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

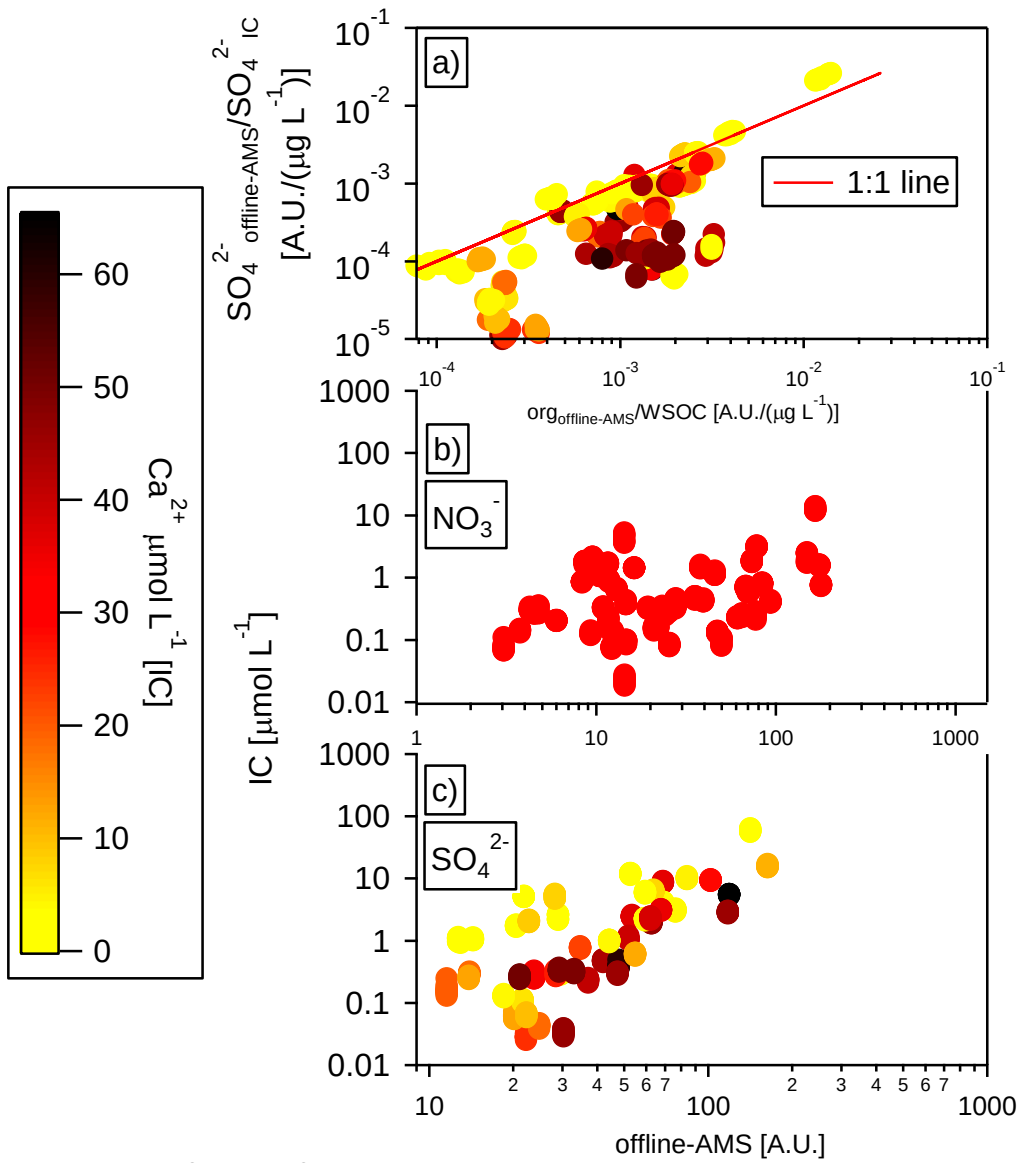


Figure D1. a) $SO_4^{2-}{}_{AMS}$:$SO_4^{2-}{}_{IC}$ correlation with $WSOC_{AMS}$:$WSOC_{IC}$. b) and c) NO_3^- and SO_4^{2-}: Offline-AMS comparison with IC.

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 extraction of a single component k to be independent of the presence of the other organics (k forms its own 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.

