We would like to thank Reviewer# 1 and #2 for very helpful comments and suggestions. All comments and suggestions have been considered. Point by point responses to these comments are listed below. The line numbers correspond to the final revised document (with accepted track changes) unless otherwise stated.

Reviewer #1

General comments The manuscript presents UHRMS data of PM2.5 aerosol samples collected in central Amazonia during both the ’wet’ and ’dry’ seasons. Several tracer compounds corresponding to sources of biogenic and anthropogenic organic aerosol (OA) have been tentatively identified and UHRMS visualisation tools such as the Kendrick mass defect, Van Krevelen diagrams etc have been plotted to obtain further information of the differences in molecular composition between samples. The work presented is very interesting, in particular the demonstrated change in the OA chemical composition with increasing number of incident fires. As the authors note, higher time resolution filter collection would have led to a better understanding of the various factors affecting aerosol sources and formation at their sampling site. However, I do believe that this work offers new and interesting data. The manuscript did however suffer from inaccuracies with the incorrect use of acronyms, spelling mistakes and wording, which is in some places, is rather difficult to understand. I recommend that this manuscript should be published but only after the manuscript has been thoroughly checked for errors and the comments below, particularly specific comments 1 and 2 which are imperative to the work, have been addressed.

Specific comments 1. Lines 259 - 272. It appears that the identification of all of the tracer compounds with the exception of IEPOX (which is mentioned in the experimental section) was confirmed using only UHRMS via MS2. As stated, UHRMS does not differentiate between structural isomers. The authors should be able to confirm if the tracer compounds are present in the samples through the comparison of the ion fragmentation patterns to the literature or authentic standards of the tracer compounds, providing the fragmentation patterns are not too ’messy’ (i.e. multiple fragmentations of different structural isomers). I suspect the fragmentation data of m/z 161.0456 consisting of four tracer compounds and possibly other structural isomers may be particularly difficult to interpret and this should be mentioned in the manuscript. In addition, I don’t believe the authors can attribute the entire ion abundance of m/z 203.05611 to 3-MBTCA, unless the fragmentation data shows no indication of any other possible structural isomers. Finally, how do the authors know that other possible structural isomers are not largely contributing to the ion abundance of the other tracer compounds (i.e. $C_6H_5NO_4$ etc)? The authors need to provide more justification/evidence for the identification of these compounds and the use of their ion abundances in Figure 4.

We would like to mention that we were very cautious in interpreting direct infusion results. For instance, we initially stressed the fact that ‘*The structural or isomeric information is not directly obtained from the direct infusion analysis….*’ (lines 259-260 in the original text). We also stated that *‘It must be noted that due to competitive ionisation of analytes in the direct infusion ESI analysis of the samples with a very complex matrix (i.e., aerosol extracts), the ion intensities do not directly reflect the concentration of the molecules in the sample; therefore, data shown in this work is semi-quantitative’* (lines 157-160). *Moreover, we stated that ‘Direct infusion analysis suffers from competitive ionisation in the complex matrices and thus comparing ion intensities across samples has* *to be done with caution'* (lines 275-277). It must be noted that there is a large number of publications indicating and justifying the use of the direct infusion analysis for semi-quantitative purposes (see review by Nizkorodov et al., 2011). While the term "nitrophenols" does include all possible isomers, we agree that other compounds may contribute to a molecular formula assigned as 3-MBTCA. As confirmed by the LC/MS analysis of selected samples the compound assigned to $C_8H_{12}O_6$ molecular formula corresponds to MBTCA. However, in the revised version of the manuscript we emphasised again that molecular formula assigned as 3-MBTCA may also include other compounds: *'Moreover, other compounds with similar molecular composition present in the aerosol matrix may also contribute to the ion intensities of the discussed above compounds.'* We also replaced 'MBTCA' by a molecular formula $C_8H_{12}O_6$ in the Fig 4.

2. Were the IOP1 and IOP2 samples analysed on the same day, or was the detector variation of the UHRMS monitored during the analysis period? The UHRMS will vary in sensitivity. Running samples days or weeks apart may result in a variation in the amount of species observed due to fluctuations in the UHRMS sensitivity (e.g. as the mass spectrometer becomes 'dirty', the detector sensitivity will decrease, affecting the ion intensity and subsequently the amount of species observed). This is particularly important in Figures 2 and 4 where the molecular formulae and ion abundances, respectively, are compared. If the samples were not analysed at the same time or if the detector sensitivity was not monitored, the authors would not be able to compare the ion abundance of the tracer compounds as shown in Figure 4. This is also likely to affect the comparison of the molecular formulae in Figure 2. The work presented here would then be only qualitative (rather than semi-quantitative). Were any attempts made to account for variations in detector sensitivity?

The instrument was routinely calibrated before the analysis. It must be noted that in the current study we used a nanoESI source where each sample is processed using a separate ESI tip and nozzle, so there is no carryover between samples. All samples were analysed in a random order and within 48-hours after extraction (to minimise possible methylation; therefore, the observed differences could not be attributed to the instrument contamination.

3. Line 217 states that the number of molecular formulae of species containing CHO increased by ∼ 20% from IOP1 to IOP2, but Figure 2 shows that this increase is within the standard deviation of the three replicate measurements. Please can the authors state in line 217 that this ∼ 20% difference is based on the average number of molecular formulae. Can the authors demonstrate that these differences are statistically significant? Do the ratios of the compounds classes differ between wet and dry season?

The t-test demonstrated that there is a significant difference for individual subgroups (e.g., CHO) between two compared seasons ($p=0.0092$ and $p=0.00007$).

As requested the following statement has been added to the text: '*The Student's t-test showed that the observed difference for CHO ($p=0.0092$) and CHON ($p=0.00007$) subgroups between two seasons is statistically significant.'*

4. The experimental section needs to be separated into sections to make it clearer. Currently, the direct infusion flow rate follows the LC-MS parameters after UHRMS has already been discussed (line 140). Sub-headings such as 'LC-MS analysis', 'ESIUHRMS analysis' and 'data processing' would make the experimental section easier to understand.

As suggested the subheadings ‘*Direct infusion UHRMS analysis*’ and ‘*LC-MS analysis*‘ have been added to the experimental section.

5. Line 157, competitive ionisation is not the only reason why ion intensities do not re- flect the concentration of the compounds when using ESI. The ionisation efficiency of species will also vastly differ depending largely on their chemical structure and composition (see Oss et al (2010)). Please can the authors acknowledge this in the manuscript?

This phenomenon, which is true for all mass spectrometry ionisation techniques, including hard ionisation techniques such as electron ionisation, is already covered by the matrix statement in the same sentence. As suggested by the reviewer, Oss et al reference has been added to the text.

6. The authors use very strict molecular formulae constraints, along with other parameters such as O/C ratio ≥ 1.3, 0.3 ≤ H:C ratio etc. This is likely to remove a large proportion of the observed peaks from further analysis. I understand why the authors have done this, but please can they include the percentage of the observed mass spectral peaks which are assigned molecular formulae in the manuscript (i.e. 57 % of the observed mass spectral peaks were assigned a molecular formulae using the constraints; as shown in Woznaik et al 2008)?

As suggested, we added a table SI2 (analogous to that in Wozniak et al., 2008) showing % occurrence of formula groups to all peaks assigned molecular formulae in the mass spectra during the two sampling periods.

7. The authors refer to Kourtchev et al (2013) and (2015) for further details regarding the processing of the UHRMS data. From these papers, it appears that the background ions are subtracted from samples. If so, please can the authors include this in this manuscript? This is an important part of the data processing and needs to be mentioned in this manuscript too.

As suggested, the following statement has been added to the text: ‘*The background spectra obtained from the procedural blanks were also processed using the rules mentioned above. The formulae lists of the background spectra were subtracted from those of the ambient sample and only formulae with a sample/blank peak intensity ratio ≥ 10 were retained*’. Lines 179-182 (see revised text)

8. Line 228-229. The authors state that the daytime %RH during IOP1 was 89%. This seems a little high based on the data shown in Figure SI2. Please state in the manuscript whether this is the average %RH or maximum. Also, have the authors calculated the %RH only during the filter sampling time periods? Given that the authors are justifying why there is increased number of organonitrates in the IOP1 samples the %RH should only refer to the filter sampling time periods.

Yes, the values correspond to the maximum and the minimum RH during the filter sampling periods. This has been now clarified in the text.

9. Line 251 states that wet deposition of aged or processed aerosol cannot be only reason for the observed differences in OSc. If the aerosol had wet deposited, how would the authors of sampled this?

Unfortunately, we do not fully understand this remark but we assume that this comment is a misunderstanding: we did not collect any precipitation or aerosol deposited due to wet deposition but only particles that were not scavenged by cloud or rain droplets.

10. Can the authors add the isoprene gas-phase measurements into Figure 4? Use of replicate figures with 'a' (benzene overlaid) and 'b' (isoprene overlaid) may prevent the data from looking too busy.

We intentionally did not show isoprene data in this manuscript because it will be published as an independent work.

11. Line 443. Can the authors give more justification as to why they think the observation of these highly oxygenated species are likely to be associated with molecules produced through homogenous photochemical ageing reactions? Compounds with ~ 10 oxygen atoms are likely to be of relatively low volatility residing mainly in the particulate phase. Heterogeneous reactions would seem likely here.

In this sentence we are referring to a literature study which also observed highly oxygenated species and suggested that they could be produced through homogenous photochemical ageing reactions. The exact formation mechanism for these species is still highly debatable as for most of them there are no chemical standards. We agree that heterogeneous reactions could possibly also lead to formation of such compounds. To clarify this, we removed the word 'homogeneous' from the statement.

12. Can the authors show the data points from IOP1 and IOP2 in different colours/shapes in Figure SI3?

As suggested different markers were used for IOP1 and IOP2 data points in Fig SI3.

13. Can the authors draw the categorises/sources of aerosol (i.e. SV-OOA, BBOA etc) onto Figure 3 as shown in Kroll et al (2011). This will make the data much easier to visualise when describing in the results section.

The main emphasis of the figure 3 is to show the shift in the carbon oxidation state from dry to wet seasons in organic aerosol *throughout the whole mass range*. Addition of the categories would make the plot very busy and difficult to visualise the shift in the OSc. However, as suggested by the reviewer, we added these categories to another carbon oxidation state plot in the revised Fig. 7 and updated the figure caption accordingly.

Reference - Oss et al., (2010) Anal. Chem. 82. 2865-2872

Technical corrections 1. OH should be written as OH or 'OH radical'

Corrected

2. Line 67, the use of 'participate in heterogeneous chemical reactions in the atmosphere' doesn't make an awful lot of sense in this sentence, re-word or remove.

The word 'heterogeneous' has been removed

3. Line 70, for the most part, precursor and oxidant types will determine the composition of SOA formed, which will in turn determine the light absorbing properties of the SOA. Remove ’precursor and oxidant types’ from this sentence or re-word.

As suggested by the reviewer ’precursor and oxidant types’ has been removed from this statement.

4. NOx should be written as NOx (use of subscript)

Corrected

5. Line 75, remove ’for example’, this sentence does not follow the above.

We disagree with this comment, reaction between anthropogenic nitrogen oxides (NO_x) and sulfur dioxide (SO_2) with a range of BVOCs leading to formation of organic nitrates is an example of anthropogenic/biogenic interactions discussed in the above sentence.

6. Line 89, ’UHRMS have a mass resolution...’ should be, ’UHRMS has a mass resolution....’

We are discussing several techniques here (e.g. Fourier transform ion cyclotron resonance MS and Orbitrap MS). To address the reviewer’s comment, we replaced ‘MS’ with ‘mass spectrometers’.

7. Line 97, need a comma after Shanghai.

Corrected

8. Line 98, this sentence would read better as; ’UHRMS has proven to be extremely useful or a value tool/technique for assessing.....’

As suggested, the sentence has been changed to ‘*UHRMS has proven to be extremely useful in assessing chemical properties of the SOA’*

9. Line 104, Martin et al 2015 is not in the references, do you mean Martin et al 2016?

Corrected

10. Line 105, the T3 site is 69.4 km from Manaus (Martin et al 2016), not 70 km. Change to ∼ 69 km or 69.4 km.

Corrected

11. Line 113, could you make this a little clearer? ’...passed over the single large city (Manaus)’

Corrected

12. Supplementary material Table SI1, is there a reason why the time is reported as, for example, 7H47? If not, change column header to ’Time (UTC, HH:MM)’ and remove ’H’. 13.

Corrected

Line 121, this sentence reads as if the sampling flow rate changes during sample collection. Re-word.

The sentence has been change to *‘The airflow through the sampler was approximately 10 L min^{-1}’*.

14. Line 123, how were the samples stored at -4 ËŽC?

The sentence has been extended to *'…and stored in the freezer at −4°C until analysis.'*

15. Line 127, 'optima' is the name of the product not the grade, the grade is LC-MS. Change.

Optima is a trademark name of the Fisher LC/MS grade solvents which is different from the regular LC/MS grade solvents available on the market. The 'TM' and 'LC/MS' have been added to the revised version of the manuscript.

16. Line 128, how was the sample reduced to a volume of 200 µL, via a nitrogen line or evaporator? If the latter, please give details of manufacturer etc.

The sentence has been extended to *'..using a nitrogen line'*.

17. Line 154, define CID

Corrected

18. Line 154, 'MSMS' should be written as MS/MS or MS2

Corrected

19. Line 156, include the word 'time' in 'chromatographic elution' (i.e. chromatographic elution time or retention time)

Added

20. Line 184, define E/N before abbreviating

'a field density ratio' has been added before *'E/N'*

21. Supplementary material SI1, explain what 'MP14-06' etc (displayed on the figures) refers to in the figure caption.

As suggested, this has been now clarified in the figure caption: *'72 h back air mass history ('footprints') arriving at the T3 station for the periods of the analysed filters (labelled as e.g., MP14-06, MP14-16, MP14-17).'* Lines 37-39

22. Line 209, states that the majority of ions were associated with molecules less than 500 Da but Figure 1 only goes up to m/z 500. Either show the full m/z scan range in Figure 1 or re-phase Line 209 (e.g. the majority of species were observed between m/z 100 to 400).

The line 209 (in the original text) has been rephrased to *'……were associated with molecules below 500 Da although the measured mass goes up to 900 Da.'*

23. Line 212, 'fragile compounds'. Why are some compounds fragile? Please expand

The sentence has been extended to *'…(e.g. highly oxygenated compounds)'*.

24. Line 213 is difficult to read. Re-word.

As suggested the sentence has been rephrased to: '*The largest group of identified molecular formulae in all samples were attributed to molecules containing CHO atoms only (1051±141 formulae during IOP2 and 820±139 during IOP1), followed by CHON (537±71 during IOP2 and 329±71 during IOP1), CHOS (183±34 during IOP2 and 137±31 during IOP1) and CHONS (37±11 during IOP2 and 28±10 during IOP1) (Fig. 2).*'

25. Figure 2; include 'IOP1' and 'IOP2' next to 'wet' and 'dry' season respectively in Figure 1 or the opposite in the figure caption.

