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, California 94954
[b] Geometric Tools, LLC, 2909 181st Ave NE, Redmond, Washington 98052
[c] University of Helsinki, Dept. of Physics, P.O. Box 33 (Yliopistonkatu 4), 00014 University of Helsinki, Helsinki, Finland
[d] U.S. EPA, Office of Research and Development, Research Triangle Park, North Carolina
*corresponding author

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.

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

KEYWORDS

Receptor modeling, air pollution, water pollution, positive matrix factorization, EPA PMF

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., 2014; Paatero et al., 2003; 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 **X** of dimensions *n* by *m*, in which *n* samples and *m* chemical species were measured. Rows and columns of **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 each individual sample that solve the chemical mass balance between measured species concentrations and factor profiles (Equation 1):

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

where e_{ij} is the residual for each sample/species and c_{ij} is the modeled solution of $\boldsymbol{x}_{ij}$. Entire matrices are denoted by capital bold-face letters. Columns of the factor contribution matrix **G** may be denoted by $\mathbf{g}_k$, and similarly rows of factor profile matrix **F** denoted by $\mathbf{f}_k$.

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* (Equation 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 (**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 **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 **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 5^{th} and 95^{th} 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 5^{th} and 95^{th} 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 5^{th} percentile of the minimums and 95^{th} 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 **X**) – (numbers of elements in **G** and **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.

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 (NH_3) 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 NH_3, 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 meters 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 PM_1, 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 ($C_3H_7^+$ fragment) and 55 ($C_4H_7^+$ 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 ($C_2H_4O_2^+$ 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 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 Figure 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., r2 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/Qexpected 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 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 Figure 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 5^{th} and 95^{th} 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 5^{th} 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/Qexpected*

289 going from six to seven modelled factors indicate that the seven-factor solution is not stable and
290 likely should not be used.

291 **3.2 Milwaukee Water Results**

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

309 **3.3 Las Vegas HR-AMS Results**

310 As described by Brown et al. (2012), four factors were determined: LV-OOA, HOA, SV-OOA
311 and BBOA. On average, HOA made up 28% of the organic matter (OM), had an abundance of
312 *m/z* 43, and peaked during the morning and evening commute periods coincident with peak
313 traffic volume. LV-OOA, indicated by an abundance of *m/z* 44, was highest in the afternoon and
314 accounted for 25% of the OM. BBOA occurred in the evening hours, was predominantly from
315 the residential area to the north, and on average constituted 10% of the OM. SV-OOA accounted
316 for the remaining 37% of the OM, and had an abundance of *m/z* 57 and 55. The HOA and
317 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 Figure 5 and EE diagnostics are summarized in Table 4, Supplemental Table 6, and in Figure 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 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., 2005; Paatero et al., 2002). 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 that 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 towards 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 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 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.

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).

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 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 **X**) – (numbers of elements in **G** and **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 **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*.

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.

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.

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

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.

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*detection limit, even if the original measured values have been available. It appears that this practice has

462 no proven advantages when species with low S/N are downweighted. On the other hand, it may
463 be demonstrated that the substitution practice prevents uncertainty estimation, introduces hard-
464 to-estimate bias, and occasionally gives rise to ghost factors. In general, it is a modeling error if
465 BDL values are replaced by a fraction of the detection limit. If such a replacement has been
466 done, then EE of PMF results should not be attempted because none of the available EE methods
467 is able to estimate the bias error incurred in results by censoring BDL values. If EE is
468 nevertheless attempted in the presence of such censoring, then a clear warning about
469 questionable validity of quoted error estimates must be included in the paper. Instead of
470 substitution methods, PMF modeling using ME-2 directly can be optimized by applying a
471 specific errormodel code to censored data values (Paatero, 2000). In this way the known
472 information, e.g., that a measurement is somewhere between zero and the detection limit, can be
473 conveyed to ME-2 without any substitutions that would likely bias the results.

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

480 *Treatment of "total mass."* In the Sacramento example, uncertainties for total $PM_{2.5}$ mass were
481 used as reported, rather than being further downweighted (e.g., (Reff et al., 2007); (Kim et al.,
482 2005)). $PM_{2.5}$ mass should be downweighted if there are likely significant measurement artifacts,
483 e.g., gaseous species adsorbing onto filters, or if there are sources that may emit $PM_{2.5}$ mass but
484 none of the measured species, in which case a factor containing only mass could be determined.
485 In these cases, the inclusion of "full strength" total mass does not help in interpretation of
486 solutions, and could lead to erroneous results.

487 *Use of constraints.* EPA PMF and ME-2 allow users to constrain or "pull" elements in their
488 solution. One common reason for pulling is the attempt to align an oblique "edge" in G-space
489 plots. However, an oblique edge may sometimes be justified because factors in atmospheric or
490 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.

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 5^{th} and 95^{th} 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 factors' identifying species.

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*.

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 5^{th} and 95^{th}), 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, source activity, etc.,

519 rather than a poor solution. BS-DISP may yield factor swaps even if BS and DISP diagnostics
520 are positive, and can be used to identify which factors are more certain than others. One
521 drawback of BS-DISP is its computation time. A large run (e.g., thousands of samples and a
522 hundred species) may take tens of hours on a modern PC. Future work may need to focus on
523 optimization of the algorithms in ME-2 to help significantly decrease the run time.

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

535 For all EE methods, the interval ratio of each factor's identifying species can provide an
536 understanding of the relative certainty of each factor's identity. However, each data set is unique,
537 so results will vary. It must be emphasized that, in some cases, a satisfactory analysis cannot be
538 performed with any number of factors. The following contrived example illustrates this situation.
539 With four factors, the result is rotationally unique and all three EE methods indicate small
540 uncertainties. However, with four factors, the fit is not satisfactory as indicated, e.g., by poor
541 total mass reconstruction, by poor factor interpretability, etc. With five factors, fit is good but
542 rotational uncertainty is very large and/or there are frequent factor swaps between factors four
543 and five. What should the scientist do in this situation? It would be wrong to only report either
544 the four-factor or the five-factor results while ignoring the presence of the alternative solution.
545 The information in the data set confirms that four factors are not enough. However, the
546 information is not sufficient for quantitative determination of five factors. If additional
547 information cannot be inserted for obtaining rotational uniqueness, then the two sets of
548 inconclusive results (using four and five factors, respectively), as well as the impact of

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.

