Author's response:

We thank the Referees for the careful revision and comments which helped improving the overall quality of the manuscript.

A point-by-point answer (in regular typeset) to the referees' remarks (in the *italic typeset*) follows, while changes to the manuscript are indicated in blue font.

In the following page and lines references refer to the manuscript version reviewed by the anonymous referees.

Anonymous Referee #1

Received and published: 27th February 2017

General Comments:

This manuscript reported the chemical composition of $PM_{2.5}$ over al full year in an urban site in France, based on offline filter analysis. The focus is to study the sources of organic aerosol by performing PMF analysis on AMS data. The authors did very careful analysis to optimize PMF results. The authors found that BBOA is the dominant OA source in winter and OOA is the main OA source in summer. The authors also compared the offline-AMS results with online-AMS results, but the measurements were not performed simultaneously. The levoglucosan/BBOA ratio was found to evolve over time, which was attributed to different types of biomass burning combustions, instead of the photochemical aging. Overall, the analysis is adequate, the conclusions are generally solid, but none of the results are earth shattering or unexpected given the preexisting literature. I recommend accepting manuscript after major revisions.

We thank Anonymous Referee #1 for the careful review and inputs which helped improving the overall quality of our work and its impact. We recognize that this work presents results relative only to one station, showing an expected pattern of source seasonality, e.g. dominance of biomass smoke in winter and SOA in summer. Despite this, we strongly believe that the work is incremental compared to existing literature and it presents innovative approaches for data analysis that can be used in future works and novel aspects regarding the composition of different aerosol sources. The following novel approaches and findings can be highlighted:

- ➔ The work introduces new methods for source apportionment validation such as the systematic comparison of the PMF factor mass spectra with literature profiles using cosine similarity. This approach quantitatively examines how variable the mass spectral profiles extracted by PMF are for each of the different sources, and how distinguishable profiles from different sources are, allowing a more robust validation of the identified factors. In our opinion such systematic analysis of ME-2 model outputs should become a standard for the optimization and validation of source apportionment results.
- ➔ While the high contribution of biomass burning aerosols during winter may not be surprising, this work presents one of the first identification of the origins of this fraction. Till now, the biomass burning fraction detected in Europe, based on the analysis of specific markers (e.g. levoglucosan), is often related to residential heating. Here, by combining several techniques (AMS/PMF and molecular speciation), we

could clearly distinguish emissions from residential heating and agricultural burning to this fraction, with the latter process found to be very important during the land clearing period, at least in this region of Europe. Therefore, this work offers on the one hand analysis techniques that can be applied in the future to distinguish between different biomass burning emission processes, and on the other hand it unveil one of the reasons behind the observed variability in biomass burning composition (e.g. markers ratios).

➔ The work also reveals that both online- and offline-AMS PMF tend to apportion rapidly formed SOA components to primary PMF factors, rather than to the OOA factors. We show that this is especially the case for nitrocatechols, formed from the oxidation of lignin-pyrolysis derived compounds in biomass burning fumes and is therefore most important during winter when lignin rich biomass is burned for residential heating. By examining the oxidation rates of these compounds towards OH and NO_3 radicals we show that these compounds have a lifetime of minutes in the atmosphere, which explains the apportionment of their oxidation products to directly emitted primary aerosols. This has implication on the technical separation between primary and secondary aerosols in the atmosphere.

➔ This study provides for the first time to the best of our knowledge, yearly contributions of industrial sources to the organic aerosol using AMS measurements and identify possible tracers (Se, fluoranthene, pyrene, phenantrene) which can be utilized in future studies. Although industrial sources can significantly change characteristics depending on the industrial processes involved; petrochemical activities, shipping and metallurgical industries, such as those encountered in the studied area, are widespread and may represent a significant aerosol source in many other industrialized areas. Here, we demonstrated that these processes may be efficiently traced by offline-AMS.

➔ In addition, other novel technical aspects presented in this study could be highlighted, including the determination of the recovery/water solubility of industrial emissions and the first identification of the interference of inorganic carbonates to OA measurements. While this interference might be most important for the offline analysis, it can also influence online AMS measurements, e.g. for studies conducted at dusty locations using a $PM_{2.5}$ aerodynamic lens, or biomass burning direct emission studies which can be affected by high concentrations of carbonates in the ashes.

Source Apportionment

Major comments

1) Instrument inter comparison.

For the filter analysis, some ions and species (such as SO4, NO3, WSOC, etc) are quantified by more than one techniques. It is natural to include the instrument inter-comparison in the manuscript. In one previous study of the authors (Bozzetti et al., 2017), the comparison of SO4 concentration between offline-AMS and IC has a non-linear relationship. Does the non-linear relationship also exist in current study? Have the authors investigated more about the non-linear relationship since the previous study? Also, I wonder if the AMS_SO4/IC_SO4 ratio is similar to AMS_OC / WSOC ratio?

We need to state that major ions in this work were quantified only by IC, while WSOC was quantified only by TOC analysis of the aqueous filter extracts. Offline-AMS did not provide

quantitative concentrations of major ions or WSOC for reasons that shall become clear in the following. In this study, similarly to Bozzetti et al. (2017), we observed a non-linear relation between IC and offline-AMS SO_4^{2-} and NO_3^- (Fig. D1). The causes of the observed non-linear relations are the following:

I) Transmission efficiency through the aerodynamic lens. The nebulization of differently concentrated filter extracts generates aerosol particles characterized by different size-distributions, i.e. the nebulization of more concentrated filter extracts generates larger particles. In case of highly diluted filter extracts, the generated aerosol particles approach the low cut size of the AMS aerodynamic lens. This yields lower transmission efficiencies for diluted extracts in comparison to more concentrated solutions. Also, the influence of particle size on the detection implies that the relationship between the concentrations for a certain species measured by IC and offline AMS is not constant, but depends on the abundance of other species. In other words, for two filter extracts characterized by the same e.g. NO_3^- concentrations, the extracts characterized by the highest SO_4^{2-} + NH_4^+ + Cl^- + organic concentration will show a higher NO_3^- sensitivity in the offline-AMS analysis due to the nebulization of bigger particles better transmitted through the AMS aerodynamic lens. This implies that applying a simple transmission efficiency correction (i.e. only function of one component, e.g. NO_3^-) won't provide accurate results.

II) Scatter in the correlations between NO_3^- or SO_4^{2-} from IC and NO_3^- or SO_4^{2-} from offline-AMS can derive from the AMS electron impact fragmentation of organo-NO_3^- or organo-SO_4^{2-} which leads to the formation of fragments attributed to inorganic NO_3^- or SO_4^{2-}. As a result a certain fraction of AMS-NO_3^- and AMS-SO_4^{2-} is of organic origin.

III) Another source of scatter in the correlation between IC SO_4^{2-} and offline-AMS SO_4^{2-}, is the presence of refractory SO_4^{2-} salts (e.g. Na_2SO_4 and $CaSO_4$ which are detectable by IC, but not by AMS. Also, the water solubilization of non-refractory SO_4^{2-} salts (e.g. $(NH_4)_2SO_4$)) can lead to the formation of refractory SO_4^{2-} species due to the possible recombination of SO_4^{2-} with other cations in solution (e.g. Na^+, Mg^{2+}, and Ca^{2+}).

Because of the points I), II) and III) also the $SO_4^{2-}{}_{AMS}$:$SO_4^{2-}{}_{IC}$ ratio differs from the $WSOC_{AMS}$:$WSOC_{IC}$ ratio (Fig. D1). For these reasons only inorganic ion concentrations from IC were reported in the main text. Fig. D1 a) shows that relation between $SO_4^{2-}{}_{AMS}$:$SO_4^{2-}{}_{IC}$ and $WSOC_{AMS}$:$WSOC_{IC}$ tends to deviate from a 1:1 line for high Ca^{2+} concentrations. Specifically, SO_4^{2-} is less efficiently detected than WSOC by offline-AMS in presence of high Ca^{2+} concentrations, indicative of the probable recombination of SO_4^{2-} with Ca^{2+} in solution, leading to the formation of refractory SO_4^{2-} salts (e.g. $CaSO_4$) not detected by AMS

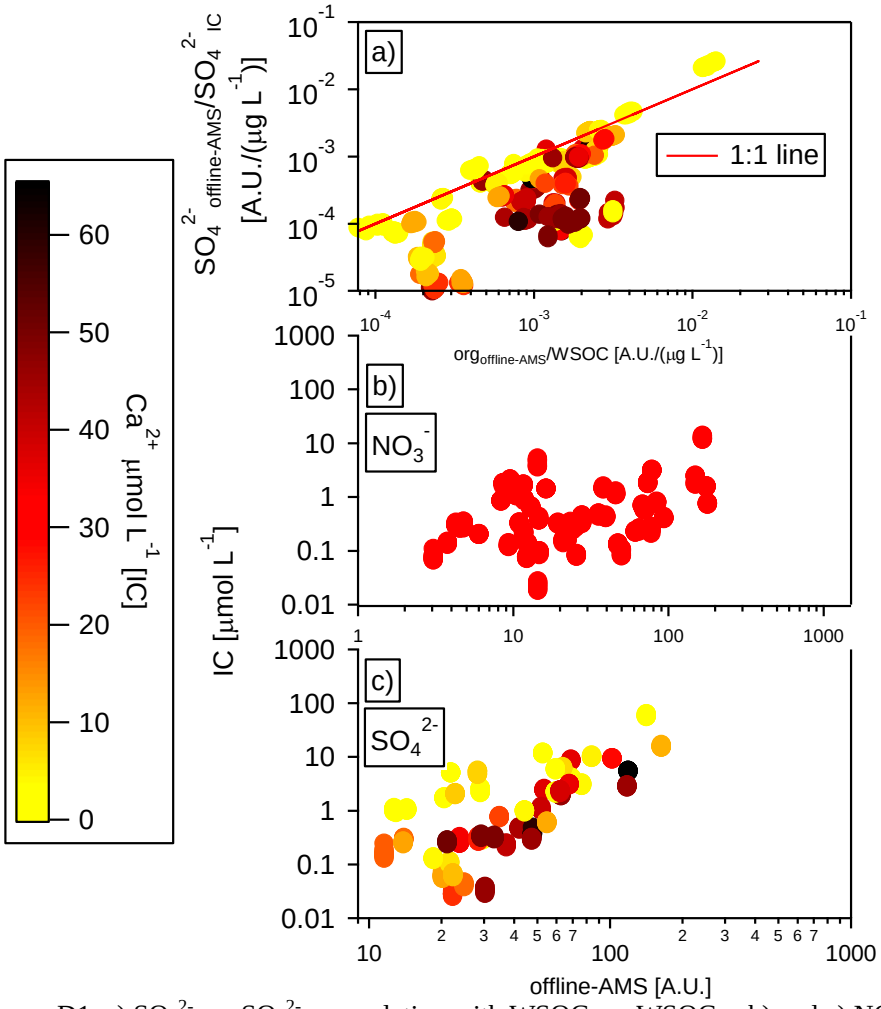


Figure D1. a) $SO_4^{2-}{}_{AMS}$:$SO_4^{2-}{}_{IC}$ correlation with $WSOC_{AMS}$:$WSOC_{IC}$. b) and c) NO_3^- and SO_4^{2-} : Offline-AMS comparison with IC.

2) *The interpretation of recovery ratio and extract the water solubility of OA factors. If I understand correctly, the recovery ratio is a function of nebulizer efficiency (i.e., species loss during nebulization), AMS collection efficiency (i.e., include the lens transmission efficiency. The collection efficiency at the vaporizer may or may not be included, depending on whether the CE has been applied), and the water solubility of OA factors. Since the species are internally mixed in the solution and in the nebulized particles, the nebulizer efficiency and AMS collection efficiency should be the same for all OA factors. Thus, the recovery ratio only depends on the water solubility of OA factors. Figure S12 shows that OOA has the largest recovery ratio and HOA has the smallest recovery ratio, which is consistent with that OOA is more water soluble than HOA. Thus, this provides a potential opportunity to estimate the water-solubility of OA factors. For example, could the authors make use of the instrument inter-comparison (i.e., SO4 between offline*

AMS and IC) to correct for the efficiency of offline AMS system, as similarly done in Xu et al. (2016)? Then, use the model proposed by Psichoudaki and Pandis (2013) to relate the fraction of a compound extracted in WSOC as a function of compound water-solubility at dissolution equilibrium. The model in Psichoudaki and Pandis (2013) is designed for filter extraction analysis. In this way, more useful information about OA factors can be extracted.

We agree with the reviewer's interpretation of the factor recoveries. As OA factors are expected to be internally mixed in the nebulized particles, their nebulization efficiency and AMS collection efficiency should be the same and hence their recovery would be governed by their water solubility. Daellenbach et al. (2016) stated that the calculated factor recoveries are consistent with the water solubility of these fractions, with HOA being barely water soluble (~13%), BBOA moderately water soluble (65%) and OOA almost entirely water soluble (90%). These factor recoveries have been revaluated in this study (see "Offline-AMS source apportionment optimization" section) and are consistent with the estimates of Daellenbach et al. (2016) (Fig. S12), based on collocated ACSM and offline-AMS measurements at another site. In addition, here the recovery for industrial OA has been assessed to be similar to that of BBOA (69%).

Based on the reviewer comment we have assessed the link between the factor recoveries and the water solubility of the compounds therein. We have assumed each of the components to comprise a single average surrogate in equilibrium between the aqueous solution and an ideal solution of water insoluble organic species. The adaptation of the equations in Psichoudaki and Pandis (2013) to our problem yield the following expression of the surrogate extracted fraction/recovery. We will refer to this fraction as R'_k, to draw a distinction between measured and calculated recoveries.

$$R'_k = \frac{m_k + V_W S_k + m_{OA}(1 - f_{WSOC}) - \left(\left(m_k + V_W S_k + m_{OA}(1 - f_{WSOC})\right)^2 - 4 m_k V_W S_k\right)^{1/2}}{2 m_k}$$

Here, m_k and m_{OA} are the total mass of a factor k and of the organic aerosol on the extracted sample. V_W is the volume of water used for extraction, f_{WSOC} the fraction of water soluble organics and the S_k the water solubility of the average surrogate compound representative of the bulk composition of the component k. This formulation should provide a highest estimate of S_k compared for example to considering the components in k to be insulated from the organic phase. Using this formulation, we estimate the recoveries obtained under our conditions to be consistent with S_k values of 10^{-3} g L^{-1}, 10^{-2} g L^{-1} and 10^{-1}g L^{-1}, for HOA, BBOA/COA/INDOA, and OOA, respectively. We have also assessed the sensitivity of R'_k towards the bulk aerosol composition, by varying f_{WSOC}, the total organic aerosol concentrations and the contribution of the factor of interest within the observed ranges. This sensitivity analysis suggests that for a similar solubility, the variability in the extraction conditions may influence the recoveries by 10 percentage points on average (see the upper and lower curves in Figure D2). These variations are relatively small, within our confidence

interval of the determined recovery parameters. We note that the extraction procedure adopted here favors the compounds' partitioning into the aqueous phase, given the high extraction volume compared to the sampled volume per extracted filter fraction: ~0.5 cm^3 m^{-3} vs. 0.1 cm^3 m^{-3} in the other studies (Psichoudaki and Pandis, 2013 and references therein). Under these conditions, all typical functionalized compounds would be extracted (Cappelli et al., 2013; Meylan and Howard, 1994a,b; Meylan et al., 1996)

We also note that the model used here is rather simplistic and the different components are expected to comprise a suite of compounds with a wide range of water solubility. This can be expressed in a solubility basis set by analogy to the volatility basis set (VBS). This simplification implies on the one hand that the solubility values provided here are only weighted average values for the solubility of different compounds contained in these components. On the other hand, the model provided here would significantly over-predict the sensitivity of the recoveries to the extraction procedure adopted (filter loading, bulk OA solubility and extraction volume). Again by analogy to the VBS, most of the compounds contained in one component may be either water soluble or insoluble under most of the extraction conditions, and only a minor fraction of semi-soluble compounds would be sensitive to the extraction procedure. Therefore, we note that the data we present here cannot be directly extrapolated to other studies and establishing a solubility basis set for the different components would require significantly varying the extraction conditions of the different samples followed by an assessment of the recovery, which is beyond the scope of this study.

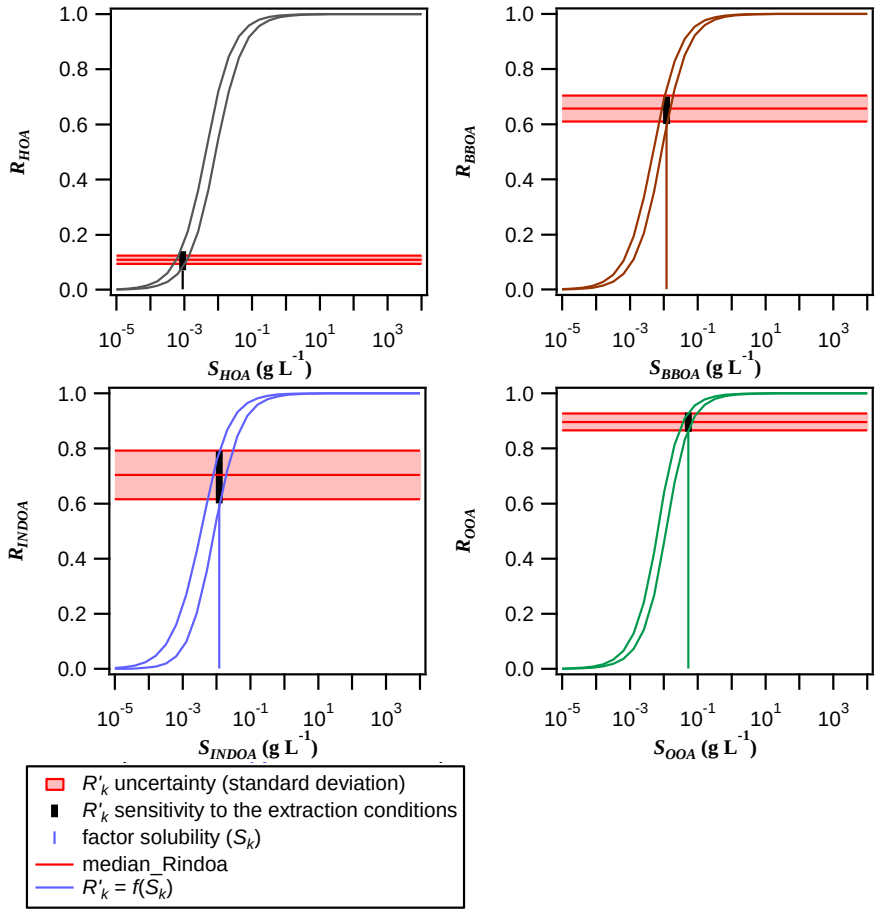


Figure D2. Sensitivity of the calculated factor recoveries R'_K to the factor solubility S_K (g L^{-1}). Vertical lines define the factor solubility calculated from the median factor recoveries (R'_K, horizontal lines) determined in this work .

This discussion was added to the revised SI.

3) The selection of PMF solutions is very careful, but some related descriptions require more clarifications.

I) Did the author constrain all OA factors or only COA and HOA? More importantly, it should be clearly stated that how the anchor profiles are selected? If the authors use the average reference mass spectra of HOA and COA as anchor profiles, would it make the analysis easier?

For both offline- and online-AMS we constrained HOA and COA profiles from Mohr et al. (2012) and Crippa et al. (2013b) respectively. The HOA profile from Mohr et al. (2012) was selected for offline-AMS consistently with Daellenbach et al. (2016), since the same factor recovery distributions were applied in this work. The same profile was applied to online-AMS for consistency. Overall, as discussed in the SI, the HOA profiles from literature show high cosine similarities with each other's, suggesting that the AMS mass spectral fingerprints from

traffic exhaust are relatively stable from station to station and consistent also with direct emission studies, making the selection of the constrained factor profiles not crucial. In addition, for practical reasons, the profile from Mohr et al. (2012) is the most useful because of the low amounts of missing ions.

More variability instead is observed among COA literature profiles. For COA we selected the profile from Crippa et al. (2013b) which showed the lowest $f C_2H_4O_2^+$ value among the considered ambient literature spectra. This guaranteed a better separation of COA from BBOA, as $C_2H_4O_2^+$ is strongly related to levoglucosan fragmentation (Alfarra et al., 2007). The use of average literature profiles is practically not straightforward because different HR peak fittings are performed in literature studies which yields different peak list and increases the amounts of missing variables. While this is indeed an issue for the PMF analysis, for the cosine similarity calculations we have overcome this issue by retaining only fragments associated with a small variability among the literature profiles. In this way, the generated profiles were characterized by a smaller number of fragments compared to the original literature spectra. This hampers the utilization of average spectra in the *a*-value approach, because anchor values for the missing/discarded fragments have to be assumed. For these reasons the calculated average profiles were not constrained.

The discussion about the choice of the reference spectra was added in the main text at P10 L23:

For both offline- and online-AMS the constrained HOA profiles were from Mohr et al. (2012), while the COA profiles were from Crippa et al. (2013b). The HOA profile from Mohr et al. (2012) was selected for offline-AMS consistently with Daellenbach et al. (2016), since the same factor recovery distributions were applied in this work. The same profile was applied to online-AMS for consistency. Overall, as discussed in the SI, the HOA profiles from literature showed high cosine similarities with each other's, indicating that the AMS mass spectral fingerprints from traffic exhaust are relatively stable from station to station and consistent also with direct emission studies, making the selection of the constrained factor profiles not crucial. More variability instead is observed among COA literature profiles. For COA we selected the profile from Crippa et al. (2013b) which showed the lowest $f C_2H_4O_2^+$ value among the considered ambient literature spectra (Crippa et al., 2013b; Mohr et al., 2012). This guaranteed a better separation of COA from BBOA, as $C_2H_4O_2^+$ is strongly related to levoglucosan fragmentation (Alfarra et al., 2007).

and P12 L21:

As already mentioned, the HOA and COA profiles were constrained using an *a*-value approach. Consistently with online-AMS we constrained the profiles from Mohr et al. (2012) and Crippa et al. (2013b) respectively. Unconstrained PMF runs for offline-AMS did not resolve HOA and COA factors.

II) Did the authors constrain industry-related OA (INDOA) in the offline analysis? The INDOA factor is resolved in whole year offline dataset and 2008 July online dataset, but not in 2011 February online dataset. Have the authors tried to constrain the INDOA for the 2011 February online dataset?

The INDOA mass spectrum was not constrained in the offline-AMS source apportionment. The INDOA factor profile displayed in Fig. 5 was resolved by the PMF model and represents

the INDOA mass spectrum (WSINDOA) of the water-soluble fraction. As mentioned in the manuscript, El Haddad et al. (2013) resolved by unit mass resolution online-AMS PMF an industrial profile at the same location during summer. This factor, similarly to offline-AMS showed abruptly changing contributions correlating with PAH concentrations. In particular El Haddad et al. (2013) reported simultaneous INDOA and AMS-PAHs increasing concentrations associated with wind directions from W/SW. By contrast, as discussed in the manuscript (section 3.1), the AMS-PAH variability was well explained by the BBOA factor and did not show increasing concentrations with wind directions oriented from W/SW (225°-270°, Fig. S14). For these reasons we preferred a more conservative approach without constraining of an INDOA factor since no clear evidences of significant contributions were found. Constraining an HR industrial profile, which is currently lacking in the literature and has similar features as other more important primary sources (e.g. COA and HOA), would result in much more uncertainties than currently is the case, where already selection criteria of the PMF solutions were set to well separate the primary sources.

III) The application of cluster analysis to select PMF solutions is very nice, but the description of cluster analysis is not clear in the main text. For example, what does "PMF solutions" refer to in Page 10 Line 33? Please be more clear that "PMF solution" is a full set of solution (i.e., including both time series and mass spectra of all OA factors). I suggest to remove some descriptions from the SI to the main text.

As suggested by anonymous reviewer #1 we modified the text as follows:

P10 L33: From the HOA and COA *a*-value sensitivity analysis we obtained a set of 121 PMF solutions each one including both factor profiles and factor time series.

In addition, we introduced a summary of the PMF solution optimization for both online- and offline-AMS in the main text following anonymous reviewer #1 and #2 suggestions (see reply to anonymous reviewer #2).

IV) After all the discussions on the optimization of PMF solutions, it is not clear what is the final PMF results. Did the authors use the average of all retained PMF solutions? It would be useful to comment on how the finally optimized PMF solution is different from that by using PMF2 solver without any constrain, since the PMF2 solver is most widely used method.

As described in the manuscript (P12 L6-8 for online-AMS, and P14 L 21-24 for offline-AMS), the source apportionment results for both offline- and online-AMS represent the average of the retained PMF solutions. For online-AMS, we identified a subset of HOA and COA *a*-value combinations optimizing the resolution of the two factors. This was performed by systematically analyzing the COA diurnal cycles using cluster analysis and systematically comparing PMF mass spectra with literature profiles using cosine similarity. The selected COA and HOA *a*-value combinations were subsequently randomly resampled when exploring

the model rotational ambiguity by performing bootstrap PMF runs. The average of the bootstrap runs represented our source apportionment final results.

Similarly to online-AMS, for offline-AMS we performed bootstrap PMF runs by randomly resampling COA and HOA *a*-value combinations. The PMF analysis in this case provided water-soluble factor concentrations. We subsequently selected the solutions in two steps. The first selection step was based on factor correlations with external tracers (6 criteria listed at P13, L3-12). Subsequently we applied factor recoveries combinations (Daellenbach et al., 2016) to the retained PMF solutions in order to rescale the water-soluble factor concentrations to corresponding total OC concentration. Only solutions showing unbiased OC residuals for all seasons together and for summer and winter separately were retained. The average of the retained solutions represented the offline-AMS final source apportionment results.

A summary of online- and offline-AMS source apportionment optimization strategies were added at P10 L11, and at P12 L13, as also requested by anonymous reviewer #2 (see answer to the 3rd major comment).

Unconstrained PMF runs for offline-AMS did not resolve HOA and COA factors. In the case of online-AMS source apportionment, leaving COA and/or HOA unconstrained enabled resolving COA only by increasing the number of factors (>5 factor solutions) while in the 4 factor solutions we observed a splitting of an OOA factor which could not be attributed to specific processes. Unconstrained PMF yielded HOA and COA time series well correlating with the constrained solutions (see Fig. below); however in the unconstrained case, HOA and COA factor profiles showed higher fCO_2^+ in comparison with literature studies (Crippa et al., 2013b; Mohr et al., 2012; Bruns et al., 2015; Docherty et al., 2011; Setyan et al., 2012; He et al., 2010) and in comparison with the constrained PMF runs. This in turn resulted in higher HOA and COA concentrations, with background night concentrations 2-3 times higher than in the constrained solutions, possibly indicative of mixing with oxidized aerosols. Similar differences between constrained and unconstrained PMF runs were also observed in Elser et al. (2016). Also, the HOA:NO_x ratio ($\mu g\ m^{-3}/\mu g\ m^{-3}$) matched typical literature values reported for France (0.02 Favez et al., 2010) in the constrained PMF case (0.023), while for the unconstrained approach it showed higher values (0.033). This discussion has been added in the main text P10 L23.

Using an *a*-value approach, we constrained HOA and COA profiles from Mohr et al. (2012) and Crippa et al. (2013b) respectively. Leaving COA and/or HOA unconstrained enabled resolving COA only by increasing the number of factors (>5 factor solutions) while in the 4 factor solutions we observed a splitting of an OOA factor which could not be attributed to specific processes. Unconstrained PMF yielded HOA and COA time series well correlating with the constrained solutions; however in the unconstrained case, HOA and COA factor profiles showed higher fCO_2^+ in comparison with literature studies (Crippa et al., 2013b; Mohr et al., 2012; Bruns et al., 2015; Docherty et al., 2011; Setyan et al., 2012; He et al., 2010,) and in comparison with the constrained PMF runs. This in turn resulted in higher HOA and COA concentrations, with background night concentrations 2-3 times higher than in the constrained solutions, possibly indicative of mixings with oxidized aerosols. Similar differences between constrained and unconstrained PMF runs were also observed in Elser et al. (2016). Also the HOA:NO_x ratio ($\mu g\ m^{-3}/\mu g\ m^{-3}$) matched typical literature values reported for France (0.02 Favez et al., 2010) in the constrained PMF case (0.023), while for the unconstrained approach it showed higher values (0.033).

