We would like to thank the reviewer for his/her helpful comments which improved the manuscript considerably. Answers to comments are written in blue. Changes in the manuscript are marked with red.

Interactive comment on "Exploring sources of biogenic secondary organic aerosol compounds using chemical analysis and the FLEXPART model" by Johan Martinsson et al.

Anonymous Referee #2
Received and published: 16 March 2017

Reviewer comments: acp-2017-90, Exploring sources of biogenic secondary organic aerosol compounds using chemical analysis and the FLEXPART model

The authors quantify a number of organic compounds in PM2.5 samples collected on filters in rural Sweden, in particular several acids, di-acids, and organosulfates. By modeling (using FLEXPART) the landcover types that influenced each sample, the authors are able to use principle component analysis to connect landcover to composition and qualitatively determine the importance of different landcover types. This is, in general, a valuable and important goal, and a reasonable approach to doing so.

The conclusions of this work add to the body of work demonstrating the importance of coniferous forests to European aerosol loadings, and the work is generally scientifically rigorous with an honest assessment of capabilities and limitations. There are certain places in this work where the authors could dig deeper, discuss further implications, or further clarify and frame the proper interpretation of the data; these are discussed below and may make this analysis stronger, but these are generally only minor revisions.

General comments:
The crux of the scientific contribution of this work is the PCA, in particular the connection between landcover and composition, and the correlations between some of the straight-chain di-acids. However, the results of the PCA are not particularly surprising (biogenic products come from forests), while some substantial lingering questions that are not wholly addressed by the authors:

1) The main conclusions revolve around PC1, which includes a large majority of the measured components, as well as most of the explained variability. As noted by the authors, the compounds include both monoterpene and isoprene products, which are known to be dominant emissions from different tree species –including further components may split this out (or not), did the authors consider trying other PC solutions or parameters?

We performed several PCAs by varying the number of rotated factors from 2 to 6. We then judged the interpretability of the PCAs by attempts to associate logical and physical explanations to the extracted factors. In this dataset, the subjective best interpretation was observed by using VARIMAX rotation for 4 extracted factors. We have added some information regarding this issue in the "2.4.3 Principal Component Analysis (PCA)" section.

2) For PC2 (and PC4 to a lesser extent), no compounds really correlate with these cover types – was there lower loadings in general, or was PM2.5 just comprised of compounds not measured in this work?

We, as authors, do not really understand this comment and question. The PCA loadings are displayed in Table 7, and they vary depending on which PC you study.

We found detectable concentrations of all presented compounds. Hence, the case might be that the strongly contributing land-cover types in PC2 and PC4 do not contribute to the observed PM-species.

3) There are some biogenic compounds that don't really fall strongly in to any of the PCs (OS250, both NOS) – any thoughts about why that is or how those are different from all the others that co-vary? In the case of pinonic acid, for instance, previous work (Thompson et al., http://dx.doi.org/10.1080/02786826.2016.1254719) has shown it is mostly in the gas phase, so filter samples may be mostly gas-phase artifacts which may make it different than the other lower volatility products.

Formation of NOS depends on availability of precursors, including NOx, which could affect their variation compared to other tracers. OS250 is a product of alpha/beta-pinene and it is not clear why the correlation to other OS is low.

We agree that a major fraction of pinonic acid is found in the gas phase, though the partitioning is expected to be somewhat shifted at the lower temperatures in Sweden compared to the study of Thompson et al. in Alabama during summer, favoring partitioning to the particle phase in this work. Previous work of Kristensen et al. (http://dx.doi.org/10.1016/j.atmosenv.2015.10.046) showed that as much as 80% of pinonic acid collected with a high volume sampler could be due to gas phase adsorption, but if gaseous and particulate products are transported in the same air masses this should only affect the variation to a minor degree.

4) Was any attempt made to consider the age of the particle as it relates to landcover?

That is, a coniferous forest (or pixel) near the site will emit pinene, but not be turned in to pinic acid before the site, while a farther forest (or pixel) might. So all pixels are not created equal, and in many ways these factors are probably a driving force in variability. For instance, is it not possible for likely that the covariance of isoprene and monoterpene products that put them in the same PC is due to chemical processes, not differences in emission from landcover?

This is a very good idea. It is a bit out of the scope of this study but we have added some sentences regarding this approach in the outlook.

Still, monoterpenes and isoprene are, as far as the authors are aware, derived from different types of forests (more specifically, plants). Hence, any similarity in emission patterns may portrait emissions from mixed forests while differences may indicate emissions from very specific land-cover types.

5) Throughout the work, the authors classify the di-acids into two groups, anthropogenic and fatty acid, but it's not clear why that is and given their structures why they wouldn't just get binned. Unless it is in the introduction and I missed it. This is especially confusing given that the authors go on to demonstrate that they co-vary, or at least if they are two different groups (adipic vs. others) it is not in the breakdown pre-supposed.

We have added some information on the precursor sources to these acids in the results and discussion section.

Specific comments:
P. 2 L. 2 – Why start the paper with a comment about PM10 when the rest of the paper is about PM2.5?

This has been changed.

P. 2 L. 9 – “Four main categories: : :” This is an odd breakdown, they aren’t really symmetrical categories. Isoprene isn’t really a “category” it’s just one compound, and there is no clear split between “other reactive” and “less reactive”. So the categories are ‘individual compound’, ‘class of compounds’, ‘compounds of a certain poorly defined property’, ‘compounds of a different certain poorly defined property’. It should be classes, or sources, or properties, or some unifying ‘type’ of categories. I can deal with isoprene being treated as a ‘class’ if there is generally other symmetry, but the switch of ‘class’ to ‘property’ is asymmetric.

Very good comment. We have changed “categories” to “classes”. We have also re-phrased the sentence and removed the last two “categories”.

P. 3 L. 1 – “BVOCs constitute more than 50% of all atmospheric VOCs: : :” – If I’m not mistaken, that is low, generally BVOCs are more like 90%.

The numbers given are correct according to the cited references.

P. 3 L. 9 – Use “Influence” instead of “enhancement”. While the influence of sulfate is generally enhancing, the presence of OSs only points to influence, they do not necessarily indicate that BSOA mass would have been lower in the absence of anthropogenic influence, just that it would have been different composition.

This has been corrected.

P. 4 L. 15 – Eluent A is 0.1% acetic acid in what? Water?

Yes, in water. This has been clarified.

P.5. L. 14 – “A formal source apportionment would require a precise accounting of these factors, which is extremely complicated and is clearly out of the scope of this study.” This is a subjective sentence that could be re-worded to more precisely state the reasons for not providing more discussion or detail into the impact of the factors discussed in the previous sentence. Even if “out of the scope,” some discussion of these factors would greatly enhance the discussion and interpretation of these results, see General Comment 4. Change to something more like “A formal source apportionment that includes a precise accounting of these factors is out of the scope of this study which is focused on landcover types, but some discussion of these factors is included where relevant.”

Good comment. We have changed this in accordance to the reviewers comment.

P. 5. L. 29 – put a separation in 100000, either a space as used elsewhere, or change all of them to commas (my preference, as an American: : :) or periods

This has been corrected.

P. 6 L. 1 – Typo in “ocean”

This has been corrected.

P. 6 L. 14 – Change to “PCA was: : :”

This has been corrected.

P.6 L. 20 – Should be “noting” instead of “noticing”.

This has been corrected.

P.6 L. 21 – “most probably” is fairly informal, and “has” is the wrong tense. Can probably just combine this and the next sentence into one sentence.

This has been corrected.

P. 6 L. 26 – see General Comment 5, why are the di-acids split into different categories?

We have added some information on the precursor sources to these acids in the results and discussion section.

P. 7 L. 15 – Here, and in general, the discussion and analysis would be bolstered by testing for the effect of excluding these days with peaks. Does the correlation become more like that of the other acids? Does it fall into the same PC as the other acids? In other words, is the big difference of this acid just these two peaks? And if so, is there any indication in the back trajectories or composition of what might be causing these peaks?

We have removed the concentration peaks in adipic acid and re-analyzed the data and the PCA. The outcome of this re-analysis is stated in section “3.3 Connection between surface type and measured species”.

P. 7 L. 29 – The intuitive interpretation of these data is of course what the authors not should not be interpretation, that the land cover exposure is a measure of the contribution of the landcover to aerosol production. It would help for the authors to re-iterate what the proper interpretation is, since it is not wholly clear (note that in the methods section, the back trajectories are discussed in the subheading of “Source Apportionment”). This clarification may help shape the discussion somewhat.

We, as authors, are not sure what the intention of the reviewer’s comment is here.

In the mentioned line we are discussing the exposure contribution of the “Other” category. We are stating that this category has significant impact on the exposure and a deeper discussion on the “Other” category follows a few rows down: “Further, the category “other” is also increased during this particular period…..”

P. 8 L. 22 – Should be “Methods section”

This has been corrected.

P. 9 L. 6 – It would be interesting to see if adipic ended up in this factor if the 2 spikes were excluded (see comment below about PC4)

We have removed the concentration peaks in adipic acid and re-analyzed the data and the PCA. The outcome of this re-analysis is stated in section “3.3 Connection between surface type and measured species”.

P. 9 L. 12 – Interpreting 0.21 has a meaning contribution is probably somewhat overinterpreting. Perhaps this is a place where exploring other correlations or factors may be worth discussion.

We agree with the reviewer that drawing conclusions from a loading of 0.21 is dangerous and may lead to false interpretations. Hence, we have removed the concluding sentence stating that broad-leaved forest may contribute to carboxylic acid production.

We have added more discussion regarding the carboxylic acids further down in the same section.

P. 9 L. 16 – Again, how much of this is due to those two spikes? To speculate for a moment, is it not possible that the landcover types associated with PC4 just happen to be co-located with some strong point source of adipic acid, so it is all due to an unidentified covariance?

We have removed the concentration peaks in adipic acid and re-analyzed the data and the PCA. The outcome of this re-analysis is stated in section “3.3 Connection between surface type and measured species”.

P. 10 L. 9 – Remove “totally.” Again, a comprehensive implementation may be out scope, but some more discussion of these factors is warranted in the PC analysis, particularly results that are difficult to explain by landcover.

We have removed “totally”. Further, we have added more discussion in the 3.3 section. Especially, regarding the carboxylic acids and the effect of removing concentration peaks of adipic acids and the followed re-analysis.