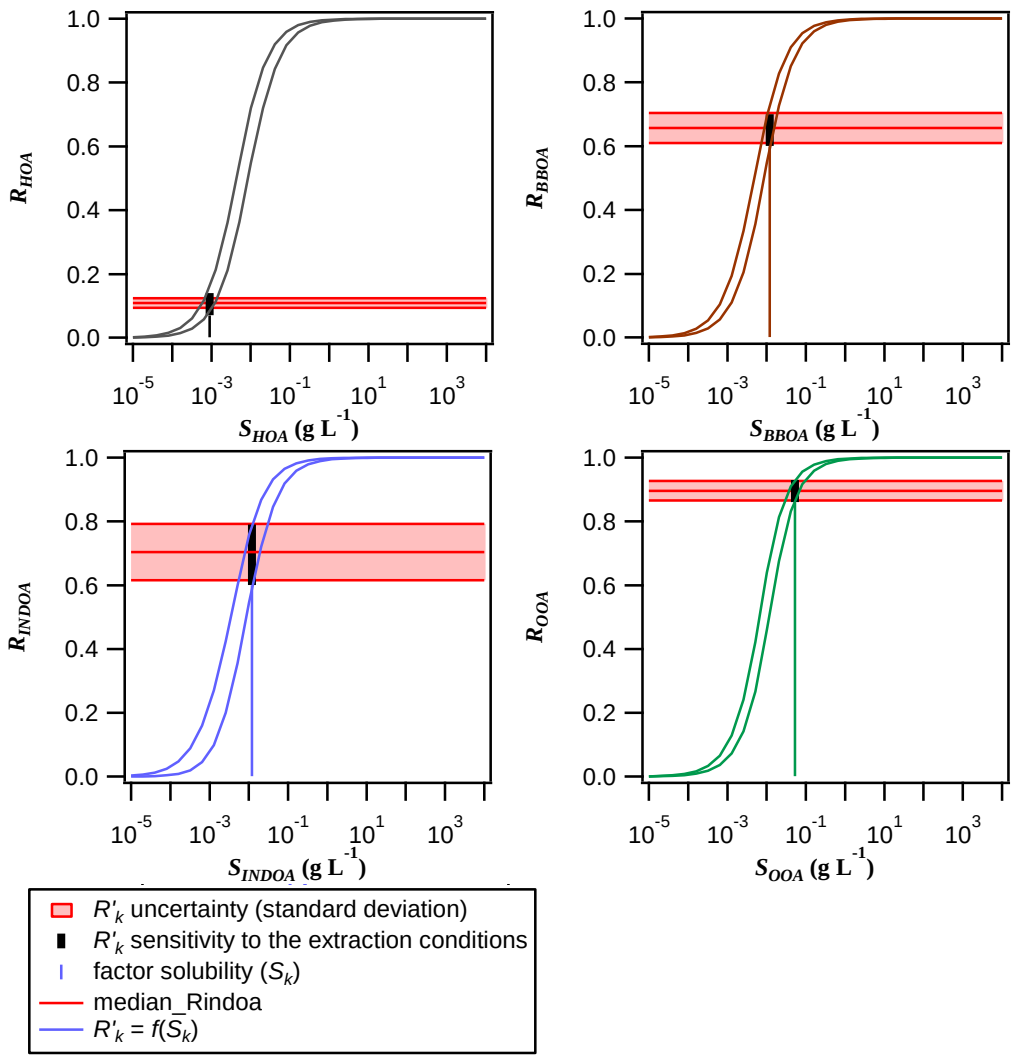


Figure D2. Sensitivity of the calculated factor recoveries R'_K to the factor solubility S_K (g L^{-1}). Vertical lines define the factor solubility calculated from the median factor recoveries (R'_K, horizontal lines) determined in this work .

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 $fC_2H_4O_2^+$ 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 $fC_2H_4O_2^+$ 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 (μg m^{-3}/μg 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 (μg m^{-3}/μg 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).

