Mapping gas-phase organic reactivity and concomitant secondary organic aerosol formation: chemometric dimension reduction techniques for the deconvolution of complex atmospheric datasets

by

K. P. Wyche *et al.*

The authors are grateful to both referees for their excellent, thoughtful and insightful reviews, their comments were most welcome and indeed were very useful for manuscript improvements to be implemented. The following document addresses all of the points raised by both reviewers and explicitly details all of the correction made to the manuscript (revised manuscript attached separately with "track changes" documenting alterations) in accordance with both of the referee's suggestions.

Author response key:

Plain, green text denotes referee comments

Italicised, back text, prefixed with "Author Response:" denotes direct author response to preceding referee comment

Plain, blue text, inside " " marks, denotes author changes to the manuscript in response to requests or recommendations from the reviewer

Anonymous Referee #1

Received and published: 26 February 2015

The manuscript of Wyche et al. represents an exciting step forward in the analysis of complex data sets, relevant to the formation of secondary organic aerosol (SOA). The authors use three statistical analysis approaches (PCA, HCA, and PLS-DA) for dimension reduction. The methods are applied to data obtained from a number of chamber studies involving different precursors grouped into four categories; application of the methods reproducibly resulted in successful classification of the gas- and particle-phase composition spectra by precursor category. While the authors suggested the potential for such approaches in ambient data mining, the results were not overstated and potential limitations were acknowledged. The paper was well written and the methodology was easy to follow. This paper should be of great interest to ACP readers involved in acquisition and interpretation of such data sets, as well as to readers interested in improving model representation of SOA.

Author Response: The authors thank the referee for their summary and recommendation.

Content

Studies have shown that limonene for example typically has higher SOA yields than other cyclic monoterpenes (Lee et al., 2006, JGR 111 D17305; Fry et al., 2014, ES&T 48: 11944-11953). Though the oxidation products of the two cyclic monoterpenes were separable from the other categories by the statistical approaches, did the results also suggest differences between the two precursors that may help explain such observations?

Author Response: To a certain degree the monoterpenes, limonene and α-pinene were separable within the monoterpene group, with three out of the four α-pinene experiments located to the upper and right region of the cluster. The referee is correct, this distribution within the statistical space results from the differences in respective gas-

phase organic compositions of the two precursor types, which then in turn impact SOA yield and composition. As such the text has been amended in Section 5, such that there are now 3 distinct discussion sections: 5.1 Mapping chemistry (the original text), 5.2 Mapping within a class (i.e. the following text supplied in response to this specific comment by referee 1) and 5.3 Mapping reactivity (i.e. additional text supplied in response to the following comment from referee 1 – see below – and comment 14) from referee 2). The additional text has been inserted on page 1676, after line 2, to answer the referees' question and to include the references that the referee has suggested:

"*5.2 Mapping within a class*

Within the monoterpene group there is a small degree of separation between the limonene and α-pinene experiments, with three out of the four α-pinene experiments located to the upper and right region of the monoterpene cluster. This distribution/separation within the group may be a consequence of precursor-specific reaction pathways; for instance, although structurally similar, α-pinene and limonene react at somewhat different rates with respect to both OH and O_3 (Atkinson and Arey, 2003). Over a fixed time period, such system reactivity will govern the degree of oxygenated content present within a closed analyte matrix and may facilitate the isolation of specific reaction pathways. Furthermore, the separation of such similar gas-phase precursors within a class cluster may help us to elucidate differences in resultant SOA yield and composition (*e.g.* limonene tends to have a larger SOA yield than α-pinene; Lee et al., 2006; Fry et al., 2014). It therefore may be possible with the use of larger and more detailed data sets to employ loading information to determine the importance of certain products to SOA composition. However, additional data to those reported here would be required to fully test this hypothesis"

Along the same lines, within a category, does the spread/distance of clusters/specificity indicate dependence on VOC/NO levels, RH, or other factors thought to influence SOA formation?

Author Response: *This is another interesting point, which we were originally intending to reserve for future work. However, following this comment and comment 13) from referee 2, we have now included a small separate analysis of the toluene data, which was conducted over a reasonable range of VOC/NO_x conditions (low, medium, high NO_x), to demonstrate how the experiments of one specific precursor type distribute/separate within the original class on account of system driving specifics, in this case, VOC/NO_x ratio. This new information has been included in a new section, Section 5.3, following the original text on page 1676, after line 2, along with an additional Figure (10), shown here:*

"5.3 Mapping reactivity

In order to explore how the PCA technique can be used to investigate product distributions driven by certain starting conditions, a separate analysis was conducted on the five toluene experiments. In this instance we investigate the product distribution dependency on initial VOC/NO_x ratios. The VOC/NO_x ratios employed nominally represent "low", "medium" and "high" NO_x conditions, with values of roughly 11 (i.e. low NO_x, NO_x limited; two experiments), 4 (i.e. moderate NO_x, two experiments) and 1 (i.e. high NO_x, VOC limited; one experiment), respectively (e.g. see work by Wagner et al., 2003). The resultant PCA loadings bi-plot (produced using the methodology described in Section 3) is given in Figure 10.

From inspection of the PCA loadings bi-plot in Figure 10, it is clear that the toluene oxidation spectra distribute in statistical space according to their respective VOC/NO_x ratios. Figure 10 shows the low NO_x, high VOC/NO_x ratio experiments grouped in the lower right-hand quadrant of the PCA space, principally influenced by loadings representing toluene (m/z 93 and 77, parent and fragment ions, respectively; note m/z 93 off-scale in Figure 10) and cresol (m/z 109). Summed spectra containing lager quantities of precursor would suggest the presence of a less reactive environment, which is the case here, where low NO_x levels in the NO_x limited regime, result in low [O_3] (via $NO_2 \rightarrow O(^3P)$ photolysis) and low [OH] (via $O_3 \rightarrow O(^1D)$ photolysis). Similarly, the *relatively large contribution from cresol to the low NO_x summed spectra, originates from a larger net cresol concentration across the experiment on account of low system reactivity (i.e. loss via reaction with OH).*

The moderate NO_x, medium VOC/NO_x experiments group uniquely in the lower left-hand quadrant of the PCA space in Figure 10, principally on account of loadings representing benzaldehyde (m/z 107) and the ring-opening products, citraconic anhydride (m/z 113), 4-oxo-2-pentenal, maleic anhydride and/or angelicalactone (m/z 99) and methyl glyoxal (m/z 73). The greater abundance of higher generation, ring-opening products implies a more reactive environment than that formed under low NO_x conditions. Larger net benzaldehyde concentrations originate from greater system reactivity and greater abundance of NO to fuel the RO_2 + NO reaction.

The high NO_x, low VOC/NO_x ratio experiment is sited in the left-hand half of Figure 10, on account of it possessing higher system reactivity (with respect to the low NO_x experiments) and the resultant presence of ring-opening product ions (as the case for the moderate NO_x experiments). However, the low VOC/NO_x ratio experiment is uniquely displaced into the upper region of the PCA space owing to a large contribution from 2-butenedial and/or 2(5H)-furanone (m/z 85, off scale in Figure 10) to the summed spectra (the yields of both of which are likely to be important under high NO_x conditions, owing to reaction through the RO_2 + NO channel).”

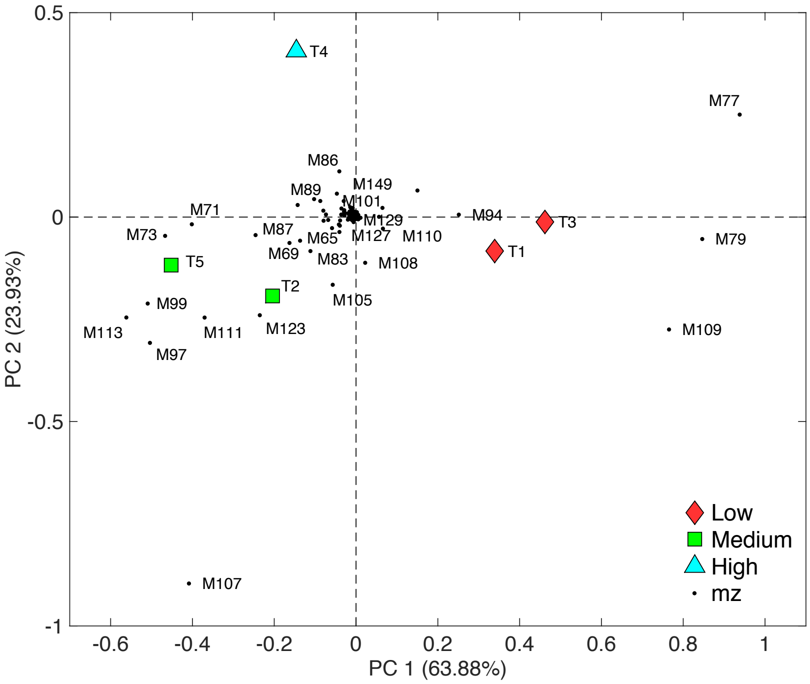


Figure 10: PCA loadings bi-plot of the second *vs.* first principal components derived from the PCA analysis of the toluene experiments. Experiments were conducted under low NO_x, high VOC/NO_x ratio (red diamonds), moderate NO_x, medium VOC/NO_x ratio (green squares) and high NO_x, low VOC/NO_x ratio (blue triangle) conditions. For clarity, the scale has been set to show the bulk of the data, hence *m/z* 93 and 85 are not shown.

Lee et al. reported a number of m/z fragments for oxidation products of isoprene, monoterpenes, and a sesquiterpene detected by PTR-MS. It would be interesting to know if the same conclusions could be ascertained from the reported m/z fragments in Lee et al.; namely that there are significant relationships between the fragments of “like” precursors. There may not be enough data to perform a full analysis as was done in the subject manuscript, but the authors may think about whether their results are similar/supported by Lee et al.

Author Response*: The referee is indeed correct, there are similarities between the fragments of products of “like” precursors. Here, the main similarities are between the* *monoterpenes limonene and α-pinene, and myrcene and linalool. However, we have highlighted this information for the major oxidation products in Section 4.2, on page 1666, line 15 onwards, e.g. "*The monoterpene groupings are influenced by the presence of m/z 107, 151 and 169 (primary aldehydes, piononaldehyde and limononaldehyde) and 139 (primary ketone, limonaketone) ions in their mass spectra" *and* "Helping to separate the straight chain from cyclic monoterpenes are m/z 95 and 93, dominant features in both the myrcene and linalool spectra (relative abundance 10–24% for m/z 93). m/z 93 has previously been identified as a major fragment ion of first generation myrcene and linalool products 4-vinyl-4-pentenal and 4-hydroxy-4-methyl-5-hexen-1-al, respectively (Shu et al., 1997; Lee et al., 2006)"

However, following this comment from referee 1 and Technical Comment 5) from referee 2, the text has been amended as follows on page 1666, line 28, to link the findings of this work to that of Lee at al., 2010:

"A list of major ions contributing to the separation of spectra in statistical space is given in Table 3 along with potential identities and precursors. It is worthy of note here that these ions and the overall fragmentation patterns observed in this study are largely in-line with those reported by Lee et al. (2006), in their comprehensive PTR-MS analysis of a wide range of BVOC precursors and their associated oxidation products."

It is suggested that the authors carefully review their citations. There are several places in which widely cited previous work is omitted in favor of self-citations.

Author Response: We thank the referee for pointing this out and we have amended our citations as follows, however we would like to highlight that in some instances we include numerous citations to our previous work in order to explain specifics of the experiments that are employed here.

One example can be found on p. 1655, line 22, regarding the importance of findings from chamber studies over the past decade. The authors do not cite the work coming out of the Caltech chamber (as early as the late 1990's); the parameters from Hoffman et al., 1997 and Griffin et al., 1999 still are widely used in chemical transport models (see also papers by Cocker, D.).

Restricted to the last 10 years, lead authors on Caltech chamber publications include: Ng., N.; Kroll, J.; Surratt, J. and others.

Much influential work has also been conducted over the last 10 years in the Carnegie Mellon chamber; lead authors include: Presto, A., Hildebrandt, L., Grieshop, A., and others.

Author Response: The referee is indeed correct, this is an oversight on our part and the following citations and references have been added to the amended text on,

Page 1655, line 22:
"A significant proportion of the findings gained regarding SOA over the last decade and more have come from atmospheric simulation chamber experiments, intermediate in complexity between classical single-process experiments and the fully open system (for various different chamber systems and VOC precursors, see for example, Pandis et al., 1991; Odum et al., 1996; Hoffmann et al., 1997; Griffin et al., 1999; Glasius et al., 2000; Cocker et al., 2001; Jaoui and Kamens, 2003; Kleindienst et al., 2004; Presto et al., 2005; Bloss et al., 2005; Rohrer et al., 2005; Ng et al., 2006, 2007; Dommen et al., 2006; Surrat et al., 2006; Grieshop et al., 2007; Chan et al., 2007; Wyche et al., 2009; Hildebrandt et al., 2009; Rickard et al., 2010; Camredon et al., 2010; Chhabra et al., 2011; Hennigan et al., 2011; Jenkin et al., 2012)."

Page 1658, line 6:
"all have previously been shown to form SOA under simulation chamber conditions (e.g. Hoffmann et al., 1997; Griffin et al., 1999; Glasius et al., 2000; Jaoui and Kamens, 2003; Presto et al., 2005; Ng et al., 2006; Surratt et al., 2006; Dommen et al., 2006; Lee et al., 2006; Hallquist et al., 2009; Alfarra et al., 2013, and references therein)."

Page 1665, line 6:

"Full details describing the underlying chemical and physical mechanisms at play within such experiments can be found elsewhere (e.g. Larsen et al., 2001; Bloss et al., 2005; Paulsen et al., 2005; Surratt et al., 2006 and 2010; Wyche et al., 2009; 2014; Camredon et al., 2010; Rickard et al., 2010; Eddingsaas et al., 2012b; Hamilton et al., 2011; Jenkin et al., 2012; Alfarra et al., 2012, 2013; and references therein)."

Page 1670, line 9:

"For further details regarding β-caryophyllene oxidation products, see for example Lee et al. (2006), Winterhalter et al. (2009), Hamilton et al. (2011), Chan et al. (2011), Li et al. (2011) and Jenkin et al. (2012) and references therein, and Sect. 5."

Page 1670, line 18:

"For further details regarding α-pinene and limonene oxidation products, see for example Larsen et al. (2001), Jaoui et al. (2003), Capouet et al. (2004), Jenkin (2004), Jaoui et al. (2006), Lee et al. (2006), Ng et al. (2006), Camredon et al. (2010) and Hamilton et al. (2011) and references therein."

Page 1675, line 12:

"β-caryophyllene readily forms particulate matter on oxidation (e.g. Jaoui et al., 2003; Lee et al., 2006; Winterhalter et al., 2009; Alfarra et al., 2012; Chen et al., 2012)"

References added to manuscript list:

Pandis, S. N., Paulson, S. E., Seinfeld, J. H., and Flagan, R. C.: Aerosol Formation in the Photooxidation of Isoprene and Beta-Pinene, Atmos. Environ., 25(5-6), 997–1008, doi: 10.1016/0960-1686(91)90141-S, 1992.

Odum, J. R., Hoffmann, T., Bowman, F., Collins, D., Flagan, R. C., and Seinfeld, J. H.: Gas/particle partitioning and secondary organic aerosol yields, Environ. Sci. Technol., 30, 2580–2585, 1996.

Hoffmann, T., Odum, J. R., Bowman, F., Collins, D., Klockow, D., Flagan, R. C., and Seinfeld, J. H.: Formation of Organic Aerosols from the Oxidation of Biogenic Hydrocarbons, J. Atmos. Chem., 26, 189-222, 1997.

Griffin, R. J., Cocker, D. R., III, Flagan, R. C., and Seinfeld, J. H.: Organic aerosol formation from the oxidation of biogenic hydrocarbons, J. Geophys. Res., 104, 3555–3567, 1999.

Glasius, M., Lahaniati, M., Calogirou, A., Bella, D. D., Jensen, N. R., Hjorth, J., Kotzias, D., and Larsen, B. R.: Carboxylic acids in secondary aerosol from oxidation of cyclic monoterpenes by ozone, Environ. Sci. Technol., 34, 1001–1010, 2000.

Cocker III, D. R., Mader, B. T., Kalberer, M., Flagan, R. C., and Seinfeld, J. H.: The effect of water on gas-particle partitioning of secondary organic aerosol: II. m-xylene and 1,3,5-trimethylbenzene photooxidation systems, Atmos. Env., 35, 6073-6085, 2001.

Jaoui, M. and Kamens, R. M.: Gaseous and particulate oxidation products analysis of a mixture of α-pinene + β-pinene/O3/air in the absence of light and α-pinene + β-pinene/NOx/air in the presence of natural sunlight, J. Atmos. Chem., 44, 259–297, 2003.

Jaoui, M., Leungsakul, S., and Kamens, R. M.: Gas and particle products distribution from the reaction of beta-caryophyllene with ozone, J. Atmos. Chem., 45, 261–287, 2003.

Kleindienst, T. E., Conver, T. S., McIver, C. D., and Edney, E. O.: Determination of secondary organic aerosol products from the photooxidation of toluene and their implications in ambient PM2.5, J. Atmos. Chem., 47, 79–100, 2004.

Presto, A. A., HuffHartz, K. E., and Donahue, N. M.: Secondary Organic Aerosol Production from Terpene Ozonolysis. 2. Effect of NOx Concentration, Environ. Sci. Technol., 39, 7046–7054, 2005.

Rohrer, F., Bohn, B., Brauers, T., Bruning, D., Johnen, F. –J., Wahner, A., and Kleffmann, J.: Characterisation of the photolytic HONO-source in the atmosphere simulation chamber SAPHIR, Atmos. Chem. Phys., 5, 2189-2201, 2005.

Jaoui, M., Corse, E., Kleindienst, T. E., Offenburg, J. H., Lewandowski, M., and Edney, E. O.: Analysis of Secondary Organic Aerosol Compounds from the Photooxidation of *d*-Limonene in the Presence of NO_x and their Detection in Ambient $PM_{2.5}$, Environ. Sci. Technol., 40(12), 3819-3828, doi: 10.1021/es052566z, 2006.

Ng, N. L., Kroll, J. H., Keywood, M. D., Bahreini, R., Varutbangkul, V., Flagan, R. C., Seinfeld, J. H., Lee, A., and Goldstein, A. H.: Contribution of first- versus second-generation products to secondary organic aerosols formed in the oxidation of biogenic hydrocarbons, Environ. Sci. Technol., 40, 2283–2297, 2006.

Dommen, J., Metzger, A., Duplissy, J., Kalberer, M., Alfarra, M. R., Gascho, A., Weingartner, E., Prevot, A. S. H., Verheggen, B., and Baltensperger, U.: Laboratory observation of oligomers in the aerosol from isoprene/NOx photooxidation, Geophys. Res. Lett., 33, L13805, doi:10.1029/2006gl026523, 2006.

Grieshop, A. P., Donahue, N. M., and Robinson, A. L.: Is the gas-particle partitioning in alpha-pinene secondary organic aerosol reversible?, Geophys. Res. Lett., 34, L14810, doi:10.1029/2007GL029987, 2007.

Chan, A. W. H., Kroll, J. H., Ng, N. L., and Seinfeld, J. H.: Kinetic modeling of secondary organic aerosol formation: effects of particle- and gas-phase reactions of semivolatile products, Atmos. Chem. Phys., 7, 4135–4147, doi:10.5194/acp-7-4135-2007, 2007.

Ng, N. L., Kroll, J. H., Chan, A. W. H., Chhabra, P. S., Flagan, R. C., and Seinfeld: Secondary organic aerosol formation from *m*-xylene, toluene, and benzene, Atmos. Chem. Phys., 7, 3909–3922, 2007.

Hildebrandt, L., Donahue, N. M., and Pandis, S. N.: High formation of secondary organic aerosol from the photo-oxidation of toluene, Atmos. Chem. Phys., 9, 2973–2986, doi:10.5194/acp-9-2973-2009, 2009.

Winterhalter, R., Herrmann, F., Kanawati, B., Nguyen, T. L., Peeters, J., Vereecken, L., and Moortgat, G. K.: The gas-phase ozonolysis of β-caryophyllene (C15H24). Part I: an experimental study, Phys. Chem. Chem. Phys., 11, 4152–4172, 2009.

Chan, M. N., Surratt, J. D., Chan, A. W. H., Schilling, K., Offenberg, J. H., Lewandowski, M., Edney, E. O., Kleindienst, T. E., Jaoui, M., Edgerton, E. S., Tanner, R. L., Shaw, S. L., Zheng, M., Knipping, E. M., and Seinfeld, J. H.: Influence of aerosol acidity on the chemical composition of secondary organic aerosol from β-caryophyllene. Atmos. Chem. Phys., 11, 1735–1751, doi:10.5194/acp-11-1735-2011, 2011.

Chhabra, P. S., Ng, N. L., Canagaratna, M. R., Corrigan, A. L., Russell, L. M., Worsnop, D. R., Flagan, R. C., and Seinfeld, J. H.: Elemental composition and oxidation of chamber organic aerosol, Atmos. Chem. Phys., 11, 8827–8845, doi:10.5194/acp-11-8827-2011, 2011.

Hennigan, C. J., Miracolo, M. A., Engelhart, G. J., May, A. A., Presto, A. A., Lee, T., Sullivan, A. P., McMeeking, G. R., Coe, H., Wold, C. E., Hao, W. –M., Gilman, J. B., Kuster, W. C., de Gouw, J., Schichtel, B. A., Collet Jr, J. L., Kreidenweis, S. M., and Robinson, A. L.: Chemical and physical transformations of organic aerosol from the photo-oxidation of open biomass burning emissions in an environmental chamber, Atmos. Chem. Phys., 11, 7669-7686, 2011.

Li, Y. J., Chen, Q., Guzman, M. I., Chan, C. K., and Martin, S. T.: Second-generation

products contribute substantially to the particle-phase organic material produced by β-caryophyllene ozonolysis, Atmos. Chem. Phys., 11, 121–132, doi:10.5194/acp-11-121-2011, 2011.

Chen, Q., Li, Y. L., McKinney, K. A., Kuwata, M., and Martin, S. T.: Particle mass yield from β-caryophyllene ozonolysis, Atmos. Chem. Phys., 12, 3165-3179, doi: 10.5194/acp-12-3165-2012, 2012.

Editorial

p 1655, line 18: "Atmospheric chemistry" as a system is awkward. One suggestion: "The chemistry of the atmospheric system is highly nonlinear. . ."

Author Response: The text has been altered accordingly

---

p 1656, line 9: The "CIR-TOF-MS" abbreviation can be introduced on line 4 and used exclusively here.

Author Response: The text has been altered accordingly

---

p. 1658, section 2 heading: The adjective "Experimental " needs a verb, or needs to be changed to a noun.

*Author Response: The text has been altered to read "**Experimental Details**"*

---

p. 1661, line 8: Remove "/" prior to photolytic

Author Response: The text has been altered accordingly

---

p. 1666, line 19: Change "caryophyllon" to "caryophyllene"

Author Response: The text has been altered accordingly

---

p. 1667, line 26: It is not clear what is meant by “oxidized atmospheres”.

Author Response: For clarity, the text has been altered to read,

“As a further test of the technique to distinguish between and to classify VOCs and the matrix of oxidized organic compounds that may derive from their atmospheric chemistry, test data from an anthropogenic system was introduced into the model.”

p. 1669, line 10-13: It is suggested the sentence be reworded to more clearly indicate that data are available for only two experiments per each of the two noncyclic monoterpenes. As written the focus is on “only two types of precursor”, which is the same as the cyclic monoterpenes. In summary, it is the number of experiments, not the number of precursors that is likely affecting the result.

Author Response: the text has been altered to read,

“The greater spread in confidence of the noncyclic monoterpene group is once again likely to result to some extend from the low number of repeat experiments employed (i.e. only two each for myrcene and linalool).”

Can the panel and text sizes be increased for Figs. 7 and 8? They are hard to read (even w/significant expansion-175%)?

Author Response: Figures 7a, 7b, 8a and 8b have been altered as requested, with increased panel and text sizes. We will also work with the editing team to ensure in typesetting that all figures are sufficiently clear.

The authors may consider more clearly indicating the difference between Figs. 7 and 8 in the figure panels (e.g., just adding a AMS and LC-MS header).

Author Response: Figures 7a, 7b, 8a and 8b have been altered as requested, with additional headers to indicate LC-MS/MS or cTOF-AMS data.

The authors may consider using consistent colors for the different categories in each of the plots (generally the same in PCA plots, but not in dendorgrams).

Author Response: *We thanks the referee for pointing this out, but we would prefer to leave the dendrogram colors as there are, as they represent the best combination for the clearest description of the data, i.e. the dendrogram colors were chosen specifically to optimize clarity of view within the figure. The colors and symbols are the same for all PCA and PLS-DA plots, e.g. isoprene is always represented by a blue diamond.*

Caption Figure 10: Remove "/" after abundance

Author Response: *The text has been altered accordingly*

Anonymous Referee #2

Received and published: 10 March 2015

Overview:

Wyche et al. describe development of a chemometric mapping of gas-phase and particle-phase matrices from oxidation of several relevant compounds: isoprene, terpenes (i.e. linalool, myrcene, limonene, a-pinene, B-caryophyllene), toluene, as well as the oxidation products of fig and birch trees. This is done via principal component analysis (PCA), hierarchical cluster analysis (HCA), and positive least squares-discriminant analysis (PLS-DA) on the mass spectra of the oxidation products from several systems. PCA is performed on gas-phase observations of chamber data using chemical ionization reaction time-of-flight mass spectrometry for the gas-phase to separate distinct regions associated with oxidation of isoprene, cyclic monoterpenes, B-caryophyllene, single chain monoterpenes, and toluene. Addition of mesocosm data (plant emission oxidation experiments) shows expected mapping of oxidation products from fig and birch trees as related to the expected emissions being isoprene-dominated and cyclic monoterpene-dominated, respectively. HCA analysis is performed, supporting the relation of the mesocosm oxidation products with associated precursor oxidation schemes represented in chamber data. Similar PCA and HCA analysis applied to the particle-phase data (liquid chromatography-ion trap mass spectrometry and compact time-of-flight aerosol mass spectrometry) also shows separation by precursor type. Using a zero-dimensional box model simulation based on the Master Chemical Mechanism for alpha-pinene photooxidation, it is argued that the employed statistical deconvolution techniques could be applicable for determining precursor type and potential mechanisms from ambient data on the basis that “model mass spectra” simulated under increasingly complex (closer to ambient conditions) are generally well-captured by “model mass spectra” under conditions more like a typical chamber experiment.

Reviewer's recommendation:

The article overall is well-written and is novel in that it seems to provide great potential for use as identifying chemistry from a particular precursor type given complex sets of mass spectra from oxidation products in the gas and particle phases. The weakest argument is that made for use of this technique to elucidate chemical mechanisms in addition to identifying precursor type, as the specificity for separating mechanistic pathways is not fully demonstrated in the current analyses. Still, the content is appropriate for ACP readers and I would recommend publication after the following comments are addressed.

Author Response: The authors thank the referee for their summary and recommendation.

General Comments:

1) Abstract: The reader would benefit from more specific result statements included here, similar to the lines included in the Introduction section p. 1657, lines 7-15.

Author Response: The following text has been added as requested,

Abstract, line 18, page 1653,

"Results show that "model" biogenic oxidative systems can be successfully separated and classified according to their gaseous oxidation products."

Abstract, line 26, page 1653,

"More specifically, the addition mesocosm data from fig and birch tree experiments shows that isoprene and monoterpene emitting sources, respectively, can be mapped onto the statistical model structure and their positional vectors can provide insight into their biogenic sources and controlling oxidative chemistry."

2) P. 1656, lines 20-25: Discussion of currently used statistical techniques is rather cursory. There is not even mention of the commonly used positive matrix factorization

technique widely used for AMS data analysis. Additional details of why the presented technique is novel/necessary should be discussed.

Author Response: *We have amended the introduction as follows; to include a more detailed discussion of other more commonly used statistical techniques, including positive matrix factorization. The additional details regarding why the technique presented here is novel are addressed in reply to Referee 2, point 3) below:*

"Similar approaches using statistical analyses have been recently applied to both detailed and broad ambient aerosol composition data (e.g. (Heringa et al., 2012;Paglione et al., 2014)), particularly in the context of source apportionment (e.g. (Alier et al., 2013)). Different methods have been attempted by several groups to deconvolve organic aerosol spectra measured by the Aerosol Mass Spectrometer (AMS) in particular (e.g. Zhang et al., 2005, 2007; Marcolli et al., 2006; Lanz et al., 2007). Zhang et al. (2005) applied a custom principal component analysis (CPCA) method to extract two distinct sources of organic aerosols in an urban environment using linear decomposition of AMS spectra and later applied a Multiple Component Analysis technique (MCA, an expanded version of the CPCA) to separate more than two factors in datasets from 37 field campaigns in the Northern Hemisphere (Zhang et al., 2007). Marcolli et al. (2006) applied a hierarchical cluster analysis method to an ambient AMS data set, and reported clusters representing biogenic VOC oxidation products, highly oxidised organic aerosols and other small categories. Receptor modelling techniques such as Positive Matrix Factorization (PMF) employ similar multivariate statistical methods in order to deconvolve a time series of simultaneous measurements into a set of factors and their time-dependent concentrations (Paatero and Tapper, 1994; Paatero, 1997). Depending on their specific chemical and temporal characteristics, these factors may then be related to emission sources, chemical composition and atmospheric processing. For example, Lanz et al. (2007) and Ulbrich et al., (2009) applied PMF to the organic fraction of AMS datasets and were able to conduct source apportionment analysis identifying factors contributing to the composition of organic aerosol at urban locations. Slowik et al. (2010), combined both particle-phase AMS and gas-phase proton transfer reaction mass spectrometry data for the PMF analysis of urban air, and were able to successfully obtain “regional transport, local traffic, charbroiling and oxidative process” factors. By combining the two datasets, Slowik and colleagues were able to acquire more in-depth information regarding the urban atmosphere than could be derived from the analysis of each of the sets of measurements on their own.

Because receptor models require no a priori knowledge of meteorological conditions or emission inventories, they are ideal for use in locations where emission inventories are poorly characterised or highly complicated (e.g. urban areas), or where atmospheric processing plays a major role. However, because all of the values in the profiles and contributions are constrained to be positive, the PMF model can have an arbitrary number of factors and the user must select the “best” solution that explains the data. This subjective step of PMF analysis relies greatly on the judgment and skill of the user.”

3) The authors mention p. 1657, lines 16-25, the potential for this technique to be used on ambient data sets, but the paper would be stronger references were provided for which similar statistical analyses are being done to map out oxidation chemistry related to certain precursors as is done here, or utilize available and published field data in the described analyses of the paper to prove the point. For example, why would this technique be useful over positive matrix factorization techniques, which now includes a way of identifying SOA formation contributions from isoprene?

Author Response: *The referee is indeed correct; the paper would be stronger if we could have included ambient data in the statistical analyses in the same way that we included the mesocosm data. Unfortunately, we do not have such data available and have tried to address the issue to some extend using the MCM simulations to model a representatively complex system. It should be noted that we could find no useable published field data (as the referee suggests). A large amount of such data comprises solely the compounds of interest and not the detailed mass spectra, and where mass spectra were available, they were incompatible/not fit for purpose, either being e.g. flux* *data and/or comprising PTR-Quadruple-MS data recorded in selected ion mode.*

We feel (as was pointed out by referee 1) that we have been fully open regarding the potential limitations of this methodology and have attempted to address them to the best of ability.

We feel that taking the technique from first principle, to single precursor and then to mesocosm scenarios, achieves the goals that were set out and that application to the "real world" would represent the next logical step, for our research and for other colleagues; given the results presented in this proof of concept study, using this approach successfully in the field is the ultimate goal.

Regarding Positive matrix factorization (PMF) and the uniqueness of the method presented here; PMF is an extremely valuable technique now used successfully for PM source apportionment. The ensemble methodology described here however is not intended as a replacement, far from it, it simply presents a different approach to isolate composition (and potentially therefore, sources) and moreover to identify and map different chemical pathways. PMF generally utilizes data solely from particle phase measurements (however, PMF with both gas and particle phase data have been reported, i.e. Slowik et al., ACP, 2010) and hence usually from one analytical technique. However, the ensemble method discussed here utilizes data from both the gas- and particle-phases and data from different analytical techniques. Furthermore, it is not designed simply to comprise a type of source-receptor/source apportionment model; the objective is rather compositional isolation to instruct on potential sources AND on underlying chemical processes.

The methodologies described here are fundamentally simpler to use and require fewer base assumptions (there is less importance placed on user skill to determine the correct input parameters, number of factors to employ etc); the techniques work on fewer data, require no reference datasets and no prior knowledge of sources is required.

We thank the referee for this useful comment and have altered the introduction to highlight these distinctions, with the addition of the following text to line 6, page 1657:

"Unlike other statistical techniques such as PMF, the ensemble methodology presented here does not require the use of additional external databases (comprising information regarding different environments/reference spectra), is simpler to use and less labour intensive, and places less importance on user skill in the production of accurate and meaningful results. Moreover, the primary focus of techniques such as PMF is on source identification/separation, whereas here the focus is placed on compositional isolation."

---

4) P. 1659, lines 9-12: While the precursors were reacted to near completion in the chamber experiments, how can one verify that the oxidation scheme went to completion to match, for example, major oxidation products in the MCM schemes presented in Figure 9. That is, should not theoretically the PCA analysis for the chamber experiments ideally match the MCM PCA if the chamber experiments are covering the same range as the simulated chemistry?

First of all, we must reiterate as was pointed out in the manuscript, the MCM simulations are not idealized, and their inclusion is not for the purpose of making a direct model-measurement comparison. Owing to a lack of in field ambient spectra, the simulations are designed to exemplify how the gas phase spectra could change if the system complexity was increased. The model is not included for direct comparison with the measurements, it is intended as a guide to help bridge the gap between what we have reported from "model" and "mesocosm" experiments and the real world. This is stated on page 1677, line 16.

However, we can verify that precursor oxidation went to completion to match the measurements, because the model was purposefully constrained to measured α-pinene concentration, i.e. the primary reactivity in the model was constrained to match the primary reactivity in the measurements. Furthermore, from our previous work (see Camredon et al., 2010) the modeled and measured temporal profiles of the precursor and first (and later) generation products are in reasonably good agreement, suggesting that the reactivities of the systems are comparable (i.e. that the model system is *chemically dynamic in a similar way as you would see in the chamber). Use of the MCM under such single precursor chamber conditions has been evaluated in details on several previous occasions, for example see work by Bloss et al., 2005 (toluene), Camredon et al., 2010 – supplementary information (α-pinene), Rickard et al., 2010 (1,3,5-TMB) and Jenkin et al., 2012 (β-caryophyllene). The reader has now been informed of this with the addition of the following text to page 1677, line 18:*

"For work regarding the evaluation of the MCM with respect to single VOC precursor chamber experiments (including model-measurement intercomparison), see for example, Bloss et al., 2005 (toluene), Metzger et al., 2008 and Rickard et al., 2010 (1,3,5-TMB), Camredon et al., 2010 (α-pinene) and Jenkin et al., 2012 (β-caryophyllene)."

---

5) P. 1659, Section 2.2: What impacts on the comparison of the chemical oxidation would be expected for the systems, considering that these systems were run across several environmental chambers?

Author Response: *Each (Teflon) chamber system will have its own specific background wall chemistry, which arguably changes by various degrees from day-to-day and even between experiments (i.e. cleanliness of chamber will vary). For certain chambers, e.g. EUPHORE (Bloss et al., 2005; Zador et al., J. Atmos. Chem., 55:147-166, 2006) and PSISC (Metzger et al., 2008), detailed auxiliary mechanisms have been constructed and tested. These chemical mechanisms are based on similar reactions and fundamental properties and go some way towards explaining the intricacies of environment chamber experiments. From our experience, where rates of reaction (influenced by radical chemistry associated with the chamber walls) can change between chambers for a give set of experimental conditions, the overall composition of the gas phase does not vary widely for the types of experiments studied. This is inherent in the data reported here, as the spectra collected are compositionally similar between chambers. The primary additives generated from the chamber walls that could influence organic gas phase composition are nitrogen-containing species such as HONO. However, for the* *experiments reported here (with the exception of the mesocosm experiments), initial [NO_x] added > [NO_x] liberated from chamber walls.*

In addition to the wall chemistry, the light intensity will vary from chamber to chamber, however, EUPHORE employs natural solar radiation and the MAC and PSISC are designed to approximate the solar spectrum, as detailed in the work referenced appropriately in the experiment section 2.2, page 1659.

In summary, we expect the impacts to be fairly minimal and within acknowledged uncertainty bounds.

---

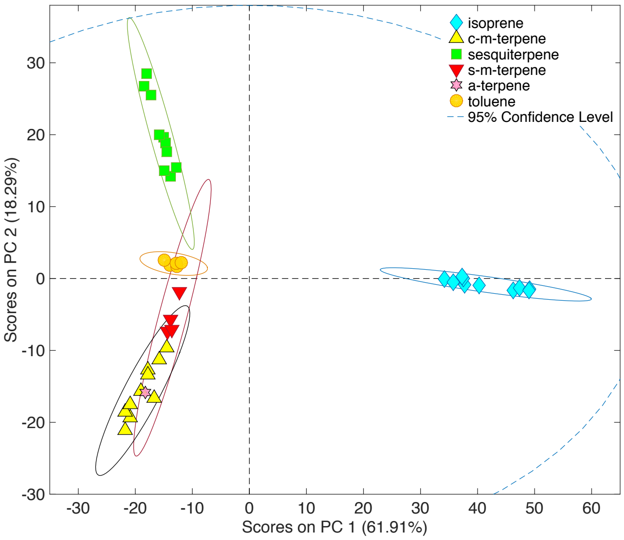
6) P. 1667, lines 2-4: Aromatic compounds are also observed to be emitted from the biosphere, (e.g. aromatics like toluene as well as aromatic monoterpenes; see Guenther et al., 2012 and references within; doi:10.5194/gmd-5-1471-2012). Would the authors expect addition of these compounds to map onto the same region as toluene/monoterpenes, and would this not change the interpretation that the analysis can separate anthropogenic and biogenic precursor sources vs. just precursor structure? Do the authors have potential ideas for why there is overlap with toluene and the single chain monoterpenes? Would single chain sesquiterpenes (e.g. farnasene) also map out separately from B-caryophyllene as the cyclic and straight chain monoterpenes do?

Author Response: *Guenther et al., (2012) estimated global annual emissions of ~ 535 Tg yr^{-1} for isoprene and a sum of ~ 147 Tg yr^{-1} for non-aromatic terpenes; total sesquiterpene emissions were of the order 29 Tg yr^{-1}. In contrast aromatic terpene emissions were estimated in a group of ~ thirty "other" monoterpenes with total annual emissions ~ 15 Tg yr^{-1}, i.e. only 2 % of Σ(isoprene, major non-aromatic terpenes and sesquiterpenes). Estimated annual toluene emissions are not individually reported by Guenther et al., but are assumed to be relatively minor, reported with an "additional 11 stress VOC(s)", with a total emissions value of 7.8 Tg yr^{-1}. Sindalevora et al. (2014) report mean isoprene, Σ(monoterpenes), Σ(sesquiterpenes) and toluene annual global emissions to be of the order 594, 95, 20 and 1.5 Tg yr^{-1}, respectively, i.e. toluene* *emissions are only ~ 0.2 % of Σ(isoprene, major non-aromatic terpenes and sesquiterpenes). In summary, biogenic aromatic compounds comprise only a small fraction of total average terpene emissions and it is likely therefore that their oxidation products would be present in such low abundance compared to those discussed within our work, that they would not cause severe spectral/statistical interference.*

Under exceptional circumstances where aromatic terpenes are present in significant abundance, it is possible that their compositional spectra could group in the vicinity of those of other monoterpenes, on account of them possessing spectra with some similar features, e.g. Lee et al. (2006; DOI: 10.1029/2006JD007050) investigated the PTR-MS spectra of methyl chavicol and reported the presence of a parent ion of m/z 149 and major product ions of m/z 151, 137, 121 and 109. Of these ions m/z 149, 151 and 137 do not appear in toluene PTR-MS spectra and moreover, the ions of m/z 151, 137 and 109 are also significant contributors to monoterpene PTR-MS oxidation spectra.

Experiments conducted with aromatic terpenes did not form the focus of this initial work and as such no data is available here to test the exact positioning of aromatic terpenes in the statistical space, however, in addition to the findings of Guenther, Sindalevora and Lee et al., the hypothesis that aromatic "like" BVOCs would not group with toluene was tested here with data from α-terpinene photooxidation. α-terpinene is not an aromatic terpene, but unlike α-pinene and limonene it does contain a 6-member carbon ring containing 2 C=C double bonds.

Results: α-terpinene oxidation produces a significant amount of α-terpinaldehyde, which like pinonaldehyde (from α-pinene photooxidation) possesses a spectral fingerprint including ions of m/z 169, 151, 123 and 107. As with other non-aromatic terpenes, α-terpinene photooxidation PTR-MS spectra also contains ions of m/z 139 (c.f. limonaketone from limonene) and 171 (also observed in α-pinene phtooxidation spectra), as well as more unique features, such as an m/z 143 ion of significant abundance (also see Lee et al., 2006, DOI: 10.1029/2006JD007050). In combination, such features help to group the α-terpinene photooxidation spectrum in the vicinity of other terpenes, *separate in statistical space from the aromatic toluene photooxidation spectra.*



The above figure and a short description matching the above text have been added to a new Supplement document and section 4.3 text has been amended accordingly in order to direct the reader appropriately:

"A brief discussion regarding aromatic BVOCs is provided separately in the supplementary material."

We do not believe that there is significant overlap between the straight chain terpenes and toluene; the overlap observed is from the high (95%) confidence limits, the magnitude of which is discussed in the main text of the paper. Also as discussed in the paper, the toluene cluster is located close to the straight chain monoterpene cluster owing to similar (but not identical) ion content (but not abundance), including primarily m/z 93. Toluene separates from the straight chain monoterpenes owing to unique spectral features such as m/z 99 and 113, resulting from ring-cleavage and the subsequent production of e.g. angelica lactone and hexenedione (and isomers thereof), *respectively.*

We have not yet performed experiments with single chain sesquiterpenes such as farnasene, however it is likely that, as with the cyclic and non-cyclic monoterpenes, single chain sesquiterpene spectra would group in the vicinity of cyclic sesquiterpenes, owing to their similar composition and chemistry, but uniquely displaced as a consequence of their structure (e.g. unsaturated, double bond content).

---

7) P. 1669, lines 17-20: Why are the birch trees sensitivity cited to be low due to only two repeat experiments, but the fig trees experiments also only having two repeat experiments have high sensitivity and specificity?

Author Response: *As stated in the main text, page 1669, Section 4.4, lines 17 – 20, the low birch sensitivity may result from a combination of (i) a low number of repeat experiments AND (ii) the aging trees producing lower emissions during the last experiment. As we state in the current text, the low sensitivity for birch trees is likely to result from a combination of these two factors.*

---

8) Section 4.5: Why do the authors choose to do separate PCA analyses on the gas-phase data versus the particle-phase data? Considering that there would be overlap for semi-volatiles (e.g. B-caryophyllinc acid as stated p. 1670, line 10), should not a PCA analysis (though mindful of the different analysis technique) be more ideal (as well as for interpreting ambient data) to map regions of any chemistry associated with various precursor types? Should it be inferred that the separation of analysis on gas vs. particle-phase data is the recommended approach to apply this type of analysis to ambient data? I would be curious to see how the analyses would compare between gas-phase only, particle-phase only, and gas and particle-phase together.

Author Response: *What the referee suggests here is an excellent idea, and one that we were hoping to execute in future work, however, such an undertaking will be extremely* *complex and a step beyond what we are trying to present here. To begin with, this task would require knowledge of (i) ALL of the species present in both the gas and aerosol phases (which the community does not possess yet) and (ii) the EXACT fragmentation patterns and ionisation routes for all of these species under each of the ionisation regimes used. If these data were available, one would then have to (re-) construct mass spectra representing the gas and particle phase matrices for each of the experiments employed, with no detection method specific alteration to the mass of each parent compound (e.g. protonation, adduct formation, fragmentation etc) or instrument biases. At present it is not possible to do this accurately, however we are steadily working towards obtaining such knowledge. Moreover, methodology such as Parallel Factor Analysis would perhaps be more suitable to perform such an investigation, which again is a whole new piece of work in itself. The purpose of this work is to introduce these new methods of working with large atmospheric data sets, partially the aim of which is to reduce computational load, not extend it, which would be required in this instance.*

---

9) Figure 7a: The plot is produced using LC-MS particle data, but what biases might be at work here based on this technique's sensitivity to certain organic aerosol products?

Author Response: In any analytical method there will be a degree of bias towards particular species. The LC-MS was operated in negative ionisation mode and so here we are looking at the species with acidic protons that can be removed easily. This leads to certain biases, as the ionisation efficiency is different for different structures. However, this plot shows the key ions that allow the mass spectral patterns of the SOA from different precursors to be separated based on the observed chemical speciation. The same analysis in positive mode (which targets different functional groups) will likely lead to some different ions being important for differentiating the precursors. However all samples were analysed using the same methods and so the impact of any systematic biases are minimised.

---

10) Figure 8a and accompanying discussion p. 1671, lines 15- p. 1672, line 3: The meaning of M43, M44 should be discussed in context as to whether the relative aging of the oxidation systems were actually comparable. Although the precursors were near complete reaction, did the oxidation develop enough such that B-caryophyllene schemes could take on additional spectral contributions from M43, M44? That is, was their sufficient oxidant such that secondary reactions (producing characteristic oxidation products and associated ions) could occur equally across precursor systems?

Author response: *the meaning of M43 and M44 has been discussed in the text (e.g. see page 1671, line 15 onwards). As can be seen from the Experimental section (and Table 1) and references therein, the MAC/AMS experiments were designed to be as comparable as possible, i.e. experimental conditions were kept similar between individual experiments, with light conditions, relative humidity and initial VOC/*NO_x *ratios kept roughly the same from experiment-to-experiment. The reader is guided to a full discussion of these experiments throughout Sections 2 and 4 (e.g. Alfarra et al., 2012 and 2013) to ensure that they are fully apprised of the experiments employed.*

The main difference between the oxidation systems investigated was in terms of reactivity of the various precursor compounds with respect to O_3 *and OH. Indeed, we purposefully chose a sufficient span of precursors to cover a range of reactivities and as such, each system exhibited a different overall reactivity, but on the time scales of our experiments, the systems were kept as comparable as possible. The references to our previous work have been made in part to account for/explain this.*

There was indeed sufficient oxidant throughout all experiments. As described in Alfarra et al., throughout each of the experiments discussed, O_3 *was produced through system photochemistry. Once formed* O_3 *was able to react appreciably with each of the target BVOC precursors. During transit through each respective* O_3*-alkene reaction mechanism and in the presence of water vapor, OH was also produced. As can be seen in Alfarra et al. (2013), with the exception of* β*-caryophyllene, each experiment was characterized by constant net* O_3 *production (and hence concomitant OH production), meaning there was* *always sufficient oxidant to allow the formation and reaction of second generation products; this can also be seen in the VOC temporal profiles presented in Alfarra et al., 2013. β-caryophyllene was observed to consume O_3 as it was produced and net O_3 production was observed only during the latter stages of the experiments; however during the net O_3 production stage, there was again sufficient oxidant to allow the formation and reaction of secondary products, this can be clearly seen in our companion work on β-caryophyllene (Jenkin et al., 2012).*

Furthermore, from measurement of the m/z 44 and 43 ions as a function of time in the AMS data, SOA chemical transformation was observed and fully reported, again in Alfarra et al., 2013.

However for further clarity, the reader has been directed to the work by Hamilton et al. Jenkin et al. and Alfarra et al. again, following the text describing particle phase results, in section 4.5, page 1672:

"For full details regarding the chemical evolution of the particle phase and the particle-phase specific experiments in general, see Hamilton *et al*. (2011), Jenkin *et al*., (2012) and Alfarra *et al.* (2012 and 2013)".

---

11) P. 1677, lines 25-27: Can the authors give a statistic that relates the similarity of the "model mass spectra" to the experimental data set from the chambers? If this model mass spectra generated from the MCM is similar enough to the experimental MS obtained, why not run the PCA on MCM generated model mass spectra for each system and see if the results are similar to the PCA for the gas and particle-phase data? Or at least, in Discussion section, when trying to relate ions characteristic of particle-phase oxidation products from grouped precursors, can a metric be used for testing similarity of the MCM model mass spectra expected and the weight of ions affecting the various B-caryophyllene and myrcene mappings for example?

Author Response: A comparison between the model and measured mass spectra is not *straightforward and would first require knowledge of all of the species present in the gas phase and the exact fragmentation patterns and ionisation routes for each of these species. Furthermore, one would need to understand all of the isobaric instances in the measured mass spectra; a difficult task when poorly characterised ions fragment and overlap. One would then have to (re-) construct mass spectra representing the gas phase matrices for each of the experiments employed, with no detection method specific alteration to the mass of each parent compound (e.g. protonation, adduct formation, fragmentation etc) or instrument biases (a simplified, first approximation attempt at this was made for α-pinene in Camredon et al., 2010). As we did with Wyche et al., 2009 and Rickard et al., 2010, we plan to work on a follow up/companion modelling paper to the current work, in which we perform/present more complex and in-depth work than is possible to include here, including the statistical analyses of modelled mass spectra as suggested by the referee. Unfortunately, to include all of this in one piece of work would make the product far too large and also divert from the original aims of the paper. The purpose of the modelling exercise performed here is simply to help us take steps towards understanding how useful these techniques are for application in a real world environment, and the discussion presented does just that. What the referee proposes here is an excellent idea, but a substantial and separate piece of work entirely. (Note, that the MCM scheme currently does not yet include mechanisms for myrcene or linalool, and is not directly coupled to partitioning code).*

---

12) P. 1678, lines 26-29: This argument would be strengthened if the authors did similar MCM analyses as in Fig. 10 for the other precursor types represented in Fig. 9 and provided metrics for the similarity of the model mass spectra with that observed in the chamber experiments.

Author Response: As already stated, the purpose of performing the MCM analysis was to help us move the technique towards a "real-world" environment and to give some indication as to whether the techniques successfully employed in chamber and mesocom *environments could be useful in the chemically more complex real world atmosphere. As pointed out by the referees, we are fully open regarding potential limitations in moving towards the real atmosphere, yet we believe at this stage of our work, the non-idealized simulations reported, provide a sufficiently chemically complex model with which to support the arguments made.*

We feel that performing the simulations using the α-pinene mechanism was the most appropriate of the four options available, i.e. α-pinene, limonene, isoprene and β--caryphyllene (note, as yet no mechanisms are available in the MCM for myrcene and linalool oxidation). α-pinene was chosen primarily because it is the most well characterised of the available options and moreover both limonene and β-caryophyllene are recent additions to the framework and therefore less well understood. Moreover, to perform the analyses proposed by the referee would be a substantial undertaking and extend the current work too greatly. It is the intention of the authors to work on a follow up/companion modelling paper (as we did with Wyche et al., 2009 and Rickard et al., 2010), where we perform/present more complex and in-depth work than is possible to include here, including statistical analyses on the modelled mass spectra.

As described in point 8), a comparison between the model and measured mass spectra would first require knowledge of all of the species present in the gas-phase and the exact fragmentation patterns and ionisation routes for each of these species. Furthermore, one would need to understand all of the isobaric instances in the measured mass spectra; a difficult task when poorly characterised ions fragment and overlap. However, this is not the current focus of this work, indeed one of the major points of the presented work is to employ the "mass spectral fingerprints" rather than speciated detail.

---

13) P. 1679, line 12: Authors claim that the gas-phase oxidation products of each structural type can be grouped "according to the controlling chemistry and the products formed." While it is evident that the products formed (as interpreted by the accompanying MS) informs this grouping, what chemical controls are derived from the

PCA analysis? I think of e.g. varying NO_x and O_3 levels as different chemical regimes/controls on the chemistry, but this point is not emphasized in the PCA mapping of the various experiments grouped by precursor. Where is this shown/taken from the figures? Or, please clarify what is meant here.

Author Response: The referee is indeed correct, under the current analysis we have not demonstrated this and the text has been amended accordingly. To this end we have included a small separate analysis of the toluene data, which was conducted under a reasonable range of VOC/NO_x conditions (low, medium, high NO_x). This new information has been included in a new section, Section 5.3 along with an additional Figure (10), shown here. Both the additional text and the new Figure 10 were provided following the second comment made by referee 1 (refer to Response to Referee 1).

---

14) P. 1679, lines 13-15: In line with comment above, this sentence is rather vague. No description thus far has really emphasized "decoding of mechanisms" by starting with an ensemble of summed mass spectra. I agree you can get to the precursor compounds based on tracer ions indicative of chemistry from a particular precursor, but what in the PCA analysis maps to particular mechanistic pathways? Do the authors envision additional mappings in the PCA plots within the precursor groups to show oxidation from one oxidant versus another, for example, which would be evident in unique tracer compounds from associated pathways?

Author Response: The referee is correct that this requires some clarification and we believe that the additional section regarding toluene/NO "mapping", as discussed above under point 14), goes some way towards achieving this.

In a sense this sentence was pointing to the questions of whether, if you have similar product ions, would the PCA more closely cluster the different mechanistic pathways? The referee is correct in the summation that we envisage additional mappings. The text has been clarified as follows to point to this:

"Indeed, a major potential strength of the data analysis methodology described here, could lie in the decoding of mechanisms into pathways (i.e. separation within a group on account of different underlying chemistry) and consequently linking chemical pathways to precursor compounds."

Furthermore, following similar comments made by referee 1, a small additional sub-section (5.2) was added on page 1676, after line 2, (see Reply to Referee 1, first comment):

"*5.2 Mapping within a class…*"

Technical Comments:

1) P. 1659, Experiment design: It is not clear in Table 1 which experiments were conducted in which chamber.

Author Response: *Table 1 and the accompanying caption have been altered to indicate which experiments were conducted in which chamber.*

---

2) P. 1661, lines 14-16: Since it is not clear which experiments were conducted in which chamber, as in the proceeding comment, please also make clear here which experiments actually had particle-phase data. Later inspection of Figures 7-8 warrants the reader questioning of why the mesocosm experiments and isoprene experiments are not included in the analysis of the particle-phase data.

Author Response: *As Table 1 and the accompanying caption have now been altered to indicate which experiments were conducted in which chamber (as per Referee 2, Technical Comment 1), the reader can now easily see which experiments were conducted in the MAC and therefore which had accompanying particle-phase data. No data was available for certain mesocosm experiments, owing to the low detection limits imposed by the experimental set up and for those experiments with particle-phase data, insufficient repeat experiments were available to use in the statistical analysis presented* *here.*

3) Figure 2 would benefit from caption description of the mass spectral ions populating the figure.

Author Response: Figure 2 caption has been altered accordingly to include a description of the major mass spectral ions populating the figure.

4) P. 1669, line 13: Change “been” to “be”

Author Response: The text has been altered accordingly.

5) For aid in interpretation of the ions mapped onto figures 2 and 7a, it may be helpful to include a table that lists all ions (LC-MS/MS, CIR-Tof-MS, and AMS), their chemical assignment, and precursor type, so the information is more readily available than filtering through the text for these ion assignments.

Author Response: The requested table has been produced (now Table 3) and inserted and the relevant main text and captions have been altered accordingly.

6) Figures 7b and 8b: It would be more helpful if the color scheme used for the cyclic monoterpenes was consistent with that used in Figure 5.

Author Response: We thank the referee for pointing this out, but we would prefer to leave the dendrogram colors as there are, as they represent the best combination for the clearest description of the data, i.e. the dendrogram colors were chosen specifically to optimize clarity of view within the figure.