6. 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.

REFERENCES

Alfarra M.R., Prévôt A.S.H., Szidat S., Sandradewi J., Weimer S., Lanz V.A., Schreiber D., Mohr M., and Baltensperger U. (2007) Identification of the mass spectral signature of organic aerosols from wood burning emissions. *Environ. Sci. Technol.*, 41, 5770-5777, doi: 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., and Worsnop D.R. (2003) Quantitative sampling using an Aerodyne aerosol mass spectrometer - 2. Measurements of fine particulate chemical composition in two U.K. cities. *Journal of Geophysical Research-Atmospheres*, 108, (D3), Feb 4.

Brown S.G., Lee T., Norris G.A., Roberts P.T., Collett J.L., Jr., Paatero P., and Worsnop D.R. (2012) Receptor modeling of near-roadway aerosol mass spectrometer data in Las Vegas,

Nevada, with EPA PMF. *Atmospheric Chemistry and Physics*, 12, 309-325, doi: 10.5194/acp-12-309-2012. Available at http://www.atmos-chem-phys.net/12/309/2012/.

Bzdusek P.A., Lu J., and Christensen E.R. (2006) PCB Congeners and Dechlorination in Sediments of Sheboygan River, Wisconsin, Determined by Matrix Factorization. *Environ. Sci. Tech.*, 40, 120-129, (1). Available at 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., and Worsnop D.R. (2007) Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrometry Reviews*, 26, 185-222, doi: 10.1002/mas.20115 (2). 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., and 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, doi: doi: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., and Jimenez J.L. (2006) Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. *Analytical Chemistry (Washington)*, 78, 8281-8289, doi: 10.1021/ac061249n (24).

Donahue N.M., Kroll J.H., Pandis S.N., and Robinson A.L. (2012) A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution. *Atmos. Chem. Phys.*, 12, 615-634, doi: 10.5194/acp-12-615-2012.

Hopke P.K. (2008) The use of source apportionment for air quality management and health assessments. *Journal of Toxicology and Environmental Health, Part A*, 71, 555-563.

Jayne J.T., Leard D.C., Zhang X.F., Davidovits P., Smith K.A., Kolb C.E., and Worsnop D.R. (2000) Development of an aerosol mass spectrometer for size and composition analysis of submicron particles. *Aerosol Science & Technology*, 33, 49-70, (1-2).

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. (2003) Ambient aerosol sampling using the Aerodyne Aerosol Mass Spectrometer. *Journal of Geophysical Research-Atmospheres*, 108, (D7), Apr 15.

618 Jimenez J.L., Canagaratna M.R., Donahue N.M., Prévôt A.S.H., Zhang Q., Kroll J.H., DeCarlo
619 P.F., Allan J.D., Coe H., Ng N.L., Aiken A.C., Docherty K.S., Ulbrich I.M., Grieshop
620 A.P., Robinson A.L., Duplissy J., Smith J.D., Wilson K.R., Lanz V.A., Hueglin C., Sun
621 Y.L., Tian J., Laaksonen A., Raatikainen T., Rautiainen J., Vaattovaara P., Ehn M.,
622 Kulmala M., Tomlinson J.M., Collins D.R., Cubison M.J., Dunlea E.J., Huffman J.A.,
623 Onasch T.B., Alfarra M.R., Williams P.I., Bower K., Kondo Y., Schneider J., Drewnick
624 F., Borrmann S., Weimer S., Demerjian K., Salcedo D., Cottrell L., Griffin R., Takami
625 A., Miyoshi T., Hatakeyama S., Shimono A., Sun J.Y., Zhang Y.M., Dzepina K., Kimmel
626 J.R., Sueper D., Jayne J.T., Herndon S.C., Trimborn A.M., Williams L.R., Wood E.C.,
627 Middlebrook A.M., Kolb C.E., Baltensperger U., and Worsnop D.R. (2009) Evolution of
628 organic aerosols in the atmosphere. *Science*, 326, 1525-1529, doi:
629 10.1126/science.1180353, December 11.

630 Kim E., Hopke P.K., Kenski D.M., and Koerber M. (2005) Sources of fine particles in a rural
631 midwestern U.S. area. *Environ. Sci. Technol.*, 39, 4953-4960, (13).

632 Kim E. and Hopke P.K. (2007) Source identifications of airborne fine particles using positive
633 matrix factorization and U.S. Environmental Protection Agency positive matrix
634 factorization. *J. Air Waste Manage.*, 57, 811-819, doi: 10.3155-1047-3289.57.7.811 (7),
635 July.

636 Kroll J.H., Donahue N.M., Jimenez J.L., Kessler S.H., Canagaratna M.R., Wilson K.R., Altieri
637 K.E., R.Mazzoleni L., Wozniak A.S., Bluhm H., Mysak E.R., Smith J.D., Kolb C.E., and
638 Worsnop D.R. (2011) Carbon oxidation state as a metric for describing the chemistry of
639 atmospheric organic aerosol. *Nature Chemistry*, 3, 133-139, doi: 10.1038/NCHEM.948.

640 Lanz V.A., Alfarra M.R., Baltensperger U., Buchmann B., Hueglin C., Szidat S., Wehrli M.N.,
641 Wacker L., Weimer S., Caseiro A., Puxbaum H., and Prévôt A.S.H. (2008) Source
642 attribution of submicron organic aerosols during wintertime inversions by advanced
643 factor analysis of aerosol mass spectra. *Environ. Sci.Technol.*, 42, 214-220, doi:
644 10.1021/es0707207 (1).

645 Lanz V.A., Prévôt A.S.H., Alfarra M.R., Weimer S., Mohr C., DeCarlo P.F., Gianini M.F.D.,
646 Hueglin C., Schneider J., Favez O., D'Anna B., George C., and Baltensperger U. (2010)
647 Characterization of aerosol chemical composition with aerosol mass spectrometry in
648 Central Europe: an overview. *Atmos. Chem. Phys.*, 10, 10453-10471, doi: 10.5194/acp-
649 10-10453-2010, November 8.

