1522, doi:10.5194/acp-7-1503-2007, 2007.

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Lanz, V. A., Alfara, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., Szidat, S., Wehrl, M. N., Wacker, L., Weimer, S., Caseiro, A., Puxbaum, H., and Prevot, A. S. H.: Source attribution of submicron organic aerosols during wintertime inversions by advanced factor analysis of aerosol mass spectra, *Environ. Sci. Technol.*, 42, 214–220, doi:10.1021/es0707207, 2008.
- Lanz, V. A., Prévôt, A. S. H., Alfara, M. R., Weimer, S., Mohr, C., DeCarlo, P. F., Gianini, M. F. D., Hueglin, C., Schneider, J., Favez, O., D'Anna, B., George, C., and Baltensperger, U.: Characterization of aerosol chemical composition with aerosol mass spectrometry in Central Europe: an overview, *Atmos. Chem. Phys.*, 10, 10453–10471, doi:10.5194/acp-10-10453-2010, 2010.
- Matson, P., Lohse, K. A., and Hall, S. J.: The globalization of nitrogen deposition: consequences for terrestrial ecosystems, *Ambio*, 31, 113–119, 2002.
- Matthew, B. M., Middlebrook, A. M., and Onasch, T. B.: Collection efficiencies in an Aerodyne aerosol mass spectrometer as a function of particle phase for laboratory generated aerosols, *Aerosol Sci. Tech.*, 42, 884–898, doi:10.1080/02786820802356797, 2008.
- Middlebrook, A. M., Bahreini, R., Jimenez, J. L., and Canagaratna, M. R.: Evaluation of composition-dependent collection efficiencies for the Aerodyne aerosol mass spectrometer using field data, *Aerosol Sci. Tech.*, 46, 258–271, doi:10.1080/02786826.2011.620041, 2012.
- Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Peñuelas, J., Jiménez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prévôt, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, *Atmos. Chem. Phys.*, 12, 1649–1665, doi:10.5194/acp-12-1649-2012, 2012.
- Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data, *Environ. Sci. Technol.*, 45, 910–916, doi:10.1021/es102951k, 2011a.
- Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P. L., Onasch, T. B., Sueper, D., Worsnop, D. R., Zhang, Q., Sun, Y. L., and Jayne, J. T.: An aerosol chemical speciation monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol, *Aerosol Sci. Tech.*, 45, 770–784, doi:10.1080/02786826.2011.560211, 2011b.

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Paatero, P.: Least squares formulation of robust non-negative factor analysis, *Chemometr. Intell. Lab.*, 37, 23–35, 1997.
- Paatero, P.: The multilinear engine – a table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model, *J. Comput. Graph. Stat.*, 8, 854–888, 1999.
- Paatero, P.: User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, University of Helsinki, Finland, 2007.
- Paatero, P. and Tapper, U.: Analysis of different modes of factor analysis as least squares fit problems, *Chemometr. Intell. Lab.*, 18, 183–194, 1993.
- Paatero, P. and Tapper, U.: Positive matrix factorization – a nonnegative factor model with optimal utilization of error-estimates of data values, *Environmetrics*, 5, 111–126, 1994.
- Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, *Chemometr. Intell. Lab.*, 60, 253–264, 2002.
- Paatero, P. and Hopke, P. K.: Rotational tools for factor analytic models, *J. Chemometr.*, 23, 91–100, 2009.
- Peng, R. D., Dominici, F., Pastor-Barriuso, R., Zeger, S. L., and Samet, J. M.: Seasonal analyses of air pollution and mortality in 100 US cities, *Am. J. Epidemiol.*, 161, 585–594, doi:10.1093/aje/kwi075, 2005.
- Reche, C., Viana, M., Amato, F., Alastuey, A., Moreno, T., Hillamo, R., Teinila, K., Saarnio, K., Seco, R., Penuelas, J., Mohr, C., Prevot, A. S. H., and Querol, X.: Biomass burning contributions to urban aerosols in a coastal Mediterranean City, *Sci. Total. Environ.*, 427, 175–190, doi:10.1016/j.scitotenv.2012.04.012, 2012.
- Sandradewi, J., Prevot, A. S. H., Szidat, S., Perron, N., Alfarra, M. R., Lanz, V. A., Weingartner, E., and Baltensperger, U.: Using aerosol light absorption measurements for the quantitative determination of wood burning and traffic emission contributions to particulate matter, *Environ. Sci. Technol.*, 42, 3316–3323, doi:10.1021/es702253m, 2008.
- Slowik, J. G., Vlasenko, A., McGuire, M., Evans, G. J., and Abbatt, J. P. D.: Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site, *Atmos. Chem. Phys.*, 10, 1969–1988, doi:10.5194/acp-10-1969-2010, 2010.
- Sun, Y.-L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W.-N., Bae, M.-S., Hung, H.-M., Hogrefe, O., Frank, B., Rattigan, O. V., and Lin, Y.-C.: Characterization of the sources and processes of organic and inorganic aerosols in New York city with a high-resolution time-of-

