We are thankful to the reviewers for the time, effort, and expertise they have dedicated to reading and reviewing our manuscript. We found their insightful feedback and constructive comments to be helpful in shaping and enhancing the quality of our work. Our point-by-point replies to the comments are provided below. Referee comments are given in black, and our replies are given in blue.

**Reviewer 1**

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 the 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.

Reply: We express our gratitude to the reviewer for dedicating their time to thoroughly read and analyze our work, and for offering valuable comments that have contributed to improving the overall quality and readability of our manuscript. The specific responses to the comments can be found in the subsequent section.

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).

Reply: Thank you for bringing the related articles to our attention. We found them to best suit the discussion in the conclusion section and therefore have added the following to the revised manuscript:

"Studies have demonstrated the effectiveness of CALIOP in even retrieving height-resolved cloud microphysical properties (Zhang et al., 2019; Zang et al., 2021), which can be directly coupled with corresponding CCN estimates to investigate ACIs."

2. Title. The retrieval of CCN based on CALIPSO have been tried by previous studies. Please remove “first”. It's controversial.

Reply: Thank you for your suggestion. We agree that there are other CCN retrieval algorithms that exist for application to lidar measurements, either in their development phase or validation phase. However, at the time of submitting the manuscript (and even now), we did not find any global CCN data set based on CALIPSO. We would therefore like to keep the title unmodified.
3. Line 66. Please shortly explain how CALIPSO classify the aerosol types to help others understand these aerosol categories clearly.

Reply: Thank you for your suggestion. We have added the following text to the updated manuscript in Section 2.1.

"These aerosol types are classified based on the type of underlying surface, aerosol optical properties such as the column-integrated attenuated backscatter coefficient and the initial estimated particle depolarization ratio at 532 nm, and the height of the retrieval (Kim et al., 2018)."

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.

Reply: Thank you for your suggestion. We are aware of the change in CALIPSO’s orbit from A-train to C-train following the orbit change of CloudSat. However, there is currently no known issue in the CALIPSO retrieval due to this transition. Please visit CALIPSO’s official page on the instrument status for more details, accessible at "https://www.calipso.larc.nasa.gov/tools/instrument_status/inst_status-2018.php" (last accessed on 24 May 2023).

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.

Reply: Thank you for your suggestion. We understand your point here. However, we design the paragraphs so that each of them conveys a single or very related concept. This results in paragraphs that are sometimes too long or too short. For example, in this case, the first paragraph presents a general overview of the section, so that the readers have an idea of what to expect and where to find further details. The second paragraph discusses the CCN-retrieval algorithm briefly, which has been described comprehensively in our previous articles (Choudhury and Tesche, 2022; Choudhury et al., 2022). Therefore, we have decided to keep the length of the paragraphs unchanged.

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.

Reply: Thank you for pointing it out. The upper integration limit is 15 µm, which also coincides with that of the AERONET size distribution. We have included this information in the updated manuscript as

"In accordance with the AERONET size distributions, the upper radius limit for integrating the size distributions is set at 15 µm."
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.

Reply: Thank you for your suggestion. We understand your point here. However, such a flowchart is already given in Choudhury and Tesche (2022) and Aravindhavel et al. (2023), highlighting the key steps described in the pre-processing and CCN-retrieval algorithm. The only additional information we can add here is the post-processing of the data, which we believe is not significant.

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?

Reply: Yes, we exclude all profiles containing cloudy pixels, not just the bins. Therefore, profiles with aerosol layers above clouds are also not considered in estimating the CCN data. As per Tackett et al. (2018), this filtering ensures the best quality in level 3 aerosol product. We understand that using this strict filtering may result in a relatively less number of profiles. However, it has been found that CALIPSO underestimates the optical properties of aerosols present above clouds (Kacenelenbogen et al., 2014). For aerosols present below clouds, the retrievals are not trustworthy because of signal attenuation (Tackett et al., 2018). To address the sampling issue resulting from the strict filtering, we have opted for a coarser grid and combined the daytime and nighttime CALIPSO retrievals, as stated in Section 3.2.1 of the manuscript.

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.