650 Lee T., Sullivan A.P., Mack L., Jimenez J.L., Kreidenweis S.M., Onasch T.B., Worsnop D.R.,
651 Malm W., Wold C.E., Hao W.M., and Jeffrey L. Collett J. (2010) Chemical smoke
652 marker emissions during flaming and smoldering phases of laboratory open burning of
653 wildland fuels. *Aerosol Science & Technology*, 44, i-v, doi:
654 10.1080/02786826.2010.499884 (9).

655 McCarthy M.C., Aklilu Y.-a., Brown S.G., and Lyder D.A. (2013) Source apportionment of
656 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., and 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, 2443-2449, doi: 10.1021/es8011518 (7).

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., and Worsnop D.R. (2010) Organic aerosol components observed in Northern Hemispheric datasets from Aerosol Mass Spectrometry. *Atmos. Chem. Phys.*, 10, 4625-4641, doi: 10.5194/acp-10-4625-2010.

Norris G., Duvall R., Brown S., and 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. and 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. (1997) Least squares formulation of robust non-negative factor analysis. *Chemometrics and Intelligent Laboratory Systems*, 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., Hopke P.K., Song X.H., and Ramadan Z. (2002) Understanding and controlling rotations in factor analytic models. *Chemometrics and Intelligent Laboratory Systems*, 60, 253-264.

Paatero P., Hopke P.K., Hoppenstock J., and Eberly S.I. (2003) Advanced factor analysis of spatial distributions of $PM_{2.5}$ in the eastern United States. *Environ. Sci. Technol.*, 37, 2460-2476, (11), June 1.

Paatero P., Hopke P.K., Begum B.A., and 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., and Norris G.A. (2014) Methods for estimating uncertainty in factor analystic solutions. *Atmospheric Measurement Techniques*, 7, 781-797, doi: 10.5194/amt-7-781-2014. Available at http://www.atmos-meas-tech.net/7/781/2014/amt-7-781-2014.html.

693 Poirot R.L., Wishinski P.R., Hopke P.K., and Polissar A.V. (2001) Comparative application of
694 multiple receptor methods to identify aerosol sources in northern Vermont. *Environ. Sci.*
695 *Technol.*, 35, 4622-4636, (23), November.

696 Reff A., Eberly S.I., and Bhave P.V. (2007) Receptor modeling of ambient particulate matter
697 data using positive matrix factorization: review of existing methods. *J. Air Waste*
698 *Manage.*, 57, 146-154.

699 Sage A.M., Weitkamp E.A., Robinson A.L., and Donahue N.M. (2008) Evolving mass spectra of
700 the oxidized component of organic aerosol: results from aerosol mass spectrometer
701 analyses of aged diesel emissions. *Atmos. Chem. Phys.*, 8, 1139-1152.

702 Soonthornnonda P. and Christensen E.R. (2008) Source apportionment of pollutants and flows of
703 combined sewer wastewater. *Water Research*, 42, 1989-1998, doi:
704 http://dx.doi.org/10.1016/j.watres.2007.11.034 (8-9), April. Available at
705 http://www.sciencedirect.com/science/article/pii/S0043135407007300.

706 Ulbrich I.M., Canagaratna M.R., Zhang Q., Worsnop D.R., and Jimenez J.L. (2009)
707 Interpretation of organic components from Positive Matrix Factorization of aerosol mass
708 spectrometric data. *Atmos. Chem. Phys.*, 9, 2891-2918.

709 Zhang Q., Jimenez J.L., Canagaratna M.R., Ulbrich I.M., Ng N.L., Worsnop D.R., and Sun Y.
710 (2011) Understanding atmospheric organic aerosols via factor analysis of aerosol mass
711 spectrometry: a review. *Anal. Bioanal. Chem.*, doi: 10.1007/s00216-011-5355-y
712 (October).

713

714

715

Figure Captions

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

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

Figure 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.

Figure 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.