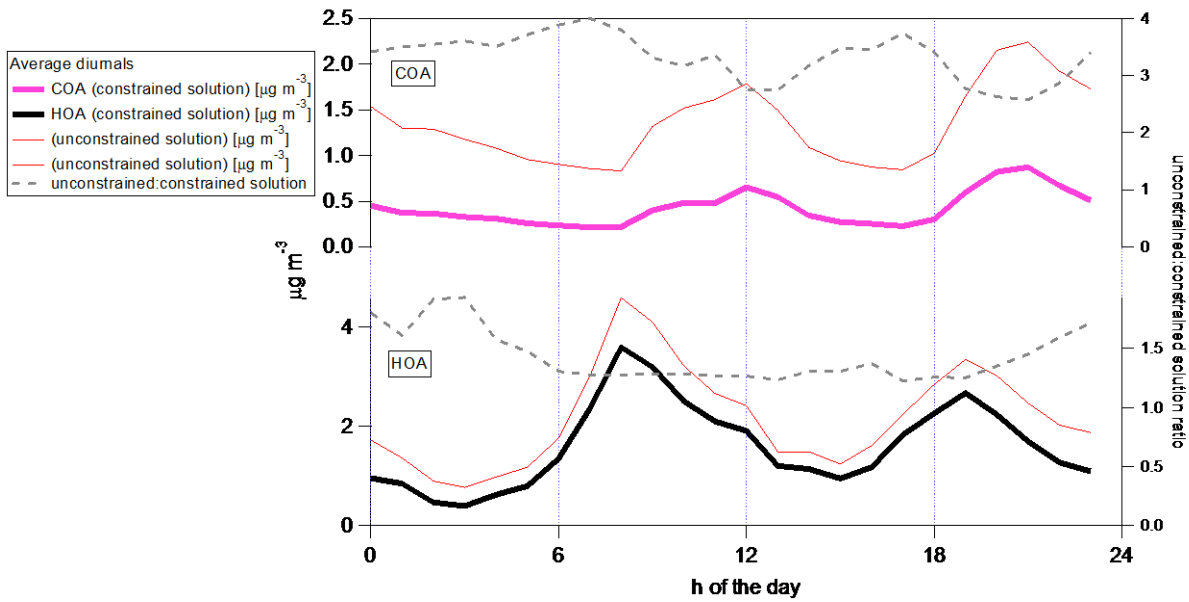


Fig. D3. Comparison of COA and HOA diurnal cycles from constrained and unconstrained PMF solutions.

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

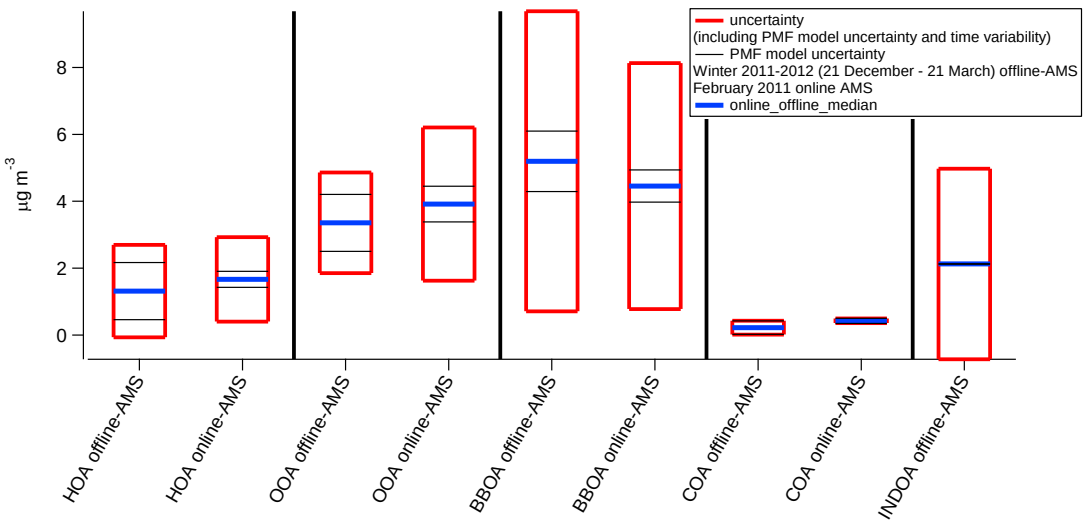
Table 1. Monitoring periods.

