Atmos. Chem. Phys. Discuss., 6, S1954–S1961, 2006 www.atmos-chem-phys-discuss.net/6/S1954/2006/ © Author(s) 2006. This work is licensed under a Creative Commons License.



ACPD

6, S1954–S1961, 2006

Interactive Comment

# Interactive comment on "Cluster analysis of the organic peaks in bulk mass spectra obtainedduring the 2002 New England Air Quality Study with an Aerodyne aerosol massspectrometer" by C. Marcolli et al.

# C. Marcolli et al.

Received and published: 20 July 2006

We thank the referee for his/her thoughtful comments, which we have addressed in detail below.

**General Comments** 

This paper applies a hierarchical clustering algorithm (HCA) to data collected using an Aerodyne Aerosol Mass Spectrometer (AMS) off the coast of New England during the summer of 2002. This is the first time such an approach has been taken using AMS data, and the subject matter is certainly appropriate for ACP. However, there are some



EGU

problems associated with the methods, and in general, I do not believe that sufficient evidence is presented to support the conclusions made. The paper is mostly well written, with an appropriate abstract.

Response: We gave good evidence in the paper that hierarchical cluster analysis was able to extract different categories of spectra from a large dataset. The conclusion that some of the categories represent organic matter from biogenic sources was based on both spectral and gas phase correlations. At the moment a better validation of these conclusions is not possible, since AMS reference spectra from further sources would be needed to assess more closely the specificity of the AMS organic mass spectral patterns. We rather view this paper as a starting point for further analysis than as an end point where final conclusions can be drawn.

Specific Comments In the abstract, the wording of 'as much as 5 ug/m3 organic aerosol mass - 17% of the total organic mass' makes it sound as if the total organic aerosol mass was always approximately 30 ug/m3, which is clearly incorrect (line 17 in the abstract).

Response: We have rephrased this sentence to: Taken together, the second through the fifth most common categories represent on average 17% of the total organic mass that stems likely from biogenic sources during the ship's cruise.

The paper should, without a doubt, provide support to the use of a single-particle data analysis technique for an instrument that yields data that represents ensemble averages of particles in the atmosphere. By using an ensemble instrument, different types of particles may be masked by one another through averaging - and then again by clustering.

Response: We discuss the caveats of the cluster analysis now in section 3.8 of the revised version.

There is a problem with equation (2) and its discussion. The dot product is one only if

### ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

the vectors are normalized. In addition, equation (2) shows the cross product, not the dot product of the vectors.

Response: We normalized the spectra to unity length and applied the dot product to the normalized spectra. We make this clearer in the revised manuscript. We also replaced the cross by the raised point to indicate the dot product in equation (2).

Section 3.1 (lines 20+ on page 4608). More detail is needed on how nitrate peaks are handled. I do not believe it is appropriate to include all of the peak from m/z 30 as part of an organic study. Some of this no doubt comes from inorganic nitrate - even if it is not ammonium nitrate (though my gut tells me even a tiny bit of that will form even in sulfate rich/ammonia poor conditions). The study is conducted in the polluted marine boundary layer where high concentrations of nitric acid will react with sea salt particles in the size domain in which the AMS can collect particles, yielding sodium nitrate. Is this nitrate potentially due to sodium nitrate? Nitric acid may also sorb to particles - yielding 'pure' nitric acid in the aerosol phase. Could this not also be a source of signal at m/z 30? The authors, at least in this manuscript (one other in preparation is cited) have not provided enough evidence to support inclusion of the peak at m/z 30 as 'organic.'