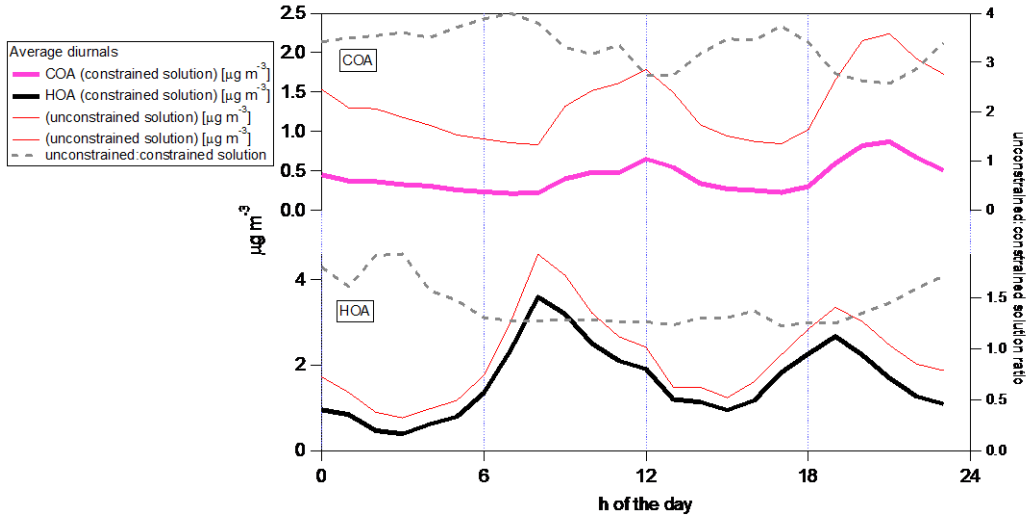


Fig. D3. Comparison of COA and HOA diurnal cycles from constrained and unconstrained PMF solutions.

Figure D3 was added to the SI as Fig. S5.

Following the reviewer's remark, we introduced in the main text a summary of the online-AMS optimization procedure (P10 L11).

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) Constrain of both the HOA and COA factors profiles adopting an *a*-value approach. *a*-value sensitivity analysis (121 PMF runs performed scanning all the COA and HOA *a*-value combinations, a-value scanning steps: 0.1).

IV) Classification of the 121 PMF runs based on the cluster analysis of the COA diurnal cycles. Selection of the best clusters, and corresponding PMF solutions.

V) PMF rotational ambiguity exploration. 100 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously varying the COA and HOA *a*-value combinations (using only the optimal *a*-value combinations identified from step IV). The average of the 100 bootstrap runs represented the online-AMS source apportionment average solution. The corresponding standard deviation represents the source apportionment uncertainty.

In a similar way we introduced a summary of the offline-PMF source apportionment optimization (P12 L13):

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) PMF rotational ambiguity exploration. 1080 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously varying the COA and HOA *a*-value combinations. PMF solutions were retained based on the correlation of the PMF factors with external tracers. The PMF solutions retrieved from this step are relative to the water-soluble fraction. The corresponding water-soluble OC factor concentrations were determined by dividing the water-soluble OM factor concentrations (PMF output) by the OM:OC ratio determined from the corresponding factor mass spectrum.

IV) Retained water-soluble OC PMF solutions from step (III) were rescaled to the total OC concentrations by applying factor recoveries. Factor recoveries were fitted (using *a*-priori information) to match total OC. Only PMF solutions and factor recoveries fitting OC with yearly and seasonally homogenous residuals were retained. The average of the retained PMF solutions represented the average source apportionment results. The corresponding standard deviation represented the source apportionment uncertainty.

4) The comparison between online and offline measurements.

I) It would be useful to include a table to summarize the sampling periods of online and offline measurements. Page 20 Line 4, it is very misleading to claim this study as the first comparison between HR online AMS and offline AMS, since the online and offline measurements are not simultaneous. Please rephrase.

We rephrased P20, L3-6 as:

In this study, we present one of the first OA source apportionment studies conducted over an entire year in the Mediterranean region. This work represents also the first comparison between HR online-AMS and HR offline-AMS source apportionments conducted at the same location, albeit in two different periods.

Table 1 was also added at P8, L5

Table 1. Monitoring periods.

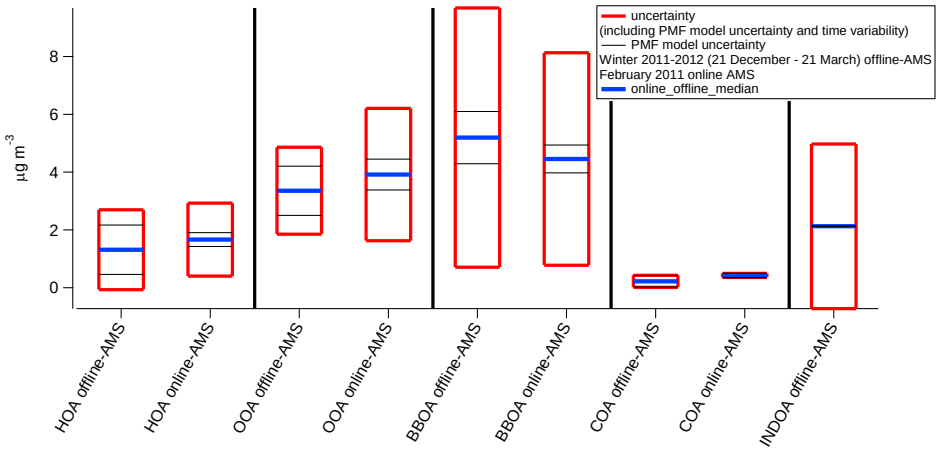
Online-AMS	Offline-AMS
28 January 2011 – 02 March 2011	30 July 2011 – 20 July 2012

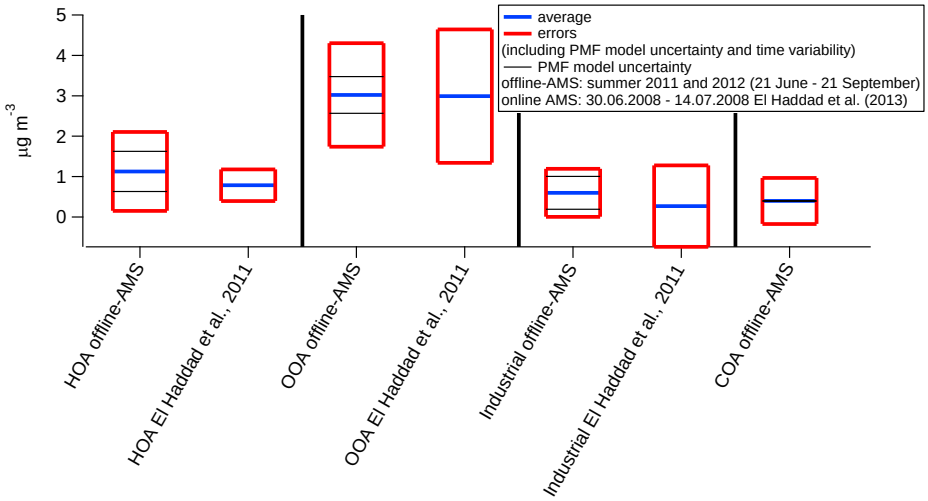
II) In Figures 7 and 8, how are winter and summer defined for offline AMS measurements? How many filters are included in winter and summer? The comparison in concentration looks generally good, but there are many

disagreements as well. For example, INDOA is resolved in offline dataset, but not in 2011 February online dataset. I suggest to include all factors in the figures, instead of only including the overlapped factors.

As suggested, we inserted the INDOA factor resolved by offline-AMS in Fig. 7 and we added the COA factor in Fig. 8.

In the offline-AMS source apportionment winter is defined from 21 December to 21 March. Summer is from 21 June to 21 September. These information are added in Fig. 7 and Fig. 8 legends. Both summer and winter are represented by 16 composite samples.





IV) Page 17 Line 24-25, HOA from offline analysis does not correlate with NOx. This is consistent with 2008 July online dataset (as noted in the main text), but not consistent with 2011 February online dataset (i.e., R = 0.86 in Page 16 Line 6). Please comment on this discrepancy. Btw, does OOA from offline analysis correlate with NH_4^+?

These observations are correct. During summer 2008, the correlation between NO_X and HOA is weak, although statistically significant, due to several reasons:

1/ In 2008, the authors have demonstrated that COA, which could not be separated by PMF, interferes with HOA. This interference influences the correlation between NO_X and HOA.

2/ We have observed in many datasets (including online ACSM measurements) a weaker correlation between NO_X and HOA concentrations, which we tentatively attributed to the rapid oxidation of NO_2 in summer.

3/ For the 2008 campaign, we have observed an increase in NO_X concentrations during industrial episodes, which worsen the correlation between HOA and NO_X.

Similar reasons apply for the offline-AMS.

By contrast during February 2011, we resolved a COA factor and no clear evidences of significant industrial contributions were observed. Therefore traffic could be well separated and can be considered as the main NOx source, which would explain the better correlation between NO_X and HOA.

OOA showed a low correlation with NH_4^+ (R = 0.3), suggesting that the OOA factor resolved in this work might originate from aging processes not following the seasonal formation of inorganic secondary components (NH_4NO_3 and $(NH_4)_2SO_4$). El Haddad et al. (2013) resolved by online-AMS PMF two different OOA factors for a summer monitoring campaign conducted at the same location: LVOOA and SVOOA. In that work only LVOOA correlated with SO_4^{2-}, but not SVOOA which had the same temporal behavior of terpene first generation products and therefore a probable biogenic origin. Also, ^{14}C measurements revealed that during summer the largest fraction of OOA had a non-fossil origin (El Haddad et al., 2013) indicating that biogenic emissions are expected to dominate OOA. A correlation between NH_4^+ and offline-AMS OOA would only be observed if the formation rates of OOA and ammonium sulfate or nitrate are very similar. In summer, biogenic emissions are expected to dominate OOA, but are associated with low emissions of SO_2 and NO_x. Therefore, for this study we do not expect a correlation between NH_4^+ and OOA.

IV) Were filter samples collected during 2011 February? If not, how are levoglucosan and vanillic acid measured for this period (Page 23 Line 23-24)? If yes, could the authors analyze them using offline-AMS and compare to simultaneous online AMS?

As also indicated in the manuscript (P7-8 L29-4) and in Table S1, filters were also collected during February 2011. This batch of filter samples was defined as "Batch B2", while the set of filters collected during the yearly cycle monitoring campaign was defined as "Batch 1". A subset of the same analyses conducted for the Batch 1 was carried out also on Batch 2 (see Table S1), however, no filter material remained for offline-AMS analysis.

Minor comment.

1. *Eqn. (11). In Bozzetti et al. (2017), there is a term TEOC (traffic emission OC) in the equation. Please discuss the rationale to replace TEOC with WSHOA/(RHOA*OM/OC) in current study. Also, WSOC is mentioned in Page 14 Line 17, but it is not clear how the WSOC is used in determining Rk.*

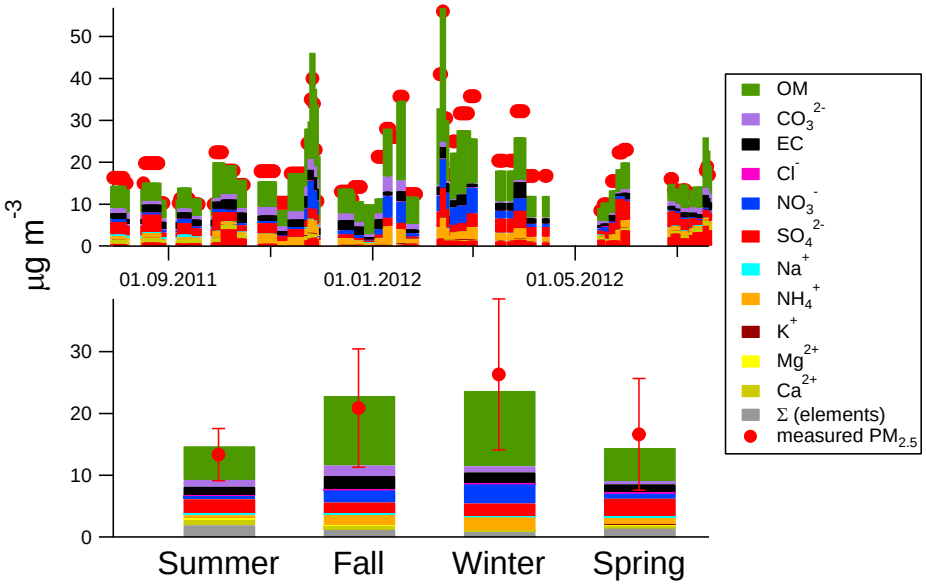
In Bozzetti et al. (2017), the traffic exhaust concentration was estimated using a chemical mass balance approach assuming hopanes measured by GC-MS to be a unique traffic tracer. This was necessary as constraining an HOA AMS profile returned a non-significant traffic exhaust correlation with typical traffic tracers such as hopanes and NOx. The traffic exhaust factor (TEOC) was not defined as HOA, because it was not identified from the AMS mass spectrum, but estimated from hopanes concentrations. In this work, as previously discussed, we resolved an HOA profile associated with a significant correlation with NOx (Criterion #2, P13, L5). The PMF factor was defined as "HOA", because it was resolved by constraining an Hydrocarbon-like organic aerosol AMS spectrum.

P14, L17. WSOC is indirectly contained within eq. 11. The sum of the PMF factors divided by their corresponding OM:OC ratios corresponds to WSOC (including also the PMF residuals within the sum). This is because the offline-AMS PMF input matrices were rescaled to WSOM = $(WSOC \cdot OM:OC)_i$ (P*10*, L1-3). Here $WSOC_i$ was determined by TOC analysis, and $OM:OC_i$ was determined from offline-AMS analysis. Since WSOC measurements might be affected by measurement biases we perturbed the WSKOC = $WSKOA/(OM:OC)_{WSKOA}$ PMF time series assuming a possible WSOC measurement bias of 5%.

For the sake of clarity we added (= $WSOA_{TOC} \cdot (OM:OC)_{\text{offline-AMS}i}$ at P10, L2, and we replaced at P14, L17 $WSKOA_i$, with $WSKOC_i/(OM:OC)_{WSKOC}$. We also clarified that the sum of $WSKOC_i/(OM:OC)_{WSKOC}$ for all PMF factors corresponds to $WSOC_i$.(P17, L20): (we note that the sum of the $WSKOC_i/(OM:OC)_{WSKOC}$ terms equals $WSOC_i$ neglecting the PMF residuals).

2. *Figure 1. Please show the detailed time series of stacked PM2.5 compositions and the measured total PM2.5 concentrations.*

   Following anonymous reviewer #1's suggestion we corrected figure 1



3. *Figure 10. How and why are the data smoothed?*

Ratios are smoothed using a running average weighted by the BBOC concentrations. The variability in this ratio for each data point (weighted standard deviation) is displayed as a range, in order not to hide any of the data. Non-smoothed ratios are very noisy because of low concentrations, which hide the general trend in the data and averaging was necessary to show the general trends in the ratios.

4. *The discussions on PAH sources are confusing. Based on the discussions in Page 16 Line 14-29, it is likely that PAH is mainly from biomass burning for 2011 February. However, in Page 18 Line 10-31, PAH is mainly from industry based on offline analysis on the whole year. Please clarify related discussions. Also, does AMS-PAH agree with filter PAH?*

The discussion at P18 L10-31 is not in contradiction with the section at P16 L14-19. In section P18 L10-31 we discuss that among the PAHs measured by GC-MS pyrene, fluoranthene and phenantrene are overwhelmingly emitted by industrial processes; however pyrene, fluoranthene and phenantrene represent only a minor fraction of the PAH total mass measured by GC-MS. In order to avoid misleading interpretation we inserted the following text at P18 L15.

We note that phenanthrene, pyrene, and fluoranthene together represent 9.6%$_{avg}$ of the PAHs mass quantified by GC-MS, indicating that PAHs are overwhelmingly emitted by BBOA.

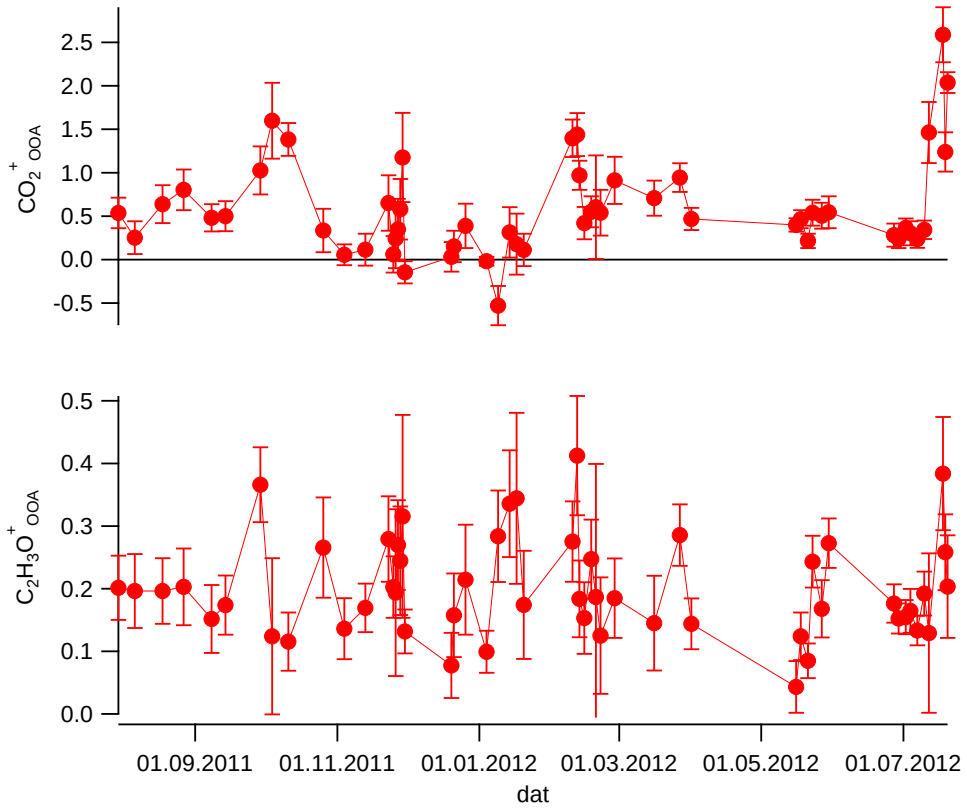
Comparison between AMS-PAHs and GC-MS PAH.
AMS-PAHs concentrations were found on average 19% higher than the sum of GC-MS quantified PAHs, and showed a significant correlation (R = 0.68). The concentration discrepancy can originate from different causes:

a) GC-MS quantified PAHs do not represent the total PAHs mass.
b) AMS-PAHs RIE could significantly differ from the average organic RIE (1.4) assumed in this work.
c) PAHs might be formed on the AMS vaporizer surface from the pyrolysis of refractory organic compounds.

This discussion was inserted in the SI.

5. *In Figure S18, why are there so many negative $CO_2^+{}_{OOA}$ values?*

Overall, only two points over 54 showed negative values within the estimated uncertainty. We added a "0 line" to Fig. S18 in order to guide the eye. These two points are related to the highest primary aerosol concentrations (due to the highest estimated BBOA concentrations), therefore they are affected by large uncertainties (possibly underestimated in our uncertainty model), because they were estimated as the difference between the measured CO_2^+ and the PMF modelled CO_2^+ from primary sources.



6. *In the abstract, please mention that 216 filter were collected over a full year, but only 58 filters were analyzed.*

P1, L24-25 corrected as follows:

In total 216 $PM_{2.5}$ (particulate matter with an aerodynamic diameter <2.5 μm) filter samples were collected over 1 year from August 2011 to July 2012. These filters were used to create 54 composite samples which were analyzed by offline-AMS.

7. *Page 9 Line 7. What is the difference in signal between measurement blank and a sample with relatively low loading?*

Considering the liquid extract associated with the lowest offline-AMS organic signal, we observed an average signal/blank ratio of 102.

8. *Page 14 Line 3, R_{COA} is not defined yet.*

R_{COA} represents the COA factor recovery. This information was added to the manuscript (P14, L3)

Anonymous Referee #2

Received and published: 28th February 2017

Bozzetti et al. describe year-long, offline AMS measurements of filters collected in Marseille. The authors perform source apportionment analysis to the filters to demonstrate changing contributions of BBOA, OOA, HOA, and INDOA. The authors compare this analysis to previous studies (e.g. El Haddad et al. 2013) and winter-time measurements conducted using a high resolution aerosol mass spectrometer. The authors find good agreement between the online and offline methods, and observe significant contributions from residential biomass burning during winter months. The authors provide additional analysis of the biomass burning factor and attribute changes in burning markers to differences in burning activities throughout the year. The authors also observe enhancements in methyl-nitrocatechol, which suggests secondary processing of the biomass burning emissions.

Overall, the paper is very well written, the methods are clear, and the interpretations of the data are reasonable. The PMF solutions, in particular, are incredibly detailed and thoroughly rationalized. The paper provides another example of the utility of off-line AMS analysis, which may serve as a useful low(er)-cost means for monitoring aerosol composition. While the study tends to confirm results previously observed in Marseille, it provides useful observations related to the seasonal changes in biomass burning markers. My biggest concerns relate to the over simplification of biomass burning sources, particularly to the assignment of periods described as lignin and cellulose burning. Upon addressing my comments, I recommend the manuscript for publication.

We thank Anonymous Referee #2 for his thorough review, which helped improving the overall quality of our work. We have inserted a more careful discussion of the complexity of BBOA emissions following the suggestions in the comments below.

Major comments

> *Page 25, Lines 3 – 10: I'm persuaded by the argument that differences in biomass burning markers could be related to changing fuel types; however, I would recommend that the authors refrain from suggesting that the differences are strictly related to cellulose vs. lignin combustion. This also pertains to Fig. 11, which highlights periods of "lignin-combustion" and "cellulose-combustion." In reality, agricultural waste burning, open burning, prescribed burning, etc is the combustion of mixtures of lignin and cellulose-rich fuels; therefore, attributing changes in tracers to one plant structure or another downplays the complexity of biomass burning. I recommend the authors reframe the discussion to focus on changes in human activity,*

i.e. periods of increased prescribed burning, periods of increased residential heating, etc. Within that discussion, the authors may describe the differences in fuel composition, keeping in mind that mixed fuels (as well as burning conditions, fuel moisture content, etc) will contribute to the variability of biomass burning tracers.

While we do agree with the reviewer that the two periods would relate to open agricultural burning and residential heating, referring to these periods as such would imply that we know with certainty the patterns of biomass combustion in the region. However this is not the case and in France domestic green waste burning, with the exception of agricultural burning, is prohibited. Aerosol from agricultural burning was previously modelled for southern France (Dernier van der Gon et al., 2015, Fountoukis et al., 2014) and we suspect that this practice occurs in this region, without knowing the extent of it or the contribution of prohibited domestic green waste combustions. Our results indicate the importance of green waste combustion and provide means for tracing it. By referring to the periods in the manuscript as cellulose rich and lignin rich biomass combustion, we did not intend to oversimplify the biomass burning processes, and we do agree that all biomass contain both cellulose and lignin. Instead, we have carefully chosen these designations as they reflect most our observations, increase of lignin pyrolysis products over cellulose pyrolysis products during coldest days. Based on these observations we could only speculate that these differences are due to a change in the biomass burning pattern. Therefore, we prefer keeping the same designations purely based on the observations and not to use the nomenclature proposed by the reviewer which would be based on the further interpretation of the data. Nevertheless, we added additional explanation for using these terms at P25, L3-10.

In this study we related the evolution of the BB composition over the cold season to the combustion of cellulose-rich and lignin-rich fuels, considering that lignin and cellulose are contained in different ratios in different biomass fuels. This designation should not be considered as an oversimplification of the combustion processes or of the fuel complexity, but rather as a classification of the BB aerosol based on our observations of increasing lignin pyrolysis products over cellulose pyrolysis products during the coldest days.

The Fig. 11 legend was modified as follows:

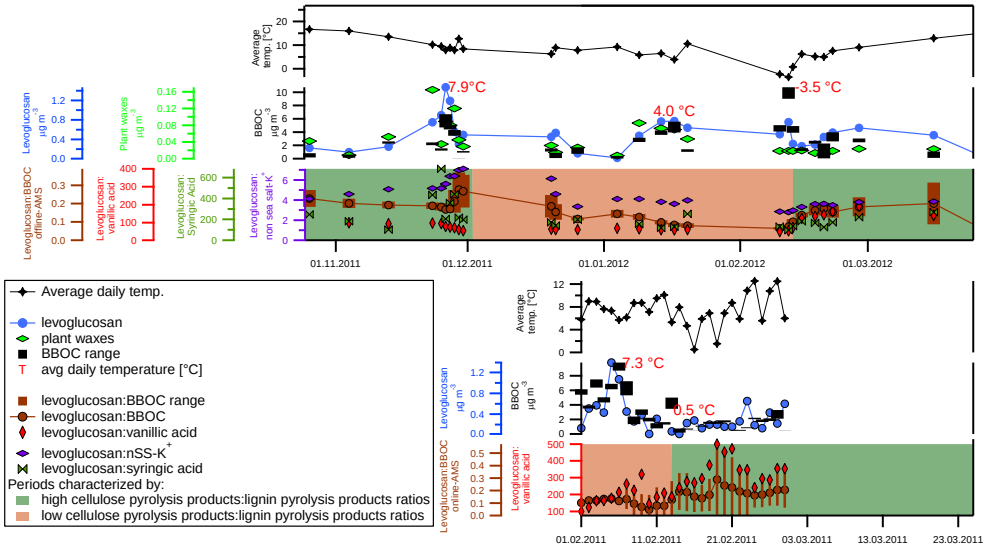
Periods characterized by:

-high cellulose pyrolysis products:lignin pyrolysis products ratios

-low cellulose pyrolysis products:lignin pyrolysis products ratios

Page 24, Lines 20-32: The authors discuss plant waxes, but do not provide any figures or correlations with other BBOA tracers. Are these supposed to be incorporated into Fig 11 (suggested at page 25, line 1)? I believe these traces need to be shown in the figure if they are to be discussed and attributed to different fuels

Following anonymous reviewer #2's suggestion, we added the plant waxes trace in Fig. 11 and modified the figure to make it clearer.



The following text was added to Fig. 11 caption:

The plant waxes concentrations were determined from GC-MS measurements of alkanes with an odd number of carbons (Li et al., 2010). As discussed in the main text the spike observed in late autumn could be related to incomplete green waste combustion.

Page 10, line 4: The source apportionment performed on AMS data is very detailed; however, at times, I had difficulty following the progression due to the amount of detail provided. This discussion should be included in the main text, however I think it would be useful if the authors could provide a brief listing of the steps taken to perform this analysis at the beginning of this section. For example, one could add "… In order to optimize the source separation, we performed sensitivity analyses on PMF solutions by (1) selecting number of factors, (2) constraining HOA and COA, (3) cluster anlaysis, (4)…. ". As the reader continues reading your method, they could reference this listing and follow the logical progression more easily.

Following the suggestion of anonymous reviewers #1 and #2 we introduced in the main text a summary of the online-AMS optimization procedure (P10 L11).

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) Constrain of both the HOA and COA factors profiles adopting an *a*-value approach. *a*-value sensitivity analysis (121 PMF runs performed scanning all the COA and HOA *a*-value combinations, a-value scanning steps: 0.1).

