Response to Referee #1

General comments

This work examines the PMF solutions for a partitioning-corrected SVOC data set in which total (gas + particle) concentrations calculated from aerosol-phase measurements are used for the input matrix, rather than the aerosol-phase measurements. This procedure is intended to (partially) remove the temperature dependence of the observations, leading to a more consistent estimation of PMF factors across seasons or periods with large variations in temperature. This is a well-written manuscript with important consequences for source apportionment for the atmospheric chemistry community and is therefore recommended for publication in ACP.

Response:

We appreciate the reviewer's recommendation for publication and comments. We have considered each comment carefully and given responses below.

Major comments

1. Section 2.3: It seems that the solutions are artificially restricted to the original (base case) solutions; possibly limiting the usefulness of the total PMF approach. Were the solutions that were thrown out (not matched) correspond to the extreme values of the sampled distributions, or not necessarily? Were they physically plausible to be considered, even though they did not match? Also, are the factors classified according to profile or strengths, or both (what is the input to the classification model)? In either case, it seems counter-intuitive to classify new solutions to old solutions you are trying to differentiate these solutions from, unless the inter-factor variability is much greater than the variability between G/P partitioning-corrected and uncorrected solutions, in which case the added value of G/P partitioning corrections is seems diminished.

Response:

We agree with the referee that the solutions thrown out (not matched) could be interesting to explore. Those unmatched solutions could reflect factor (or source) patterns for specific periods (e.g., cold, warm and hot) or events (e.g., wildfire). In this work, the bootstrap technique added to PMF model is used to estimate the bias and variability in factor contributions due to random sampling error (Hemann et al., 2009). The selection of base case solution for this refined model should satisfy two criteria: physically interpretable and high factor matching rate (at least higher than 50%). If the PMF solution generated from the original data set could not meet either of the two criteria, we will not use it as base case solution. So the generation of base case solution is also possibly from replicate data sets.

In this work, the PMF solution derived from the original data set is physically plausible and the factor matching rate is very high (79.9%). The high factor matching rate reflects the uniqueness of base case factors, robustness of the base case solution to input data set (Xie et al., 2012). If we use one of those unmatched bootstrap solutions as the base case, most of those bootstrapped solutions could not be matched to the base case.

The factors of those replicate data set solutions are classified according to factor profile, not the strength, which was discussed in detail by Hemman et al (2009). Multilayer Feed Forward Neural Networks were trained to perform pattern recognition, to classify factors of replicate solutions to base case solution. This approach avoids the sensitivity to outliers that is a problem for linear correlation.

Only those bootstrapped solutions for which factors could be uniquely matched are used to estimate the factor contribution uncertainty, and the matched factors have similar factor contribution time series as their corresponding factors in the base case solution. Then the inter-factor variability only determines the uncertainty of factor contribution due to random sampling error, but not the factor contribution time series. As shown in Figure 4, the G/P partitioning-corrected and uncorrected solutions are different in factor contribution time series. As a result, the added value of G/P partitioning corrections will not be diminished by inter-factor variability.

2. Section 3.4: Is the factor that you were additionally able to resolve with the total PMF significant to your interpretation of sources in the region (disregarding the loss of two factors in comparison to the base-case PMF)?

Response:

Yes. In the total SVOC-based PMF solution of this study, the additionally resolved factor is due to the separation of nitrate and sulfate. The nitrate and sulfate were combined in one factor (inorganic ion) in the particle-only PMF solution. In Table 3, we can see that both the nitrate and sulfate contribute a big part of $PM_{2.5}$. They reflect the formation of secondary ions.

3. Figs. 3 and 4: Assuming that the total PMF solution is the correct one, how much might you overestimate your average contribution from each corresponding source factor (in terms of OM) if you only used particle-phase measurements with the PMF without the vapor concentration corrections (the base case)? In the base case, is the mass apportioned to its own factor, incorrectly apportioned to another factor, or remains unapportioned?

Response:

As we mentioned at the end of section 3.4, the higher contributions of carbonaceous factors from particle-only solution compared to that from total SVOC-based solution is attributed to the fact that OC1 fraction was excluded for PMF analysis in the current study, and OC1 accounted for a big part (47.6%) of total OC.

In this work, OC fractions were used instead of total OC for PMF analysis of the total SVOC (gas + particle phase) data set. OC fractions (OC1 – OC4 and PC) represent the organic carbon measured at four different temperature steps and a pyrolized carbon adjustment. However, OC1 was excluded from PMF analysis since it was measured during the first temperature step (340 °C) and most likely influenced by G/P partitioning. While total OC was used for the particle only-based PMF analysis. Therefore, we only discuss the similarity in factor contribution time series between G/P partitioning-corrected and uncorrected solutions, but not the difference in absolute factor contributions.

