We thank the reviewer for taking the time to read the article and for valuable comments on our paper. To facilitate the revision process, we have copied the reviewer comments (in black text) and our responses are in blue font. We have responded to all the reviewer comments and made alterations to our paper (**in bold text**).

Reviewer 1 # The quality and language of the manuscript have been significantly improved. However, the way in which the ACSM data is used to interpret different PMF factors still needs revision. I would recommend the manuscript be published once the following comments have been addressed.

Major Comments

1. If the dataset measured from the two sites has been analysed using PMF by Rivas et al. (2020), what will be the novelty of this study? Please provide a short statement in the Introduction.

AR: The end of the introduction chapter has now been modified to the following form to state the novelty of this study better although Rivas et al. (2020) have used partially the same data.

The PMF has been applied to the data from these two sites earlier in a study by Rivas et al., (2020). However, they used data from January 2007 to December 2016 whereas data in this study is from January 2015 to June 2019. Additionally, Rivas et al., (2020) used NO_2, NO, SO_2, CO, and O_3 data in addition to PNSD data in the PMF input files to separate the sources. In contrast, in this study only the PNSD data was included in the PMF input files and the results were later compared to the other measurement data. In addition to this, the novelty of this study arises from how the data was handled from the two nearby sites with strongly overlapping aerosol sources by adding the data from the two sites to the same data matrix horizontally as columns instead of doing two separate PMF analyses.

2. Lines 352 – 373 (SCA factor): m/z 60 has been used to diagnose the secondary combustion aerosol (SCA) factor. If m/z 60 was equivalent to Org60 in Table 4, the analysis for SCA was very confusing. According to the table, Org60 has a lower Pearson correlation coefficient (0.24) with SCA, compared with Org57 (0.36). When the authors mention the weak correlation between SCA and Org 57 in line 355, it is unclear why m/z 60 has a good correlation with SCA in Line 367.

AR: Authors agree that the SCA factor explanation was confusing and efforts to make it more understandable have been made. The paragraph explaining SCA is now as follows:

The SCA factor had a peak particle size of 44.7 nm at both sites and was interpreted as a secondary aerosol originating from combustion processes (i.e., of liquid fuel such as diesel, oil solid fuel such as biomass and coal, or gas). SCA had relatively weak correlations with the primary traffic emissions (e.g., NO_x, BC, CO, m/z 57) data, as could be expected for atmospherically processed aerosol. The strongest Pearson correlation coefficient of 0.56 was observed between SCA and NO_x at the UB site (Table 4). SCA and NO_x at the UB site also had similar diurnal patterns on working days and

weekends (Fig. 6). The highest SCA peak was seen approximately 3 hours later than for the TRA1 factor, indicating that the SCA factor included traffic emissions that had been aged/processed a couple of hours in the atmosphere. SCA was found to have an evening peak in addition to the morning rush hour peak (Fig. 4e). The evening peak was more pronounced during weekends, which indicates possible contributions from biomass combustion (Fig. 4f). In an earlier study, BC originating from biomass combustion was shown to contribute 15 ± 14% at the SC and between 41 ± 14 and 46 ± 15% of the BC in residential/detached house areas (Helin et al., 2018). To support this, the diurnal trends of SCA and organic fragments at m/z 60 (Q-ACSM) at the SC were plotted (Fig. 6). The fragments at m/z 60, particularly its fraction of the total OA, have widely been used as a marker for primary wood combustion emissions (Alfarra et al., 2007). The shape of the m/z 60 diurnal profile was similar to the SCA diurnal profile during the workdays and weekends strengthening the assumption of wood combustion contribution to SCA. An important thing to note is that the overall correlation with the m/z 60 was still relatively low (Table 4). The similar rush hour peak of m/z 60 to that of SCA was slightly surprising as the m/z 60 is usually related to biomass combustion and not traffic. The annual variation of SCA is small (Fig. 4g). Likely because although during the wintertime, the amount of biomass combustion increases the amount of sunlight is low, limiting SOA formation, whereas during summer, the amount of biomass burning is lower, but the amount of sunlight increases, thus enhancing SOA formation. In contrast, traffic emissions remain stable throughout the whole year.

In addition, in the whole text, ORG_{xx} has been changed to m/z xx to make the text more uniform. Additionally, in Tables 3 and 4 the ORG_{xx} form has also been changed to m/z xx. Fig. 6 was also updated as the Fig. 6b was found to have a mistake in the title.

Minor Comments

1. Line 65: What concentrations decrease? Please clarify it.

AR: The sentence is now changed to:

They stated that in urban environments, the majority of particles are in the size range of 10-100 nm, and the PNC decrease approximately by a factor of 100 when the particle size increases from 100 nm to 1000 nm.

2. Lines 124 – 128: Please provide the IE values for NO3 and the RIE for NH4 and SO4 in the paper.

AR: The following discussion about the values is now added to the text:

IE(NO_3) varied over the years and the final correction of the NRPM1 mass was done against the mass concentrations derived from DMPS data as described by Barreira et al., (2021). The RIE for SO_4 varied from 0.51-0.61 and for NH_4 3.8-5.32.

3. Section 2.4: Lots of details about processing data have been covered in the main text. To

improve the readability of the section, please only include the necessary parts in the main text and move the details into the supplement.