Reply: Usually, a significant level of signal attenuation is caused by clouds. Aerosols, whose concentrations mostly decrease with height, usually do not degrade the signal as much as clouds when present at higher levels. This is the reason why the CALIOP signal usually penetrates the atmosphere, reaching and detecting surfaces even when multiple aerosol layers are present (Tackett et al., 2018). Therefore, we believe that it is not required to account for such an effect in the CCN-retrieval algorithm.

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?

Reply: We understand your concern. However, as stated in the response to Comment 8, the retrievals below clouds have quality issues due to signal attenuation (Tackett et al., 2018) and are hence excluded from the CCN data. Further, the CCN concentrations below clouds may not be the best for studying aerosol-cloud interactions (ACIs), as they are modulated by whether or not the cloud is precipitating (CCN sink). Therefore, we believe that the CCN present in the vicinity of clouds close to the base is the most relevant in the context of ACIs, which are well represented in our data set.
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.

Reply: We agree with you. The signal-to-noise ratio (SNR) of CALIOP is comparatively lower for the daytime retrievals, which may lead to faint aerosol layers being undetected (Winker et al., 2009, 2013), as already stated in Line 220 of the unrevised manuscript. We decided to combine the daytime and nighttime retrievals to get a decent sample size sufficient to yield regionally representative CCN data at a monthly resolution. We also believe that it will not impact the mean state of the CCN variations significantly, which is governed by the presence of thick aerosol layers, except in regions that are associated with clean conditions. We have already discussed its consequences in Section 3.3 of the manuscript. The undetected faint aerosol layers during the daytime measurements may lead to an underestimation of CCN concentrations, which may somewhat compensate for the intrinsic overestimation of the CCN-retrieval algorithm that arises from the algorithm approximations. Having said that, we have compiled the daytime and nighttime CCN files separately and are currently working on the comparison, which will be published in a separate study.

12. Line 173. How about the number of profiles at the 2°x5°grid for 1 month globally? Such as total profile, cloud-free profile and available aerosol profile. This information (figure) is very useful for reader.

Reply: Thank you for your suggestion. The data set includes two similar 3D metrics, namely ‘Na’ (number of aerosol samples) and ‘N’ (total number of clear air and aerosol samples) detected by CALIPSO within a grid box. Therefore, the maximum value of N in a grid column in a month is equivalent to the total number of profiles used to produce the weighted average CCN. A difference between N and the total number of profiles would arise when the aerosol-related quality flags have impacted many bins within a column. We decided to use and save the N and Na metrics, as they serve a key role in converting the monthly data to seasonal and annual weighted means.

While showing a figure on the number of profiles used in producing the CCN data is important, we believe that the number of days observed (NDO in the manuscript) in a month has better physical meaning in terms of understanding the representativeness of the monthly data set. We already discussed NDO using texts (see Lines 173–176 and 196–198 of the unmodified manuscript) and a figure (see Figure 1) in the manuscript.

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?

Reply: The algorithm is independent of the altitude of the measurements. However, over 8 km, there were not many aerosol bins detected by CALIPSO, with a contribution to the total CCN in terms of magnitude of about 0.7%. Therefore, we excluded such high-altitude retrievals from the data, which would have otherwise increased the data size unnecessarily. We have now included the following in the updated manuscript:

"As the retrievals over an altitude of 8 km constitute about 0.7% of the total tropospheric CCN, to reduce the overall data size, all parameters are limited to a maximum altitude of 8 km."
As stated in Section 3 of the manuscript, the algorithm is designed for application to the CALIPSO level 2 profile product (Choudhury and Tesche, 2022; Choudhury et al., 2022). However, because of the very narrow cross-swath coverage of the lidar (highly collimated beams), the distance between two consecutive CALIPSO overpasses is remarkably high compared to passive sensors like MODIS (Moderate Resolution Imaging Spectroradiometer). In other words, we will not have a CALIPSO overpass in all the grid boxes in a day unless we use a very coarse grid for constructing daily global data. Even going to a monthly resolution with a grid of 2°X5° covers a maximum of about 17 days in the tropics, which increases as we go towards the poles. Therefore, using solely CALIPSO observations to produce a daily average would yield unrealistic results. Also, instantaneous global data is not possible as CALIPSO is a polar-orbiting satellite.