IV) Classification of the 121 PMF runs based on the cluster analysis of the COA diurnal cycles. Selection of the best clusters, and corresponding PMF solutions. Best clusters were selected based on the analyses of the average cluster spectra of COA, HOA and BBOA. In other words, clusters characterized by average factor mass spectra (and corresponding clusters) not statistically different from the corresponding average literature profiles were retained.

V) PMF rotational ambiguity exploration. 100 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously varying the COA and HOA *a*-value combinations (using only the optimal *a*-value combinations identified from step IV). The average of the 100 bootstrap runs represented the online-AMS source apportionment average solution. The corresponding standard deviation represents the source apportionment uncertainty.

In a similar way we introduced a summary of the offline-PMF source apportionment optimization (P12 L13):

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) PMF rotational ambiguity exploration. 1080 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously varying the COA and HOA *a*-value combinations. PMF solutions were retained based on the correlation of the PMF factors with external tracers. The PMF solutions retrieved from this step are relative to the water-soluble fraction.

IV) Retained water-soluble PMF solutions from step (III) were rescaled to the total OM concentrations by applying factor recoveries. Factor recoveries were fitted (using *a*-priori information) to match total OC. Only PMF solutions and factor recoveries fitting OC with yearly and seasonally homogenous residuals were retained. The average of the retained PMF solutions represented the average source apportionment results. The corresponding standard deviation represented the source apportionment uncertainty.

Page 10, lines 24-26: For readers who may be unfamiliar with the a-value sensitivity analysis, it would be useful to explain here why one might apply this analysis to HOA and COA components. The authors mention that lack of acceptable tracers for COA emissions (line 31), but it may also be helpful to discuss that other studies have observed improved resolution of HOA after constraining these factors (e.g. Canonaco et al. (2013)), or that these two factors may exhibit similarities in the mass spectrum and/or diurnal profile, as demonstrated in the cluster analysis described in the SI.

We thank reviewer #2 for the suggestion. Since a similar argument was raised also by anonymous reviewer #1, we introduced a paragraph elucidating the improvements of the constrained PMF solution in comparison to the unconstrained approach (see 3rd comment, subsection I).

> *Page 11, Lines 24 – 28 and Page SI 13, Lines 18-22: My understanding from reading this section is that the authors rejected clusters 4 and 5 primarily based on mass spectrum similarities with reference spectra, or by similarities with other factors (e.g. COA with HOA). From my untrained eye, it also appears that cluster 3 exhibits a strong correlation with cluster 4 (Fig S7 and Table S3, R = 0.93). Similarly, the correlation between Cluster 3 and NOx (R=0.57) is not substantially different from that of Cluster 4 and NOx (0.64). Would this also be grounds to reject cluster 3? Or, are the authors placing more weight on similarities mass spectra as opposed to similarities in temporal profiles? Personally, I believe similarities in mass spectra is a more important criterion, but other readers may disagree.*

As anonymous reviewer #2 correctly mentioned, the cluster selection was based on mass spectral similarities. In absence of a reliable COA tracer, retaining a cluster based on the COA diurnal cycles was more subjective than accepting/rejecting a cluster based on the resemblance of the mass spectra with literature profiles. For these reasons we opted for the selection of the clusters based on mass spectral analyses. Indeed the average COA diurnal of cluster 3 and 4 are quite well correlated, however the COA diurnal cycle for cluster #3 showed lower background values at night and a more pronounced peak at noon, while for cluster #4 this peak was almost not visible. Overall, after applying the last selection step (Fig. S8), almost no solution from cluster #3 was retained, i.e. most of the solutions associated to cluster #3, were attributed to cluster #4 for more than 5% of the *k*-means initiations, confirming the resemblance between the two clusters.

> *Page 14, lines 10-12: The authors indicate that a 5th factor was resolved by source apportionment of the offline measurements, but not by online measurements. The authors note later in the manuscript (page 22, lines 4-5) that this was previously discussed, however I can't find this discussion in the PMF description. Please clarify.*

The authors referred to section 3.1: P16, L18-29. This reference was added for clarity (P22 L4).

> *Page 20, Lines 3-4: Since online and offline AMS measurements were not conducted simultaneously, I don't agree that you can make a direct comparison. Please revise.*

We rephrased P20 L 3-6 as follows:

In this study, we present one of the first OA source apportionments conducted over an entire year in the Mediterranean region. This work represents also the first comparison between HR

online-AMS and HR offline-AMS source apportionments conducted at the same location, despite in two different periods. Previous studies (Daellenbach et al., 2016) reported a comparison between offline-AMS and online-ACSM results.

Page 23, Lines 1-14. I'm confused about what message the authors are trying to convey with this discussion. Are the authors trying to attribute nitrocatechol formation to a chemical process, or is the focus to show that offline measurements can't capture the chemical evolution of these tracers due to their high reactivity and the low time resolution of offline analysis?

The discussion at P23, L1-14 aims at demonstrating that rapidly formed secondary aerosol compounds (such as nitrocatechols) are likely apportioned by PMF to primary factors rather than to secondary or OOA factors. Such fast SOA formation is not captured by offline-AMS, but it is also probably not captured by online-AMS, unless in the proximity of the source or directly within the plume. In order to better convey the message we added the following text at the end of the discussion (P23, L14).

Overall these findings suggest that rapid SOA formation is not well captured by PMF and rapidly formed SOA compounds (such as nitrocatechols) can be systematically attributed by PMF to factors commonly considered as "primary" (BBOA in this case).

Page 24, Lines 7 – 17: The authors mention that the levoglucosan:nss-K ratio was 3.35 in winter at line 8, but then describe a minimum ratio in Jan/Feb of 6.3. I'm assuming this is a mistake, since I observe a minimum of ~3-4 from Fig. 11.

We thank anonymous reviewer #2 for reporting this mistake, we corrected the values as follows:

PL24 L7-8: showed lower average values in summer (0.23) than in winter (3.14).

P24 L 12-13: shows higher values during March and late autumn (up to 7.11) and lower in January, February (minimum = 2.79; Fig. 11)

Figure 6: The x-axis is very difficult to read. The authors could remove the year from the dates, or average the collection interval to present a single value rather than a range.

In order to better display the x-axis we displayed sticks for individual filters from each composite, instead of one stick per composite.

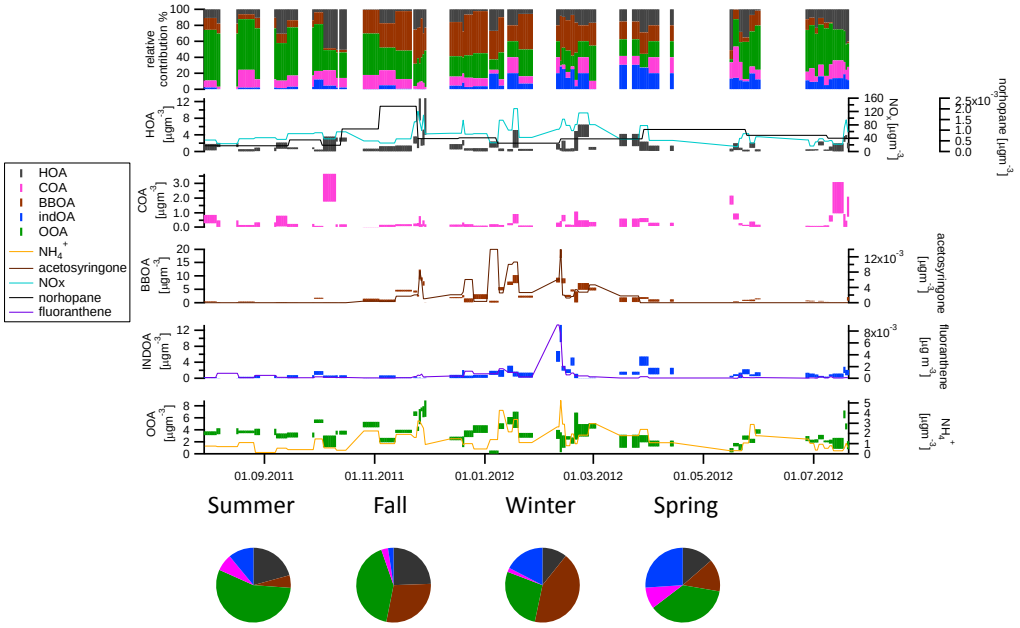


Figure 10: I find this figure to be misleading. The authors note that the reader should only consider the monthly changes and not the day-to-day behavior since these measurements were not performed simultaneously; however, as a reader, my first intuitive response is to believe these measurements were conducted at the same time. Only after reading and interpreting the caption do I understand what the authors are conveying. In my opinion, the temporal profiles should be placed on separate axes or presented differently.

Following the suggestions of anonymous reviewer #2 we display the two traces on different time axes (bottom and top). In order to better identify the tracers, we used different markers and different colors.

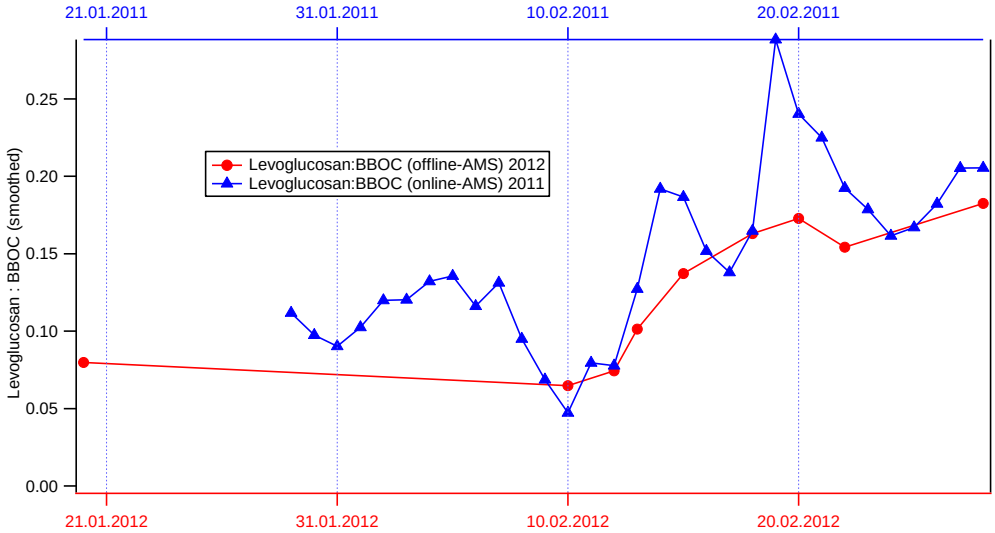


Figure 11: I find Figure 11 to be very difficult to read. The marker sizes are quite small, similar in shape (circles), and displayed on top of similarly colored, easter-egg-like backgrounds (lignin combustion period, cellulose-combustion period). Personally, I think the figure is too busy and I struggled to immediately grasp its message. To simplify the figure, the authors could remove the measurements from 2011 and plot them separately in the supplemental information since these appear to be auxiliary evidence.

Following the suggestion of anonymous reviewer #2, and #3 we corrected Fig. 11 as follows:

- Background colors were changed
- Markers were changed
- Other modifications suggested by other reviewers were introduced

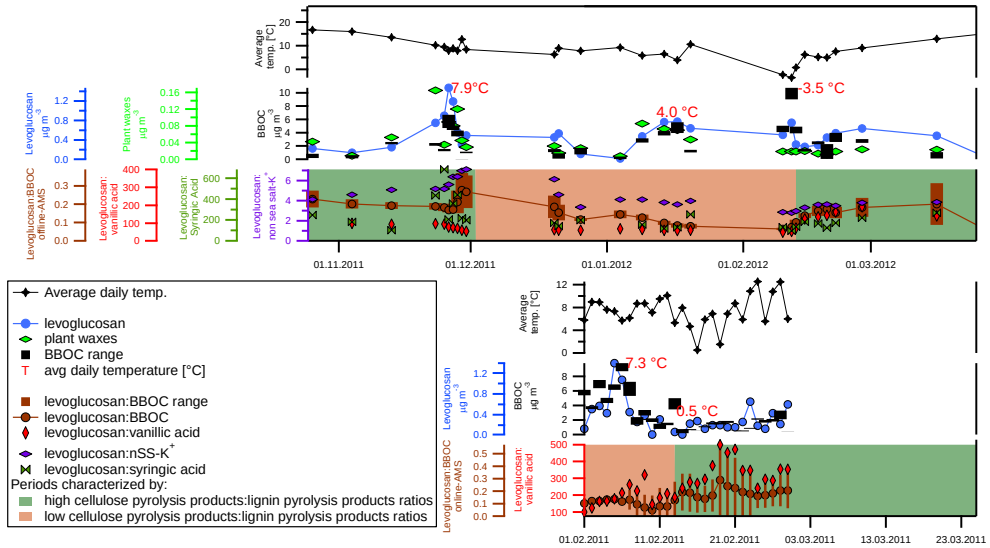
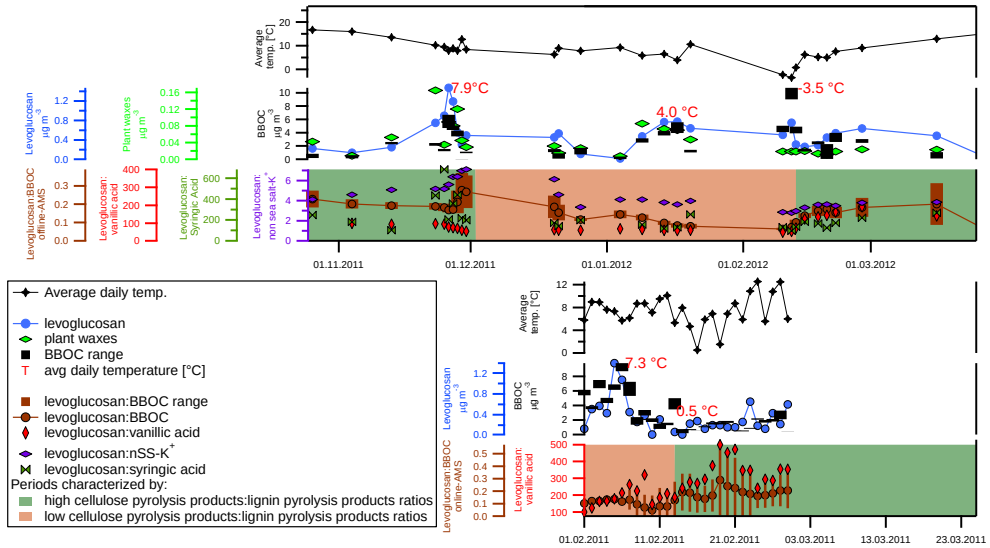


Figure 11: I'm confused how the authors derive the levoglucosan:vanillic acid ratio from the 2011 measurements. Are the authors using AMS tracers, or were filters collected? I'm assuming these represent AMS tracers since the scale is drastically different than that for 2012.

We thank anonymous reviewer #2 for finding this issue, Fig. 11 was corrected as follows (see below). As also indicated in the manuscript (P7-8 L29-4) and in Table S1, filters were also collected during February 2011. This batch of filter samples was defined as "Batch B2", while the set of filters collected during the yearly cycle monitoring campaign was defined as "Batch 1". A subset of the same analyses conducted for the Batch 1 was carried out also on Batch 2 (see Table S1). Both levoglucosan and vanillic acid were measured by chemical derivatization GC-MS (El Haddad et al., 2009; Favez et al., 2010)



Minor Comments

Page 5, lines 6 -7: How frequently were the filters extracted from the sampler and stored? Weekly? Daily? I'm curious how long each sample was "aged" in the sampler prior to freezer storage.

Filters were collected from the sampler every Monday, Wednesday and Friday and immediately freezer stored.

Page 11, Line 2: I do not see a reference to Elser et al. 2015 in the bibliography.

We thank anonymous reviewer #2 for finding this error. The reference in the main text was corrected to Elser et al., 2016.

Page 14, Line 3: Do the authors refer to the Pearson's R for COA, or its factor recovery?

R_{COA} referred to the COA factor recovery, this was clarified in the main text as also suggested by anonymous reviewer #1.

Page14, lines 5-7: I believe the authors mixed up KOA and WSKOA?

Corrected as suggested.

Page 18, Lines 6-7, Page 23, Line 24, Fig. 11: At times, I'm confused as to whether the authors are referring to levoglucosan and vanillic acid measured by GC-MS or AMS. Can the authors please specify?

According to anonymous reviewer #2 we clarified in the text that levoglucosan and vanillic acid measurements are from GC-MS.

Editorial Comments

Page 20, Line 3: Add "of" between "one" and "the"

Corrected as suggested

Anonymous Referee #3

Received and published: 1st March 2017

General Comments:

The authors demonstrate the use of positive matrix factorization (PMF) to the watersoluble, offline AMS spectra to reveal the contribution of the different organic aerosol (OA) sources (hydrocarbon-like OA (HOA), cooking OA (COA), biomass burning OA (BBOA), oxygenated OA (OOA), and an industry-related OA (INDOA) in Marseille, France. They also make comparison between online AMS and offline AMS source apportionment to further show the application of offline AMS measurements for OA source analysis. The paper is very well written, the experimental approach and the data analysis are very clear. I only have one question about the definition of watersoluble, offline AMS spectra.

The authors have a detailed description of how they extract the filters in the experimental section, page 5, line 11" One punch per filter 12 sample (from 5 to 25 mm diameter depending on the filter loading and on the number of 13 punches per composite sample) was prepared for analysis. Punches from the same composite 14 sample were extracted together in 15 mL of ultrapure water (18.2 Mï ¸ A°U cm, total organic 15 carbon < 5ppb, 25_C) in an ultrasonic bath for 20 min at 30_C. After extraction, filters were 16 vortexed for 1 min, and the resulting liquids were filtered with 0.45 ïA¸ m nylon membrane 17 syringe filters."

My questions are: How do we define water soluble AMS spectra? Will the water soluble AMS spectra strongly depend on the filer extraction method (e.g, the volume of water and the temperature used for extraction, and sonication time)? How would these factors affect the composition of aqueous extracts and the water soluble AMS spectra?

Without applying the same filter extraction approach, how could we compare the water soluble AMS spectra and source apportionment analysis in different studies? There is a possibility that the filters could be extracted in different ways in different studies.

Since the offline AMS measurements could be a very useful tool for OA source apportionment, as demonstrated in this work, the authors further elaborate and address these issues in the manuscript.

We thank Anonymous Referee #3 for the review and inputs.

Water soluble AMS spectra are defined as the mass spectra collected from the AMS analysis of the nebulized aqueous filter extracts. Indeed extraction conditions can affect the water soluble AMS spectra. Bozzetti et al. (2017) reported the comparison of water-soluble AMS spectra collected from the atomization of filter extracts using two different nebulizers. Results showed that the collected spectra were not different within the measurements repeatability; however more comparisons (between AMS spectra collected for filter extracted under different conditions) are required.

More importantly, the water extraction conditions may indeed affect the PMF factor water-solubility; therefore factor recoveries relative to filters extracted in different conditions might be different, and therefore should be re-determined from a comparison between the water-soluble OA source apportionment and a well-established OA source apportionment method, as in Daellenbach et al. (2016). Bozzetti et al. (2017) stated that in absence of a well-established OA source apportionment method to be adopted as a reference, the factor recoveries from Daellenbach et al. (2016) can be assumed as a first guess, but their applicability needs to be verified. Overall we do not expect a large factor recovery sensitivity to the water extraction conditions. This is confirmed by the results of the solubility analysis we conducted (see answer to comment 2 from anonymous reviewer #1). Results indicate the factor recoveries to vary by 10% on average when changing the water extraction conditions, where this bias is well within our factor recovery uncertainties. Not surprisingly, the factor recoveries estimated for offline-AMS (Daellenbach et al., 2016; Bozzetti et al. 2016; Bozzetti et al., 2017) and from PILS-AMS (Xu et al., 2017) are in good agreement.

Anonymous Referee #4

Received and published: 8th March 2017

General Comments:

This paper presents results from an offline analysis of aerosol samples collected in Marseille France using HR-AMS for the organic and other offline techniques for the inorganic, elemental carbon, and metals analyses. Using PMF, source apportionment was carried out and the results from the offline study were compared to PMF analysis of a smaller sample set of on-line AMS data. While the two time periods did not overlap between the offline and online data sets, good comparisons were observed for the different organic factors between offline and online AMS. The authors found increased BBOA in the winter and that OOA was dominant in the summer. The authors also observed changes in the levoglucosan:BBOC ratio over the course of the winter-spring months that was opposite what would be expected if it was driven only by changes in photochemical processing from the winter to the early spring. They attribute the observed change in the BBOC to changing sources with more domestic wood burning in the cold winter months and more agricultural biomass burning in the early spring. Overall, the analysis is thorough and the conclusions are explained well and match previous studies in similar areas. I recommend publishing the manuscript after some minor revisions.

We thank anonymous referee #4 for the review and the comments.

Specific Comments:

1. *A blank was run between each sample with ultra-pure water to monitor memory effects. However, the minimum concentration needed to make aerosols that will dry down to a size observable in the AMS is relatively large. Thus, some material may be staying behind in the atomizer that would not be observed. Were any blanks with a clean salt solution ran to test for this carry-over between runs? Were the samples analyzed in any particular order or randomly?*

Filters were measured without following a specific order. As anonymous reviewer #4 suggested, nebulizing solutions of inorganic salts would provide a more representative blank. However, we have previously showed that the organic blank measurements collected by ultrapure water nebulization provide a comparable blank estimate to the organic blanks determined from the nebulization of $(NH_4)_2SO_4$ (Bozzetti et al., 2017).

2. *Page 16, line 18 "The AMS-PAHs:HOA ratio was : : :" the text gets a bit confusing here. I think the same point is being made and supported with different pieces of data*

but I recommend breaking the text up into a few paragraphs to make the discussion clearer.

According to anonymous reviewer #4 we subdivided the section into paragraphs as follows.

The AMS-PAHs:HOA ratio was 0.0020, while the AMS-PAHs:BBOA was 0.0028.

In general, industrial emissions can be an important source of PAHs at this location as discussed in El Haddad et al. (2013). In presence of an industrial contribution, the BBOA *vs*. AMS-PAHs correlation would decrease. In this work the correlation between AMS-PAHs and the $C_2H_4O_2^+$ fragment, typically related to levoglucosan fragmentation (Alfarra et al., 2007), was high (R=0.87) and no AMS-PAHs spike was observed without a simultaneous increase of $C_2H_4O_2^+$ (Fig. S14). Moreover the industrial-related OA factor resolved by El Haddad et al. (2013) was clearly associated to wind directions from W/SW (225°-270°), while in this work wind directions were oriented from W/SW only for 7% of the monitoring time, furthermore without being associated to any significant increase in the AMS-PAHs concentration (Fig. S15), indicating the absence of clear industrial episodes.

The BBOA diurnal cycle, similarly to AMS-PAHs, showed higher values at night than during the day (Fig. 4). In addition, the BBOA highest concentrations were detected at night and associated to slow wind speeds from the E/NE which is consistent with the night land breeze direction. Moreover, strong enhancements of the BBOA factor concentrations were perceived when the wind direction shifted to the E/NE (typically around 18 o'clock during the monitoring period) suggesting that BBOA could be transported from the valleys nearby Marseille (Fig. S17).

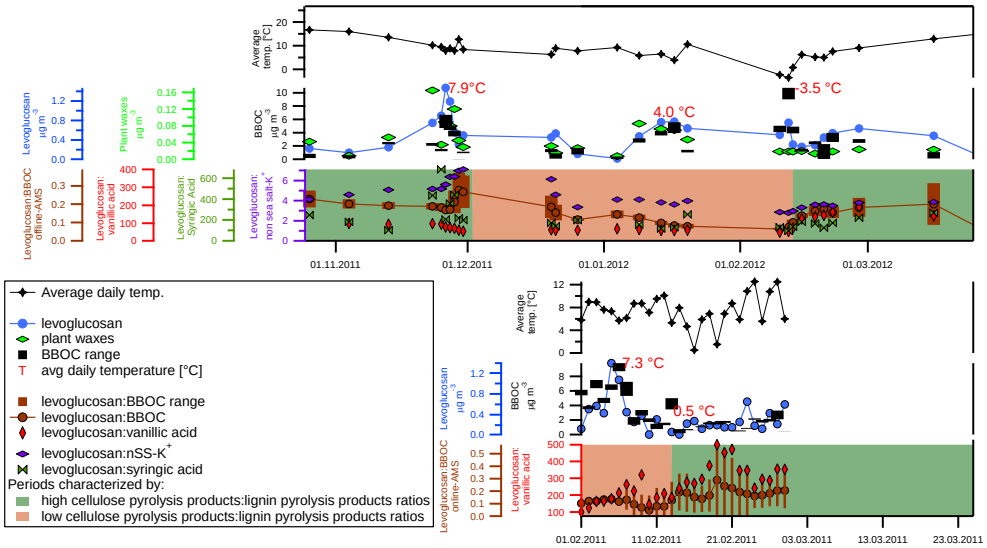
We calculated the BBOC time series by dividing the BBOA concentrations by the $OM:OC_{BBOA}$ ratio calculated from the average BBOA HR spectrum (1.60). The average BBOC:levoglucosan ratio [$\mu g\ m^{-3}/\mu g\ m^{-3}$] was 0.15, comparable to other European studies (Zotter et al., 2014; Herich et al., 2014; Minguillón et al., 2011).

3. *The analysis presented here used IC to measure the salts in the samples but the AMS also quantifies some of those ions. How do the measurements between the two techniques compare?*

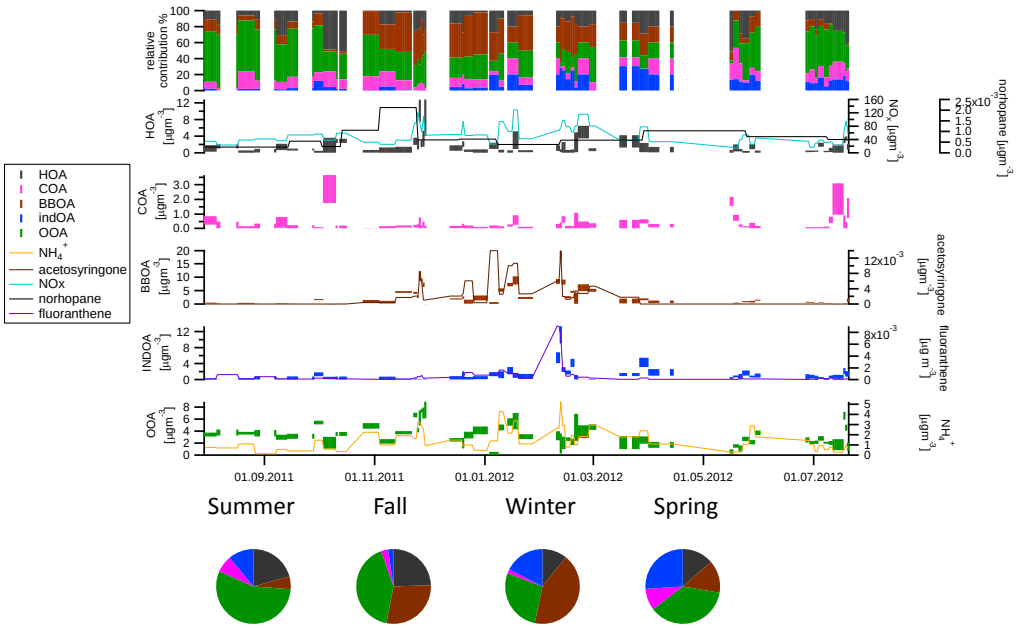
Since the same question was raised by anonymous reviewer #1, we kindly refer anonymous reviewer #4 to the answer to the major comment I) by anonymous reviewer #1.

4. *Figure 11 is very hard to read, both the colors used and the sizes of the markers should be adjusted.*

Following the suggestion of anonymous reviewer #4 we changed the background colors and the marker sizes.