AR: The authors realize that a lot of details about processing the data have been given in the article however the data analysis is the novelty of this study and preprocessing the data is important to understand the process therefore authors think that significant amounts of the data can not be moved to supplemental material. However, the author removed some unnecessary repetition from the section.

4. Line 355: Please provide a short description of Org57. This is important for readers who are not familiar with AMS or ACSM measurements.

AR: The m/z 57 is now shortly described in the text:

Of these m/z 57 (C_4H_{9+}, C_3H_5O+) is a part of the HOA mass fraction that is linked to traffic exhaust emissions (Crilley et al., 2013; Crippa et al., 2012; Daellenbach et al., 2016; Mohr et al., 2012).

Also, the new references are added to the reference list:

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

Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Peñuelas, J., Jiménez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prévôt, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, Atmos. Chem. Phys., 12, 1649–1665, https://doi.org/10.5194/acp-12-1649-2012, 2012.

Crilley, L.R, Ayoko, G.A., Jayaratne, E.R., Salimi, F., Morawska, L.: Aerosol mass spectrometric analysis of the chemical composition of non-refractory PM1 samples from school environments in Brisbane, Australia, Sci. Total Environ., 458–460, 81-89. https://doi.org/10.1016/j.scitotenv.2013.04.007, 2013.

Daellenbach, K. R., Bozzetti, C., Křepelová, A., Canonaco, F., Wolf, R., Zotter, P., Fermo, P., Crippa, M., Slowik, J. G., Sosedova, Y., Zhang, Y., Huang, R.-J., Poulain, L., Szidat, S., Baltensperger, U., El Haddad, I., and Prévôt, A. S. H.: Characterization and source apportionment of organic aerosol using offline aerosol mass spectrometry, Atmos. Meas. Tech., 9, 23–39, https://doi.org/10.5194/amt-9-23-2016, 2016.

5. Line 395: What filters were used in the Niemi et al (2009)? Please clarify.

AR: Now the filter material Teflon is added to the text:

Additionally, Niemi et al. (2009) did not report high concentrations of NO_3 during the LRT episodes, likely because of evaporation losses of ammonium nitrate from the Teflon filters.

6. Lines 379 – Lines 386 (SecA factor): Please discuss why the SecA factor has moderate correlation coefficients with Org57 (0.52) and Org60 (0.5) in Table 4.

AR: The following paragraph discussing the possible reason for this is now added to the text:

In addition, the SecA had moderate correlations with m/z 57 and m/z 60 (Table 4). The reason for this is likely to be the high total amount of organics as the m/z 57 and m/z 60 do not refer to the relative fraction of the total organics but absolute concentrations of the mass fraction and therefore they might be elevated with the higher total organic mass in particles.

7. Lines 387 – Lines 398 (LRT factor): Please discuss why the LRT factor has considerable correlation coefficients with Org60 (0.64) in Table 4.

AR: The following discussion about the correlation is added to the text:

The LRT factor had a moderate correlation with m/z 60. This is likely to be caused by the large PM of LRT particles and therefore higher total organic mass. Additionally, the correlation with m/z 60 might indicate the contribution of remote biomass burning to the LRT factor.

Technical Comment

1. Lines 16 – 17: Please rephrase "… a novel approach to positive matrix factorization (PMF)"

AR: It is now rephrased as follows:

This study intended to develop the source apportionment of urban aerosols by utilising a novel approach to positive matrix factorization (PMF).

2. Line 56: The abbreviation for particle number size distribution should be PNSD. Please correct it here, but also other places where NSD are being in use.

AR: OK done.

3. Lines 136 – 137: Provide the manufactures for TEI42S and TEI49.

AR: The manufacturer of the instruments is TEI (Thermo Environmental Instruments) this is now written open in the text.

4. Line 182: Please clarify: "This was also necessary because using too large data files was not possible in EPA PMF 5.0."

AR: This is now clarified in the text:

This was also necessary because growing the size of data files over a certain point would cause EPA PMF 5.0 program to crash during the analysis because of the program running out of memory.

5. Line 365: Is m/z 60 the same as the Org60 in Table 4? If so, please make it consistent. Same case for m/z43 in Line 383

AR: This has been done as a part of the answer to major comment 2.

6. Line 430: Is Fig. 8 redundant here?

AR: The Fig. 8 is not redundant here but the sentence is missing other words and it is now corrected to the following form:

This is in contrast to for example TRA1, which shows little month-to-month variation, and the concentrations stay relatively stable between the different months in Fig. 7 and Fig. 8.