Corrected

26. Supplementary material Table SI1, make clear which samples are from wet and dry season.

This has been now clarified in the Table SI1 footnote*: The samples MP14_06 to MP14_28 correspond to 'wet' (IOP1) period and MP14_128 to MP14_153 to 'dry' (IOP2) period.*

27. Line 223, Table SI1, NOy should be written as NOy (use of subscript).

Corrected

28. Line 227 and Figure SI2 caption, define 'RH'.

Defined

29. Figure SI2 caption, what is 'ARM'? Define. Should this be in the references?

The ARM has been defined and the link to the website is provided: '*Figure SI2. Relative humidity (RH) at the T3 sampling site during (a) IOP1 and (b) IOP2 The arrows indicate sample collection periods. Atmospheric Radiation Measurement (ARM) data source http://www.archive.arm.gov.* '

30. Figure SI2, what are the dashed lines displaying? Explain in caption.

The explanation has been added to the figure legend: *'The continuous dashed line indicates the lowest and highest RH vales during both seasons'*

31. Line 228, use of 'IOP1' then 'wet season'. Please use either wet and dry or IOP1 and IOP2.

For consistency we have added IOP2 to this sentence: *'In this respect, while night time maximum RH during both filter sampling periods was very similar (~90%), day-time RH during IOP1 was higher (89%) compared to that from the IOP2 period (66%) (Fig. SI2).'*

32. Line 230 and elsewhere, 'OSc' should be written as 'OSc'

We do not understand this comment.

33. Figure 3, please give a starting number of carbon atoms on the x-axis or start from zero.

Done

34. Line 256, move reference to the end of the sentence.

Corrected

35. Line 273, SOx should be written as SOx (use of subscript).

Corrected

36. Line 390, change to ’a reduced number of ’ or ’decreased number of’

As suggested we replaced ‘*reduced*’ by ‘*decreased*’

37. Lines 400 and 401 are difficult to understand. Be more precise (e.g. ....difference in OSc is more pronounced with compounds containing more than 7 carbon atoms). ’Affected ions’? Re-word.

Yes, in this sentence we are discussing figure 7 and thus OSc differences associated with the carbon atoms in the molecular formulae. For clarity we expanded this sentence to *‘The difference in OSc becomes even more pronounced with increased numbers of carbons (e.g. >7 carbon atoms) in the detected molecular formulae’.*

38. Line 375, change ’nitroartomatic’ to ’nitro-aromatic’

Done

39. Line 376, ’overplayed’?

Changed to ‘overlaid’

Reviewer #2

General comments: This manuscript deals with the molecular characterization of PM2.5 aerosol collected in Manaus, Brazil, which is impacted by regional biomass burning, mainly during the dry season, and anthropogenic pollution from the city. Advanced analytical ultra-high resolution MS-based tools are applied, which allow a comprehensive MS data evaluation and identification of molecular formulae. Several comments were already formulated in a first review, with which I could agree. This review will therefore be limited to additional comments. The manuscript contains indeed interesting and novel data on Amazonian fine aerosol, which could be elaborated and is worth publishing after suitable revision. What I miss in the manuscript is a comparison with previous studies dealing with the detailed molecular characterization of Amazonian fine aerosol.

A first example: the 2-methyltetrols have been measured in several studies (e.g., Claeys et al., ACP 10, 9319-9331, 2010); the highest levels were observed during the dry period which is characterized by biomass burning (and higher particle concentrations of sulfuric acid). This observation is in agreement with the results obtained in the present study, taking into account that 2-methyltetrol sulfates were converted to 2-methyltetrols during the GC/MS procedure with prior trimethylsilylation?

In the current study we concentrated on direct infusion mass spectrometry analysis and therefore we intentionally limited our comparison mainly to the literature that employed similar techniques. As suggested by the reviewer comparison to previous studies by Claeys et al. (2010) has been added to the text: *‘This is also in agreement with previous studies from Amazon where the highest levels of 2-methyltetrols were observed during the dry period which was characterised by biomass burning* *(and higher particle concentrations of sulfuric acid) (Claeys et al., 2010). Considering that Claeys et al (2010) employed alternative GC/MS procedure with prior trimethylsylilation, 2-methyltetrol sulfates were converted to 2-methyltetrols and not detectable as separate OS compounds.'* Lines 339-344 (see revised text)

A second example: A study on Amazonian biomass burning aerosol (Claeys et al., Environ. Chem., 9, 273-284, 2012) using LC/MS led to the molecular characterization of several strongly UV-absorbing nitro-aromatic compounds, with 4-nitrocatechol and isomeric methyl-nitrocatechols being the most abundant ones. Nitrocatechols are mentioned in the current manuscript but no mention is made of the methyl-nitrocatechols, which are very important markers for biomass burning secondary organic aerosol (SOA), formed from m-cresol emitted during the fires.

As suggested discussion of methyl-nitrocatechols has been added to the text: '*… and methyl-nitrocatechols ($C_7H_7NO_4$, m/z 168.03023) are important markers for biomass burning OA, formed from m-cresol emitted during biomass burning (Iinuma et al., 2010)*' Lines 283-285

In addition, also several biogenic SOA markers were identified in the study of Claeys et al. (2012), including MBTCA, terebic acid, terpenylic acid, 2-hydroxyterpenylic acid, and azelaic acid.

Please see our response below.

Specific comments Lines 129-131: Part of the samples was used for LC/MS analysis but no LC/MS results are presented in the current manuscript. It would be very relevant to provide LC/MS results and as such support molecular assignments. It would also be relevant to see whether the major compounds found in LC/MS correspond to the major ones with the semi-quantitative direct infusion approach used in the present work. Figure 1: The base peak in panel (a) is at m/z 171 (terpenylic acid?), but this ion is not discussed in the manuscript. Has this ion been assigned? There are also other abundant ions of the CHO type in the region below m/z 200 which merit attention, such as m/z 157 (terebic acid?) and m/z 187 (2-hydroxyterpenylic acid, or azelaic acid?), and are likely biogenic SOA markers.

To be able to compare mass spectral intensities from different aerosol samples, as well as to minimise matrix effect, we aimed to have similar aerosol loading of the sample extracts for the direct infusion analysis. Since we aimed for approximately 0.3 µg μL^{-1} of particulate matter in each sample extract, there was enough aerosol material only for direct infusion analysis for most of the samples. Only a few aerosol samples had sufficient loading for both direct infusion and LC/MS analyses. Therefore, the later technique was only used to confirm the assignments of a few marker compounds observed in direct infusion analysis. This now has been clarified in the text:

'Depending on the aerosol loading of the analysed samples a part (1/2 to whole) of the filter was extracted in methanol (Optima R LC/MS grade, Fisher Scientific) in a chilled ice slurry, filtered through a Teflon filter (0.2 µm, ISODiscTM Supelco) and reduced by volume using a nitrogen line to achieve approximately 0.3 µg of aerosol per µL methanol. Several samples with the highest aerosol loading were divided into two parts for both direct infusion and LC/MS analyses, while the samples with the lowest loading were only analysed using direct infusion analysis.' Lines 126-132

With direct infusion analysis we identified more than a thousand molecular formulae, therefore the main emphasis was placed on a bulk molecular composition of the OA rather than identification of all possible marker compounds and which would be a study different to the one presented here. The discussion of the individual compounds was limited to a few known marker compounds that corresponded to the most intense ions.

The base peak in the panel (a) at *m/z* 173.0454 corresponds to $C_7H_{10}O_5$, which is different from that of terpenylic acid. Unfortunately, this molecule was not clearly identified by the LC/MS analysis. With regards to the other abundant ions of the CHO type in the region below *m/z* 200, an ion at *m/z* 187.0612 corresponds to a deprotonated molecular formula $C_8H_{12}O_5$, which is neither 2-hydroxyterpenylic acid nor azelaic acid. An ion at *m/z* 157.01425 corresponds to a deprotonated molecular formula $C_6H_6O_5$, which is also different from terebic acid. Unfortunately, due to the absence of standards for the above mentioned compounds, the discussion of these molecular formula would be highly speculative.

Lines 209-229: As already mentioned above, LC/MS results would be very useful to support the molecular assignments, more useful in my opinion than MS/MS data, which in the case of 2-methyltetrolsulfates provide limited structural information (only the bisulfate anion). Quite some emphasis is given to the number of molecular formulae containing CHO, CHON, CHOS, and CHONS. More emphasis could be given to the molecular characterization of the major species, taking into account that LC/MS analysis has been performed and reference can be made to the literature. This type of information will be of great interest to readers dealing with molecular characterization.

Please see our response above.

Lines 279-286: Here, the origin of benzene is discussed and it is argued that benzene has mainly an anthropogenic origin because it correlates well with CO. It is not very clear what is meant by “anthropogenic origin”. Biomass burning for domestic purposes (e.g., cooking) in urban locations can also be regarded as an anthropogenic activity and this must be clarified in the manuscript. Benzene could very well have mainly a biomass burning origin. More detailed insights could be obtained by measuring other aromatic compounds, such as cresols, and acetonitrile, which are characteristic for biomass burning; a good correlation between benzene and cresols/acetonitrile would point to a biomass burning origin. A differentiation between an anthropogenic and a tropical biomass burning origin cannot easily be made and will remain problematic. See the following article and references cited therein: Iinuma et al., Environ. Sci. Technol. 2010, 44, 8453–8459.

We agree that the correlation between benzene and cresols/acetonitrile would provide more information on the aerosol sources; unfortunately, cresols/acetonitrile data is not available for our study. In Manaus for heating and cooking purposes people mainly use natural gas; therefore, a significant contribution from these activities to the biomass burning OA at the site is highly unlikely. This statement has been added to the text: *‘In Manaus natural gas is mainly used for heating and cooking and therefore, the contribution from these activities to biomass burning OA at our site is highly unlikely’.* Lines 310-312

Lines 371-373: It would be relevant to mention 4-nitrocatechol and isomeric methylnitrocatechols in the group of nitroaromatic compounds, since they are characteristic of biomass burning SOA; see Iinuma et al., Environ. Sci. Technol. 2010, 44, 8453– 8459.

As suggested by the reviewer, the following statement has been added to the text: *'Nitro-aromatic compounds, such as nitrophenols (DBE=5) and N-heterocyclic compounds including 4-nitrocatechol and isomeric methyl-nitrocatechols are often observed in the OA from the biomass burning sources (Kitanovski et al., 2012a,b; Iinuma et al., 2010) and have been suggested as potential contributors to light absorption by brown carbon (Laskin et al., 2015).'* Lines 405-409

Lines 393-396: In addition, the CHON molecules identified by LC/MS in biomass burning OA from Amazonia showed O/C ratios below 0.7, i.e., 4-nitrocatechol (C6H5O4N; O/C = 0.67), isomeric methyl-nitrocatechols (C7H7O4N; O/C = 0.57), and isomeric dimethyl-nitrocatechols (C8H9O4N; O/C = 0.50). Ref: Claeys et al., Environ. Chem. 9273-284, 2012.

As suggested, the following statement has been added to the text: *'In addition, the CHON molecules identified by LC/MS in biomass burning OA from Amazonia showed O/C ratios below 0.7, i.e., 4-nitrocatechol ($C_6H_5NO_4$; O/C = 0.67), isomeric methyl-nitrocatechols ($C_7H_7NO_4$; O/C = 0.57), and isomeric dimethyl-nitrocatechols ($C_8H_9NO_4$; O/C = 0.50) (Claeys et al., 2012).'* Lines 434-437

Lines 436-437: Species with molecular formulae C5H10O7S (m/z 213) could also be due to organosulfates formed from the green leaf volatiles 2-E-pentenal, 2-Ehexenal, and 3-hexenal, and have recently been characterized as isomeric 3-sulfooxy2-hydroxypentanoic acid and 2-sulfooxy-3-hydroxypentanoic acid. Ref. Shalamzari et al., ACP 16, 7135-7148, 2016. See also the corresponding discussion document, where the issue is raised that C5H10O7S species could be oxidation products of isoprene.

The following statement has been added to the text: '*This molecular formula could also be associated with organosulfates (e.g., isomeric 3-sulfooxy-2-hydroxypentanoic acid and 2-sulfooxy-3-hydroxypentanoic acid) formed from the green leaf volatiles 2-E-pentenal, 2-E-hexenal, and 3-hexenal(Shalamzari et al., 2016)'* Lines 482-485

Figure 5: What do the large grey circles between m/z 120 – 240 represent? Please, explain in the legend of the figure and discuss in the main text.

As suggested by the reviewer the following explanation has been added to the text and the figure: '*The largest grey circles in Fig 5(a-c) correspond to the ions at m/z 133.01425 (with neutral molecular formula $C_4H_6O_5$), m/z 187.0612 ($C_8H_{12}O_5$), m/z 201.07685 ($C_9H_{14}O_5$), m/z 203.05611 ($C_8H_{12}O_6$), m/z 215.05611 ($C_9H_{12}O_6$) with DBE<6.'* Lines 367-370 and 881-883

FigureSI5: What do the large grey circles at around m/z180 and 280 represent in panel (a)? What do the large grey circles between m/z 140 and 180 represent in panels (b) and (c)? Please, explain in the legend of the figure and discuss in the main text.

As suggested by the reviewer the following explanation has been added to the text and the figure legend:

'The largest grey circles in Figure SI5a correspond to the ions at m/z 187.11357 with a neutral molecular formula $C_9H_{17}NO_3$ and m/z 281.26459 with a neutral molecular formula $C_{18}H_{35}NO$. The largest grey circles in Figure SI5 b and c correspond to the ions at m/z 154.0146, m/z 168.03023 and m/z 152.03532 with neutral molecular formulae $C_6H_5NO_4$, $C_7H_7NO_4$ and $C_7H_7NO_3$, respectively'. Lines 398-402 (main text) and 85-89 (SI).

Figure SI6: It is evident from these figures that CHO compounds are present at significant abundances in the natural background. What species do the large grey circles represent in panels (a–c)?

Figure SI6 shows exactly the same ion distribution as in Figure 5, but expressed using aromaticity index. The large circles in both figures correspond to the same molecular formula; therefore, to avoid repetition we added the explanation for these ions to only one of these figures.

Figure SI7: I wonder what the large yellow circles (panels (b) and (c)) between m/z 150 and 200 represent. Do they correspond to m/z 168 (C7H6NO4) compounds, due to isomeric methyl-nitrocatechols, which are expected to be very prominent and most abundant in the samples from the dry biomass burning period? Looking at panel (a) I wonder what the large grey circles around m/z 190 and 380 represent? Please, explain in the legend of the figure and discuss it in the main text.

As suggested by the reviewer the following explanation has been added to the figure legend: '*The largest grey circles in panel 'a' correspond to ions at m/z 186.11357 and m/z 280.26459 with neutral molecular formulae* $C_9H_{17}NO_3$ *and* $C_{18}H_{35}NO$*, respectively. The yellow circles in panels 'b' and 'c' correspond to the ions at m/z 154.0146, m/z 168.03023 and m/z 152.03532 with molecular formulae* $C_6H_5NO_4$*,* $C_7H_7NO_4$ *and* $C_7H_7NO_3$*, respectively, which are known biomass burning marker compounds (see discussion in the main text)'.* Lines 105-110 (SI)

I found Figures SI5, SI6 and SI6 the more interesting figures in the manuscript, but unfortunately they ended up in the supplement. Please, consider to include them in the main text, perhaps leaving out some other figures and putting some emphasis on methyl-nitrocatechols, specific SOA markers for biomass burning. Other interesting (but less abundant) biomass burning SOA markers are m/z 182 (C8H8NO4) compounds, corresponding to isomeric dimethyl-nitrocatechols.