1



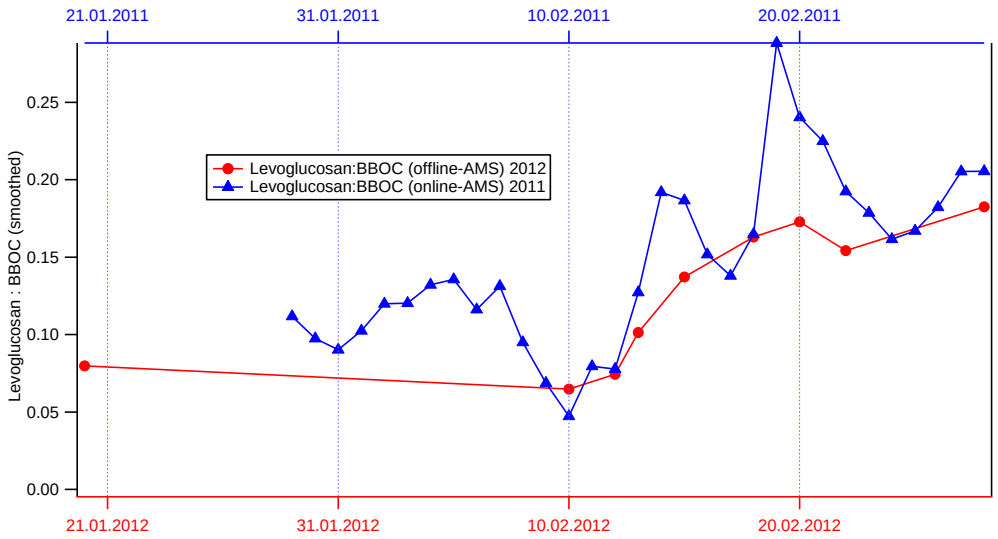
2
3
4 5. *Figure 6, change the black to grey in the pie chart or clarify in the caption.*
5
6 Following anonymous reviewer #4 suggestion, we changed the color of HOA in the pie plots
7 to dark grey.
8
9

6. Figure 6, the time series for BBOA is shown with acetosyringone but no mention of this is made in the text.

Acetosyringone is listed in Table S1. We added the acetosyringone correlation with BBOA at P18 L6-7.

7. Figure 10 shows dates on the x-axis for 2012 but 2011 online AMS data is overlaid using the same axis. I suggest changing the axis labels to 10.02.20xx or just the month and day to highlight this difference in collection times.

Fig. 10 was modified according to the suggestions of anonymous reviewer #2. The two traces were displayed on two different axes and labelled with different colors.



References (not already included in the main text):

Cappelli, C. I., Manganelli, S., Lombardo, A., Gissi, A., and Benfenati, E.: Validation of quantitative structure–activity relationship models to predict water-solubility of organic compounds, Sci Tot. Environ., 463-464, 781-789, 2013.

Meylan W. M., and Howard, P. H.: Upgrade of PCGEMS water solubility estimation method (draft). U.S. Environmental Protection Agency, Office of pollution Prevention and Toxics, Washinghton DC, 1994a.

Meylan W. M., and Howard, P. H.: Validation of water solubility estimation methods using K_{OW} for application in PCGEMS & EPI. Final report. U.S. Environmental Protection Agency, Office of pollution Prevention and Toxics, Washinghton DC, 1994b.

Meylan W. M., Howard, P. H., and Boethling, R. S.: Improved method for estimating water solubility from octanol/water partition coefficient. Environ. Toxicol. Chem. 15, 100–106, doi:10.1002/etc.5620150205, 1996.

Xu, L., Guo, H., Weber, R. J., and Ng, N. L.: Chemical Characterization of Water-Soluble Organic Aerosol in Contrasting Rural and Urban Environments in the Southeastern United States, Environ Sci Technol, 51, 78-88, 2017.

List of all relevant changes made in the manuscript

In the following page numbers and lines refer to the originally submitted version of the manuscript.

1) A discussion about the choice of the reference spectra was added in the main text at P10 L23 and P12 L21
2) A comparison between constrained and unconstrained PMF results was inserted at P10 L23
3) Summaries of online- and offline-AMS PMF optimization procedures were added at P10 L11, and at P12 L13 respectively.
4) We inserted Table 1 at P8 L5 to summarize the time coverage of the online- and offline-AMS source apportionments
5) Figs. 1, 6, 7, 8, 10 and 11 were modified following reviewers' suggestions.
6) We inserted a paragraph at P25 L3-10 in order to clarify the use terms cellulose-rich and lignin-rich BBOA.

Organic aerosol source apportionment by offline-AMS over a full year in Marseille

Carlo Bozzetti[1], Imad El Haddad[1], Dalia Salameh[2,7], Kaspar Rudolf Daellenbach[1], Paola Fermo[3], Raquel Gonzalez[3], Maria Cruz Minguillón[4], Yoshiteru Iinuma[5], Laurent Poulain[5], Emanuel Müller[6], Jay Gates Slowik[1], Jean-Luc Jaffrezo[7], Urs Baltensperger[1], Nicolas Marchand[2], and André Stephan Henry Prévôt[1]

[1]{Paul Scherrer Institute (PSI), 5232 Villigen-PSI, Switzerland }

[2]{Aix Marseille Univ., CNRS, LCE, 13331 Marseille, France}

[3]{Università degli Studi di Milano, 20133 Milano, Italy}

[4]{Institute of Environmental Assessment and Water Research (IDAEA), CSIC, 08034 Barcelona, Spain}

[5]{Leibniz Institute für Troposphärenforschung, 04318 Leipzig, Germany}

[6]{Eawag, 8600 Dübendorf, Switzerland}

[7]{Université Grenoble Alpes, CNRS, LGGE, 38000 Grenoble, France}

Correspondence to: A. S. H. Prévôt (andre.prevot@psi.ch)

N. Marchand (nicolas.marchand@univ-amu.fr)

Abstract

We investigated the seasonal trends of OA sources affecting the air quality of Marseille (France) which is the largest harbor of the Mediterranean Sea. This was achieved by measurements of nebulized filter extracts using an aerosol mass spectrometer (offline-AMS). In total 216 $PM_{2.5}$ (particulate matter with an aerodynamic diameter <2.5 μm) filter samples were collected over 1 year from August 2011 to July 2012. These filters were used to create 54 composite samples which were analyzed by offline-AMS. The same samples were also analyzed for major water-soluble ions, metals, elemental and organic carbon (EC/OC), and organic markers, including n-alkanes, hopanes, polyaromatic hydrocarbons (PAHs), lignin and cellulose pyrolysis products and nitrocatechols. The application of positive matrix

Formatted: Font: (Default) Times New Roman, 11 pt, English (U.S.)

Formatted: Font: (Default) Times New Roman, English (U.S.)

Formatted: Font: (Default) Times New Roman

factorization (PMF) to the water-soluble AMS spectra enabled the extraction of five factors, related to hydrocarbon-like OA (HOA), cooking OA (COA), biomass burning OA (BBOA), oxygenated OA (OOA), and an industry-related OA (INDOA). Seasonal trends and relative contributions of OA sources were compared with the source apportionment of OA spectra collected from the AMS field deployment at the same station but in different years and for shorter monitoring periods (February 2011 and July 2008). Online- and offline-AMS source apportionment revealed comparable seasonal contribution of the different OA sources. Results revealed that BBOA was the dominant source during winter representing on average 48% of the OA, while during summer the main OA component was OOA (63% of OA mass on average). HOA related to traffic emissions contributed on a yearly average 17% to the OA mass, while COA was a minor source contributing 4%. The contribution of INDOA was enhanced during winter (17% during winter and 11% during summer), consistent with an increased contribution from light alkanes, light PAHs (fluoranthene, pyrene, phenantrene) and selenium, which is commonly considered as an unique coal combustion and coke production marker. Online- and offline-AMS source apportionments revealed evolving levoglucosan:BBOA ratios, being higher during late autumn and March. A similar seasonality was observed in the ratios of cellulose combustion markers to lignin combustion markers, highlighting the contribution from cellulose rich biomass combustion, possibly related to agricultural activities.

1 Introduction

Outdoor particulate air pollution is estimated to be responsible for approximately 3.3 million premature deaths each year worldwide, and this number is projected to double by 2050 (Lelieveld et al., 2015). Organic aerosols (OA) can contribute up to 90% of the PM_1 (Jimenez et al., 2009), therefore understanding their main emission sources and formation processes is a key prerequisite for the development of appropriate mitigation policies.

In the Mediterranean basin, sources and trends of OA remain scarcely investigated, despite their deleterious impact in such a densely populated region. The Mediterranean region is characterized by an intense photochemistry during summer. Not surprisingly, the majority of the OA source apportionment studies conducted in the region using aerosol mass spectrometry (AMS) focused on the summer period (e.g., El Haddad et al., 2013; Minguillón et al., 2011, 2016; Hildebrandt et al., 2011). Through positive matrix factorization (PMF)

techniques, these studies revealed that during summer the oxygenated organic aerosol (OOA) fraction formed by oxidation of gaseous precursors, represented the largest part of OA. Amongst these studies, the field deployment of the AMS in Marseille, the largest port in the Mediterranean, has demonstrated that this instrument is well suited for quantifying the contribution of industrial emissions (El Haddad et al., 2013). In that work, the industrial OA factor was identified by the high correlation with heavy metals and AMS-polycyclic aromatic hydrocarbons (AMS-PAHs), moreover strong increments of the industrial factor concentrations were systematically observed when winds shifted to the west/south west, consistent with back-trajectory analysis highlighting the transport of industrial emissions from an industrial pole. Overall the industry-related OA contributed on average 7% of the bulk OA mass (El Haddad et al., 2011; 2013). However, these results were limited to 2 weeks of measurements during summer while the contribution of industrial emissions during the rest of the year remains unknown.

There is a general paucity of AMS and ACSM datasets in the Mediterranean region during winter. Exceptions include AMS campaigns (Mohr et al. 2012; Hildebrandt et al., 2011) covering a few weeks during late winter-early spring, and studies with an aerosol chemical speciation monitor (ACSM) (e.g., Minguillón et al., 2015, covering three weeks of monitoring). The measurement of organic markers and elements (e.g., Salameh et al., 2015; Reche et al., 2012) at different stations indicate a substantial contribution from biomass burning (BB). However, the sources and chemical composition of this fraction and its evolution during the year remain uncertain. Modelling results within the European Monitoring and Evaluation Programme (EMEP) have shown that the South of France, together with Portugal, can be a major hot spot in Europe for OA during February-March, possibly due to agricultural fires (Dernier van der Gon et al., 2015; Fountoukis et al., 2014). In this region, BBOA can derive from various processes such as agricultural land clearing activities, wildfires, and domestic heating, and therefore may have a variable chemical composition.

The current study capitalizes on the AMS measurements of offline samples collected over one year (2011-2012), in Marseille, an ideal environment for the characterization of urban emissions from biomass burning, traffic and industrial activities and their transformation under high photochemical activity. The source apportionment results obtained from PMF applied to the OA mass spectra are corroborated using a comprehensive set of offline measurements including elemental and organic carbon (EC/OC) measurements, as well as

measurements of elements by inductively coupled plasma mass spectrometry (ICP-MS), of molecular markers by gas chromatography mass spectrometry (GC-MS) and ultra-performance liquid chromatography mass spectrometry (UPLC-MS), and of major ions by ion chromatography (IC). We mainly focus on the sources and trends of winter OA and therefore we additionally analyzed an online AMS dataset acquired at the same location during the winter of the previous year. The comparison of online- and offline-AMS data, and organic marker concentrations enables an in-depth characterization of OA sources in Marseille, and in particular the identification of the main processes by which biomass smoke is emitted and transformed in this region.

2 Methods

2.1 Site description

Marseille is the second largest city in France with more than 1 million inhabitants (2010). It hosts the largest harbor in France and in the Mediterranean Sea. Many port-related industries, especially petrochemical companies, are located in a big cluster. These facilities are situated about 40 km NW from the city and include steel facilities, coke production plants, oil storing, refining plants, and several shipyards. The Marseille commercial harbor is located in the vicinity of this industrial cluster and represents the third-largest harbor of the world for crude oil storage and treatment. During summer, typical wind patterns in the city of Marseille favor the transport of polluted air masses from the industrial cluster to the city, including the sea breeze and the light Mistral wind from the Rhone valley. At night, the land breeze may transport air masses from an agricultural valley located east of the sampling site. A more detailed description of wind patterns in Marseille can be found in Drobinski et al. (2007) and Flaounas et al. (2009). The sampling location is classified as an urban background station and is situated in the urban park “Cinq Avenue” in a traffic-free zone near the city center (43°18’20’’ N, 5°23’40’’E; 64 m a.s.l.).

2.2 Yearly cycle dataset

Sample collection.

In total, 216 24-h (from midnight-to-midnight) integrated $PM_{2.5}$ pre-baked (500°C for 3 h) quartz fiber filters (150 mm diameter, Tissuquartz) were collected between 30 July 2011 and 20 July 2012 using a High-Volume sampler (Digitel DA80) operated at 500 L min^{-1} (Batch 1). Filter samples were subsequently wrapped in aluminum foil, sealed in polyethylene bags and stored at -18°C.

Offline-AMS analysis. This work discusses the offline-AMS analysis of 55 composite samples (created from the batch of 216 $PM_{2.5}$ filters collected) which were analyzed by Salameh et al. (~~*in prep*~~*submitted*~~,~~) for major ions, molecular markers and elements (Table S1). A thorough description of the offline-AMS analysis can be found in Daellenbach et al. (2016). One punch per filter sample (from 5 to 25 mm diameter depending on the filter loading and on the number of punches per composite sample) was prepared for analysis. Punches from the same composite sample were extracted together in 15 mL of ultrapure water (18.2 MΩ cm, total organic carbon < 5ppb, 25°C) in an ultrasonic bath for 20 min at 30°C. After extraction, filters were vortexed for 1 min, and the resulting liquids were filtered with 0.45 μm nylon membrane syringe filters.

The generated liquid extracts were atomized in air using a custom-made two-nozzle nebulizer. The generated aerosol was dried using a silica gel diffusion drier and then measured by a high resolution time-of-flight AMS (HR-ToF-AMS, running in V-mode). In the AMS, particles are flash vaporized (600°C) and the resulting gas is then ionized by electron impact (EI, 70eV), yielding quantitative mass spectra of the non-refractory submicron aerosol components, including OA, NO_3^-, SO_4^{2-}, NH_4^+, and Cl^-. A detailed description of the AMS operating principles, calibration protocols, and analysis procedures are provided by DeCarlo et al. (2006). In total about 10 mass spectra (mass range 12-300 Da, 60 sec averaging time) were collected per composite sample. Between each sample, a measurement blank was recorded via nebulization of ultra-pure water to minimize and monitor the possible memory effects of the system. In total five mass spectra were collected per each measurement blank. Offline-AMS data were processed and analyzed using the HR-ToF-AMS analysis software SQUIRREL (SeQUential Igor data RetRiEvaL) v.1.52L and PIKA (Peak Integration by Key Analysis) v.1.11L for IGOR Pro software package (Wavemetrics, Inc., Portland, OR, USA). HR

analysis of the mass spectra was performed in the mass range 12-115 Da and in total 217 ion fragments were fitted.

The interference of NH_4NO_3 on the CO_2^+ signal was corrected according to Pieber et al. (2016) as follows:

$$\boldsymbol{CO_{2,real} = CO_{2,meas} - \left(\frac{CO_{2,meas}}{NO_{3,meas}}\right)_{NH_4NO_3,pure} * NO_{3,meas}} \quad (1)$$

where the $\left(\frac{CO_{2,meas}}{NO_{3,meas}}\right)_{NH_4NO_3,pure}$ correction factor was 2.5% as determined from aqueous NH_4NO_3 measurements conducted regularly during the measurement period.

Other offline measurements. A complete list of the measurements performed can be found in Table S1. To summarize, major ions (Ca^{2+}, Mg^{2+}, K^+, Na^+, NH_4^+, NO_3^-, SO_4^{2-}, Cl^-, oxalate, malate, succinate, and malonate) were measured by ion chromatography (IC) according to the methodology described by Jaffrezo et al. (1998). A subset of the filters was selected for CO_3^{2-} quantification following the method described by Karanasiou et al. (2011). The method encompasses the fumigation of the filter samples with HCl. The CO_2 evolved by this acidification of the carbonates deposited on the filters is detected by thermal optical transmittance determination. The CO_3^{2-} measurements agreed fairly well with the CO_3^{2-} estimate from ion balance calculations based on IC data (Fig. S1). In the following discussion, ion concentrations from filter samples always refer to the IC measurements unless otherwise specified.

Elemental and organic carbon (EC and OC) were determined for each filter by thermal optical transmittance using a Sunset Lab analyzer (Birch and Cary, 1996) following the EUSAAR2 protocol (Cavalli et al., 2010). The CO_3^{2-} concentration determined from the IC ion balance was then subtracted from OC concentration. The water soluble OC (WSOC) was measured with a total organic carbon analyzer (TOC) following the methodology described in Bozzetti et al. (2016a) and references therein. Before the analyses, the liquid extracts were treated with a 2 M HCl solution for 1-30 min to remove the inorganic C fraction. Total nitrogen (TN) was determined using a TOC analyzer combustion tube. The NO_2 generated from the water soluble (WS) N decomposition was detected by a chemiluminescence TNM-1 unit detector. Organic markers were measured via GC-MS analysis, following the methodology described in El Haddad et al. (2009; 2011), Favez et al. (2010) and Piot et al. (2012). In total 15 different polycyclic aromatic hydrocarbons (PAH), 19 alkanes (C19-C36), 8 hopanes, 5

phtalate esters, levoglucosan, 6 lignin pyrolysis compounds, 6 fatty acids, and 3 sterols were determined (Table S1). 33 chemical elements (Table S1) were quantified using ICP-MS according to the procedure described in Chauvel et al. (2010) and the modifications suggested in El Haddad et al. (2011). A subset of 20 composite samples was selected for the quantification of methyl-nitrocatechol isomers (Table S1) via Ultra Performance Liquid Chromatography coupled with an Electro Spray Ionization – ToF - MS (UPLC-ESI-ToF-MS), following the procedure described in Iinuma et al. (2010).

2.3 Intensive winter campaign

A HR-ToF-AMS was deployed at the same station (urban park "Cinq Avenue") between 25 January 2011 and 2 March 2011 to monitor the real-time NR-PM_1 aerosol chemical composition. Although February 2011 is not included in the sampling period covered by offline-AMS, these online measurements nonetheless provide a good opportunity to compare the separation, relative contributions and winter seasonal trends of the OA sources retrieved by the offline- and online-AMS source apportionment procedures. Summer offline-AMS results were instead compared with online-AMS source apportionment results reported by El Haddad et al. (2013). The AMS was operated with an averaging time of 8 minutes, and in total 5633 mass spectra were collected during the monitoring period. We performed an ionization efficiency (IE) calibration by NH_4NO_3 nebulization, and the resulting IE value of $1.76 \cdot 10^{-7}$ was applied to the dataset. The standard relative ionization (RIE) efficiency was assumed for organics (1.4), SO_4^{2-} (1.2), NH_4^+ (4) and Cl^- (1.3), while the collection efficiency (CE) was estimated using the composition dependent collection efficiency model (Middlebrook et al., 2012). Total AMS-PAHs were estimated from AMS data according to Dzepina et al. (2007).

Similarly to offline-AMS, online-AMS data were also processed and analyzed using HR-ToF-AMS Analysis software SQUIRREL (SeQUential Igor data RetRiEvaL) v.1.52L and PIKA (Peak Integration by Key Analysis) v.1.11L for IGOR Pro software package (Wavemetrics, Inc., Portland, OR, USA). HR analysis of the mass spectra was performed in the mass range 12-115 Da and in total 215 ion fragments were fitted.

A NO_x analyzer was run in parallel to the AMS to monitor the real-time NO_x concentration. A set of pre-baked (500°C for 3h) 24-h integrated $PM_{2.5}$ filter samples was also collected during this campaign (Batch 2) following the same sampling and storage procedure described in

Section 2.2. Filters were analyzed for major ions, metals, elemental and organic carbon (EC/OC), and organic markers, including n-alkanes, hopanes, polyaromatic hydrocarbons (PAHs), lignin and cellulose pyrolysis products using the techniques previously described in Section 2.2 (Table S1).

Table 1. Monitoring periods.

Online-AMS	Offline-AMS
28 January 2011 – 02 March 2011	30 July 2011 – 20 July 2012

Formatted: Font: (Default) Times New Roman, Font color: Auto, Not

Formatted Table

Formatted: Font: (Default) Times New Roman, 11 pt, Font color: Auto, Not Highlight

2.4 Source apportionment

Implementation. The online- and offline-AMS source apportionment results discussed in this work were obtained from PMF analysis (Paatero and Tapper, 1994) of AMS spectra using the Multilinear Engine (ME-2, Paatero 1999). The Source Finder toolkit (SoFi, Canonaco et al., 2013, v.5.1) for Igor Pro (Wavemetrics, Inc., Portland, OR, USA) served as interface for data input and results evaluation. PMF is a multilinear statistical tool used to describe the variability of a multivariate dataset as the linear combination of static factor profiles times their corresponding time series, as described in Eq. 2:

$$x_{i,j} = \sum_{z=1}^{p} g_{i,z} \cdot f_{z,j} + e_{i,j} \quad (2)$$

Here $x_{i,j}$, $g_{i,z}$, $f_{z,j}$, and $e_{i,j}$ represent respectively elements of the data matrix, factor time series matrix, factor profile matrix and residual matrix, while subscripts i,j and z denote time elements, variables (in our case AMS fragments), and discrete factor numbers, respectively. p represents the total number of factors selected by the user for current given PMF solution. The PMF algorithm returns only $g_{i,z}$ and $f_{z,j}$ values ≥ 0, and solves Eq. 2 by minimizing the object function Q, defined as:

$$Q = \sum_i \sum_j \left(\frac{e_{i,j}}{s_{i,j}}\right)^2 \quad (3)$$

Here $s_{i,j}$ is an element of the error input matrix. PMF is subject to rotational ambiguity, i.e. different **G**·**F** combinations characterized by the same *Q* can exist. The ME-2 implementation of the PMF algorithm offers an efficient exploration of the solution space by directing the solution toward environmentally-meaningful rotations by constraining the factor profile elements $f_{z,j}$ for one or more *z* factors. In the *a*-value implementation of ME-2, the elements of the factor profile matrix **F** (in our case AMS fragments) are forced to predefined values $f_{z,j}$ allowing a certain variability defined by the *a*-value, such that the modelled element $f_{z,j}'$ satisfies Eq. 4:

$$\frac{(1-a)f_{z,n}}{(1+a)f_{z,m}} \leq \frac{f_{z,n'}}{f_{z,m'}} \leq \frac{(1+a)f_{z,n}}{(1-a)f_{z,m}} \quad (4)$$

where *n* and *m* represent any two arbitrary variables in the normalized **F** matrix. A complete description of the *a*-value approach can be found elsewhere (Canonaco et al., 2013).

For the offline–AMS source apportionment, the PMF input data matrix was constructed as follows: each composite sample is represented by approximately 10 time points *i*, corresponding to the ~10 mass spectra collected per filter sample (section 2.4). Each point of the data matrix is subtracted by the average corresponding measurement blank.

The error matrices were instead constructed as follows. For online-AMS source apportionment, the error matrix elements $s_{i,j}$ were calculated according to Allan et al. (2003) and Ulbrich et al. (2009), and included the uncertainty deriving from electronic noise, ion-to-ion variability at the detector, and ion counting statistics. $s_{i,j}$ included also a minimum error which was applied according to Ulbrich et al. (2009). For the offline-AMS source apportionment, the error term $\delta_{i,j}$ was calculated in the same way, but a further term ($\sigma_{i,j}$) including the blank subtraction uncertainty was propagated according to Eq. 5:

$$s_{i,j} = \sqrt{\delta_{i,j}^2 + \sigma_{i,j}^2} \quad (5)$$

Finally for both online- and offline-AMS we applied a down-weighting factor of 3 to all variables with an average signal to an average error ratio lower than 2 (Ulbrich et al., 2009). No variable with an average signal to error value lower than 0.2 was detected.

Dust and ash can contain significant amount of inorganic CO_3^{2-}. Both the IC balance and the CO_3^{2-} measurements revealed non-negligible contributions from CO_3^{2-} in the $PM_{2.5}$ fraction (Fig. S1). Preliminary PMF results also resolved a factor correlating with Ca^{2+} (SI) which was characterized by high fCO_2^+, suggesting a possible solubilisation of CO_3^{2-} from dust which

could affect the OA mass spectral fingerprint. Overall, as discussed in the SI, we could not achieve a clear inorganic dust separation using PMF, therefore we opted for a correction of the PMF input matrices. The measured pH of our filter extracts was never >8, therefore we can exclude the presence of CO_3^{2-} in the extracts, but rather assume all solubilized CO_3^{2-} to exist as HCO_3^-. Direct measurements of nebulized standard $NaHCO_3$ aqueous solutions, revealed that thermal decomposition of HCO_3^- on the AMS vaporizer (600°C) releases CO_2 (Fig. S2). Currently no HCO_3^- correction for the OA spectra is implemented in the standard AMS fragmentation table (Aiken et al., 2008), therefore the measured CO_2^+ signal needs to be subtracted from the OA AMS spectra. Offline-AMS PMF input matrices were corrected for HCO_3^- and rescaled for $WSOM_i$ ($= WSOC_{TOC} \cdot (OM:OC)_{offline\text{-}AMS,i}$) according to the procedure described in the Supplementary information, SI.

*Online-AMS source apportionment optimization***.** In the following we describe the optimization of the online-AMS source apportionment results. In order to optimize the source separation we performed sensitivity analyses on PMF solutions. We adopted different optimization strategies for online- and offline-AMS source apportionments (SI) as we encountered dissimilar mixings between sources. This is not surprising as the two methods are characterized by different time resolution and different monitoring time extension (1 year for offline-AMS, 1 month for online-AMS), which in turn result in different variabilities apportioned by the PMF algorithm (daily for online-AMS *vs.* seasonal for offline-AMS).

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) Constrain of both the HOA and COA factors profiles adopting an *a*-value approach. *a*-value sensitivity analysis (121 PMF runs performed scanning all the COA and HOA *a*-value combinations, a-value scanning steps: 0.1).

IV) Classification of the 121 PMF runs based on the cluster analysis of the COA diurnal cycles. Selection of the best clusters, and corresponding PMF solutions.

V) PMF rotational ambiguity exploration. 100 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously

Formatted: Font: (Default) Times New Roman, English (U.S.)

Formatted: English (U.S.)

Formatted: Font color: Auto, Englis (U.S.)

Formatted: Line spacing: 1.5 lines Adjust space between Latin and Asia text, Adjust space between Asian tex and numbers

Formatted: Space Before: 6 pt, Lin spacing: 1.5 lines, Numbered + Lev 1 + Numbering Style: I, II, III, … + Start at: 1 + Alignment: Left + Align at: 1.25 cm + Indent at: 2.52 cm, Adjust space between Latin and Asia text, Adjust space between Asian tex and numbers

Formatted: Font color: Auto, Englis (U.S.)

Formatted: Font color: Auto, Englis (U.S.)

Formatted: List Paragraph, Numbe + Level: 1 + Numbering Style: I, II, III, … + Start at: 1 + Alignment: Lef Aligned at: 1.25 cm + Indent at: 2. cm, Adjust space between Latin and Asian text, Adjust space between As text and numbers

varying the COA and HOA *a*-value combinations (using only the optimal *a*-value combinations identified from step IV). The average of the 100 bootstrap runs represented the online-AMS source apportionment average solution. The corresponding standard deviation represents the source apportionment uncertainty.