Reviewer 2# General Comments
It needs to be acknowledged that significant improvement has been made in the revised version of the manuscript. A more thorough introduction is now present, while the authors made an effort to address the comments made by the reviewers, regarding the presentation of the solution selection rationale, as well as regarding the interpretation of the obtained results. Nevertheless, I find it troubling that the authors are still not providing any proof whatsoever that their proposed new methodology is advantageous in figuring out the sources of particles affecting the two sites when compared to other known practices. I would like to stress this and ask the authors to take action to address it. If it is not proven that combining data from two sites in a PMF leading to a common timeseries has a positive effect in the source apportionment results, both from an algorithmic as well as from a physical meaning standpoint, the question about the validity and quality of the results will be overshadowing the whole effort. It is the authors themselves that in lines 429-430 in the "Conclusions and summary" section of the revised manuscript state that "the novel method of attaching simultaneous data from two sites seems to improve the detection of various factors and could be useful in locations where NSD data is available from more than one site", yet no proof of that improvement is provided in the article. Not very helpful in the review process itself is the fact that the authors when replying to my comments from the previous round of reviews, provide unclear and contradicting responses, while in certain occasions they address the comments rather poorly. For instance:

• In their response to General Comment #1, the authors state that "We tested different approaches for the data analysis (with and without additional data, analyzing stations separately, and different numbers of factors)" but a few sentences below they say that "Analyzing this large data set even with one tool was really time-consuming and challenging and thus it was not feasible to use several different approaches". So, what is the case here? Did they or did they not test different approaches?

AR: The authors did test different approaches to the PMF. These ways included doing PMF separately for the two stations using a wide range of factor numbers and also doing PMF with added BC data as columns to input matrixes to PMF. Also, different numbers of factors were tested for horizontally joined PMF. This is reflected in Fig. 2 which shows the mean residuals for factor numbers of 2 to 10. The answer to the first round of reviewer comments: "Analysing this large data set even with one tool was really time-consuming and challenging and thus it was not feasible to use several different approaches" refers to the fact that results from these other methods could not be analysed as thoroughly as the results from horizontally joined PMF 5 factor solution. Meaning that the diurnal profiles, residual analyses and correlations with auxiliary data were not studied for all of the tested approaches to PMF

• When replying to General Comment #2. In the first sentence they state "The authors were not able to do the analysis separately for the stations", and immediately after that they state that "The results were not satisfactory. We tested many different approaches and found this to be the best solution." What I understand from these sentences is that the analysis was done, but the authors deemed it not satisfactory. Proof of that needs to be provided in the revised version.

AR: The authors agree that the answer to General Comment #2 was not very clear. The author's statement "The authors were not able to the analysis separately for the stations. The results were not satisfactory" was intended to mean just that the authors tested doing the PMF analyses separately for the stations but the results were not good enough. The factors were meant to be found in this article using only PNSD data in the PMF input files and only comparing them to the auxiliary data afterwards. By using this approach when the PMF was done separately for the SC and UB, especially at the SC all the factors seemed to be correlating with traffic. This did not seem realistic. Now the main reason why the results were deemed not satisfactory is also stated in the text:

The decision to join the data together horizontally instead of doing the PMF analysis separately for the SC and UB was made because the idea in this study was to use only PNSD data as input data in PMF analysis and only to use the auxiliary data for identification of the factors. When the PMF analyses were done separately for the stations without additional data, the PMF was able to split the measured PNSD into factors at each site. However, seemingly any number of factors could be fitted, with PMF only fractioning the measured PNSD to more sub-modes. Additionally, when PMF analysis was performed separately the attained factors had different modes and mode concentration in between the stations in all cases. This is not likely to resemble reality as the stations reside less than 1 km from each other and therefore somewhat similar background and long-range transport factors would be expected. Adding the data from the two sites together horizontally forces the PMF to find a common time series between the stations. This is beneficial in finding the common factors for the UB and SC as the time series of the common factors can be expected to be similar because of the small distance between the stations. On the other hand, joining the data horizontally does not force the same factor profiles for both sites. An additional problem of doing the PMF separately for the stations was that all factors at the SC site seemed to correlate strongly with the traffic diurnal cycle indicating that the traffic emissions are split between the different factors. The figures showing the 4, 5, and 6-factor solutions of the PMF analyses done separately for SC and UB are presented in

supplemental material S3, S4, and S5 respectively. The negative side of merging the data horizontally can be expected to lower the total explained a fraction of the PNSD while the common time series are forced for both of the stations.

The following figures are added to the supplement:

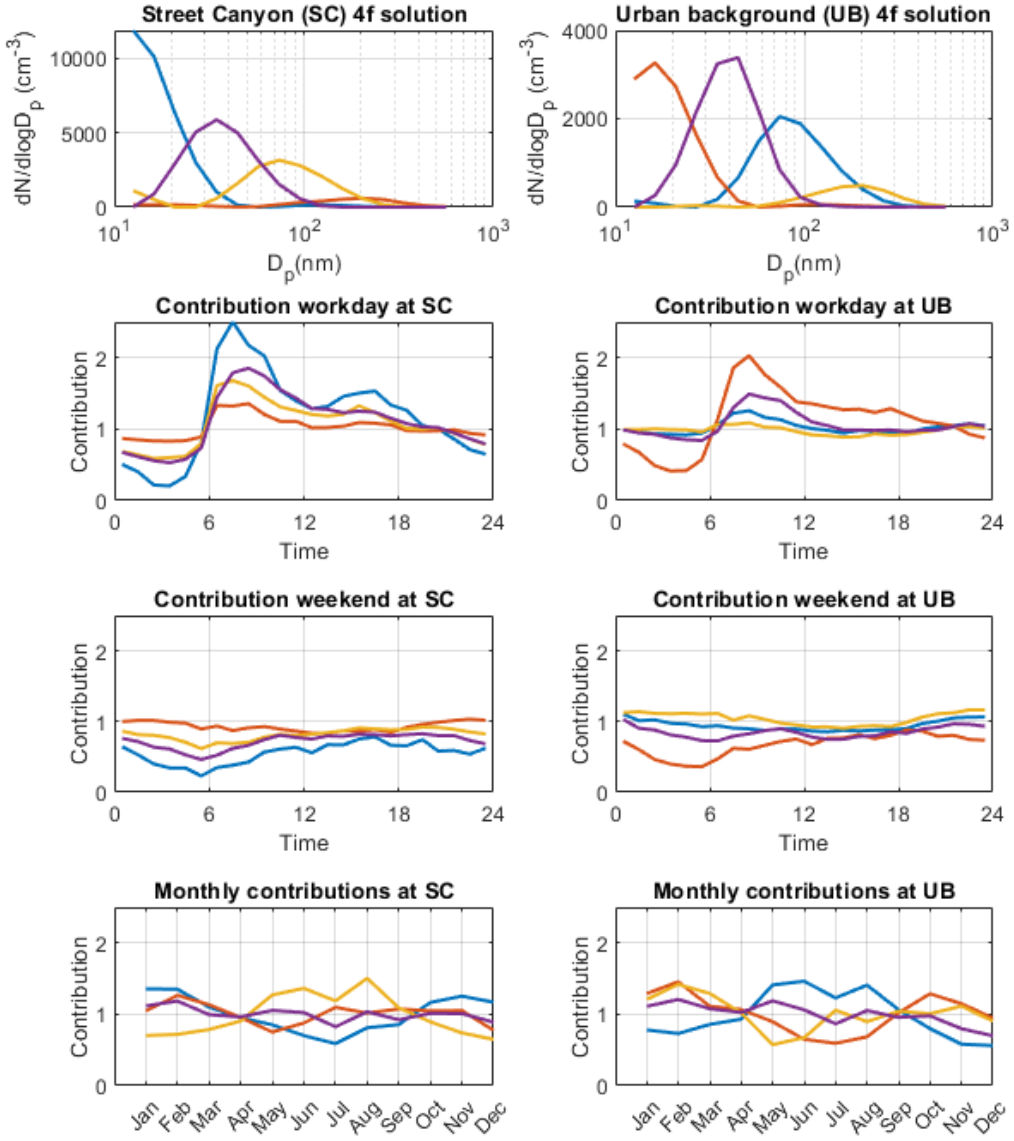

Figure S3: Four-factor solutions, contributions during workdays, weekends, and different months from PMF analyses done separately for SC (left column) and UB (right column).

