The authors would like to thank the referees for their valuable comments. We have responded to all comments and have revised the manuscript accordingly. Details of our responses to each comment are shown below.

Anonymous Referee #1

In their manuscript, von der Weiden-Reinmüller et al. compare mobile and stationary aerosol and gas-phase measurements of background air and the emission plume of the European megacity Paris. In particular, they use the mobile laboratory MoLa to measure the submicron aerosol chemical composition, aerosol number concentrations and size distributions, and atmospheric trace gases at various fixed stations, to measure cross sections of the emission plume, and to carry out quasi-Lagrangian experiments following the emission plume from the Paris center to the surroundings. The authors use fresh pollution markers such as black carbon and hydrocarbon-like organic aerosol (HOA) to identify and characterize the emission plume, then compare pollutant concentrations in the emission plume and background air, study the spatial distribution of pollutants in the emission plume, and finally investigate aging of the organic aerosol in the emission plume.

The manuscript is well structured and clearly written. The presented results are an important contribution to the MEGAPOLI project, and will help designing future field experiments investigating urban emission plumes and their interpretation. The authors show that the mobile laboratory MoLa is a valuable tool in studying urban emission plumes and their impact on air quality in surrounding areas. The supplementary material nicely complements the manuscript. I recommend publication of this manuscript in ACP after taking into account the following comments:

a) In section 2.2, please add some more information about the instrumentation, e.g. a table with the types of instruments used in this study.

Reply by the authors: A short table (now Table 1) summarizing the instruments and measured variables used for the measurements presented in this study was added to the manuscript. A table with more information on each of the instruments operated within MoLa was already presented in von der Weiden-Reinmüller et al., 2014 (AMT 7,p279-299). Instead of repeating all this information we reference this paper here.

b) p. 11259, l. 9/10: Please give an estimate or quantify the mass fraction typically contributed by m/z > 100.

Reply by the authors: Due to the strong fragmentation in the hard evaporation/ionization process in the aerosol mass spectrometer the majority of signal is found at small m/z. Typically in the order of < 1% of the total signal and about 5 % of the organics signal is found for m/z > 100. This information was added to the text (Page 11259, Line 10). Only when extremely large organic aerosol fractions of polyaromatic hydrocarbons are measured (which do not show such strong fragmentation in the AMS), the fraction of signal at m/z > 100 might be well above 10%. However, such a situation was not encountered during this work.

c) p. 11261: How would PMF results change if local pollution events were removed from the data before positive matrix factorization as compared to removing data when OC mass concentrations are larger than 50 ug m-3. Also, please discuss the uncertainties of the changes in the composition of the organic aerosol (p. 11272) taking into account a 30 % uncertainty in ambient AMS data for absolute

mass concentrations of individual factors derived by PMF. How will this affect the significance of the plume contribution estimates shown in Fig. 3, the changes in the average organic mass spectra presented in Fig. S4 (supplementary material), and the correlation coefficients shown in Tab. 2?

Reply by the authors: Before performing the PMF calculations individual data points with extraordinarily large organics concentrations (outliers) should be removed from the dataset because otherwise the PMF results will be dominated by these few data points and biased results will be generated. There is no general rule where the threshold should be set. In order to remove the most "severe" peaks but not removing large fractions of the dataset we arbitrarily chose to set the threshold to 50 µg/m3 as described in the text. This threshold has nothing to do with local contamination events at the first place. As also described in the text all data identified to be contaminated by local sources have been removed from the data set after the PMF calculations have been finished. In order to make this clearer we slightly modified the text in the respective locations (section 2.4.3).
In order to determine uncertainties of the PMF calculations several variations have been made to the input data set. For example PMF has been performed with and without removal of such outliers. PMF has been performed on the full data set collected during MEGAPOLI and on the mobile and stationary data sets separately. PMF has been performed with a large variety of "seed" and "fpeak" values. All these calculations have shown that "seed" variations resulted in only very small (~5%) variations of the results while variations due to "fpeak" variations are in the order of ~20%. For the more robust factors (HOA and aged OOA) the variations due to removal or not of the outliers are in the order of 5-15% while for the other PMF factors these variations can be up to 50%.
All these efforts have shown that especially for less "robust" factors there can be significant uncertainty in the mass concentration of these factors due to the decisions made by the person who performs the analysis. However, this is not the point of the reviewer's comment. The major question is whether or to which extent these uncertainties affect conclusions drawn from such data.
Generally, the result of a PMF analysis is a rather subjective decision. Often several very different solutions are possible. It is the operator's duty to base the decision of the selected solution on as many as possible external facts like comparison with time series of other variables, comparison of resulting mass spectra with spectra from data bases, etc. The mass concentrations of the various PMF factors resulting from such a process have uncertainties in the order of the values presented above.
There is an additional uncertainty of AMS mass concentrations in the order of 20-30% mainly due to imperfect calibrations and assumptions needed to convert raw data into mass concentrations of individual species. In this work we mainly base our conclusions on the comparison of species or PMF factor mass concentrations measured at different locations. All these concentrations were calculated with the same calibration and correction factors and with the same PMF settings, i.e., only additional uncertainties from the PMF analysis itself remain. Due to this fact the uncertainty of the relative concentrations or the variation of the concentrations is much lower than the uncertainty of the absolute concentrations or the uncertainty of concentrations derived using different PMF settings. We estimate these remaining uncertainties to be in the order of 5-10% and therefore we do not expect them to affect the significance of our conclusions.
To avoid a long discussion of all the effects and consequences we added a short sentence stating that the relative uncertainties of the AMS mass concentrations are significantly lower than those of the absolute values and thus do not affect the significance of the results (section 2.4.3).