~~W~~For online-AMS we selected a 4-factor solution based on residual analysis. We investigated the time dependent $Q(t)/Q_{exp}(t)$ evolution when increasing the number of factors. Q/Q_{exp} is defined as the ratio between Q (as defined in Eq. 3) and the remaining degrees of freedom of the model solution (Q_{exp}) calculated as $i \cdot j-(j+i)p$ (Canonaco et al., 2013). A decrease of the Q/Q_{exp}, from lower to higher order solutions indicates an improvement in the variation explained by the model. In particular we calculated the $\Delta(Q/Q_{exp}(t))$ obtained as the difference between the $Q/Q_{exp}(t)$ for a (z)–factor solution minus the $Q/Q_{exp}(t)$ value obtained from the (z-1)–factor solutions, where z indicates the number of factors. We observed a large reduction of $\Delta(Q/Q_{exp}(t))$ until 4 factors (Fig. S4). Higher order solutions provided only minor contributions to the explained variability and in terms of solution interpretability resulted in a splitting of primary sources which could not be unambiguously associated to specific aerosol sources or processes.

Using an *a*-value approach, we constrained HOA and COA profiles from Mohr et al. (2012) and Crippa et al. (2013b) respectively. Leaving COA and/or HOA unconstrained enabled resolving COA only by increasing the number of factors (>5 factor solutions) while in the 4 factor solutions we observed a splitting of an OOA factor which could not be attributed to specific processes. Unconstrained PMF yielded HOA and COA time series well correlating with the constrained solutions; however in the unconstrained case, HOA and COA factor profiles showed higher fCO_2^+ in comparison with literature studies (Crippa et al., 2013b; Mohr et al., 2012; Bruns et al., 2015; Docherty et al., 2011; Setyan et al., 2012; He et al., 2010,) and in comparison with the constrained PMF runs. This in turn resulted in higher HOA and COA concentrations, with background night concentrations 2-3 times higher than in the constrained solutions, possibly indicative of mixings with oxidized aerosols (Fig. S5). Similar differences between constrained and unconstrained PMF runs were also observed in Elser et al. (2016). Also the HOA:NO_x ratio ($\mu g\ m^{-3}/\mu g\ m^{-3}$) matched typical literature values reported for France (0.02 Favez et al., 2010) in the constrained PMF case (0.023), while for the unconstrained approach it showed higher values (0.033).

Formatted: Font color: Auto, English (U.S.)

Formatted: Font color: Auto, English (U.S.)

Formatted: Font color: Auto, English (U.S.)

Formatted: Font color: Auto, English (U.S.)

Formatted: Font color: Auto, English (U.S.)

For both offline- and online-AMS the constrained HOA profiles were from Mohr et al. (2012), while the COA profiles were from Crippa et al. (2013b). The HOA profile from Mohr et al. (2012) was selected for offline-AMS consistently with Daellenbach et al. (2016), since the same factor recovery distributions were applied in this work. The same profile was applied to online-AMS for consistency. Overall, as discussed in the SI, the HOA profiles from literature showed high cosine similarities with each other, indicating that the AMS mass spectral fingerprints from traffic exhaust are relatively stable from station to station and consistent also with direct emission studies, making the selection of the constrained factor profiles not crucial. More variability instead is observed among COA literature profiles. For COA we selected the profile from Crippa et al. (2013b) which showed the lowest $f C_2H_4O_2^+$ value among the considered ambient literature spectra (Crippa et al., 2013b; Mohr et al., 2012). This guaranteed a better separation of COA from BBOA, as $C_2H_4O_2^+$ is strongly related to levoglucosan fragmentation (Alfarra et al., 2007).

An *a*-value sensitivity analysis was performed by scanning all possible *a*-value combinations for HOA and COA given by an *a*-value range 0-1 with a step size of 0.1. In order to optimize the source apportionment results, we retained only the PMF solutions satisfying an acceptance criterion described hereafter.

PMF factors were associated to specific aerosol emissions/processes based on mass spectral features, diurnal cycles, and time series correlations with tracers. The identified factors were associated to traffic (HOA), cooking (COA), biomass burning (BBOA), and oxygenated organic aerosol (OOA). A thorough interpretation of the PMF factors will be discussed in the Results section. Given the absence of widely accepted tracers for COA emissions, the optimization of the COA contributions was based on the analysis of the COA diurnal cycles. From the HOA and COA *a*-value sensitivity analysis we obtained a set of 121 PMF solutions each one including both factor profiles and factor time series. PMF solutions obtained ~~from the COA and HOA *a*-value sensitivity analysis (121 PMF runs in total)~~in this way were categorized according to a cluster analysis of the normalized COA diurnal cycles (Elser et al. ~~2015~~ 2016 and references therein). The *k*-means clustering approach enables classifying the PMF solutions into *k* clusters, by minimizing a cost function (C):

$$C=\Sigma_{i,z}((x_i-\mu_{z,i})^2) \quad (6)$$

where C represents the sum of the Euclidian distances between each observation (x_i) and its respective cluster centre (μ_{zi}), according to Eq. 6.

Formatted: Space Before: 6 pt, Adjust space between Latin and Asia text, Adjust space between Asian tex and numbers

Formatted: Font color: Auto, Englis (U.S.)

Formatted: Font color: Auto, Englis (U.S.)

The number of clusters (k) that best represents the data is a critical choice in order to perform a proper cluster analysis. The addition of a cluster (k+1) on one hand adds complexity to the solution, while on the other hand decreases the cost function. A typical strategy to select the right number of clusters is to explicitly penalize the addition of new clusters by using Bayesian information criteria. This approach consists in adding a penalty term to Eq. 6 proportional to the number of clusters (k):

$$C' = \Sigma_{i,z}((x_i - \mu_{z,i})^2) + k \cdot ln(D) \quad (7)$$

where D denotes the dimensionality of the clusters (24 in our case, as we consider diurnal cycles with hourly time resolution). In this study the C' function showed the minimum at 5 clusters (Fig. ~~S5~~S6). The absence of convexity properties (i.e. several local minima can exist and the solution strongly depends on the initialization) represents a possible drawback of the k-means algorithm, therefore 100 random initializations of the k-means algorithm were conducted.

The best clusters are selected based on a novel statistical analysis of the HOA, COA and BBOA average cluster spectra (SI). Briefly, a cluster was retained if the HOA, COA, and BBOA average cluster spectra were not statistically different from the average reference HOA, COA, and BBOA spectra from literature (Crippa et al., 2013b; Mohr et al., 2012; Bruns et al., 2015; Docherty et al., 2011; Setyan et al., 2012; He et al., 2010, Table S3). A complete description of the best clusters selection is reported in the SI (Figs. ~~S4~~S6-~~S7~~S10). Overall, three clusters were retained and 2 were rejected. Finally, we retained only the PMF solutions that were attributed to the 3 best clusters in more than 95% of the k-means random initializations (Fig. ~~S8~~S9).

In order to explore the rotational ambiguity of our PMF model we performed 200 PMF runs by initiating the PMF algorithm using different input matrices. The 200 different input matrices were generated using a bootstrap approach (Davison and Hinkley, 1997; Brown et al., 2015). In short, the bootstrap approach creates new input matrices by randomly resampling mass spectra (i elements) from the original input matrices. Note that some mass spectra are resampled multiple times, while others are not represented at all. On average we randomly resampled 63±1% of the original spectra per bootstrap PMF run. Finally, each bootstrap PMF run was initiated by randomly varying the HOA and COA a-values using the {a-value HOA; a-value COA} combinations previously selected as optimal from the cluster analysis (Fig. ~~S9~~S10). Only solutions showing a higher COA diurnal correlation with the three selected clusters than with the two rejected clusters were retained. In this way we

rejected 3.7% of the solutions. In the following we present the average bootstrap solution. The source apportionment uncertainty is calculated as the variability of the retained bootstrap PMF runs.

Offline-AMS source apportionment optimization.

~~1.~~ In this section we discuss the optimization of the offline-AMS source apportionment. The PMF input matrices included 217 ions and 538 time elements deriving from about 10 AMS mass spectral repetitions collected for each of the 54 composite samples.

In order to optimize the source separation, we performed sensitivity analyses on PMF solutions according to the following scheme:

I) Selection of the number of factors based on residual analysis.

II) Qualitative evaluation of the unconstrained PMF solution in comparison with the constrained PMF solutions (*a*-value approach: COA and/or HOA constraints)

III) PMF rotational ambiguity exploration. 1080 bootstrap (Davison and Hinkley, 1997; Brown et al., 2015) PMF runs were performed by simultaneously varying the COA and HOA *a*-value combinations. PMF solutions were retained based on the correlation of the PMF factors with external tracers. The PMF solutions retrieved from this step are relative to the water-soluble fraction. The corresponding water-soluble OC factor concentrations were determined by dividing the water-soluble OM factor concentrations (PMF output) by the OM:OC ratio determined from the corresponding factor mass spectrum.

IV) Retained water-soluble OC PMF solutions from step (III) were rescaled to the total OC concentrations by applying factor recoveries. Factor recoveries were fitted (using *a*-priori information) to match total OC. Only PMF solutions and factor recoveries fitting OC with yearly and seasonally homogenous residuals were retained. The average of the retained PMF solutions represented the average source apportionment results. The corresponding standard deviation represented the source apportionment uncertainty.

Formatted: Normal, No bullets or numbering

Formatted: Font color: Auto, English (U.S.)

Formatted: Line spacing: 1.5 lines

Based on analysis of the PMF residuals, we selected a 5-factor solution to explain the variability of our dataset (Fig. ~~S10~~S11). Similar to online-AMS, we monitored the decrease in Q/Q_{exp} when increasing the number of factors (z). In this study, a large Q/Q_{exp} decrease was observed until 5 factors. We also observed a clear $\Delta Q/Q_{exp}$ structure removal until 5 factors, with higher order solutions leading to additional factors that were not attributable to specific aerosol sources/processes. The 5 separated factors included HOA, COA, BBOA, OOA, and industry-related OA (INDOA). The complete validation of the PMF factors will be discussed in the Results section.

As already mentioned, the HOA and COA profiles were constrained using an *a*-value approach. Consistently with online-AMS we constrained the profiles according to Mohr et al. (2012) and Crippa et al. (2013b) respectively. Unconstrained PMF runs for offline-AMS did not resolve HOA and COA factors.

Formatted: Font color: Auto, English (U.S.)

Formatted: Font color: Auto, English (U.S.)

To explore the rotational ambiguity of our PMF model we performed 1080 bootstrapped PMF runs. In this case we performed a higher number of bootstrap runs than online-AMS because the COA and HOA *a*-value combinations could not be separately optimized because the offline-AMS method cannot resolve diurnal patterns. Each PMF run was also initiated using different input matrices. As previously mentioned the input matrices contained about 10 mass spectral repetitions per filter sample, therefore the bootstrap algorithm was implemented to randomly resample 54 filters samples, each one with all the corresponding mass spectral repetitions. The final generated matrices included 54 samples; note that some filter samples could be resampled more times, while others were not resampled at all. On average 63±5% of the original samples were resampled. Finally, each of the PMF runs was initiated by randomly varying the HOA and COA *a*-values. The optimal PMF solutions were selected based on 6 acceptance criteria including:

1) Significantly (p=0.05) positive Pearson correlation coefficient R between BBOA and levoglucosan.
2) Significantly positive R between HOA and NO_x.
3) Significantly positive R between INDOA and Se.
4) BBOA correlation with levoglucosan (R) significantly higher than the correlation between COA and levoglucosan.

5) HOA correlation with NO_x significantly higher than the correlation between COA and NO_x.
6) COA correlation with Se significantly smaller than the correlation between INDOA and Se.

Criteria 1-3 analyse the correlation between factor and marker time series. The significance of a correlation was determined by calculating the Fisher transformed correlation coefficient *z* (Garcia, 2011):

$$z = 0.5 * \ln\left(\frac{1+R}{1-R}\right) = \arctan(R) \quad (8)$$

where *R* is the Pearson correlation coefficient between factor and marker time series. Subsequently we conducted a *t*-test to verify the significance (α=0.95) of the correlation:

$$t = \frac{R}{\sqrt{\frac{1-R^2}{N-2}}} \quad (9)$$

Here, *N* represents the number of samples (54). For a confidence interval of 95% the minimum significant correlation was *R* = 0.23. For criteria 4-6, in order to evaluate whether HOA, BBOA and INDOA correlated significantly better than COA with their corresponding markers, we compared the *z* values obtained between each factor and its corresponding tracer (e.g. BBOA and levoglucosan) and between COA and the same tracer (e.g. levoglucosan), using a standard error on the *z* distribution of $1/\sqrt{N-3}$ (Zar, 1999).

In total, we retained 1.5% of the PMF runs. The criteria that discarded the largest number of solutions were the ones based on COA (4-6) correlation with tracers of other sources. This suggests that for this dataset the COA separation from other sources was particularly difficult due to the absence of high temporal resolved data which aids the separation of a distinct COA diurnal cycle. Moreover, this separation is also complicated by the small COA contribution estimated by both online- and offline-AMS source apportionments (on average 0.4 μg m^{-3} as discussed in the following sections). Furthermore, the relatively small ~~R_{COA}~~ COA factor recovery (R_{COA} median 0.54) hampers the COA apportionment by offline-AMS.

The PMF performed on offline-AMS mass spectra returned water-soluble OA factor concentrations, $WSKOA_i$. To rescale the water-soluble OA concentration to the total OA, ~~$WSKOA_i$~~, we used the factor recoveries (R_k) reported by Daellenbach et al. (2016) for the HOA, COA, BBOA, and OOA factors (R_{HOA}, R_{COA}, R_{BBOA}, R_{OOA}).

Formatted: Font: Not Italic, Font color: Auto

Formatted: Font: Italic, Font color: Auto

Formatted: Font: Italic, Font color: Auto, Subscript

Formatted: Font: Italic, Font color: Auto

$$KOA_i = WSKOA_i/R_{KOA} \quad (10)$$

This is the first offline-AMS study where an INDOA factor was identified. Therefore, we determined the INDOA recovery (R_{INDOA}) in this study, by performing a single parameter fit according to Eq. 11:

$$OC_i = \frac{WSHOA_i}{\frac{OM}{OC}_{HOA} \cdot R_{HOA}} + \frac{WSCOA_i}{\frac{OM}{OC}_{COA} \cdot R_{COA}} + \frac{WSBBOA_i}{\frac{OM}{OC}_{BBOA} \cdot R_{BBOA}} + \frac{WSOOA_i}{\frac{OM}{OC}_{OOA} \cdot R_{OOA}} + \frac{WSINDOA_i}{\frac{OM}{OC}_{INDOA} \cdot R_{INDOA}} \quad (11)$$

500 different fits were performed for each of the retained PMF solutions. Moreover each fit was initiated using different R_{KOA} combinations randomly selected from the R_{KOA} combinations determined by Daellenbach et al. (2016) and reported in Bozzetti et al. (2016a). In order to account for possible WSOC and OC systematic measurement biases, each fit was initiated by also perturbing the OC_i, $WSKOA_i/(OM:OC)_{WSKOC}$ ~~$WSKOA_i$~~, and R_{KOA} inputs assuming for each parameter a possible bias of 5%, corresponding to the WSOC and OC measurement accuracy (we note that the sum of the $WSKOC_i/(OM:OC)_{WSKOC}$ terms equals $WSOC_i$ neglecting the PMF residuals). Finally the input OC_i was randomly perturbed within its measurement uncertainty assuming a normal distribution of the errors. Among the performed fits we retained the recovery combinations and factor time series associated with OC_i unbiased residuals (residual distribution centered on 0 within the 1st and 3rd quartiles) for all seasons together and for summer and winter separately (Fig. ~~S11~~S12). Accordingly, we retained 13% of the solutions. All the retained factor recovery combinations can be found at http://doi.org/10.5905/ethz-1007-75. The median INDOA recoveries were estimated as 0.69 (1st quartile 0.65, 3rd quartile 0.73, Fig. ~~S12~~S13), while the retained R_{KOA} for the other sources were consistent within the quartiles with the R_{KOA} values reported by Daellenbach et al. (2016) despite their input value was perturbed as described above. The variability of the retained solutions is considered as our best estimate of the source apportionment uncertainty, which accounts for offline-AMS repeatability, model rotational uncertainty explored bootstrapping the input matrices and scanning the HOA and COA a-value sensitivity, and R_{KOA} uncertainties. Overall, for a generic factor KOA, we estimated the corresponding average relative uncertainty as follows: we calculated the campaign averages of the KOA concentrations for each of the v retained PMF solutions ($\overline{KOA_v}$). The relative uncertainty of the KOA concentration was calculated as the standard deviation of $\overline{KOA_v}$ divided by its average.

Formatted: Italian (Italy)

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto, English (U.S.)

Formatted: Font: (Default) Times New Roman, Font color: Auto, Englis

Formatted: Font: (Default) Times New Roman, Font color: Auto, Englis

Formatted: Font: (Default) Times New Roman, Font color: Auto, Englis

Field Code Changed

We also explored a 4-factor solution without constraining the COA profile. In this case we performed 100 bootstrap PMF runs by randomly varying the HOA *a*-value. Results revealed the COA separation (in the 5-factor solution with COA constrained) affected the HOA separation more than the other factors (BBOA, OOA, INDOA). Overall, when comparing the 4- and 5-factor solutions (without and with COA constrained, respectively) HOA showed not statistically different concentrations within our estimated source apportionment uncertainty for 85% of the samples, BBOA and OOA for 96%, and INDOA for 94%. This is probably due to the high similarity between COA and HOA spectra (SI, "Best cluster selection" Section), which are both characterized by high contributions from hydrocarbons.

3 Source apportionment validation

Figure 1 displays the stacked seasonal average concentrations of the measured $PM_{2.5}$ components (ions measured by IC, elements measured by ICP-MS, EC by the EUSAAR method, and OM estimated as the sum of the offline-AMS PMF factors). Higher concentrations were observed during winter than in summer due to the enhanced contributions of NO_3^- and OM. NO_3^- increased during winter and autumn due to NH_4NO_3 partitioning into the particle phase at lower temperatures. OM concentrations were higher during winter due to the strong BBOA contributions.

Overall OM was the dominant $PM_{2.5}$ component over the whole year highlighting the importance of studying its sources. OM represented 46% of the total mass with higher relative contributions during winter (51%) than in summer (37%). SO_4^{2-} represented the second most abundant $PM_{2.5}$ component, contributing on average 12% of the mass. Among the other components, EC contributed 9% of the mass, NO_3^- 9% ($13\%_{avg}$ during winter and $3\%_{avg}$ during summer), NH_4^+ 8%, the sum of the elements 7% (3% during winter and 13% during summer, possibly because of dust resuspension), CO_3^{2-} 6%, Ca^{2+} 2%. K^+, Cl^-, Na^+, and Mg^{2+} individually did not exceed 1% of the mass. In the following, subscripts *avg*, and *med* denote average and median values, respectively.

3.1 Online-AMS source apportionment validation

PMF factors were associated to aerosol sources/processes based on mass spectral features (Fig. 2), correlation with tracers (Fig. 3), and diurnal cycles (Fig. 4). The HOA was well-correlated with NO_x (R=0.86), with peaks during rush hours (centered on 8h and 19h) and

higher concentrations during the first half of the campaign. The average HOA:/NO_x ratio (μg m^{-3}/μg m^{-3}) was 0.~~024~~023, consistent with Favez et al. (2010). The COA diurnal variation showed two peaks at lunch and dinner time (12.00 and 21.00), as observed in other cities (Elser et al., 2016; Mohr et al., 2012). The BBOA factor profile showed the highest $fC_2H_4O_2^+$ and $fC_3H_5O_2^+$ contributions among the apportioned factors. Previous studies (Alfarra et al., 2007) associated the high $fC_2H_4O_2^+$ and $fC_3H_5O_2^+$ contributions in BBOA AMS spectra to the fragmentation of anhydrous sugars from cellulose pyrolysis. The BBOA time series was well-correlated with levoglucosan (R=0.74) and AMS-PAHs (R=0.88). Note that AMS-PAHs are not unique BBOA tracers, but in general they derive from combustion sources (see SI for the comparison between AMS-PAHs and GC-MS PAHs). In this specific dataset they could partially derive from traffic, although from the AMS-PAHs multilinear regression, we estimated that 79% of the AMS-PAHs are related to BBOA and 21% to HOA, indicating that BBOA dominates the PAH emissions. The AMS-PAHs:HOA ratio was 0.0020, while the AMS-PAHs:BBOA was 0.0028.

In general, industrial emissions can be an important source of PAHs at this location as discussed in El Haddad et al. (2013). In presence of an industrial contribution, the BBOA *vs*. AMS-PAHs correlation would decrease. In this work the correlation between AMS-PAHs and the $C_2H_4O_2^+$ fragment, typically related to levoglucosan fragmentation (Alfarra et al., 2007), was high (R=0.87) and no AMS-PAHs spike was observed without a simultaneous increase of $C_2H_4O_2^+$ (Fig. ~~S14~~S15). Moreover the industrial-related OA factor resolved by El Haddad et al. (2013) was clearly associated to wind directions from W/SW (225°-270°), while in this work wind directions were oriented from W/SW only for 7% of the monitoring time, furthermore without being associated to any significant increase in the AMS-PAHs concentration (Fig. S16~~5~~), indicating the absence of clear industrial episodes.

The BBOA diurnal cycle, similarly to AMS-PAHs, showed higher values at night than during the day (Fig. 4). In addition, the BBOA highest concentrations were detected at night and associated to slow wind speeds from the E/NE which is consistent with the night land breeze direction. Moreover, strong enhancements of the BBOA factor concentrations were perceived when the wind direction shifted to the E/NE (typically around 18 o'clock during the monitoring period) suggesting that BBOA could be transported from the valleys nearby Marseille (Fig. ~~S17~~S18).

We calculated the BBOC time series by dividing the BBOA concentrations by the $OM:OC_{BBOA}$ ratio calculated from the average BBOA HR spectrum (1.60). The average BBOC:levoglucosan ratio [$\mu g\ m^{-3}/\mu g\ m^{-3}$] was 0.15, comparable to other European studies (Zotter et al., 2014; Herich et al., 2014; Minguillón et al., 2011).

The OOA profile showed the most oxidized mass spectral fingerprint with an O:C ratio of 0.67 in comparison to the values of 0.35 retrieved for BBOA, 0.12 for COA and 0.03 for HOA. The OOA time series was well correlated with the NH_4^+ time series (R=0.86), suggesting a probable secondary origin of the OOA factor (Lanz et al., 2008). The OOA diurnal cycle was flat, suggesting OOA to be representative of regionally-transported oxygenated aerosols, consistent with the conclusions of El Haddad et al. (2013).

3.2 Offline-AMS source apportionment validation

PMF factors from the offline-AMS dataset were related to aerosol sources/processes based on mass spectral features (Fig. 5), seasonal trends and correlation with tracers (Fig. 6). A comparison of the online-AMS and offline-AMS factor profiles is reported in the SI. In the following, for a generic k factor, we calculated the corresponding KOC_i time series dividing KOA_i by the OM:OC ratio determined from the average HR-AMS factor profile.

During summer, when biomass burning contributions to EC are low, HOA correlated well with EC (R = 0.76) and yielded an HOC:EC (Hydrocarbon-like OC = $HOA/(OM:OC)_{HOA}$) ratio of 0.64, similar to other European studies (El Haddad et al., 2009 and references therein). Over the whole year, the retained PMF solutions showed an HOA correlation with NO_x (R) spanning between 0.23 and 0.49. These low correlations are comparable to the ones found by El Haddad et al. (2013) at the same station by online-AMS. In this case, the relatively low HOA correlation with NO_x is probably due to the low R_{HOA} (median 0.11) which together with the low HOA concentration (1.5 $\mu g\ m^{-3}{}_{avg}$, Results section), results in small water-soluble HOA concentrations, leading to an uncertain HOA apportionment. This was already reported in previous offline-AMS studies (Daellenbach et al., 2016; Bozzetti et al., 2016a). Although the HOA variability could not be well captured, the estimated HOA concentration was corroborated by the average HOA/NO_x (0.02 $\mu g\ m^{-3}/\mu g\ m^{-3}$) which was found to be consistent with El Haddad et al. (2013) for the same station and with Favez et al. (2010) for an alpine location in France.

BBOA was identified from its mass spectral features, with the highest $fC_2H_4O_2^+$ and $fC_3H_5O_2^+$ contributions among the apportioned factors, consistent with the findings of Alfarra et al. (2007). BBOA correlated well with ~~the~~ biomass combustion tracers measured by GC-MS, such as levoglucosan (R=0.76), acetosyringone (R=0.71) and vanillic acid (R=0.84). The winter average levoglucosan:BBOC [μg m^{-3}/μg m^{-3}] ratio was equal to 0.12, consistent with other studies in Europe (Zotter et al., 2014; Herich et al., 2014; Minguillón et al., 2011).

The fourth factor (INDOA) was related to industrial emissions due to the high correlation with light alkanes (C19-C22, 0.77≤R≤0.86), Se (R=0.54), Pb (R=0.44) and with some PAHs such as pyrene (R=0.74), fluoranthene (R=0.77), and phenantrene (R=0.74). Among the measured PAHs, pyrene, fluoranthene and phenantrene showed the lowest correlations with levoglucosan (Table S1, R = 0.31, 0.29, and 0.27 respectively), suggesting that these particular PAHs were overwhelmingly emitted by INDOA rather than BBOA. We note that phenanthrene, pyrene, and fluoranthene together represent 9.6%$_{avg}$ of the PAHs mass quantified by GC-MS, indicating that PAHs are overwhelmingly emitted by BBOA. While Se is considered to be a unique coal marker in the literature (Weitkamp et al., 2005; Park et al., 2014), in Marseille this source is likely related to coke and steel production facilities (El Haddad et al., 2011). The average INDOA OM:OC (1.60) was intermediate between the OM:OC ratios of HOA (1.23) and COA (1.28), and those of BBOA (1.85) and OOA (1.82). El Haddad et al. (2013) resolved an industrial OA factor at the same station by online-AMS PMF. In that work the authors suggested a probable contribution of oxygenated OA to the resolved industrial factor, probably deriving from (photo)chemical aging during the transport from the industrial facilities to the receptor site occasionally accompanied by new particle formation processes within the industrial plume (as observed by the increased ultrafine particle number concentration associated to W/SW wind directions) . Considering the average wind speed from W/SW (0.8 km/h), and the distance between the receptor site and the Marseille commercial harbor (~ 40 km) we estimate an aging time of several hours, which could lead to a more oxidized fingerprint in comparison to the fresh primary emissions (Huang et al., 2014). Overall this factor explained the largest fraction of the variability of S- and Cl-containing organic fragments such as C_2HSO^+, CH_2SO^+, $CH_3Cl_2^+$, $CH_4SO_3^+$, $C_3H_3SO_2^+$, and $C_7H_{16}^+$.

The last factor was defined as OOA as it showed a highly oxygenated fingerprint with the largest CO_2^+ fractional contributions (fCO_2^+) among the apportioned factors (14%, in

Formatted: Font color: Auto, English (U.S.)

comparison with 11% for BBOA, 2% for HOA, and 1% for COA and INDOA). This factor showed on average the largest contributions over the year. Overall, the OOA:NH_4^+ ratio was 2.3_{avg}, in line with the values reported by Crippa et al. (2014) for 25 different European sites (2.0_{avg}; minimum value 0.3; maximum 7.3).