Any change in input data set will lead to a different PMF solution. If we assume that the G/P partitioning corrected solution is correct and all bulk components (sulfate, nitrate, EC and OC fractions) could be apportioned to specific source related factors, the particle-only solution will overestimate or underestimate the PM mass in each corresponding source factor. However, OC1 could not be used for the PMF analysis of G/P partitioning corrected data set; a big part of OC fractions and EC could not be apportioned to specific source factors and were lumped in the Bulk carbon factor. So the current G/P partitioning corrected solution needs to be refined with more information on gas phase OC and specific source related organic markers, and we cannot determine if there was overestimation or incorrect apportionment of PM mass for the particle-only solution at this time.

4. Given the uncertainties mentioned in the estimation of the total concentrations, there are improvements observed in the consistency of sub-data PMF solutions. The authors state this but

in a less direct way, but is the conclusion of the manuscript that an uncertain G/P correction is better than no correction?

Response:

It is true, we did not conclude that an uncertain G/P correction is better than no correction, but we have demonstrated that using total SVOC data set can remove the influence of G/P partitioning on PMF analysis.

The objective of this manuscript is to demonstrate that G/P partitioning corrected data set could be used instead of uncorrected data set for PMF analysis, so as to avoid the influence of temperature dependent G/P partitioning. However, the method used to estimate gas phase SVOCs concentrations needs to be verified, and refined if necessary to decrease uncertainties, by field measurements. After the adjustment of the gas phase SVOC estimation, we will likely conclude that an uncertain (much lower than that in this manuscript) G/P correction is better than no correction, but are not able to make that definitive conclusion at this time. The adjustment of the gas phase SVOC estimation will be included in our later work.

5. Just a comment regarding one of the additional, unconsidered dependencies that the authors raise, but to a first order, the temperature-dependence of the kinetics of VOC oxidation is smaller than G/P partitioning according to Pathak et al. (2007; J. Geophys.Res.), but a) they only considered dark ozonolysis and not the full possibility of photochemistry, b) PMF separation may still be sensitive to the relative variations among compounds for its separation.

Response:

We appreciate the referee's comments on that unconsidered dependency, which was mentioned but not studied in this work. The degradation of SVOCs in the atmosphere is another process that could affect the SVOC-based PMF analysis. Different SVOC compounds vary in degradation rates, and the speed of degradation changes across the sampling period due to the variation in temperature, solar radiance, oxidants (e.g. ozone, hydroxyl radical) concentrations and so on. Therefore, ambient temperature is one of those factors impacting SVOCs degradation.

We will add ", which is partly dependent on ambient temperature (Crounse et al., 2011; Pathak et al., 2007) and not considered in this work." at the end of page 5215, line 20.

Minor comments:

1. Fig. 1: why are there lines connecting the species? This should be a barplot instead to show that the chemical species are not connected together in series.

Response:

We included the lines in Figure 1 to help the reader to identify the difference in particle phase fractions among the three periods. Figure 1 could not be represented with as much clarity in the form of a barplot or with the connecting lines removed.

2. Fig. 3: "filter-based" should possibly be renamed as "original" or "particle only" or something else as the total SVOC-based analysis is also based on filter measurement. Also in the text, the labels "original" and "base-case" are used alongside "filter-based". It would be better to define this label up front and use it throughout (but "filter-based" is discouraged as it is not discriminating enough).

Response:

We agree with the referee and “particle only-based” will be used instead of “filter-based” throughout the manuscript. The term “base case” corresponds to the base case solution derived from original data set, but not replicate data sets. The base case solution is compared to all bootstrapped solutions for factor matching rate calculation. So the term “base case” is not related to “particle only” or “filter-based”.

In section 2.3, the original data set is used to derive base case solution, while bootstrapped solutions are generated from replicate data sets. Thus, the term “original” is used to differentiate the base case solution and bootstrapped solutions, and has nothing to do with terms of “particle only” or “filter-based”. To be clear, we have changed “……….the original solution based on the observed data (known as the base case).” (page 5207, lines 10-11) into “…..the base case solution derived from the original data set.”

References

Crounse, J. D., Paulot, F., Kjaergaard, H. G., and Wennberg, P. O.: Peroxy radical isomerization in the oxidation of isoprene, Physical Chemistry Chemical Physics, 13, 13607-13613, 10.1039/c1cp21330j, 2011.

Hemann, J. G., Brinkman, G. L., Dutton, S. J., Hannigan, M. P., Milford, J. B., and Miller, S. L.: Assessing positive matrix factorization model fit: a new method to estimate uncertainty and bias in factor contributions at the measurement time scale, Atmospheric Chemistry and Physics, 9, 497-513, 2009.

Pathak, R. K., Stanier, C. O., Donahue, N. M., and Pandis, S. N.: Ozonolysis of α-pinene at atmospherically relevant concentrations: Temperature dependence of aerosol mass fractions (yields), Journal of Geophysical Research: Atmospheres, 112, D03201, 10.1029/2006jd007436, 2007.

Xie, M., Hannigan, M. P., Dutton, S. J., Milford, J. B., Hemann, J. G., Miller, S. L., Schauer, J. J., Peel, J. L., and Vedal, S.: Positive Matrix Factorization of PM2.5: Comparison and Implications of Using Different Speciation Data Sets, Environmental Science & Technology, 46, 11962-11970, 10.1021/es302358g, 2012.