d) p. 11273: The mode in the aerosol number size distributions at a diameter of 10 nm (Fig. 4) may be an artifact as stated e.g. by Freutel et al. (2013; Atmos. Chem. Phys., 13, 933–959): »>The comparison shows a mode in the number distribution measured by the FMPS around 10 to 15 nm which is likely an artefact due to the inversion algorithm used for this instrument (A. Wiedensohler, personal communication, 2012).«< Please include this information and briefly discuss it.

Reply by the authors: We included this information in the manuscript and added after "The large error bars especially of the background distribution indicate that there was a large temporal and spatial variation in the occurrence of small aerosol particles." (p11274 in the first version of the manuscript): "In addition the particle mode around 10 nm might be affected by artefacts due to the inversion algorithm used for this instrument (A. Wiedensohler, personal communication, 2012)."

e) p. 11275/76: I was distracted by the CO2 data in Fig. 5 (especially in Fig. 5b) when the discussion in the text focused primarily on the comparison of measured and modeled HOA concentrations. Therefore, I suggest removing the CO2 example from Fig. 5. You may add an additional figure showing the good agreement between the plume shape and spatial extent of other measured fresh pollution markers.

Reply by the authors: Here we do not agree with the statement that the discussion focuses primarily on the HOA concentrations. Indeed, the discussion starts with a comparison of modelled primary organic matter and measured HOA concentrations (p11276 in the first version of the manuscript). However, one page later (p11277) also the CO2 distribution across and along the emission plume is discussed in almost similar extent. Since, in addition, Figure 5 seems not to be overloaded to us we decided to keep the CO2 data in this figure.

Technical comments:

p. 11251, l. 17: Please clarify what "the second process" refers to.

Reply by the authors: To make this clearer we modified the sentence to: "The aerosol particle size distribution in plume air masses was influenced by nucleation and growth due to coagulation and condensation in summer, while in winter only the latter process (i.e. particle growth) seemed to be initiated by urban pollution."

p. 11267, l. 10: Change "average O/C ratios are shown" to "average O/C ratios in organic aerosol are shown".

Reply by the authors: We changed the text according to the suggestion.

p. 11281, l.17: Replace "regression coefficient" by "correlation coefficient".

Reply by the authors: The reviewer is correct that "regression coefficient" is not the adequate term. We changed "regression coefficient" into "coefficient of determination", which is the square of the correlation coefficient and which we used here.

p. 11283, l.14: Change "secondary pollutants as particulate sulfate" to "secondary pollutants such as particulate sulfate".

Reply by the authors: We changed the text according to the suggestion.

Fig. 1: In the figure caption, replace "squares" by "diamonds".

Reply by the authors: We replaced "squares" by "diamonds" as suggested.

Tab. 1: The columns showing the plume contributions in summer and winter are hard to read. Please split these two columns into four columns, showing the absolute concentrations and the relative contributions in separate columns, respectively.

Reply by the authors: We changed the table (now Table 2) according to the suggestion. The second and third columns are now both split into two columns each.

Tab. 2: Replace "regression coefficient" by "correlation coefficient" in the caption of Tab. 2.

Reply by the authors: We changed "regression coefficient" into "coefficient of determination" as discussed above.

Anonymous Referee #3

The manuscript shows a study analyzing the measurements done using a mobile Lab during two deployments (winter, summer) on the Paris metropolitan area. It shows a very thorough explanation on how the measurements were performed and how they were post processed. The measurements were classified as background or affected by the Paris plume, with the methodology of classification also thoroughly described.

The classification was used to contrast plume and background conditions, analyze the dilution characteristics of the plume and to explore transformation processes within the plume. The manuscript is well written, the contents are in the scope of the Journal and I believe it is a contribution to the field. This is why I recommend publication after minor revisions.

