**Atmospheric
Chemistry
and Physics**
Discussions

Interactive comment on "Assessment and application of clustering techniques to atmospheric particle number size distribution for the purpose of source apportionment" *by* F. Salimi et al.

Anonymous Referee #1

Received and published: 8 July 2014

General comments:

This manuscript provides valuable information about different clustering methodologies that can be applied to particle number size distributions (PNSD). It analyses the performances of 4 different clustering techniques using the Dunn index and the Silhouette width. The k-means clustering technique is chosen and thus applied to the data set, resulting in 5 clusters. They are further classified in 3 categories: regional background particles, photochemically induced nucleated particles and vehicle generated particles.

The paper is well written and structured and the references to other studies on similar subject are accurate.

Specific comments:

There are two issues that are not touched in the paper, and in my opinion are relevant. First, it would be interesting to investigate the relationships between clusters and air masses during the sampling period, as well as wind speed and wind direction. Although this is a complex study including 25 sites, an estimation of the influence of these parameters would be appreciated. Second, in Figure 4 a very high density of local peaks is observed for all clusters in the lower sizes. What is the error associated to the method used (GAM) for these sizes, and in general?

Page 15270 lines 5-8: a 20 nm shift in the PNSD from Cluster 5 to Cluster 4 should not be directly attributed to the influence of biomass burning particles (100-200 nm in size). A log-Normal fitting of these two clusters PNSD should be performed in order to confirm this hypothesis. Ideally a 100-200 nm mode should be found in Cluster 4 that would be attributed to biomass burning aerosols.

Page 15270 from line 19 until the end of the manuscript: this should be a new Conclusions section.

Technical corrections:

Page 15264 line 5: what does FAQ stand for?

Page 15266 lines 20-21: should probably be Figs. 3, 4 and 5 instead of 1 and 2

Page 15267: references to Table 1 in the description of Cluster 1 and 2 are misleading

Page 15268 line 19: "less" should be substituted by "lower"

Page 15270 line 5: "higher" should be substituted by "larger"

Page 15270 line 10: "occurrence" should be substituted by "prevalence", as traffic

emissions also occur during the biomass burning period, although the PNSD might be dominated by biomass burning particles

Page 15270 line 23: "Clusters 3 and 5" should be "Clusters 4 and 5"

Page 15279: Caption of Figure 3: abbreviation of solar radiation (SR) should be included

―――――――――――――――

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 15257, 2014.

C4567