flight aerosol mass spectrometer, *Atmos. Chem. Phys.*, 11, 1581–1602, doi:10.5194/acp-11-1581-2011, 2011.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data, *Atmos. Chem. Phys.*, 9, 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.

Watson, J. G.: Visibility: science and regulation, *JAPCA J. Air Waste Manage.*, 52, 628–713, doi:10.1080/10473289.2002.10470813, 2002.

Wold, S., Esbensen, K., and Geladi, P.: Principal component analysis, *Chemometr. Intell. Lab.*, 2, 37–52, doi:10.1016/0169-7439(87)80084-9, 1987.

Zhang, Q., Alfarra, M. R., Worsnop, D. R., Allan, J. D., Coe, H., Canagaratna, M. R., and Jimenez, J. L.: Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry, *Environ. Sci. Technol.*, 39, 4938–4952, doi:10.1021/es048568l, 2005.

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

AMTD

6, 6409–6443, 2013

SoFi, an Igor based interface

F. Canonaco et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



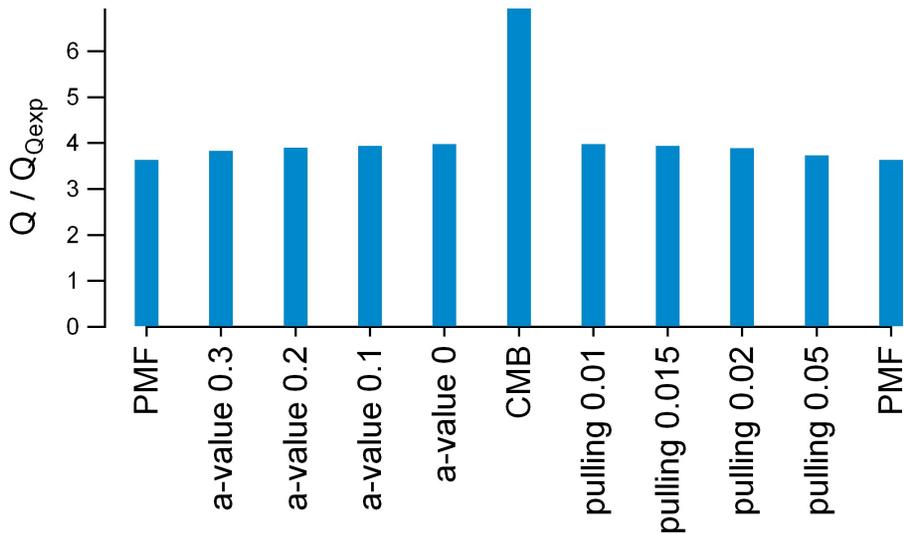


Fig. 1. Values of Q/Q_{exp} for different model runs. The CMB run, for which all factor profiles have been fixed, is almost double the other values.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



