

Contents lists available at ScienceDirect

Science of the Total Environment

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

Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results

Steven G. Brown^a, Shelly Eberly^b, Pentti Paatero^c, Gary A. Norris^{d,*}^a Sonoma Technology, Inc. 1455 N. McDowell Blvd, Suite D, Petaluma, CA 94954, United States^b Geometric Tools, LLC, 2909 181st Ave NE, Redmond, WA 98052, United States^c University of Helsinki, Dept. of Physics, Rikalan tie 6, 00970, Helsinki, Finland^d U.S. EPA, Office of Research and Development, Research Triangle Park, NC, United States

HIGHLIGHTS

- We provide examples and interpretation of new error estimation methods in EPA PMF.
- Multiple error estimation methods improve understanding of rotational uncertainty.
- Multiple error estimation methods provide range of uncertainty in factor profiles.

ARTICLE INFO

Article history:

Received 5 August 2014

Received in revised form 9 January 2015

Accepted 11 January 2015

Available online 13 March 2015

Editor: Eddy Y. Zeng

Keywords:

Receptor modeling

Air pollution

Water pollution

Positive matrix factorization

EPA PMF

ABSTRACT

The new version of EPA's positive matrix factorization (EPA PMF) software, 5.0, includes three error estimation (EE) methods for analyzing factor analytic solutions: classical bootstrap (BS), displacement of factor elements (DISP), and bootstrap enhanced by displacement (BS-DISP). These methods capture the uncertainty of PMF analyses due to random errors and rotational ambiguity. To demonstrate the utility of the EE methods, results are presented for three data sets: (1) speciated PM_{2.5} data from a chemical speciation network (CSN) site in Sacramento, California (2003–2009); (2) trace metal, ammonia, and other species in water quality samples taken at an inline storage system (ISS) in Milwaukee, Wisconsin (2006); and (3) an organic aerosol data set from high-resolution aerosol mass spectrometer (HR-AMS) measurements in Las Vegas, Nevada (January 2008). We present an interpretation of EE diagnostics for these data sets, results from sensitivity tests of EE diagnostics using additional and fewer factors, and recommendations for reporting PMF results. BS-DISP and BS are found useful in understanding the uncertainty of factor profiles; they also suggest if the data are over-fitted by specifying too many factors. DISP diagnostics were consistently robust, indicating its use for understanding rotational uncertainty and as a first step in assessing a solution's viability. The uncertainty of each factor's identifying species is shown to be a useful gauge for evaluating multiple solutions, e.g., with a different number of factors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Multivariate receptor modeling tools are widely used for examining patterns in environmental data. Positive matrix factorization (PMF) is one such model and data analysis tool that decomposes a matrix of speciated sample data into two matrices—factor contributions and factor profiles—in order to understand the factors or sources impacting the speciated sample data (Poirot et al., 2001; Paatero et al., 2003, 2014; Reff et al., 2007; Ulbrich et al., 2009; Zhang et al., 2011; Hopke,

2008; Kim and Hopke, 2007; McCarthy et al., 2013; Brown et al., 2012). The two most common PMF programs are PMF2 and ME-2, in addition to EPA PMF, a freely available graphical user interface (GUI) developed by the U.S. Environmental Protection Agency (EPA) that uses the ME-2 program. The detailed methods of these programs have been documented elsewhere (Paatero, 1997; Paatero and Tapper, 1994), and are summarized below.

A speciated data set can be viewed as a data matrix \mathbf{X} of dimensions n by m , in which n samples and m chemical species were measured. Rows and columns of \mathbf{X} and of related matrices are indexed by i and j , respectively. The goal of multivariate receptor modeling, for example with PMF, is to identify the number of factors p , the species profile f of each factor, and the amount of mass g contributed by each factor to

* Corresponding author.

each individual sample that solves the chemical mass balance between measured species concentrations and factor profiles (Eq. (1)):

$$x_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij} = c_{ij} + e_{ij} \quad (1)$$

where e_{ij} is the residual for each sample/species and c_{ij} is the modeled solution of x_{ij} . Entire matrices are denoted by capital bold-face letters. Columns of the factor contribution matrix \mathbf{G} may be denoted by g_{ik} , and similarly rows of factor profile matrix \mathbf{F} denoted by f_{kj} .

In PMF, factor elements are constrained so that no sample can have a significantly negative contribution. PMF allows each data value to be individually weighted. This feature allows analysts to adjust the influence of each data point, depending on the confidence in the measurement. The PMF solution minimizes the object function Q (Eq. (2)) via a conjugate gradient algorithm, based upon the estimated data uncertainties (or adjusted data uncertainties) u_{ij} .

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left[\frac{X_{ij} - \sum_{k=1}^p g_{ik} f_{kj}}{U_{ij}} \right]^2 \quad (2)$$

In some scientific disciplines, data rows are normalized (e.g., to sum = 1) before the data matrix is analyzed by PMF models. In such a scaled matrix the errors of data values are not uncorrelated. However, such normalization does not require any special attention before fitting the data with PMF: For any row i , the fitted values g_{ik} take care of any error in normalization (i.e., of the correlated part of the error) on row i , so that fitting the values f_{kj} only deals with the original uncorrelated errors.

ME-2 is the underlying program used to solve the PMF problem in EPA PMF (Norris et al., 2014), the GUI that feeds the data and user specifications to ME-2. ME-2 then performs the iterations via the conjugate gradient algorithm until convergence to a minimum Q value is obtained. The minimum Q may be global or local; users can attempt to determine whether Q values are global or local by using different starting points for the iterative process and comparing the minimum Q values reached. Output from ME-2 is then fed back through EPA PMF and formatted for users to use and visualize. The latest version of EPA PMF, version 5.0, includes a revised signal/noise (S/N) calculation and new error estimation (EE) methods. S/N details are in the EPA PMF Fundamentals and User Guide (Norris et al., 2014), while details of EE methods are summarized here and presented fully in Paatero et al. (2014). In conjunction with the new EPA PMF version, this manuscript demonstrates the utility of the multiple EE methods over multiple case studies (speciated fine particulate matter [PM_{2.5}] data from Sacramento; water quality data from Milwaukee; and high-time-resolution aerosol mass spectrometer [HR-AMS] data from Las Vegas) and provides recommendations for reporting PMF results. These case study examples using different ambient data sets are presented to complement the theory behind new EE methods in EPA PMF and ME-2, and synthetic data analyses presented in a companion paper (Paatero et al., 2014). Together, they show the use of multiple EE methods.

2. Methods

2.1. Error estimation in EPA PMF and ME-2

As described in detail in Paatero et al. (2014) variability in the PMF solution has traditionally been assessed via bootstrapping (BS), where multiple PMF solutions are generated by using a series of data sets that are resampled versions of the original data set. EPA PMF performs BS by randomly selecting non-overlapping blocks of consecutive samples (block size is suggested by the software or by the user) and creating

a new input data file of the selected samples, with the same dimensions (i.e., number of samples and number of species) as the original data set. PMF is then run on the new resampled data set, and each BS factor is mapped to a base run factor by comparing factors' contributions (\mathbf{G} matrix) for those samples included in the resampled data set. The BS factor is assigned to the base factor with which the BS factor has the highest uncentered correlation, above a user-specified threshold. If no base factors have a correlation above the threshold for a given BS factor, that factor is considered “unmapped.” If more than one BS factor from the same run is correlated with the same base factor, they will all be mapped to that base factor. This process is repeated for as many BS runs as the user specifies. In this manner, an understanding of the uncertainty of the apportionment of each species in each factor is found.

EPA PMF and ME-2 now have two additional EE methods: displacement (DISP) analysis and bootstrapping with displacement (BS-DISP). The three methods are complementary and can be used to understand the uncertainty of a PMF solution:

1. BS intervals include effects from random errors and partially include effects of rotational ambiguity. If the user misspecifies data uncertainties, this modeling error usually has minimal impact in BS results.
2. DISP intervals include effects of random errors and rotational ambiguity. If the user specifies too-small data uncertainties, then this modeling error results in DISP intervals that are too short. Specifying too-large data uncertainties, e.g., when a species is made weak intentionally, results in DISP intervals that are too long.
3. BS-DISP intervals include effects of random errors and rotational ambiguity. If the user misspecifies data uncertainty, BS-DISP results are more robust than for DISP since the DISP phase of BS-DISP does not displace as strongly as DISP by itself.