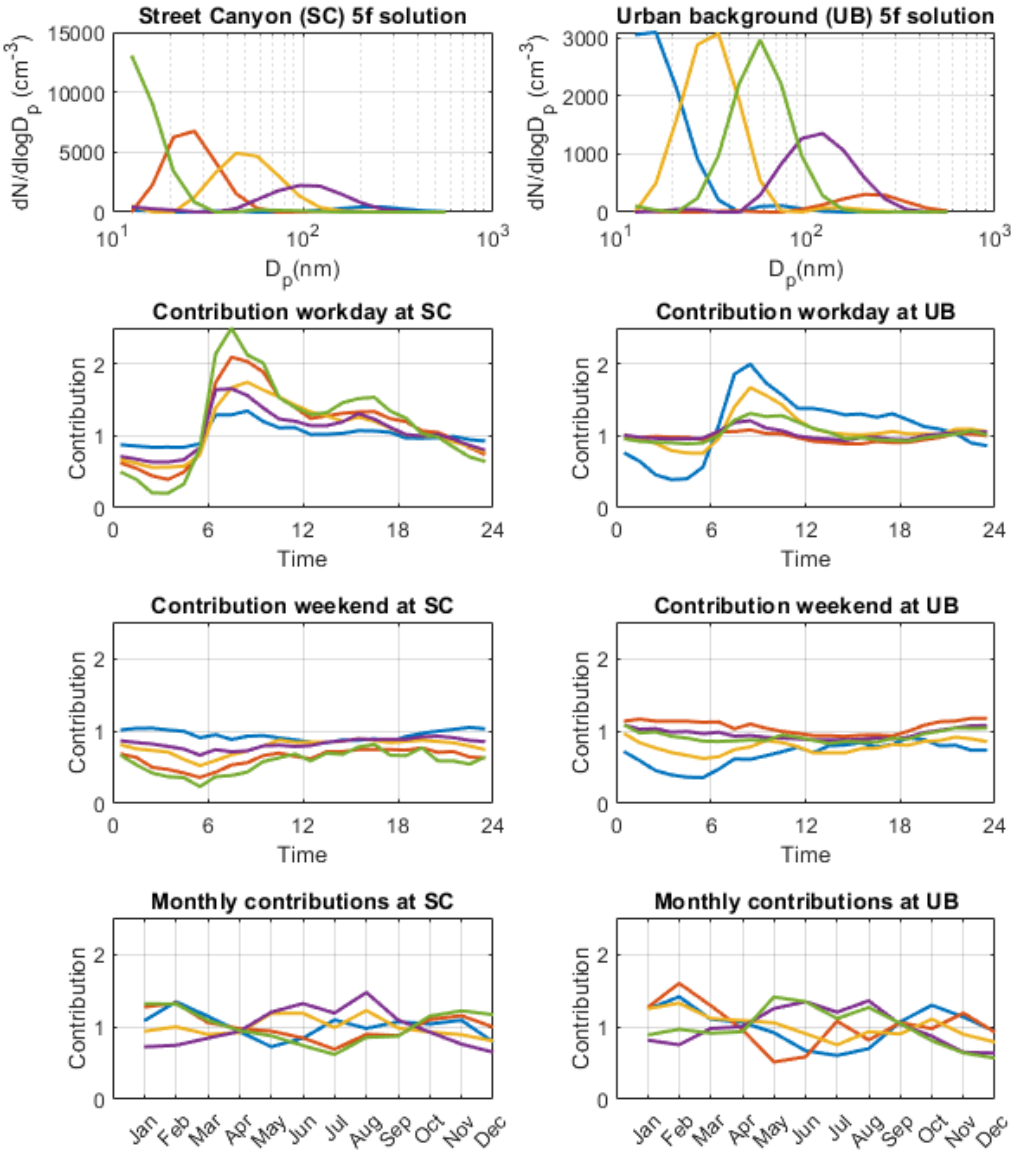

Figure 4S: Five-factor solutions, contributions during workdays, weekends, and different months from PMF analyses done separately for SC (left column) and UB (right column).

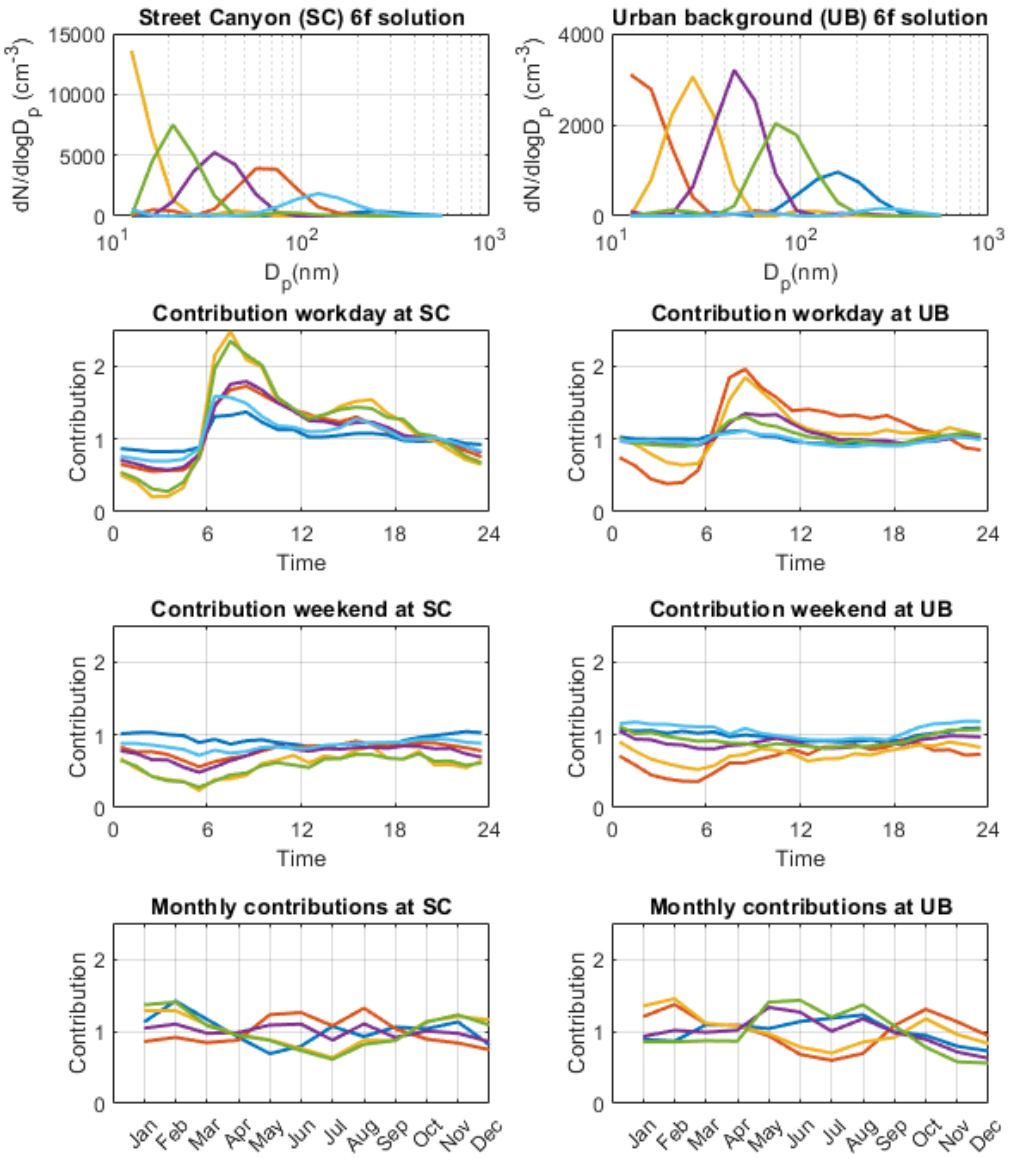

Figure S5: Six-factor solutions, contributions during workdays, weekends, and different months from PMF analyses done separately for SC (left column) and UB (right column).