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

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

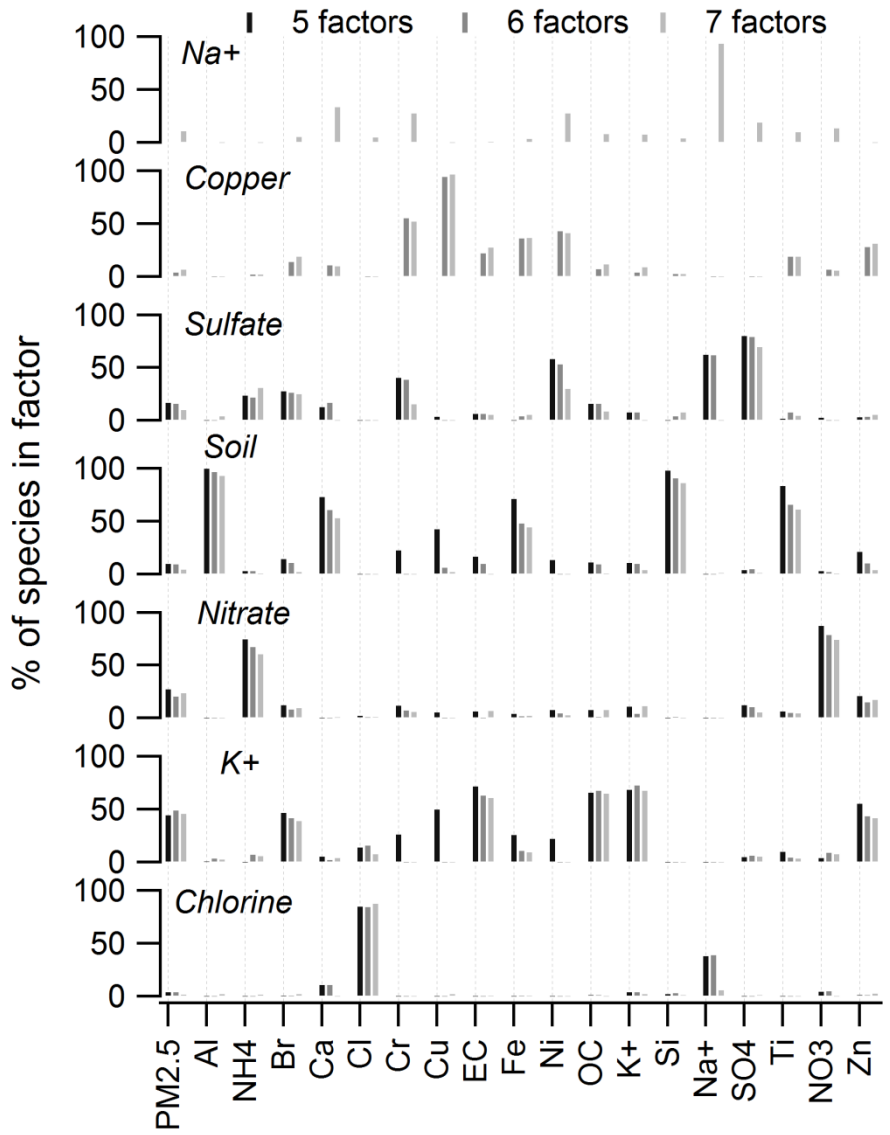

Figure 1. Factor profiles (fraction of each species by factor) using Sacramento $PM_{2.5}$ data, for five, six and seven factors.

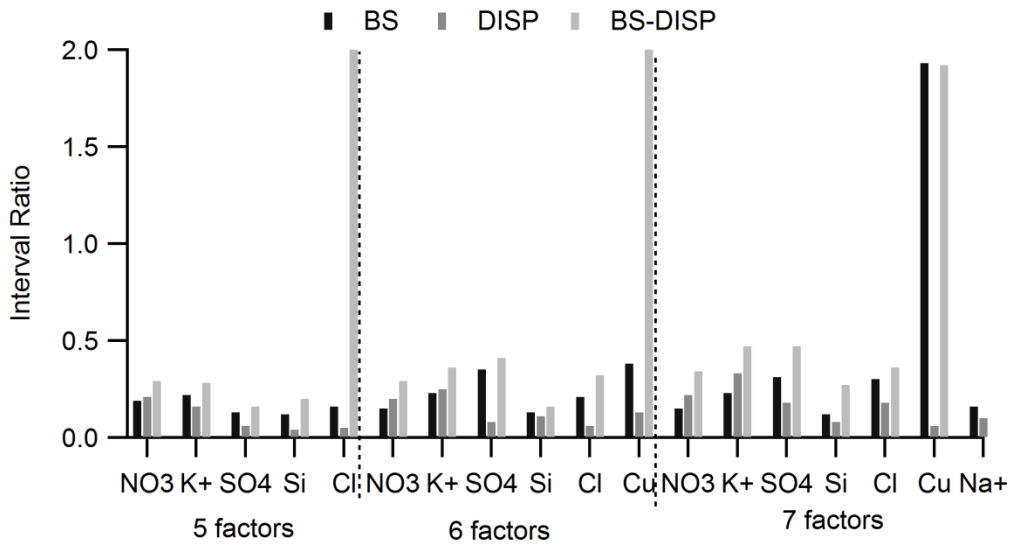

Figure 2. Interval ratios of identifying species by EE method and number of factors (5, 6, 7) for Sacramento $PM_{2.5}$ data.

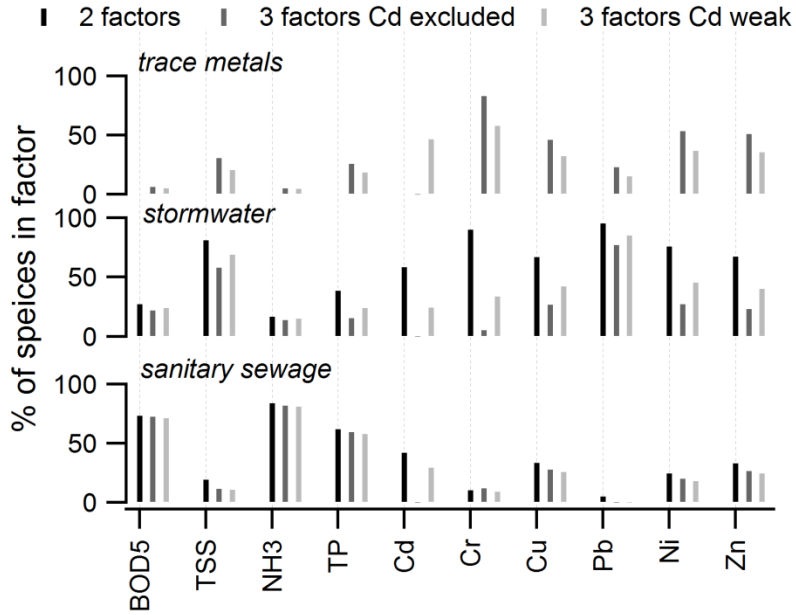

Figure 3. Factor profiles (fraction of each species by factor) using Milwaukee water data, for two factors, three factors with Cd excluded, and three factors with Cd weak.

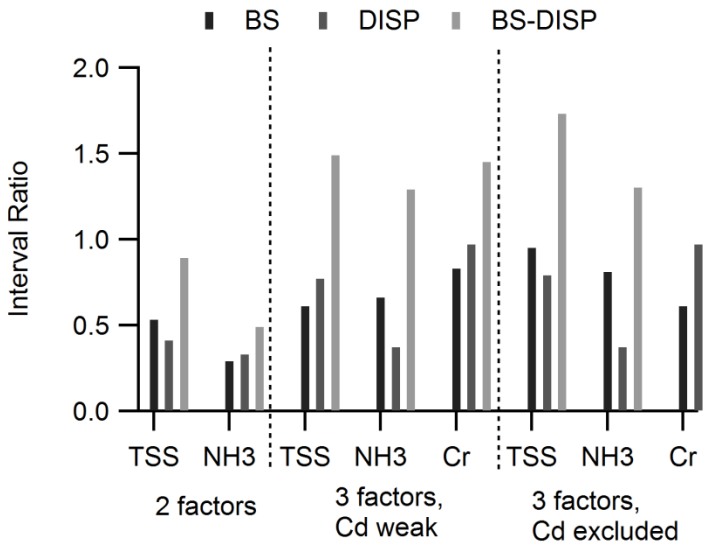

Figure 4. Interval ratios of identifying species by EE method and run (two factors, three factors with Cd weak, and three factors with Cd excluded) for Milwaukee water data.

