

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2021-743-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2021-743

Anonymous Referee #2

Referee comment on "Understanding aerosol microphysical properties from 10 years of data collected at Cabo Verde based on an unsupervised machine learning classification" by Xianda Gong et al., Atmos. Chem. Phys. Discuss.,
<https://doi.org/10.5194/acp-2021-743-RC2>, 2022

Review of "An unsupervised machine-learning-based classification of aerosol microphysical properties over 10 years at Carbo Verde"

The manuscript investigates the microphysical properties of aerosols based on a large data set at a remote site. They have used an unsupervised algorithm to classify the properties and investigated further, based on the corresponding air mass history also. The study presents a valuable data set for a long duration and follows a novel technique. However, the organization/focus of the manuscript is confusing with respect to the title, along with some other concerns. The results and discussion section could have been made more comprehensive and precise, say, the LAC results are disturbing the continuity between the particle NSD and CCN discussion unnecessarily. Sections 4 and 5 could have merged to form the summary and conclusions. One major concern is regarding the estimation of the effective hygroscopicity parameter and their further interpretations. The paper is worth publishing in the journal of Atmospheric Chemistry and Physics after considering the following aspects.

General Comments

- The effective hygroscopicity parameter (Petters and Kreidenweis, 2007) represent the hygroscopicity for that dry diameter. When the critical diameter (obtained from the back-integration of the NSD) is considered as the dry diameter, the corresponding K value should indicate the minimum hygroscopicity of the aerosol system at that supersaturation, since all the particles above that critical diameter should activate as CCN at lower SS itself. So how is the claim in Line 1 on Page 8 valid? Also, how well does the estimated K represent a multi-modal aerosol system, say having a distinct nucleation and accumulation mode as mentioned in the study for the moderate dust periods? Based on these discussions, what is the relevance of the claim in L9-10 in P17? What is the 'overall average K value' (L2, P9), an average of the K values for all the

supersaturations? If so, how it can represent the overall hygroscopicity of the aerosol system?

- There is confusion with the data availability. Each parameter seems to have different periods of availability. It is mentioned in the Introduction (not even in the Experiment and Methods section), and so hard to follow during the Results and Discussion. How much period does Fig. 1 represent? It will be better if the measured parameters along with their observation period are presented as a table.
- The absorption coefficient is corrected using a theoretically (Mie) derived scattering coefficient assuming a 'less absorbing' marine aerosol system. However, the same study highlights the seasonal presence of dust aerosols. In that case, how relevant is the scattering correction applied to the reported absorption coefficient values?
- The introduction needs a thorough revision. The authors should clearly specify the objectives and relevance of this paper systematically. Why the unsupervised MI is preferred in this study as mentioned in L3, P3? The data strength and location details can be moved to the later (Experiment and Methods) section.
- Another concern is the lack of appropriate references which might have enriched the discussions more. A few examples are;

Section 2.4: studies like Furutani et al., 2008; Jayachandran et al., 2017; 2021, etc has followed this approach at other parts of the globe

Studies like Nair et al., (2020) have investigated the CCN characteristics during the mixing of distinct air masses based on the clustering of aerosol NSD, which are not cited or discussed.

- The LAC data is mentioned and a monthly mean picture is shown. But, no more discussions on that! Any reasons?

Specific comments

L13, P2: the 'physics, chemistry, and biology' usage seems too qualitative!

L5-6, P3: What is the relevance of this statement?

Figure 2 is specified as the monthly mean, but there is no such information about Figure 1. Is it hourly mean?

Fig. 7: What is the significance of linear scaling apart from the logarithmic one? Not clear from the text.

L7-11, P13: Confusing. It is obvious that the nucleation mode particles contribute less to the volume as the mass distributions. As seen in the figure, the nucleation mode aerosol system will have a large total aerosol concentration. But it is mentioned that the NPF happens in marine air mass with low particle concentrations. Please justify this statement.

L28, P15: during 'this' previous study??

L7-9, P16: Not clear.