Replies to Referee #1

The authors would like to thank Anonymous Referee #1 for his/her comments as well as for the grammar corrections.
The Referee's comments are copied below and followed by our replies.

General: Not enough information is presented regarding the factorisation of the HNMR data (the AMS PMF analysis is covered by the Saarikoski et al. paper). The authors state their reasons for not using the solution sets with 6+ factors in section 3.4, however these results are not shown graphically, which would be very useful to see. These could very easily be included as supplementary material. Additionally, the factorisation also presumably needed to employ an estimate of measurement precision but the authors give no explanation for how this data was generated. The fact the Q/Qexp values in figure 5 are very high (around 30) may suggest that these uncertainty estimates are too low and therefore potentially wrong.

Authors Reply. We accept the Referee's suggestions to improve the discussion about the NMR factor analysis methodology. We will treat these aspects in the revised sections 2.3 and 3.4 and in a new section of the supplementary material, which will report the diagnostics (the Q function) as well as alternative solutions with a greater number of factors. Most importantly, it should be noted that Figure 5 of the original manuscript reported the Q/Qexp for the MCR and NMF methods, not PMF. These algorithms do not account for the measurement uncertainty and the function (Q) to minimize is the sum of squares:

$$Q^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{i,j} - g_{i,k} * f_{k,j})^2$$

The fact that the above sum is not normalized by the measurement uncertainties means that Q is not adimensional, and its actual value depends on the units used in the sample concentration matrix. Therefore, in general the ratio Q/Qexp, where Qexp is defined here analogously than for PMF, will *not* converge to 1. We acknowledge the fact that the use of the ratio Q/Qexp as a diagnostic for methods other than PMF can be misleading. For this reason, we now simply report the variation of Q with the number of factors as a diagnostic for the factor analysis performed using the MCR and NMF techniques. Figure 5 of the original manuscript was moved in supplementary material and its new version now is Fig. S1.

P33354, L8: The use of an exclusively flaming fire presents an issue because real domestic fires produce emissions through a combination of flaming and smouldering. Were smouldering fires investigated? If not, why not?

Authors Reply. The exact procedure was the following: the fire was started and kept flaming for about 90 minutes by several add-ons of beech logs. At the end, the fire was smoldering. All phases were combined in one filter sample. We will add these specifics in the revised version of section 2.4.

P33362, L16: This concluding sentence seems to be a little pointless. The high degree of substitution in organic particulates is something that is already well established, so I do not see the need to view this as confirmation.

Authors Reply. Accepted. We removed the sentence.

P33372, L1: The Jimenez and Ng references are perhaps not the most appropriate because they do not deal with biomass burning specifically. Jolleys et al. (doi: 10.1021/Es302386v) and Cubison et al. (doi: 10.5194/acp-11-12049-2011) may be more appropriate.

Authors Reply. Accepted.

Figures 1, 3 and 8: The regressions presented presumably use standard least squares fits, however given that it is not apparent which of the two measurements is the most accurate, orthogonal distance regression would probably be more appropriate.

Authors Reply. We agree. We re-calculated the regression line equations with a total least squares (orthogonal) method in Fig. 1 and 8. With respect to the previous analysis, the new regression lines brought to slightly different conclusions about the relationship between WSOC and the OC concentrations accounted for by the AMS factors. The new slopes in Figure 1a are of about 1 (1.15 and 0.91 with and without BBOA, respectively), thus suggesting a substantial overlap between WSOC and the AMS factors for the oxygenated organics, leaving HOA the only plausible contributor to the water-insoluble organic carbon.

We kept the regression lines calculated with the ordinary least squares (OLS) regression in Figure 3c, because the result was not significantly different and also because in that case our x data (O_3/NO_2) are affected by a much smaller instrumental uncertainty than the y quantities, and the use of OLS seemed more appropriate (Smith et al., American Journal of Physical Anthropology, 140:476–486, 2009).