With DISP, each fitted element in \mathbf{F} in the base PMF solution is “displaced” from its fitted value far enough so that Q increases by a predetermined amount called dQ^{max} . Each such extended displacement is interpreted as the upper or lower interval estimate of the perturbed variable, thus yielding an uncertainty estimate for each species in each factor profile. The uncertainty estimate consists of both data uncertainties (data noise) and rotational ambiguity. Depending on the data set in question, one or the other may be more significant. In EPA PMF, only “strong” species are actively displaced in DISP. Since “weak” variables have their uncertainty increased by a factor of 3 and DISP intervals are directly related to species uncertainties, these result in large error estimates for weak species. Hence, only “strong” species are displaced in DISP.

BS-DISP was developed to combine BS's strength of robustness to data errors and DISP's strength of capturing rotational uncertainty. BS-DISP is a combination of BS and DISP methods in which each resampled data set is decomposed into profile and contribution matrices and then fitted elements in \mathbf{F} are displaced. The collection of all results from the process of resampling, decomposing, and displacing is then summarized to derive uncertainty estimates. Intuitively, this process may be viewed as follows: each BS resample results in one solution that is randomly located within the rotationally accessible space. Then, the DISP analysis determines an approximation for the rotationally accessible space around that solution. Taken together, all the approximations of rotationally accessible spaces for randomly located solutions represent both the random uncertainty and the rotational uncertainty for the modeled solution to the complete data set.

2.2. EE results and diagnostics in EPA PMF

With three EE methods, there is copious output generated by ME-2 and EPA PMF for evaluating PMF solutions. In the results presented here, all diagnostics and data are readily available in the output of EPA PMF. For DISP, the focus is on the number of swaps at the lowest dQ^{max} level and the percent change in Q ($\%dQ$), where swaps occur if factors change so much that they exchange identities, indicating a

“not-well-defined (NWD) solution” (Paatero et al., 2014). In addition, the minimum and maximum values for each species that were reached during displacement for the lowest dQ^{max} level are used as the DISP uncertainty range for each factor profile. In DISP, only “strong” species are used, since DISP is very sensitive to higher data uncertainties, e.g., those that occur when a user makes a species “weak” in EPA PMF, where data uncertainty is tripled. With BS, the fraction of BS runs mapped to the base solution by factor is assessed to understand the reproducibility of the base solution. The 5th and 95th percentiles are used as the BS uncertainty range for each factor profile. All species are used in BS, since the input data are being resampled. BS-DISP diagnostics include the overall fraction of BS-DISP runs that met all BS-DISP criteria, including factor swaps, decrease in Q , and lack of mapping via BS. Ideally, the identifying species for each factor, e.g., silicon and calcium for a soil factor, should be activated in DISP and BS-DISP. For DISP and BS-DISP, results for actively displaced species are considered to be the most reliable; other species, left passive in DISP or BS-DISP, may have error intervals that are smaller than if they were actively displaced.

Since species concentrations can often span multiple orders of magnitude, it can be difficult to compare EE results among solutions or species using concentration units. As presented by Paatero et al. (2014), the interval ratio can be used to compare results among species and is defined as the length of the given species' EE interval divided by the interval midpoint. With this method, the maximum interval ratio is 2, and is indicative of more uncertain results. For DISP, endpoints of the uncertainty interval for a specific F factor element are the minimum and maximum values for that factor element observed in all displacements and are output by ME-2 in the DISPres file, with one file for each dQ^{max} . For BS, the endpoints of the uncertainty interval for a factor element are the 5th and 95th percentile values for that factor element from all BS resamples, calculated by EPA PMF using the PMF_ab_boot.dat file output from ME-2. For BS-DISP, each BS resample is displaced and minimum and maximum values are calculated for each factor element as described for DISP. EPA PMF then calculates the 5th percentile of the minimums and 95th percentile of the maximums, which are used as the lower and upper bounds for BS-DISP EE; the minimums and maximums for each BS-DISP iteration are output by ME-2 in the BSDISPres file, with one file for each dQ^{max} value. In the examples presented here, the interval ratios of the identifying species for each factor are

compared among species and runs. If the interval ratios of the identifying species of a given factor are large, the identification of that factor is more uncertain than factors whose identifying species' interval ratios are lower.

2.3. Example data sets: Sacramento PM_{2.5}, Milwaukee water, and Las Vegas HR-AMS data

Table 1 summarizes the settings used for EPA PMF analysis for all three data sets presented here; details for each data set follow the table. Summary statistics of input parameters for each data set are provided in Supplemental Information. Multiple runs with different numbers of factors were run for each data set; PMF and EE diagnostics for each run are provided by data set, including figures showing how factor profiles and EE intervals changed with increasing number of factors. $Q_{expected}$ was calculated for each scenario, equal to (number of non-weak data values in \mathbf{X}) – (numbers of elements in \mathbf{G} and \mathbf{F} , taken together). For example, for five factors, 642 samples, and 19 strong species, this equals $(642 * 19) - ((5 * 642) + (5 * 19))$, or 8893. For six factors, $Q_{expected}$ would be 8232, and for seven factors it would be 7571. Weak data values were excluded due to their minor influence on $Q_{expected}$. $Q/Q_{expected}$ was calculated by species, as the sum of the squared scaled residuals (i.e., Q_{true}) divided by (overall $Q_{expected}$ divided by number of non-weak species); the denominator is simply the overall $Q_{expected}$ divided equally across species.

For the Sacramento data set, 24-hour speciated PM_{2.5} data for the Del Paso Manor monitoring site in Sacramento, California, operated as part of the chemical speciation network (CSN), were downloaded from EPA's Air Quality System (AQS) database, for July 2003 through February 2009. Concentrations and uncertainties were used as reported, to provide an example of using commonly available data; no censoring of data below detection was done, nor were samples with missing data included. Supplemental Table 1 provides summary statistics of input data. Nine aberrant samples, when Fourth of July or New Year's Day fireworks occurred, were excluded, leaving 642 samples for PMF analysis. The 19 species with S/N greater than 1 were included, including PM_{2.5} mass. For simplicity in this example, no species were made weak. Between five and seven factors were run.

Table 1
Summary of EPA PMF settings for all three data sets.

Parameter	Sacramento	Milwaukee	Las Vegas
Data type; averaging timeframe	PM _{2.5} CSN; 24-hr	Water quality; grab samples	HR-AMS; 20-min
N species	19	9–10	120
N samples	642	53	1405
N factors	5 to 7	3	3 to 5
Treatment of missing data	No missing data included	No missing data included	No missing data included
Treatment of data below detection limit (BDL)	Data used as reported, no modification or censoring of BDL data	Data used as reported, no modification or censoring of BDL data	Data used as reported, no modification or censoring of BDL data
Treatment of concentrations equal to or less than zero	Data used as reported, no modification or censoring of data < 0	Data used as reported, no modification or censoring of data < 0	Data used as reported, no modification or censoring of data < 0
Lower limit for normalized factor contributions g_{ik}	–0.2	–0.2	–0.2
Robust mode	Yes	Yes	Yes
Constraints	None	None	None
Seed value	Random	Random	Random
N bootstraps in BS	400	400	400
r^2 for BS	0.8	0.8	0.8
BS block size	3	1	4
DISP dQ^{max}	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32
DISP active species	All (no species were made weak)	All non-weak	All non-weak
N bootstraps; r^2 for BS in BS-DISP	50; 0.8	50; 0.8	50; 0.8
BS-DISP active species	K ⁺ , SO ₄ , NO ₃ , Si, Na + ^a , Cl ^b , Cu	NH ₃ , TSS, Cr	m/z 43, 44, 57, 60
BS-DISP dQ^{max}	0.5, 1, 2, 4	0.5, 1, 2, 4	0.5, 1, 2, 4
Computer run times ^c for DISP, BS-DISP	1 h, 3 h	<1 h, <1 h	9 h, 16 h

^a Only used in 7-factor solution.

^b Only used in 6- and 7-factor solutions.

^c Windows 7 64-bit, 3.1 GHz processor, 4 GB RAM.

