## Response to reviewer comment (Rev. 2):

The revised paper contains changes according to comments of all of the three reviewers. The abstract was revised according to the remarks of the three reviewers. As recommended, single paragraphs were shifted to new positions for better understanding. The entire introduction was restructured. In section 2 the discussion on weather conditions in Beijing was extended, specifications for stopping conditions of the NMF algorithm were included, the role of NMF factors and weights was clarified and the subsection for relevance of the NMF factors was improved. In Section 3 interpretation of the resulting pattern was improved and extended regarding weather conditions and NMF weights. The subsections in chapter 3 were combined to form at least two subsections in the revised paper.

**Figures and Tables:** We changed the content of the figures as follows. Figure 1 was replaced by the temporal mean, which is considered useful for comparison with the NMF factors and their patterns. Since dust storms do not play a role anymore, when shifting the NMF volume application to a separate study, the former graphs of 2 exemplary days with and without dust storm influence are removed. Figure 2 is new, comprising information about wind direction and speed in the study period. The former figure 2 becomes now figure 3. The former figure 3 becomes figure 4 (recomputed because of the use of mean instead of median values). The former figure 4 becomes figure 5, where time series of the NMF factors have been added for all 5 factors. The new figure 6 is the upper part of the former version. Since volume NMF results are omitted, the lower part does not play a role. Figure captions have been adapted and the content of Table 1 was changed according to the changes in the manuscript.

The authors thank the reviewer for its useful and reasonable comments to improve the manuscript.

In the following we give answers to the comments by reviewer 2:

## General Comments:

*Rev. 2:*
*1) Given the large number of observations (864 daily profiles), I would have expected that the factors extracted out of the NMF methodology should have had a better physical meaning than the ones described in section 3.1 and 3.2. For instance NMF-N1 is a mixture of three sources: this is a poor result for a data-rich-driven factor analysis where usually separate traffic, industrial and residential sources are resolved when PSD are used (Gu et al., 2011, using PMF methodology were able to resolve local and long range dust, fresh and aged traffic sources in a medium size city in Germany).*

**Recent studies on PSD data might have done better in finding out more different sources. This study is based on a completely new approach. It is not finally developed. However, we aimed at a new method that provides new information and a new view on the subject. We regard the method appropriate and successful to obtain physically interpretable patterns that may be linked to particle sources, solely based on particle number size data (without any chemical composition and additional pollutant measurements).**
**The value of r restricts the number of separate patterns, but as explained in the corresponding section r=5 was considered best in this special case. More patterns could be obtained, but with increasing difficulty in interpretation. Nevertheless, we revised the interpretation of the NMF patterns for better understanding, making use of meteorological parameters and seasonal means of the NMF weights.**

*I see no 'soil dust' source identified despite Beijing being subject to PM loads from dust storms. The authors should strive to identify those sources in the context of analysis of the NMF-V factors which is not carried out in depth, despite evidence like Figure 5. For instance, since NMF-V factors are reflecting the sources' mass contributions they could be compared with results from receptor modeling studies carried out at Beijing.*

**We would like to address the issue of dust storms in another study concerning NMF on particle volume data following the advices of reviewer 1. Regarding particle volume, dust storm patterns are observed, since corresponding particles are larger. The presented NMF patterns for particle number comprise mainly variations of fine particulate matter.**

*I am also concerned with the apparent discontinuity of PSD temporal profiles at midnight which has not been observed in similar studies (Gu et al, 2011, cited in the MS). Is it possible that such a gap is due to some artifact in NMF implementation? Is it possible to add a constraint to NMF to enforce an agreement at midnight?*

**We thank the reviewer for this remark. We added a short paragraph in the section 2.3.2 addressing this issue in the revised paper, because we understand that at a first glance this feature of the NMF results is confusing. But when superimposing the different NMF factors additively, that feature of the PSDs is not important anymore, since the factors are overlayed. Differences before and after midnight, appearing in a single factor, are corrected by the pattern of another one. Adding a constraint in the NMF algorithm forces the factors to a certain shape. But this restriction is not intended and therefore, we omitted it.**

*Rev. 2:*
*2) The analysis of time-varying source contributions is superficial. I would recommend looking at workday-weekend and seasonal variation to confirm source identification and discriminate among traffic, industrial, long range and residential sources (Yue et al, 2008, cited in the MS). Since in the Abstract section's first sentence: "increasing traffic density and a changing car fleet" authors are suggesting potential trends in traffic impacts at Beijing, they should either remove that sentence from the Abstract or conduct a trend analysis for the traffic sources, after they are better resolved within the NMF methodology.*

**We analyzed seasonal variations in terms of mean contributions per season to support the interpretation of the patterns. A more detailed analysis on time-varying contributions of the patterns will be added in future studies.**
**Discussion on traffic was extended in the introduction section (including references) as well as concerning the interpretation of NMF-2.**

*Rev. 2:*
*3) The discussion on wind trajectories of potential sources reaching the monitor site should be expanded with more quantitative meteorological data to confirm source identification and potential location. I understand that Beijing and adjacent cities is part of a cluster with _ 180 million inhabitants so a proper source attribution is a challenging issue and if the manuscript approach can provide new insights this would add in more value for regulatory purposes, which is a key goal stated in the Abstract section.*

**Thank you for this remark. A further figure for mean wind direction and speed was included to follow the advice of the reviewer and discussion was improved in the weather data section. However, a detailed analysis of wind trajectories was omitted, because this would be beyond the scope of this article.**

*Rev. 2:*
*4) The manuscript is poorly edited. The introduction should be focused on how different processes and emission sources lead to PSD with different size modes (in number and volume) because those are the kind of results that will be discussed. The health effects literature review should be reduced to a minimum. There are many vague, incomplete sentences scattered across the manuscript; I'd recommend hiring an English editor to improve on this.*

**Thank you for this remark. We followed the advice to improve the introduction extending the discussion on particle processes and emission sources and reducing the discussion regarding health effects of PM. The language has been improved.**

## Specific Comments:

*Abstract*
*P13017, first sentence: if trends in traffic emissions are hypothesized, they should be discussed in the respective section. Otherwise rephrase the sentence. (See also comments of the Conclusions section below).*

**The sentence was changed in the abstract and the traffic issue addressed in the introduction section.**

*P13017, second sentence (lines 4-7): if one goal is to identify sources and their locations for a better regulatory policy, then this should be answered in the manuscript conclusions.*

**We changed the sentence in the abstract, since it was indeed a bit misleading regarding the inetention of this article. Later studies could start at this point considering results for a better regulatory policy.**

*P13017 third paragraph (lines 17-28): the sources (factors) obtained with NMF should be clearly identified and their mean contribution to number (or volume) quoted. The NMF-V factors should be compared with chemical composition receptor model results for Beijing.*

**The corresponding paragraph in the abstract was adjusted. Mean contributions of the factors are now presented in brackets. Volume NMF is omitted and shiftet to another study.**

*Introduction*
*P13018-13019: reduce discussion on PM-associated health effects and focus on the sources and processes that lead to distinctive PSD in ambient particulate matter. Summarize previous findings for Beijing focusing on source apportionment and PSD measurements, to help following discussion of results.*

**We followed the advices (see above) and included references on source apportionment for Beijing.**

*P13020: be more succinct on discussing factor analysis approaches, leaving details for section 2.3.*

**Thank you for the remak. We adjusted the paragraph, but decided to maintain this brief introduction to PMF and NMF in the introduction.**

*P13022-23: A map of Beijing showing the typical wind patterns would be helpful for source identification (by season for instance) of city and regional scale sources.*

**We included an additional figure for wind speed and direction.**

*P13024: Explain why you chose a maximum number of iterations set to 1000. What were the typical numbers of iterations required for convergence in NMF? More explanation is needed to understand the stopping criteria used.*

**Thanks for this advice. We improved this imprecise statement including discussion on our stopping criteria.**

*P13026 lines 18-28. The choice of number of factors (r) was made on inspection of resulting PSD profiles to avoid combination of sources (factors). However the results for NMF-N1 are identified as a mixture of three sources. Reconcile this.*

**We considered the choice of r carefully. The value of r restricts the number of separate patterns. The number of NMF factors was chosen as 5, because in this case the best physical interpretability of the patterns was obtained. For more factors, the patterns become too local and an assignment to sources is more difficult. Unfortunately, some processes are mixed together in the r=5 case. We extended our interpretation of the patterns and their related sources.**

*P13027 lines 1-11. It is unclear which was the final weighing of data for PSD (if any), especially when considering volume data. This should be stated for the results quoted in the manuscript.*

**Volume and weighted NMF are shiftet to another paper. Therefore they are omitted now in the revised paper.**

*P13027 lines 13-16. It seems curious that NMF-V factors could not be physically interpreted. What about correlating them with available pollutant measurements? What about the time dependency of pattern amplitude? Workday - weekend differences?*

**You are right. Thank you for this advice. Our aim was to keep the section about volume NMF short and just emphasize the dust storm relationship. Due to your and the other reviewers' comments we decided to omit the whole section. In another paper we are going to provide a more detailed discussion of this issue.**

*2.3.3 Relevance of NMF factors.*
*P13027 lines 26-27. Why was variance assumed to be constant for all factors? Was this assumption verified ex-post? Please justify.*

**Maybe the phrase was confusing. Therefore we changed the paragraph including reasons for our methodology. We intended to explain, that we set our focus on the mean of the factors, not on the variance. Because of the non-negativity constraint a usual standardization could not be applied. Hence, we used the diagonl scaling step.**

*P13028 I don't understand the scaling using equations (4) and (7) after NMF has converged. It seems more convenient to normalize W to 1.0 (like any probability density function) and leave the amplitudes H pick up the time-varying source contribution in units of particle number (or volume) concentration. In this way both factors can be analyzed on their own. The use of equations (4) and (7) is non-standard in factor analysis so comparisons with other studies (like traditional receptor modeling versus the NMF-V sources) is not direct.*

**We changed the section in the manuscript in order to clarify our intention in applying this special normalization strategy. As stated there, a standardization with mean 0 and variance 1 (as it is done for probability density functions) is not possible, since non-negativity is required. There are some**