Figure 5: This figure appears to be of a low quality. The final version should be in a vector format (e.g. eps)

Authors Reply. The quality was improved and the figure was moved into supplementary material.

Replies to Referee #2

The authors would like to thank Anonymous Referee #2 for his/her comments.

The Referee's comments followed by our replies are listed below.

Comment: I have one comment on factor analysis of NMR. The number of samples is 17, which appears too few for factor analysis. The high Q/Qexpected (≈20, Fig. 5) for 5-factor solution indicates a large underestimation of error which should be addressed in the text.

Authors Reply. The definition of a minimum dataset size in factor analysis is an issue not yet resolved univocally: even if there are a lot of different theoretical rules (Arrindell & van der Ende (1985), Velicer & Fava (1998), and MacCallum et al. (1999) have reviewed many of these recommendations), many studies have demonstrated that the general rules of thumb of the minimum sample size are not always valid and useful (MacCallum et al., 1999; Preacher & MacCallum, 2002). The minimum level of N (dataset size) was object of a very high number of studies that demonstrated its dependency on other aspects of design, such as: the *communality* of the variables (percent of variance in a given variable explained by all the factors jointly and interpreted as the reliability of the indicator) (Hogarty et al., 2005; MacCallum et al., 2001; Costello & Osborne, 2005); the *degree of overdetermination of the factor* (or number of factors/number of variables) (Preacher & MacCallum, 2002; MacCallum et al., 1999); etc.

For the above reasons, we chosen the most appropriate factor number on the comparison between the outcomes of distinct factor analysis methods. The best agreement was achieved by far for the 5-factor solution, and this remains the most realistic NMR spectral deconvolution in this experiment.

In fact, the present study is based on a short timeline of samples (# = 17), but the variability in NMR composition was high, and this is witnessed by the poor correlations between the spectral profiles (except between the two "polysubstituted aliphatics"). In turn, the variability in the chemical composition was influenced by the great variability in the weather and atmospheric dynamic conditions which characterized this field campaign.

In respect to the meaning of the diagnostic Q/Qexp we copy here our reply to one of the Refereee#1's criticism:

It should be noted that Figure 5 of the original manuscript reported the Q/Qexp for the MCR and NMF methods, not PMF. These algorithms do not account for the measurement uncertainty and the function (Q) to minimize is the sum of squares:

$$Q^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{i,j} - g_{i,k} * f_{k,j})^2$$

The fact that the above sum is not normalized by the measurement uncertainties means that Q is not adimensional, and its actual value depends on the units used in the sample concentration matrix. Therefore, in general the ratio Q/Qexp, where Qexp is defined here analogously than for PMF, will *not* converge to 1. We acknowledge the fact that the use of the ratio Q/Qexp as a diagnostic for methods other than PMF can be misleading. For this reason, we now simply report the variation of Q with the number of factors as a diagnostic for the factor analysis performed using the MCR and NMF techniques.

Additional references

Arrindell, W. A., & van der Ende. J.: An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement, 9*, 165 – 178, 1985.

Costello, A. B., & Osborne, J. W.: Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*(7), 2005. from http://pareonline.net/pdf/v10n7a.pdf

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford K. R.: The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement, 65*, 202-226, 2005.

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong S.: Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research, 36*, 611-637, 2001.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong S.: Sample size in factor analysis. *Psychological Methods, 4*, 84-99, 1999.

Preacher, K. J., & MacCallum, R. C.: Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Sizes. *Behavior Genetics, 32*, 153-161, 2002.

Velicer, W. F., & Fava, J. L.: Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*, 231-251, 1998.