Response: Inorganic nitrate was measured using ion chromatography for samples analyzed online with PILS (only submicron particles) and offline with impactors (both sub and supermicron particles) and the results of these measurements are reported in Quinn and Bates (2003), Brown et al. (2004), and Bates et al. (2005). Essentially very little inorganic nitrate was measured in submicron particles, with some nitrate detected in the supermicron particles in the form of sodium nitrate (impactor data). Comparisons of the AMS nitrate with PILS nitrate during times when submicron sodium and nitrate were detected with PILS, the AMS nitrate was low. Hence, the AMS was probably not detecting significant amounts of sodium nitrate. This is expected since the AMS vaporizer temperature was 550 C which is not high enough to volatilize sodium nitrate. On average for the entire study, the AMS nitrate and PILS nitrate were not well correlated,

#### ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

**Printer-friendly Version** 

Interactive Discussion

implying that the two methods were not detecting the same species. The study average ammonium to sulfate mole ratio is  $1.5 \pm 0.2$  from the data shown in Figure 2 and is  $1.7 \pm 0.3$  from the submicron impactor data (Bates et al., 2005). Since this ratio is less than 2, the submicron particles were acidic and should contain very little ammonium nitrate. Along with the high 30/46 ratio, this evidence supports the possibility that the AMS nitrate signal could be due to some other species and is the reason we included m/z 30 in the organic cluster analysis. This is now explained in the revised manuscript.

On page 4610, lines 4 and 5 - why are these m/z ratios chosen? What do these values represent?

Response: These m/z ratios represent the most intense organic m/z peaks in the spectra. We state this now in the revised manuscript and also have a table of dot products to compare the reference spectra with the categories.

In the following paragraph, it is not clear to me why removing the peak at m/z 44 is necessary, especially since it is a dominant contributor to signal.

Response: We repeated the correlation without m/z 44 for two reasons. i) because we are sampling particles continuously the measured spectra represent an ensemble of particles rather than particles from one source with one history. Many of the minor categories spectra might therefore be influenced by a background of the dominant category 1. The influence of this category will be largest in the dominant mass, which is at m/z 44. ii) Particles from one source that are sampled at different times might differ in their degree of oxidation which influences the intensity at m/z 44. To give the variability due to different degrees of oxidation less weight, we repeated the correlation without m/z 44.

Also in this paragraph, no where is it specified what type of oxidation system was used in the alpha-pinene study of Alfarra et al. Is the use of a different oxidant the reason for the differences in the spectra shown for that study and that of Bahreini et al.?

# ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Response: The oxidants for the alpha-pinene chamber spectra were indeed different (hydroxyl radical and ozone), and this is now clarified in the manuscript.

Page 4611, line 1+ - the authors state that the average delta patterns for the top 5 categories are similar to each other and more similar to those from biogenic precursors than those from anthropogenic precursors. Two comments: One, for non mass-spec specialists, a little bit more detail on what a delta pattern represents would be helpful. Two, I suggest changing the word anthropogenic to aromatic - to my knowledge, the delta patterns of SOA from anthropogenic species other than aromatics are not known.

Response: We have extended the explanation of the delta patterns in the revised manuscript by stating: "In addition to subtle differences in the mass spectra of the low m/z peaks from various precursors, the high m/z fragments may also exhibit distinct mass spectral patterns. These can be analyzed by the ion series or delta analysis technique which uses the delta value, delta = m/z-14n+1 (where n is the "nominal" number of CH2 groups), as an indication of the functional groups in the molecule. Bahreini et al. (2005) previously used this technique to examine the delta patterns of oxidation products from many types of precursors (terpenoid, cycloalkene, and m-xylene). In that study, the larger carbon fragments (C5-C6 and C7-C15) of secondary organic aerosols from different biogenic (terpenoid) precursors showed negative delta values which was in sharp contrast to the positive ones observed for cycloalkene and m-xylene secondary organic aerosols representing anthropogenic precursors." Since the delta pattern analysis by Bahreini et al. (2005) is based on terpenoid, cycloalkene, and aromatic precursors we keep the nomenclature (biogenic vs. anthropogenic) introduced by these authors.

In general, I find the discussion of relatively minor categories not particularly meaningful, especially since they were observed during low mass periods when relatively fewer particles were being sampled. I would suggest focusing more on the most important categories.

## **ACPD**

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

**Printer-friendly Version** 

Interactive Discussion

Response: We agree with the reviewer that the major categories are the most important ones. We list some of the minor categories in Table 3 and mention them shortly in one paragraph. We still would like to show them in the revised manuscript to point out the diversity of the categories. Moreover, in future work, spectra might be measured that resemble the ones of the minor categories and correlations would become possible.