• Regarding the authors response to General Comment #3. I fully understand the challenges surrounding the process of introducing, and for the first time applying a new methodology. I guess it requires thorough knowledge and understanding of existing approaches in the authors part, but more importantly, the proposed method needs to be able to perform equally or outperform existing approaches, if not in the general case, at least when applied in a specific dataset. After the new method's performance is presented, then the reader can really focus on the method's results and their significance in understanding the atmospheric processes under investigation. In the form that the manuscript is now, the reader is left uninformed on whether this proposed way of treating PNSD data from different types of

stations is worth the effort. I genuinely think that horizontally combining the datasets is a brilliant idea. If it really works similarly or even better to single station PMF then it could be recommended to everybody interested in contrasts between different types of stations. I would like to urge the authors to perform single station PMF tests and present their comparison to the combined PMF here. They could do it in a smaller but still representative portion of the dataset of their choosing, if they feel that a full comparison is time consuming. If the authors are still reluctant on performing such tests, at least they could perform a comparison with the overlapping years presented in Rivas et al. (2020). In any case, I believe that any advancements and/or limitations related to the new methodology need to be discussed in the text.

AR: The PMF analyses were redone for both stations separately for factor numbers of 4-6. The results from these comparisons are now presented in supplemental material as stated in the author's response to the previous comment of the reviewer. The discussion about the benefits and drawbacks of the method are now discussed in the text. This modified text has been attached to this document as a response to the previous comment of the reviewer.

• With General Comment #4, along with several comments from Reviewer 1, the authors were given the chance to provide sufficient information on the PMF quality. They partly addressed the matter, but still some features can be added. The following should be added in the supplementary material:

AR: Answering some of these questions afterwards is a slightly challenging as EPA PMF only gives 5 text files as an output after the analysis is done and some information is lost when the program is turned off. These files include residual matrixes, contributions of factors at each given time, diagnostics, source profiles and base run comparisons.

o a comparison (regression) between the total (Ntot) particle number concentration (PNC) measured at each site and the respective PMF reconstructed PNC.

AR: This is now done, and the figures have been added to the supplements:

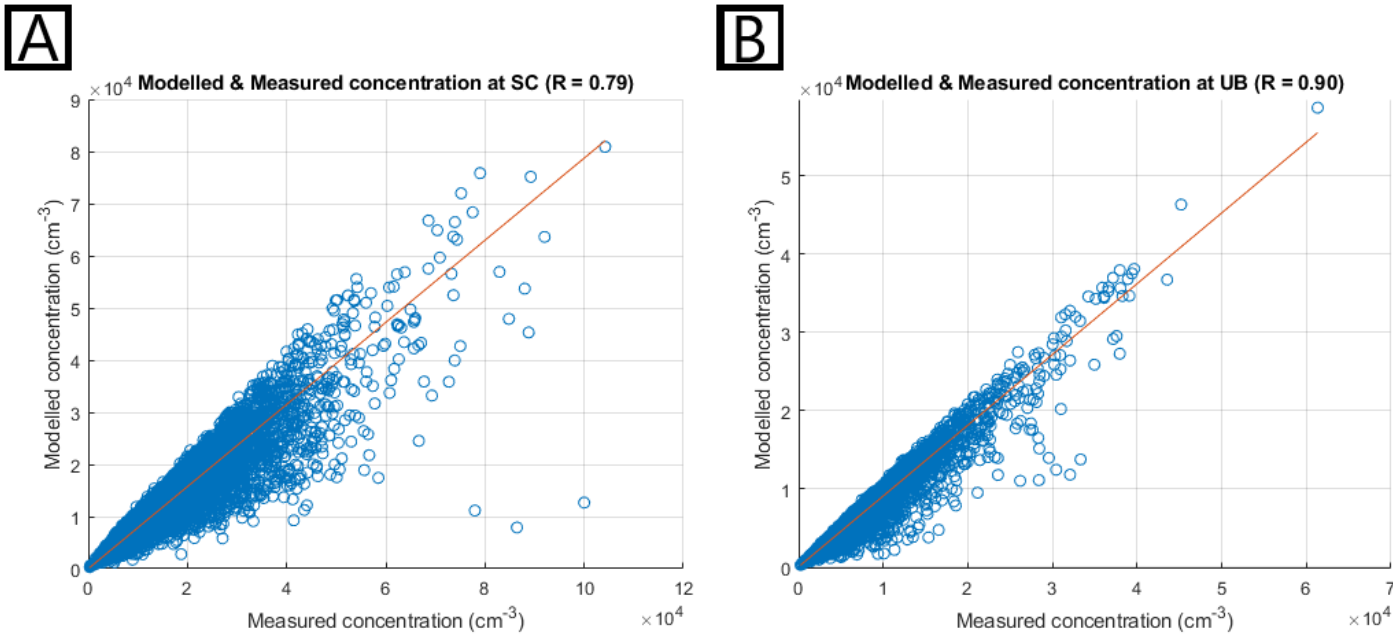

Figure 9S: Regression plots between modelled and measured interpolated concentrations at SC (A) and UB (B).