Replies to Referee #3

The authors would like to thank Anonymous Referee #3 for his/her through review.
The Referee's comments are copied below and followed by our replies

Detailed Comments:
P33348, L23: It says that investigating the correlations between high-resolution tracers and met data is the "main advantage" of PMF. It's not clear what high-resolution tracers is. And what is it the main advantage compared too? The main purpose of using PMF on aerosol measurements is clearly to identify source contributions. Please reword to reflect exactly what is meant.
Authors Reply. Accepted. We reworded the sentence into the following: *PMF is becoming common for organic source apportionment (Zhang et al., 2005; Lanz et al., 2007) and factor analysis of AMS data using PMF in its diverse implementations (including e.g. ME-2) has extracted biomass burning organic aerosol (BBOA) factors in multiple sites (Aiken et al., 2010; Elsasser et al., 2012; Crippa et al., 2013a etc.).*

P33356, L24: "So called "OOA-a". Check other AMS-PMF records. I think this is typically called OOA-1 or LV-OOA. If so you should refer as that if using "so called" and then note that Saarikoski et al. calls it OOA-a
Authors Reply. True. We corrected the text.

P33357, L17-18: Why use conversion factors to calculate organic carbon loadings measured by the AMS? An HR-ToF-AMS was used for those measurements, for which high-resolution fitting calculates elemental concentrations directly? And Saarikoski et al. clearly did the HR analysis as it's presented throughout that paper.
Authors Reply. The so "conversion factors" were indeed the elemental ratios calculated by Saarikoski et al. We rephrased the sentence it in this way:
In the following discussion, all concentrations of AMS factors for organic matter will be converted (using the OM/OC and H/C ratios reported by Saarikoski et al. (2012)) to µg/m^3 of organic carbon (µgC m^{-3}) for comparison with thermal analyses and of organic carbon and organic hydrogen for comparison with the NMR spectroscopic data.

P33359, L10: Is that necessarily true? Couldn't the higher hydrogen measurements in the AMS compared to the NMR be just as likely due to the AMS measuring the non-water-soluble fraction of aerosol? Both are probably contributing factors.
Authors Reply. This does not hold for the data shown in Figure 1d where the AMS concentrations refer only to the sum of factors representative for the WSOC with two possible representations: a) OOAs + NOA, and b) OOAs + NOA + BBOA. The HOA was left out in this calculation.

P33360, L8-9: "H-C-O" is not a hydroxyl group. I'm assuming this is shown this way to indicate the carbon-attached H that NMR can actually measure. Perhaps this is standard notation in NMR, but it's chemically wrong as written. Please reword. It would probably be clearest if you note the functional group being quantified, the C-OH group in R-C(Hx)-OH, and then explain what NMR measures, the first H in "R-C(Hx)-OH". Otherwise, non-NMR people may be confused here.
Authors Reply. True. We changed it into "alkoxy groups" which is chemically correct..

P33361, L7: Whey is O3/NO2 used as a photochemical aging marker. Please provide references where this has been used or explain why it is being used rather than more commonly used indicators such as Ox=NO2+O3, NOx/NOy, or VOC ratios.
Authors Reply. The O3/NO2 ratio is not a universal tracer for photochemical age but it is a suitable one for this specific study. For most of the days, the diurnal cycles of NO2 and O3 resulted

anticorrelated (Figure S9 in Saarikosky et al. 2012) because not only of the nighttime titration of ozone by NOx but also of the daytime development of a convective mixing layer. The morning increase of the mixing layer height was responsible for the dispersion of NOx which had accumulated overnight and for the entrainment of ozone from residual layers. Such entrained layers were characterized by aged organic aerosols and a relatively high sulfate content (Fig. 10 in Saarikosky et al. 2012). We used the O3/NO2 as a compact metric to trace both the photochemical and the atmospheric dynamic processes in this specific experiment.

P33361, L8-11: Confusing sentence. Reword.
Authors Reply. We omitted the sentence as unnecessary for the discussion.

P33362, L1-3 & P33362, L3-6: How can you estimate the amount of carboxylic acids from the AMS? You cannot assume that for each CO2+ you have one carboxylic acid if this is what is being assumed here. Some ends up as CO+ while others can end up as part of a larger organic fragment. This fragmentation depends on the type of acid (see Fig 8 in Duplissy et al.: www.atmos-chem-phys.net/11/1155/2011/)
Authors Reply. The authors acknowledge that an exact relationship between $C_2H_3O^+$ and CO_2^+ signals and carbonyl and carboxylic acid concentration has not yet been clearly established as the laboratory studies provided ambiguous results in this respect (Duplissy et al., 2011). However, it is possible that the fragmentation patterns of the specific carboxylic acids are less important in the ambient aerosol due to its greater molecular complexity. In fact, the CO2+ fragment has already been used as a proxy for total carboxylic acids in ambient samples (Takegawa et al., Aerosol Sci. Technol. 41, 418-437, 2007). We therefore partially accept the Referee's criticism and we delete the statement about the presumable degree of substitution based on AMS measurements and rephrased the sentence: "*However the carboxylic acid concentrations determined by the NMR method are lower than those from AMS in most cases (Fig. 4) and their correlation is not good (R = 0.19). Such discrepancy could be explained by the unavoidable loss of volatile carboxylic acids during the chemical derivatization procedure, or by the formation of CO_2^+ fragments in the AMS vaporizer from groups other than carboxylic acids, like esters and peroxides*".

P33365, L9-11: I think the authors are saying that ammonium nitrate was higher due to increased aerosol water content? This seems highly speculative. Couldn't higher ammonium nitrate be due to many other factors such as greater photochemical processing (HNO3 gas), cooler temperatures, higher NH3, different aerosol sources transported to site? Please clarify this connection, remove or provide stronger evidence.
Authors Reply. Those days (period III) were characterized by low Tmax and relative humidity frequently close to 100% (see Figure 1 in Saarikoski et al., 2012), therefore the ambient conditions were favorable for the formation of particulate nitrate. We made the statement less speculative: "*Period III of the campaign was characterized by a stably stratified atmosphere, low maximum temperatures and high humidity (Saarikoski et al., 2012), with conditions favorable for the formation of particulate nitrate which indeed showed the highest concentrations of the campaign in those days. The correlation between F5 and aerosol nitrate suggests that the low-molecular weight amines occurred in the aerosol mainly as aminium nitrate and that its formation was regulated by gas/particle partitioning similarly than for ammonium nitrate*".

Table 4: Why are Persons correlation coefficients (R) shown in Table 4 but R-squared values shown in the figures? Unless good reason, seems it should be consistent throughout.
Authors Reply. We converted all coefficients to R-values for sake of consistency through the text.

Figure 3a: It would be useful to show where the factors for this study lie on this plot. Consider adding these points.

Authors Reply. We accepted the suggestion of the Referee and added the functional group composition for the five factors to Figure 3a.

Figure 4: It looks like if one made correlation plots for the bottom two panels there may be no correlation. This lack of correlation seems to be skipped over and would seem important addressing the agreement of the methods and when/why they might be different. Can the authors add text discussing this.

Authors Reply. We did not discuss the lack of correlation between the AMS and the NMR concentrations in this dataset, because the database itself is very small and the coefficient of variation of the substitution degrees (amount of functional group over total WSOC) is also very small: between 17% and 21% (calculated for each concentration series as the ratio of the standard deviation to the mean), with the exception of the COOH groups (as fraction of WSOC) determined by NMR which vary to a greater extent (53%) because one sample showed a substitution degree double than the other four. The fact that the NMR results for the COOH groups are not really in line with the AMS measurement is already highlighted in the text. In respect to the C=O, we believe that the small sample size and especially the small coefficients of variation do not allow to assess the correlation between the concentrations provided by the two methods. The results provided in this study are more oriented to compare the *average* functionalization degree measured by the NMR derivatizaion technique with that inferred by the AMS mass fragment concentrations.

Figure 8: Add R2 values to the figs or captions. On page 33367, lines 9-11, it says the correlations were very good but no statistical measure given.

Authors Reply. Corrected.