Previous offline-AMS (Bozzetti et al., 2016a; Bozzetti et al., 2016b; Daellenbach et al., 2016) and online-ACSM studies (e.g., Canonaco et al., 2015) conducted in Switzerland and Lithuania reported the separation of two OOA factors characterized by different seasonal trends and different $C_2H_3O^+$:CO_2^+ ratios. In particular, the OOA factor characterized by the highest $C_2H_3O^+$:CO_2^+ ratio contributed mostly during summer and was linked to secondary OA from biogenic emissions. Here we calculated a $(C_2H_3O^+:CO_2^+)_{OOA}$ ratio by subtracting the $C_2H_3O^+$ and CO_2^+ contributions deriving from primary sources, from the measured $C_2H_3O^+$ and CO_2^+ (Canonaco et al., 2015):

$$\left(\frac{C_2H_3O^+}{CO_2^+}\right)_{OOA,i} = \frac{C_2H_3O^+_{meas,i} - HOA_i \cdot fC_2H_3O^+_{HOA} - BBOA_i \cdot fC_2H_3O^+_{BBOA} - INDOA_i \cdot fC_2H_3O^+_{INDOA} - COA_i \cdot fC_2H_3O^+_{COA}}{CO_2^+{}_{meas,i} - HOA_i \cdot fCO_2^+{}_{HOA} - BBOA_i \cdot fCO_2^+{}_{BBOA} - INDOA_i \cdot fCO_2^+{}_{INDOA} - COA_i \cdot fCO_2^+{}_{COA}} \quad (12)$$

Overall, $C_2H_3O^+_{OOA}$, and $CO_2^+{}_{OOA}$ did not show a clear seasonality (Fig. S19), which hampered the separation of two OOA sources. Even though another OOA factor was not separated, El Haddad et al. (2013) estimated for the same location during summer a substantial contribution of secondary biogenic aerosol using ^{14}C measurements (no measurements conducted in other seasons). As a consequence the OOA factor resolved in this work explains both secondary biogenic and aged/secondary anthropogenic sources. The absence of a clear increase in the $(C_2H_3O^+:CO_2^+)_{OOA}$ ratio in Marseille during summer could be explained by the large emissions of anthropogenic secondary OA (SOA) precursors during winter, leading to a different $(C_2H_3O^+:CO_2^+)_{OOA}$ seasonality in comparison with previous offline-AMS studies (Daellenbach et al., 2016, Bozzetti et al. 2016a), which were conducted either at rural sites characterized by different types of vegetation, or in smaller urban areas. In general, several parameters affect the biogenic SOA concentrations and their separation, e.g. intensity of the biogenic precursor sources, air masses photochemical age, and NO_x concentrations. All those parameters were different in Marseille from previous offline-AMS studies which were conducted in central/northern Europe.

4. Results and discussion

4.1 OA source apportionment results and uncertainties.

In this study, we present one of the first OA source apportionments conducted over an entire year in the Mediterranean region. This work represents also the first comparison between HR online-AMS and HR offline-AMS source apportionments conducted at the same location, despite in two different periods. Previous studies (Daellenbach et al., 2016) reported a comparison between offline-AMS and online-ACSM results.

~~In this study, we present one the first OA source apportionments conducted over an entire year in the Mediterranean region and the first comparison between HR online-AMS and offline-AMS source apportionments, as Daellenbach et al. (2016) compared offline-AMS results to online-ACSM data in Zurich.~~ Although related to different years and size-fractions (PM_1 online-AMS, $PM_{2.5}$ offline-AMS), the offline-AMS source apportionment returned average seasonal factor concentrations not statistically different to online-AMS for both winter (Fig. 7) and summer (comparison with El Haddad et al., 2013, Fig. 8). We note that the total OC concentration quantified by online-AMS for PM_1 and by the thermal-optical procedure used for the offline-AMS source apportionment of $PM_{2.5}$ was not different on a seasonal scale considering our uncertainty which includes time variability and measurements uncertainties.

Both online and offline-AMS source apportionment revealed that BBOA was the largest OA source during winter. Offline-AMS source apportionment estimated an average BBOA concentration during winter 2011-2012 of 5.2 μg m^{-3}_{avg}, representing 43%$_{avg}$ of the OA. Similarly, online-AMS source apportionment revealed a BBOA concentration of 4.4$_{avg}$ μg m^{-3} (corresponding to 42% of OA) during February 2011. During summer, the offline-AMS BBOA concentration dropped to an average of 0.3 μg m^{-3}_{avg} representing 5% of the OA. Not surprisingly such low BBOA contributions were not resolved by online-AMS source apportionment during summer (El Haddad et al., 2013). On average the offline-AMS BBOA relative uncertainty was 9%. As a comparison, the online-AMS BBOA average relative uncertainty was 6%. Overall for both online- and offline-AMS, the BBOA contributions were the least uncertain among the primary sources, possibly because of the high loadings and the distinct seasonal and diurnal BBOA variability in comparison with the other separated factors. A comparison between the offline- and online-AMS source apportionment uncertainties can be carried out with the caveat that the online-AMS source apportionment uncertainties estimated in this work should be considered as a low estimate as they do not account for the

Formatted: Font color: Auto, English (U.S.)

Formatted: Line spacing: 1.5 lines, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

AMS mass error deriving mostly from CE, and particle transmission. This source of uncertainty affects the total OA mass but not the relative contribution of the factors. By contrast, the OA mass uncertainty was accounted for in the offline-AMS source apportionment as the OA mass was rescaled to external measurements (WSOC and OC), the uncertainty of which was propagated in the final source apportionment error (Section 2.4).

On a yearly scale, the offline-AMS source apportionment revealed that OOA was the largest OA source, with the highest relative contributions during summer due to the reduced BBOA emissions. The OOA concentration during summer was estimated from offline-AMS at 3.0 μg m^{-3}_{avg}, corresponding to 55% of the OA mass. El Haddad et al. (2013) also reported OOA to be the dominant OA fraction during summer with a similar average concentration of 2.9 μg m^{-3}. During winter, the OOA concentration was estimated by online-AMS to be 3.9_{avg} μg m^{-3} corresponding to 38% of the OA, while the OOA relative uncertainty was 4%. As a comparison, the OOA relative uncertainty from offline-AMS was $6\%_{avg}$. The offline-AMS source apportionment revealed similar OOA concentrations during winter (3.4 μg m^{-3}_{avg} corresponding to $27\%_{avg}$ of the OA). Even though during winter the OOA concentration was higher than in summer, possibly due to partitioning and due to the shallower boundary layer, the relative contribution decreased because of the strong BBOA contributions.

HOA is one of the most uncertain factors, with an average relative uncertainty of 39% estimated from offline-AMS and 10% from online-AMS analysis, where the larger uncertainty observed for offline-AMS derives mostly from the low R_{HOA} and from the lower time resolution which does not capture the traffic diurnal variability. On average, the HOA concentration predicted by offline-AMS was 1.5 μg m^{-3}, corresponding to 17% of the OA. The estimated HOA concentration by online-AMS during February 2011 was 1.6 μg m^{-3}_{avg} (16% of OA). These values are higher than the ones of El Haddad et al. (2013) who estimated a traffic contribution of 0.8 μg m^{-3}_{avg} during July 2008.

The COA contributions were only minor (average of 0.3 μg m^{-3}), representing on average 4% of the OA mass according to the offline-AMS source apportionment. The online-AMS winter source apportionment returned similar concentrations with 0.4 μg m^{-3}_{avg}, equivalent to $4\%_{avg}$ of the OA. Overall, due to the low concentrations, the COA contributions were uncertain in both source apportionments (6% for online-AMS, 73% for offline-AMS).Similarly to HOA, the larger uncertainty observed for offline-AMS was most possibly due to the low $R_{COA,}$ and the low time resolution which did not enable the COA separation based on the diurnal

variability. The summer COA contribution was not resolved from HOA by El Haddad et al. (2013), possibly because the COA reference mass spectrum was not constrained and because of the lack of HR data which typically aid the separation of the two sources.

Finally, the INDOA factor concentration estimated from offline-AMS was on average 2.1 μg m^{-3} during winter and 0.6 μg m^{-3}_{avg} during summer, where this seasonal trend was driven by a strong episode that occurred during early February. The offline-AMS relative uncertainty was estimated as 17%. As previously discussed (Section 3.1), this factor was not separated by online-AMS analysis (February 2011) because of the absence of clear events, which in the offline-AMS dataset were observed only over a short period during January-February 2012. An industrial factor was instead resolved by El Haddad et al. (2013) during summer 2008, with an average concentration of 0.3 μg m^{-3}_{avg}. In that study, the industrial OA factor was also characterized by a low background intercepted by ten-fold spiking episodes.

From the sum of the offline-AMS factor concentrations we estimated the total OM mass. Using this OM and the measured OC we calculated the OM:OC ratio to be 1.40 on average. Specifically, during winter this ratio was 1.55, which is consistent with the online-AMS values determined from the HR-AMS spectra (median = 1.52, 1^{st} quartile = 1.46; 3^{rd} quartile 1.59). The bulk OM:OC variability was driven by the source variabilities. Indeed the relative contribution of the most oxidized source (OOA) was higher during summer (mostly due to the absence of BBOA), however also the relative contributions of the less oxidized sources (such as HOA and COA) were higher during summer mostly due to low BBOA contributions. The BBOA mass spectrum instead was associated with intermediate OM:OC ratios comprised between the values of COA and OOA, and therefore influenced less strongly the bulk OM:OC ratio. Overall the combination of these effects led to a higher bulk OM:OC during winter.

4.2 Insights into the BBOA origin during winter

Methyl-nitrocatechols measurements showed high correlations with BBOA (Fig. 9, R=0.95) and no correlation with OOA (R=0.06, offline-AMS source apportionment). Similarly high correlations were already observed in other studies (e.g. Poulain et al., 2011). This large correlation difference suggests that the variability of the methyl-nitrocatechols is likely explained by the BBOA source. However, methyl-nitrocatechols are secondary compounds deriving from the nitration of catechols which can be either directly emitted by wood

combustion (Schauer et al., 2001), or generated by OH· oxidation of cresols directly released by wood combustion (Iinuma et al., 2010). *m*-cresol/NO_x photooxidation experiments (Iinuma et al., 2010) revealed a total contribution of all methyl-nitrocatechol isomers to the catechol SOA of approximately 10%. Assuming ~~that~~ methyl-nitrocatechols ~~are~~ to be entirely apportioned ~~entirely~~ to the BBOA factor, we estimate a methyl-nitrocatechol-SOA contribution to BBOA on the order of 8%, indicating that part of the BBOA factor is of secondary origin. Previous studies (Atkinson and Arey, 2003) revealed an *o*-cresol life-time in the atmosphere of 2.4 minutes towards NO_3, and 3.4 h towards OH (at 298 K, dark conditions). This would suggest that such fast SOA formation can be better traced by the high time resolution online-AMS source apportionment (8 minutes) than by the offline-AMS with 24 h time resolution, and in any case only in the BB plume or in the vicinity of the emission source. Nevertheless we did not observe statistically different ratios (within 1σ, error calculated as the time variability) of OOA:NH_4^+ (1.5_{avg} and 1.25_{avg} for the offline-AMS and online-AMS source apportionments, respectively), OOA:BBOA (0.65_{avg} and 0.89_{avg} respectively), and levoglucosan:BBOC (1.13_{avg} and 1.15_{avg} respectively, Fig. 10) during winter, suggesting that despite the different time resolutions, the online and offline methods provide a comparable BBOA-SOA separation. Overall these findings suggest that rapid SOA formation is not well captured by PMF and rapidly formed SOA compounds (such as nitrocatechols) can be systematically attributed by PMF to factors commonly considered as "primary" (BBOA in this case).

Formatted: Font color: Auto, English (U.S.)

Formatted: Font: (Default) Calibri, Font color: Auto

Both online- and offline-AMS source apportionment revealed for the two different winter seasons a comparable temporal evolution of the levoglucosan:BBOC ratio (Fig. 10, and Fig. 11). This ratio showed typical literature values for domestic wood combustion in Europe during January and early February (0.05-0.2, Zotter et al., 2014; Herich et al., 2014; Minguillón et al., 2011), while during late autumn and March (Fig. 11) it increased up to 0.3, highlighting an evolution of the BBOA chemical composition. A similar seasonal trend was observed for the levoglucosan:vanillic acid, levoglucosan:syringic acid, and levoglucosan:non sea salt-K^+ (nss-K^+, calculated according to Seinfeld and Pandis, 2006) ratios (Fig. 11). Although the online dataset was limited to one month of measurements, the levoglucosan:vanillic acid ratio also showed a statistically significant increasing trend from early February to the beginning of March (confidence interval of 95%, Mann-Kendall test). These results suggest the occurrence of different types of biomass combustions during low temperature winter days compared to late autumn and early spring: as levoglucosan derives

from cellulose pyrolysis (>300°C), while vanillic and syringic acid result from lignin combustion (Simoneit et al., 1998, Sullivan et al., 2008). Different reactivities / volatilities of BBOA markers may complicate this analysis. For this reason we discuss in the following the levoglucosan stability, and propose that the major driver of the observed seasonal trends is the occurrence of different BBOA combustions.

Previous studies revealed the levoglucosan reactivity toward OH· radical oxidation (Hennigan et al., 2010) both in gas and aqueous phase (Hoffmann et al., 2010). In the following we analyze the levoglugocan and nss-K^+ time series in order to investigate the possible effects of levoglucosan chemical stability and different types of biomass combustions on the seasonal evolution of the levoglucosan:nss-K^+ ratio. During summer nss-K^+ derives mostly from dust, while levoglucosan is depleted by both photochemistry (Hennigan et al., 2010) and low BBOA emissions. Not surprisingly the levoglucosan:nss-K^+ ratio showed lower average values in summer (0.23) than in winter (3.14)~~showed lower values in summer (0.25) than in winter (3.35)~~. During winter nss-K^+ is considered to be mostly emitted by BBOA, and consistently in our dataset it shows a good correlation with BBOA tracers (R=0.66 with syringic acid). Overall, the levoglucosan:nss-K^+ ratio during the cold season manifests a behavior that is opposite to the photochemical activity (with temperature considered as a proxy) as it shows higher values during March and late autumn (up to ~~15.7~~7.11) and lower in January, February (minimum = ~~6.3~~2.79; Fig. 11) when temperature is lower and photochemistry is less intense. For these reasons we relate the winter levoglucosan:nss-K^+ variability to different types of combustion rather than to a levoglucosan depletion due to photochemistry. Furthermore we observed the highest levoglucosan concentrations (late autumn) simultaneously with the highest relative humidity (89%) values, suggesting the depletion of levoglucosan by OH· radical oxidation in aqueous phase to be not significant (Hoffmann et al., 2010).

A similar winter seasonal behavior was observed also for plant waxes. Plant waxes concentrations were estimated from high molecular weight *n*-alkanes (C24-C35) according to the methodology described by Li et al. (2010). This methodology is based on the observation that alkanes from epicuticular waxes preferentially contain an odd number of carbon atoms (Aceves and Grimalt, 1993; Simoneit et al., 1991). This was observed for a large variety of plants including broad leaf trees, conifers, palms, shrubs, grasses, and groundcover (Hildemann et al., 1996 and references therein). Waxes showed the highest concentrations

Formatted: Font: (Default) Times New Roman, Font color: Auto, English (U.S.)

during late autumn (up to 0.16 μg m^{-3}) and in May (up to 0.17 μg m^{-3}), while the minima were observed during winter (minimum 0.007 μg m^{-3}). In general, high molecular weight *n*-alkanes are typically detected in atmospheric aerosol in significant amounts during the growing season. In a similar way, Hildemann et al. (1996) estimated the highest plant waxes concentrations in April-May in Los Angeles and Pasadena where the climate is similar to Marseille. Similarly we observed the highest concentrations during May. However, comparable plant waxes concentrations were observed also in late autumn during the period characterized by the highest levoglucosan:lignin combustion tracers (Fig 11), suggesting a possible emission from open combustion of green wastes.

Taken together the above observations suggest the occurrence of combustion of cellulose-rich material during March and late autumn, compared to lignin rich biomass burning for residential heating during January. The combustion of cellulose-rich material is possibly related to agricultural waste burning at the beginning and at the end of the agricultural cycle. The occurrence of emission of biomass plumes due to land clearing episodes during March has already been reported in other parts of Europe (Ulevicius et al., 2016), and has been previously modelled for southern France (Dernier van der Gon et al., 2015, Fountoukis et al., 2014).

In this study we related the evolution of the BB composition over the cold season to the combustion of cellulose-rich and lignin-rich fuels, considering that lignin end cellulose are contained in different ratios in different biomass fuels. This designation should not be considered as an oversimplification of the combustion processes or of the fuel complexity, but rather as a classification of the BB aerosol based on our observations of increasing lignin pyrolysis products over cellulose pyrolysis products during the coldest days.

We note that BB is described in our PMF models by only one factor which therefore potentially represents a combination of several types of biomass burning sources. Increasing the number of factors did not lead to an unambiguous separation of different BBOA sources, however, the comparison with source-specific markers revealed a real BBOA composition evolution over the winter season with higher cellulose to lignin combustion tracer ratios observed during late autumn and early spring in comparison to January/February. This hypothesis of at least two types of BB sources (one linked to domestic heating, another to agricultural activities) is also supported by the direct PMF analysis of the organic and inorganic markers measured for batch 1 (Salameh et al., *submitted*)

5 Conclusions

$PM_{2.5}$ filter samples were collected during an entire year (August 2011 to July 2012) at an urban site in Marseille, France. Filter samples were analyzed by water extraction followed by nebulization of the liquid extracts and subsequent measurement of the generated aerosol with an HR-ToF-AMS (Daellenbach et al., 2016).

PMF analysis was conducted on the offline-AMS mass spectra and on online-AMS data collected at the same station during February 2011. Offline-AMS source apportionment results were also compared with a previous online-AMS source apportionment study of two weeks during July 2008 at the same location (El Haddad et al., 2013). The methods returned statistically similar seasonal factor concentrations, although different years and size fractions were considered (PM_1 for online-AMS, $PM_{2.5}$ for offline-AMS). OOA was the major source of OA during summer representing on average 55% of the OA mass, while BBOA was the dominant OA source during winter contributing on average 43% of the OA. Smaller contributions were estimated for HOA, INDOA and COA, representing 17%, 12%, and 4% of the OA mass, respectively. The contribution of primary anthropogenic sources (HOA + BBOA + COA + INDOA) was substantial over the year ($62\%_{avg}$ of OA), with larger absolute and relative contributions during winter (73% of OA_{avg}) associated with an intense biomass burning activity.

Coupling offline- and online-AMS data with molecular markers showed increasing levoglucosan:BBOC ratios during the late winter-early spring period in both 2011 and 2012. This trend was also observed for the ratios between cellulose and lignin combustion markers (e.g. levoglucosan:vanillic acid), with ratios approaching more typical domestic wood combustion European values during January/early February, and values characterized by higher values of cellulose-combustion markers during late autumn and March indicative of the influence of different types of fuels, possibly related to agricultural-related activities.

From the offline-AMS source apportionment, we observed a high BBOA correlation with nitrocatechols deriving from the nitration of catechols directly emitted by biomass combustion. These secondary components are rapidly formed in the atmosphere in presence of $NO_3\cdot$ (life time of a few minutes). Overall, despite the different time resolution, online- and offline-AMS provided a comparable SOA-BBOA separation during winter. Nevertheless, in case of fast SOA formation (relative to the time scale of the online-AMS time resolution, or

relative to the transport time to the receptor site) this separation can be hindered, and further efforts are needed to improve the SOA separation from BBOA.

Authors contributions:

Experimental design: A. S. H. Prévôt, I. El Haddad, N. Marchand, and C. Bozzetti

Filter collections and online-AMS measurements: N. Marchand and D. Salameh.

Offline-AMS analysis: C. Bozzetti, E. Müller, K. R. Daellenbach, and I. El Haddad

WSOC and TN analysis: P. Fermo and R. Gonzalez

Sunset, and IC analysis: J.-L. Jaffrezo

GC-MS analysis: D. Salameh, and N. Marchand

Nitrocatechols analysis: Y. Iinuma

Carbonates analysis: M. C. Minguillón

Data analysis and validation: C Bozzetti, and I. El Haddad

PMF analysis: C. Bozzetti

All authors read and commented on the manuscript.

Acknowledgements

CB thanks the Lithuanian–Swiss Cooperation Programme "Research and Development" project AEROLIT (Nr. CH-3-.MM-01/08). JGS acknowledges the support of the Swiss National Science Foundation (Starting Grant No. BSSGI0 155846). IE-H acknowledges the support of the Swiss National Science Foundation (IZERZ0 142146). This work has also been supported by the MED program (APICE, grant number 2G-MED09-026: http://www.apice-project.eu/), the French Environment and Energy Management Agency (ADEME) and Provence Alpes Cote d'Azur Region (PACA). Part of the OC/EC analysis carried out in MRS has been supported by the French national CARA program. This program is directed by O. Favez (INERIS; http://www.ineris.fr/). He is gratefully acknowledged.

References

Aceves, M. and Grimalt, J. O.: Seasonally dependent size distributions of aliphatic and polycyclic aromatic hydrocarbons in urban aerosols from densely populated areas, Environ. Sci. Technol., 27(13), 2896–2908, doi:10.1021/es00049a033, 1993.

Aiken, A. C., DeCarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffmann, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P. J., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S. H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., and Jimenez J. L. O/C and OM:OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry. Environ. Sci. Technol. 42, 4478-4485, 2008.

Aksoyoglu, S., Keller, J., Barmpadimos, I., Oderbolz, D., Lanz, V. A., Prévôt, A. S. H., and Baltensperger U.: Aerosol modelling in Europe with a focus on Switzerland during summer and winter episodes, Atmos. Chem. Phys., 11, 7355–7373, 2011.

Aksoyoglu, S., Keller, J., Ciarelli, G., Prévôt, A. S. H., and Baltensperger, U.: A model study on changes of European and Swiss particulate matter, ozone and nitrogen deposition between 1990 and 2020 due to the revised Gothenburg protocol, Atmos. Chem. Phys., 14, 13081-13095, doi:10.5194/acp-14-13081-2014, 2014.

Alfarra, M. R., Prévôt, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D., Mohr, M., and Baltensperger, U.: Identification of the mass spectral signature of organic aerosols from wood burning emissions, Environ. Sci. Technol., 41, 5770–5777, 2007.

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer: 1. Techniques of data interpretation and error analysis, J. Geophys. Res.-Atmos., 108, 4090, 2003.

Atkinson, R., and Arey, J.: Atmospheric degradation of volatile organic compounds, Chem. Rev., 103, 4605-4638, 2003.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G.,

Hirtl, M., Joffre, S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A., Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N., Rao, S. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmos. Chem. Phys., 14, 317-398, doi:10.5194/acp-14-317-2014, 2014.

Birch, M. E. and Cary, R. A.: Elemental carbon-based method for monitoring occupational exposures to particulate diesel exhaust, Aerosol Sci. and Tech., 25, 221–241, 1996.

Bozzetti, C., Daellenbach, K., R., Hueglin, C., Fermo, P., Sciare, J., Kasper-Giebl, A., Mazar, Y., Abbaszade, G., El Kazzi, M., Gonzalez, R., Shuster Meiseles, T., Flasch, M., Wolf, R., Křepelová, A., Canonaco, F., Schnelle-Kreis, J., Slowik, J. G., Zimmermann, R., Rudich, Y., Baltensperger, U., El Haddad, I., and Prévôt, A. S. H.: Size-resolved identification, characterization, and quantification of primary biological organic aerosol at a European rural site, Environ. Sci. Technol., 50, 3425-3434, doi:10.1021/acs.est.5b05960, 2016~~a~~.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, Atmos. Chem. Phys. 17, 117–141, 2017.

Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A., Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, Sci. Tot. Environ. 518-519, 626-635, 2015.

Bruns, E. A., Krapf, M., Orasche, J., Huang, Y., Zimmermann, R., Drinovec, L., Močnik, G., El-Haddad, I., Slowik, J. G., Dommen, J., Baltensperger, U. and Prévôt, A. S. H.: Characterization of primary and secondary wood combustion products generated under different burner loads, Atmos. Chem. Phys., 15, 2825–2841, 2015.

Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M. J., DeCarlo, P. F., Kolb, C. E., Davidovits, P. and Worsnop, D. R.: Chemical and microphysical characterization of ambient aerosols with the Aerodyne aerosol mass spectrometer, Mass Spectrom. Rev. 26:185-222, 2007.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data, Atmos. Meas. Tech., 6, 3649-3661, 2013.

Canonaco, F., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: Seasonal differences in oxygenated organic aerosol composition: implications for emissions sources and factor analysis. Atmos. Chem. Phys. 15, 6993-7002, 2015.

Cavalli, F., Viana, M., Yttri, K. E., Genberg, J., and Putaud, J. P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, Atmos. Meas. Tech., 3, 79-89, 2010.

Chauvel, C., Bureau, S., and Poggi, C.: Comprehensive chemical and isotopic analyses of basalt and sediment reference materials, Geostand. Geoanalyt. Res., 35, 125–143, doi: 10.1111/j.1751-908X.2010.00086.x, 2010.

Crippa, M., Canonaco, F., Slowik, J. G., El Haddad, I., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Temime-Roussel, B., Abidi, E., Poulain, L., Wiedensohler, A., Baltensperger, U., and Prévôt, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, Atmos. Chem. Phys., 13, 8411-8426, doi:10.5194/acp-13-8411-2013, 2013a.

Crippa, M., El Haddad, I., Slowik, J. G., DeCarlo, P.F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., L., Sciare, J., Baltensperger, U., and Prévôt, A. S. H.: Identification of marine and continental aerosol sources in Paris using high resolution aerosol mass spectrometry, J. Geophys. Res., 118, 1950-1963, 2013b.

Daellenbach, K. R., Bozzetti, C., Krepelova, A., Canonaco, F., Huang, R.-J., Wolf, R., Zotter, P., Crippa, M., Slowik, J., Zhang, Y., Szidat, S., Baltensperger, U., Prévôt, A. S. H., and El Haddad, I.: Characterization and source apportionment of organic aerosol using offline aerosol mass spectrometry, Atmos. Meas. Tech., 9, 23-39, 2016.

Davison, A. C. and Hinkley, D. V.: Bootstrap Methods and Their Application, Cambridge University Press, Cambridge, UK, 582 pp., 1997.

DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R., and Jimenez, J. L.:

Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, Anal. Chem., 78, 8281–8289, 2006.

Denier van der Gon, H. A. C., Bergström, R., Fountoukis, C., Johansson, C., Pandis, S. N., Simpson, D., and Visschedijk, A. J. H.: Particulate emissions from residential wood combustion in Europe – revised estimates and an evaluation, Atmos. Chem. Phys., 15, 6503-6519, doi:10.5194/acp-15-6503-2015, 2015.

Docherty, K. S., Aiken, A. C., Huffman, J. A., Ulbrich, I. M., DeCarlo, P. F., Sueper, D., Worsnop, D. R., Snyder, D. C., Peltier, R. E., Weber, R. J., Grover, B. D., Eatough, D. J., Williams, B. J., Goldstein, A. H., Ziemann, P. J., and Jimenez, J. L.: The 2005 Study of organic aerosols at riverside (SOAR-1): instrumental intercomparisons and fine particle composition, Atmos. Chem. Phys., 11, 12387-12420, 2011.

Dockery, D. W., Luttmann-Gibson, H., Rich, D. Q., Link, M. S., Mittleman, M. A., Gold, D. R., Koutrakis, P., Schwartz, J. D., and Verrier, R. L.: Association of air pollution with increased incidence of ventricular tachyarrhythmias recorded by implanted cardioverter defibrillators, Environ. Health Perspect. 113, 670-674, 2005.

Drewnick, F., Hings, S. S., DeCarlo, P. F., Jayne, J. T., Gonin, M., Fuhrer, K., Weimer, S., Jimenez, J. L., Demerjian, K. L., Borrmann, S., and Worsnop, D. R.: A new time-of-flight aerosol mass spectrometer (ToF-AMS) – instrument description and first field deployment, Aerosol Sci. Tech., 39, 637–658, 2005.

Drobinski, P., Said, F., Arteta, J., Augustin, P., Bastin, S., Brut, A., Caccia, J. L., Campistron, B., Cautenet, S., Colette, A., Coll, I., Corsmeier, U., Cros, B., Dabas, A., Delbarre, H., Dufour, A., Durand, P., Guenard, V., Hasel, M., Kalthoff, N., Kottmeier, C., Lasry, F., Lemonsu, A., Lohou, F., Masson, V., Menut, L., Moppert, C., Peuch, V. H., Puygrenier, V., Reitebuch, O., and Vautard, R.: Regional transport and dilution during high-pollution episodes in southern France: Summary of findings from the Field Experiment to Constraint Models of Atmospheric Pollution and Emissions Transport (ESCOMPTE), J. Geophys. Res.-Atmos., 112, D13105, doi:10.1029/2006JD007494, 2007.