A very large diversity of the data was produced in this work resulting in a large number of figures, so we had to be very selective which figures could be kept in the man text. Our justification was based on the fact that some of the figures in the SI would be rather challenging for a general reader with little mass spectrometry background. Therefore, we prefer to keep SI5 and SI6 in the SI.

As suggested by the reviewer the following explanation has been added to the text: *'It is worth mentioning that aerosol samples affected by biomass burning contained another interesting ion at m/z 182.04588 with a neutral molecular formula* $C_8H_9NO_4$*, possibly corresponding to biomass burning OA markers isomeric dimethyl-nitrocatechols (Kahnt et al., 2013)*'. Lines 409-412 (main text)

Lines 473 – 477: Here, the authors indicate that future work is needed to better understand the quantitative contributions of the various factors to the aerosol composition at the T3 site and they suggest to analyze samples with higher sampling resolution. A better approach would be to also measure specific marker compounds more quantitatively by LC/MS or other methods, including biogenic SOA markers, and primary and secondary biomass burning markers, and apply a receptor modelling technique. See, for example, the recent study by de Oliveira Alves et al. (Atmos. Environ., 120, 277285, 2015), where for a site in western Amazonia, i.e., Porto Velho, a distinction could be made between contributions from biomass burning, fossil fuel combustion and a mixed source to the PM10 mass.

The following statement has been added to the text: *'The analysis of aerosol samples with higher sampling resolution or quantifying specific marker compounds and applying receptor modelling techniques (Alves et al., 2015) would allow separating these sources in more detail and thus improve understanding of the aerosol formation sources at the site.'* Lines 519-523 (main text)

Molecular composition of organic aerosols in central Amazonia: an ultra-high resolution mass spectrometry study

I. Kourtchev[1]*, R.H.M. Godoi[2], S. Connors[1], J.G. Levine[3], A. Archibald[1,4], A.F.L. Godoi[2], S.L. Paralovo[2], C.G.G. Barbosa[2], R.A.F. Souza[5], A.O. Manzi[6], R. Seco[7], S. Sjostedt[8], J.-H. Park[9], A. Guenther[7,10], S. Kim[7], J. Smith[11,12], S.T. Martin[13,14] and M. Kalberer[1]*

[1]Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK

[2]Environmental Engineering Department, Federal University of Parana, Curitiba, Brazil

[3]School of Geography Earth & Environmental Sciences, University of Birmingham, Birmingham, B15 2TT, UK

[4]NCAS climate, University of Cambridge, Cambridge, CB2 1EW, UK

[5]State University of Amazonas, Av. Darcy Vergas, 1200, 69065-020, Manaus-AM, Brazil

[6]Instituto Nacional de Pesquisas da Amazônia (INPA), Clima e Ambiente (CLIAMB), Manaus-AM, Brazil

[7]Department of Earth System Science, University of California, Irvine CA 92697, USA

[8]NOAA ESRL Chemical Sciences Division, Boulder CO, USA.

[9]National Institute of Environmental Research, Republic of Korea

[10]Pacific Northwest National Laboratory, Richland WA, USA

[11]Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder CO, USA
[12]Dept of Chemistry, University of California, Irvine CA, USA

[13] Sch. Eng. & Appl. Sci., Harvard University, Cambridge, MA 02138 USA

[14]Dep. Earth & Planetary Sciences, Harvard University, Cambridge, MA 02138 USA

*Corresponding authors: I. Kourtchev (ink22@cam.ac.uk) and M. Kalberer (mk594@cam.ac.uk)

Abstract

The Amazon basin plays key role in atmospheric chemistry, biodiversity and climate change. In this study we applied nanoelectrospray (nanoESI) ultrahigh resolution mass spectrometry (UHR-MS) for the analysis of the organic fraction of $PM_{2.5}$ aerosol samples collected during dry and wet seasons at a site in central Amazonia receiving background air masses, biomass burning and urban pollution. Comprehensive mass spectral data evaluation methods (e.g., Kendrick Mass Defect, Van Krevelen diagrams, carbon oxidation state and aromaticity equivalent) were used to identify compound classes and mass distributions of the detected species. Nitrogen and/or sulfur containing organic species contributed up to 60% of the total identified number of formulae. A large number of molecular formulae in organic aerosol (OA) were attributed to later-generation nitrogen- and sulfur-containing oxidation products, suggesting that OA composition is affected by biomass burning and other, potentially anthropogenic, sources. Isoprene derived organo sulfate (IEPOX-OS) was found as the most dominant ion in most of the analysed samples and strongly followed the concentration trends of the gas-phase anthropogenic tracers confirming its mixed anthropogenic-biogenic origin. The presence of oxidised aromatic and nitro-aromatic compounds in the samples suggested a strong influence from biomass burning especially during the dry period. Aerosol samples from the dry period and under enhanced biomass burning conditions contained a large number of molecules with high carbon oxidation state and an increased number of aromatic compounds compared to that from wet. The results of this work demonstrate that the studied site is influenced not only by biogenic emissions from forest but also by biomass burning and potentially other anthropogenic emissions from the neighboring urban environments.

Keywords: organic aerosol, ultra-high resolution mass spectrometry, molecular composition, IEPOX-OS, Amazon.

Introduction

The Amazon basin plays key role in atmospheric chemistry, biodiversity and climate change (Keller et al., 2009; Andrea et al., 2015). The Amazon rainforest is an important source of Biogenic Volatile Organic Compound (BVOC) emissions to the atmosphere (Greenberg et al., 2004; Alves et al., 2015), which give rise to secondary organic aerosol (SOA) through reaction with atmospheric oxidants (i.e. O_3, $OH\cdot$ and $NO_3^{\cdot}$) (e.g., Martin et al., 2010). SOA particles scatter and absorb solar and terrestrial radiation, influence cloud formation, participate in ~~heterogeneous~~ chemical reactions in the atmosphere, and thus are suggested to play an important role in climate change (Andreae and Crutzen, 1997; Haywood and Boucher, 2000; Hallquist et al., 2009; Pöschl et al., 2010). Aerosol optical properties, which govern the ability to absorb solar radiation, strongly depend on SOA composition~~, precursor and oxidant types~~ (Laskin et al., 2015). It has been shown that organic nitrates, nitrooxy-organosulfates and organic sulfates may contribute to light absorption by SOA (e.g., Song et al., 2013; Jacobson, 1999; Lu et al., 2011; Laskin et al., 2015). Chemical interactions between anthropogenic and biogenic aerosol precursors can play a significant role in the formation of SOA (Goldstein et al., 2009; Hoyle et al., 2011; Kleinman et al., 2015). For example, anthropogenic nitrogen oxides (NO_x) and sulfur dioxide (SO_2) are shown to react with a range of BVOCs leading to formation of organic nitrates (e.g., Roberts, 1990; Day et al., 2010; Fry et al., 2014), nitroxy-organosulfates and organosulfates (Surratt et al., 2008; Budisulistiorini et al., 2015). Much remains to be explored in terms of the molecular diversity of these compounds in the atmosphere.

Formatted: Subscript

A comprehensive knowledge of aerosol molecular composition, which in turn leads to better understanding of aerosol sources, is required for the development of effective air pollution mitigation strategies. However, identification of the organic aerosol composition, remains a major analytical challenge (Noziere et al., 2015). Organic aerosol is composed of thousands of organic compounds, which cover a wide range of physical and chemical properties (Goldstein and Galbally, 2007) making it difficult to find a single analytical technique for a detailed chemical analysis at the molecular level. Methods based on ultrahigh resolution mass spectrometry (UHRMS) have shown great potential in solving this longstanding problem. UHR mass spectrometers~~MS~~ (e.g., Fourier transform ion cyclotron resonance MS and Orbitrap MS) have ~~a~~ mass resolution power that is at least one order of magnitude higher (≥100 000) than conventional MS and high mass accuracy (<5 ppm) and thus, when coupled with soft ionisation techniques (e.g., electrospray ionisation (ESI)), can provide a detailed molecular composition of the organic aerosol (Nizkorodov et al., 2011, Noziere et al., 2015). Direct infusion ESI-UHRMS has been applied successfully for the analysis of aerosol samples from remote (e.g., boreal forest in Finland, Pico Island of the Azores archipelago), rural (e.g., Millbrook, USA; Harcum, USA; K-Puszta, Hungary) and urban (e.g., Cambridge, UK, Birmingham, UK, Cork, Ireland, Shanghai, China and Los Angeles, USA) locations (Wozniak et al., 2008; Schmitt-Kopplin et al., 2010; Kourtchev et al., 2013; 2014; Tao et al., 2014; Dzepina et al., 2015). UHRMS has proven to be extremely useful ~~This technique is extremely useful~~ in assessing chemical properties of the SOA.

The aim of this study was to investigate the detailed molecular composition of organic aerosol from a site that received air masses from a wide range of origins, including the background atmosphere of Amazonia, biomass burning and urban pollution plumes. The measurements were performed as a part of the *Observations and Modeling of the Green Ocean Amazon* (GoAmazon2014/5) campaign (Martin et al., 201~~5~~6). The location of the research site where aerosol was collected for this study is ~~~7~~69~~0~~ km downwind of Manaus (population 2 million), intersected background and polluted air with day-to-day variability in the position of the Manaus plume. The study designed served as a laboratory for investigating anthropogenic perturbations to biogenic processes and atmospheric chemistry.

Methods

Sampling site

Aerosol sampling was conducted at site "T3" of GoAmazon2014/5 located at -3.2133º and -60.5987º.°35'55 32'' W. The T3 site is located in the pasture area, ~2.5 km from the rainforest. The air masses arriving to the sampling site often passed over the single large city (Manaus) in the region. Detailed descriptions of the site and instrumentation are provided in Martin et al. (2015).

PM$_{2.5}$ aerosol samples were collected on 47 mm polycarbonate filters Nuclepore, using a Harvard impactor (Air Diagnostics, Harrison, ME, EUA) with flow rate of 10 L min^{-1} from 5 to 26 March 2014 and 5 Sept to 04 Oct of 2014, which were during Intensity Operating Periods 1 and 2 (IOP1 and IOP2) of GoAmazon2014/5, respectively, corresponding to the traditional periods of wet and dry seasons of Amazonia. The sampling durations are shown in the Table SI1. The airflow through the sampler was approximately 10 L min^{-1} ~~for about 24-36 h per sample~~. After collection, the aerosol samples were transferred into Petri dishes and stored in the freezer at −4°C ~~-~~until analysis.

Aerosol Sample Analysis

Fifteen samples, 5 from IOP2 and 10 from IOP1, were extracted and analysed using a procedure described elsewhere (Kourtchev et al., 2014; Kourtchev et al., 2015). Depending on the aerosol loading of the analysed samples, a part (1/2 to whole) of the filter was extracted in methanol (Optima TM LC/MS grade, Fisher Scientific) in a chilled ice slurry, filtered through a Teflon filter (0.2 µm, ISODiscTM Supelco) and reduced by volume using a nitrogen line to achieve approximately 0.3 µg of aerosol per µL methanol. Several samples with the highest aerosol loading were divided into two parts for both direct infusion and LC/MS analyses while the samples with the lowest loading were only analysed using direct infusion analysis.~~Briefly, ½ of filters were in methanol (Optima~~ TM LC/MS ~~grade, Fisher Scientific) in a chilled ice slurry, filtered through a Teflon filter (0.2 µm, ISODiscTM Supelco), and reduced by volume~~ using a nitrogen line ~~to approximately 200 µL. The sample was divided into two parts for direct infusion~~

~~and LC/MS analyses.~~ The LC/MS portion was further evaporated to 20 μL and diluted to 100 μL by aqueous solution of formic acid (0.1%). The final extracts were analysed as described in Kourtchev et al. (2013) using a high-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Fisher, Bremen, Germany) equipped with ESI and a TriVersa Nanomate robotic nanoflow chip-based ESI (Advion Biosciences, Ithaca NY, USA) sources. The Orbitrap MS was calibrated using an Ultramark 1621 solution (Sigma-Aldrich, UK). The mass accuracy of the instrument was below 1 ppm. The instrument mass resolution was 100 000 at *m/z* 400. The ion transmission settings were optimised using a mixture of camphor sulfonic acid (20 ng μL^{-1}) glutaric acid (30 ng μL^{-1}), and *cis*-pinonic acid (30 ng μL^{-1}) in methanol and Ultramark 1621 solution.

Direct infusion UHRMS analysis

The ionisation voltage and back pressure of the nanoESI direct infusion source were set at -1.4 kV and 0.8 psi, respectively. The inlet temperature was 200 ^{0}C and the sample flow rate was approximately 200–300 nL min^{-1}. The negative ionisation mass spectra were collected in three replicates at two mass ranges (*m/z* 100–650 and *m/z* 150–900) and processed using Xcalibur 3.1 software (Thermo Fischer Scientific Inc.). Similar to our preceding studies (Kourtchev et al., 2015) the average percentage of common peaks between analytical replicates was ~80%. This is also in agreement with literature reports for similar data analysis (Sleighter et al., 2012). The identification of IEPOX organosulfates was performed by comparing MS fragmentation patterns and chromatographic elution time with a synthesised IEPOX-OS standard which was provided by Dr Surratt from University of North Carolina. It must be noted that due to competitive ionisation of analytes in the direct infusion ESI analysis of the samples with a very complex matrix (i.e., aerosol extracts), the ion intensities do not directly reflect the concentration of the molecules in the sample (Oss et al., 2010); therefore, data shown in this work is semi-quantitative.