The following text was also added to the manuscript:

Also, the figures showing the regressions between the modelled and measured concentrations after interpolation are presented in supplemental material S9.

o what was the unexplained fraction of Ntot? Any temporal trends there?

AR: The average unexplained fraction for SC was 12.7 % & and 6.7 % for UB. There were slight temporal trends with the unexplained factor increasing over the daytime hours at SC. In the monthly plot, no clear trend was observed. The unexplained percentages are now presented in the text:

The average difference between these was 12.7 % at SC and 6.7 % at UB and the temporal patterns for the difference have been presented in supplemental material S10.

Additionally, a figure showing the diurnal variation of the unexplained parts during workdays and weekends with monthly unexplained fractions is presented in the supplement:

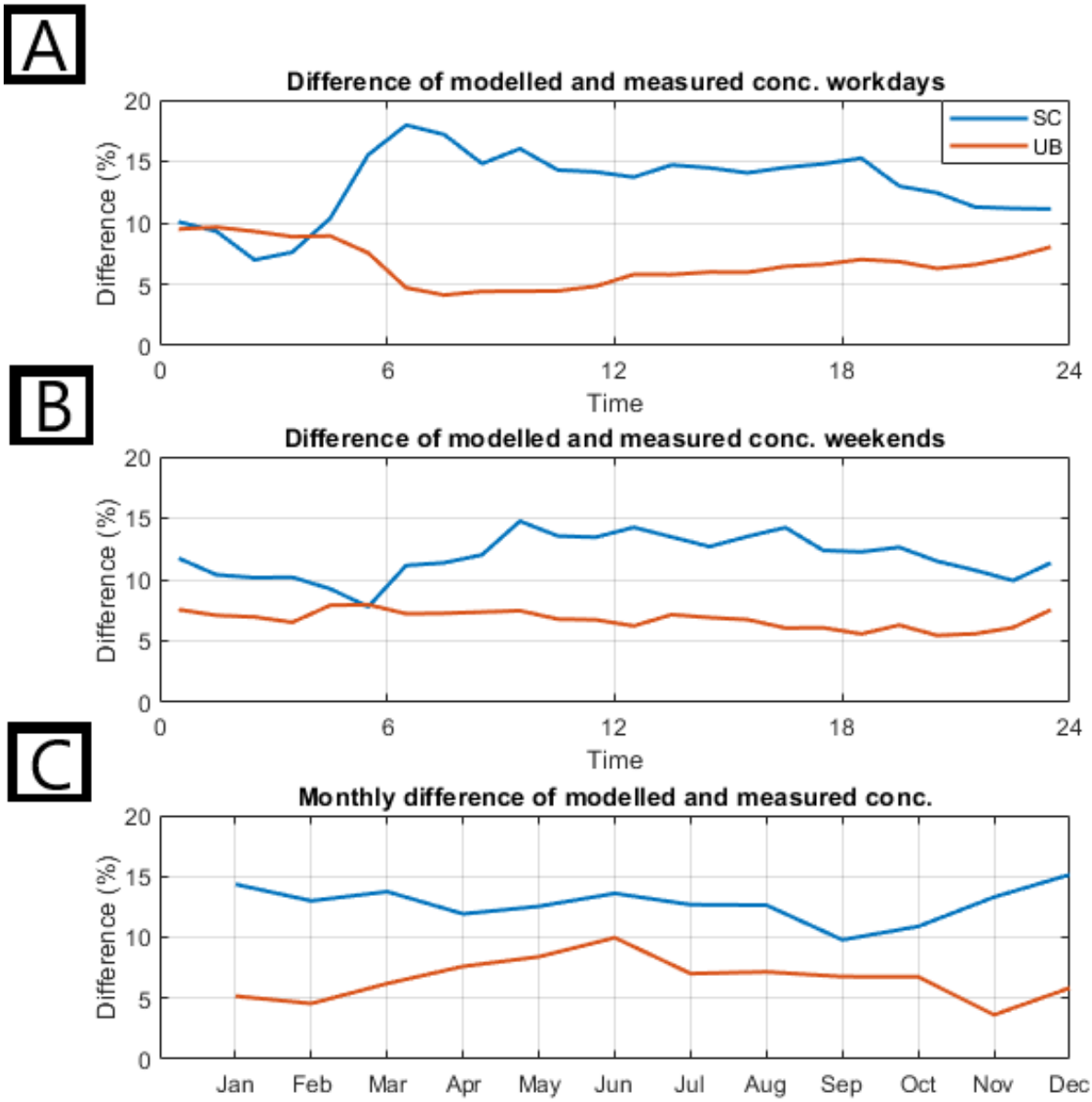

Figure S10: Average differences between modelled and measured concentrations for workdays (A), weekends (B) and months (C)at SC and UB.

o a short description of the DISP process results (sawps, dQ percentage etc)

AR: The DISP process was run by PMF and there were no swaps or dQ drops. This is also now mentioned in the text:

The results of the displacement analysis showed no drop in Q values or swaps in any of the analyses.

o how many random seed runs were performed for the presented solution?

AR: The number of seed runs used was 5. Using a larger number caused EPA PMF 5.0 to crash. However, The analysis was run several times and the results were similar at all of them. The number of used seeds is now also mentioned in the text:

The robustness of the solution was tested using five different random seeds as starting points. Performing analysis with a larger seed number sometimes caused the program to crash. In addition to using random sees a displacement analysis was performed on the solutions.

o was there any variability between each run's factor contribution to Ntot?