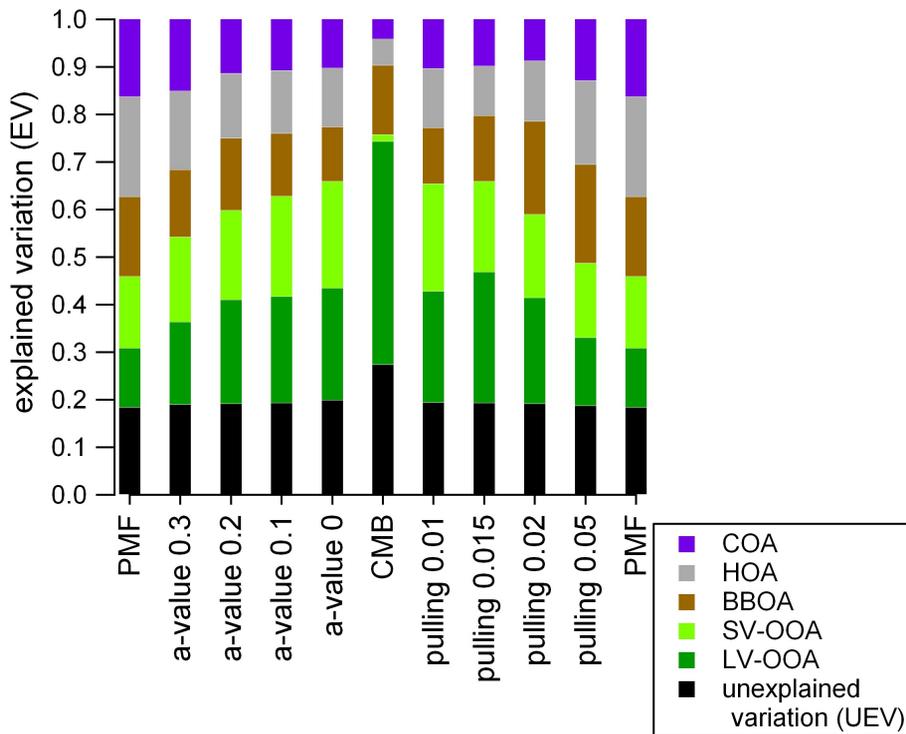


Fig. 2. Explained variation (EV) for each factor and total unexplained variation (UEV) over the different model runs. The abbreviations in the legend refer to the retrieved components, i.e. COA (cooking organic aerosol), HOA (hydrocarbon-like organic aerosol), BBOA (biomass burning organic aerosol), SV-OOA (semi-volatile organic aerosol) and LV-OOA (low-volatility organic aerosol). The same abbreviations are used through all figures of this manuscript.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



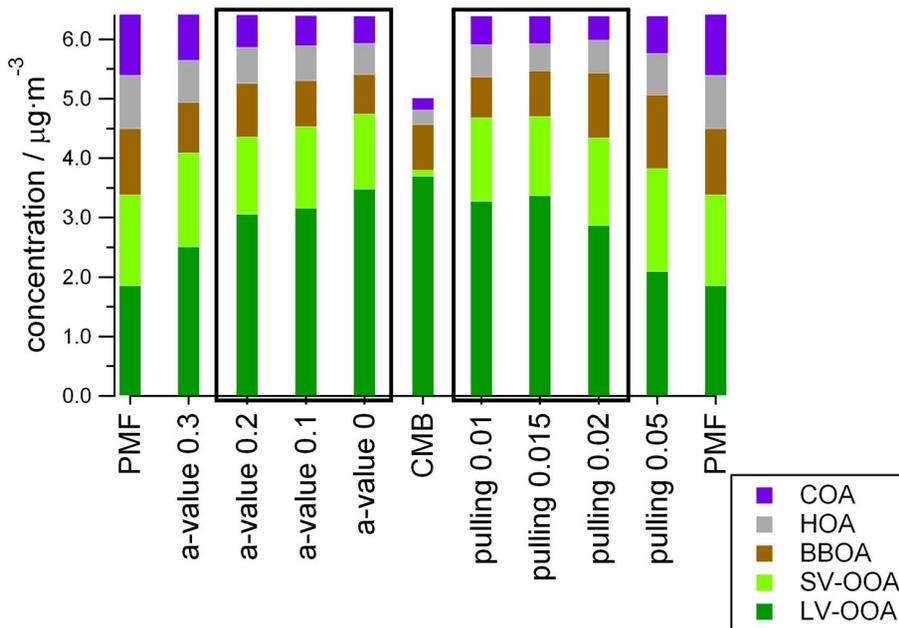


Fig. 3. Mean mass concentrations for the five factors for the model runs. The results reported within the two rectangles represent environmentally reasonable solutions.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