L~~/~~C-MS analysis

LC-MS ESI parameters were as follows: spray voltage -3.6 kV; capillary temperature 300 ^{0}C; sheath gas flow 10 arbitrary units, auxiliary gas flow 10; sweep gas flow rate 5; S-lens RF level 58 %. LC/(-)ESI-MS analysis was performed using an Accela system (Thermo Scientific, San Jose, USA) coupled with LTQ Orbitrap Velos MS and a T3 Atlantis C18 column (3 µm; 2.1 x 150 mm; Waters, Milford, USA). The sample extracts were injected at a flow rate of 200 µL min^{-1}. The mobile phases consisted of 0.1% formic acid (v/v) (A) and methanol (B). The applied gradient was as follows: 0–3 min 3% B, 3–25 min from 3 to 50% B (linear), 25–43 min from 50 to 90% B (linear), 43–48 min from 90 to 3% B (linear), and kept for 12 min at 3% B. The Collision Induced Dissociation (CID) settings for MS/MS analysis are reported in Kourtchev et al (2015). ~~The identification of IEPOX organosulfates was performed by comparing MS fragmentation patterns and chromatographic elution with a synthesised IEPOX-OS standard which was provided by Dr Surratt from University of North Carolina. It must be noted that due to competitive ionisation of analytes in the direct infusion ESI analysis of the samples with a very complex matrix (i.e., aerosol extracts), the ion intensities do not directly reflect the concentration of the molecules in the sample; therefore, data shown in this work is semi-quantitative.~~

High resolution MS data analysis

The direct infusion data analysis was performed using procedures described in detail by Kourtchev et al. (2013). Briefly, for each sample analysis, 60–90 mass spectral scans were averaged into one mass spectrum. Molecular formulae assignments were made using Xcalibur 3.1 software using the following constraints ^{12}C≤100, ^{13}C≤1, ^{1}H≤200, ^{16}O≤50, ^{14}N≤5, ^{32}S≤2, ^{34}S≤1. The data processing was performed using a Mathematica 8.0 (Wolfram Research Inc., UK) code developed in-house that utilises a number of additional constraints described in previous studies (Kourtchev et al., 2013; Kourtchev et al., 2015). Only ions that appeared in all three replicates were kept for evaluation. The background spectra obtained from the procedural blanks were also processed using the rules mentioned above. The formulae lists of the background spectra were subtracted from those of the ambient (or chamber) sample and only formulae with a sample/blank peak intensity ratio ≥ 10 were retained

The Kendrick Mass Defect (KMD) is calculated from the difference between the nominal mass of the molecule and the exact KM (Kendrick, 1963). Kendrick mass of the CH_2 unit is calculated by renormalising the exact IUPAC mass of CH_2 (14.01565) to 14.00000.

Benzene and isoprene measurements

For benzene and isoprene analysis we used a high-resolution selective-reagent-ionisation proton transfer reaction time-of-flight mass spectrometer (SRI-PTR-TOF-MS 8000, Ionicon Analytik, Austria). A description of the PTR-TOF-MS instrument and the data reduction process used are provided elsewhere (Graus et al. 2010; Müller et al. 2013). Background of the instrument was measured regularly by passing ambient air through a platinum catalyst heated to 380 ºC. Sensitivity calibrations were performed by dynamic dilution of VOCs using several multi-component gas standards (Apel Riemer Environmental Inc., Scott-Marrin, and Air Liquide, USA). The calibration cylinders contained acetaldehyde, acetone, benzene, isoprene, *α*-pinene, toluene and trichlorobenzene, among others. During IOP1, the instrument was operated with H_3O^+ reagent ion and at a drift tube pressure of 2.3 mbar, voltage of 600 V, and temperature of 60 ºC, corresponding to a~~n~~ a field density ratio E/N ratio of 130 Td (E being the electric field strength and N the gas number density; 1 Td = 10^{-17} V cm^{-2}). During IOP2, the reagent ion was NO^+ and the drift tube settings were 2.3 mbar, 350 V, and 60 ºC, resulting in an E/N ratio of 76 Td. The sampling was done with 1 min time resolution and the instrument detection limit for benzene and isoprene were below 0.02 and 0.04 ppbv, respectively.

Air mass history analysis

Air mass history analysis was done for the sampling period using the Numerical Atmospheric-dispersion Modeling Environment (NAME) model, developed by the UK Met Office (Maryon et al., 1991). NAME is a Lagrangian model in which particles are released into 3D wind fields from the operational output of the UK Met Office Unified Model meteorology data (Davies et al., 2005). These winds have a horizontal resolution of 17 km and 70 vertical levels up that reach ~80 km. In addition, a random walk technique was used to model the effects of turbulence on the trajectories (Ryall and Maryon, 1998). To allow the calculation of air mass history for the average sampling time (which varied between samples, 24, 36 or 48 hours), 10 000 particles per hour were released continuously from the T3 site. The trajectories travelled back in time for 3 days with the position of the particles in the lowest 100 m of the model atmosphere recorded every 15 min. The particle mass below 100 m was integrated over the 72 h travel time. The air mass history (‘footprints’) for the periods of the analysed filters are shown in Figure SI1. The majority of the three-day air mass footprints originated from the east, although wind direction showed variability nearer to the sampling site on some occasions e.g., sample MP14-17 (Fig. SI1). Almost all air masses pass over Manaus and therefore highlight this city as a potential source. Some air masses also pass over Manacapuru, but this is rare and the corresponding time-integrated concentrations are lower than the equivalent Manaus values.

Results and discussions

Figure 1 shows mass spectra from two typical samples collected during IOP1 and IOP2. The majority of the ions were associated with molecules below 500 Da although the measured mass goes up to 900 Da. Although ESI is a ‘soft’ ionisation technique resulting in minimal fragmentation, we cannot exclude the possibility that some of the detected ions correspond to fragments, also in light of the many relative fragile compounds (e.g., highly oxygenated compounds) that constitute OA. The largest group of identified ~~number of~~ molecular formulae in all samples were attributed to molecules containing CHO atoms only (1051±141 formulae during IOP2 and 820±139 during IOP1), followed by CHON (537±71 during IOP2 and 329±71 during IOP1), CHOS (183±34 during IOP2 and 137±31 during IOP1) and CHONS (37±11 during IOP2 and 28±10 during IOP1) (Fig. 2).The number of molecular formulae containing CHO and CHON subgroups increased by ~20% from IOP1 to IOP2 period; however, rather insignificant increase was observed for CHOS and CHONS subgroups. The Student's t-test showed that the observed difference for CHO (p=0.0092) and CHON (p=0.00007) subgroups between two seasons is statistically significant. This is consistent with the observed increase in odd reactive nitrogen species (NO_y) from IOP1 to IOP2 (Table SI1). Organic nitrates are believed to form in polluted air through reaction with nitrogen oxides during day and from reaction of NO_3^- with BVOCs during nighttime (Day et al., 2010; Ayres et al., 2015). The average concentration of NO_y during IOP1 was found to be on almost two times higher, which is possibly reflected in the increased number of organonitrates in the aerosol samples from IOP2. Moreover, the increase in the number of organonitrates during IOP2 is consistent with the recent studies, which demonstrated that organonitrates groups in aerosol particles may hydrolyse under high RH conditions (Liu et al., 2012). In this respect, while night time maximum RH during both filter sampling periods was very similar (~90%), day-time RH during IOP1 was higher (89%) compared to that from the IOP2~~'dry'~~ period (66%) (Fig. SI2).

Formatted: Subscript

Formatted: Subscript

Carbon oxidation state (OS_C)introduced by Kroll et al. (2011) can be used to describe the composition of a complex mixture of organics undergoing oxidation processes. OS_C was calculated for each molecular formula identified in the mass spectra using the following equation:

$$OS_C = -\sum_i OS_i \frac{n_i}{n_C} \qquad \text{(Eq. 1)}$$

where OS_i is the oxidation state associated with element i, n_i/n_C is the molar ratio of element i to carbon within the molecule (Kroll et al., 2011).

Figure 3 shows overlaid OSc plots for two samples from IOP1 and IOP2. Consistent with previous studies, the majority of molecules in the sampled organic aerosol had OS_C between –1.5 and +1 with up to 30 (nC) carbon atoms throughout the selected mass range (*m/z* 100-650) (Kroll et al., 2011 and the references therein). The molecules with OS_C between –1 and +1 with 13 or less carbon atoms (nC) are suggested to be associated with semivolatile and low-volatility oxidised organic aerosol (SV-OOA and LV-OOA) produced by multistep oxidation reactions. The molecules with OSc between -0.5 and -1.5 with 7 or more carbon atoms are associated with primary biomass burning organic aerosol (BBOA) directly emitted into the atmosphere (Kroll et al., 2011). The cluster of molecules with OS_C between –1 and –1.5 and nC less than 10 could be possibly associated with OH radical oxidation products of isoprene (Kourtchev et al., 2015), which is an abundant VOC in Amazon rain forest (Rasmussen and Khalil, 1988; Chen et al., 2015). The isoprene daytime average was above 1.5 ppbv during both seasons, with hourly campaign-averages reaching up to 2.3 and 3.4 ppbv for IOP1 and IOP2, respectively. In general, aerosol samples from IOP1 contained less oxidised molecules compared to those from IOP2. Wet deposition of aged or processed aerosol during wet (i.e., IOP 2) sampling period cannot be the only reason for the observed differences in OSc. It has been shown that different oxidation regimes to generate SOA (e.g., OH radical vs. ozonolysis) can result in significantly different OSc of SOA (Kourtchev et al., 2015). For example, the SOA component from OH initiated oxidation of α-pinene as well as BVOC mixtures had a molecular composition with higher OSc throughout the entire molecular mass range ~~(Kourtchev et al., 2015)~~ compared to that obtained from ozonolysis reaction (Kourtchev et al., 2015).

Figure 4 shows the distribution of ion intensities for selected tentatively identified tracer compounds for anthropogenic, biogenic and mixed sources in all 15 samples. The structural or isomeric information is not directly obtained from the direct infusion analysis; therefore, the identification of the tracer compounds was achieved by comparing MS/MS fragmentation patterns from authentic standards and published literature. The tracer compounds include anhydrosugars, structural isomers with a molecular formula $C_6H_{10}O_5$ at *m/z* 161.0456 corresponding to levoglucosan, mannosan, galactosan and 1,6-anhydro-*β*-*D*-glucofuranose, which are regarded as marker compounds for biomass burning (Simoneit et al., 1999; Pashynska et al., 2002; Kourtchev et al., 2011). Nitrocatechols, with a molecular formula $C_6H_5NO_4$ (*m/z* 154.01458) are attributed to mixed anthropogenic sources, e.g., biomass and vehicular emissions ~~sources~~and methyl-nitrocatechols ($C_7H_7NO_4$, *m/z* 168.03023) are important markers for biomass burning OA, formed from *m*-cresol emitted during biomass burning (Iinuma et al., 2010). ~~.~~ 3-methyl-1,2,3-butanetricarboxylic acid (3-MBTCA), with a molecular formula $C_8H_{12}O_6$ at *m/z* 203.05611, is an OH-initiated oxidation product of *α*- and *β*-pinene (Szmigielski et al., 2007), and regarded as a tracer for processed or biogenic SOA. Finally, isoprene epoxydiol organosulfate ester (IEPOX-OS), with a molecular formula $C_5H_{12}O_7S$ at *m/z* 215.0231, is shown in Figure 4. From studies in mid latitude environments it has been suggested that IEPOX-OS is formed through reactions between SO_x and isoprene oxidation products (Pye et al., 2013; Budisulistiorini et al., 2015) and thus can be used to observe the extent of SO_2 aging effects on the biogenic SOA. Direct infusion analysis suffers from competitive ionisation in the complex matrices and thus comparing ion intensities across samples has to be done with caution. Moreover, other compounds with similar molecular composition present in the aerosol matrix may also contribute to the ion intensities of the discussed above compounds. ~~Nevertheless, a~~All selected tracers showed very similar variations with benzene concentration that was measured in the gas-phase using PTR-MS (Fig. 3). Benzene, generally regarded as an anthropogenic species, has various sources including industrial solvent production, vehicular emissions and biomass burning (Hsieh et al., 1999; Seco et al., 2013; Friedli et al., 2001). Recent studies indicated that vegetation (leaves, flowers, and phytoplankton) emits a wide variety of benzenoid compounds to the atmosphere at substantial rates (Misztal et al., 2015). However, considering that benzene concentration correlated very well with another anthropogenic tracer CO (R^2=0.77, Figure SI3) during IOP1 and IOP2 periods, it is rather likely that the observed benzene concentrations were mainly due to anthropogenic emissions. During the sampling period, irrespectively of the season, air masses passed over the large city Manaus and small municipalities located near the T3 site (Figure SI1). It must be noted that due to rather low sampling resolution time (≥24h) the molecular composition of all analysed samples is likely to be influenced by clean air masses and anthropogenic plumes from these urban locations which usually last only a few hours per day and thus individual urban plume events cannot be identified with the data analysed here. In Manaus natural gas is mainly used for heating and cooking and therefore, the contribution from these activities to biomass burning OA at our site is highly unlikely. During IOP1 much

Formatted: Subscript

lower incidents of forest fires were observed compared to that during IOP2 (Martin et al., 2016). For example, a number of forest fires in the radius of 200 km from the sampling site varied between 0 to 340 fires (http://www.dpi.inpe.br/proarco/bdqueimadas/). This is reflected in the ion intensities of the particle phase biomass burning markers, i.e., anhydrosugars ($C_6H_{10}O_5$) and nitrocatechols ($C_6H_5NO_4$) and gas-phase benzene concentrations, which were significantly lower during IOP1 compared to that from IOP2, when on average more fires are observed.

It should be noted that ion intensities for anhydrosugars ($C_6H_{10}O_5$) and nitrocatechols ($C_6H_5NO_4$) showed very good correlation ($R^2>0.7$) suggesting that nitroca~~aha~~techols, observed at the sampling site, are mainly associated with biomass burning sources. The highest ion intensities of these tracer compounds were observed during two periods: 7-9 September 2014 (sample MP14-128) and 27-28 September 2014 (sample MP14-148) with the later one coinciding with highest incident of fires (340 fires). Although during 7-9 September (sample MP14-128) a significantly lower number (22 fires) of fires was observed compared to the period of 27-28 September 2014, lower wind speed occurring during 7-9 September suggests that high intensity of the biomass burning markers could be due to the biomass burning emissions from nearby sources. Between the T3 sampling site and Manaus (about 20 km east of the site), there are a number of small brick factories, which use wood to fire the kilns (Martin et al., 2016) and thus they are an additional local wood burning source besides the forest and pasture fires.

Interestingly the sample MP14-148 had the highest ion intensity corresponding to IEPOX-OS (Fig. 4), which also coincided with the strong increase of the ion intensity at *m/z* 96.95987 corresponding to $[HSO_4]^-$. This is consistent with organosulfates formation mechanism through reactive uptake of isoprene epoxydiols (IEPOX) in the presence of acidic sulfate seed (Surratt et al., 2010; Lin et al., 2012; 2013). A similar relationship between sulfate and organosulfates concentrations has been observed previously in field studies in the Southeastern US (Surratt et al., 2007, 2008, 2010; Lin et al., 2012, 2013). This is also in agreement with previous studies from Amazon where the highest levels of 2-methyltetrols were observed during the dry period which was characterised by biomass burning (and higher particle concentrations of sulfuric acid) (Claeys et al., 2010). Considering that Claeys et al (2010) employed alternative GC/MS procedure with prior trimethylsylilation, 2-methyltetrol sulfates were converted to 2-methyltetrols and not detectable as separate OS compounds. It should be noted that the 27-28 September period (sample MP14-148) was marked by a very strong increase in the CO concentration (Fig. SI4). In mid-latitude environments it has been suggested that the production of anthropogenic SOA in an air mass, as it travels from an urban source region, can be estimated by using a relatively inert pollution tracer, such as CO occurring in the air mass (De Gouw et al., 2005; Hoyle et al., 2011). At T3 sampling site, highest CO concentrations are observed in air masses affected by biomass burning. Therefore, it is possible that organic aerosol in the sample MP14-148 has experienced the highest contribution from biomass burning as well as other anthropogenic activities.

To investigate the influence of anthropogenic activities (i.e., biomass burning) on a detailed molecular composition of organic aerosol at the T3 site we compared samples from the periods with the lowest (9 fires), moderately high (254 fires) and the highest (340 fires) incidents of fires occurring within 200 km around the site.