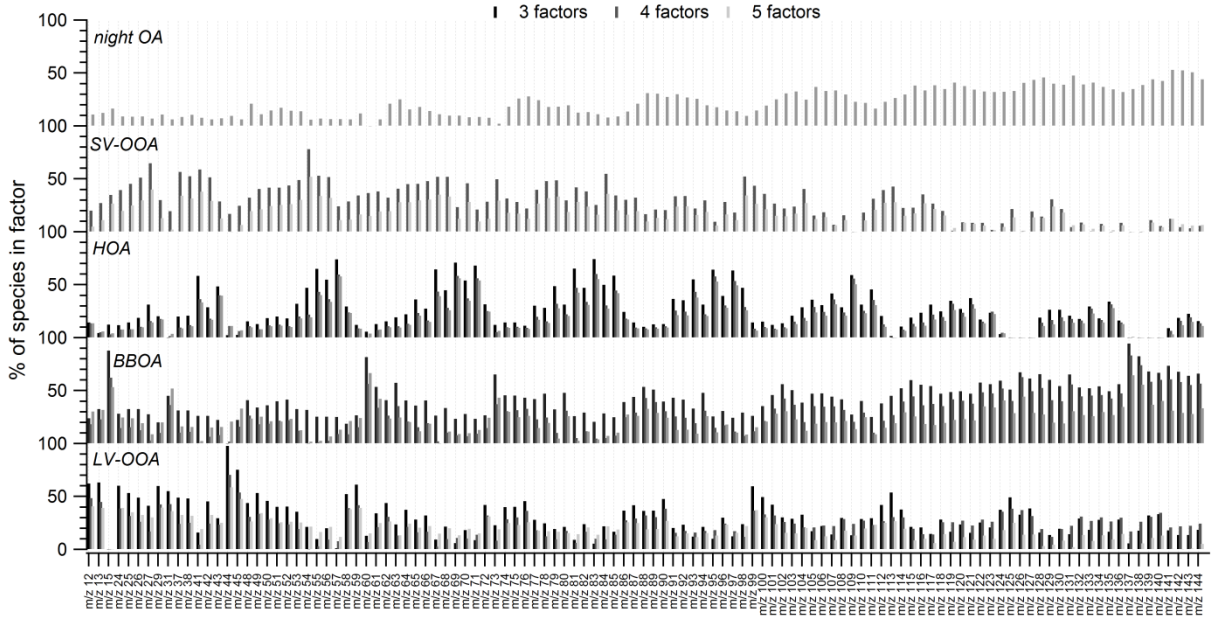

Figure 5. Factor profiles (fraction of each species by factor) using Las Vegas AMS data, for three, four and five factors.

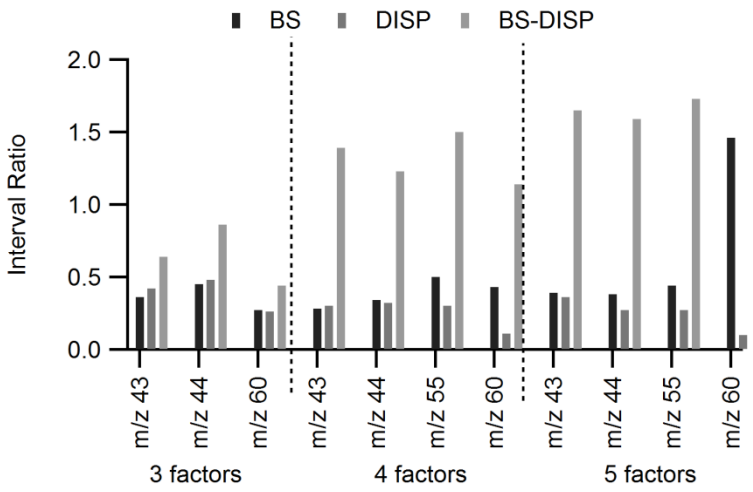

Figure 6. Interval ratios of identifying species by EE method and number of factors (3, 4, 5) for Las Vegas data.

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 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 species	K^+, SO_4, NO_3, Si, $Na+^*$, Cl^{**}, 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*** for DISP, BS-DISP	1 hour, 3 hours	< 1 hour, < 1 hour	9 hours, 16 hours

* Only used in 7-factor solution

** Only used in 6- and 7-factor solutions

*** Windows 7 64-bit, 3.1 GHz processor, 4 GB RAM

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}	65793	52241	45651
Q_{robust}	48929	40604	35572
$Q_{robust}/Q_{expected}$	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%

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%

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

Diagnostic	**3 factors**	**4 factors**	**5 factors**
$Q_{expected}$	137351	135833	134315
Q_{true}	920029	607201	402907
Q_{robust}	454015	340269	264455
$Q_{robust}/Q_{expected}$	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

Supplemental Information

Revised Signal/Noise Calculation in EPA PMF

The signal/noise (S/N) calculation in EPA PMF has been revised for version 5.0. Previously, S/N of a given species was essentially the sum of the concentration values divided by the sum of the uncertainty values. While reasonable, this could lead to different problems in certain specific situations. Artificially high S/N values would be obtained in species with a handful of high concentration events, resulting in an S/N that may actually be higher than another species' S/N with a more consistent signal. More seriously, artificially low S/N values could appear for species with a few missing values. Missing values are usually downweighted by very large uncertainty values, typically (much) larger than the largest concentration values in the species in question. If uncertainties for missing values were inflated prior to input into EPA PMF, such inflated uncertainty values will inflate the noise in S/N calculations, resulting in an S/N that actually will be small enough to cause the classification of a strong variable as "weak." The latter problem has repeatedly been observed in practical work. In addition, the presence of slightly negative concentration values, not uncommon in environmental data, could artificially decrease S and hence the S/N of a species.

In the revised calculation, only concentration values that exceed their uncertainty contribute to the signal portion of the S/N calculation, since the concentration value is essentially equal to the sum of signal and noise, and therefore signal is the difference between concentration and uncertainty.