For the Milwaukee data set, water samples of combined sewer outflows (wastewater plus stormwater), taken from an inline storage system (ISS) in Milwaukee and described in detail in Soonthornnonda and Christensen (2008) and Bzdusek et al. (2006), were explored; additional information is available at <http://v3.mmsd.com/wastewatertreatment/deep-tunnel>. Samples were collected from multiple sites on one day and were analyzed for trace metals (Cd, Cr, Cu, Pb, Ni, Zn) via EPA method 6010 (Inductively Coupled Plasma-Atomic Emission Spectrometry), and for biological oxygen demand (BOD), total suspended solids (TSS), total phosphorus (TP), and ammonia (NH₃) via standard methods (American Public Health Association methods 5210B, 2540D, and EPA methods 365.1 and 350.1). Supplemental Table 2 provides summary statistics of input data. Ten species across 53 samples were used, with three factors, as described by Soonthornnonda and Christensen (2008). Since Cd values were one of two values across all samples, indicating that the concentrations were near or below the detection limit, Cd was made weak in one run and excluded in another. Active species in BS-DISP were NH₃, Cr, and TSS.

For the Las Vegas data set, HR-AMS measurements were made outdoors next to a classroom and play yard at Fyfe Elementary School, 18 m from the US 95 freeway soundwall in Las Vegas, Nevada, during January 2008. Analysis of these data, including development and assessment of a PMF solution, are described in Brown et al. (2012). The HR-AMS measures non-refractory PM₁, including detailed spectra of organic aerosol (OA), which can then be decomposed by PMF (Allan et al., 2003; Jayne et al., 2000; Jimenez et al., 2003; DeCarlo et al., 2006; Ulbrich et al., 2009; Zhang et al., 2011; Lanz et al., 2010). Unlike traditional OA analysis, where aerosol is collected on filters over the course of many hours and later analyzed for individual molecules by gas chromatography–mass spectroscopy (GC–MS) or other techniques, the HR-AMS provides a high-time-resolution quantification of OA via mass spectra. The OA is characterized by a detailed mass spectrum of individual mass-to-charge ratios *m/z*; a time series of these mass spectra can then be decomposed via PMF to understand the OA composition (Ng et al., 2010; Ulbrich et al., 2009; Jimenez et al., 2009). Factors range in both volatility and degree of oxidation (Donahue et al., 2012; Jimenez et al., 2009; Kroll et al., 2011): low-volatility oxygenated organic aerosol (LV-OOA), semi-volatile oxygenated organic aerosol (SV-OOA), hydrocarbon-like organic aerosol (HOA), and biomass burning organic aerosol (BBOA). LV-OOA is highly oxygenated and is commonly found in all ambient OA; it typically has a high amount of *m/z* 44 (COO + fragment). HOA is generated by primary emissions, has low oxygen content, and is typically composed of saturated fragments such as *m/z* 43 (C₃H₇⁺ fragment) and 55 (C₄H₇⁺ fragment). In between HOA and LV-OOA in terms of oxidation is SV-OOA, which is composed of a mixture of the fragments in LV-OOA and HOA. BBOA is also somewhere between HOA and LV-OOA, depending on the combustion conditions (e.g., residential biomass burning versus wildfires) and aging of the OA, but is associated with *m/z* 60 (C₂H₄O₂⁺ fragment), which is derived from and is proportional to the biomass burning tracer molecule levoglucosan (Alfarra et al., 2007; Canagaratna et al., 2007; Lee et al., 2010).

In Brown et al. (2012), results using two-minute averaged data were presented. To reduce ME-2 run time and EPA PMF processing time, 20-minute averaged data were used in this analysis. A 75% completeness requirement was used for each 20-minute average, resulting in 1405 samples used in EPA PMF. Fragments with S/N less than 5 were made weak, meaning their uncertainties were increased by a factor of three; of the total 113 fragments (*m/z* up to 140), 12 were downweighted by this scheme. Supplemental Table 3 provides summary statistics of input data. Results were essentially identical between the two-minute data set in Brown et al. (2012) and the 20-minute data set used here, with the correlations (*r*²) between the factor profiles and factor contributions of the two data sets ranging between 0.98 and 0.999. EE results for the four-factor solution presented in Brown et al. (2012), as well as for three and five factors, are discussed here. No censoring of data below

detection or substitution of missing data was done. Active species for BS-DISP were *m/z* 43, 44, 55, and 60.

3. Results and discussion

3.1. Sacramento PM_{2.5} data

Five to seven factors were used with the Sacramento data; profiles are shown in Fig. 1 and results are summarized in Table 2 and Supplemental Table 4. Profiles identified for the five-factor solution included: (1) nitrate; (2) chlorine with sodium; (3) sulfate; (4) biomass burning/potassium (K); and (5) soil (Si, Ca, Fe). Moving to six factors, copper, chromium, and nickel moved out of the burning and soil factors to a new copper/metals factor. At seven factors, sodium ion separated from the chlorine factor into its own factor. PM_{2.5} mass, ammonium, elemental carbon (EC), organic carbon (OC), K, Si, sulfate and nitrate were well predicted (i.e., *r*₂ observed/predicted greater than 0.8) with five to seven factors. At six factors, iron was better predicted (0.80 with six factors versus 0.71 with five factors), and at seven factors, aluminum, calcium, and sodium ion were well predicted. In moving from five to six factors, there was a decrease in *Q/Q_{expected}* from 5.5 to 4.93, and a smaller decrease when moving from six to seven factors (4.93 to 4.63). When changes in *Q* become small with increasing factors, it can indicate that there may be too many factors being fit, suggesting here that six factors may be the optimal solution.

With five factors, all factors but Cl were mapped in 100% of BS runs (Cl was mapped 86% of runs), there were no swaps with DISP, and 100% of the BS-DISP runs were successful. Results were generally stable at six factors as well, with all factors mapped in BS in 100% of runs except for the copper/metals factor (mapped on 88% of runs). No swaps occurred with DISP, and all BS-DISP runs were successful. At seven factors, the solution was less stable. The new sodium factor was mapped with BS in 72% of the runs and copper/metals in 78% of the runs, while other factors were mapped in 100% of runs. There were no swaps in

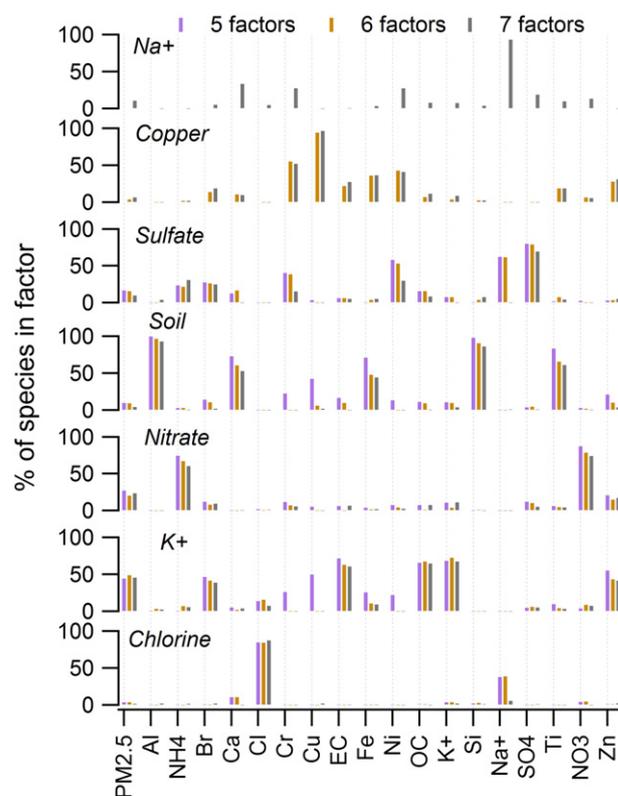


Fig. 1. Factor profiles (fraction of each species by factor) using Sacramento PM_{2.5} data, for five, six and seven factors.

Table 2
Summary of PMF and EE diagnostics by run for Sacramento PM_{2.5} data.