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

Reply: Unfortunately, we do not include an uncertainty indicator in the data set. However, as stated in Section 3.3 of the manuscript, the uncertainty associated with the CCN concentrations can go as high as a factor of 2 (+100% and -50%). As the CCN concentrations are weighted averages with values ranging many orders of magnitude, we initially planned to include CCN percentiles in the data. However, as the data is 3D and the CCN concentrations are reported for 5 aerosol types, including percentiles nearly tripled the size of the data. Therefore, we decided to omit such a metric.

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

Reply: Good point. However, we expect the data to be different from AOD, as the latter is an extensive property. Coarser dust aerosols, even though present in small numbers, can have a higher optical depth compared to finer pollution aerosols. Therefore, we compared our data set with the modelled CCN concentrations in the manuscript.

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.

Reply: We understand your point. As this is a data descriptor paper, we wanted to focus on explaining the production, usage, and limitations of the data set. As an example of potential application, we compared the data with global model simulations and demonstrated the ability of the data to capture spatial and seasonal CCN variations for aerosols that are relevant to ACI (generally present in the boundary layer below 2 km) in Section 4 of the manuscript. Further studies on the vertical structure of aerosol-type-specific CCN are outside the scope of the current paper but will be discussed comprehensively in future research, as discussed in the usage notes in Section 5 of the manuscript.
Minor Comments:

1. Line 12. It should be revised from “warm clouds” to “clouds”.
   Reply: Corrected.

2. Line 21 and 147. Please add references for these comments.
   Reply: Added.

3. Line 64. Please explain the “(NASA/LARC/SD/ASDC, 2018)”. The reader may be not known CALIPSO well.
   Reply: It is the standard way of citing CALIPSO level 2 aerosol profile product which is used in this work. Please see "https://asdc.larc.nasa.gov/project/CALIPSO/CAL_LID_L2_05kmAPro-Standard-V4-20_V4-20/citation" (last accessed on 24 May 2023) for more details.

4. Line 120. The paragraph is not structured well. Please number all criterions about quality control to let reader can understand easily
   Reply: We apologize for the confusion. However, we have already listed all the quality control criteria in Table 1. We are therefore not modifying the paragraph.

5. Line 152. Please add more detailed descriptions about “kappa parameterization”.
   Reply: Added. We now have included the parametrization equation and explained how we compute the extinction coefficients at different relative humidities in Section 3.1.3.

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).
   Reply: Thank you for your suggestion. We have cited the article in the updated manuscript.

7. Please add longitude and latitude labels for all figures.
   Reply: Thank you for your suggestion. We have added the required labels in all the figures.

8. The caption of all figures is not detailed and clear. Such as time range? Altitude?
   Reply: Thank you for pointing it out. We missed the altitude range in the Figure 2 caption, which is now added. However, other required information is already given in the figure captions. Figure 1 depicts the NDO (dimension of lat x lon), which does not have an altitude dimension. Figures 3 and 4 captions include the altitude range. From either the caption or the definition of the
plotted parameters (e.g., climatology plots), it is clear that Figures 1, 3, and 4 use the complete data set.

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.

Reply: We apologize for the confusion. Thank you for pointing it out. We have modified the color bars used in the figures as per your suggestions.

10. I do not understand why you copy all figures at supplement information again.

Reply: We apologize for the confusion. The supplementary file includes the Matlab live scripts (in pdf format) used to produce the plots from the monthly gridded data presented in the paper. It is intended to help users get started with the data.

11. Table 1. The comment of extinction QC is not correct

Reply: The extinction QC flag depicts the final status of the extinction solution for each aerosol layer (Tackett et al., 2018). Retrievals with extinction QC values of 0, 1, 16, and 18 have the highest confidence as they are associated with no or little change in the initial assumed or computed lidar ratio. The comment in Table 1 for extinction QC in our manuscript says "low confidence in extinction retrievals for other values", which we believe is 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.

Reply: Already addressed in the response to the Comment 12 under the Major comments section.
The authors describe a data set consisting of global, gridded, height-resolved cloud condensation nuclei (CCN) concentrations retrieved from spaceborne lidar measurements. Such data sets are of importance, e.g., as input and validation data for models, e.g. global circulation models. The authors thoughtfully provide potential applications of their data set and kindly invite readers to use the described data set for further important studies. In general, the manuscript is well written and comprehensible. Some parts are elucidated in more detail, while other parts lack a bit of information which might be of interest to the reader as well (e.g., see comment on dust separation). This data description paper fits well into the scope of Earth Syst. Sci. Data and I suggest publication after minor revisions.