Page 4614, first paragraph of section 3.4. Why is the peak for category 1 much broader than the organic aerosol distribution with the peak in the afternoon if both represent SOA of a given type? Why does the category 1 distribution also feature a peak at night?

Response: The broader peak in category 1 than in the organic mass can be explained if one assumes further processing of the organics in the aerosol: organic species become oxidized in the gas phase and the products partition to the particulate phase, which explains the increase in organic mass. Organics in the particulate phase are further oxidized leading to an increase in category 1. Alternatively, the composition of particles could be different when the boundary layer changes. We discuss this further in the revised manuscript. We considered the peak of category 1 at night as too weak to be significant and did not look for an explanation for it.

Page 4615, the authors discuss correlation with gas phase species of some of their clustered categories. However, they use rather vague descriptive words (i.e., well, to a lesser degree). Please provide the numbers as it is difficult to really look at correlations on a time series plot on which multiple variables are plotted (Figure 11).

Response: Although a scatter plot might be useful to quantify the correlation, the sampling times are quite different - 5 minutes every half hour for the gas phase calculation and hourly averages for the AMS data. Furthermore, changes in relative occurrence are not always directly related to gas phase concentrations for many reasons. We qualify the agreement as quite good given all the assumptions and uncertainties connected with the biogenic mass derived from the categories. A scatter plot and r2 statistics

## ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

**Printer-friendly Version** 

Interactive Discussion

would rather direct the attention away from the specific time periods when the agreement was good or bad and imply the possibility of a full agreement which can not be achieved given all the assumptions that had to be made.

A key issue with the paper is stated on this page (lines 14-15): 'implying that during this time period also the particulate organic mass represented by category 1 could have biogenic sources.' Elsewhere, they discuss the conversion of categories other than 1 to category 1. Therefore, the authors can not say what fraction of their SOA is biogenic vs. anthropogenic. In addition, the converse may be true - that there may be some unidentified anthropogenic that contributes to categories 2-5 (especially given the ensemble type of measurement). Therefore, it is unknown what the bounds on their estimate are. In addition, comparing to deGouw et al. certainly gives credence to their ideas. However, it must be even more strongly underscored that the method of deGouw completely ignores monoterpenes as an SOA source.

Response: In section 3.6 we assume indeed that categories 2-5 are entirely biogenic. We agree with the reviewer that this introduces a bias in the direction of too high biogenic. We think that this bias is overcompensated by the assumption that all of category 1 is anthropogenic. In the revised manuscript we state the caveats to the analysis in section 3.6 more completely by adding: On the other hand, a part of the mass especially in category 2 but also in categories 3-5 might be anthropogenic. The assumption that the mass in categories 2-5 is totally biogenic might therefore lead to an overestimate of the biogenic organic mass at times when these categories are abundant.

In section 3.6, method 1 is not appropriate.

Response: Method 1 has been omitted from the text.

Section 3.8 should be removed.

Response: In the revised manuscript, we have added to this section a better discussion of the limitations and caveats of hierarchical cluster analysis applied to an AMS dataset

#### ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

**Printer-friendly Version** 

Interactive Discussion

as asked for by referees 2-4. We regard a discussion of the method as necessary.

Comments on Figures:

Why are two different plates required for Figure 1?

Response: The top panel shows the isoprene and the bottom panel the monoterpene emissions. The resolution for the two species is different. It is not possible to merge these two emission fields into one panel.

Figure 2 (and elsewhere) - high-resolution wind direction measurements may not necessarily represent air mass history. I suggest use of some back trajectories.

Response: For this region in the summer time, there is a link between wind direction, synoptic meteorology, and particulate matter. We have added a discussion of this to the methods section where Figure 2 is introduced to justify using only the wind direction in our analysis.

The text in Figures 5 and 7 is small and difficult to read.

Response: In the revised manuscript, we have enlarged the text in Figure 7. It is difficult to enlarge the text in Figure 5. We think that the whole Figure should be enlarged in the final version.

Interactive comment on Atmos. Chem. Phys. Discuss., 6, 4601, 2006.

ACPD

6, S1954–S1961, 2006

Interactive Comment

Full Screen / Esc

**Printer-friendly Version** 

Interactive Discussion