Journal: Earth System Science Data

Title: A first global height-resolved cloud condensation nuclei data set derived from spaceborne lidar measurements

Manuscript Number: essd-2023-91

General Comments:

This paper releases a long-time series (2006-2021) multilayer dataset of global cloud condensation nuclei concentration (CCN) with the spatial (2°×5°) and temporal (1 month) resolution based on CALIPSO spaceborne lidar. Moreover, the CCN of different aerosol types is also provided based on the developed method by authors. Further, the simple comparison is provided with the results from model. This dataset is important, has large potential to enhance the investigation of aerosol-cloud interaction. The expression and structure of paper is not enough clear now, and should be improved, especially for data process. And the vertical information of estimated CCN should be highlighted, which still missing in this paper. Therefore, I suggest this manuscript can be considered for publication after revision.

Major Comments:

1. Introduction. There are some previous works about global retrieval of cloud particle concentration number based on CALIPSO. Author should mention this information. For example, Zhang et al. 2019 (doi.org/10.1364/OE.27.034126) and Zang et al. 2021 (doi.org/10.1364/OE.427022).
2. Title. The retrieval of CCN based on CALIPSO have been tried by previous studies. Please remove “first”. It's controversial.
3. Line 66. Please shortly explain how CALIPSO classify the aerosol types to help others understand these aerosol categories clearly.
4. Line 70. The CALIPSO track was change after September 2018. This information should be mentioned in paper. This may be result in some error for the retrieved CCN after September 2018.
5. The length of paragraphs varies so much. For example, Paragraph 1 of section 2.3 is only 3 lines, but following paragraphs are beyond 10 lines. The too long

paragraph possibly let reader miss important information. Please adjust all paragraphs with the proper length throughout this paper.

6. Line 107. What about the upper radius limit of the integration? As many studies already find the CCN and GCCN have even opposite influence on cloud, it's better to indicate the limitation of retrieval for particle size.

7. Line 113. I suggest the writer to add a flow chart containing the steps from Pre-processing to Output data records. The current methodology section is too long and not structured well. This flow chart can be used to clearly introduce the methodology section.

8. Line 130. Authors mention that "we filter out the profiles or columns with cloudy pixels". Do you mean that all profiles containing cloud are removed? Or just remove the vertical bins containing cloud? Additionally, how about the aerosol occurred above cloud?

9. Line 130. CALIPSO signal is gradually attenuated when detecting aerosol and cloud. For the profiles containing multiple aerosol layers, how to consider the effect of attenuation to CCN retrieval, especially for the lower aerosol layer.

10. Line 130. It seems the authors reject all profiles with cloudy pixels. However, the CCN under the cloud is of most interest to scientists. LiDAR can penetrate through some clouds (not too thick) and thus able to obtain aerosol information under the cloud. How about these cloud-penetrated profiles? Do you try it?

11. A main concern (Line 170). The signal-to-noise of CALIPSO have much difference during daytime and nighttime so that aerosol have been retrieved during daytime and nighttime separately (Mao et al. 2022, doi.org/10.5194/acp-22-10589-2022). Do you retrieve the CCN during daytime and nighttime separately? Or consider the uncertainty from this difference.

12. Line 173. How about the number of profiles at the 2°×5°grid for 1 month globally? Such as total profile, cloud-free profile and available aerosol profile. This information (figure) is very useful for reader.

13. Line 200. Does the algorithm have the ability to invert the CCN above 8 km? Can this algorithm be used in smaller time intervals? For example, instantaneous, daily average?

14. Line 208. Does the dataset include uncertainty indicators? At least standard deviation for monthly average data. It is important for others to use.

15. Line 237. Is there a similar pattern to the global distribution of AOD? which can prove the reliability of the CCN product by another way
16. Another main concern. Authors said the retrieved 3D information of CCN based on CALIPSO. This is one of main advantage of the product comparing the resulted CCN based on MODIS. But I do not see that the vertical CCN is highlighted in this paper. It is strange. It is necessary to show the global CCN at different altitude. The profile of CCN also should be shown over several typical regions.

Minor Comments:

1. Line 12. It should be revised from "warm clouds" to "clouds".
2. Line 21 and 147. Please add references for these comments.
3. Line 64. Please explain the "(NASA/LARC/SD/ASDC, 2018) ". The reader may be not known CALIPSO well.
4. Line 120. The paragraph is not structured well. Please number all criterions about quality control to let reader can understand easily
5. Line 152. Please add more detailed descriptions about "kappa parameterization".
6. Line 218. The retrieval of instantaneous faint aerosol (or undetected aerosol) has been explored by (Mao et al. 2022, doi.org/10.5194/acp-22-10589-2022) .
7. Please add longitude and latitude labels for all figures.
8. The caption of all figures is not detailed and clear. Such as time range? altitude?
9. Figure 2-4. The color bar should be improved. I suggest the non-value grid is shown as white color, and low-value grid is shown as deep blue color. Because I can't distinguish these white grids are related to low-value or non-value.
10. I do not understand why you copy all figures at supplement information again.
11. Table 1. The comment of extinction QC is not correct
12. Table 2. Adding the information about profile number at each grid into product is better. Such as the number of total profiles, cloud-free profiles and available aerosol profiles. A profile counting available aerosol bins at each grid is also helpful for people who want use this product.