



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

Goutam Choudhury^{1,2} and Matthias Tesche¹

¹Leipzig Institute for Meteorology (LIM), Leipzig University, Stephanstrasse 3, 04103 Leipzig, Germany

²Department of Geography and Environment, Bar-Ilan University, 5290002 Ramat Gan, Israel

Correspondence: Goutam Choudhury (goutam.choudhury@biu.ac.il)

Received: 11 March 2023 – Discussion started: 24 March 2023

Revised: 29 June 2023 – Accepted: 15 July 2023 – Published: 22 August 2023

Abstract. We present a global multiyear height-resolved data set of aerosol-type-specific cloud condensation nuclei concentrations (n_{CCN}) estimated from the spaceborne lidar aboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite. For estimating n_{CCN} , we apply the recently introduced Optical Modelling of the CALIPSO Aerosol Microphysics (OMCAM) algorithm to the CALIPSO Level 2 Aerosol Profile product. The estimated n_{CCN} are then gridded into a uniform latitude–longitude grid of $2^\circ \times 5^\circ$, a vertical grid of resolution 60 m from the surface to an altitude of 8 km, and a temporal resolution of 1 month. The data span a total of 186 months, from June 2006 to December 2021. In addition, we provide a 3D aerosol-type-specific climatology of n_{CCN} produced using the complete time series. We further highlight some potential applications of the data set in the context of aerosol–cloud interactions. The complete data set can be accessed at <https://doi.org/10.1594/PANGAEA.956215> (Choudhury and Tesche, 2023).

1 Introduction

Airborne aerosols can serve as cloud condensation nuclei (CCN) to form liquid cloud droplets and influence the properties of clouds. For instance, an increase in the number of aerosols that can act as CCN may result in more but smaller cloud droplets at a constant cloud liquid water content. As a result, the effective surface area of the cloud available to interact with incoming solar radiation increases, enhancing the cloud albedo and cooling the Earth (Twomey, 1974). Furthermore, it may take longer for these smaller droplets to grow large enough to form precipitation, leading to an increase in cloud lifetime and cloud cover and, thus, imposing an additional cooling effect (Albrecht, 1989). Understanding and quantifying such aerosol–cloud interactions (ACIs) is necessary to predict and mitigate their potential impacts on the Earth’s radiation budget (Forster et al., 2021).

Only a fraction of aerosols, depending primarily on their size and, to a lesser extent, chemical composition, can act as CCN (Dusek et al., 2006). This dependency is further regulated by atmospheric water vapour supersaturation, which is a function of meteorological parameters like updraught ve-

locity, humidity, temperature, and pressure (Seinfeld et al., 2016). The supersaturation near the cloud base of boundary layer clouds is usually about 0.15 %–0.20 %. Under such conditions, continental and marine aerosols with a dry radius greater than 50 nm ($n_{50,dry}$) and dust aerosols with a dry radius greater than 100 nm ($n_{100,dry}$) form the reservoir of most favourable CCN (Mamouri and Ansmann, 2016; Choudhury and Tesche, 2022b). The concentration of such particles varies with geographical location and, as per the findings from surface in situ measurements, can range from about 10 cm^{-3} under pristine conditions to as much as 10^5 cm^{-3} in polluted urban air masses (Schmale et al., 2018). Similar orders of magnitude of variation in CCN concentrations are also seen in the vertical dimension, with the highest concentrations close to the surface and only a few particles in the upper troposphere (Brock et al., 2021; Zhang et al., 2022). Given the wide range of CCN concentrations, global information on the horizontal and vertical distribution of CCN is necessary to accurately quantify the impact of ACIs on climate (Bellouin et al., 2020; Quaas et al., 2020).

Today, there is a lack of observation-based global CCN data. While climate model outputs can be used to study the global distribution of CCN, they have been identified to underestimate CCN at the surface by a normalized mean error of about 40 % to 80 % depending on the geographical location (Fanourgakis et al., 2019a). In situ measurements, despite being performed continuously with a high temporal resolution, are localized to specific land sites and do not cover oceans. Also, such measurements are usually limited to the surface, as it is not feasible to operate airborne measurements over longer periods. Here, satellite remote sensing, specifically based on spaceborne lidar, emerges as the best way to obtain a height-resolved, long-term, and observation-based global picture of atmospheric CCN concentrations.

The polar-orbiting Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite includes the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument, which provides height-resolved information on aerosol optical properties for different aerosol types (Winker et al., 2009). Choudhury and Tesche (2022a) present an algorithm that uses the spaceborne lidar measurements and the aerosol microphysics included in the CALIPSO aerosol model (Omar et al., 2009) to estimate height-resolved cloud-relevant aerosol number concentrations ($n_{50,\text{dry}}$ and $n_{100,\text{dry}}$) and CCN concentrations. The results are found to be consistent with in situ measurements at various geographic locations over land and ocean covering different aerosol environments (Choudhury et al., 2022; Choudhury and Tesche, 2022b; Aravindhavel et al., 2023). This highlights the potential of CALIOP measurements with respect to producing an unprecedented global CCN data set from satellite observations.

In this paper, we present a global monthly 3D data set of the number concentration of CCN (n_{CCN}) for different aerosol types that has been constructed by applying the CCN retrieval algorithm of Choudhury and Tesche (2022a) to more than 15 years (June 2006 to December 2021) of CALIPSO Level 2 Aerosol Profile data. We also provide seasonal and annual climatologies of n_{CCN} estimated from the monthly data. All data sets are produced at a uniform latitude–longitude grid of $2^\circ \times 5^\circ$ resolution and a vertical resolution of 60 m, consistent with the grid of the CALIPSO Level 3 monthly product (Tackett et al., 2018). The data records are given in Network Common Data Form (NetCDF) format separately for each year and can be accessed at <https://doi.org/10.1594/PANGAEA.956215> (Choudhury and Tesche, 2023). Details regarding the input data and algorithm used to produce the CCN data set are discussed in Sect. 2. Section 3 provides the description of the pre- and post-processing steps involved in the production of the gridded data and elaborates on the physical meaning and application of all of the variables stored in the data. The estimated seasonal and annual n_{CCN} climatologies and a comparison of the annual averaged n_{CCN} with the global climate model outputs from Fanourgakis et al. (2019a) for the year 2011 are

presented in Sect. 4. We conclude the paper with potential applications of the global CCN data set in Sect. 6.

2 Data and CCN retrieval

2.1 CALIPSO aerosol profile product

The CALIPSO polar-orbiting satellite has been operational since June 2006. CALIOP, a nadir-viewing elastic-backscatter lidar, is the principal payload on CALIPSO. It measures vertical profiles of aerosol and cloud properties at two wavelengths: 532 and 1064 nm. In this work, we use aerosol optical properties at 532 nm from the CALIPSO Level 2 version 4.20 (4.21 since July 2020) aerosol profile product (NASA/LARC/SD/ASDC, 2018), such as the aerosol extinction coefficient (α), backscatter coefficient (β), and depolarization ratio (δ). All parameters are provided at an along-track horizontal resolution of 5 km with a vertical resolution of 60 m for tropospheric aerosols. We further use the atmospheric volume description flag to access the information on aerosol type, which contains seven categories: marine, desert dust, polluted continental, clean continental, elevated smoke, polluted dust, and dusty marine. These aerosol types are classified based on the type of underlying surface; the 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). Additionally, quality control flags are utilized to screen the data for the most reliable retrievals (further details given in Sect. 3). We also use the auxiliary profiles of ambient relative humidity (RH) and pressure in the Level 2 data product. These are derived from the Global Modelling and Assimilation Office (GMAO) Global Data Assimilation System (GDAS; Molod et al., 2015). The Level 2 Aerosol Profile data considered in this study cover a total of 186 months in the time period from June 2006 to December 2021 (data for February 2016 are not available). Note that CALIPSO underwent an orbit adjustment in September 2018 to synchronize its path with that of the CloudSat satellite. Although this orbital shift resulted in a slight variation in the geographic region observed by CALIPSO, there are currently no known issues associated with CALIPSO's retrieval quality as a result of this transition.

2.2 Model data

We use the modelled n_{CCN} from Fanourgakis et al. (2019a) for comparison to CALIPSO estimates. The data are comprised of annual average surface n_{CCN} at different supersaturations for the year 2011 with a latitude–longitude resolution of $1^\circ \times 1^\circ$ and were obtained from the output of 15 global models. Details on the model configurations and aerosol emission inventories used for modelling n_{CCN} are described comprehensively in Fanourgakis et al. (2019a). The data can be downloaded from

<https://doi.org/10.5281/zenodo.3265866> (Fanourgakis et al., 2019b).

2.3 CCN retrieval algorithm

The CCN retrieval algorithm used in this work is based on the Optical Modelling of the CALIPSO Aerosol Microphysics (OMCAM) algorithm. The algorithm is described comprehensively in Choudhury and Tesche (2022a) and Choudhury et al. (2022). In this section, we briefly review the main steps involved in OMCAM.

The OMCAM algorithm first estimates the aerosol size distribution by using (i) CALIPSO-derived aerosol-type-specific extinction coefficients, (ii) their corresponding microphysical properties (normalized volume size distributions and refractive indices), and (iii) an optical modelling package called Modelled Optical Properties of ensembles of Aerosol Particles (MOPSMAP; Gasteiger and Wiegner, 2018) for light-scattering calculations. The microphysical properties of continental, dust, and smoke aerosols are taken from CALIPSO's aerosol model (Omar et al., 2009). For marine aerosols, the microphysical properties are derived from Sayer et al. (2012) because it yields aerosol number concentrations that agree better with airborne in situ measurements (Choudhury et al., 2022). OMCAM first selects a normalized volume size distribution and refractive index based on the aerosol type identified by CALIPSO. The algorithm then scales the normalized size distribution linearly to reproduce the CALIPSO-estimated extinction coefficient as follows:

$$\frac{dV(r)}{d \ln r} = S_V \cdot \sum_{i=1}^2 \frac{v_i}{\sqrt{2\pi} \ln \sigma_i} \exp\left(\frac{-(\ln r - \ln \mu_i)^2}{2 \ln \sigma_i^2}\right), \quad (1)$$

where

$$S_V = \frac{\alpha}{\alpha_n}.$$

Here, S_V is called the volume scaling factor and is estimated from the ratio of α , the CALIOP-derived extinction coefficient, and α_n , the extinction coefficient calculated from the normalized size distribution and refractive index using MOPSMAP. For modelling α_n , we treat continental (clean continental, polluted continental, and smoke) and marine aerosols as spheres and use Mie scattering theory. We consider desert dust aerosols as spheroids and use a combination of the T matrix and the improved geometric optics method, depending on the value of the aerosol size parameter, to model α_n . The values of normalized size distribution parameters, such as the standard deviation (σ_i), volume fraction (v_i), and mean radius (μ_i) of i th (fine and coarse) mode, are listed in the Table A1. After the scaling step, the volume size distribution is converted to a number size distribution, which is then integrated starting at 50 nm to compute $n_{50, \text{dry}}$ for continental and marine aerosols and at 100 nm to compute $n_{100, \text{dry}}$ for desert dust aerosols. This is because aerosols

within these size ranges have been identified to act as CCN for a supersaturation of 0.15 %–0.20 % in several studies using in situ measurements (Koehler et al., 2009; Rose et al., 2010; Deng et al., 2011; Kumar et al., 2011; Mamouri and Ansmann, 2016). In accordance with the Aerosol Robotic Network (AERONET) size distributions, the upper radius limit for integrating the size distributions is set at 15 μm . Further, the CCN concentrations at higher supersaturation can be estimated following the simple CCN parameterization given by Mamouri and Ansmann (2016):

$$n_{\text{CCN}} = f_{\text{ss}} \cdot n_{j, \text{dry}}, \quad (2)$$

where j represents the lower radius limit of the size-distribution integration and f_{ss} is the enhancement factor with values equal to 1.0, 1.35, and 1.7 for supersaturations of 0.15 %–0.20 %, 0.25 %, and 0.40 %, respectively. While the simple parameterizations based on aerosol size and type do not consider the aerosol chemistry stringently, they have been shown to yield reasonable n_{CCN} estimates, particularly when applied to spaceborne lidar measurements (Choudhury and Tesche, 2022a, b). Further information on the limitations and known issues of the algorithm are discussed in Sect. 3.3.

3 Methodology

In this section, we discuss the complete methodology used to apply the CCN retrieval algorithm to the CALIPSO profile data to generate a global gridded monthly CCN data set. It includes a number of preprocessing stages that are implemented prior to applying the CCN retrieval algorithm, followed by a set of post-processing steps to convert the n_{CCN} profiles to global gridded data.

3.1 Preprocessing

3.1.1 Quality screening

The CALIPSO Level 2 Aerosol Profile product includes a number of quality control flags that are intended to filter out unreliable retrievals that may result from clouds misclassified as aerosols and errors pertaining to the extinction coefficient retrieval (Tackett et al., 2018). The cloud–aerosol discrimination (CAD) score specifies the confidence in the classification of aerosol and cloud for each data bin (Liu et al., 2009). We only select data which have a CAD score in the range $[-100, -20]$, as they correspond to a high confidence in the aerosol classification (Liu et al., 2009). To screen retrievals with extinction-coefficient-related retrieval issues, we use the “extinction uncertainty” and the “extinction QC” (extQC) metrics. Data bins with an extinction uncertainty of -99.99 km are not considered and nor are the bins present below them, as the extinction retrieval uncertainty may propagate to solutions at lower altitudes. The extQC flag describes how the extinction coefficient is generated for each Level 2 sample. Following Tackett et al. (2018), we

Table 1. Quality screening applied to the CALIPSO Level 2 Aerosol Profile product.

| Quality flags | Valid range or values | Comments |
|-----------------------------------|---------------------------------|---|
| Cloud aerosol discrimination flag | $=[-100, -20]$ | Bins outside the bound have low confidence in aerosol classification |
| Extinction QC flag | $= 0, 1, 16, \text{ and } 18$ | Low confidence in extinction retrievals for other values |
| Extinction uncertainty flag | $\neq -99.99 \text{ km}^{-1}$ | Range bins below the affected one are also rejected |
| Feature-type flag 1 | reject if $= 2$ | Reject a profile if it has at least one cloudy bin |
| Feature-type flag 2 | reject if $= 0 \text{ or } > 4$ | Reject any samples if invalid, surface, subsurface, or attenuated |
| Feature-type flag 3 | if $= 1$, set $\alpha = 0$ | Extinction coefficient of clear air is assumed to be 0 km^{-1} |
| Minimum laser energy filter | reject if < 0.08 | Deteriorated data quality in the South Atlantic Anomaly region since mid-2017 |

only consider bins with extQC values of 0, 1, 16, and 18, because they represent high confidence in the retrieved extinction coefficient. We also use the minimum laser energy (MLE) parameter at 532 nm, which was recently introduced in the version 4.20 update, to identify and screen columns affected by low-laser-energy shots ($\text{MLE} < 0.08$) that result in higher noise and degraded retrieval quality. Additionally, we filter out the profiles or columns with cloudy pixels, as clouds can impede aerosol retrievals underneath them due to signal attenuation. All of the quality-screening criteria used in this work are listed in Table 1. Readers are advised to refer to Tackett et al. (2018) for further information on the production and functions of the quality-screening flags.

3.1.2 Dust separation

Before applying the CCN retrieval algorithm to the quality-screened Level 2 data, the extinction coefficients of dust mixtures (polluted dust and dusty marine) need to be separated into dust and non-dust components. This is done by using the particle depolarization ratio and first separating the backscatter coefficient into dust and non-dust components following the methodology described in Tesche et al. (2009). 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). This is a rather straightforward technique that assumes aerosols to be externally mixed. It has been implemented and tested in several studies based on ground-based and spaceborne lidar retrievals (Mamouri and Ansmann, 2015, 2016; Choudhury et al., 2022; Choudhury and Tesche, 2022b). The extinction coefficients of dust and non-dust aerosol components are then estimated by multiplying the respective separated backscatter coefficients by the lidar ratios corresponding to the aerosol type. Following this methodology, the polluted dust aerosol mixture is separated into polluted continental and desert dust components, and the dusty marine aerosol mixture is separated into desert dust and marine components. The lidar ratio for these aerosol types is taken from Table 2 of Kim et al. (2018).

3.1.3 Hygroscopicity correction

Hydrophilic aerosols may grow in size under moist conditions and, thus, may result in a higher extinction coefficient relative to dry conditions, even if the aerosol number concentrations remains the same (Zieger et al., 2013). Further, the normalized size distributions used in OMCAM have been derived mostly under dry conditions (Omar et al., 2009), and the CCN parameterization requires the dry-aerosol number concentration. Therefore, the extinction coefficient is first corrected for aerosol hygroscopicity before applying the OMCAM algorithm. This is done by using the aerosol-type-specific growth factors (ratio of ambient and dry extinction coefficients) given in Choudhury et al. (2022) at different RH values. The growth factors for continental and marine aerosols are estimated at different RH values using the κ parameterization (Petters and Kreidenweis, 2007), which expresses the hygroscopic growth of an aerosol particle with a dry radius r_{dry} in terms of a hygroscopicity parameter κ as follows:

$$\frac{r_{\text{wet}}(\text{RH})}{r_{\text{dry}}} = \left(1 + \kappa \cdot \frac{\text{RH}}{1 - \text{RH}}\right)^{\frac{1}{3}}. \quad (3)$$

Here, r_{wet} is the wet radius of a particle at a given relative humidity (RH). The ratio between r_{wet} and r_{dry} is used to modify the dry normalized size distributions of aerosols at different RH values, followed by the computation of their respective extinction coefficients using the MOPSMAP package. A globally averaged κ value of 0.3 is used for polluted continental, clean continental, and smoke aerosols, whereas a value of 0.7 is used for marine aerosols (Andreae and Rosenfeld, 2008). We treat desert dust as non-hygroscopic and do not modify the extinction coefficients for dust samples. This approach has been shown to yield dry extinction coefficients and dry number concentrations that are consistent with in situ measurements (Choudhury et al., 2022; Choudhury and Tesche, 2022a). Further details on the hygroscopicity correction are given in Choudhury et al. (2022).

3.2 Post-processing

We use the quality-screened CALIPSO Level 2 data obtained by following the steps outlined in Sect. 3.1.1 in the OMCAM

algorithm to estimate n_{CCN} profiles at a supersaturation of 0.20 %. We further limit the n_{CCN} in the output data to a supersaturation of 0.20 %, as it can be easily converted to higher supersaturations by using the enhancement factors as shown in Eq. (2). In this section, we discuss the spatial and temporal grid configuration used to convert the n_{CCN} profiles to 3D gridded data. We further describe all of the parameters stored in the output data files.

3.2.1 Gridding, sampling, and averaging

The CALIPSO Level 2 Aerosol Profile data have a very high vertical resolution of 60 m and an along-swath resolution of 5 km. However, because of the negligible divergence of lasers, CALIPSO has a narrow cross-track coverage. Further, the distance between two consecutive overpasses is inversely proportional to the latitude because of CALIPSO's polar orbit with an inclination of 98.36°. Therefore, a grid box of a certain dimension close to the Equator will include fewer overpasses compared with one close to the poles. Due to this factor and to produce a regionally representative global gridded data with a sufficient aerosol sampling frequency in each grid cell, we choose a monthly temporal resolution and a horizontal grid resolution of $2^\circ \times 5^\circ$ to process the n_{CCN} profiles. The vertical resolution is, however, kept unchanged at 60 m. We further consider both the daytime and night-time CALIPSO overpasses to compute the average n_{CCN} to increase the sampling frequency. This grid configuration is also used in monthly CALIPSO Level 3 products (Tackett et al., 2018) and is suggested by Choudhury and Tesche (2022b) to compile a global n_{CCN} data set. Using this grid, we estimated the number of days observed with valid n_{CCN} (value ≥ 0 ; hereafter abbreviated as NDO) samples in each grid box for each month and found them to be as high as 17 d in the tropics and 31 d at the poles. Note that the NDO estimated here is different from the “Days_Of_Month_Observed” parameter in the CALIPSO Level 3 product, as the latter also considers valid cloud retrievals.

We further produced a climatological n_{CCN} data set using the same horizontal and vertical grid configuration. An alternative approach would be to consider a higher horizontal resolution of $1^\circ \times 1^\circ$, as used in Amiridis et al. (2015), to construct a climatology of aerosol and cloud properties using the CALIPSO Level 2 data. However, using such a grid results in low aerosol sampling, especially in the tropics. This is shown in Fig. 1 based on more than 15 years of CALIPSO measurements. While the former gives a maximum of 2340 d observed in the tropics, the latter results in a significantly lower maximum of 591 d. Therefore, we use the coarser grid to produce the n_{CCN} climatology that ensures the ample aerosol sampling required for a realistic and regionally representative data set.

For averaging the n_{CCN} for each latitude, longitude, and altitude grid cell, we follow the methodology used in Tackett et al. (2018) for producing the CALIPSO Level 3 aerosol

products. We assign the clear-air Level 2 bins with an extinction coefficient of 0 km^{-1} , which leads to an n_{CCN} of 0 cm^{-3} , and compute the average n_{CCN} (\bar{n}_{CCN}) for each grid cell as follows:

$$\bar{n}_{CCN} = \frac{\sum_{i=1}^{N_a} n_{CCN,i}}{N_a + N_{ca}}. \quad (4)$$

Here, N_a is the number of Level 2 samples with an aerosol extinction coefficient $> 0 \text{ km}^{-1}$ and N_{ca} is the number of clear-air samples. This averaging scheme is used to generate the global monthly gridded data set as well as the n_{CCN} climatology.

3.2.2 Output data records

Following the gridding and averaging scheme discussed in the previous section, we produce the global n_{CCN} data sets at a monthly temporal resolution. Table 2 provides a list of all of the parameters included in the output data and their description. Along with the averaged n_{CCN} , we further provide the total number of Level 2 samples with an aerosol extinction coefficient $\geq 0 \text{ km}^{-1}$ ($N = N_a + N_{ca}$) and with an extinction coefficient > 0 (N_a). The n_{CCN} , N , and N_a are also provided separately for each aerosol type. Note that both N and N_a may not be equal to the sum of the contributions from all of the aerosol types, specifically when aerosol mixtures are present. They are useful in computing annual and seasonal averages of n_{CCN} using Eq. (4). We further provide the NDO for each month and suggest using the data only when $\text{NDO} > 10$ (coverage of a minimum of 30 % of days a month). Average pressure and temperature are also provided for each latitude, longitude, and altitude grid cell. The pressure values can be used to convert the data from height coordinates to pressure coordinates. As the retrievals over an altitude of 8 km constitute about 0.7 % of the total tropospheric CCN, all parameters are limited to a maximum altitude of 8 km in order to reduce the overall data size. For ease of accessibility across different platforms, we provide the data in Network Common Data Form (NetCDF) format with a medium level of data compression (deflate level of 5). These NetCDF files are accessible at <https://doi.org/10.1594/PANGAEA.956215> (Choudhury and Tesche, 2023) and are tested to work with tools and software like Climate Data Operators (CDO), netCDF Operators (NCO), and ncdump. The netCDF files with monthly gridded data are provided separately for each year from 2006 to 2021. The n_{CCN} climatology is provided in a separate netCDF file, with a data structure and nomenclature similar to the monthly data. The file also includes n_{CCN} climatologies for the boreal winter (December, January, and February), spring (March, April, and May), summer (June, July, and August), and autumn (September, October, and November) seasons.

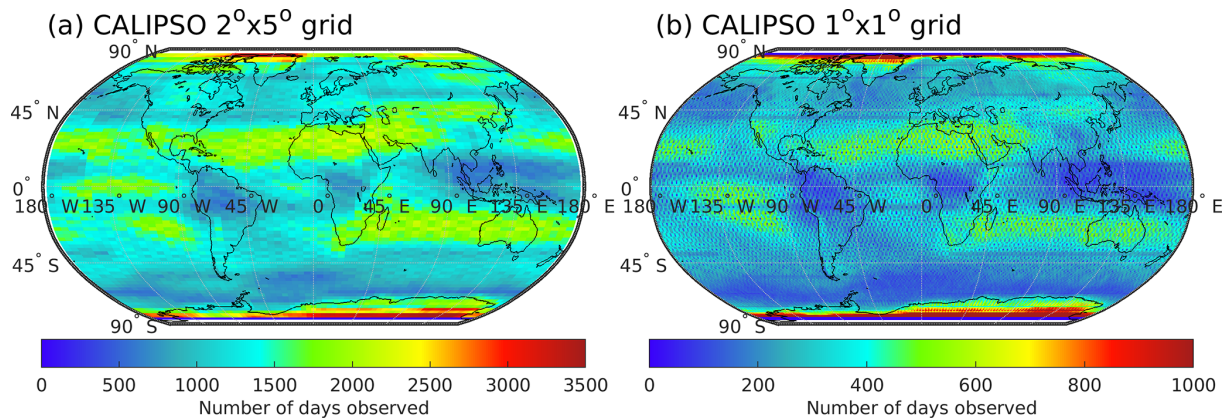


Figure 1. Global map of the number of days with valid CCN observations ($n_{\text{CCN}} \geq 0 \text{ cm}^{-3}$) for horizontal grid configurations of $2^\circ \times 5^\circ$ (a) and $1^\circ \times 1^\circ$ (b) produced from more than 15 years of CALIOP measurements.

Table 2. Description of the NetCDF data records.

| Parameter | Unit | Description |
|---------------------------------|--------------------|--|
| lon | $^\circ \text{ E}$ | Longitude midpoint |
| lat | $^\circ \text{ N}$ | Latitude midpoint |
| altitude | km | Altitude midpoint above mean sea level |
| Monthly files | | |
| time | days | Days since 1 January 2000 |
| CCN | cm^{-3} | Cloud condensation nuclei (CCN) concentration at 0.20 % supersaturation |
| CCN_std | cm^{-3} | Standard deviation of CCN |
| CCN_j $j = m, d, pc, cc, es$ | cm^{-3} | Type-specific CCN: <i>m</i> – marine, <i>d</i> – dust, <i>pc</i> – polluted continental, <i>cc</i> – clean continental, and <i>es</i> – elevated smoke |
| CCN_j_std | cm^{-3} | Standard deviation of CCN_j |
| <i>N</i> | unitless | Level 2 bin counts with an aerosol extinction coefficient ≥ 0 |
| <i>N_j</i> | unitless | <i>N</i> for different aerosol types |
| <i>Na</i> | unitless | Level 2 bin counts with an aerosol extinction coefficient > 0 |
| <i>Na_j</i> | unitless | <i>Na</i> for different aerosol types |
| <i>P</i> | hPa or mbar | Pressure |
| <i>T</i> | $^\circ \text{ C}$ | Temperature |
| DMO | unitless | Number of days observed in a month with $N > 0$ |
| Climatology file | | |
| CCN_cl | cm^{-3} | Climatology of CCN |
| CCN_cl_j | cm^{-3} | Type-specific CCN climatology |
| CCN_cl_sn | cm^{-3} | Seasonal climatology of CCN |
| CCN_cl_sn_j | cm^{-3} | Type-specific seasonal CCN climatology |
| NDO | unitless | Number of days observed with $N > 0$ |

3.3 Uncertainty, known issues, and validation

The uncertainty in the estimated n_{CCN} can arise from the uncertainties in the input parameters, such as the aerosol extinction coefficient, type-specific normalized size distribution, and the ambient RH, as well as from the dust-separation technique and the CCN parameterizations used in the algorithm. Choudhury and Tesche (2022a) studied the sensitivity of the cloud-relevant aerosol number concentrations ($n_{j,\text{dry}}$) to the variations in the size distributions for each aerosol type at different RH values. They reported a variation of a factor of 1–1.5 depending on the aerosol type. Combining this with the uncertainties associated with other previously mentioned sources, the overall uncertainty associated with the output n_{CCN} is found to be between a factor of 2 and 3. Such a range is reasonable for the spaceborne retrieval of n_{CCN} (Shinozuka et al., 2015; Mamouri and Ansmann, 2016), as atmospheric CCN concentrations can potentially vary by orders of magnitude in space and time (Schmale et al., 2018).

There are some known issues associated with CALIOP's retrieval algorithm and the CCN parameterizations used to produce the global data set. First, faint aerosol layers with an extinction coefficient $< 0.001 \text{ km}^{-1}$ (optical depth < 0.01) may not exceed the signal-to-noise ratio required to be detected by CALIOP (Tackett et al., 2018; Mao et al., 2022). The background noise due to solar radiation further impacts the feature detection, especially for the daytime retrievals (Winker et al., 2009, 2013). Such layers may therefore be classified as clear air by CALIOP's feature classification algorithm and assigned with a zero extinction coefficient. This may result in an underestimation of the average extinction and, thus, the n_{CCN} , particularly in grid cells comprising a clean environment (rural continental sites and higher altitudes). Second, the simple aerosol-type-specific parameterizations implemented in this study assume that all particles over a certain minimum radius are CCN active. This may lead to an overestimation of n_{CCN} that may also compensate for the underestimation due to the undetected aerosol layers. Third, CALIOP cannot distinguish between polluted continental and smoke aerosol layers that occur below a layer-top height of 2.5 km (Kim et al., 2018). Therefore, isolated smoke layers that do not extend above a height of 2.5 km may get classified as polluted continental aerosols, leading to an overestimation of n_{CCN} by about 13.6%. As a result, we anticipate that the CCN retrieval algorithm will overestimate the CCN in areas often influenced by smoke aerosols below 2.5 km. Additionally, strong signal attenuation caused by optically thick aerosol layers located above may lead to increased uncertainties in the retrievals of layers below. Nevertheless, it is anticipated that these retrievals will be filtered out during the quality-screening process.

As mentioned earlier in Sect. 3.2.1, we combine daytime and night-time CCN retrievals to achieve an optimal sampling frequency. The daytime retrievals usually have a lower signal-to-noise ratio compared with night-time retrievals,

which may result in higher retrieval uncertainty (Young et al., 2013; Tackett et al., 2018). By comparing the daytime and night-time CCN climatology (refer to Fig. S1 in the Supplement), we observe higher values over continents in the former. This observation is expected because anthropogenic emissions are more prominent during the day. The values over oceans are comparable in both cases. However, it is important to note that this concurrence may be attributed to long-term averaging used in computing the climatologies. Therefore, a more detailed comparison is required at various temporal scales (instantaneous, monthly, seasonal, and annual) to accurately quantify the effect of merging daytime and night-time retrievals. Such an investigation is beyond the scope of the present study and will be a subject of future analysis.

The n_{CCN} values estimated using the OMCAM algorithm have been validated using independent ground-based and airborne in situ measurements for different aerosol environments (Choudhury et al., 2022; Choudhury and Tesche, 2022b; Aravindhavel et al., 2023). Even though the uncertainty associated with the CCN retrieval from CALIOP can be as high as a factor of 3, the validation results indicate a very good agreement with in situ measurements, with a normalized mean bias of $\approx 22\%$ and a correlation coefficient of ≈ 0.7 . Such consensus has not yet been achieved for spaceborne estimations of aerosol and CCN number concentrations, accentuating the potential of the spaceborne-lidar-derived global n_{CCN} data set.

4 Results

4.1 Comparison with global model outputs

A comparison of the CALIOP-estimated average global n_{CCN} with the model outputs for the year 2011 is shown in Fig. 2. Excluding the data within the Antarctic Circle (latitudes $< -66.5^\circ \text{ N}$), CALIOP-estimated CCN concentrations are, on average, larger than the models, with a normalized difference $(100 \times (\text{MODEL} - \text{CALIOP})/\text{CALIOP})$ of -65.6% , 89.4% , and 4.5% for the median, minimum, and maximum of all of the models, respectively. The larger values are expected, as the models have been identified to underestimate the n_{CCN} when compared with in situ measurements at continental surface stations with a mean error of 60% (Fanourgakis et al., 2019a). The best agreement with CALIOP is found for the maximum of all of the models, particularly in the Northern Hemisphere. For retrievals over the Southern Ocean with latitudes $< 45^\circ \text{ S}$, the modelled n_{CCN} are significantly lower than that of CALIOP with a maximum of 171.9 cm^{-3} . Recent in situ measurements show that the n_{CCN} in such regions can even exceed 500 cm^{-3} (Humphries et al., 2021), which is not observed in the model outputs, although it is well captured by CALIOP. This highlights the potential of the CALIOP-derived global n_{CCN} data set to validate model outputs at locations where in situ observations

are sparse or even non-existent. Please note that the comparison presented here is intended to highlight the data set's unique potential for application and that validation studies have already been published (Choudhury et al., 2022; Choudhury and Tesche, 2022b; Fanourgakis et al., 2019a). Readers are encouraged to use the full potential of the open-source data set to conduct further comprehensive model evaluation studies.

4.2 CCN climatology

Figure 3 shows the global n_{CCN} climatology computed from the monthly data set averaged for altitudes below 2 km using Eq. (4) and the corresponding averaged n_{CCN} profile. The global average n_{CCN} is estimated to be $365.12 \pm 649.92 \text{ cm}^{-3}$, with $683.25 \pm 1172.95 \text{ cm}^{-3}$ over land and $303.2 \pm 382.33 \text{ cm}^{-3}$ over ocean. The highest concentrations are found over regions influenced by pollution and dust aerosols, such as South and East Asia. Polluted continental aerosols with an average n_{CCN} of $124.3 \pm 459 \text{ cm}^{-3}$ have the maximum contribution to the total CCN below 2 km, followed by marine ($110.8 \pm 175.12 \text{ cm}^{-3}$), dust ($108.9 \pm 256.39 \text{ cm}^{-3}$), and elevated smoke ($21 \pm 74.61 \text{ cm}^{-3}$) aerosols. Clean continental aerosols with a global average of $0.12 \pm 0.34 \text{ cm}^{-3}$ contribute the least to the total CCN (not shown). Over land, the average CCN value is estimated to be $683.25 \pm 1172.95 \text{ cm}^{-3}$, with maximum contributions coming from polluted continental ($419.4 \pm 900.9 \text{ cm}^{-3}$), dust ($203.67 \pm 435.45 \text{ cm}^{-3}$), and smoke ($43.6 \pm 117.93 \text{ cm}^{-3}$) aerosols. The concentrations over oceans are less than half of those over land with an average of $303.2 \pm 382.33 \text{ cm}^{-3}$, with marine ($181.93 \pm 221.2 \text{ cm}^{-3}$), dust ($84.68 \pm 206.8 \text{ cm}^{-3}$), and polluted continental ($27.33 \pm 114.17 \text{ cm}^{-3}$) aerosols being the major contributors. Although pollution aerosols contribute the most to the total CCN, it is evident that dust has the most widespread coverage, encompassing nearly the entire globe, indicating its significance in ACIs, even in pristine aerosol environments far away from the continents.

When considering the vertical distribution of n_{CCN} , the highest values are observed near the surface, and these values decrease exponentially as the altitude increases. The majority of marine CCN (97 %) and continental CCN (78 %) are predominantly located at altitudes below 2 km. On the other hand, smoke and dust CCN extend into the free troposphere, with approximately 60 % and 33 % located above 2 km altitude, respectively. Interestingly, smoke CCN exhibit an opposite trend, with concentrations increasing with height and reaching a maximum between 2 and 3 km before decreasing at higher altitudes. Land-based aerosols make up the majority of free-tropospheric CCN, with 68 % located above 2 km, compared with 32 % for aerosols over oceans. They exhibit a relatively higher contribution to the global CCN across all altitude levels. It is important to note that this variation may not be observed in localized regions over oceans that are

more frequently affected by dust and smoke transported from nearby continents, for instance, the west of Africa.

4.2.1 Seasonal climatology

Figure 4 illustrates the global map presenting the seasonal climatologies of total n_{CCN} for altitudes below 2 km, along with the corresponding profiles of spatially averaged type-specific n_{CCN} . The global average n_{CCN} reaches its maximum during winter with a value of $376 \pm 800.91 \text{ cm}^{-3}$. Similar to the global climatology in Fig. 4, the average winter n_{CCN} is significantly higher over land ($725.53 \pm 1499.36 \text{ cm}^{-3}$) compared with over ocean ($321.61 \pm 427.20 \text{ cm}^{-3}$). Strong seasonality in n_{CCN} is observed in regions influenced by Northern Hemisphere summer monsoon, such as Asia and West Africa, with minimum values occurring during the summer. This pattern is likely attributed to the wet scavenging of aerosol particles associated with cloud droplet formation and precipitation during the monsoon season. The seasonality of n_{CCN} is also evident in regions like East and Central Africa, possibly resulting from changes in local biomass burning patterns (Myhre et al., 2003; van der Werf et al., 2017) that peak during the dry summer (Southern Hemisphere winter) season.

Furthermore, seasonality in the n_{CCN} profiles is observed for all of the aerosol types, except for marine aerosols. During boreal winter, CCN concentrations are predominantly limited to altitudes below 2–3 km (84 %–94 %), with the highest near-surface concentrations compared with other seasons. The vertical distribution of CCN for all of the aerosol types gradually expands to higher altitudes with the transition to warmer spring and summer seasons. Although near-surface n_{CCN} are at their lowest during summer, they contribute the most to the free-tropospheric CCN, accounting for 35 % of CCN at altitudes higher than 2 km. This is followed by spring (29 %), autumn (24 %), and winter (16 %), highlighting the substantial impact of the boundary layer depth on modulating the vertical extent of CCN throughout all seasons.

It is worthwhile noting that, in Figs. 3 and 4, we have limited our focus to the spatial variations in low-level aerosols (altitude < 2 km) and have only presented the profiles of spatially averaged n_{CCN} . The climatology data include information up to an altitude of 8 km for various aerosol types, which offers a unique opportunity to investigate the altitudinal variations in type-specific n_{CCN} across different seasons and regions throughout the globe. Such studies, however, are outside the scope of this paper, which is aimed at introducing a new global CCN data set, and will be the focus of a future publication.

5 Code and data availability

The CALIPSO Level 2 Aerosol Profile product can be downloaded from <https://doi.org/10.5067/CALIOP/>

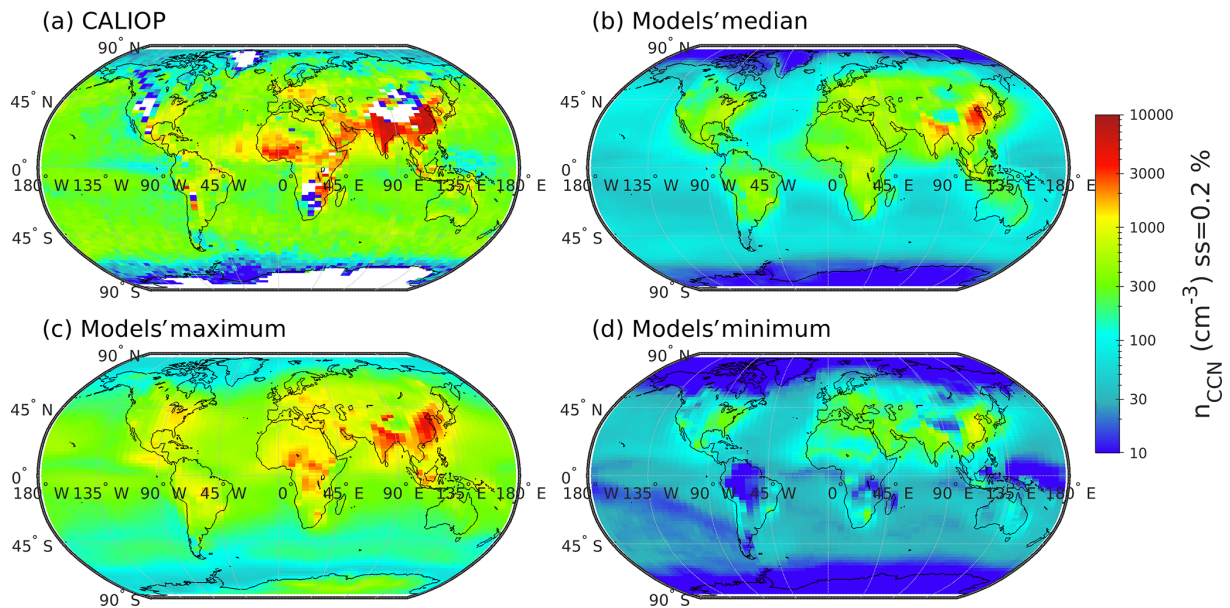


Figure 2. Comparison of global n_{CCN} maps at a supersaturation of 0.20% estimated from CALIOP between altitudes of 500 m and 1 km (a) and from the outputs of 15 global models at the surface (b–d) for the year 2011. Panels (b), (c), and (d) represent the median, maximum, and minimum modelled n_{CCN} , respectively.

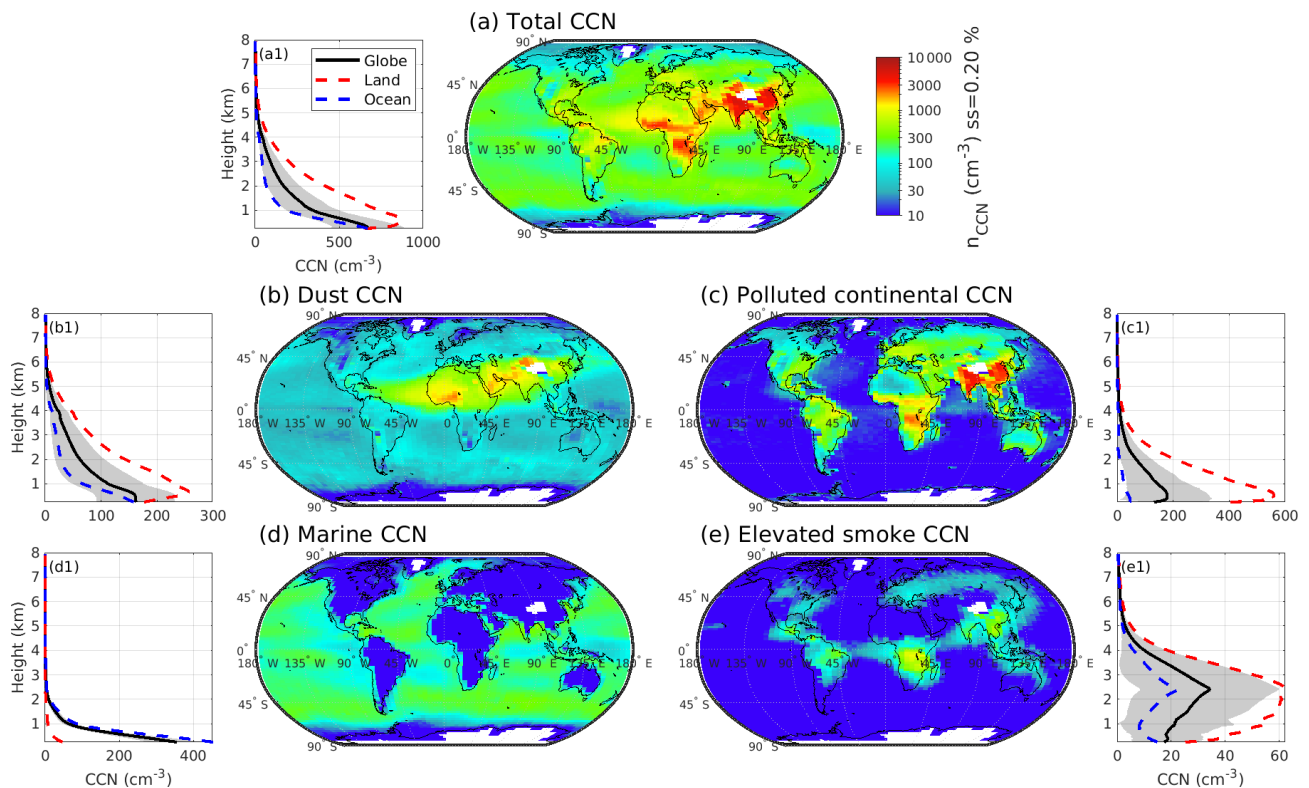


Figure 3. Global climatology of n_{CCN} at a supersaturation of 0.20% estimated from the monthly CCN data set averaged for altitudes below 2 km for all aerosol types (a), and the results separated according to mineral dust (b), polluted continental (c), marine (d), and elevated smoke (e) aerosol types, respectively. The sub-panels adjacent to each panel depict the vertical variation in n_{CCN} averaged over the globe (solid black line), land (dashed red line), and ocean (dashed blue line). The semitransparent grey patch represents half of the standard deviation in globally averaged n_{CCN} .

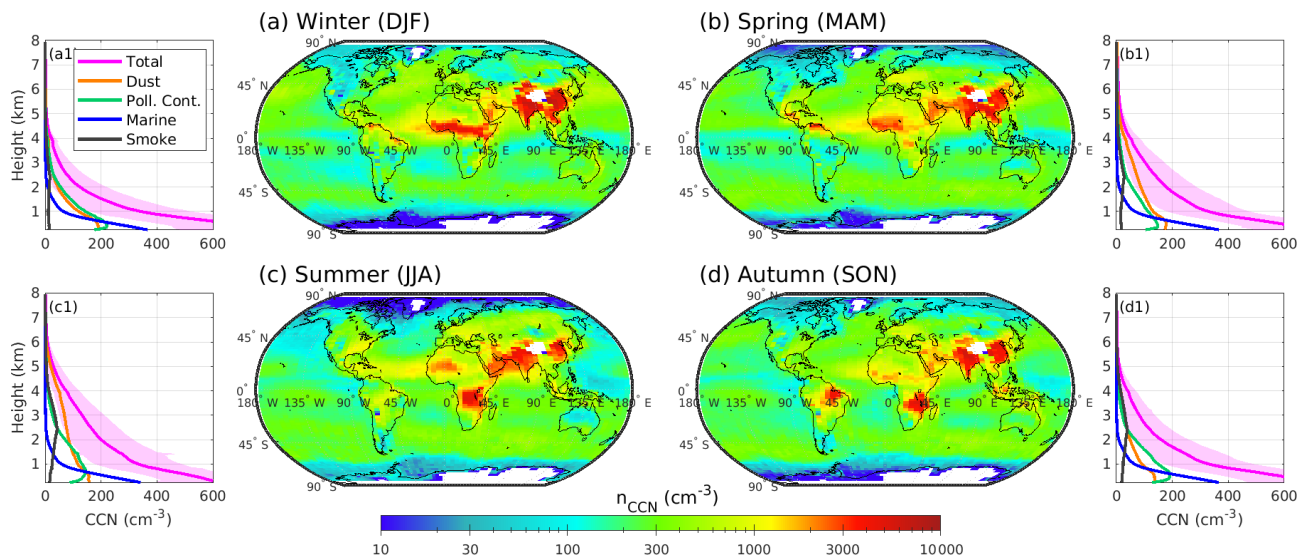


Figure 4. Seasonal climatology of n_{CCN} at a supersaturation of 0.20 % and at altitudes below 2 km for winter (a), spring (b), summer (c), and autumn (d). The vertical variation in seasonal n_{CCN} is depicted in the adjacent sub-panels for dust (orange line), polluted continental (green line), marine (blue line), and smoke (black line) aerosol types. The magenta profile in the sub-panels represent the total n_{CCN} , and the semitransparent patch highlights the associated standard deviation scaled to half of its original value for better visualization.

(NASA/LARC/SD/ASDC, 2018). The data produced in this work are available at <https://doi.org/10.1594/PANGAEA.956215> (Choudhury and Tesche, 2023). The model data used for comparison can be found at <https://doi.org/10.5281/zenodo.3265866> (Fanourgakis et al., 2019b). The codes used to generate the data were written in MATLAB and can be provided by the corresponding author upon reasonable request. The MATLAB script used to generate the plots shown in the paper is provided in the Supplement.

6 Conclusion and usage notes

We present a first aerosol-type-specific global CCN data set derived from spaceborne lidar measurements at a horizontal latitude–longitude resolution of $2^\circ \times 5^\circ$, a vertical resolution of 60 m, and a temporal resolution of 1 month. The data set spans more than 15 years, from June 2006 to December 2021, or a total of 186 months. We further use the complete time series to construct a global n_{CCN} climatology for different aerosol types. The climatologies are also reported at a seasonal timescale. These data sets are aimed at replacing the currently used satellite-derived optical proxies for CCN. Readers are encouraged to utilize the full potential of the global data set for the following purposes:

- The data can be used to investigate the horizontal and vertical distributions of CCN for various aerosol types as well as to study their trends and variations across monthly, seasonal, and annual timescales. In Sect. 4.2 of this paper, we briefly discuss the spatial distribution

of n_{CCN} for different aerosol types for altitudes limited below 2 km, where we find the global average n_{CCN} to be 365.12 cm^{-3} and identify various CCN hotspots. This approach can be expanded to include the variations along the height and time dimensions. In addition, the type-specific CCN concentrations can be used to identify the sources and sinks of CCN as well as to investigate the contributions of the long-range transport of aerosols with a high residence time (e.g. dust and smoke) to CCN in the atmosphere.

- Recent studies have demonstrated that spaceborne aerosol, cloud, and radiation measurements at a monthly temporal resolution can be used to estimate the radiative forcing associated with ACIs (Wall et al., 2022; Chen et al., 2022). The monthly data presented in this paper can be similarly coupled with cloud and radiation measurements to quantify ACIs. 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. Furthermore, the type-specific CCN data can be used to accurately quantify the anthropogenic component of present-day CCN (sum of polluted continental and smoke components), which has previously been estimated from model simulations of past climate (Bellouin et al., 2020; Quaas et al., 2020). This creates a unique opportunity to quantify ACIs by solely using satellite observations independent of model simulations.

– Due to the lack of any observation-based global CCN data, model evaluation studies usually rely on in situ measurements, which are only available for a limited time period at certain point locations. Further, the observations over oceans and the height-resolved measurements restricting the model evaluation to specific land sites are sparse. However, validating global model outputs with respect to estimating height-resolved CCN concentrations over both land and ocean is necessary, as they form the key components of ACIs and, thus, the future climate predictions. With the height-resolved CCN data available for more than 15 years, a comprehensive model evaluation can be performed at different height levels, even for regions over oceans far from continents. In Sect. 4.1 of the paper, we compare the annual average n_{CCN} estimated from the monthly data with the simulations from 15 global climate models at the surface for the year 2011, highlighting the data's usefulness with respect to assessing model outputs. This idea can be further expanded to other models by incorporating the temporal and altitudinal distribution of CCN for different aerosol types encompassing a longer time span.

The above list is by no means exhaustive. Users can also use this data set and check for closure with other spaceborne retrievals of cloud-relevant CCN or aerosol number concentrations. For instance, comparisons can be made with (i) column-integrated aerosol number concentrations derived from polarimetric observations (Hasekamp et al., 2019), (ii) aerosol number and CCN concentrations estimated from spaceborne lidar using AERONET-derived extinction-to-number concentration conversions (Mamouri and Ansmann, 2016), (iii) CCN at cloud base derived from passive spaceborne sensors by treating clouds as CCN chambers (Rosenfeld et al., 2016), or with any future algorithms or data.

Appendix A

Table A1. Lognormal, bimodal volume size distribution parameters of different aerosol types. ν , μ , and σ represent the volume fraction, mode radius, and standard deviation, respectively. The subscripts “f” and “c” represent the fine and coarse modes of the size distribution, respectively.

| Aerosol type | Size distribution parameters | | | | | |
|----------------------|------------------------------|---------|---------|---------|------------|------------|
| | ν_f | ν_c | μ_f | μ_c | σ_f | σ_c |
| Marine | 0.14 | 0.86 | 0.1137 | 1.8756 | 1.6487 | 2.0544 |
| Dust | 0.223 | 0.777 | 0.1165 | 2.8329 | 1.4813 | 1.9078 |
| Polluted continental | 0.531 | 0.469 | 0.1577 | 3.547 | 1.5257 | 2.065 |
| Clean continental | 0.050 | 0.950 | 0.20556 | 2.6334 | 1.61 | 1.8987 |
| Smoke | 0.329 | 0.671 | 0.1436 | 3.726 | 1.5624 | 2.1426 |

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/essd-15-3747-2023-supplement>.

Author contributions. GC programmed the algorithm, processed the data, prepared the plots, and wrote the manuscript. MT contributed to improving the algorithm and the output data set as well as to revising the manuscript.

Competing interests. The contact author has declared that neither of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors acknowledge the CALIPSO science team for sharing the CALIPSO data. We are grateful to the AERIS/ICARE Data and Services Center for providing access to the CALIPSO data utilized in this work.

Financial support. This research has been supported by the Franco-German Fellowship Programme on Climate, Energy, and Earth System Research (Make Our Planet Great Again – German Research Initiative, MOPGA-GRI, grant no. 57429422) of the German Academic Exchange Service (DAAD), funded by the German Ministry of Education and Research.

Review statement. This paper was edited by Luis Millan and reviewed by two anonymous referees.

References

- Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, *Science*, 245, 1227–1230, <https://doi.org/10.1126/science.245.4923.1227>, 1989.
- Amiridis, V., Marinou, E., Tsekeri, A., Wandinger, U., Schwarz, A., Giannakaki, E., Mamouri, R., Kokkalis, P., Binietoglou, I., Solomos, S., Herekakis, T., Kazadzis, S., Gerasopoulos, E., Proestakis, E., Kottas, M., Balis, D., Papayannis, A., Kontoes, C., Kourtidis, K., Papagiannopoulos, N., Mona, L., Pappalardo, G., Le Rille, O., and Ansmann, A.: LIVAS: a 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET, *Atmos. Chem. Phys.*, 15, 7127–7153, <https://doi.org/10.5194/acp-15-7127-2015>, 2015.
- Andreae, M. O. and Rosenfeld, D.: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, *Earth-Sci. Rev.*, 89, 13–41, <https://doi.org/10.1016/j.earscirev.2008.03.001>, 2008.
- Aravindhavel, A., Choudhury, G., Prabhakaran, T., Murugavel, P., and Tesche, M.: Retrieval and validation of cloud condensation nuclei from satellite and airborne measurements

- over the Indian Monsoon region, *Atmos. Res.*, 290, 106802, <https://doi.org/10.1016/j.atmosres.2023.106802>, 2023.
- Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, *Rev. Geophys.*, 58, e2019RG000660, <https://doi.org/10.1029/2019RG000660>, 2020.
- Brock, C. A., Froyd, K. D., Dollner, M., Williamson, C. J., Schill, G., Murphy, D. M., Wagner, N. J., Kupc, A., Jimenez, J. L., Campuzano-Jost, P., Nault, B. A., Schroder, J. C., Day, D. A., Price, D. J., Weinzierl, B., Schwarz, J. P., Katich, J. M., Wang, S., Zeng, L., Weber, R., Dibb, J., Scheuer, E., Diskin, G. S., DiGangi, J. P., Bui, T., Dean-Day, J. M., Thompson, C. R., Peischl, J., Ryerson, T. B., Bourgeois, I., Daube, B. C., Commane, R., and Wofsy, S. C.: Ambient aerosol properties in the remote atmosphere from global-scale in situ measurements, *Atmos. Chem. Phys.*, 21, 15023–15063, <https://doi.org/10.5194/acp-21-15023-2021>, 2021.
- Chen, Y., Haywood, J., Wang, Y., Malavelle, F., Jordan, G., Partridge, D., Fieldsend, J., De Leeuw, J., Schmidt, A., Cho, N., Oreopoulos, L., Platnick, S., Grosvenor, D., Field, P., and Lohmann, U.: Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover, *Nat. Geosci.*, 15, 609–614, <https://doi.org/10.1038/s41561-022-00991-6>, 2022.
- Choudhury, G. and Tesche, M.: Estimating cloud condensation nuclei concentrations from CALIPSO lidar measurements, *Atmos. Meas. Tech.*, 15, 639–654, <https://doi.org/10.5194/amt-15-639-2022>, 2022a.
- Choudhury, G. and Tesche, M.: Assessment of CALIOP-Derived CCN Concentrations by In Situ Surface Measurements, *Remote Sensing*, 14, 3342, <https://doi.org/10.3390/rs14143342>, 2022b.
- Choudhury, G. and Tesche, M.: Global multiyear 3D dataset of cloud condensation nuclei derived from spaceborne lidar measurements, PANGAEA [data set], <https://doi.org/10.1594/PANGAEA.956215>, 2023.
- Choudhury, G., Ansmann, A., and Tesche, M.: Evaluation of aerosol number concentrations from CALIPSO with ATom airborne in situ measurements, *Atmos. Chem. Phys.*, 22, 7143–7161, <https://doi.org/10.5194/acp-22-7143-2022>, 2022.
- Deng, Z. Z., Zhao, C. S., Ma, N., Liu, P. F., Ran, L., Xu, W. Y., Chen, J., Liang, Z., Liang, S., Huang, M. Y., Ma, X. C., Zhang, Q., Quan, J. N., Yan, P., Henning, S., Mildenberger, K., Sommerhage, E., Schäfer, M., Stratmann, F., and Wiedensohler, A.: Size-resolved and bulk activation properties of aerosols in the North China Plain, *Atmos. Chem. Phys.*, 11, 3835–3846, <https://doi.org/10.5194/acp-11-3835-2011>, 2011.
- Dusek, U., Frank, G. P., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D., Drewnick, F., Hings, S., Jung, D., Borrmann, S., and Andreae, M. O.: Size Matters More Than Chemistry for Cloud-Nucleating Ability of Aerosol Particles, *Science*, 312, 1375–1378, <https://doi.org/10.1126/science.1125261>, 2006.
- Fanourgakis, G. S., Kanakidou, M., Nenes, A., Bauer, S. E., Bergman, T., Carslaw, K. S., Grini, A., Hamilton, D. S., Johnson, J. S., Karydis, V. A., Kirkevåg, A., Kodros, J. K., Lohmann, U., Luo, G., Makkonen, R., Matsui, H., Neubauer, D., Pierce, J. R., Schmale, J., Stier, P., Tsigaridis, K., van Noije, T., Wang, H., Watson-Parris, D., Westervelt, D. M., Yang, Y., Yoshioka, M., Daskalakis, N., Decesari, S., Gysel-Beer, M., Kalivitis, N., Liu, X., Mahowald, N. M., Myriokefalitakis, S., Schrödner, R., Sfakianaki, M., Tsimpidi, A. P., Wu, M., and Yu, F.: Evaluation of global simulations of aerosol particle and cloud condensation nuclei number, with implications for cloud droplet formation, *Atmos. Chem. Phys.*, 19, 8591–8617, <https://doi.org/10.5194/acp-19-8591-2019>, 2019a.
- Fanourgakis, G., Kanakidou, M., Nenes, A., et al.: Data for the “Evaluation of global simulations of aerosol particle and cloud condensation nuclei number, with implications for cloud droplet formation” (Version v1), Zenodo [data set], <https://doi.org/10.5281/zenodo.3265866>, 2019b.
- Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth’s energy budget, climate feedbacks, and climate sensitivity, in: *Climate Change 2021: The Physical Science Basis, Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirant, A., Connors, S. L., Pean, C., Berger, S., Caud, C., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, L. B. R., Maycock, T. K., Waterfield, T., Yelekci, O., Yu, R., and Zhou, B., Cambridge University Press, 923–1054, <https://www.ipcc.ch/report/ar6/wg1/> (last access: 11 August 2023), 2021.
- Gasteiger, J. and Wiegner, M.: MOPSMAP v1.0: a versatile tool for the modeling of aerosol optical properties, *Geoscientific Model Development*, 11, 2739–2762, <https://doi.org/10.5194/gmd-11-2739-2018>, 2018.
- Hasekamp, O. P., Gryspeerdt, E., and Quaas, J.: Analysis of polarimetric satellite measurements suggests stronger cooling due to aerosol-cloud interactions, *Nat. Commun.*, 10, 5405, <https://doi.org/10.1038/s41467-019-13372-2>, 2019.
- Humphries, R. S., Keywood, M. D., Gribben, S., McRobert, I. M., Ward, J. P., Selleck, P., Taylor, S., Harnwell, J., Flynn, C., Kulkarni, G. R., Mace, G. G., Protat, A., Alexander, S. P., and McFarquhar, G.: Southern Ocean latitudinal gradients of cloud condensation nuclei, *Atmos. Chem. Phys.*, 21, 12757–12782, <https://doi.org/10.5194/acp-21-12757-2021>, 2021.
- Kim, M.-H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu, Y., Liu, Z., Poole, L. R., Pitts, M. C., Kar, J., and Magill, B. E.: The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm, *Atmos. Meas. Tech.*, 11, 6107–6135, <https://doi.org/10.5194/amt-11-6107-2018>, 2018.
- Koehler, K. A., Kreidenweis, S. M., DeMott, P. J., Petters, M. D., Prenni, A. J., and Carrico, C. M.: Hygroscopicity and cloud droplet activation of mineral dust aerosol, *Geophys. Res. Lett.*, 36, L08805, <https://doi.org/10.1029/2009GL037348>, 2009.
- Kumar, P., Sokolik, I. N., and Nenes, A.: Cloud condensation nuclei activity and droplet activation kinetics of wet processed regional dust samples and minerals, *Atmos. Chem. Phys.*, 11, 8661–8676, <https://doi.org/10.5194/acp-11-8661-2011>, 2011.
- Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C., and Hostetler, C.: The CALIPSO Lidar Cloud and Aerosol Dis-

- crimination: Version 2 Algorithm and Initial Assessment of Performance, *J. Atmos. Ocean. Tech.*, 26, 1198–1213, <https://doi.org/10.1175/2009JTECHA1229.1>, 2009.
- Mamouri, R. E. and Ansmann, A.: Estimated desert-dust ice nuclei profiles from polarization lidar: methodology and case studies, *Atmos. Chem. Phys.*, 15, 3463–3477, <https://doi.org/10.5194/acp-15-3463-2015>, 2015.
- Mamouri, R.-E. and Ansmann, A.: Potential of polarization lidar to provide profiles of CCN- and INP-relevant aerosol parameters, *Atmos. Chem. Phys.*, 16, 5905–5931, <https://doi.org/10.5194/acp-16-5905-2016>, 2016.
- Mao, F., Shi, R., Rosenfeld, D., Pan, Z., Zang, L., Zhu, Y., and Lu, X.: Retrieving instantaneous extinction of aerosol undetected by the CALIPSO layer detection algorithm, *Atmos. Chem. Phys.*, 22, 10589–10602, <https://doi.org/10.5194/acp-22-10589-2022>, 2022.
- Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, *Geosci. Model Dev.*, 8, 1339–1356, <https://doi.org/10.5194/gmd-8-1339-2015>, 2015.
- Myhre, G., Grini, A., Haywood, J. M., Stordal, F., Chatenet, B., Tanré, D., Sundet, J. K., and Isaksen, I. S. A.: Modeling the radiative impact of mineral dust during the Saharan Dust Experiment (SHADE) campaign, *J. Geophys. Res.-Atmos.*, 108, 8579, <https://doi.org/10.1029/2002JD002566>, 2003.
- NASA/LARC/SD/ASDC: CALIPSO Lidar Level 2 Aerosol Profile, V4-20, NASA [data set], https://doi.org/10.5067/CALIP/CALIPSO/LID_L2_05KMAPRO-STANDARD-V4-20, 2018.
- Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kitaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmos. Ocean. Tech.*, 26, 1994–2014, <https://doi.org/10.1175/2009JTECHA1231.1>, 2009.
- Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, *Atmos. Chem. Phys.*, 7, 1961–1971, <https://doi.org/10.5194/acp-7-1961-2007>, 2007.
- Quaas, J., Arola, A., Cairns, B., Christensen, M., Deneke, H., Ekman, A. M. L., Feingold, G., Fridlind, A., Gryspeerd, E., Hasekamp, O., Li, Z., Lipponen, A., Ma, P.-L., Mülmenstädt, J., Nenes, A., Penner, J. E., Rosenfeld, D., Schrödner, R., Sinclair, K., Sourdeval, O., Stier, P., Tesche, M., van Diedenhoven, B., and Wendisch, M.: Constraining the Twomey effect from satellite observations: issues and perspectives, *Atmos. Chem. Phys.*, 20, 15079–15099, <https://doi.org/10.5194/acp-20-15079-2020>, 2020.
- Rose, D., Nowak, A., Achtert, P., Wiedensohler, A., Hu, M., Shao, M., Zhang, Y., Andreae, M. O., and Pöschl, U.: Cloud condensation nuclei in polluted air and biomass burning smoke near the mega-city Guangzhou, China – Part 1: Size-resolved measurements and implications for the modeling of aerosol particle hygroscopicity and CCN activity, *Atmos. Chem. Phys.*, 10, 3365–3383, <https://doi.org/10.5194/acp-10-3365-2010>, 2010.
- Rosenfeld, D., Zheng, Y., Hashimshoni, E., Pöhlker, M. L., Jefferson, A., Pöhlker, C., Yu, X., Zhu, Y., Liu, G., Yue, Z., Fischman, B., Li, Z., Giguzin, D., Goren, T., Artaxo, P., Barbosa, H. M. J., Pöschl, U., and Andreae, M. O.: Satellite retrieval of cloud condensation nuclei concentrations by using clouds as CCN chambers, *P. Natl. Acad. Sci. USA*, 113, 5828–5834, <https://doi.org/10.1073/pnas.1514044113>, 2016.
- Sayer, A. M., Smirnov, A., Hsu, N. C., and Holben, B. N.: A pure marine aerosol model, for use in remote sensing applications, *J. Geophys. Res.-Atmos.*, 117, D05213, <https://doi.org/10.1029/2011JD016689>, 2012.
- Schmale, J., Henning, S., Decesari, S., Henzing, B., Keskinen, H., Sellegri, K., Ovadnevaite, J., Pöhlker, M. L., Brito, J., Bougiatioti, A., Kristensson, A., Kalivitis, N., Stavroulas, I., Carbone, S., Jefferson, A., Park, M., Schlag, P., Iwamoto, Y., Aalto, P., Äijälä, M., Bukowiecki, N., Ehn, M., Frank, G., Fröhlich, R., Frumau, A., Herrmann, E., Herrmann, H., Holzinger, R., Kos, G., Kulmala, M., Mihalopoulos, N., Nenes, A., O’Dowd, C., Petäjä, T., Picard, D., Pöhlker, C., Pöschl, U., Poulain, L., Prévôt, A. S. H., Swietlicki, E., Andreae, M. O., Artaxo, P., Wiedensohler, A., Ogren, J., Matsuki, A., Yum, S. S., Stratmann, F., Baltensperger, U., and Gysel, M.: Long-term cloud condensation nuclei number concentration, particle number size distribution and chemical composition measurements at regionally representative observatories, *Atmos. Chem. Phys.*, 18, 2853–2881, <https://doi.org/10.5194/acp-18-2853-2018>, 2018.
- Seinfeld, J. H., Bretherton, C., Carslaw, K. S., Coe, H., DeMott, P. J., Dunlea, E. J., Feingold, G., Ghan, S., Guenther, A. B., Kahn, R., Kraucunas, I., Kreidenweis, S. M., Molina, M. J., Nenes, A., Penner, J. E., Prather, K. A., Ramanathan, V., Ramaswamy, V., Rasch, P. J., Ravishankara, A. R., Rosenfeld, D., Stephens, G., and Wood, R.: Improving our fundamental understanding of the role of aerosol–cloud interactions in the climate system, *P. Natl. Acad. Sci. USA*, 113, 5781–5790, <https://doi.org/10.1073/pnas.1514043113>, 2016.
- Shinozuka, Y., Clarke, A. D., Nenes, A., Jefferson, A., Wood, R., McNaughton, C. S., Ström, J., Tunved, P., Redemann, J., Thornhill, K. L., Moore, R. H., Latham, T. L., Lin, J. J., and Yoon, Y. J.: The relationship between cloud condensation nuclei (CCN) concentration and light extinction of dried particles: indications of underlying aerosol processes and implications for satellite-based CCN estimates, *Atmos. Chem. Phys.*, 15, 7585–7604, <https://doi.org/10.5194/acp-15-7585-2015>, 2015.
- Tackett, J. L., Winker, D. M., Getzewich, B. J., Vaughan, M. A., Young, S. A., and Kar, J.: CALIPSO lidar level 3 aerosol profile product: version 3 algorithm design, *Atmos. Meas. Tech.*, 11, 4129–4152, <https://doi.org/10.5194/amt-11-4129-2018>, 2018.
- Tesche, M., Ansmann, A., Müller, D., Althausen, D., Engelmann, R., Freudenthaler, V., and Groß, S.: Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008, *J. Geophys. Res.-Atmos.*, 114, D13202, <https://doi.org/10.1029/2009JD011862>, 2009.
- Twomey, S.: Pollution and the planetary albedo, *Atmos. Environ.*, 8, 1251–1256, [https://doi.org/10.1016/0004-6981\(74\)90004-3](https://doi.org/10.1016/0004-6981(74)90004-3), 1974.
- van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, *Earth Syst. Sci. Data*, 9, 697–720, <https://doi.org/10.5194/essd-9-697-2017>, 2017.
- Wall, C. J., Norris, J. R., Possner, A., McCoy, D. T., McCoy, I. L., and Lutsko, N. J.: Assessing effective ra-

- diative forcing from aerosol–cloud interactions over the global ocean, *P. Natl. Acad. Sci. USA*, 119, e2210481119, <https://doi.org/10.1073/pnas.2210481119>, 2022.
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, *J. Atmos. Ocean. Tech.*, 26, 2310–2323, <https://doi.org/10.1175/2009JTECHA1281.1>, 2009.
- Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A., and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, *Atmos. Chem. Phys.*, 13, 3345–3361, <https://doi.org/10.5194/acp-13-3345-2013>, 2013.
- Young, S. A., Vaughan, M. A., Kuehn, R. E., and Winker, D. M.: The Retrieval of Profiles of Particulate Extinction from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Data: Uncertainty and Error Sensitivity Analyses, *J. Atmos. Ocean. Tech.*, 30, 395–428, <https://doi.org/10.1175/JTECH-D-12-00046.1>, 2013.
- Zang, L., Rosenfeld, D., Mao, F., Pan, Z., Zhu, Y., Gong, W., and Wang, Z.: CALIOP retrieval of droplet effective radius accounting for cloud vertical homogeneity, *Opt. Express*, 29, 21921–21935, <https://doi.org/10.1364/OE.427022>, 2021.
- Zhang, R., Wang, Y., Li, Z., Wang, Z., Dickerson, R. R., Ren, X., He, H., Wang, F., Gao, Y., Chen, X., Xu, J., Cheng, Y., and Su, H.: Vertical profiles of cloud condensation nuclei number concentration and its empirical estimate from aerosol optical properties over the North China Plain, *Atmos. Chem. Phys.*, 22, 14879–14891, <https://doi.org/10.5194/acp-22-14879-2022>, 2022.
- Zhang, Y., Zhao, C., Zhang, K., Ke, J., Che, H., Shen, X., Zheng, Z., and Liu, D.: Retrieving the microphysical properties of opaque liquid water clouds from CALIOP measurements, *Opt. Express*, 27, 34126–34140, <https://doi.org/10.1364/OE.27.034126>, 2019.
- Zieger, P., Fierz-Schmidhauser, R., Weingartner, E., and Baltensperger, U.: Effects of relative humidity on aerosol light scattering: results from different European sites, *Atmos. Chem. Phys.*, 13, 10609–10631, <https://doi.org/10.5194/acp-13-10609-2013>, 2013.