

Geosci. Model Dev. Discuss., author comment AC1
<https://doi.org/10.5194/gmd-2021-191-AC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Jingmin Li et al.

Author comment on "An aerosol classification scheme for global simulations using the K-means machine learning method" by Jingmin Li et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-191-AC1>, 2021

We thank the reviewer for his/her remarks on our study. We would like to clarify one misleading description in the manuscript and address the key issue on data standardization in this short comment.

The referee raises a correct point that the data analyzed in our study do not follow a Gaussian distribution in many cases. Our statements at lines 210-211 ("...assuming the deviation of the data from their respective mean to follow a Gaussian distribution...") was indeed misleading and we will correct it in the revised manuscript.

We would like to point out, however, that the purpose of this standardization in our study is to weight the different input quantities equally with respect to each other, before applying the K-means algorithm. It is not necessary to assume that the data follows a normal distribution to apply such standardization. The standardization procedure can be applied to any data, since it does not change the underlying distribution of the data.

We agree that the choice of the variance applied for standardization could potentially have an effect on the clustering (e.g, due skewed distributions or simply due the relatively large number of sample points potentially causing complex multimodal kinds of distributions). We will try to investigate the effects of alternative scaling methods on our results and include a more in-depth analysis in the revised version of the manuscript.

A more detailed reply to the reviewer's comments will be provided during the final response phase.