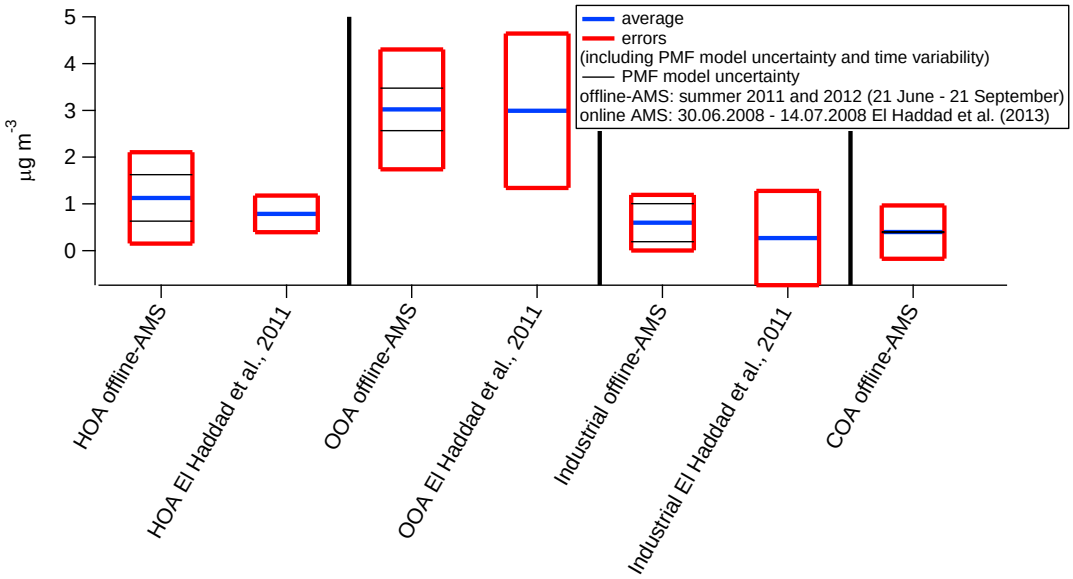
Online-AMS	Offline-AMS
28 January 2011 – 02 March 2011	30 July 2011 – 20 July 2012

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.

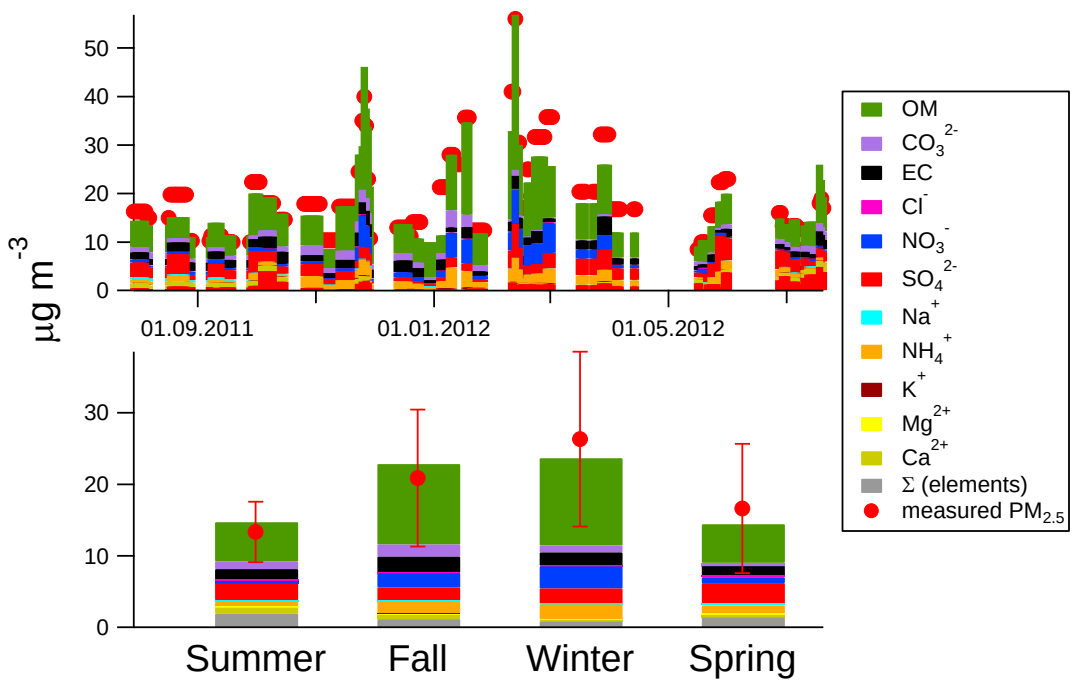
1. *Eqn. (11). In Bozzetti et al. (2017), there is a term TEOC (traffic emission OC) in the equation. Please discuss the rationale to replace TEOC with WSHOA/(RHOA*OM/OC) in current study. Also, WSOC is mentioned in Page 14 Line 17, but it is not clear how the WSOC is used in determining Rk.*

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.