Diagnostic	5 factors	6 factors	7 factors
$Q_{expected}$	8893	8232	7571
Q_{true}	65,793	52,241	45,651
Q_{robust}	48,929	40,604	35,572
$Q_{robust}/Q_{expected}$	5.5	4.93	4.63
Species with $Q/Q_{expected} > 6$	PM _{2.5} , Ca, Cr, Cu, Ni, Si, Zn	PM _{2.5} , Ca, Cr, Cu, Ni, Zn	PM _{2.5} , Cr, Cu, Ni, Zn
DISP %dQ	<0.1%	<0.1%	<0.1%
DISP swaps	0	0	0
Factors with BS mapping < 100%	Cl factor 86%	Cu factor 94%	Na factor 72%, Cu factor 80%
BS-DISP % cases with swaps	0	0	28%

DISP, but 28% of BS-DISP runs were rejected due to factor swaps. Thus, while additional species had better observed/predicted diagnostics with seven factors, these additional factors appeared less stable than the factors found in the five- and six-factor solutions.

As seen in Supplemental Table 4 and Fig. 2, DISP error estimate intervals, expressed as interval ratios to be comparable across species of differing magnitudes, are quite low for key species, indicating little rotational ambiguity in the solutions. Ratios are generally highest for the model with seven factors, indicating modestly higher uncertainty for these key species with seven factors. For BS, interval ratios are generally consistent for a given factor for all three model runs, with the exception of the copper/metals factor at seven factors. Here, the BS interval of copper is relatively large, spanning an order of magnitude (0.00012 to 0.0065 μg between the 5th and 95th BS percentiles), resulting in a very high EE for copper in this factor, which suggests a poorly defined factor. The BS mapping indicates some instability in the chlorine factor at five factors, the copper/metals factor at both six and seven factors, and modest instability of the sodium-only factor with seven factors. The instability of the chlorine, copper/metals, and sodium factors are further seen with BS-DISP, where the interval ratios for the key species are high for factors with low BS mapping (chlorine at five factors, copper at six factors, and both sodium and copper at seven factors). In these three cases, BS-DISP interval ratios approach or are equal to two, since the BS-DISP 5th percentile for these species/factor combinations is at or near zero. The combination of poor BS mapping of two of seven factors, the very high EE intervals from both BS and BS-DISP for both copper and sodium factors, and the small change in $Q/Q_{expected}$ going from six to seven modeled factors indicate that the seven-factor solution is not stable and likely should not be used.

3.2. Milwaukee water results

As further described by Soonthornnonda and Christensen (2008), three factors were determined: (1) stormwater was characterized by

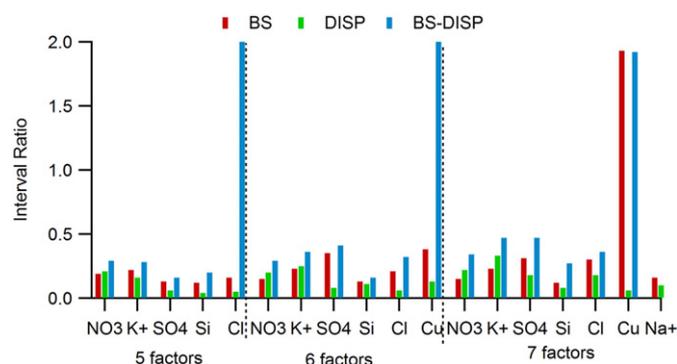


Fig. 2. Interval ratios of identifying species by EE method and number of factors (5, 6, 7) for Sacramento PM_{2.5} data.

high amounts of TSS and Pb; (2) sanitary sewage was characterized by high BOD, TP, and ammonia; and (3) high-metals-content stormwater, likely from sewer sediment erosion, was characterized by high concentrations of metals such as Cr. Table 3, Supplemental Table 5 and Figs. 3 and 4 summarize the results. With three factors and Cd included as weak, all species were relatively well predicted, with $Q/Q_{expected}$ values all less than 2 except for BOD. BS results showed 100% mapping for two factors and 86% mapping for the trace metals factor; DISP had no swaps and 98% of BS-DISP cases were successful. Upon removing Cd, other species were not any better predicted, $Q/Q_{expected}$ was similar, and BS mapping was poorer compared to the run with Cd included (BS mapping 72% for metals factor). In both scenarios, the poorer BS mapping of the metals factors is likely due to its more intermittent signal across the samples, relative to the other two more consistent sources. In addition, the small overall size of the matrix (10 species, 53 samples) is likely at the extreme lower end of a viable size for PMF applications, which may lead to some instability. Despite these limitations, the factors when including Cd are stable and as reported by Soonthornnonda and Christensen (2008), also compare very well with chemical mass balance (CMB) results, further solidifying their interpretability.

3.3. Las Vegas HR-AMS results

As described by Brown et al. (2012), four factors were determined: LV-OOA, HOA, SV-OOA and BBOA. On average, HOA made up 28% of the organic matter (OM), had an abundance of m/z 43, and peaked during the morning and evening commute periods coincident with peak traffic volume. LV-OOA, indicated by an abundance of m/z 44, was highest in the afternoon and accounted for 25% of the OM. BBOA occurred in the evening hours, was predominantly from the residential area to the north, and on average constituted 10% of the OM. SV-OOA accounted for the remaining 37% of the OM, and had an abundance of m/z 57 and 55. The HOA and LV-OOA factors were nearly identical to those found in other studies; correlations of the profiles with Pittsburgh factor profiles (Ulbrich et al., 2009) were 0.99. The HOA factor profile is very similar to pure diesel exhaust (Mohr et al., 2009). The BBOA factor had typical tracer fragments of m/z 60 and 73, which are produced during AMS analysis of levoglucosan and related anhydrosugars produced during biomass combustion (Lanz et al., 2008; Lee et al., 2010). The SV-OOA factor profile was similar to that of aged diesel exhaust (Sage et al., 2008).

Factor profiles are shown in Fig. 5 and EE diagnostics are summarized in Table 4, Supplemental Table 6, and in Fig. 6. For the four-factor base solution, BS resamples reproduced 100% of the base factors. There were no factor swaps with DISP and, as also seen in the other data set examples, only an extremely low change in Q (less than 0.1%) was observed. However, in BS-DISP, 46% of the runs had swaps. BS-DISP interval ratios were also the largest among EE methods. The BS-DISP interval ratio was highest for SV-OOA across all EE methods. These results suggest that the SV-OOA factor is more uncertain than the other factors.

With three factors, only HOA, LV-OOA and BBOA were identified, 100% of the BS resamples identified these 3 factors, 100% of the BS-DISP runs were accepted, and no swaps occurred with DISP. EE interval ratios are generally lower with three factors than with four factors, with the exception of DISP interval ratios that were higher when using three instead of four factors. This may indicate that using three factors distorts the solution so that the three factors also accommodate parts of the omitted fourth factor SV-OOA. When four factors are used, DISP intervals are smaller and BS results similar, suggesting at least four factors are needed, despite the modest BS-DISP results.

With five factors, an additional "night OA" factor is found that occurs on most evenings coincident with BBOA and SV-OOA. However, this night OA factor is only found with 80% of the BS resamples, while the other factors are mapped in 100% of the runs. With BS-DISP and five factors, 44% of the runs were accepted and there were no swaps with DISP. These results indicate that the five-factor solution, and in particular the

Table 3
Summary of PMF and EE diagnostics by run for Milwaukee water quality data.

Diagnostic	2 factors	3 factors, with Cd weak	3 factors, excluding Cd
$Q_{expected}$	353	291	291
Q_{true}	1014	652	635
Q_{robust}	1004	647	630
$Q_{robust}/Q_{expected}$	2.86	2.25	2.17
Species with $Q/Q_{expected} > 2$	Biological oxygen demand, total suspended solids, Cr, Pb	Biological oxygen demand, total suspended solids	Biological oxygen demand, total suspended solids
DISP %dQ	<0.1%	<0.1%	<0.1%
DISP swaps	0	0	0
BS mapping	100%	Metals factor 88%	Metals factor 74%, stormwater factor 92%
BS-DISP % cases with swaps	0%	8%	8%

night OA factor, is much less certain than the four-factor solution. The modest BS-DISP results with four factors suggest that there is some factor interdependence and rotational ambiguity, confirmed by the oblique, slanting edges seen in the G-space plots (Paatero et al., 2002, 2005). In these results, the oblique edges in the G-space plots could not be straightened out by applying customary rotational techniques, e.g., by applying Fpeak or by pulling points along the edges. Thus, they indicate a “modeling error” in the analysis, such as variation in true source profiles during the monitoring campaign, or presence of data outliers that block the rotations that would be needed for straightening the edges.