Dzepina, K., Arey, J., Marr, L.C., Worsnop, D.R., Salcedo, D., Zhang, Q., Onasch, T.B., Molina L.T., Molina M. J., Jimenez, J.L.: Detection of particle-phase polycyclic aromatic hydrocarbons in Mexico City using an aerosol mass spectrometer, International Journal of Mass Spectrometry, 263, 152–170, 2007.

El Haddad, I., D'Anna, B., Temime-Roussel, B., Nicolas, M., Boreave, A., Favez, O., Voisin, D., Sciare, J., George, C., Jaffrezo, J.-L., Wortham, H., and Marchand, N.: Towards a better understanding of the origins, chemical composition and aging of oxygenated organic aerosols: case study of a Mediterranean industrialized environment, Marseille, Atmos. Chem. Phys., 13, 7875–7894, 2013.

El Haddad, I., Marchand, N., Dron, J., Temime-Roussel, B., Quivet, E., Wortham, H., Jaffrezo, J. L., Baduel, C., Voisin, D., Besombes, J. L., and Gille, G.: Comprehensive primary particulate organic characterization of vehicular exhaust emissions in France, Atmos. Environ., 43, 6190–6198, 2009.

El Haddad, I., Marchand, N., Wortham, H., Piot, C., Besombes, J.-L. , Cozic, J., Chauvel, C., Armengaud, A., Robin D., and Jaffrezo, J.-L.: Primary sources of $PM_{2.5}$ organic aerosol in an industrial Mediterranean city, Marseille, Atmos. Chem. Phys., 11, 2039-2058, 2011.

Elser, M., Huang, R.-J., Wolf, R., Slowik, J. G., Wang, Q., Canonaco, F., Li, G., Bozzetti, C., Daellenbach, K. R., Huang, Y., Zhang, R., Li, Z., Cao, J., Baltensperger, U., El-Haddad, I., and Prévôt, A. S. H.: New insights into PM2.5 chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry, Atmos. Chem. Phys., 16, 3207-3225, doi:10.5194/acp-16-3207-2016, 2016.

Favez, O., El Haddad, I., Piot, C., Boréave, A., Abidi, E., Marchand, N., Jaffrezo, J.-L., Besombes, J.-L., Personnaz, M.-B., Sciare, J., Wortham, H., George, C., and D'Anna, B.: Inter-comparison of source apportionment models for the estimation of wood burning aerosols during wintertime in an Alpine city (Grenoble, France), Atmos. Chem. Phys., 10, 5295–5314, doi:10.5194/acp-10-5295- 2010, 2010.

Flaounas, E., Coll, I., Armengaud, A., and Schmechtig, C.: The representation of dust transport and missing urban sources as major issues for the simulation of PM episodes in a Mediterranean area, Atmos. Chem. Phys., 9, 8091–8101, doi:10.5194/acp-9-80912009, 2009.

Fountoukis, C., Megaritis, A. G., Skyllakou, K., Charalampidis, P. E., Pilinis, C., Denier van der Gon, H. A. C., Crippa, M., Canonaco, F., Mohr, C., Prévôt, A. S. H., Allan, J. D., Poulain, L., Petäjä, T., Tiitta, P., Carbone, S., Kiendler-Scharr, A., Nemitz, E., O'Dowd, C., Swietlicki, E., and Pandis, S. N.: Organic aerosol concentration and composition over Europe: insights from comparison of regional model predictions with aerosol mass spectrometer factor analysis, Atmos. Chem. Phys., 14, 9061-9076, doi:10.5194/acp-14-9061-2014, 2014.

Garcia, E.: A Tutorial on Correlation Coefficients, http://web.simmons.edu/~benoit/lis642/a-tutorial-on-correlation-coefficients.pdf, pp. 7-8, 2011.

He, L.-Y., Lin, Y., Huang, X.-F., Guo, S., Xue, L., Su, Q., Hu, M., Luan, S.-J., and Zhang, Y.-H.: Characterization of high-resolution aerosol mass spectra of primary organic aerosol emissions from Chinese cooking and biomass burning, Atmos. Chem. Phys., 10, 11535-11543, doi:10.5194/acp-10-11535-2010, 2010.

Hennigan, C. J., Sullivan, A. P., Collett, J. L., and Robinson, A. L.: Levoglucosan stability in biomass burning particles exposed to hydroxyl radicals, Geophys. Res. Lett., 37, L09806, doi:10.1029/2010GL043088, 2010.

Herich, H., Gianini, M. F. D., Piot, C., Mocnik, G., Jaffrezo, J. L., Besombes, J. L., Prévôt, A. S. H., and Hueglin, C.: Overview of the impact of wood burning emissions on carbonaceous aerosols and PM in large parts of the Alpine region, Atmos. Environ., 89, 64–75, doi:10.1016/j.atmosenv.2014.02.008, 2014.

Hildebrandt, L., Kostenidou, E., Lanz, V. A., Prevot, A. S. H., Baltensperger, U., Mihalopoulos, N., Laaksonen, A., Donahue, N. M., and Pandis, S. N.: Sources and atmospheric processing of organic aerosol in the Mediterranean: insights from aerosol mass spectrometer factor analysis, Atmos. Chem. Phys., 11, 12499-12515, doi:10.5194/acp-11-12499-2011, 2011.

Hildebrandt, L., Kostenidou, E., Mihalopoulos, N., Worsnop, D. R., Donahue, N. M., and Pandis, S. N.: Formation of highly oxygen-ated organic aerosol in the atmosphere: Insights from the Finokalia Aerosol Measurement Experiments, Geophys. Res. Lett., 37, L23801, doi:10.1029/2010GL 045193, 2010.

Huang, R.-J., Zhang, Y., Bozzetti, C., Ho, K.-F., Cao, J., Han, Y., Daellenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., Zotter, P., Wolf, R., Pieber, S. M., Bruns, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., and Prévôt, A. S. H.: High secondary aerosol contribution to particulate pollution during haze events in China, Nature, 514, 218-222, 2014.

Hildemann, L., M., Rogge, W. F., Cass, G. R., Mazurek, M. A., and Simoneit, B. R. T.: Contribution of primary aerosol emissions from vegetation-derived sources to fine particle

concentrations in Los Angeles. J. Geophys. Res., 101, 19,541-19,549, doi: 10.1029/95JD02136, 2016.

Hoffmann, D., Tilgner A., Iinuma, Y., and Herrmann, H.: Atmospheric stability of levoglucosan: a detailed laboratory and modeling study, Environ. Sci. Technol., 44, 694–699, 2010.

Iinuma, Y., Böge, O., Gräfe, R., and Herrmann, H.: Methyl-nitrocatechols: Atmospheric tracer compounds or biomass burning secondary organic aerosols, Environ. Sci. Technol., 44, 8453–8459, 2010.

Jaffrezo, J. L., Calas, T., and Bouchet, M.: Carboxylic acids measurements with ionic chromatography, Atmos. Environ., 32, 2705–2708, 1998.

Jayne, J. T., Leard, D. C., Zhang, X. F., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R. (2000). Development of an aerosol mass spectrometer for size and composition analysis of submicron particles, Aerosol Sci. Technol. 33, 49–70, 2000.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, Science, 326, 1525–1529, 2009.

Karanasiou, A., Diapouli, E., Cavalli, F., Eleftheriadis, K., Viana, M., Alastuey, A., Querol1, X., and Reche, C.: On the quantification of atmospheric carbonate carbon by thermal/optical analysis protocols, Atmos. Meas. Tech., 4, 2409–2419, 2011

Laden, F., Neas, L. M., Dockery, D. W., and Schwartz, J.: Association of fine particulate matter from different sources with daily mortality in six US cities, Environ. Health Perspect. 108:941-947, 2000.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prevot, A. S. H.. Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, Atmos. Chem. Phys. 7:1503-1522, 2007.

Lanz, V. A., M. R. Alfarra, U. Baltensperger, B. Buchmann, C. Hueglin, S. Szidat, M. N. Wehrli, L. Wacker, S. Weimer, A. Caseiro, H. Puxbaum, and A. S. H. Prevot: Source attribution of submicron organic aerosols during wintertime inversions by advanced factor analysis of aerosol mass spectra, Envir. Sci. Tech., 42, 214-220, 2008.

Lee, A. K. Y., Herckes, P., Leaitch, W. R., Macdonald, A. M., and Abbatt, J. P. D.: Aqueous OH oxidation of ambient organic aerosol and cloud water organics: Formation of highly oxidized products, Geoph. Res. Lett., 38, L11 805, doi: 10.1029/2011GL047439, 2011.

Li, W., Peng, Y., and Bai, Z.: Distributions and sources of *n*-alkanes in $PM_{2.5}$ at urban, industrial and coastal sites in Tianjin, China, J. Environ. Sci., 22, 1551-1557, 2010.

Lohmann, U., Broekhuizen, K., Leaitch, R., Shantz, N., and Abbatt, J.: How efficient is cloud droplet formation of organic aerosols?, Geophys. Res. Lett. 31, L05108, doi: 10.1029/2003GL018999, 2004.

Middlebrook, A.M., Bahrein, R., Jimenez, J.L., and Canagaratna, M.R.: Evaluation of composition-dependent collection efficiencies for the Aerodyne aerosol mass spectrometer using field data, Aerosol Sci. and Tech., 46, 258-271, 2012.

Mihara, T. and Mochida, M.: Characterization of solvent-extractable organics in urban aerosols based on mass spectrum analysis and hygroscopic growth measurement, Environ. Sci. Technol., 45, 9168–9174, 2011.

Minguillón M. C., Pérez N., Marchand N., Bertrand A., Temime-Roussel B., Agrios K., Szidat S., van Drooge B., Sylvestre S., Alastuey A., Reche C., Ripoll A., Marco E., Grimalt J. O., and Querol X.. Secondary organic aerosol origin in an urban environment: influence of biogenic and fuel combustion precursors. Faraday Discussions, 189, 337-359, 2016.

Minguillón, M. C., Perron, N., Querol, X., Szidat, S., Fahrni, S. M., Alastuey, A., Jimenez, J. L., Mohr, C., Ortega, A. M., Day, D. A., Lanz, V. A., Wacker, L., Reche, C., Cusack, M., Amato, F., Kiss, G., Hoffer, A., Decesari, S., Moretti, F., Hillamo, R., Teinila, K., Seco, R., Penuelas, J., Metzger, A., Schallhart, S., Muller, M., Hansel, A., Burkhart, J. F., Baltensperger, U., and Prevot, A. S. H.: Fossil versus contemporary sources of fine elemental

and organic carbonaceous particulate matter during the DAURE campaign in Northeast Spain, Atmos. Chem. Phys., 11, 12067-12084, 2011.

Minguillón, M. C., Ripoll, A., Pérez, N., Prévôt, A. S. H., Canonaco, F., Querol, X., and Alastuey, A.: Chemical characterization of submicron regional background aerosols in the western Mediterranean using an Aerosol Chemical Speciation Monitor, Atmos. Chem. Phys., 15, 6379-6391, 2015.

Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Penuelas, J., Jimenez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prevot, 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, 2012.

Mohr, C., Huffman, J. A., Cubison, M. J., Aiken, A. C., Docherty, K. S., Kimmel, J. R., Ulbrich, I. M., Hannigan, M., and Jimenez, J. L.: Characterization of primary organic aerosol emissions from meat cooking, trash burning, and motor vehicles with high resolution aerosol mass spectrometry and comparison with ambient and chamber observations, Environ. Sci. Technol., 43, 2443–2449, 2009.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemom. Intell. Lab. Syst., 37, 23–35, 1997.

Paatero, P.: The multilinear engine - A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model, J. Comput. Graph. Stat., 8, 854-888, 1999.

Paatero, P. and Tapper, U.: Positive matrix factorization - a nonnegative factor model with optimal utilization of error-estimates of data values, Environmetrics, 5, 111-126, 1994.

Park, S.-S., Cho, S. Y., Jo, M. R., Gong, B. J., Park, J. S., and Lee, S. J.: Field evaluation of a near–real time elemental monitor and identification of element sources observed at an air monitoring supersite in Korea, Atmos Pollut Res., 5, 119-128, doi:10.5094/APR.2014.015, 2014.

Pieber, S. M., El Haddad, I., Slowik, J. G., Canagaratna, M. R., Jayne, J. T., Platt, S. M., Bozzetti, C., Daellenbach, K. R., Fröhlich, R., Vlachou, A., Klein, F., Dommen, J., Miljevic, B., Jimenez, J. L., Worsnop, D. R., Baltensperger, U., and Prévôt, A. S. H.: Inorganic salt

interference on CO_2^+ in Aerodyne AMS and ACSM organic aerosol composition studies, Environ. Sci. Technol., 50, 10494-10503, DOI: 10.1021/acs.est.6b01035, 2016.

Piot, C., Jaffrezo, J.-L., Cozic, J., Pissot, N., El Haddad, I., Marchand, N., and Besombes, J.-L.: Quantification of levoglucosan and its isomers by High Performance Liquid Chromatography – Electrospray Ionization tandem Mass Spectrometry and its applications to atmospheric and soil samples, Atmos. Meas. Tech., 5, 141–148, doi:10.5194/amt-5-141-2012, 2012.

Poulain, L., Iinuma, Y., Müller, K., Birmili, W., Weinhold, K., Brüggemann, E., Gnauk, T., Hausmann, A., Löschau, G., Wiedensohler, A., and Herrmann, H.: Diurnal variations of ambient particulate wood burning emissions and their contribution to the concentration of polycyclic aromatic hydrocarbons (PAHs) in Seiffen, Germany, Atmos. Chem. Phys., 11, 12697-12713, doi:10.5194/acp-11-12697-2011, 2011.

Rocke, D. M., and Lorenzato, S.: A two-component model for measurement error in analytical chemistry, Technometrics, 37, 176-184, 1995.

Salameh, D., Detournay, A., Pey, J., Pérez, N., Liguori, F., Saraga, D., Bove, M. C., Brotto, P., Cassola, F., Massabò, D., Latella, A., Pillon, S., Formenton, G., Patti, S., Armengaud, A., Piga, D., Jaffrezo, J.-L., Bartzis, J., Tolis, E., Prati, P., Querol, X., Wortham, H., Marchand, N.: $PM_{2.5}$ chemical composition in five European Mediterranean cities: a 1-year study, Atmos. Res., 155, 102-117, 2015.

Salameh, D., Pey, J., Bozzetti, C., El Haddad, I., Detournay, A., Sylvestre, A., Canonaco, F., Armengaud, A., Piga, D., Robin, D., Prévôt, A. S. H., Jaffrezo, J.-L., Wortham, H., and Marchand, N.: Sources of PM2.5 at an urban-industrial Mediterranean city, Marseille (France): application of the ME-2 solver to inorganic and organic markers~~Sources of $PM_{2.5}$ at an urban-industrial Mediterranean city: an application of ME-2 model to $PM_{2.5}$ offline dataset from Marseille (France)~~, Environ. Pollut., *submitted.*

Schauer, J. J., Kleeman, M. J., Cass, G. R., Simoneit, B. R. T.: Measurement of emissions from air pollution sources. 3. C1-C29 organic compounds from fireplace combustion of wood, Environ. Sci. Technol., 35, 1716–1728, 2001.

Schwarze, P. E., Ovrevik, J., Lag, M., Refsnes, M., Nafstad, P., Hetland, R. B., Dybing, E. (2006). Particulate matter properties and health effects: consistency of epidemiological and toxicological studies, Hum. Exp. Toxicol., 25, 559-579, 2006.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: English (U.S.)

Seinfeld, J. H. and Pandis, S. N: Atmospheric chemistry and physics: from air pollution to climate change, John Wiley, second Edn., New York, 2006.

Setyan, A., Zhang, Q., Merkel, M., Knighton, W. B., Sun, Y., Song, C., Shilling, J. E., Onasch, T. B., Herndon, S. C., Worsnop, D. R., Fast, J. D., Zaveri, R. A., Berg, L. K., Wiedensohler, A., Flowers, B. A., Dubey, M. K., and Subramanian R.: Characterization of submicron particles influenced by mixed biogenic and anthropogenic emissions using high-resolution aerosol mass spectrometry: results from CARES, Atmos. Chem. Phys., 12, 8131-8156, 2012.

Simoneit, B. R. T., Schauer, J. J., Nolte, C. G., Oros, D. R., Elias, V. O., Fraser, M. P., Rogge, W. F., and Cass, G. R.: Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles, Atmos. Environ., 33, 173-182, 1998.

Simoneit, B. R. T., Sheng, G., Chen, X., Fu, J., Zhang, J. and Xu, Y.: Molecular marker study of extractable organic matter in aerosols from urban areas of China, Atmos. Environ. Part A. Gen. Top., 25(10), 2111–2129, doi:10.1016/0960-1686(91)90088-O, 1991.

Slowik, J. G., Vlasenko, A., McGuire, M., Evans, G. J., and Abbatt, J. P. D.: Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site, Atmos. Chem. Phys, 10:1969-1988, 2010.

Sullivan, A. P., Holden, A. S., Patterson, L. A., McMeeking, G. R., Kreidenweis, S. M., Malm, W. C., Hao, W. M., Wold, C. E., and Collett Jr., J. L.: A method for smoke marker measurements and its potential application for determining the contribution of biomass burning from wildfires and prescribed fires to ambient PM2.5 organic carbon, J. Geophys. Res., 113, D22302, doi:10.1029/2008JD010216, 2008.

Sun, Y., Zhang, Q., Zheng, M., Ding, X., Edgerton, E. S., and Wang, X.: Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles (PM2.5) with high-resolution aerosol mass spectrometry and GC-MS, Environ. Sci. Technol., 45, 4854–4861, 2011.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891-2918, 2009.

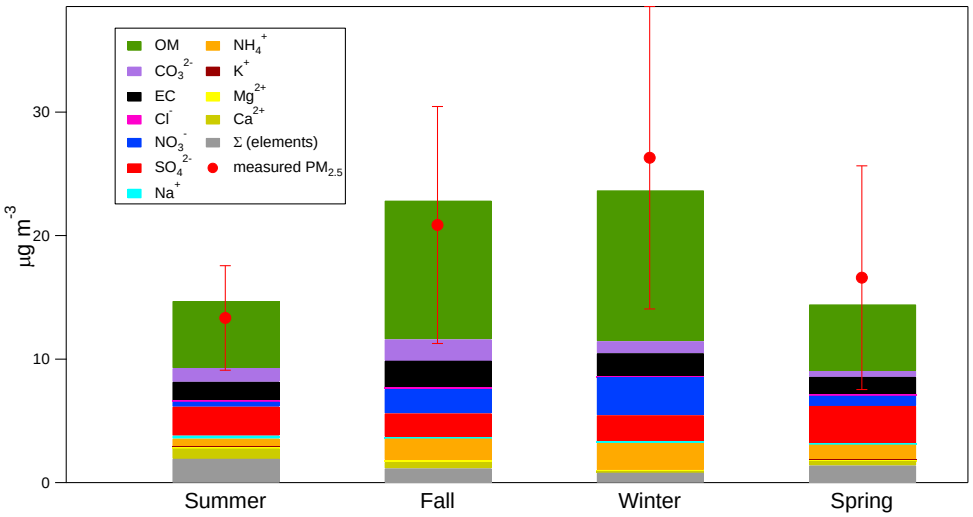
Ulevicius, V., Byčenkienė, S., Bozzetti, C., Vlachou, A., Plauškaitė, K., Mordas, G., Dudoitis, V., Abbaszade, G., Remeikis, V., Garbaras, A., Masalaite, A., Blees, J., Fröhlich, R.,

Dällenbach, K. R., Canonaco, F., Slowik, J. G., Dommen, J., Zimmermann, R., Schnelle-Kreis, J., Salazar, G. A., Agrios, K., Szidat, S., El Haddad, I., and Prévôt, A. S. H.: Fossil and non-fossil source contributions to atmospheric carbonaceous aerosols during extreme spring grassland fires in Eastern Europe. Atmos. Chem. Phys., 16, 5513-5529, 2016.

Weitkamp, E. A., Lipsky, E. M., Pancras, P. J., Ondov, J. M., Polidori, A., Turpin, B. J., and Robinson, A. L.: Fine particle emission profile for a large coke production facility based on highly time-resolved fence line measurements, Atmos. Environ., 39, 6719–6733, 2005.

Zar, J. H., Biostatistical analysis, 4th ed., 929 pp., Prentice Hall, Englewood Cliffs, New Jersey, 1999.

Zotter, P., Ciobanu, V. G., Zhang, Y. L., El-Haddad, I., Macchia, M., Daellenbach, K. R., Salazar, G. A., Huang, R.-J., Wacker, L., Hueglin, C., Piazzalunga, A., Fermo, P.,Schwikowski, M., Baltensperger, U., Szidat, S., and Prévôt, A. S. H.: Radiocarbon analysis of elemental and organic carbon in Switzerland during winter-smog episodes from 2008 to 2012 – Part 1: Source apportionment and spatial variability, Atmos. Chem. Phys., 14, 13551–13570, doi:10.5194/acp-14-13551-2014, 2014.



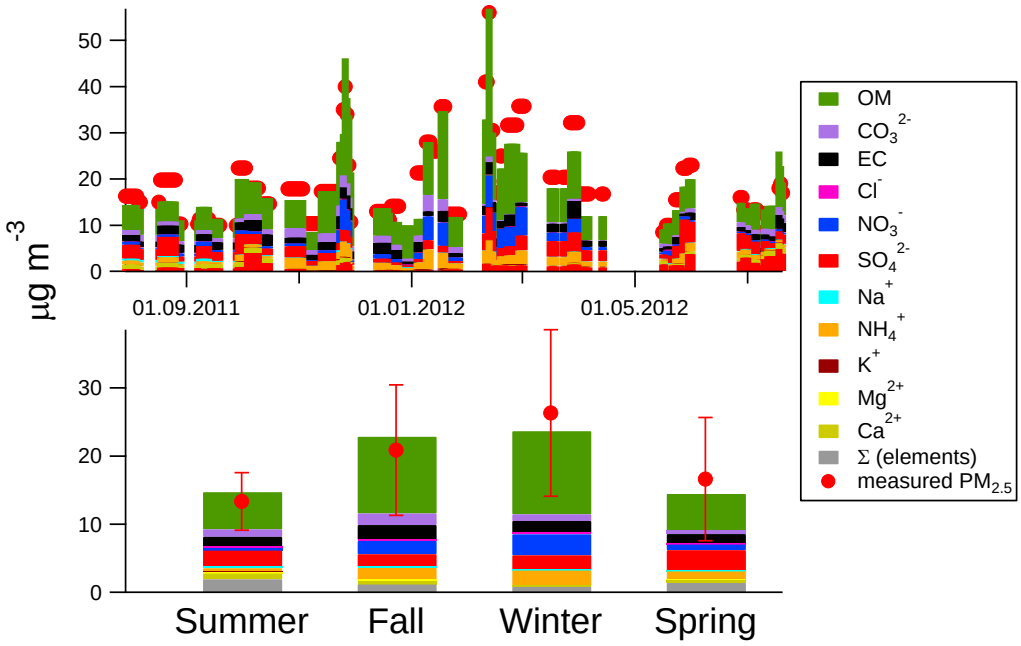


Figure 1. $PM_{2.5}$ composition: stacked average seasonal concentrations. Measured $PM_{2.5}$ error bars represent the seasonal standard deviation. OM was estimated as the sum of the offline-AMS source apportionment factors.

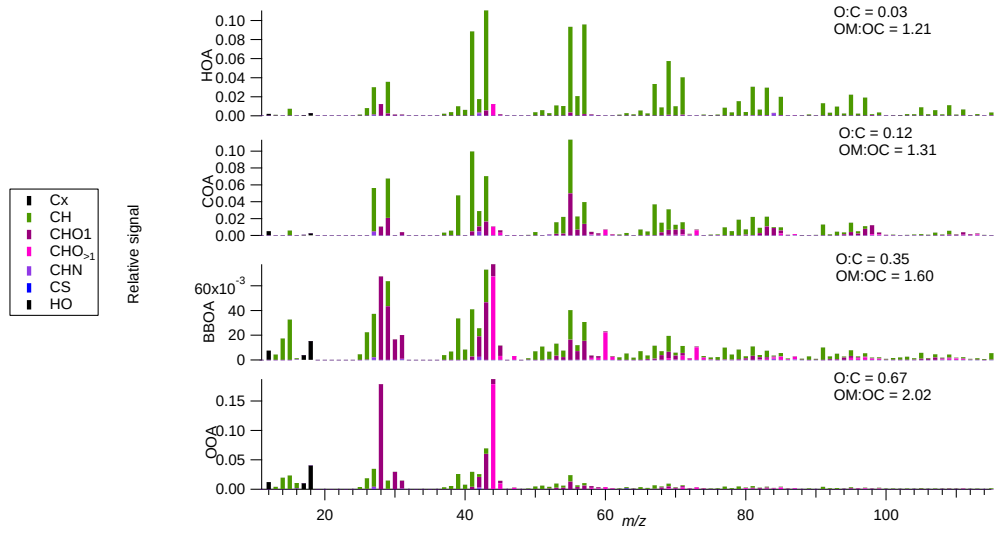


Figure 2. Online-AMS: average PMF factor mass spectra.

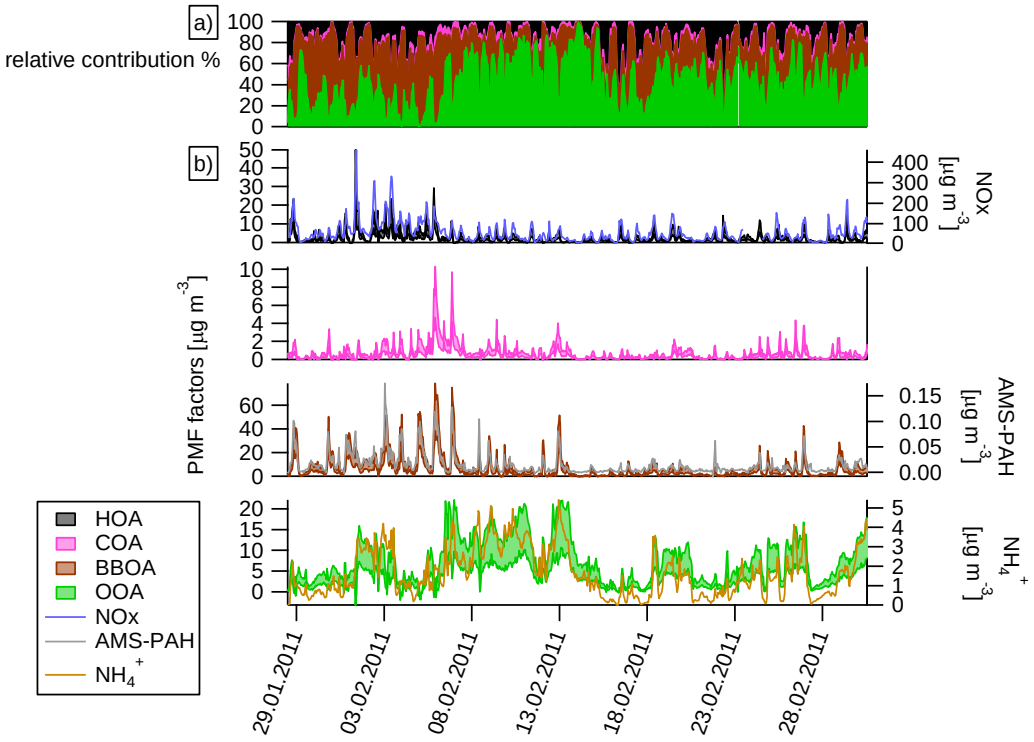


Figure 3. Online-AMS: a) PMF factors relative contributions. b) Time series of PMF factors and corresponding tracers. Shaded areas denote the model uncertainties.