AR: The data from the other seed runs is not saved when using EPA PMF 5.0 so I can not calculate this. However, the output files of PMF files show that all the runs had the same exact Qrobust values and only minimal differences in the Qtrue value indicating that the differences between the results of different runs were minimal.

o were there any tools used to address rotational ambiguity (e.g. fpeak)? I would assume that there were, so please provide the appropriate details (G-space plots etc).

AR: The rotational ambiguity was tested only based on displacement analysis. The results of displacement analysis demonstrated no rotational ambiguity so this was not further tested.

• Regarding the authors response to General Comment #6. Given that the authors have performed wind regression to the PMF results, these results should be presented at least in the supplementary material. What do the authors mean when saying that this type of analysis did not provide any additional information? Didn't the polar plots of each factor point to different contributions for different wind directions and speeds? About the SC site, I'm not sure I fully agree. For instance, Rivas et al. (2020) performed the analysis and presented it, providing some insightful information.

AR: The authors have studied the effect of wind directions on the source factor contributions. This has been done by drawing wind roses for contributions of factors from different wind directions and it is now added to the supplemental material and text referring to the figure added to the text:

These contributions were also calculated for each factor separately in supplemental material S22 for different wind directions.

The following figure was also added to the supplemental material:

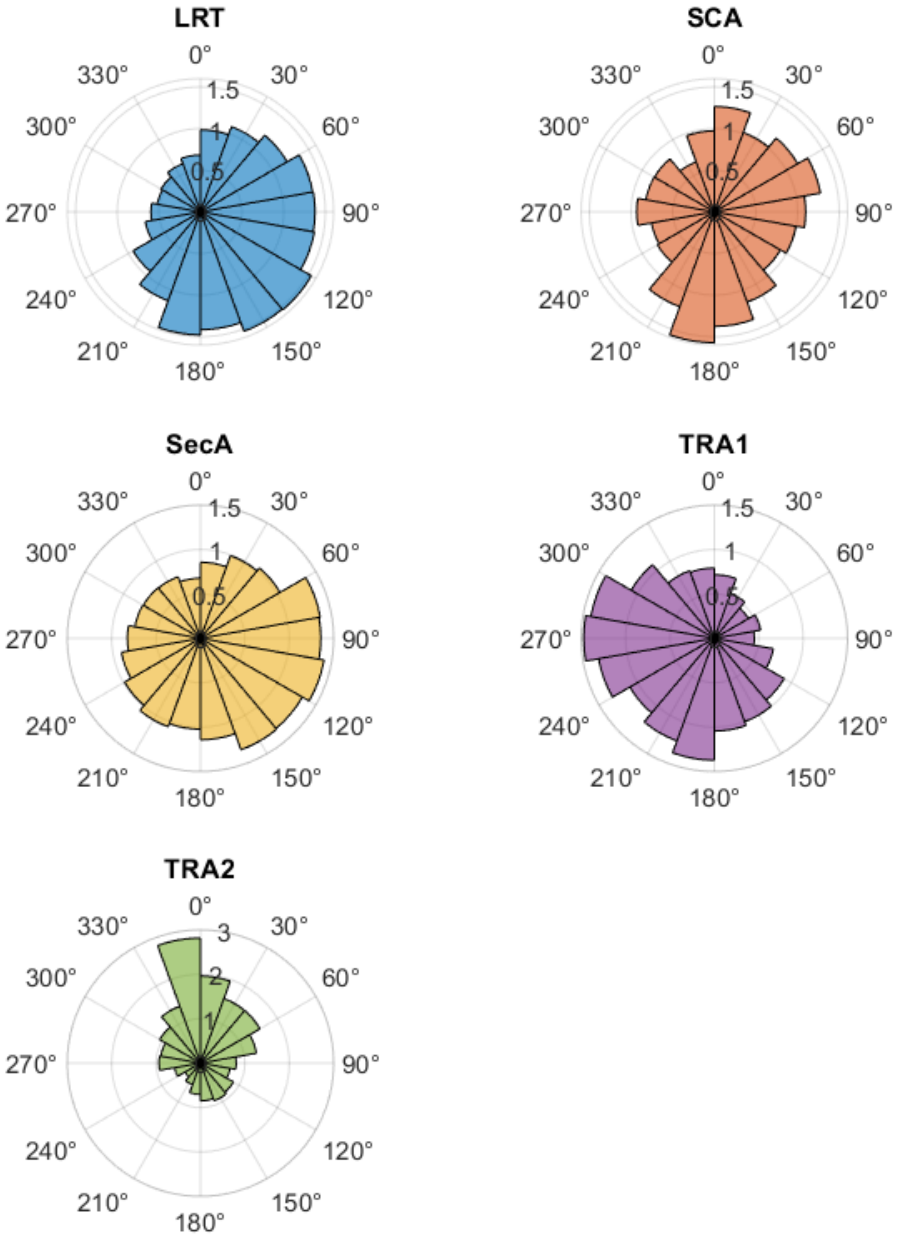

Figure S22: Scaled contributions of LRT, SCA, SecA, TRA1 and TRA2 factors from different wind directions.