3.4. Discussion

In all three data sets, there were no swaps evident in DISP, indicating that the solutions had no or few data errors and were well defined. These results differ from the synthetic data analyses shown in Paatero et al. (2014), which did have swaps in DISP but only when there were too many factors. In the ambient data examples here, even when pushed up to two or more factors above the “base” solution, swaps did not occur with DISP. This was also the case when small data sets were run, e.g., the Milwaukee water data. It seems likely that if more than a few swaps occur with an ambient data set, then there are either too many factors used or the solution is not well defined. Thus, DISP appears to be a good first-step screening tool for a PMF solution; if zero or only a few swaps occur, the user is assured that they are working toward a reasonable solution, though results with BS and BS-DISP may eventually suggest otherwise.

In the examples presented here, BS continues to be a useful EE method even though it does not account for rotational ambiguity. When

factors are not reproduced during BS resampling, it indicates potential problems with that solution. It could be that too many factors are being used, as is likely at seven factors in Sacramento PM_{2.5} or five factors in Las Vegas HR-AMS data, or that the factors with low reproducibility occur infrequently in the data. In the case where it is clear that the occurrence of a factor is dependent on other environmental conditions, such as wind direction, meteorology, or source operations, it is not surprising that many BS resamples do not identify that factor. BS results thus are useful for quantifying the uncertainty of a solution, and also for identifying factors that have a low degree of reproducibility. Such factors with low reproducibility may still be real, but require additional investigation and support for their inclusion.

BS-DISP combines BS's strength with data errors and DISP's strength with rotational uncertainty. In these examples, solutions with no swaps in DISP and more than 95% reproducibility with BS had variable BS-DISP results. For example, the four-factor solution with Las Vegas HR-AMS had no swaps in DISP, high reproducibility with BS, and low interval ratios for the identifying species in each factor. However, there were swaps in 46% of the BS-DISP runs, indicating some uncertainty with the solution. With positive BS and DISP results and clear interpretability of the factors, this amount of swaps is not fatal to the analysis, but confirms that there is some uncertainty in the solution, particularly regarding the SV-OOA factor. Since SV-OOA factors can vary widely across studies depending on the atmospheric conditions and processing of OA, and the factors span a much larger range of volatility compared to HOA and LV-OOA, the modest swapping in BS-DISP appears to confirm its larger uncertainty compared to other factors.

DISP and BS-DISP provide results for four different dQ^{max} values; the range of results for each key species/factor combination by dQ^{max} value can indicate whether uncertainties are controlled more by rotational

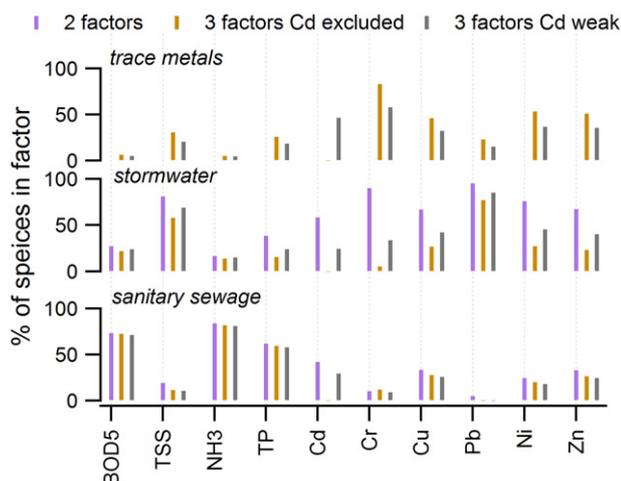


Fig. 3. Factor profiles (fraction of each species by factor) using Milwaukee water data, for two factors, three factors with Cd excluded, and three factors with Cd weak.

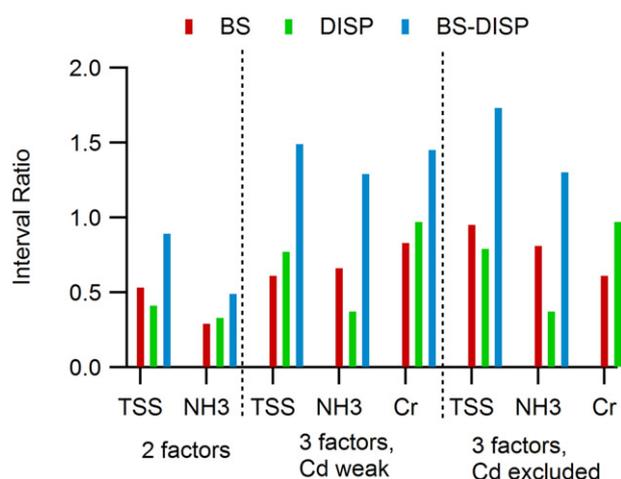


Fig. 4. Interval ratios of identifying species by EE method and run (two factors, three factors with Cd weak, and three factors with Cd excluded) for Milwaukee water data.

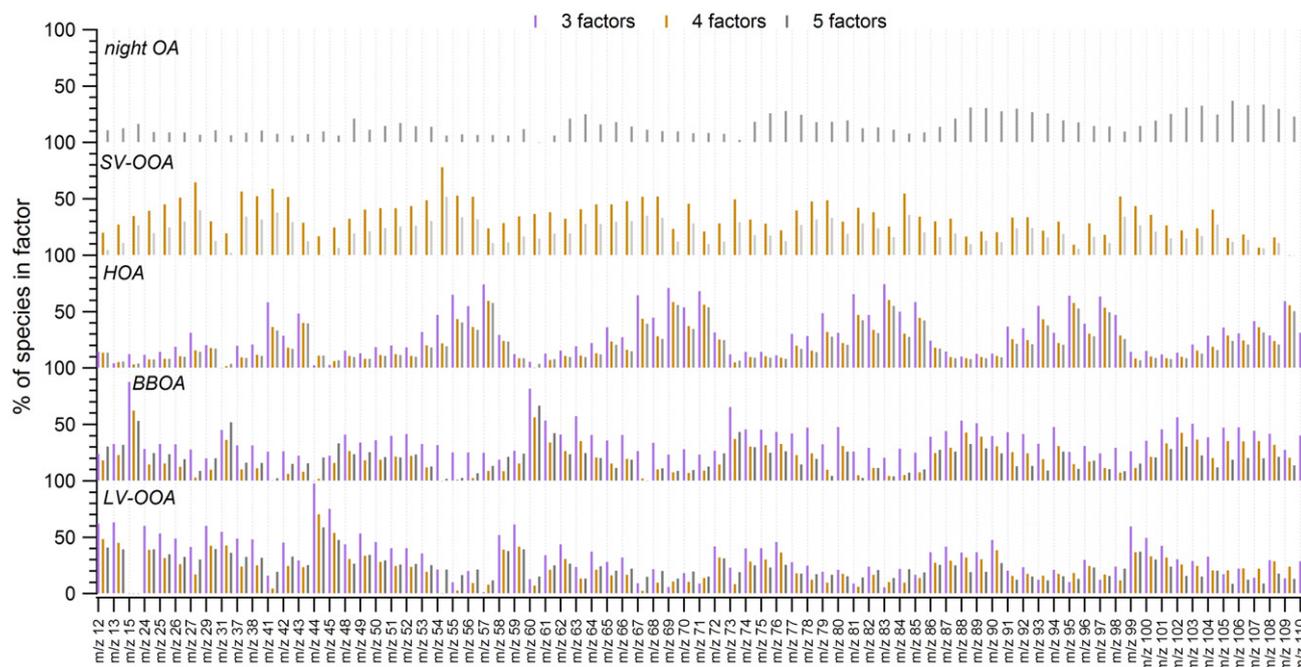


Fig. 5. Factor profiles (fraction of each species by factor) using Las Vegas AMS data, for three, four and five factors.

uncertainty or user-specified data uncertainties. Supplemental Table 7 shows the DISP intervals for the Sacramento example for key identifying species in each solution. The Sacramento example is particularly useful, because the solutions have residuals far exceeding the input uncertainties and thus have a relatively high $Q/Q_{expected}$, indicating that the input uncertainties are not accurate, the factor profiles vary over time, or both. How DISP intervals change with dQ^{max} may help explain the high $Q/Q_{expected}$ values. With five factors, nitrate intervals are nearly independent of dQ^{max} ; thus, the input uncertainties for nitrate are well represented by the solution, so that the DISP intervals are controlled predominantly by rotational uncertainty, i.e., how far rotations may proceed during DISP. In contrast, sulfate, Cl, and Si intervals increase proportionally to the square root of dQ^{max} , meaning that they double as dQ^{max} increases by a factor of 4. Thus, there is less rotational uncertainty for these factors; rather, the uncertainty is due to input uncertainty. At six factors, all species except Cl have intervals that do not increase proportional to dQ^{max} , indicating significant rotational uncertainty. With seven factors, rotational uncertainty appears to be reduced, but at the expense of much larger uncertainty intervals. These variations in results by dQ^{max} further support the earlier interpretation that seven factors are likely too many, and that there are some trade-offs in uncertainty between the five- and six-factor solutions.

Table 4
Summary of PMF and EE diagnostics by run for Las Vegas HR-AMS data.

Diagnostic	3 factors	4 factors	5 factors
$Q_{expected}$	137,351	135,833	134,315
Q_{true}	920,029	607,201	402,907
Q_{robust}	454,015	340,269	264,455
$Q_{robust}/Q_{expected}$	3.31	2.51	1.97
Species with $Q/Q_{expected} > 8$	m/z 27, 41, 43, 44, 54, 55, 56, 57, 60, 69, 80, 71, 73, 86, 95	m/z 31, 43, 55, 57, 60, 70, 73, 86	m/z 41, 43, 55, 70, 86
DISP %dQ	<0.1%	<0.1%	<0.1%
DISP swaps	0	0	0
BS mapping	100%	100%	5th factor 80%
BS-DISP % cases with swaps	0% with swaps	46% with swaps	56% with swaps

4. PMF reporting recommendations

PMF analyses involve many details about the development of the data, decisions of what data to include/exclude, determination of a solution, and evaluation of robustness of that solution; reporting of PMF solutions and analyses vary widely. In many cases, limitations on word count and other restrictions mean that authors do not include important details of their modeling efforts in published articles. As more journals publish online and allow appendices or supplemental material, more analysis details can be shared. Having a consistent base of what is reported will help all PMF users evaluate, compare, understand, and reproduce PMF analyses. Below, we provide a sample list of recommended items to report when presenting a PMF analysis. This is complementary to the AMS-specific strategy recently assembled in Crippa et al. (2013).

4.1. Q values

Report the Q_{robust} and Q_{true} values of the analysis that was deemed most useful, and note how Q or $Q/Q_{expected}$ changed under different scenarios, e.g., with a different number of factors or with different species

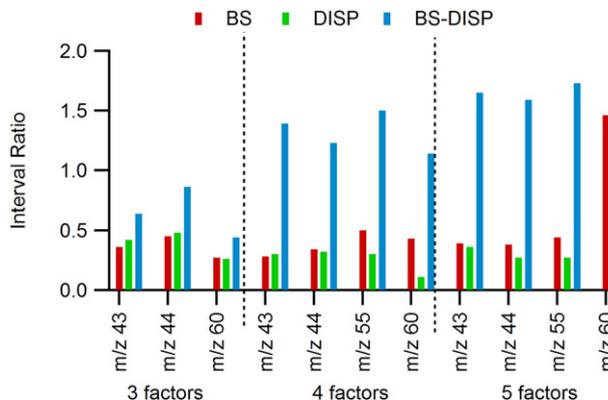


Fig. 6. Interval ratios of identifying species by EE method and number of factors (3, 4, 5) for Las Vegas data.

included/excluded. Discussion of the obtained Q value is moot if uncertainties of the input data have been fine-tuned in order to produce meaningful scaled residuals. In contrast, observing changes in Q under different scenarios is often very helpful when selecting between different modeling alternatives, such as different numbers of factors. When changes in Q become small with increasing factors, it can be indicative that there may be too many factors being fit.

$Q_{expected}$ should also be calculated, but only the “good” or non-weak variables should be taken into account. The expected value of Q is approximately = (number of non-weak data values in \mathbf{X}) – (numbers of elements in \mathbf{G} and \mathbf{F} , taken together). A downweighted weak variable has only a small, rarely significant contribution to $Q_{expected}$, and for simplicity is excluded here. If the Q value of the chosen model differs significantly from what is expected (e.g., by a factor of ten or more), then DISP error analysis becomes invalid and BS-DISP is likely questionable.

Lastly, it can be useful to report if an individual column or row of \mathbf{X} had a $Q/Q_{expected}$ ratio that was much higher than that of other columns or rows. This indicates that the column or row was not well fitted and contributes significantly more than expected to Q .

4.2. Estimated or adjusted uncertainties of input data

Uncertainties drive not only a base solution but also the BS, and in particular, the DISP and BS-DISP results. Their development and use in the PMF analysis needs to be clearly documented. This also includes documenting if extra modeling uncertainty (an adjustable parameter within PMF) was applied in the analysis.

4.3. Lower limit for G (contributions)

In EPA PMF, the lower limit of the normalized contributions is set to -0.2 , since allowing a small negative value helps PMF accept true rotations even in the presence of a large number of zero values in some G factors.

4.4. Use of robust mode

In EPA PMF, the robust mode is always used, which automatically downweights the influence of observations that have a scaled residual greater than 4. Nevertheless, use of robust mode should always be documented in publications.

4.5. Treatment of missing values

If missing data were included in the PMF analysis, they need to be treated appropriately so they do not influence the solution. Often, the median concentration of a given species is used, with an uncertainty of four times the median. The scaled residuals for these points should be inspected to ensure that they are clearly less than one. If missing data are given a standard deviation of four times the median, then the scaled residuals for these points may occasionally violate this requirement. In such cases, the analysis should be repeated so that uncertainties of missing values are increased sufficiently. Multiple statistical methods exist for replacing missing data with statistically viable values. We do not recommend these methods for PMF analyses. They are necessary for any statistical procedures that cannot accommodate missing data. With PMF analyses, inputting a sufficiently large uncertainty makes the data truly “missing”, a process that cannot be improved by using data substitutions. PMF can even be used to obtain substitution values to be used in other statistical procedures: run PMF so that the values in question (missing and/or BDL values) have sufficiently large uncertainty values associated to them. Use the fitted values (fitted by PMF to missing/BDL positions) as substitution values, and then use them as replacements to whatever values were originally present in the missing/BDL positions of the matrix.

4.6. Treatment of data below detection

In many published PMF studies, below detection level (BDL) data values have been censored, i.e., substituted by replacement values, such as $0.5 \times$ detection limit, even if the original measured values have been available. It appears that this practice has no proven advantages when species with low S/N are downweighted. On the other hand, it may be demonstrated that the substitution practice prevents uncertainty estimation, introduces hard-to-estimate bias, and occasionally gives rise to ghost factors. In general, it is a modeling error if BDL values are replaced by a fraction of the detection limit. If such a replacement has been done, then EE of PMF results should not be attempted because none of the available EE methods is able to estimate the bias error incurred in results by censoring BDL values. If EE is nevertheless attempted in the presence of such censoring, then a clear warning about questionable validity of quoted error estimates must be included in the paper. Instead of substitution methods, PMF modeling using ME-2 directly can be optimized by applying a specific error model code to censored data values (Paatero, 2000). In this way the known information, e.g., that a measurement is somewhere between zero and the detection limit, can be conveyed to ME-2 without any substitutions that would likely bias the results.

4.7. Treatment of data equal to or below zero

Data equal to or less than zero can be included in the PMF model and, if these values are genuine measured values, should not be censored by truncation to zero or transformed to positive values. If such censored values must be used because the original measured values have been discarded, then a warning must be included in the documentation. EPA PMF allows negative concentration values to be used, though input uncertainties must still be positive.

4.8. Treatment of “total mass”

In the Sacramento example, uncertainties for total $PM_{2.5}$ mass were used as reported, rather than being further downweighted (e.g., Reff et al. (2007); Kim et al. (2005)). $PM_{2.5}$ mass should be downweighted if there are likely significant measurement artifacts, e.g., gaseous species adsorbing onto filters, or if there are sources that may emit $PM_{2.5}$ mass but none of the measured species, in which case a factor containing only mass could be determined. In these cases, the inclusion of “full strength” total mass does not help in interpretation of solutions, and could lead to erroneous results.

4.9. Use of constraints

EPA PMF and ME-2 allow users to constrain or “pull” elements in their solution. One common reason for pulling is the attempt to align an oblique “edge” in G -space plots. However, an oblique edge may sometimes be justified because factors in atmospheric or environmental data are rarely truly independent. Thus pulling contributions based on G -space plots must be clearly justifiable, and should be justified and reported in detail if done. If source profiles or contributions are known for some factors or samples, and constraints are used to model these, then this information and the reason (e.g., the industrial plant was shut down and should have a contribution of zero) should be noted.

4.10. BS

Report the number of resamples analyzed and the size of percentiles of the obtained distribution of results chosen for error limits, e.g., in EPA PMF these are the 5th and 95th percentiles. Also report the percentage of BS factors assigned to each base case factor and the number of BS factors not assigned to any base case factor, and the interval ratios of each factor's identifying species.

4.11. DISP

Report species not displaced, such as those downweighted (in EPA PMF, all strong species are used in DISP), the decrease in Q , the number of factor swaps, and the interval ratios of each factors' identifying species. If factor swaps occur for the smallest dQ^{max} , it indicates that there is significant rotational ambiguity and that the solution is not sufficiently robust to be used. If the decrease in Q is greater than 1%, it is likely the case that no DISP results should be published unless DISP analysis is redone after finding the true global minimum of Q .

4.12. BS-DISP

As with BS and DISP, report the number of BS resamples analyzed, the size of percentiles chosen for error limits (in EPA PMF, these are the 5th and 95th), the species actively displaced, the decrease in Q , and the number of factor swaps. For each factor's identifying species, note the extent of the EE interval.

5. Conclusions

These examples using different ambient data sets are presented to complement the theory behind new EE methods in EPA PMF and ME-2, as well as synthetic data analyses presented in a companion paper (Paatero et al., 2014). Together, they show the use of multiple EE methods. With these ambient data sets, DISP typically had tight intervals and no factor swaps; it appears that DISP is a good screening tool for solutions, as solutions that have swaps likely have significant rotational ambiguity and should probably not be used. BS results do not typically capture rotational ambiguity, but can help identify factors that are not very reproducible, though low reproducibility may be due to other influences such as wind direction and source activity, rather than a poor solution. BS-DISP may yield factor swaps even if BS and DISP diagnostics are positive, and can be used to identify which factors are more certain than others. One drawback of BS-DISP is its computation time. A large run (e.g., thousands of samples and a hundred species) may take tens of hours on a modern PC. Future work may need to focus on optimization of the algorithms in ME-2 to help significantly decrease the run time.

Based on the results here and in Paatero et al. (2014) the different roles of DISP on one hand, and BS and BS-DISP on the other hand, may tentatively be described as follows: DISP analyzes the given data set "as is," not speculating about the reproducibility of the results in future similar measurements. Hence, uncertainty intervals given by DISP are rather short. The BS-based methods are based on resampling. Hence, they also estimate variability that follows if a few key samples are omitted from the data set. Such estimation may be relevant in predicting what might happen with the following year's data. A few key samples might not occur at all, or they might not occur on days when sampling is performed. BS takes into account this kind of uncertainty, which is not considered by pure DISP. Thus, uncertainty estimates based on BS may be much larger for such data sets, where a few matrix rows have special importance regarding rotations or regarding sources that are only observed in a small fraction of all samples.

For all EE methods, the interval ratio of each factor's identifying species can provide an understanding of the relative certainty of each factor's identity. However, each data set is unique, so results will vary. It must be emphasized that, in some cases, a satisfactory analysis cannot be performed with any number of factors. The following contrived example illustrates this situation. With four factors, the result is rotationally unique and all three EE methods indicate small uncertainties. However, with four factors, the fit is not satisfactory as indicated, e.g., by poor total mass reconstruction and by poor factor interpretability. With five factors, fit is good but rotational uncertainty is very large and/or there are frequent factor swaps between factors four and five. What should the scientist do in this situation? It would be wrong to only report either

the four-factor or the five-factor results while ignoring the presence of the alternative solution. The information in the data set confirms that four factors are not enough. However, the information is not sufficient for quantitative determination of five factors. If additional information cannot be inserted for obtaining rotational uniqueness, then the two sets of inconclusive results (using four and five factors, respectively), as well as the impact that subtracting or adding a factor has on the profiles, contributions, and EE results, should be reported.

The conclusions presented in this work are based on our experience with a limited number of synthetic and real data sets. It was not our intention to "prove" the validity or usefulness of these methods. The statistical properties of real data are so unknown and varied that a general assessment of the validity can only be reached through a long process. Successful and failed analyses of different data should be carefully reported in literature. Thus, the present conclusions should not be regarded as the final truth about EE of bilinear models. Instead, these results are the first steps toward full understanding of these complicated questions. It is essential that follow-up studies be performed with an open mind, so that general validity of our conclusions is not taken for granted in all possible situations. Lastly, we provide a recommended "best practices" list of information for users to report in their publications, which is critical as more users employ the new EE methods available in ME-2 and EPA PMF.

Disclaimer

The United States Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under contract EP-D-09-097 to Sonoma Technology, Inc. It has been subjected to Agency review and approved for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.scitotenv.2015.01.022>.

References

- Alfarra, M.R., Prévôt, A.S.H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V.A., Schreiber, D., Mohr, M., Baltensperger, U., 2007. Identification of the mass spectral signature of organic aerosols from wood burning emissions. *Environ. Sci. Technol.* 41, 5770–5777. <http://dx.doi.org/10.1021/es062289b>.
- 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., Worsnop, D.R., 2003. Quantitative sampling using an aerodyne aerosol mass spectrometer – 2. Measurements of fine particulate chemical composition in two U.K. cities. *J. Geophys. Res.-Atmos.* 108 (D3) (Feb 4).
- Brown, S.G., Lee, T., Norris, G.A., Roberts, P.T., Collett Jr., J.L., Paatero, P., Worsnop, D.R., 2012. Receptor modeling of near-roadway aerosol mass spectrometer data in Las Vegas, Nevada, with EPA PMF. *Atmos. Chem. Phys.* 12, 309–325. <http://dx.doi.org/10.5194/acp-12-309-2012> (Available at <http://www.atmos-chem-phys.net/12/309/2012/>).
- Bzdusek, P.A., Lu, J., Christensen, E.R., 2006. PCB congeners and dechlorination in sediments of Sheboygan River, Wisconsin, determined by matrix factorization. *Environ. Sci. Technol.* 40 (1), 120–129. <http://dx.doi.org/10.1021/es050083p>.
- 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 Spectrom. Rev.* 26 (2), 185–222. <http://dx.doi.org/10.1002/mas.20115> (Available at <http://onlinelibrary.wiley.com/doi/10.1002/mas.20115/abstract>).
- Crippa, M., Canonaco, F., Lanz, V.A., Äijälä, M., Allan, J.D., Carbone, S., Capes, G., Dall'Osto, M., Day, D.A., DeCarlo, P.F., Di Marco, C.F., Ehn, M., Eriksson, A., Freney, E., Hildebrandt Ruiz, L., Hillamo, R., Jimenez, J.-L., Junninen, H., Kiendler-Scharr, A., Kortelainen, A.-M., Kulmala, M., Mensah, A.A., Mohr, C., Nemitz, E., O'Dowd, C., Ovadnevaite, J., Pandis, S.N., Petäjä, T., Poulain, L., Saarikoski, S., Sellegri, K., Swietlicki, E., Tiitta, P., Worsnop, D.R., Baltensperger, U., Prévôt, A.S.H., 2013. Organic aerosol components derived from 25 AMS datasets across Europe using a newly developed ME-2 based source apportionment strategy. *Atmos. Chem. Phys. Discuss.* 13, 23325–23371. <http://dx.doi.org/10.5194/acpd-13-23325-2013> (Available at <http://www.atmos-chem-phys-discuss.net/13/23325/2013/acpd-13-23325-2013.html>).