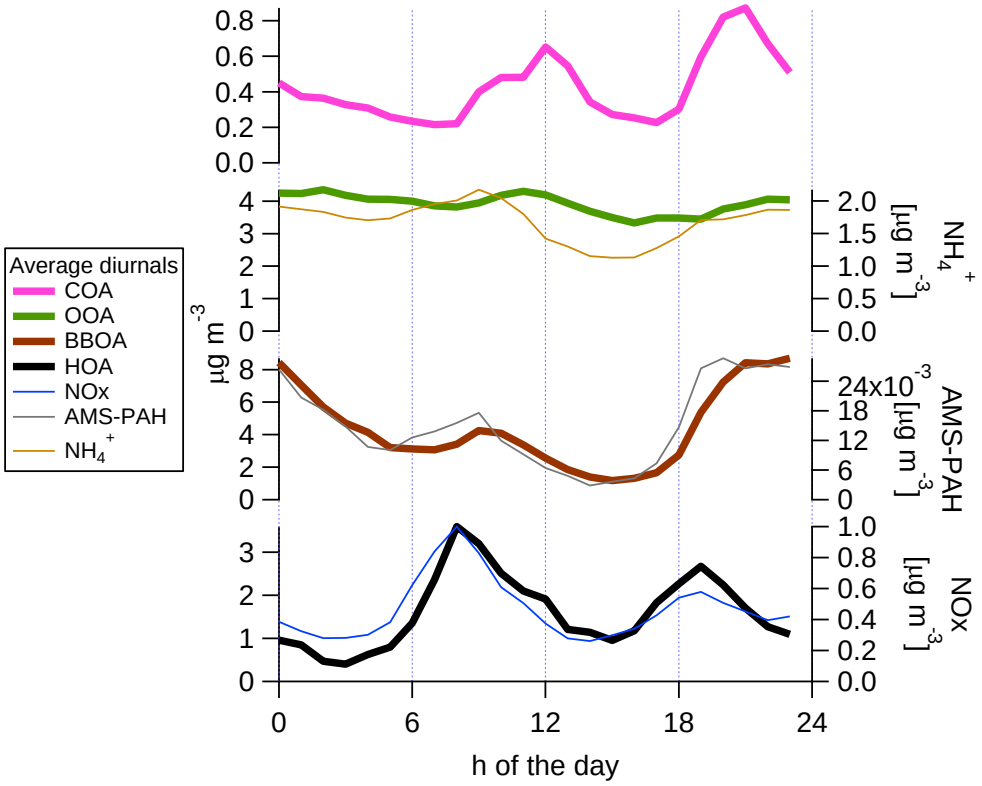


Figure 4. Online-AMS: average diurnal cycles of PMF factors and corresponding tracers.

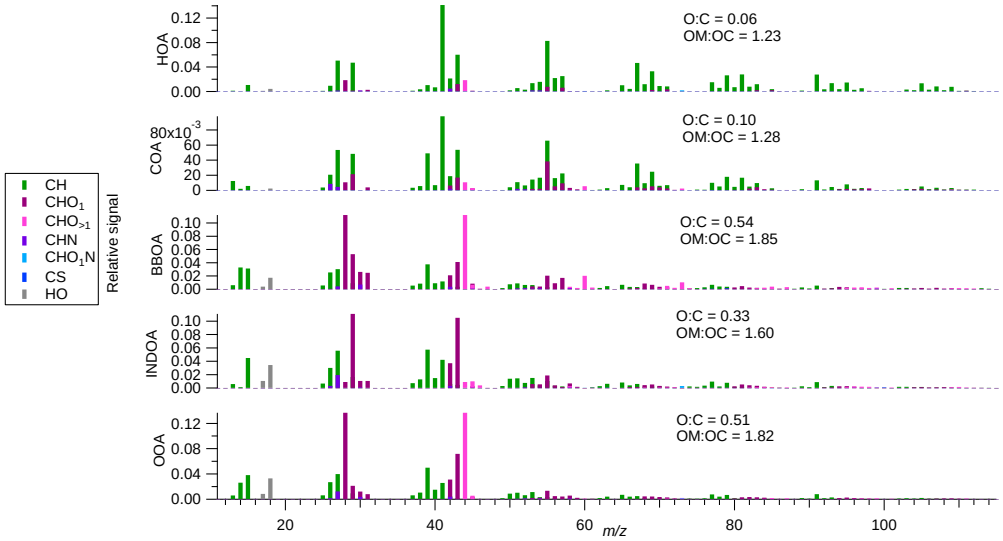
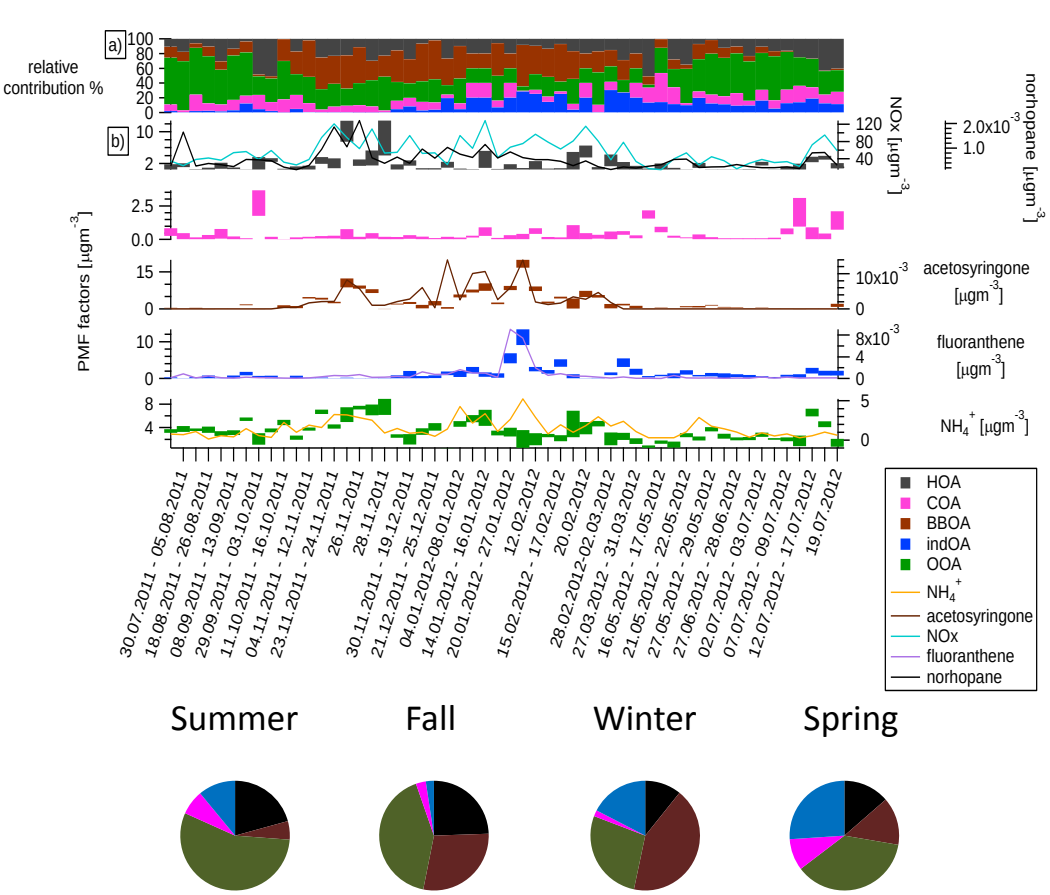
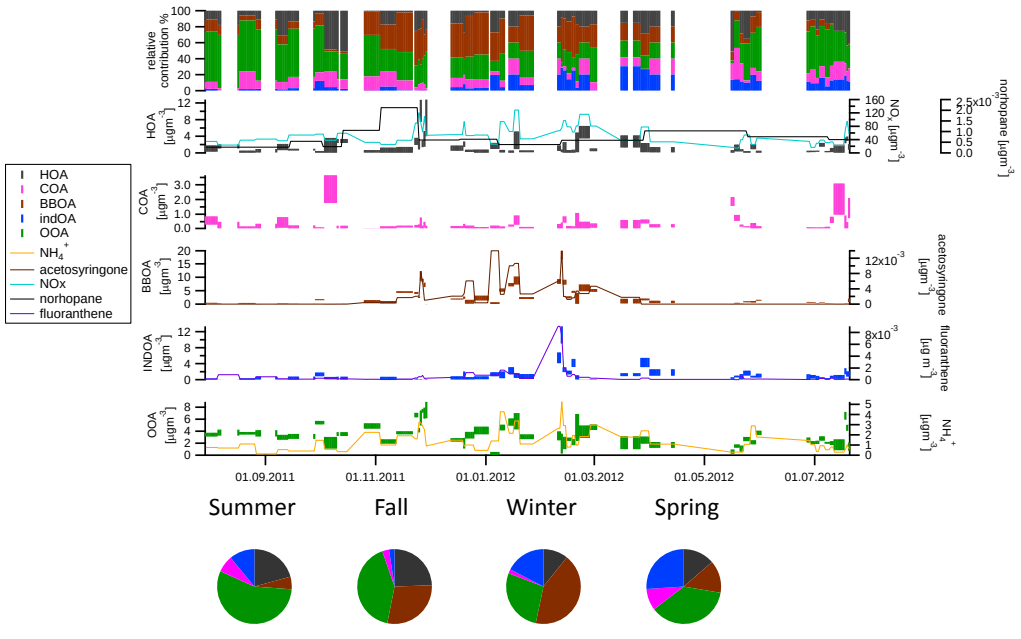


Figure 5. Offline-AMS: water soluble average mass spectra.


relative contribution %
a)
b)
PMF factors [µgm-3]
NOx [µgm-3]
norhopane [µgm-3]
acetosyringone [µgm-3]
fluoranthene [µgm-3]
NH4+ [µgm-3]
HOA
COA
BBOA
indOA
OOA
NH4+
acetosyringone
NOx
fluoranthene
norhopane
Summer
Fall
Winter
Spring

1


relative contribution %
HOA [µgm-3]
COA [µgm-3]
BBOA [µgm-3]
INDOA [µgm-3]
OOA [µgm-3]
NOx [µgm-3]
norhopane [µgm-3]
acetosyringone [µgm-3]
fluoranthene [µg m-3]
NH4+ [µgm-3]
HOA
COA
BBOA
indOA
OOA
NH4+
acetosyringone
NOx
norhopane
fluoranthene
01.09.2011
01.11.2011
01.01.2012
01.03.2012
01.05.2012
01.07.2012
Summer
Fall
Winter
Spring

2

Figure 6. Offline-AMS: a) PMF factors relative contributions. b) Time series of PMF factors and corresponding tracers. Bars denote the model uncertainties.

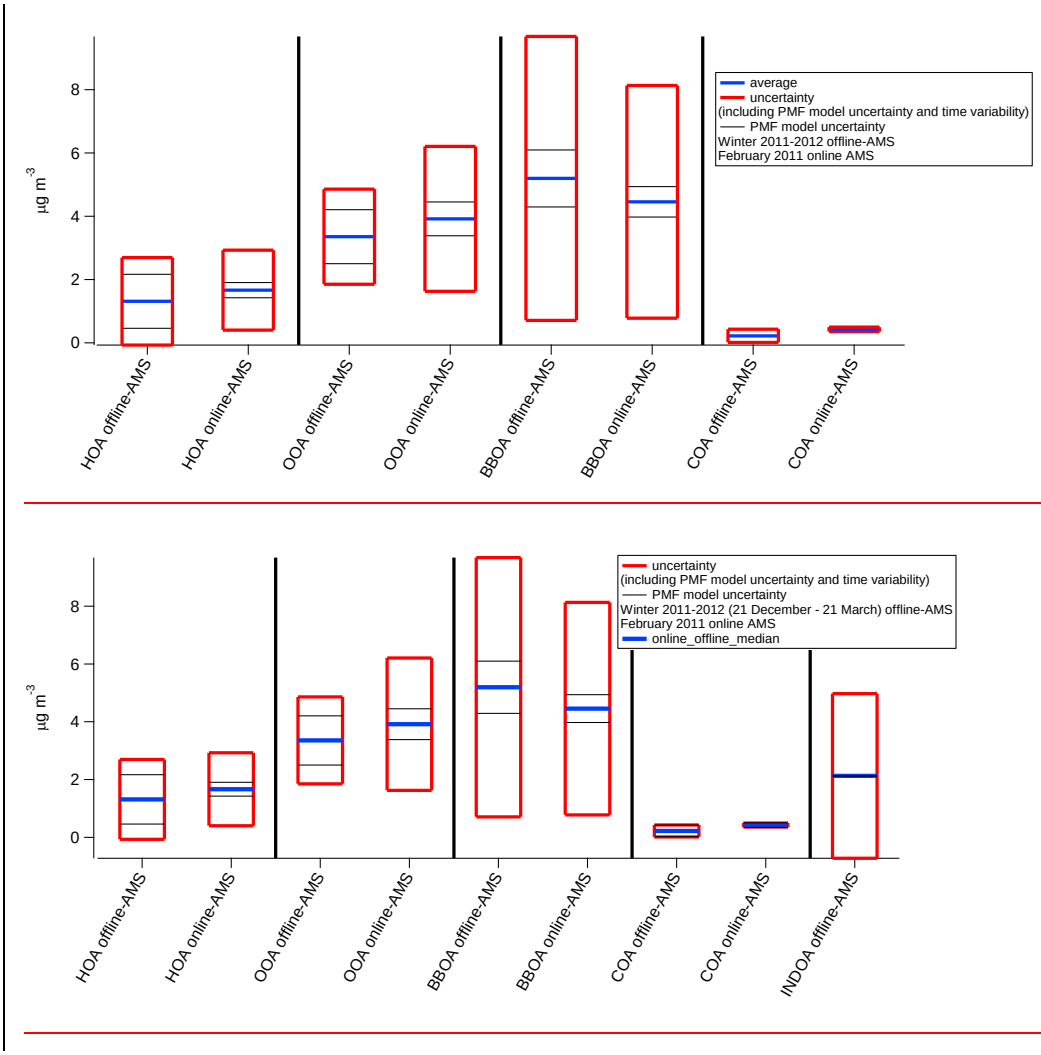
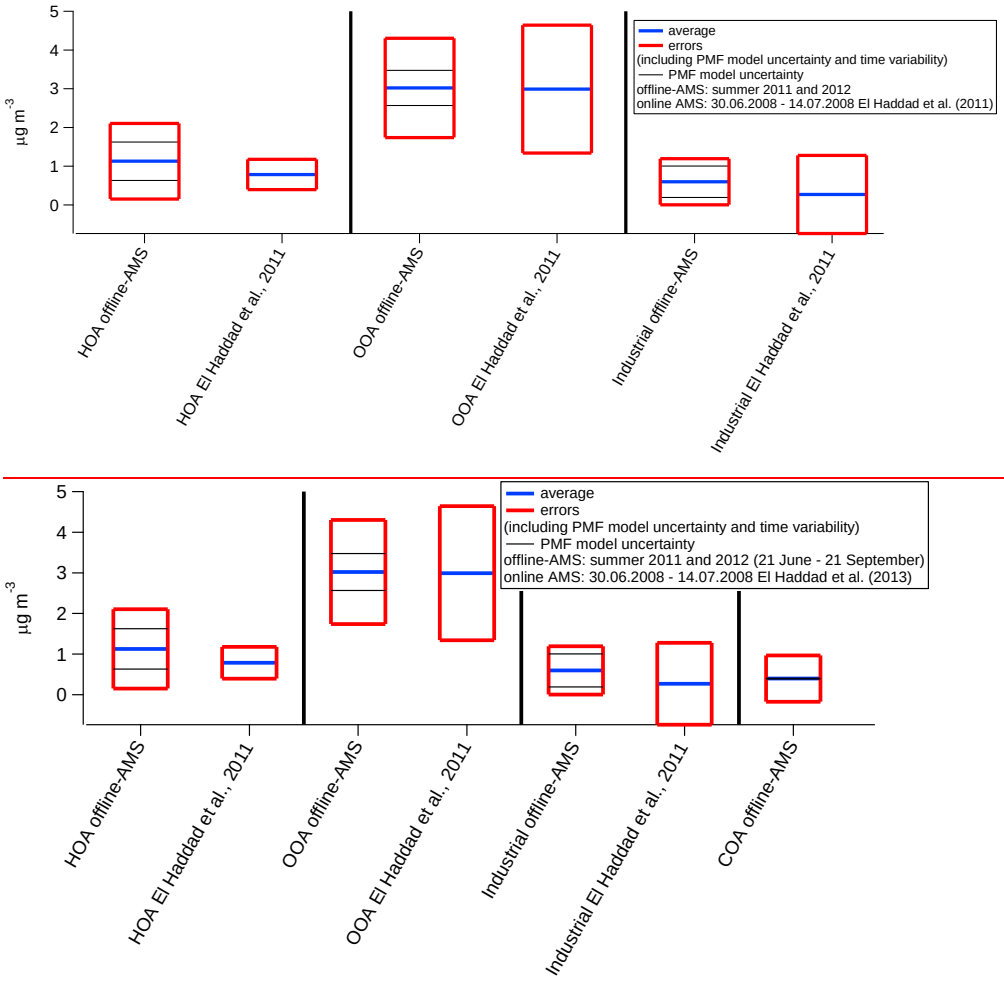


Figure 7. Online (PM_1) and offline-AMS ($PM_{2.5}$) comparison. Bars represent the error including temporal variability and model uncertainty.



1

2

3 Figure 8. Online (PM_1, El Haddad et al., 2013) and offline-AMS ($PM_{2.5}$) comparison. For
4 offline-AMS bars represent the error including temporal variability and model uncertainty.
5 For online-AMS bars represents only the temporal variability.

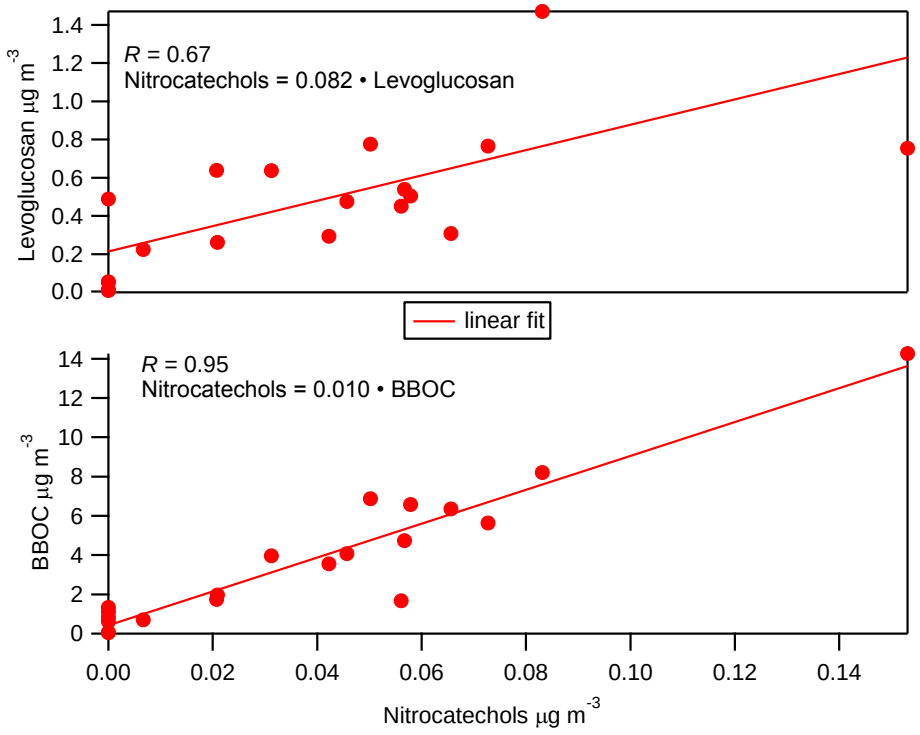


Figure 9. Correlation between the sum of nitrocatechols (Table S1) with levoglucosan and BBOC.

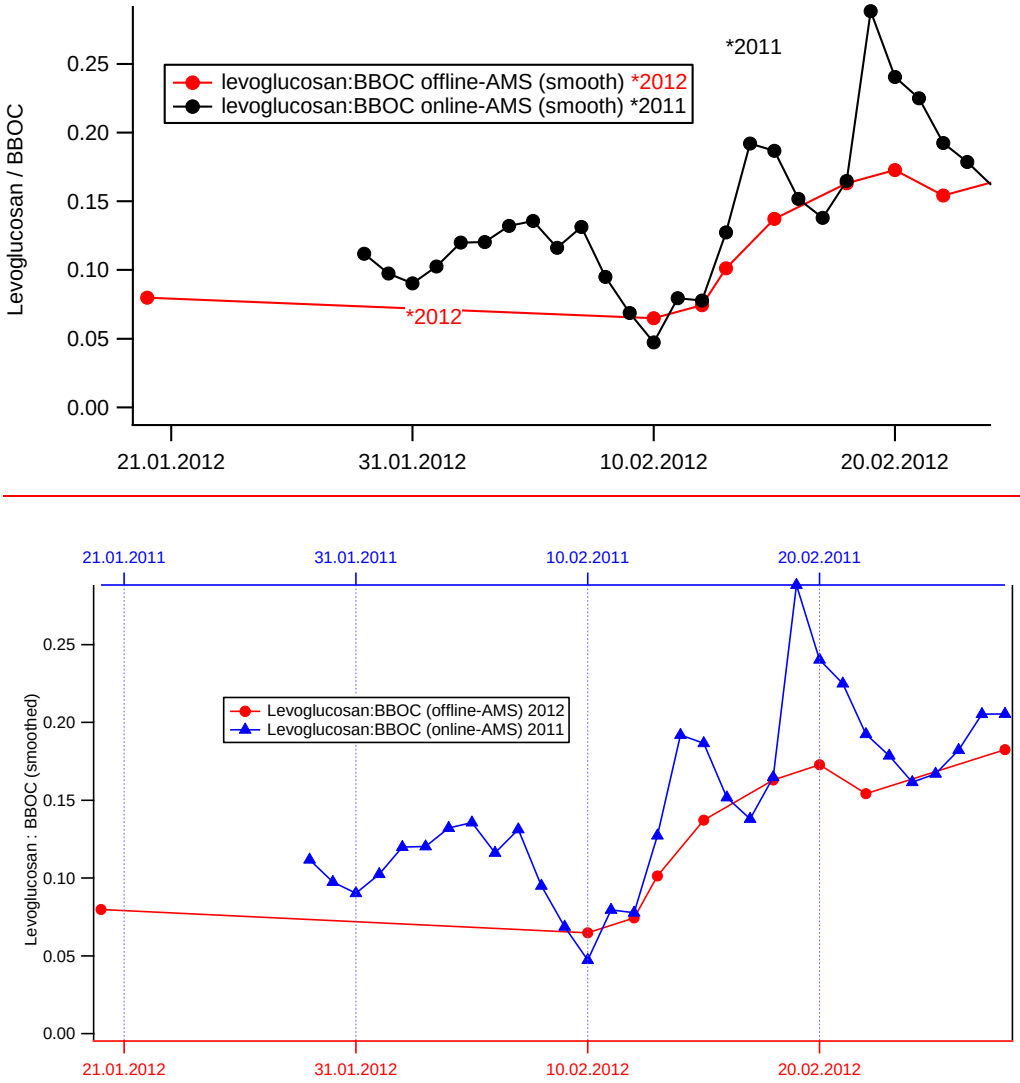


Figure 10: Offline-AMS (February 2012) and online-AMS (February 2011) smoothed time-dependent levoglucosan:BBOC ratios. We note that the levoglucosan:BBOC comparison should not be considered on a day-to-day basis where the levoglucosan:BBOC ratio in the two different years can be coincidentially equal or different, but rather on a monthly time scale where, as discussed in the manuscript, we observed a statistically significant (p=0.05) evolution of the levoglucosan:BBOC ratio which is similarily captured by the two models.

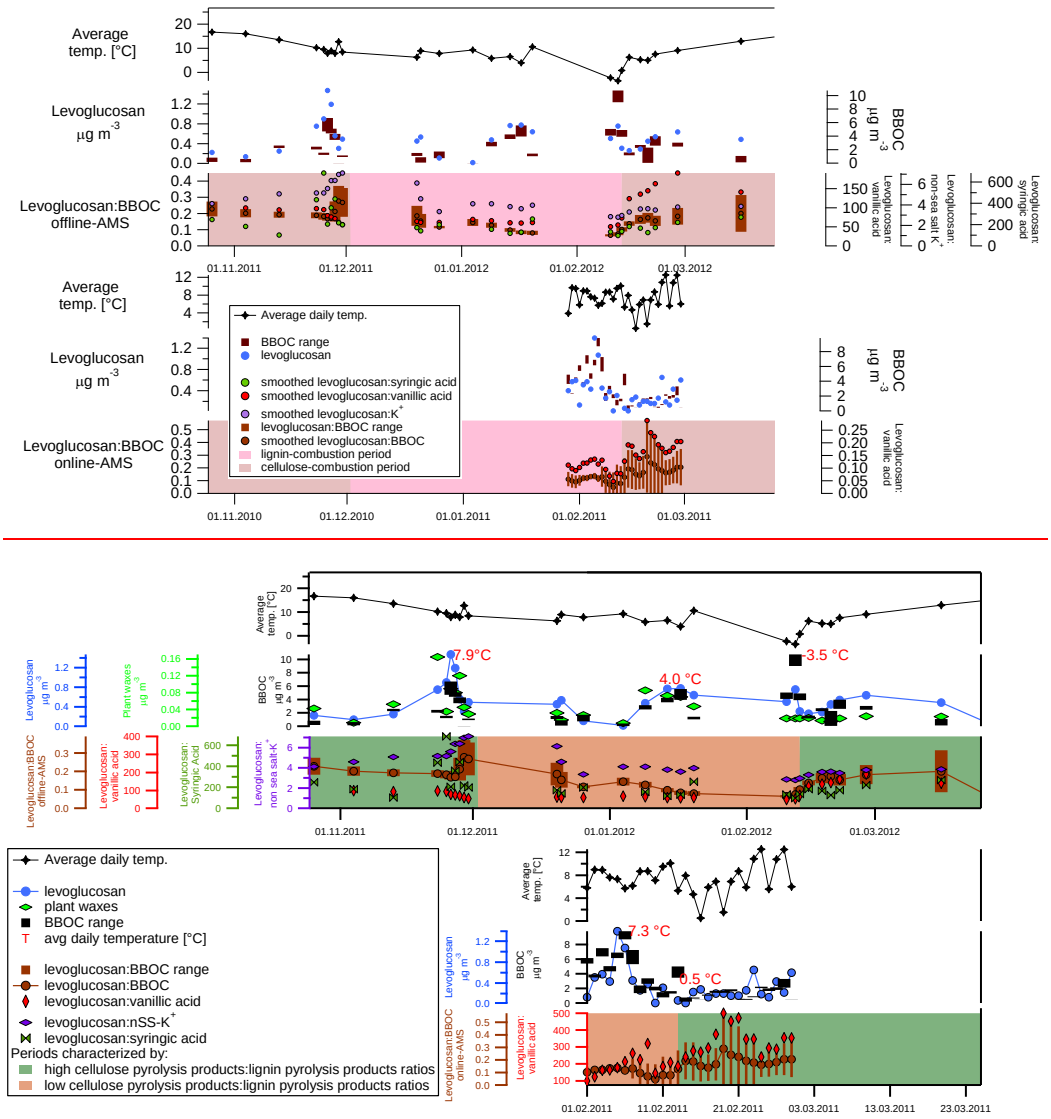


Figure 11: Online- and offline-AMS time-dependent levoglucosan:BBOC, levoglucosan:vanillic acid, levoglucosan:syringic acid, and levoglucosan:K^+ ratios. The plant waxes concentrations were determined from GC-MS measurements of alkanes with an odd number of carbons (Li et al., 2010). As discussed in the main text the spike observed in late autumn could be related to incomplete green waste combustion.

Formatted: Line spacing: 1.5 lines, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Font color: Auto, English (U.S.)

Formatted: English (U.S.)

Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers