**Atmospheric
Measurement
Techniques**
Discussions

Interactive comment on "Quantitative single particle analysis with the Aerodyne aerosol mass spectrometer: development of a new classification algorithm and its application to field data" *by* F. Freutel et al.

Anonymous Referee #1

Received and published: 6 August 2013

This paper presents a new algorithm for the processing of single particle data obtained using an Aerodyne AMS equipped by a light scattering module. This work builds on previous techniques used for the interpretation of data from this instrument and could be very useful in the future for the application to atmospheric science. The approach uses empirical classification rules based on existing mass spectral profiles. Compared to a more general-purpose clustering method (such as k-means), this has a heavily reliance on a priori assumptions on the expected compositions of ambient particles. However, because of the low signal quality of the individual particle data, it can be ex-

pected that this will produce much more decisive results and the technique is demonstrated to be open to unexpected particle types. This paper is relevant and suitable for AMT and is largely well written, barring a few issues. I recommend publication, subject to the following comments.

General: Linear correlation (Pearson's r) is not the only metric available for comparing mass spectra and it can be argued that the normalised dot product (also known as uncentred r) is more suitable. While the two metrics may produce qualitatively similar results, a particular issue with r surfaces when considering the applicability of the flowchart on figure 1. Because the exact quantitative value of r is dependent on the number of zeros (or near zeros) in the mass spectrum in addition to the magnitudes of the signal-containing peaks, the thresholds quoted in figure 1 will not necessarily remain valid if one were to use mass spectra of different sizes to the ones used here (which isn't currently specified). If r2 is to continue to be used, the authors need to be very specific about the size of the mass spectra they have used in this instance and comment on how the thresholds in figure 1 might change with different sizes. Alternatively, the thresholds could be changed to the corresponding normalised dot product values.

Page 5658, line 14: As it stands, the comment about AP240 thresholding is unlikely to make any sense to someone not already familiar with the operation of the TOF-AMS. This needs expanding or referencing.

Page 5660: Many studies of particulate engine emissions (including the cited Canagaratna et al.) focus on diesel rather than gasoline emissions, as they are arguably the more atmospherically significant. While I accept that the choice of particle generator used here will depend on availability, the authors should discuss the anticipated differences between gasoline and diesel emissions and whether this would be expected to affect the analysis presented here.

Page 5661, line 7: Is the increased bounce of larger particles due to the higher fraction

of dust and sea salt?

Page 5662, line 24: The use of glucose as a proxy for biomass and biomass burning OA needs further validation. To wit, glucose has different H:C and O:C ratios compared to species such as levoglucosan. A comparison of a glucose mass spectrum with those of biomass and biomass burning obtained from the AMS library would address this.

Page 5665, line 23: The discussion of uncertainty here does not make a distinction between accuracy and precision (i.e. systematic vs. random errors), which I think is warranted. While all effects should be considered for the reporting of a single datum, when obtaining certain comparative statistics, only precision needs to be considered, which as far as I can tell, consists only of the 'MS-retrieval' component. The authors should make the distinction between the two types of error clearer.

Page 5667, line 1: The caveat should be added that oleic acid only appears hydrocarbon-like under unit mass resolution (UMR) analysis. Subjected to higher resolution analysis, there is a clear distinction in that it contains oxygen-containing peaks such as C2H3O+.

Page 5669, line 6: It is not very informative to state that the Gaussian fits are not perfect without showing an example, but moreover, there is nothing to say that the data should conform to a Gaussian distribution anyway (the distribution for a single component is more likely to follow a Poisson distribution), so it begs the question why it was even used in the first place. If Gaussian fitting is found to be problematic (as seems to be the case), then the authors should really consider using a different fitting function or report a more robust statistic, such as the median.

Page 5569, line 28: How is the standard deviation obtained from the Gaussian fits? Is this the fitting uncertainty or the width of the distribution?

Page 5674, line 24 (also later on page 5679): The statement about unambiguity isn't strictly correct because the algorithm itself contains a lot of operator-lead decisions in

C1934

its design, so while the results are more decisive, it still not truly objective. If the authors want to claim it to be completely unambiguous, they would need to demonstrate its effectiveness against an independent data source.

Page 5676, line 16 (also later on page 5679): I'm not sure I follow the argument about the boundary layer affecting nitrate concentration or mixing state. While a suppressed boundary layer can increase the number concentration of continuously-emitted primary particulates through reduced dilution, this will not apply to pre-existing particles or secondary species like nitrate. If there is an increase in nitrate-containing particles at night, this will be because there are a larger number of particles for the nitrate to condense onto, but this will occur independently of the mass concentration of particulate nitrate. The phenomenon of particles growing into the detection regime of the instrument sounds plausible, however.

Page 5678, line 3: Remove comma after 'both'.

Figures 4 and 5: Why are these in black and white? They would be much clearer in colour.

Figure 6: Unless they would cause any major problems, I would have like to have seen the error bars included for sulphate and organics.