My main comment is that I think there is a problem in the methodology used. The main method of air mass classification used in the paper is through enhancements of fresh pollution markers. Afterwards, one of the main results reported is the increase in fresh pollutants from background to megacity air masses. So it's expected that you are going to find these results if you are using the same criteria for classification. I would like to see the methodology validated by using an independent identification of plume emissions, such as backtrajectory analysis, which the authors used but on a very limited part of the study. The authors mention they use meteorological observations and air pollution reanalysis but these are not shown. Some of these examples could be added to the supplemental material.

Reply by the authors: We understand that the reader might see such a problem in the methodology used in our work. This is mainly due to the presentation of the work and not due the work itself. Actually, the identification of the plume in the data was performed in several steps. The first step was to determine the direction of air mass transport from meteorological forecasts and from Prev'Air pollution forecasts. Based on this information the measurement route was selected. This was described in detail in von der Weiden-Reinmüller et al., 2014 (AMT 7, p279-299). After the data were validated and cleaned for local pollution influence, all data sets from all measured variables were searched for fingerprints from the emission plume in the region where approximately – according to the pollution forecasts – the plume was expected. This step was made without any preference for certain types of variables like concentrations of fresh pollution markers. However, in this process it turned out that the plume was mainly visible in the data sets of the fresh pollution markers. Therefore these data were used to determine the location of the plume and of the background without plume influence. Only after this determination average plume and background concentrations were calculated for all variables.
In the manuscript the point of view was from the results side: Only for those variables where a clear plume was visible the distribution of concentrations was shown. These were of course those markers where the largest contribution from the plume was found.
In order to avoid the impression of this apparent problem in the methodology we revised section 3.1 (Emission plume identification) accordingly. In the revised text the actual process of how the plumes were identified is now better reflected. This, we hope, now shows better that there was no preference for fresh pollution markers to identify the plume but that these markers turned out to be the best markers for the plume, simply because their concentrations are those that are most strongly enhanced in the area where, according to the Prev'Air calculations, the plume was expected.

Other comments:

Page 11265, Lines 14-20. I disagree with the statement " … while air masses further away from Paris show nearly constant background values". On Fig 1b, CO_2 shows a clear decreasing trend as you

move away from the city. The trend is not so clear in the other chemical species, but the values further away from the city seem to be the lowest. The values you are flagging as Background air masses could be City emission plume diluted after axial transport, or could be that the mobile sampler moved from the center of the plume (higher concentrations) to the side of it (more diluted) as it got away from the city.

Reply by the authors: Indeed, CO2 shows a clear decreasing trend in the distance range that was identified as "background". We agree with the reviewer that this might be confusing to the reader. We also agree with the statement that the measurements further away from the plume might have missed the plume center.
The definition of plume and background ranges was the result of intense studies of the values of the various variables as a function of distance (or direction) from Paris. Since it is not likely that different components of the plume are unmixed and show different distributions we assume that the combination of information from different pollutants gives us the best information on the extension of the plume. The definition of plume and background boundaries is therefore based on this combination of information.
Variability in the concentration levels with distance to the city has different reasons. Since only for CO2 a clear decrease with increasing distance to the city is observed and not for the other variables, we assume a different reason for this behavior. As described in the text such quasi Lagrangian measurements lasted for many hours. Since in any cases the measurements started in the morning near Paris and the most distant point was reached in the early afternoon we assume that this decrease observed for CO2 is – at least to a certain degree – an artefact due to typical diurnal variations of this trace gas.
In order to make all this clearer we reworded the paragraph that describes the plume and background definition. In this location in the text we included a statement about potential influences by diurnal variations (which was already made later in the text) and about potential influences by missing the plume (which was also already in the text) on measured pollution levels. In addition we made clearer that the definition of plume and background ranges is based on a combination of information from various pollutants.

Page 11267, Lines 22-24. You could use statistical testing to backup this claim.

Reply by the authors: As shown in Table 1 (last row, after revision now Table 2) the uncertainties of the average plume and background values of the O/C ratio are so low that within the given precision they do not differ from zero. With a decrease of 0.05 and 0.08 for summer and winter, respectively, the differences between plume and background O/C ratios are much larger than the uncertainty of these values.
We added this information to the text: "The average O/C ratios in organic aerosol confirm that during both seasons the oxidation level of the organic aerosol in emission plume air masses is significantly lower (much more than the uncertainty of the values) than that in background air masses."

Fig 5b. How confident are you that the measurement actually stayed in the direction where the plume was moving? As seen in Fig 5a, city plumes could tend to be narrower than predicted, so maybe that low background value you are finding is because the sampler went quickly out of the plume. Backtrajectory analysis could be helpful here.