Two calculations are performed to determine S/N, where concentrations that are below uncertainty are determined to have no signal, and for concentrations above uncertainty, the difference between concentration (x_{ij}) and uncertainty (s_{ij}) is used as the signal:

$$d_{ij} = \left(\frac{x_{ij} - s_{ij}}{s_{ij}} \right) \quad \text{if } x_{ij} > s_{ij}$$
$$d_{ij} = 0 \qquad \text{if } x_{ij} \le s_{ij} \tag{3}$$

S/N is then calculated as:

$$\left(\frac{S}{N} \right)_j = 1/n \sum_{i=1}^{n} d_{ij} \tag{4}$$

The result with this new S/N calculation is that species with concentrations always below their uncertainties have an S/N of 0. Species with concentrations that are twice their uncertainty values have an S/N of 1. An S/N greater than 1 often indicates a species with "good" signal, though this depends on how uncertainties are determined. Negative concentration values do not

contribute to the S/N, and species with a handful of high concentration events will not have artificially high S/N. While there are many methods to determine S/N, the one selected in the new version of EPA PMF may be more useful in environmental data analysis compared to the prior version, with the caveat that the S/N is merely one of many analyses for screening data. All S/N values reported in this paper are calculated according to equations (3) and (4).

Input Data Summary Statistics

Supplemental Table 1. Summary of Sacramento $PM_{2.5}$ input data. Concentration units in $\mu g/m^3$.

Species	Category	S/N	Min	25th	Median	75th
$PM_{2.5}$	Strong	9.78	0.8	6.7	9.9	15.825
Aluminum	Strong	1.56	0	0	0.01155	0.034825
Ammonium Ion	Strong	9.96	0.0102	0.30175	0.5355	0.98025
Antimony	Bad	0.15	0	0	0	0.00233
Arsenic	Bad	0.40	0	0	0.00008	0.00097
Barium	Bad	0.38	0	0	0	0.0051975
Bromine	Strong	2.40	0	0.0011075	0.00234	0.00413
Calcium	Strong	7.07	0	0.013475	0.02465	0.042525
Chlorine	Strong	4.10	0	0	0.01145	0.048675
Chromium	Strong	1.16	0	0	0.00034	0.00205
Cobalt	Bad	0.11	0	0	0	0.00028
Copper	Strong	4.39	0	0.00163	0.0036	0.0078025
EC	Strong	1.65	0	0.329	0.543	1.065
Iron	Strong	9.75	0	0.0395	0.06085	0.101
Lead	Bad	0.55	0	0	0.00082	0.00292
Manganese	Bad	0.84	0	0	0.00089	0.00187
Mercury	Bad	0.22	0	0	0	0.00117
Molybdenum	Bad	0.04	0	0	0	0
Nickel	Strong	1.23	0	0	0.000595	0.00139
OC	Strong	7.63	1.45	3	4.62	7.1325
Phosphorus	Bad	0.05	0	0	0	0
Potassium Ion	Bad	5.82	0	0	0.0564	0.113
Potassium	Strong	9.15	0.00058	0.0372	0.0646	0.118
Rubidium	Bad	0.13	0	0	0	0.0005225
Selenium	Bad	0.14	0	0	0	0.0007425
Silicon	Strong	6.46	0	0.0278	0.06285	0.118
Sodium Ion	Strong	2.77	0	0.062475	0.119	0.23075
Strontium	Bad	0.37	0	0	0.00009	0.001365
Sulfate	Strong	8.68	0.115	0.58275	0.9255	1.4525
Sulfur	Bad	9.92	0	0.204	0.315	0.50025
Tin	Bad	0.12	0	0	0	0.0036125
Titanium	Strong	1.26	0	0	0.001805	0.0049
Total Nitrate	Strong	9.08	0.0836	0.62675	1.125	2.45
Vanadium	Bad	0.61	0	0	0.0007	0.0019925
Yttrium	Bad	0.13	0	0	0	0.00026
Zinc	Strong	4.10	0	0.00173	0.003955	0.00753

Supplemental Table 2. Summary of Milwaukee input data. Concentration units in mg/L.

Species	Category	S/N	Min	25th	Median	75th
BOD5	Strong	4	2	6.4	15	16.5
TSS	Strong	4	6.7	27.5	41	73
NH3	Strong	4	0.095	0.485	0.81	1.1
TP	Strong	4	0.41	0.585	0.87	1.15
Cd	Bad/Weak	4	0.0009	0.0009	0.002	0.002
Cr	Strong	4	0.0009	0.0039	0.007	0.013
Cu	Strong	4	0.007	0.0125	0.017	0.026
Pb	Strong	4	0.0008	0.0088	0.019	0.049
Ni	Strong	4	0.0014	0.0023	0.003	0.006
Zn	Strong	4	0.023	0.0565	0.079	0.11

Supplemental Table 3. Summary of Las Vegas HR-AMS input data. Concentration units in μg/m^3.