Figure 5 (a-c) shows H/C ratios of CHO containing formulae as a function of their molecular mass and double bond equivalent (DBE), which shows a degree of unsaturation of the molecule, for a sample with the lowest (a) moderately high (b) and highest incidents (c) of fires. One of the obvious differences between these samples is the abundance of ions with low H/C ratios (< 1). The majority of these ions have DBE above 7 indicating that they likely correspond to oxidised aromatic compounds, which are mainly of anthropogenic origin (Kourtchev et al., 2014; Tong et al., 2016). For example, the smallest polycyclic aromatic hydrocarbon (PAH), naphthalene with a molecular formulae $C_{10}H_8$ has an H/C=0.8 and DBE=7. The number of CHO containing formulae with high DBE equivalent and low H/C increased dramatically during the days with moderately high and high incidents of fires (Fig.

5a-c), suggesting that they are mainly associated with biomass burning. The largest grey circles in Fig 5(a-c) correspond to the ions at *m/z* 133.01425 (with neutral molecular formula $C_4H_6O_5$), *m/z* 187.0612 ($C_8H_{12}O_5$), *m/z* 201.07685 ($C_9H_{14}O_5$), *m/z* 203.05611 ($C_8H_{12}O_6$), and *m/z* 215.05611 ($C_9H_{12}O_6$) with DBE<6.

Recent studies indicated that different families of compounds with heteroatoms (e.g. O, S) overlap in terms of DBE and thus may not accurately indicate the level of unsaturation of organic compounds. For example, the divalent atoms, such as oxygen and sulphur, do not influence the value of DBE, yet they may contribute to the potential double bonds of that molecule (Reemtsma 2009; Yassine et al., 2014). Yassine et al (2014) suggested using aromaticity equivalent (X_c), to improve the identification and characterisation of aromatic and condensed aromatic compounds in WSOC. The aromaticity equivalent can be calculated as follows:

$$\mathrm{X}_c = \frac{3(\mathrm{DBE}-(\mathrm{mN}_O+n\mathrm{N}_S))-2}{\mathrm{DBE}-(\mathrm{mN}_O+\mathrm{nN}_S)} \qquad \text{(Eq. 2)}$$

where ‘m’ and ‘n’ correspond to a fraction of oxygen and sulfur atoms involved in π–bond structures of a compound, which varies depending on the compound class. For example, carboxylic acids, esters, and nitro functional groups have m=n=0.5. For compounds containing functional groups such as aldehydes, ketones, nitroso, cyanate, alcohol, or ethers ‘m’ and ‘n’ are 1 or 0. Considering that ESI, in negative mode, is most sensitive to compounds containing carboxylic groups we, therefore, used m=n=0.5 for the calculation of the Xc. For molecular formulae with an odd number of oxygen or sulfur, the sum (mN_O+nN_S) in Eq. 2 was rounded down to the closest integer as detailed in Yassine et al (2014). The authors proposed that aromaticity equivalent with Xc ≥2.50 and Xc ≥2.71 as unambiguous minimum criteria for the presence of aromatics and condensed aromatics.

Expressing our data using aromaticity equivalents confirmed that the increase in the number of molecules with high DBE from the sample with the lowest to the highest incidents of fires was due to the increase in the number of aromatic and condensed aromatic compounds in the aerosol samples (Figures SI5). Considering the Yassine et al. (2014) assignment criteria for the aromatic-reach matrices, the highest number of the aromatic compounds in the Amazon samples was observed for formulae with a benzene core structure (Xc =2.50) followed by formulae with pyrene core structure (Xc = 2.83), and an ovalene core structure (Xc =2.92) as well as highly condensed aromatic structures or highly unsaturated compounds (Xc >2.93). The largest grey circles in Figure SI5a correspond to the ions at *m/z* 187.11357 with a neutral molecular formula $C_9H_{17}NO_3$ and *m/z* 281.26459 with a neutral molecular formula $C_{18}H_{35}NO$. The largest grey circles in Figure SI5b and c correspond to the ions at *m/z* 154.0146, *m/z* 168.03023 and *m/z* 152.03532 with neutral molecular formulae $C_6H_5NO_4$, $C_7H_7NO_4$ and $C_7H_7NO_3$, respectively.

Interestingly, a similar trend was observed for the molecules containing CHON subgroups (Figure SI6). A number of CHON molecules with low H/C (<1) and high DBE (≥5) almost doubled from the days with 9 to 340 fires (Figure SI7). Nitro-aromatic compounds, such as nitrophenols (DBE=5) and N-heterocyclic compounds including 4-nitrocatechol and isomeric methyl-nitrocatechols are often observed in the OA from the biomass burning sources (Kitanovski et al., 2012a,b; Iinuma et al., 2010) and have been suggested as potential contributors to light absorption by brown carbon (Laskin et al., 2015)~~Nitro-aromatic compounds, such as nitrophenols (DBE=5) and N-heterocyclic compounds are often observed in the PM from the biomass burning sources (Kitanovski et al., 2012a,b) and have been suggested as potential contributors to light absorption by brown carbon (Laskin et al., 2015)~~. It is worth mentioning that aerosol samples affected by biomass burning contained another interesting ion at *m/z* 182.04588 with a neutral molecular formula $C_8H_9NO_4$, possibly corresponding to biomass burning OA markers isomeric dimethyl-nitrocatechols (Kahnt et al., 2013). The differences in the increased number of nitro-artomatic compounds in aerosol samples affected by biomass burning are also apparent in ~~overplayed~~overlaid Van Krevelen diagrams (Figure 6), which show H/C and O/C ratios for each formula in a sample. Van

Krevelen diagrams, can be used to describe the overall composition or evolution of organic mixtures (Van Krevelen, 1993; Nizkorodov et al., 2011; Noziere et al., 2015). Organic aerosol affected by biomass burning contained significantly larger number of CHON formulae with O/C < 0.5 and H/C < 1 (Fig. 6a and b, area B) but smaller number of formulae with O/C < 0.5 and H/C > 1. (Fig. 6a and b, area A). While molecules with H/C ratios (<1.0) and O/C ratios (<0.5) (area A in Fig. 3) are generally associated with aliphatic compounds typically belong to oxidised aromatic hydrocarbons, molecules with high H/C ratios (>1.5) and low O/C ratios (<0.5) (area B in Fig. 3) (Mazzoleni et al., 2010; 2012). Although the smaller number of nitro-aliphatic compounds in the samples affected by biomass burning requires further investigation, it is possible that they were oxidised in the polluted air by NO_x and O_3 (Zahardis et al., 2009; Malloy et al., 2009), which production is significantly enhanced during fire events (e.g., Galanter et al., 2000). The majority (up to 80%) of the CHON molecules in the analysed samples have O/C ratios < 0.7 (Fig. 6). The relatively low oxygen content suggests that these molecules include ~~reduced~~ decreased nitrogen-containing compounds (Zhao et al, 2013). Although biomass burning material type is expected to result in different molecular composition, the presence of a large number of molecules with low O/C ratio is consistent with the literature. For example, most of the CHON molecules in OA from wheat straw burning in K-puszta, Great Hungarian Plain in Hungary and biomass burning at Canadian rural sites (Saint Anicet, Quebec, and Canterbury, New Brunswick) had O/C ratios below 0.7 (Schmitt-Kopplin et al., 2010). In addition, the CHON molecules identified by LC/MS in biomass burning OA from Amazonia showed O/C ratios below 0.7, i.e., 4-nitrocatechol ($C_6H_5NO_4$; O/C = 0.67), isomeric methyl-nitrocatechols ($C_7H_7NO_4$; O/C = 0.57), and isomeric dimethyl-nitrocatechols ($C_8H_9NO_4$; O/C = 0.50) (Claeys et al., 2012).

Figure 7 shows overlaid OSc plots for OA from the days with low, moderately high and high incidents of fires. During the days affected by high and moderately high number of fires, OSc was shifted towards more oxidised state for the CHO molecules containing more than 7 carbon atoms. The ~~This~~ difference in OSc becomes even more pronounced with the increased number of carbons (e.g. >7 carbon atoms) in the detected molecular formulae. Interestingly, the affected ions with high OSc do not fall into the category of the BBOA (encircled area in Fig. 7) which are associated with primary particulate matter directly emitted into the atmosphere as defined in Kroll et al (2011).

At first glance, biomass burning seems to influence the number and intensity of the CHOS containing formulae; however, the effect was at a much lower extent compared to that for the CHO and CHON molecules (see discussion above). Higher number of CHOS containing molecules was observed in the sample (MP14-148) corresponding to the highest incident of fires (Figures 8a). Interestingly, IEPOX-OS was found to be very abundant in the sample that experienced the highest incidents of fires (Figure 8a). The significant IEPOX-OS mass was previously observed during a low-altitude flight campaigns at Northern California and southern Oregon at high NO conditions (> 500 pptv) (Liao et al., 2015). The authors explained this observation by the transport or formation of IEPOX from isoprene hydroxynitrate oxidation (Jacobs et al., 2014) and higher sulphate aerosol concentrations occurring during their sampling period (Nguyen et al., 2014). This explanation is also consistent with our results. The ion at *m/z* 96.95987 corresponding $[HSO_4]^-$ in UHR mass spectra of the sample MP14-148 was three times more abundant compared that in the sample MP14-129 suggesting that particle acidity may be one of the reasons for the high abundance of the IEPOX-OS in this sample. Considering that the main sources of sulphate at T3 site are industrial pollution (e.g., power plants), natural and long range-sources, they could also be responsible for the high abundance of the sulphate and IEPOX-OS in the samples besides the overlapping biomass burning event. Noticeably, these samples not only contained a larger number of oxygenated CHOS-containing molecules with $O/C>1.2$ but also molecules with $O/C<0.6$ and H/C ranging from 0.4 to 2.2. Recent laboratory and field studies indicated the presence of a large number of aromatic and aliphatic OSs and sulfonates in OA and linked them to anthropogenic precursors (Tao et al., 2014; Wang et al., 2015; Riva et al., 2015; 2016; Kuang et al., 2016). Riva et al (2015, 2016) demonstrated formation of OSs and sulfonates in the laboratory smog chamber experiments from photooxidation of alkanes and PAHs, respectively. The authors indicated enhancement of organosulfates yields in the presence of the acidified ammonium sulphate seed and suggested that these OSs are mainly formed through reactive uptake of gas-phase epoxides. It must be noted that above cited field studies are based on measurements at the Northern Hemisphere USA and thus organosulfates formation pathways and sources may differ from that of Amazonia.

KMD plots are useful visualisation technique for identification of homologous series of compounds differing only by the number of a specific base unit (e.g., a CH_2 group). Anthropogenically affected aerosol samples have longer homologous series of molecules containing CHOS subgroups (Figure 8b). One of these longer series includes a second most intensive ion at *m/z* 213.0075 ($C_5H_{10}O_7S$). The compound with molecular formula $C_5H_{10}O_7S$ has been previously observed in the laboratory and field studies and attributed to isoprene derived organosulfates (Surratt et al., 2008; Gómez-González, 2008; Kristensen and Glassius, 2011; Nguyen et al., 2014; Hettiyadura et al., 2015). This molecular formula could also be associated with organosulfates (e.g., isomeric 3-sulfooxy-2-hydroxypentanoic acid and 2-sulfooxy-3-hydroxypentanoic acid) formed from the green leaf volatiles 2-*E*-pentenal, 2-*E*-hexenal, and 3-hexenal (Shalamzari et al., 2016). The KMD plot (Figure 8b) shows that OA from the anthropogenically affected samples contained an additional series of CHOS molecules with high KMD >0.33 that were not present in the sample from the less polluted period. Most of these ions are highly oxygenated (containing >10 oxygens) and are likely to be associate with molecules produced through ~~homogeneous~~ photochemical ageing reactions (Hildebrandt et al., 2010).

It is worth noting that in the most of the samples IEPOX-OS was not a part of any homologous series in KMD plot (e.g., Fig 8b). This observation confirms that atmospheric oxidation reactions resulting in the incorporation of S and N functional groups do not always conserve homologous series but could also lead to a wide range of possible reaction products (Rincon et al., 2012; Kourtchev et al., 2013).

Conclusions

In this study we applied direct infusion nanoESI UHR-MS for the analysis of the organic fraction of $PM_{2.5}$ samples collected IOP1 and IOP2 of GoAmazon2014/5 in central Amazonia which is influenced by both background and polluted air masses. Up to 2100 elemental formulae were identified in the samples, with the largest number of formulae found during IOP2. The distribution of several tracer compounds along with the comprehensive mass spectral data evaluation methods (e.g., Kendrick Mass Defect, Van Krevelen diagrams, carbon oxidation state and aromaticity equivalent) applied to the large UHRMS datasets were used to identify various sources of organic aerosol components, including natural biogenic sources, biomass burning and anthropogenic emissions. The distinguishable homologous series in the KMD diagram contained nitrogen-containing series included NACs, e.g., nitrocatechols, nitrophenols, nitroguaiacols and nitrosalicylic acids derived from biomass burning material. Isoprene derived IEPOX-OS was found as the most dominant ion in most of the analysed samples and strongly followed the concentration trends of the gas-phase anthropogenic tracer benzene and CO (with biomass burning as dominant tracer at the T3 site) supporting its mixed biomass burning-anthropogenic-biogenic origin. Van Krevelen, DBE and Xc distributions along with relatively low elemental O/C and H/C ratios indicated the presence of a large number of oxidised aromatic compounds in the samples. A significant number of CHO containing formulae in aerosol samples from IOP2 had higher oxidation state compared to that from IOP1 and became even more important during the days with the highest incidents of fires. Although our results suggest that the studied site is not only significantly influenced by biogenic emissions and biomass burning but also anthropogenic emissions from the neighboring urban activities, future work is needed to better understand the quantitative contributions of the various factors to the aerosol composition at the T3 site. The analysis of aerosol samples with higher sampling resolution or quantifying specific marker compounds and applying a receptor modelling techniques (Alves et al., 2015) would allow separating these sources in more detail and thus significantly improve understanding of the aerosol formation sources at the site.~~The analysis of aerosol samples with higher sampling resolution would allow separating these sources in more detail and thus significantly improve the understanding of the aerosol formation sources at the site.~~

Acknowledgment:

Research at the University of Cambridge was supported by the ERC grant no. 279405. The authors would like to thank Dr Jason Surratt (University of North Carolina) for providing a synthesised IEPOX-OS standard. O_3, CO, NO_y, RH and rain data were obtained from the Atmospheric Radiation Measurement (ARM) Climate Research Facility, a U.S. Department of Energy Office of Science user facility sponsored by the Office of Biological and Environmental Research. We acknowledge the support from the Central Office of the Large Scale Biosphere Atmosphere Experiment in Amazonia (LBA), the Instituto Nacional de Pesquisas da Amazonia (INPA), and the Universidade do Estado do Amazonia (UEA). The work was conducted under 001030/2012-4 of the Brazilian National Council for Scientific and Technological Development (CNPq).

Formatted: Subscript

Formatted: Subscript

References:

Andreae, M. O. and Crutzen, P. J.: Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry, Science, 276, 1052–1058, 1997.

Andreae, M. O., Acevedo, O. C., Araùjo, A., Artaxo, P., Barbosa, C. G. G., Barbosa, H. M. J., Brito, J., Carbone, S., Chi, X., Cintra, B. B. L., da Silva, N. F., Dias, N. L., Dias-Júnior, C. Q., Ditas, F., Ditz, R., Godoi, A. F. L., Godoi, R. H. M., Heimann, M., Hoffmann, T., Kesselmeier, J., Könemann, T., Krüger, M. L., Lavric, J. V., Manzi, A. O., Lopes, A. P., Martins, D. L., Mikhailov, E. F., Moran-Zuloaga, D., Nelson, B. W., Nölscher, A. C., Santos Nogueira, D., Piedade, M. T. F., Pöhlker, C., Pöschl, U., Quesada, C. A., Rizzo, L. V., Ro, C.-U., Ruckteschler, N., Sá, L. D. A., de Oliveira Sá, M., Sales, C. B., dos Santos, R. M. N., Saturno, J., Schöngart, J., Sörgel, M., de Souza, C. M., de Souza, R. A. F., Su, H., Targhetta, N., Tóta, J., Trebs, I., Trumbore, S., van Eijck, A., Walter, D., Wang, Z., Weber, B., Williams, J., Winderlich, J., Wittmann, F., Wolff, S., and Yáñez-Serrano, A. M.: The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols, Atmos. Chem. Phys., 15, 10723-10776, 2015.

Alves, E. G., Jardine, K., Tota, J., Jardine, A., Yáñez-Serrano, A. M., Karl, T., Tavares, J., Nelson, B., Gu, D., Stavrakou, T., Martin, S., Manzi, A., and Guenther, A.: Seasonality of isoprenoid emissions from a primary rainforest in central Amazonia, Atmos. Chem. Phys. Discuss., 15, 28867-28913, 2015.

Ayres, B. R., Allen, H. M., Draper, D. C., Brown, S. S., Wild, R. J., Jimenez, J. L., Day, D. A., Campuzano-Jost, P., Hu, W., de Gouw, J., Koss, A., Cohen, R. C., Duffey, K. C., Romer, P., Baumann, K., Edgerton, E., Takahama, S., Thornton, J. A., Lee, B. H., Lopez-Hilfiker, F. D., Mohr, C., Wennberg, P. O., Nguyen, T. B., Teng, A., Goldstein, A. H., Olson, K., and Fry, J. L.: Organic nitrate aerosol formation via NO_3 + biogenic volatile organic compounds in the southeastern United States, Atmos. Chem. Phys., 15, 13377-13392, 2015.

Budisulistiorini, S. H., Li, X., Bairai, S. T., Renfro, J., Liu, Y., Liu, Y. J., McKinney, K. A., Martin, S. T., McNeill, V. F., Pye, H. O. T., Nenes, A., Neff, M. E., Stone, E. A., Mueller, S., Knote, C., Shaw, S. L., Zhang, Z., Gold, A., and Surratt, J. D.: Examining the effects of anthropogenic emissions on isoprene-derived secondary organic aerosol formation during the 2013 Southern Oxidant and Aerosol Study (SOAS) at the Look Rock, Tennessee ground site, Atmos. Chem. Phys., 15, 8871-8888, 2015.

Chen, Q., Farmer, D. K., Rizzo, L. V., Pauliquevis, T., Kuwata, M., Karl, T. G., Guenther, A., Allan, J. D., Coe, H., Andreae, M. O., Pöschl, U., Jimenez, J. L., Artaxo, P., and Martin, S. T.: Submicron particle mass concentrations and sources in the Amazonian wet season (AMAZE-08), Atmos. Chem. Phys., 15, 3687-3701, 2015.

Claeys, M., Kourtchev, I., Pashynska, V., Vas, G., Vermeylen, R., Wang, W., Cafmeyer, J., Chi, X., Artaxo, P., Andreae, M. O., and Maenhaut, W.: Polar organic marker compounds in atmospheric aerosols during the LBA-SMOCC 2002 biomass burning experiment in Rondnia, Brazil: sources and source processes, time series, diel variations and size distributions, Atmos. Chem. Phys., 10, 9319–9331, doi:10.5194/acp-10-9319-2010, 2010.

Claeys, M., Vermeylen, R., Yasmeen, F., G´omez-Gonz´alez, Y., Chi, X., Maenhaut, W., Meszaros, T., and Salma, I.: Chemical characterisation of humic-like substances from urban, rural and tropical biomass burning environments using liquid chromatography with UV/vis photodiode array detection and electrospray ionisation mass spectrometry, Environ. Chem., 9, 273–284, doi:10.1071/EN11163, 2012.

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., and Wood, N.: A new dynamical core for the Met Office's global and regional modelling of the atmosphere, Q. J. Roy. Meteorol. Soc., 131, 1759–1782, 2005.

Day, D. A., Liu, S., Russell, L. M. and Ziemann, P. J.: Organonitrate group concentrations in submicron particles with high nitrate and organic fractions in coastal southern California, Atmos. Environ., 44, 1970-1979, 2010.

de Gouw, J. A., Middlebrook, A. M., Warneke, C., Goldan, P. D., Kuster, W. C., Roberts, J. M., Fehsenfeld, F. C., Worsnop, D. R., Canagaratna, M. R., Pszenny, A. A. P., Keene, W. C., Marchewka, M., Bertman, S. B., and Bates, T. S.: Budget of organic carbon in a polluted atmosphere: Results from the New England Air Quality Study in 2002, J. Geophys. Res.-Atmos., 110, D16305, doi:10.1029/2004jd005623, 2005.

Dzepina, K., Mazzoleni, C., Fialho, P., China, S., Zhang, B., Owen, R. C., Helmig, D., Hueber, J., Kumar, S., Perlinger, J. A., Kramer, L. J., Dziobak, M. P., Ampadu, M. T., Olsen, S., Wuebbles, D. J., and Mazzoleni, L. R.: Molecular characterization of free tropospheric aerosol collected at the Pico Mountain Observatory: a case study with a long-range transported biomass burning plume, Atmos. Chem. Phys., 15, 5047-5068, 2015.

Friedli, H. R., E. Atlas, V. R. Stroud, L. Giovanni, T. Campos, and L. F. Radke: Volatile organic trace gases emitted from North American wildfires, Global Biogeochem. Cycles, 15, 435–452, 2001.

Fry, J. L., Draper, D. C., Barsanti, K. C., Smith, J. N., Ortega, J., Winkler, P. M., Lawler, M. J., Brown, S. S., Edwards, P. M., Cohen, R. C., and Lee, L.: Secondary organic aerosol formation and organic nitrate yield from NO3 oxidation of biogenic hydrocarbons, Environ. Sci. Technol, 48, 11944–11953, 2014.

Galanter, M., Levy II, H., and Carmichael, G. R.: Impacts of biomass burning on tropospheric CO, NOx, and O3, J. Geophys. Res., 105, 6633–6653, 2000.

Graus, M., Müller, M. and Hansel, A.: High Resolution PTR-TOF: Quantification and Formula Confirmation of VOC in Real Time, J. Am. Soc. Mass Spectrom., 21, 1037-1044, 2010.

Greenberg, J., Guenther, A., Petron, G., Wiedinmyer, C., Vega, O., Gatti, L. V., Tota, J., and Fisch, G.: Biogenic VOC emissions from forested Amazonian landscapes, Global Change Biol., 10(5), 651–662, 2004.

Gómez-González, Y., Surratt, J. D., Cuyckens, F., Szmigielski, R., Vermeylen, R., Jaoui, M., Lewandowski, M., Offenberg, J. H., Kleindienst, T. E., Edney, E. O., Blockhuys, F., Van Alsenoy, C., Maenhaut, W., and Claeys, M.: Characterization of organosufates from the photooxidation of isoprene and unsaturated fatty acids in ambient aerosol using liquid chromatography/(–) electrospray ionization mass spectrometry, J. Mass Spectrom., 43, 371–382, 2008.

Goldstein, A. H. and Galbally, I. E.: Known and unexplored organic carbon constituents in the Earth’s atmosphere, Environ. Sci. Technol., 41, 1514–1521, 2007.

Goldstein, A. H., Koven, C. D., Heald, C. L., and Fung, I. Y.: Biogenic carbon and anthropogenic pollutants combine to form a cooling haze over the southeastern United States, P. Natl. Acad. Sci. USA, 106, 8835–8840, 2009.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, Th. F., Monod, A., Prevot, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9, 5155–5235, 2009.

Haywood, J. and Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review, Rev. Geophys., 38, 513–543, 2000.

Hildebrandt, L., Engelhart, G. J., Mohr, C., Kostenidou, E., Lanz, V. A., Bougiatioti, A., DeCarlo, P. F., Prevot, A. S. H., Baltensperger, U., Mihalopoulos, N., Donahue, N. M., and Pandis, S. N.: Aged organic aerosol in the Eastern Mediterranean: the Finokalia Aerosol Measurement Experiment – 2008, Atmos. Chem. Phys., 10, 4167–4186, 2010.

Hettiyadura, A. P. S., Stone, E. A., Kundu, S., Baker, Z., Geddes, E., Richards, K., and Humphry, T.: Determination of atmospheric organosulfates using HILIC chromatography with MS detection, Atmos. Meas. Tech., 8, 2347–2358, 2015.

Hoyle, C. R., Boy, M., Donahue, N. M., Fry, J. L., Glasius, M., Guenther, A., Hallar, A. G., Huff Hartz, K., Petters, M. D., Petaja, T., Rosenoern, T., and Sullivan, A. P.: A review of the anthropogenic influence on biogenic secondary organic aerosol, Atmos. Chem. Phys., 11, 321–343, 2011.

Iinuma, Y., Boge, O., Grafe, R., and Herrmann, H.: Methyl-Nitrocatechols: atmospheric tracer compounds for biomass burning secondary organic aerosols, Environ. Sci. Technol., 44, 8453–8459, 2010.

Jacobson, M. Z.: Isolating nitrated and aromatic aerosols and nitrated aromatic gases as sources of ultraviolet light absorption, J. Geophys. Res. Atmospheres, 104 (D3), 3527-3542, 1999.

Jacobs, M. I., Burke, W.J., and Elrod, M.K.: Kinetics of the reactions of isoprene-derived hydroxynitrates: Gas phase epoxide formation and solution phase hydrolysis, Atmos. Chem. Phys., 14, 8933–8946, 2014.

Kahnt, A., Behrouzi, S., Vermeylen, R., Safi Shalamzari, M., Vercauteren, J., Roekens, E., Claeys, M., and Maenhaut, W.: One year study of nitro-organic compounds and their relation to wood burning in PM10 aerosol from a rural site in Belgium, Atmos. Environ., 81, 561–568, 2013.

Kendrick E.: A Mass Scale Based on CH_2 = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds, Anal. Chem., 35 (13), 2146–2154, 1963.

Keller, M., Bustamante, M., Gash, J., and Dias, P.: Amazonia and Global Change, Vol. 186, American Geophysical Union, Wiley, Washington, D.C., 2009.

Kitanovski, Z., Grgíc, I., Vermeylen, R., Claeys, M., and Maenhaut, W.: Liquid chromatography tandem mass spectrometry method for characterization of monoaromatic nitro-compounds in atmospheric particulate matter, J. Chromatogr. A, 1268, 35–43, 2012a.

Kitanovski, Z., Grgic, I., Yasmeen, F., Claeys, M., and Cusak, A.: Development of a liquid chromatographic method based on ultraviolet-visible and electrospray ionization mass spectrometric detection for the identification of nitrocatechols and related tracers in biomass burning atmospheric organic aerosol, Rapid Commun. Mass Sp., 26, 793–804, 2012b.

Kourtchev, I., Hellebust, S., Bell, J. M., O'Connor, I. P., Healy, R. M., Allanic, A., Healy, D., Wenger, J. C., and Sodeau, J. R.: The use of polar organic compounds to estimate the contribution of domestic solid fuel combustion and biogenic sources to ambient levels of organic carbon and PM2.5 in Cork Harbour, Ireland, Sci.Total Environ., 409, 2143–2155, 2011.

Kourtchev, I., Fuller, S., Aalto, J., Ruuskanen, T. M., McLeod, M. W., Maenhaut, W., Jones, R., Kulmala, M., and Kalberer, M.: Molecular composition of boreal forest aerosol from Hyytiälä, Finland, using ultrahigh resolution mass spectrometry, Environ. Sci. Technol., 47, 4069–4079, 2013.

Kourtchev, I., O'Connor, I. P., Giorio, C., Fuller, S., Kristensen, K., Maenhaut, W., Wenger, J. C., Sodeau, J. R., Glasius, M., and Kalberer, M.: Effects of anthropogenic emissions on the molecular composition of urban organic aerosols: an ultrahigh resolution mass spectrometry study, Atmos. Environ., 89, 525-532, 2014.

Kourtchev, I., Doussin, J.-F., Giorio, C., Mahon, B., Wilson, E. M., Maurin, N., Pangui, E., Venables, D. S., Wenger, J. C., and Kalberer, M.: Molecular composition of fresh and aged secondary organic aerosol from a mixture of biogenic volatile compounds: a high-resolution mass spectrometry study, Atmos. Chem. Phys., 15, 5683-5695, 2015.

Kristensen, K. and Glasius, M.: Organosulfates and oxidation products from biogenic hydrocarbons in fine aerosols from a forest in North West Europe during spring, Atmos. Environ., 45, 4546–4556, 2011.

Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E., and Worsnop, D. R.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol, Nat. Chem. Biol., 3, 133–139, 2011.

Kleinman, L., Kuang, C., Sedlacek, A., Senum, G., Springston, S., Wang, J., Zhang, Q., Jayne, J., Fast, J., Hubbe, J., Shilling, J., and Zaveri, R.: What do correlations tell us about anthropogenic–biogenic interactions and SOA formation in the Sacramento Plume during CARES?, Atmos. Chem. Phys. Discuss., 15, 25381-25431, 2015.

Kuang, B.Y., Lin, P., Hub, M., and Yu, J.Z.: Aerosol size distribution characteristics of organosulfates in the Pearl River Delta region, China, Atmos. Environ., 130, 23-35, 2016.

Laskin, A., Laskin, J., and Nizkorodov, S. A.: Chemistry of Atmospheric Brown Carbon, Chem. Rev., 115, 4335-4382, 2015.

Liao, J., Froyd, K. D., Murphy, D. M., Keutsch, F. N., Yu, G., Wennberg, P. O., St. Clair, J. M., Crounse, J. D., Wisthaler, A., Mikoviny, T., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Ryerson, T. B., Pollack, I. B., Peischl, J., Anderson, B.E., Ziemba, L. D., Blake, D. R., Meinardi, S., and Diskin, G.: Airborne measurements of organosulfates over the continental U. S., J. Geophys. Res., 120, 2990–3005, doi: 10.1002/2014JD022378,2015.

Lin, Y. H., Zhang, Z. F., Docherty, K. S., Zhang, H. F., Budisulistiorini, S. H., Rubitschun, C. L., Shaw, S. L., Knipping, E. M., Edgerton, E. S., Kleindienst, T. E., Gold, A., and Surratt, J. D.: Isoprene Epoxydiols as Precursors to Secondary Organic Aerosol Formation: Acid-Catalyzed Reactive Uptake Studies with Authentic Compounds, Environ. Sci. Technol., 46, 250–258, 2012.

Lin, Y.-H., Knipping, E. M., Edgerton, E. S., Shaw, S. L., and Surratt, J. D.: Investigating the influences of SO_2 and NH_3 levels on isoprene-derived secondary organic aerosol formation using conditional sampling approaches, Atmos. Chem. Phys., 13, 8457-8470, 2013.

Lu, J. W., Flores, J. M., Lavi, A., Abo-Riziq, A., and Rudich, Y.: Changes in the optical properties of benzo[a]pyrene-coated aerosols upon heterogeneous reactions with NO2 and NO3, Phys. Chem. Chem. Phys., 13, 6484–6492, 2011.

Malloy, Q. G. J., Li Qi, Warren, B., Cocker III, D. R., Erupe, M. E., and Silva, P. J.: Secondary organic aerosol formation from primary aliphatic amines with NO3 radical, Atmos. Chem. Phys., 9, 2051–2060, 2009.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmos. Chem. Phys., 16, 4785-4797, 2016.

Martin, S. T., Andreae, M. O., Artaxo, P., Baumgardner, D., Chen, Q., Goldstein, A. H., Guenther, A., Heald, C. L., Mayol-Bracero, O. L., McMurry, P. H., Pauliquevis, T., Pöschl, U., Prather, K. A., Roberts, G. C., Saleska, S. R., Silva Dias, M. A., Spracklen, D. V., Swietlicki, E., and Trebs, I.: Sources and properties of Amazonian aerosol particles, Rev. Geophys, 48, RG2002, doi:10.1029/2008RG000280, 2010.

Maryon, R.H., Smith, F.B., Conway, B.J., Goddard, D.M.: The U.K. nuclear accident model, Prog. Nucl. Energy 26, 85-104, 1991.

Mazzoleni, L. R., Ehrmann, B. M., Shen, X., Marshall, A. G., and Collett Jr., J. L.: Water-soluble atmospheric organic matter in fog: exact masses and chemical formula identification by ultrahighresolution Fourier transform ion cyclotron resonance mass spectrometry, Environ. Sci. Technol., 44, 3690–3697, 2010.

Mazzoleni, L. R., Saranjampour, P., Dalbec, M. M., Samburova, V., Hallar, A. G., Zielinska, B., Lowenthal, D., and Kohl, S.: Identification of water-soluble organic carbon in nonurban aerosols using ultrahigh-resolution FT-ICR mass spectrometry: organic anions, Environ. Chem., 9, 285–297, 2012.

Müller, M., Mikoviny, T., Jud, W., D'Anna, B., and Wisthaler, A.: A New Software Tool for the Analysis of High Resolution PTR-TOF Mass Spectra, Chemometr. Intell. Lab., 127, 158-165, 2013.

Nguyen, Q. T., Christensen, M. K., Cozzi, F., Zare, A., Hansen, A. M. K., Kristensen, K., Tulinius, T. E., Madsen, H. H., Christensen, J. H., Brandt, J., Massling, A., Nøjgaard, J. K., and Glasius, M.: Understanding the anthropogenic influence on formation of biogenic secondary organic aerosols in Denmark via analysis of organosulfates and related oxidation products, Atmos. Chem. Phys., 14, 8961-8981, 2014.

Nizkorodov, S. A., Laskin, J., and Laskin, A.: Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry, Phys. Chem. Chem. Phys., 13, 3612–3629, 2011.

Nozière, B., Kalberer, M., Claeys, M., Allan, J., D'Anna, B., Decesari, S., Finessi, E., Glasius, M., Grgic, I., Hamilton, J. F., Hoffmann, T., Iinuma, Y., Jaoui, M., Kahnt, A., Kampf, C. J., Kourtchev, I., Maenhaut, W., Marsden, N., Saarikoski, S., Schnelle-Kreis, J., Surratt, J., Szidat, S., Szmigielski, R., and Wisthaler, S.: The molecular identification of organic compounds in the atmosphere: state of the art and challenges, Chem. Rev., 115, 3919–3983, 2015.

Oss, M., Kruve, A., Herodes, K., and Leito, I.: Electrospray ionization efficiency scale of organic compounds, Anal. Chem., 82, 2865–2872, 2010.

Pashynska, V., Vermeylen, R., Vas, G., Maenhaut, W., and Claeys, M.: Development of a gas chromatographic/ion trap mass spectrometric method for the determination of levoglucosan and saccharidic compounds in atmospheric aerosols. Application to urban aerosols. J. Mass Spec., 37, 1249–1257, 2002.

Pöschl, U., Martin, S. T., Sinha, B., Chen, Q., Gunthe, S. S., Huffman, J. A., Borrmann, S., Farmer, D. K., Garland, R. M., Helas, G., Jimenez, J. L., King, S. M., Manzi, A., Mikhailov, E., Pauliquevis, T., Petters, M. D., Prenni, A. J., Roldin, P., Rose, D., Schneider, J., Su, H., Zorn, S. R., Artaxo, P., and Andreae, M. O.: Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon, Science, 429, 1513–1516, 2010.

Pye, H. O. T., Pinder, R. W., Piletic, I. R., Xie, Y., Capps, S. L., Lin, Y.-H., Surratt, J. D., Zhang, Z., Gold, A., Luecken, D. J., Hutzell, W. T., Jaoui, M., Offenberg, J. H., Kleindienst, T. E., Lewandowski, M., and Edney, E. O.: Epoxide pathways improve model predictions of isoprene markers and reveal key role of acidity in aerosol formation, Environ. Sci. Technol., 47, 11056–11064, 2013.

Rasmussen, R.A. and Khalil, M.A.K.: Isoprene over the Amazon Basin. J. Geophys. Res., 93: 1417-1427. doi: 10.1029/JD093iD02p01417JD093iD02p01417, 1988.

Reemtsma, T.: Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry status and needs, J. Chromatogr. A, 1216, 3687-3701, 2009.

Riva, M., Tomaz, S., Cui, T., Lin, Y.-H., Perraudin, E., Gold, A., Stone, E. A., Villenave, E., and Surratt, J. D.: Evidence for an unrecognized secondary anthropogenic source of organosulfates and sulfonates: gas-phase oxidation of polycyclic aromatic hydrocarbons in the presence of sulfate aerosol, Environ. Sci. Technol., 49, 6654-6664, 2015.

Riva, M., Da Silva Barbosa, T., Lin, Y.-H., Stone, E. A., Gold, A., and Surratt, J. D.: Characterization of Organosulfates in Secondary Organic Aerosol Derived from the Photooxidation of Long-Chain Alkanes, Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-20, 2016.

Ryall, D. B. and Maryon, R. H.: Validation of the UK Met Office's NAME model against the ETEX dataset, Atmos. Environ., 32, 4265–4276, 1998.

Roberts, J. M.: The atmospheric chemistry of organic nitrates. Atmos. Environ., 24,243-287, 1990.

Rincón, A. G., Calvo, A. I., Dietzel, M., and Kalberer, M.: Seasonal differences of urban organic aerosol composition: an ultra-high resolution mass spectrometry study. Environ. Chem., 9, 298-319, 2012.

Schmitt-Kopplin, P., Gelencśer, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong, Y., and Gebefügi, I.: Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: Organosulfates as photochemical smog constituents, Anal. Chem., 82, 8017–8026, 2010.

Seco, R., Peñuelas, J., Filella, I., Llusia, J., Schallhart, S., Metzger, A., Müller, M., and Hansel, A.: Volatile Organic Compounds in the Western Mediterranean Basin: urban and rural winter measurements during the DAURE campaign, Atmos. Chem. Phys., 13: 4291-4306, 2013.

Shalamzari, M. S., Vermeylen, R., Blockhuys, F., Kleindienst, T. E., Lewandowski, M., Szmigielski, R., Rudzinski, K. J., Spólnik, G., Danikiewicz, W., Maenhaut, W., and Claeys, M.: Characterization of polar organosulfates in secondary organic aerosol from the unsaturated aldehydes 2-*E*-pentenal, 2-*E*-hexenal, and 3-*Z*-hexenal, Atmos. Chem. Phys., 16, 7135-7148, 2016.

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, 1999.

Surratt, J. D., Lewandowski, M., Offenberg, J. H., Jaoui, M., Kleindienst, T. E., Edney, E. O., and Seinfeld, J. H.: Effect of acidity on secondary organic aerosol formation from isoprene, Environ. Sci. Technol., 41, 5363–5369, 2007.

Surratt, J. D., Gomez-Gonzalez, Y., Chan, A. W. H., Vermeylen, R., Shahgholi, M., Kleindienst, T. E., Edney, E. O., Offenberg, J. H., Lewandowski, M., Jaoui, M., Maenhaut, W., Claeys, M., Flagan, R. C., and Seinfeld, J. H.: Organosulfate formation in biogenic secondary organic aerosol, J. Phys. Chem. A, 112, 8345–8378, 2008.

Surratt, J. D., Chan, A. W. H., Eddingsaas, N. C., Chan, M. N., Loza, C. L., Kwan, A. J., Hersey, S. P., Flagan, R. C., Wennberg, P. O., and Seinfeld, J. H.: Reactive intermediates revealed in secondary organic aerosol formation from isoprene, P. Natl. Acad. Sci. USA, 107, 6640–6645, 2010.

Song, C., Gyawali, M., Zaveri, R. A., Shilling, J. E., and Arnott, W. P.: Light absorption by secondary organic aerosol from α-pinene: Effects of oxidants, seed aerosol acidity, and relative humidity, J. Geophys. Res., 118 (20), 11741-11749, doi: 10.1002/jgrd.50767, 2013.

Szmigielski, R., Surratt, J. D., Gomez-Gonzalez, Y., Van der Veken, P., Kourtchev, I., Vermeylen, R., Blockhuys, F., Jaoui, M., Kleindienst, T. E., Lewandowski, M., Offenberg, J. H., Edney, E. O., Seinfeld, J. H., Maenhaut, W., and Claeys, M.: 3-methyl-1,2,3-butanetricarboxylic acid: An atmospheric tracer for terpene secondary organic aerosol, Geophys. Res. Lett., 34, L24811, doi:10.1029/2007GL031338, 2007.

Tao, S., Lu, X., Levac, N., Bateman, A. P., Nguyen, T. B., Bones, D. L., Nizkorodov, S. A., Laskin, J., Laskin, A., and Yang, X.: Molecular characterization of organosulfates in organic aerosols from Shanghai and Los Angeles urban areas by nanospray-desorption electrospray ionization high-resolution mass spectrometry, Environ. Sci. Technol., 48, 10993–11001, 2014.

Tong, H., Kourtchev, I., Pant, P., Keyte, I., O'Connor, I., Wenger, J., Pope, F.D., Harrison, R.M., and Kalberer, M.: FDATMOS16 Molecular Composition of Organic Aerosols at Urban Background and Road Tunnel sites using Ultra-high Resolution Mass Spectrometry, Faraday Discuss., DOI: 10.1039/C5FD00206K, 2016.

Wang, X.K., Rossignol, S., Ma, Y, Yao, L., Wang, M.Y., Chen, J.M., George, C., and Wang, L.: Identification of particulate organosulfates in three megacities at the middle and lower reaches of the Yangtze River, Atmos. Chem. Phys. Discuss., 15, 21414-21448, 2015.

Wozniak, A. S., Bauer, J. E., Sleighter, R. L., Dickhut, R. M., and Hatcher, P. G.: Technical Note: Molecular characterization of aerosol-derived water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, Atmos. Chem. Phys., 8, 5099– 5111, 2008.

Yassine, M. M., Harir, M., Dabek-Zlotorzynska, E., and Schmitt-Kopplin, P: Structural characterization of organic aerosol using Fourier transform ion cyclotron resonance mass spectrometry: aromaticity equivalent approach. Rapid Commun. Mass Sp., 28, 2445-2454, 2014.

Zahardis, J., Geddes, S., and Petrucci, G. A.: The ozonolysis of primary aliphatic amines in fine particles, Atmos. Chem. Phys., 8, 1181-1194, 2008.

Zhao, Y., Hallar, A. G., and Mazzoleni, L. R.: Atmospheric organic matter in clouds: exact masses and molecular formula identification using ultrahigh-resolution FT-ICR mass spectrometry, Atmos. Chem. Phys., 13, 12343-12362, 2013.

Figures:

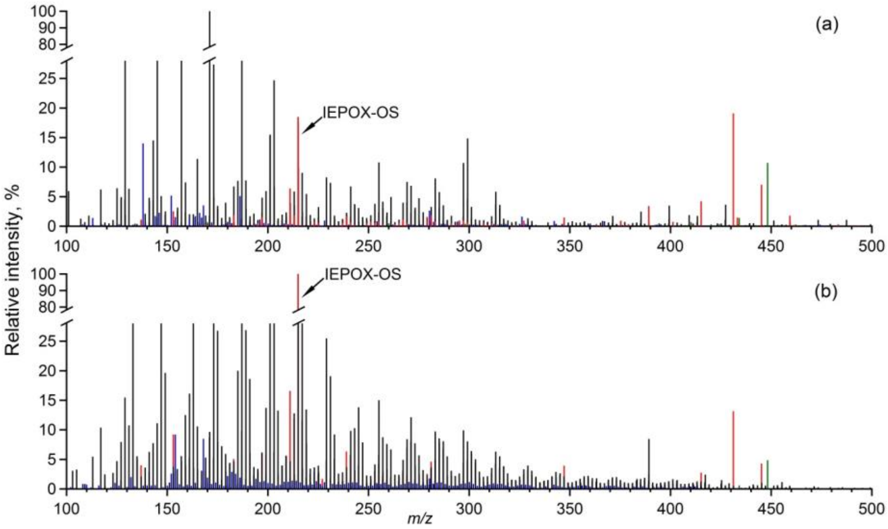


Figure 1. (-)-nanoESI-UHRMS of the representative PM2.5 samples during (a) IOP1 (b) IOP2. The line colours in the mass spectra correspond to the CHO (black), blue (CHON), CHOS (red) and CHONS (green) formulae assignments. The relative intensity axis was split to make a large number of ions with low intensities visible.

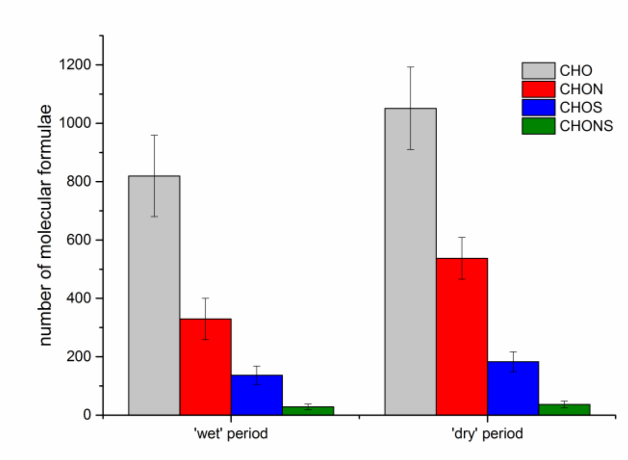


Figure 2. Average number of molecular formulae during IOP1 and IOP2. Standard deviation bars show variations between samples within individual season.

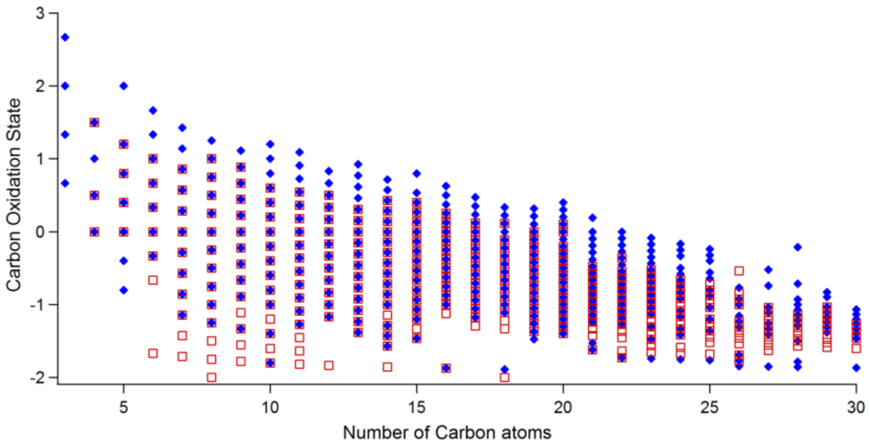
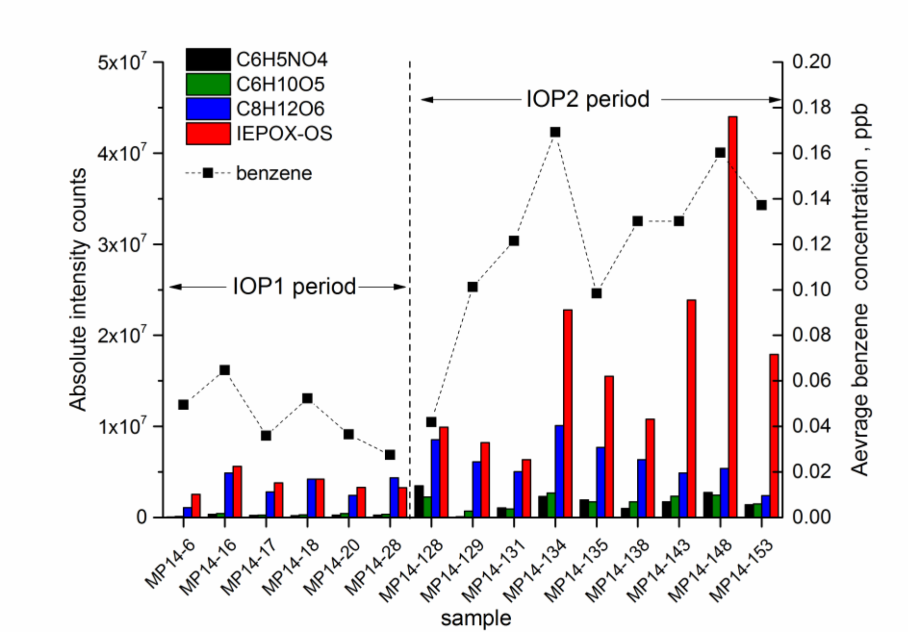


Figure 3. Carbon oxidation state plot for CHO containing formulae in organic aerosol from IOP1 (red squares) and IOP2 (blue diamonds).



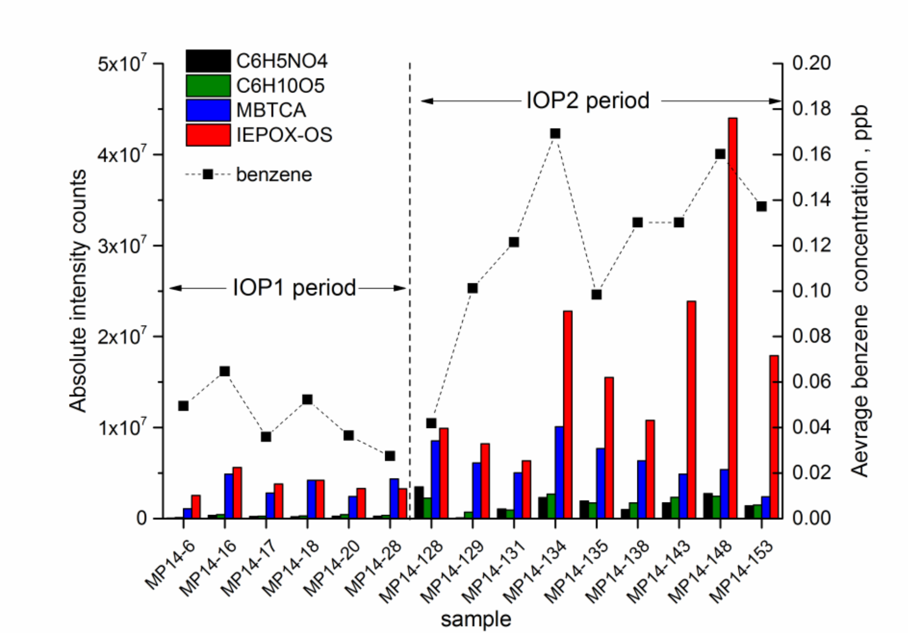


Figure 4. Ion intensity distributions (left axis) of selected tentatively identified markers in individual samples using UHRMS analysis and averaged benzene concentration (right axis)

from PTR-TOF-MS analysis. Benzene concentration was averaged for the aerosol filter sampling intervals. The UHRMS data was corrected for organic carbon load in each individual filter sample (see method section).

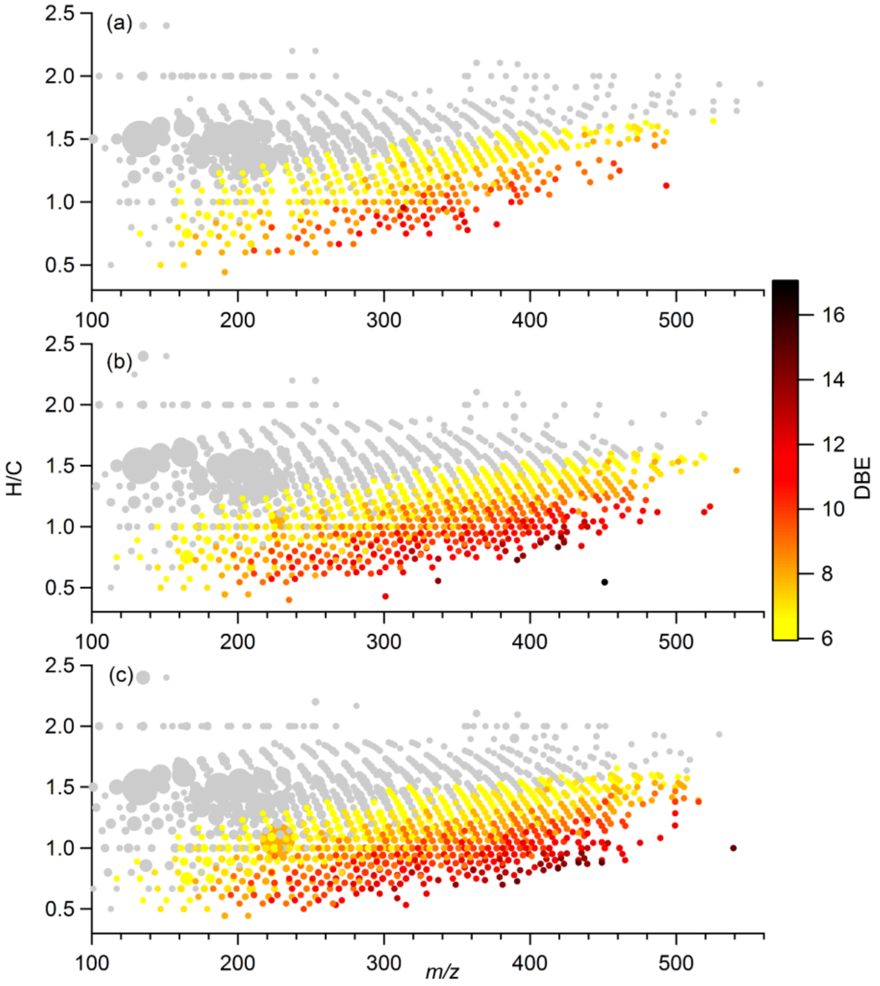


Figure 5. H/C vs m/z plot for CHO containing formulae in the samples from the periods with (a) low (b) moderately high and (c) very high incidents of fires. The marker areas reflect relative ion abundance in the sample. The colour code shows double bond equivalent (DBE) in the individual molecular formula. Molecular formulae with DBE<6 are shown as grey markers. The largest grey circles correspond to the ions at *m/z* 133.01425 (with neutral molecular formula $C_4H_6O_5$), *m/z* 187.0612 ($C_8H_{12}O_5$), *m/z* 201.07685 ($C_9H_{14}O_5$), *m/z* 203.05611 ($C_8H_{12}O_6$), and *m/z* 215.05611 ($C_9H_{12}O_6$).

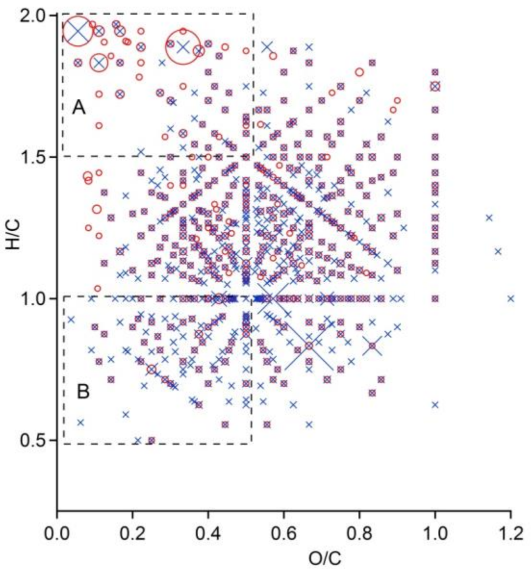
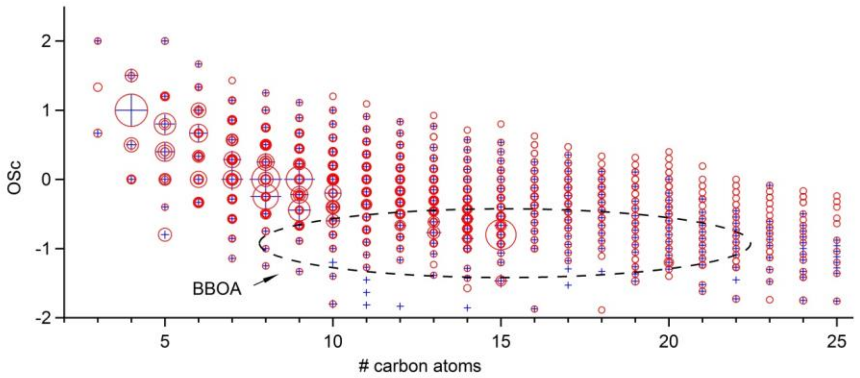


Figure 6. Overlaid Van Krevelen diagrams for CHON containing formulae in the samples from the periods with low (red markers) and very high incidents (blue markers) of fires. The marker areas reflect relative ion abundance in the sample. Areas ‘A’ and ‘B’ indicate differences in the number of ions tentatively attributed to aliphatic and aromatic species, respectively.



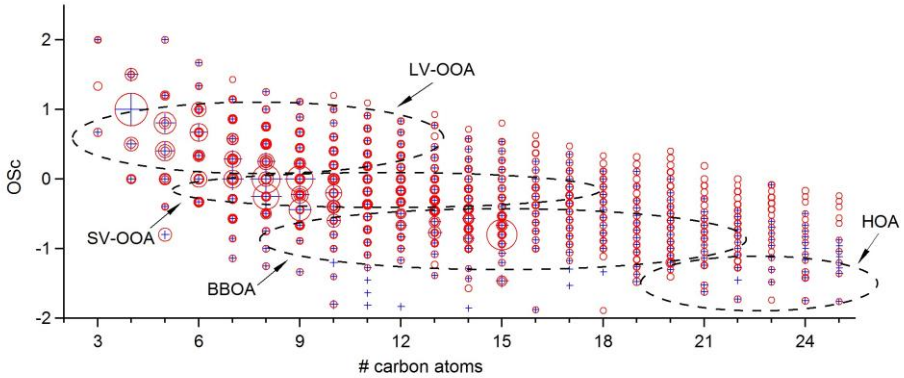


Figure 7. Overlaid carbon oxidation state (OSc) plots for CHO subgroups in the samples from the periods with low (blue markers) and very high (red markers) incidents of fires. The marker areas reflect relative ion abundance in the sample. The area marked as SV-OOA, LV-OOA, BBOA and HOA correspond to the molecules associated with semivolatile and low-volatility oxidised organic aerosol, biomass burning organic aerosol and hydrocarbon-like organic aerosol as outlined by Kroll et al. (2011).

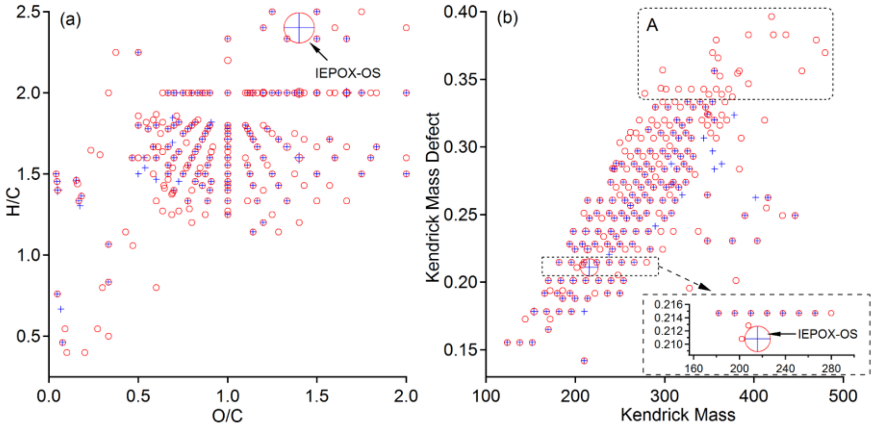


Figure 8. Overlaid Van Krevelen diagram (a) and Kendrick Mass Defect plot (b) for CHOS containing formulae in the samples from the periods with low (blue markers) and very high incidents of fires (red markers). The marker areas reflect relative ion abundance in the sample. Red markers correspond to the ions from the period with the lowest incidents of fires. Note that IEPOX-OS is not a part of any homologous series in the sample with very low incident of fires and only one additional homologue in the sample that experienced very high incident of fires (see enlarged area of the Fig 8a). Area 'A' in Kendrick Mass Defect (KMD) plot shows formulae with KMD>0.33 that are mainly present in the sample with high incident of fires.