Reply by the authors: As described in the text we compared the experimentally determined plume directions, extensions and shapes with those determined from Prev'Air forecasts and re-analyses, which is a more detailed model compared to the backtrajectory calculations we used. Since we found some discrepancies between the modelled and measured plume distributions and directions which were not systematic, we know that differences in modelled plume directions and measurement routes

do not necessarily have to mean that the measurements were performed outside the plume. Therefore we just stated in the text (p11277 line 10-12) that such effects cannot be excluded. Unfortunately we cannot give any better information on the probability whether such effects actually happened during our measurements.

Section 3.4, "Quasi-Lagrangian axial measurements:" The results of this sub-section are not showed in tables or figures. You could show them in the supplementary material.

Reply by the authors: As stated on p11278 line 28 – p11279 line 2 (first version of the manuscript) with this analysis approach we were not able to detect any significant chemical transformation processes beyond the expected dilution of the plume. The reasons for this are presented in the following lines (p11279 lines 3-15). Since on the one hand there was mainly noise left over after this analysis and no transformation above noise level could be observed, and on the other hand many calculations and therefore many data have been involved in this analysis, we think that it does not make sense to add this information to the supplementary material. In the case that an individual reader is interested in this analysis we would prefer to have a chance to discuss this directly.

The Summary states "The cross sectional profile of the plume is typically Gaussian-like while the axial decrease of fresh pollution concentrations shows an exponential shape." However, the authors present only 1 case for each. More cases need to be added (maybe just in the supplement) to backup these conclusions. Also include model results to see how representative is the case shown in the main manuscript.

Reply by the authors: We added five more measurement examples, including the corresponding model data, to the supplement (Sect. S3, Figs. S5 – S9). Three examples show the Gaussian-like cross sectional profile of the emission plume, two examples show the exponential axial decrease of fresh pollution concentrations in the emission plume.

Technical corrections:

Change Beekman 2013 to 2014

Reply by the authors: We changed this citation to "Beekman et al., in preparation for ACPD, 2014".

Page 11261, Lines 3-4. This sentence is not clear; it contradicts the sentence in lines 7-8. Please rephrase.

Reply by the authors: We do not think there is a contradiction between these two sentences, though we agree that as written, it might be confusing to the reader. The sentence in line 3-4 claims that no data were removed when the measurements were potentially contaminated by local sources – independent on the concentrations measured during these times. The sentence in line 7-8 claims that individual data points with very high concentrations – independent on the source of these concentrations – have been removed.
To make this clearer – and to include the next comment – we changed the two sentences into: "Local pollution contamination, as defined above, was not removed from the AMS data before PMF application but afterwards from the resulting factor time series." and "Therefore, data points with intense peaks in the organic time series, independent of the cause of these peaks, were removed before PMF was applied."

Page 11261, Lines 7-8: Change to “Therefore, data points with intense peaks in the organic time series were removed before PMF was applied.”

Reply by the authors: This was already covered by the response to the previous comment.

Page 11265, Line 1. Fig 1 does not show PAH, shows HOA.

Reply by the authors: Thank you for this hint. We changed the sentence accordingly.

Fig. 2. You could add the mean in the box and whisker plots as a symbol in the middle of the box/whisker. The mean is in Table 1 but I think it’s better to repeat this information in the box and whisker plots. Page 11267 Line 5 and Fig 2 Caption says that this figure has the “mean” but it’s not plotted. The box and whisker represent the data distribution, not the mean.

Reply by the authors: We added the mean values to the box and whisker plots as suggested. While doing this we realized that the values for LV-OOA in the winter measurements were not correct. We corrected these values in both, the table and the text.

List of all relevant changes made to the manuscript:

- We added a new table (now Table 1) that summarizes the instruments used for the measurements presented in this study.
- In Sect. 2.4.2 (line 229-230) we added information about the fraction of the total mass for m/z larger than 100.
- In Sect. 2.4.3 (line 276-290) we extended the discussion about the uncertainty of the PMF results.
- In Sect. 3.1 (line 374-392, line 400-403, line 416-417) we extended the description of the identification of the emission plume.
- In Sect. 3.1 (line 420-429) we extended the discussion of the data presented in Fig. 1b.
- In Sect. 3.2 (line 659-661) we added information about the uncertainty of the data shown in Fig. 4.

List of relevant changes made to the supplement:

- We added the new Sect. S3 presenting more measurement examples showing the Gaussian-like profile of cross sections of the emission plume and the exponential axial decrease of fresh pollution concentrations inside the emission plume (Figs. S5-S9).
- We added the reference Zhang et al., 2013.