Species	Category	S/N	Min	25th	Median	75th
mz12	Strong	6.64027	0.0076	0.02108	0.03222	0.0526
mz13	Strong	5.99052	0.00075	0.00497	0.0094	0.01707
mz15	Weak	4.23978	-0.0801	0.00064	0.01668	0.06972
mz24	Weak	4.59659	-0.0002	0.0018	0.00328	0.00574
mz25	Strong	7.34043	0.00026	0.0057	0.01083	0.01975
mz26	Strong	9.32493	0.00655	0.02813	0.05505	0.10129
mz27	Strong	9.79196	0.01416	0.06399	0.12737	0.23371
mz29	Weak	9.97133	0.03103	0.15862	0.24624	0.39205
mz31	Strong	7.43832	0.00048	0.01066	0.0209	0.04278
mz37	Weak	3.59783	-0.0017	0.00577	0.01134	0.0208
mz38	Strong	8.12982	0.0024	0.01011	0.01967	0.03596
mz41	Strong	9.74817	0.0164	0.09251	0.18074	0.37634
mz42	Strong	9.83868	0.00827	0.04005	0.07897	0.14489
mz43	Strong	9.98527	0.02913	0.12036	0.24026	0.45828
mz44	Strong	8.64144	-0.0074	0.08723	0.17112	0.26699
mz45	Strong	8.60323	0.00131	0.01283	0.02361	0.03864
mz48	Strong	6.94685	6.3E-05	0.00096	0.00201	0.00392
mz49	Strong	5.20763	4E-05	0.0014	0.0028	0.00509
mz50	Strong	8.50717	0.00159	0.00765	0.01551	0.02955
mz51	Strong	9.20871	0.00222	0.0111	0.02312	0.04431
mz52	Strong	8.75171	0.00143	0.00664	0.01373	0.02653
mz53	Strong	9.79787	0.00477	0.01865	0.03791	0.07122
mz54	Strong	9.67074	0.00266	0.01394	0.02697	0.05584
mz55	Strong	9.97195	0.01369	0.07951	0.15732	0.33873
mz56	Strong	9.67613	0.00362	0.02261	0.04571	0.09289
mz57	Strong	9.87256	0.00622	0.04157	0.08792	0.19425
mz58	Strong	9.07492	0.00132	0.00709	0.01523	0.02727
mz59	Weak	7.94741	0.00054	0.00286	0.00573	0.0099
mz60	Strong	8.58301	0.00082	0.00528	0.01171	0.03901
mz61	Strong	7.36074	0.00022	0.00181	0.00392	0.00892
mz62	Strong	6.94685	0.00013	0.00192	0.00401	0.00785
mz63	Strong	7.29958	-0.0057	0.0031	0.00663	0.01681
mz64	Strong	9.00247	0.00144	0.00629	0.01297	0.0255
mz65	Strong	9.8444	0.00327	0.01278	0.02562	0.05181
mz66	Strong	8.05736	0.00068	0.00448	0.00958	0.0192
mz67	Strong	9.91404	0.00489	0.02657	0.05229	0.11071
mz68	Strong	9.59442	0.00267	0.01073	0.02175	0.04491
mz69	Strong	9.9086	0.00495	0.03016	0.06224	0.13184
mz70	Strong	8.70116	0.00165	0.00908	0.01785	0.03655
mz71	Strong	9.47584	0.00225	0.01366	0.0271	0.05754
mz72	Strong	6.76481	0.00036	0.00193	0.00371	0.00711
mz73	Strong	8.57921	0.00077	0.00399	0.00902	0.0234
mz74	Strong	7.5283	0.00044	0.00222	0.00464	0.00959
mz75	Strong	6.5182	0.00012	0.00146	0.00311	0.00633
mz76	Strong	6.20329	-0.0001	0.00177	0.00371	0.00716
mz77	Strong	9.38165	0.00172	0.00994	0.0208	0.0431

mz78	Strong	7.89004	0.0009	0.00477	0.00982	0.02111
mz79	Strong	9.78718	0.00331	0.01445	0.02844	0.05865
mz80	Strong	7.72935	0.00023	0.0018	0.00384	0.00843
mz81	Strong	9.94842	0.00327	0.01739	0.0347	0.07285
mz82	Strong	9.04364	0.00086	0.00587	0.01178	0.02386
mz83	Strong	9.63347	0.00224	0.01131	0.02262	0.04842
mz84	Strong	8.44	0.00079	0.00421	0.00847	0.01759
mz85	Strong	8.88894	-0.001	0.00554	0.01081	0.02274
mz86	Strong	6.36324	-0.0108	0.00105	0.0022	0.00456
mz87	Strong	7.82926	0.00022	0.00187	0.00367	0.00743
mz88	Strong	5.46014	-3E-05	0.00061	0.00123	0.00266
mz89	Strong	7.62395	0.00029	0.00172	0.00377	0.00795
mz90	Strong	5.85027	8.1E-05	0.00071	0.00155	0.00304
mz91	Strong	9.09518	0.00156	0.0101	0.02115	0.04465
mz92	Strong	5.76976	0.00011	0.00233	0.00504	0.01036
mz93	Strong	7.99309	0.00054	0.00508	0.01104	0.0226
mz94	Strong	7.72935	0.0003	0.0024	0.00512	0.01123
mz95	Strong	9.51552	0.00155	0.0081	0.01693	0.03497
mz96	Strong	7.94093	0.00039	0.00262	0.00544	0.01051
mz97	Strong	8.8534	-0.0009	0.00464	0.00929	0.01971
mz98	Strong	8.45931	0.00047	0.00234	0.00477	0.0099
mz99	Strong	7.43477	-0.001	0.00172	0.00334	0.00594
mz100	Strong	5.22097	-0.0001	0.00063	0.00129	0.00236
mz101	Strong	7.10802	5.9E-05	0.00117	0.00229	0.00464
mz102	Strong	6.60916	4.2E-05	0.00103	0.00224	0.00502
mz103	Strong	6.74811	7.4E-05	0.0016	0.00345	0.00747
mz104	Strong	4.77613	-8E-05	0.00118	0.00254	0.00521
mz105	Strong	8.19342	0.00034	0.00346	0.00763	0.01698
mz106	Strong	6.62573	-6E-06	0.00129	0.00272	0.00589
mz107	Strong	8.5247	0.00044	0.00272	0.00586	0.01285
mz108	Strong	7.26816	0.00021	0.00122	0.00259	0.00523
mz109	Strong	9.04447	-0.0005	0.00304	0.00648	0.01318
mz110	Strong	7.07327	0.00018	0.00115	0.00243	0.00493
mz111	Strong	7.34289	-0.0008	0.00147	0.00295	0.00596
mz112	Weak	3.83745	-0.0042	0.00051	0.00114	0.00218
mz113	Strong	6.10393	-0.0032	0.00077	0.00169	0.00311
mz114	Strong	5.29001	-0.0002	0.00045	0.00096	0.00196
mz115	Strong	8.60069	0.00061	0.00291	0.00618	0.01486
mz116	Strong	6.66794	2.9E-05	0.00095	0.00211	0.00484
mz117	Strong	7.53592	0.00024	0.00161	0.00363	0.00832
mz118	Strong	6.12417	4.9E-05	0.00082	0.00185	0.00398
mz119	Strong	7.60066	0.00018	0.00155	0.00354	0.008
mz120	Strong	6.01684	-7E-05	0.00071	0.00159	0.00352
mz121	Strong	6.97556	0.00015	0.00135	0.00292	0.00661
mz122	Strong	5.29943	-7E-05	0.00064	0.0014	0.00325
mz123	Strong	6.34159	-0.0016	0.00083	0.00193	0.00445
mz124	Strong	4.72048	-0.0009	0.00037	0.00082	0.0018
mz125	Weak	3.27382	-0.0066	0.00025	0.00072	0.00152

mz126	Strong	4.55255	-0.0071	0.00037	0.0008	0.00196
mz127	Strong	6.01397	-0.0019	0.00078	0.00167	0.00368
mz128	Strong	7.18673	0.00015	0.00133	0.00299	0.00734
mz129	Strong	7.35485	0.00021	0.00126	0.00284	0.00705
mz130	Weak	3.79174	-0.0003	0.00049	0.00116	0.00262
mz131	Strong	6.81728	6.9E-05	0.00094	0.00231	0.00566
mz132	Weak	3.38582	-0.0002	0.0005	0.00109	0.00249
mz133	Strong	6.27521	6.5E-05	0.00093	0.00217	0.00511
mz134	Weak	4.44894	-0.0001	0.00045	0.00102	0.00233
mz135	Strong	5.54901	-0.0002	0.00069	0.00156	0.00352
mz136	Weak	4.51851	-0.0002	0.00032	0.00073	0.00168
mz137	Strong	6.38935	-0.001	0.00054	0.0014	0.00516
mz138	Strong	4.6425	-0.0011	0.00024	0.00055	0.0017
mz139	Strong	6.37272	-0.0022	0.00057	0.00129	0.00302
mz140	Weak	3.89419	-0.0011	0.00018	0.00041	0.00097
mz141	Strong	7.09137	-0.0001	0.00073	0.00171	0.00447
mz142	Strong	5.63898	-0.0001	0.00042	0.00098	0.00243
mz143	Strong	6.21922	4.9E-06	0.00054	0.00122	0.00302
mz144	Strong	5.02847	-8E-05	0.00033	0.00072	0.00185

Error Estimation Result Details

Supplemental Table 4. Summary of EE interval ratios by factor and key species for five-, six-, and seven-factor solutions with Sacramento PM2.5 data.

Number of factors	Factor	Key species	BS interval ratio	DISP interval ratio	BS-DISP interval ratio
5	NO_3	NO_3	0.19	0.21	0.29
6	NO_3	NO_3	0.15	0.20	0.29
7	NO_3	NO_3	0.15	0.22	0.34
5	K+/OC/EC	K+	0.22	0.16	0.28
6	K+/OC/EC	K+	0.23	0.25	0.36
7	K+/OC/EC	K+	0.23	0.33	0.47
5	SO_4	SO_4	0.13	0.06	0.16
6	SO_4	SO_4	0.35	0.08	0.41
7	SO_4	SO_4	0.31	0.18	0.47
5	soil	Si	0.12	0.04	0.20
6	Soil	Si	0.13	0.11	0.16
7	Soil	Si	0.12	0.08	0.27
5	Cl	Cl	0.16	0.05	2.00
6	Cl	Cl	0.21	0.06	0.32
7	Cl	Cl	0.30	0.18	0.36
6	Copper	Cu	0.38	0.13	2.00
7	Copper	Cu	1.93	0.06	1.92
7	Na	Na	0.16	0.10	1.96

Supplemental Table 5. Summary of EE interval ratios by factor and key species for Milwaukee water quality data.

Number of factors, scenario	Factor	Key species	BS interval ratio	DISP interval ratio	BS-DISP interval ratio
2	Stormwater	TSS	0.53	0.41	0.89
3, with Cd weak	Stormwater	TSS	0.61	0.77	1.49
3, excluding Cd	Stormwater	TSS	0.95	0.79	1.73
2	Sanitary sewage	NH3	0.29	0.33	0.49
3, with Cd weak	Sanitary sewage	NH3	0.66	0.37	1.29
3, excluding Cd	Sanitary sewage	NH3	0.81	0.37	1.30
3, with Cd weak	Metals	Cr	0.83	0.97	1.45
3, excluding Cd	Metals	Cr	0.61	0.97	2.00

Supplemental Table 6. Summary of EE Interval ratios by factor and key species for three-, four-, and five-factor solutions for Las Vegas AMS data.

Number of factors	Factor	Key Species (m/z)	BS interval ratio	DISP Interval ratio	BS-DISP interval ratio
3	BBOA	60	0.27	0.26	0.44
4	BBOA	60	0.43	0.11	1.14
5	BBOA	60	1.46	0.10	1.47
3	HOA	43	0.36	0.42	0.64
4	HOA	43	0.28	0.30	1.39
5	HOA	43	0.39	0.36	1.65
3	LV-OOA	44	0.45	0.48	0.86
4	LV-OOA	44	0.34	0.32	1.23
5	LV-OOA	44	0.38	0.27	1.59
3	SV-OOA	55	n/a	n/a	n/a
4	SV-OOA	55	0.50	0.30	1.50
5	SV-OOA	55	0.44	0.27	1.73

Supplemental Table 7. Sacramento EE intervals by N factors, key species and dQ^{max} value.

N factors	dQ^{max}	SO_4	NO_3	K	Cl	Si	Cu	Na+
5	4	0.06	0.21	0.16	0.05	0.04	n/a	n/a
5	8	0.09	0.24	0.22	0.08	0.06	n/a	n/a
5	16	0.13	0.27	0.29	0.11	0.08	n/a	n/a
6	4	0.08	0.20	0.25	0.06	0.11	0.13	n/a
6	8	0.10	0.23	0.31	0.08	0.15	0.16	n/a
6	16	0.13	0.26	0.39	0.11	0.18	0.20	n/a
7	4	0.18	0.22	0.33	0.18	0.08	0.06	0.10
7	8	0.27	0.25	0.39	0.22	0.12	0.08	0.15
7	16	0.37	0.29	0.44	0.25	0.16	0.12	0.21