Atmospheric Measurement Techniques Discussions

Interactive comment on "Cluster analysis of WIBS single particle bioaerosol data" *by* N. H. Robinson et al.

Anonymous Referee #1

Received and published: 29 November 2012

A technique using single particle measurements to classify super micrometer aerosol particles as biological (bacteria, fungal/plant spores) or non-biological is presented. The technique is a cluster analysis of single particle measurements of size, asymmetry factor, and UV stimulated fluorescence at three wavelengths in the blue. The technique is tested in the laboratory on fluorescent and non-fluorescent polystyrene latex (PSL) particles. The best technique for separation of time series data into particles classes is then applied to ambient aerosol measurements in a forest in summer 2011. These types of measurements and analysis are new and so still require a lot of development work. This paper makes a useful contribution in that regard and should be published after the authors consider the following comments/confusions/corrections.

C3103

While the writing is primarily clear there are places where it is a bit confusing. Some of that may be due to my own limitations. This paper is my introduction to cluster analysis and these types of measurements, so please forgive any naivety in the following comments/questions.

6398 and Table 2. Where are these data from? Were these measured or made up? The text and table merely state here are the data input to the cluster analysis.

6399.1: Why THE 6 major clusters retained? This is a big reduction from 13 to 6 clusters in one sentence with little justification? Why were 13 clusters even chosen. Fig. 4 shows a significant drop in R2 and rise in RMS when 6 clusters are reached.

Table 3. Now, where do these values come from, measurements? If so why are they different than Table 2? The origin of cluster C at a size not in the original data is also not clear. There is something I do not understand which separates the origins of Tables 2 and 3.

Table 3. Are all the PSL used spherical? Is so what is the origin of the large differences in asymmetry factor?

6399.11–18: It is hard to understand how two clusters (A and B) are defined as separated by 0.03 μm in diameter and both fluorescent. Is it really believed that the cluster analysis could make such a fine separation? In fact it can not as seen in Fig. 5. But Fig. 5 is a bit misleading. It only lists the diameter per cluster, when the real separation was probably based on asymmetry factor for this difference, but, again, why should these be different? See the question above about Table 3.

6399.19: Does not the "population normalised distance simple attribution" approach also have a problem with an inability to separate clusters C and D?

6400.17: The justification for the 4 cluster solution is not obvious. The RMS does not significantly rise until cluster 3 is reached. R2 also drops more steeply then as well.

6400.20-21: Here is a good explanation of how the 10 clusters were reduced to 6

based on sample size. Could this have been applied in going from 13 to 6 clusters in the example with PSL?

6401.10: Somewhere about now the authors should reference Tables 4 and 5. In fact these tables are never called out in the text.

6401.15: I do not understand how two clusters in the 6 cluster solution would be agglomerated in a 9 cluster solution. It seems the resolution would only increase with cluster size. I also am not following how C4 and D4 are agglomerated in the 6 cluster solution when they are distinctly separated in Table 5 which I thought was the 6 cluster solution?

6402.10: Do you mean E4 instead of E3?

6402.12: What does the following phrase mean, "...if FL2 280 is typical of grass smut ..."? Fl2_280 is a type of fluorescent measurement, which could be low or high, depending on the particle.

Fig. 8: The caption indicates that rainfall is displayed at the bottom of the figure, but that is not the case. The bottom panel is fungal spores.

───────────────────────────

C3105