- DeCarlo, P., Kimmel, J.R., Trimborn, A., Northway, M., 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. *Anal. Chem. (Wash.)* 78 (24), 8281–8289. <http://dx.doi.org/10.1021/ac061249n>.
- Donahue, N.M., Kroll, J.H., Pandis, S.N., Robinson, A.L., 2012. A two-dimensional volatility basis set – Part 2: diagnostics of organic-aerosol evolution. *Atmos. Chem. Phys.* 12, 615–634. <http://dx.doi.org/10.5194/acp-12-615-2012>.
- Hopke, P.K., 2008. The use of source apportionment for air quality management and health assessments. *J. Toxic. Environ. Health A* 71, 555–563.
- Jayne, J.T., Leard, D.C., Zhang, X.F., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an aerosol mass spectrometer for size and composition analysis of submicron particles. *Aerosol Sci. Technol.* 33 (1–2), 49–70.
- 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., Davidovits, P., 2003. Ambient aerosol sampling using the Aerodyne aerosol mass spectrometer. *J. Geophys. Res. Atmos.* 108 (D7) (Apr 15).
- Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prévôt, 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., Shimon, 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., Worsnop, D.R., 2009. Evolution of organic aerosols in the atmosphere. *Science* 326, 1525–1529. <http://dx.doi.org/10.1126/science.1180353> (December 11).
- Kim, E., Hopke, P.K., 2007. Source identifications of airborne fine particles using positive matrix factorization and U.S. Environmental Protection Agency positive matrix factorization. *J. Air Waste Manage. Assoc.* 57 (7), 811–819. <http://dx.doi.org/10.3155/1047-3289.57.7.811> (July).
- Kim, E., Hopke, P.K., Kanski, D.M., Koerber, M., 2005. Sources of fine particles in a rural midwestern U.S. area. *Environ. Sci. Technol.* 39 (13), 4953–4960.
- Kroll, J.H., Donahue, N.M., Jimenez, J.L., Kessler, S.H., Canagaratna, M.R., Wilson, K.R., Altieri, K.E., Mazzoleni, L.R., Wozniak, A.S., Bluhm, H., Mysak, E.R., Smith, J.D., Kolb, C.E., Worsnop, D.R., 2011. Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol. *Nat. Chem.* 3, 133–139. <http://dx.doi.org/10.1038/NCHEM948>.
- Lanz, V.A., Alfarra, M.R., Baltensperger, U., Buchmann, B., Hueglin, C., Szidat, S., Wehrl, M.N., Wacker, L., Weimer, S., Caseiro, A., Puxbaum, H., Prévôt, A.S.H., 2008. Source attribution of submicron organic aerosols during wintertime inversions by advanced factor analysis of aerosol mass spectra. *Environ. Sci. Technol.* 42 (1), 214–220. <http://dx.doi.org/10.1021/es0707207>.
- Lanz, V.A., Prévôt, A.S.H., Alfarra, M.R., Weimer, S., Mohr, C., DeCarlo, P.F., Gianini, M.F.D., Hueglin, C., Schneider, J., Favez, O., D'Anna, B., George, C., Baltensperger, U., 2010. Characterization of aerosol chemical composition with aerosol mass spectrometry in Central Europe: an overview. *Atmos. Chem. Phys.* 10, 10453–10471. <http://dx.doi.org/10.5194/acp-10-10453-2010> (November 8).
- Lee, T., Sullivan, A.P., Mack, L., Jimenez, J.L., Kreidenweis, S.M., Onasch, T.B., Worsnop, D.R., Malm, W., Wold, C.E., Hao, W.M., Collett Jr., J.L., 2010. Chemical smoke marker emissions during flaming and smoldering phases of laboratory open burning of wildland fuels. *Aerosol Sci. Technol.* 44 (9), i–v. <http://dx.doi.org/10.1080/02786826.2010.499884>.
- McCarthy, M.C., Aklilu, Y.-a., Brown, S.G., Lyder, D.A., 2013. Source apportionment of volatile organic compounds measured in Edmonton, Alberta. *Atmos. Environ.* 81, 504–516 (December, Available at <http://www.sciencedirect.com/science/article/pii/S1352231013007048>).
- Mohr, C., Huffman, J.A., Cubison, M., Aiken, A.C., Docherty, K.S., Kimmel, J.R., Ulbrich, I.M., Hannigan, M., Jimenez, J.L., 2009. Characterization of primary organic aerosol emissions from meat cooking, trash burning, and motor vehicles with high-resolution aerosol mass spectrometry and comparison with ambient and chamber observations. *Environ. Sci. Technol.* 43 (7), 2443–2449. <http://dx.doi.org/10.1021/es8011518>.
- Ng, N.L., Canagaratna, M.R., Zhang, Q., Jimenez, J.L., Tian, J., Ulbrich, I.M., Kroll, J.H., Docherty, K.S., Chhabra, P.S., Bahreini, R., Murphy, S.M., Seinfeld, J.H., Hildebrandt, L., Donahue, N.M., DeCarlo, P.F., Lanz, V.A., Prévôt, A.S.H., Dinar, E., Rudich, Y., Worsnop, D.R., 2010. Organic aerosol components observed in Northern Hemispheric datasets from Aerosol Mass Spectrometry. *Atmos. Chem. Phys.* 10, 4625–4641. <http://dx.doi.org/10.5194/acp-10-4625-2010>.
- Norris, G., Duvall, R., Brown, S., Bai, S., 2014. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide. Prepared for the U.S. Environmental Protection Agency Office of Research and Development, Washington, DC (EPA/600/R-14/108; STI-910511-5594-UG, April).
- Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. *Chemom. Intell. Lab. Syst.* 37, 23–35.
- Paatero, P., 2000. User's Guide for the Multilinear Engine Program "ME2" for Fitting Multilinear and Quasi-multilinear Models (February).
- 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., Hopke, P.K., Song, X.H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. *Chemom. Intell. Lab. Syst.* 60, 253–264.
- Paatero, P., Hopke, P.K., Hoppenstock, J., Eberly, S.I., 2003. Advanced factor analysis of spatial distributions of PM_{2.5} in the eastern United States. *Environ. Sci. Technol.* 37 (11), 2460–2476 (June 1).
- Paatero, P., Hopke, P.K., Begum, B.A., Biswas, S.W., 2005. A graphical diagnostic method for assessing the rotation in factor analytical models of atmospheric pollution. *Atmos. Environ.* 39, 193–201.
- Paatero, P., Eberly, S.I., Brown, S.G., Norris, G.A., 2014. Methods for estimating uncertainty in factor analytic solutions. *Atmos. Meas. Tech.* 7, 781–797. <http://dx.doi.org/10.5194/amt-7-781-2014> (Available at <http://www.atmos-meas-tech.net/7/781/2014/amt-7-781-2014.html>).
- Poirot, R.L., Wishinski, P.R., Hopke, P.K., Polissar, A.V., 2001. Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont. *Environ. Sci. Technol.* 35 (23), 4622–4636 (November).
- Reff, A., Eberly, S.I., Bhawe, P.V., 2007. Receptor modeling of ambient particulate matter data using positive matrix factorization: review of existing methods. *J. Air Waste Manage. Assoc.* 57, 146–154.
- Sage, A.M., Weitkamp, E.A., Robinson, A.L., Donahue, N.M., 2008. Evolving mass spectra of the oxidized component of organic aerosol: results from aerosol mass spectrometer analyses of aged diesel emissions. *Atmos. Chem. Phys.* 8, 1139–1152.
- Soonthornnonda, P., Christensen, E.R., 2008. Source apportionment of pollutants and flows of combined sewer wastewater. *Water Research* 42 (8–9), 1989–1998. <http://dx.doi.org/10.1016/j.watres.2007.11.034> (April, Available at <http://www.sciencedirect.com/science/article/pii/S0043135407007300>).
- 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. *Atmos. Chem. Phys.* 9, 2891–2918.
- Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Ulbrich, I.M., Ng, N.L., Worsnop, D.R., Sun, Y., 2011. Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review. *Anal. Bioanal. Chem.* <http://dx.doi.org/10.1007/s00216-011-5355-y> (October).