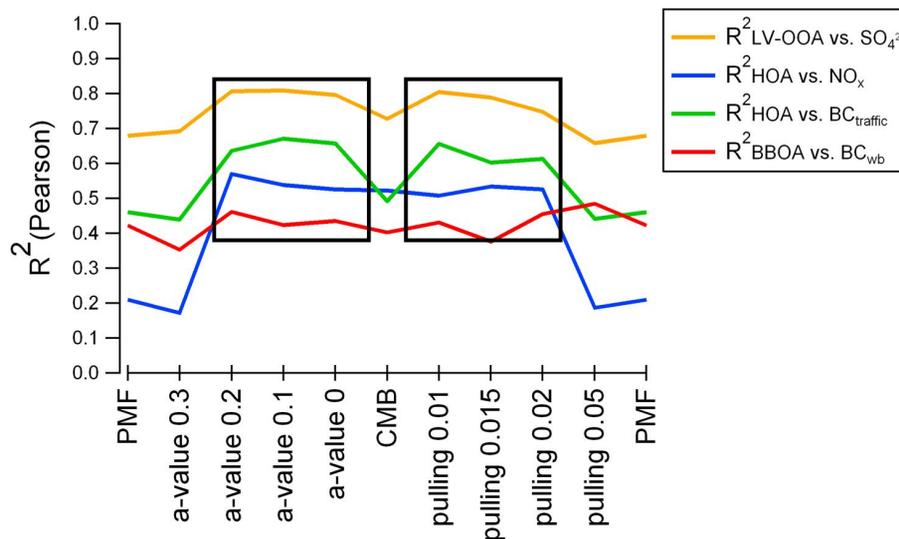


Fig. 4. Correlations R^2 (Pearson) between the time series of selected factors and the time series of external data as a function of the model runs. The results reported within the two rectangles represent environmentally reasonable solutions.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


SoFi, an Igor based interface

F. Canonaco et al.

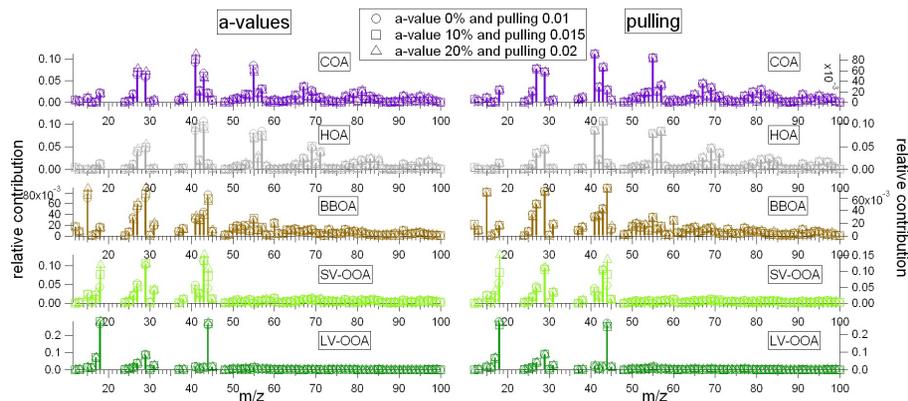


Fig. 5. Factor profiles of the five factor solutions HOA, COA, BBOA, SV-OOA and LV-OOA as a function of a value (left panel) and pulling strength s for solutions classified as environmentally reasonable. Different solutions are represented by different symbols, with circles and triangles being the most and least constrained, respectively.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


SoFi, an Igor based interface

F. Canonaco et al.

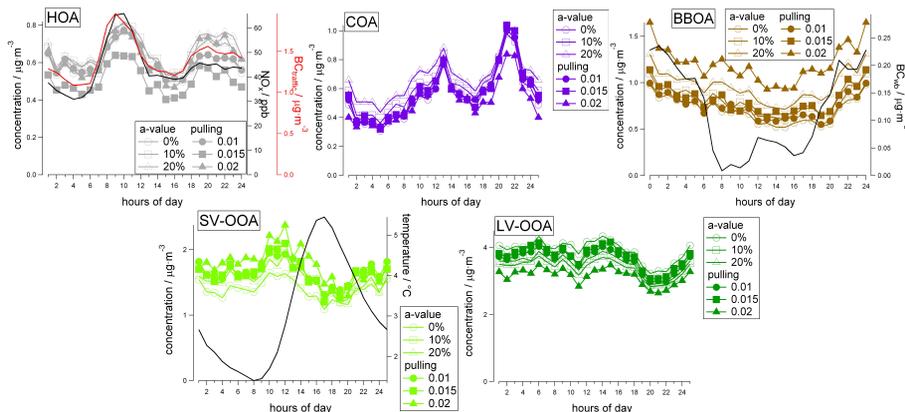


Fig. 6. Mean hourly factor mass concentrations for solutions classified as environmentally reasonable. Open and closed symbols denote a value and pulling solutions, respectively. Symbol shapes indicate the level of constraint, with circles being the most constrained and triangles the least. NO_x , $\text{BC}_{\text{traffic}}$ and $\text{BC}_{\text{wood burning}}$ are shown for comparison.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