**alternatives left (the optimality of the solution W,H must not change, i.e. W^, H^ are diagonal transformations of W,H according to equations (4)-(6)):**

**1) to constrain the columns of W to sum up to 1.**
**2) to transform the columns of W to unit vectors (sum of squared components is equal to 1).**
**3) to transform the data in the interval [0,1] by division by the maximum of each column.**
**4) method described in the manuscript.**
**5) other approaches.**

**The methods 1)-3) are not desirable for the vectors that were used in the study, since the resulting values would be extremely small (since the values of W vary between several powers of 10). Instead we intended the patterns to be interpreted in terms of temporal mean patterns contributing to the true particle burden by weights in H, not as normalized and hardly interpretable values in [0,1].**
**Hence, method 4) was regarded reasonable.**

*3. Results and Discussion*
*P13029 lines 9 –12. If the physical interpretation of the W factors depends on the time-varying H factors, why these were not extensively used to identify the sources/processes that correspond to each W found? What is the meaning of the "particle burden, measured in terms of the median value of the corresponding time series"? What are the units of H factors?*

**We based our interpretation on both factors in W and weights in H. In the revised manuscript, we added some sentences on the time-varying coefficients in H or shifted some remarks from the end of each paragraph into the middle of the paragraph to clear up the interpretation. The particle burden is built by the non-negative combination of the NMF factors by their corresponding weights each day. The weights in H do not have any unit, the components of the factors in W do ($1/cm^3$, but as dN/dlogDp values).**

*P13029 lines 15 – 19. If NMF-N1 is a "background pattern" that provides "a kind of basic load" then why is the time-varying contribution H zero at so many days? (Figure 4b) Could Beijing ever "shut down" their basic PM sources/processes for a whole day? I doubt it.*

**Of course, you are right, when stating that NMF is not always present and there are a lot of days showing a 0 NMF-1 weight.**
**But the factor can still be interpreted as the urban background pollution: NMF-1 is present for days with low wind speed or no wind. So particles from the urban background are prevalent. If the wind is more intense, the background pattern is not that important, since pre-existing particles are blown away and replaced by or interacting (coagulation, ...) with brought-in particles. Therefore the weights of the other factors are high for these days.**

*P13030 Provide further support for the identification of factors NMF-N2 and NMF-N3. NMF-N2 should be analyzed using the time-varying contribution H (comparing seasonal behavior for instance) and NMF-N3 with the help of wind trajectories to identify likely source locations. All time-varying contributions should be presented.*

**Discussion on those factors is intensified and further informations are given in the revised paper. The time varying coefficients are presented in the extended figure 5.**

*P13032 lines 17 –19 authors state that they could not identify NMF-V factors for they only change in the coarse mode. Have they tried looking at the temporal amplitudes H's for seasonal, weekly patterns? How about looking for correlations of the temporal amplitudes H's with the ones computed for the NMF-N factors? What about looking at correlations of NMF-V temporal amplitudes with*

*gaseous or particulate pollutants at nearby monitoring stations? Last, but not the least, perhaps the results do not support so many NMF-V factors after all. I do think more analysis is required here to achieve a credible explanation.*

**Thank you for detailed remarks concerning the NMF-V factors and their interpretation. Following the advice of Reviewer 1, we decided to omit volume patterns in this manuscript and analyze those patterns in separate study.**
**A detailed explanation and analysis of the patterns as well as weights will be provided there. We will consider your remarks.**

*3.4 Categorization of NMF patterns*
*P13035 lines 1-5. The identification of NMF-N4 and NMF-N5 as secondary aerosols requires further support. Are inorganic (sulfates, nitrates) and organic (SOA) aerosols included in those two factors? If so, are they mixed in both factors or separately re-solved? Is it possible to correlate those factor's time amplitudes H with ambient measurements of nitrates, sulfates, OC?*

**We revised the paragraph and included some additional information (including the introduction) that supports our classification. Unfortunately, we do not have access to nitrates/sulfates/EC/OC measurements etc.**

*4. Conclusions*
*P13036 lines 18-19. This sentence should be repeated in the abstract as the main goal of this work, being part of the first paragraph in the Abstract.*

**We included a sentence for this main goal in the abstact.**

*P13036 lines 25-26 and beginning of P13037. I do not think that the authors have made the case that they can fully combine NMF-N and NMF-V factors to identify all PM sources/processes at Beijing, especially for coarse mode particles. Certainly dust storms alone cannot explain all coarse mode data.*

**You are right. In the revised manuscript we focused on particle number and are going to address volume NMF separately. After careful analysis of particle volume patterns there might be the possibility to obtain a combined view on the particle sources via NMF-N and NMF-V.**

*P13037 lines 17-18 Time varying contributions H were used only in their mean values per season and long term means. Correlations among time-varying contributions and local pollutant concentrations and among NMF-N/V factors were not studied at all, yet they could provide support for further source identification.*

**Correlation of NMF weights and weather parameters have been calculated (some of them presented in the interpretation of the NMF factors) and time series of NMF weights were presented. Other pollutant data are not available to us.**