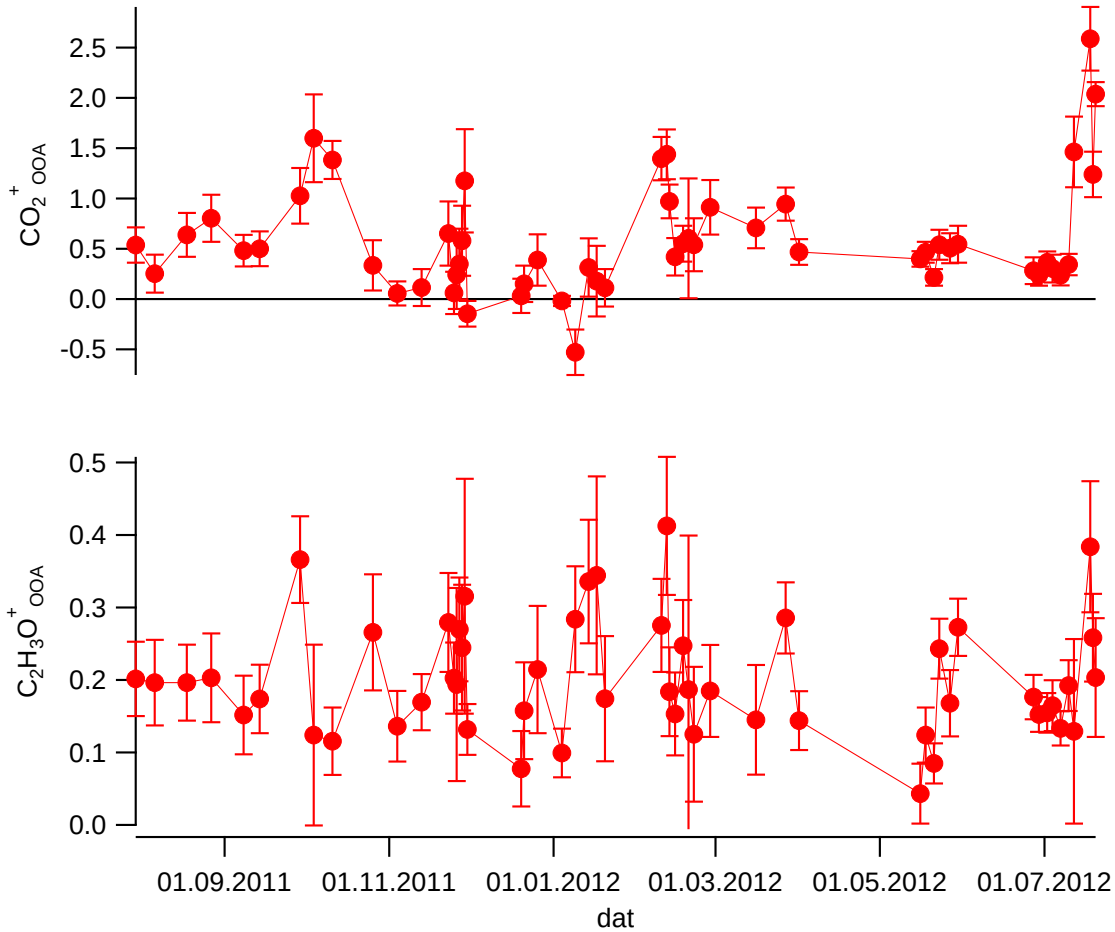
   We note that phenanthrene, pyrene, and fluoranthene together represent $9.6\%_{avg}$ of the PAHs mass quantified by GC-MS, indicating that PAHs are overwhelmingly emitted by BBOA.

   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)

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.