Reply: We are thankful to the reviewer for their thorough examination of our work and for providing valuable constructive feedback that has enhanced the overall quality and readability of our manuscript. We greatly appreciate their time and effort. The detailed responses to their comments can be found in the following section.

1. Line 62: January 1 2021 or July 1 2020? What exactly is meant here? The reference states "The V4.21 data product covers July 1, 2020 to current." (https://asdc.larc.nasa.gov/project/CALIPSO/CAL_LID_L2_05kmAPr-Standard-V4-20_V4-20) Are there significant differences between these two (sub)versions?

Reply: Thank you for pointing this out. You are correct. Version 4.21 product covers from July 1 2020 to present. There are no differences in the scientific algorithms used to generate these two products. The version bump was a result of an upgrade to the operating system used in CALIPSO’s data production system.

We have corrected the date in the updated manuscript.

2. Line 71: Why is there no data available for February 2016?

Reply: As per the CALIPSO’s official statement, a GPS anomaly resulted in no scientific data collection from 28 January 2016 till 14 March 2016. More information can be found at "https://www-calipso.larc.nasa.gov/tools/instrument_status/inst_status-2016.php" (last accessed on 24 May 2016).

3. Line 129: Add a reference where this Minimum Laser Energy parameter is described, please. It is also available for V4.21, I suppose?

Reply: Thank you for pointing it out. We have added the following web reference to the Minimum Laser Energy parameter from CALIPSO’s official webpage to the revised manuscript.

4. Line 144: I suggest to explicitly state the used lidar ratios (Kim et al., 2018) and separation thresholds (particle linear depolarization ratio of dust 0.31? non-dust 0.05?).

Reply: Thank you for your suggestion. We have added the following sentence on the depolarization ratio limits to Section 3.1.2 of the updated manuscript:

"We assume the aerosol mixture to be pure dust (non-dust) if the particle depolarization ratio is > 0.31 (< 0.05). The backscatter coefficients of aerosol mixtures with a depolarization ratio between 0.05 and 0.31 are separated using Eq. (14) of Tesche et al. (2009)."

However, we believe that the lidar ratios can be easily accessed from Kim et al. (2018). Therefore, instead of explicitly stating them, we have indicated that the values are taken from Table 2 of Kim et al. (2018). The updated text in the manuscript is as follows:

"The lidar ratio for these aerosol types is taken from Kim et al. (2018, Table 2)."

5. Line 254ff: Given the provided two digits after comma averages, please also provide standard deviations.

Reply: Thank you for your suggestion. We have now mentioned the standard deviations along with the mean CCN concentrations in the revised manuscript.

6. Similar to Referee 1, I wonder about CCN retrievals at heights >8 km AGL, and even more so, about potential application of this technique (or a similar, adapted one) on INP (ice nucleation particle) retrievals, even though, I understand that a substantial discussion of that would be beyond the scope of a data description paper.

Reply: The algorithm is independent of the altitude of the measurements. However, over 8 km, there were not many aerosol bins detected by CALIPSO, with their contribution to the total CCN in terms of magnitude being about 0.7%. Therefore, we excluded such high-altitude retrievals from the data, which would have otherwise increased the data size unnecessarily. We have now included the following in the updated manuscript:

"As the retrievals over an altitude of 8 km constitute about 0.7% of the total tropospheric CCN, to reduce the overall data size, all parameters are limited to a maximum altitude of 8 km."

Regarding the potential application of the algorithm to retrieve INP from CALIOP, we are still in the development and validation phases. However, another popular technique given by Mamouri and Ansmann (2015, 2016) has shown promising results in retrieving INP concentrations from ground-based and spaceborne lidar measurements (Marinou et al., 2019). As our paper specifically focuses on the CCN data set, we have not included this discussion in our manuscript.

7. Further, rather minor textual recommendations can be found in the attached pdf file.
Reply: Thank you for your recommendations. We have corrected all of them in the revised manuscript.

References:


