Geosci. Model Dev., 19, 10571–10593, 2026
https://doi.org/10.5194/gmd-19-10571-2026
© Author(s) 2026. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development
Open Access
EGU

Evaluation and improvement of CAMS-derived CCN number concentrations using in-situ measurements

Yannick Emanuel Anders[1], **Karoline Block**[1], **Mira Pöhlker**[1,2], **and Johannes Quaas**[1]

[1]Leipzig Institute for Meteorology, Faculty of Physics and Earth System Sciences,
Leipzig University, 04103 Leipzig, Germany
[2]Department of Atmospheric Microphysics, Leibniz Institute for Tropospheric Research (TROPOS), 04318 Leipzig, Germany

Correspondence: Karoline Block (karoline.block@uni-leipzig.de)

Received: 13 March 2026 – Discussion started: 30 March 2026
Revised: 23 June 2026 – Accepted: 7 October 2026 – Published: 9 November 2026

Abstract. Cloud condensation nuclei (CCN) are key components of aerosol-cloud interactions (ACI). Thus, a precise knowledge about their number concentrations (Nccn) is crucial for climate models and ACI studies. This study presents a comprehensive evaluation of the recently published CAMS-derived total Nccn using direct observations from 25 ground-based sites. The evaluation focuses on the representation of temporal variability, the applicability of CAMS-derived Nccn across different environments and pollution regimes and in particular, the sensitivity of CCN to supersaturation. At a mid-range supersaturation of $s = 0.4\,\%$ the CAMS-derived Nccn reproduces monthly and seasonal variability well, with monthly median Pearson correlation coefficients ranging from +0.558 to +0.911, while the representation of day-to-day variability is considerably weaker, with significant daily correlations ranging from +0.121 to +0.818. No systematic dependence of model performance on environment or pollution regime is found for this level of supersaturation. The dataset therefore appears well suited for climatological applications and studies focusing on variability beyond synoptic timescales. The analysis further reveals a bias shift in simulated Nccn with increasing supersaturation. This feature correlates to the ratio of the two dominant CCN species, likely reflecting assumptions in the underlying size distributions and/or emissions fractions. To address this issue, we developed an observation-based parameterization using a modified power law with an s-dependent k spectrum that modifies the supersaturation dependence of total Nccn without altering aerosol size distributions or species concentrations. This is applied to modeled Nccn which are grouped by environment and pollution regimes yielding different levels of best fitting supersaturations. Overall, this correction substantially reduces the median bias by 30.1 % at 0.1 % supersaturation and by 43.3 % at 1.0 % supersaturation, improving agreement of the modeled Nccn with observations at most sites. These results further enhance the applicability of the dataset for global aerosol–cloud interaction studies, leading the way to an improved version of CAMS-derived Nccn.

1 Introduction

The simulation of aerosol-cloud interactions (ACI), still a remaining source of uncertainty in climate model intercomparisons (IPCC, 2023), relies on accurate fields of cloud condensation nuclei (CCN) number concentrations (Nccn) (e.g. Seinfeld et al., 2016). Since direct observations of CCN are only possible through in-situ measurements, from ground-based facilities, air- and shipborne platforms, the spatial and temporal distribution of observed Nccn is still sparse. For that reason, most climate models and ACI studies incorporate aerosol optical properties, e.g. aerosol optical depth (AOD), as a proxy for Nccn (e.g. Quaas et al., 2020; Gryspeerdt and Stier, 2012; Bellouin et al., 2020). Derived from satellite measurements, AOD data can potentially be provided on a global domain, continuously over long time periods. However, the use of AOD as an estimate of Nccn comes with disadvantages (e.g. Stier, 2016; Jia et al., 2021): As a column-integrated optical variable, it can neither account for the vertical variability of CCN nor the variety of aerosol species that can act as CCN, depending on their size and hygroscopicity. Also, other aerosol processes, such as swelling, might result

in changes in the AOD, which are not caused by changes in the number of CCN (e.g. Quaas et al., 2009). Challenges in the satellite retrieval above cloudy or other bright surfaces reduce the global presentability of AOD as a CCN proxy. For example, Stier (2016) stated that the variability in AOD explains only 25 % of global CCN variance.

Thus, other approaches for the determination of globally resolved Nccns have been proposed. One of them is the recently published dataset by Block (2023), where Nccn are derived from aerosol mass mixing ratios from the Copernicus Atmosphere Monitoring Service (CAMS) aerosol reanalysis (Inness et al., 2019). The CAMS-derived CCN dataset contains globally and vertically resolved Nccn for 6 supersaturations ranging from 0.1 % to 1 %.

This dataset has been used in several studies but has received little systematic evaluation. In a first very brief evaluation, Block et al. (2024) compared CAMS Nccn and ground-based Nccn observations from five measurement sites at a single supersaturation of 0.4 %. Considering all data points, they found an improved overall correlation coefficient of 0.71, compared to the use of a Moderate Resolution Imaging Spectroradiometer (MODIS; Levy et al., 2013) AOD product as a CCN proxy, which yielded an overall correlation coefficient of 0.37. They have also found that observations are overestimated in their median and interquartile range for four out of five stations at this supersaturation. Notably, the largest bias has been obtained for the station with the highest CCN concentration. Whether this behavior is common to heavily polluted sites in general, has not been confirmed yet, as other sites with a similar level of pollution have not been validated yet.

A recent study by Kulkarni et al. (2025) examined the vertical distribution of CAMS-derived Nccn and compared it to airborne Nccn measurements that have been performed over the Southern Great Plains. For both sampling periods in spring and summer, they found that CAMS-derived Nccn tends to underestimate the observations. The r^2 coefficients of determination were 0.54 and 0.31, respectively for the two periods. The underestimation was found to be strongest at higher Nccn ($\geq 1400\,\mathrm{cm}^{-3}$), where the authors obtained an underprediction by a factor of 8. Also in comparison to other CCN simulating methods, CAMS-derived Nccn exhibit the lowest concentrations in the scope of this specific study.

A similar result has been found by Choudhury et al. (2025), but for data points below an altitude of 2 km and only for a single supersaturation of 0.2 %. They compared the spatio-temporal variability of CAMS-derived Nccn to another novel global CCN dataset, which has been derived from Cloud Aerosol Lidar with Orthogonal Polarization (CALIOP) measurements (Choudhury and Tesche, 2023). Overall, CALIOP-derived Nccn are 79 % higher than CAMS-derived Nccn. Their analysis revealed that the largest discrepancies were found over the oceans, especially in the Southern Hemisphere (SH), while over the Northern Hemisphere (NH), the differences between the medians of CALIOP and CAMS are smaller. For all oceanic domains, except North Atlantic, both datasets furthermore disagree on the representation of the respective seasonal cycle, while the patterns in continental regions are mostly consistent. They have also found that inter-hemispheric differences are larger in CAMS than in CALIOP.

All of the mentioned studies only focused on one single supersaturation each: 0.2 % (Choudhury et al., 2025), 0.24 % (Kulkarni et al., 2025) and 0.4 % (Block et al., 2024). In contrast, Anders et al. (2025) made use of all 6 available supersaturations in their comparison of CAMS-derived Nccn to 13 measurement sites in both continental and marine environment. For most sites they found a supersaturation between 0.4 %–0.8 % showing the best agreement between CAMS-derived CCN and observations. At some stations, higher supersaturations are less biased, while at others, low supersaturations perform better, which confirms the regional inhomogeneity that was found by Choudhury et al. (2025). Furthermore, Anders et al. (2025) revealed that CAMS-derived Nccn tend to underestimate at low supersaturations while it tends to overestimate at high supersaturations.

The bias shift identified in that study motivated the present work, which provides a substantially more extensive analysis of the CCN-supersaturation relationship in the CAMS-derived Nccn dataset, including and evaluation of its performance against an extended set of ground-based observational sites. Here we perform a systematic analysis on the variability of CAMS-derived total Nccn in time, its suitability in different environments and pollution regimes and in particular, its sensitivity to supersaturation.

2 Methodology

2.1 The CAMS-derived Nccn dataset

The Nccn dataset (Block, 2023) that is being evaluated in this study was derived from the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis (Inness et al., 2019), which is produced using the Integrated Forcasting System (IFS; Morcrette et al., 2009; Rémy et al., 2019) of the European Centre for Medium-Range Weather Forecast (ECMWF) and its data assimilation scheme (Benedetti et al., 2009). The aerosol mass mixing ratios that are simulated by the IFS are constrained with aerosol optical depth (AOD) retrievals from satellite data. Incorporated are the dark target product and deep-blue product of MODIS as well as AOD data retrieved from the Advanced Along-Track Scanning Radiometer (AATSR; Popp et al., 2016). It has to be pointed out that the assimilation is limited to latitudes between 70° S–70° N, due to the low solar illumination in polar regions (Inness et al., 2019). From these constrained aerosol mass mixing ratios, Block et al. (2024) have derived total and species-related Nccn using a simplified kappa-Köhler approach, assuming fixed hygroscopicity (kappa) values for each CCN-

relevant aerosol species and globally fixed log-normal size distributions for each aerosol component of the external mixture. As a result, total Nccn are estimated for supersaturations of $s = 0.1\,\%$, $0.2\,\%$, $0.4\,\%$, $0.6\,\%$, $0.8\,\%$, and $1.0\,\%$, while species-related Nccn, which are sulfate (SU), sea salt (SS) as well as hydrophilic fractions of black carbon (BC) and organic matter (OM), are available for $0.2\,\%$ and $0.8\,\%$. Mineral dust aerosols are excluded entirely, following the argument that they do not contribute notably to Nccn as they are large in size and have a low hygroscopicity (Che et al., 2022). Wildfires and biomass-burning events are well captured by CAMS and have large impacts on total Nccn via BC and OM, whereas aerosols related to volcanic eruptions, nitrate aerosols or stratospheric aerosols are not represented at all (Block et al., 2024). Currently, the publicly available dataset contains daily Nccn from January 2003 to December 2024 on a $0.75° \times 0.75°$ grid with 60 hybrid sigma-pressure levels from surface to top of atmosphere (Block, 2023). In the scope of the present study, only the years 2003–2021 have been used.

2.2 Observational data

The ground-based data that is used in this study to validate CAMS Nccn is provided by 25 measurement sites (Fig. 1). These stations were chosen because they measure Nccn for at least 4 different supersaturations. The full station names, the provided time periods and available supersaturations are listed in Table 1. References to all observational datasets are included in the Appendix, Table A1. As can be seen from Fig. 1, available observational sites are concentrated in Europe and North America, while Africa, Australasia and large parts of central and northern Asia remain particularly underrepresented. Furthermore, effectively only two sites (ENA/-GRW, ASI) are available to represent remote oceanic environments. This uneven spatial coverage can introduce distortions when observations from different sites are combined to derive global estimates. Therefore, oceanic and continental contributions are weighted separately in subsequent analyses, as described in Sect. 3.2.2.

7 of these 25 stations are part of the European Aerosols, Clouds, and Trace gases Research InfraStructure (ACTRIS, http://www.actris.eu/, last access: 23 June 2026): Cabauw (CES), Finokalia (FIK), Jungfraujoch (JFJ), Melpitz (MEL), Mace Head (MHD), Hyytiälä (SMR) and Vavihill (VAV). Information about these stations as well as about the stations of Seoul (SEO) in South Korea, Noto (NOT) in Japan, Utqiaġvik (formerly Barrow, BRW) in USA and Amazon Tall Tower Observatory (ATT; Pöhlker et al., 2016) in Brazil are given in Schmale et al. (2017, 2018). At each of these 11 stations, a single-column CCN counter (model CCN-100) by Droplet Measurement Technologies (DMT; Roberts and Nenes, 2005; Rose et al., 2008) was used.

The remaining 14 datasets from Ascension Island (ASI), McMurdo Station (Antarctica, AWR), Córdoba (COR), Graciosa Island (ENA and GRW, covering two time periods), Black Forest (FKB), Manacapuru (MAO), Niamey (NIM), North Slope of Alaska (NSA, same location as BRW), Oliktok Point (OLI), Nainital (PGH), Cape Cod (PVC), Point Reyes (PYE) and Southern Great Plains (SGP) are provided by the United States' Department of Energy's Atmospheric Radiation Measurement (ARM, https://www.arm.gov, last access: 23 June 2026) user facility. Either single- or dual-column DMT CCN counters have been running at the ARM measurement sites and various data products were available at the time of this analysis. With the different data products, also varying levels of data quality checks were provided. The published data quality reports for each station (also see https://www.arm.gov/data/data-quality-program, last access: 23 June 2026) have been used to remove suspect, incorrect and missing data. Furthermore, the following criteria have been applied to all datasets in the process of data checking: There must not be more than one single file per day of measurement. This file has to contain a valid time dimension and no negative Nccn or supersaturation values. A list of all removed dates is given in Table A1 in the Appendix. The measurements from ENA and GRW (Graciosa Island), as well as those from BRW and NSA (north of Alaska), were taken at the same ARM facilities.

At SGP, four different data products are available for overlapping time periods (Fig. A1), which we combined to one time series. Between two data products, which are available for the same period, we have always selected that one with the higher reported level of data quality. Thus, the AOSCCN (Aerosol Observing System Cloud Condensation Nuclei) product covers the period from 13 September 2006 to 31 December 2010, AOSCCNAVG (AOSCCN Average) from 1 January 2011 to 31 October 2012, AOSCCN1COLAVG (AOSCCN Single Column Average) from 1 November 2012 to 16 August 2017 and AOSCCN2COLAAVG (AOSCCN Dual Column Average) from 17 August 2017 to 31 December 2021.

Following Spracklen et al. (2011) who assessed an uncertainty range in observed Nccn of 5 %–40 % depending on the magnitude of concentrations, supersaturation and the type of CCN instrument used, we assume a maximum relative uncertainty of $\pm 40\,\%$ and a minimum absolute uncertainty of $\pm 20\,\mathrm{cm}^{-3}$ in observed Nccn, as was done in Block et al. (2024).

2.3 Data treatment

Please note that in the remainder of this study, $\mathrm{CCN_C}$ refer to Nccn derived from CAMS and $\mathrm{CCN_O}$ are observed Nccn. Supersaturation is abbreviated as s. In order to validate $\mathrm{CCN_C}$ with observed data, we processed and homogenized both datasets. The model height level is chosen such that the geopotential height of the nearest grid point to the measurement site matches the measurement altitude approximately. To reduce representativeness errors associated with

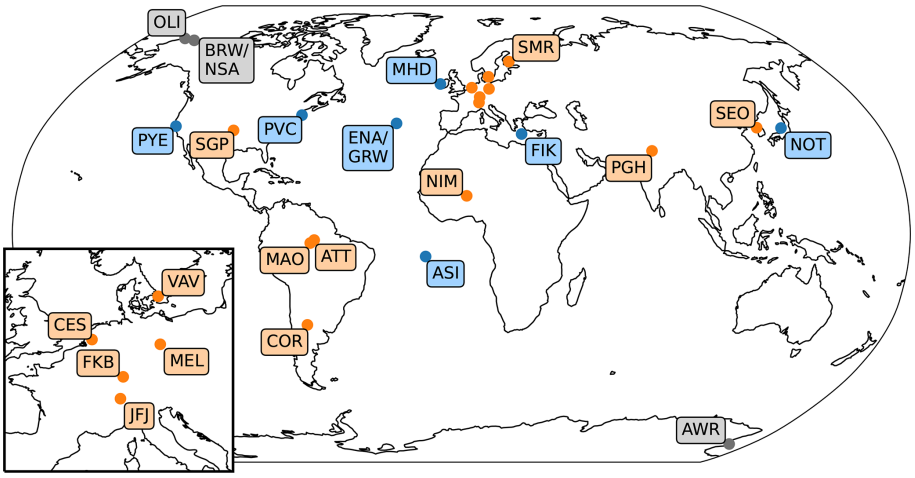


Figure 1. Locations of ground-based measurement sites. BRW and NSA as well as ENA and GRW share the same location, respectively. Continental stations in orange, marine in blue and polar stations in grey. The full station names, provided time periods and available supersaturations of each measurement site are listed in Table 1.

comparing point-based observations to gridded model output, a horizontal mean of the nine nearest CAMS grid points surrounding each measurement site was used. This spatial averaging mitigates the influence of sub-grid variability and small spatial displacement errors. Using only the nearest grid point resulted in systematically lower correlation coefficients (not shown).

As a last step of preparation, for some stations the CAMS-derived Nccn are normalized from ambient temperature and pressure (ATP) to standard temperature and pressure (STP). This is done by using the ideal gas law for standard conditions ($T_0 = 273.15\,\mathrm{K}$, $p_0 = 1013\,\mathrm{hPa}$), as e.g. in Holanda et al. (2020) and Andrews et al. (2025):

$$\mathrm{Nccn_{STP}} = \mathrm{Nccn_{ATP}} \left(\frac{T_0}{T} \frac{p}{p_0} \right). \tag{1}$$

Observational data of BRW, CES, FIK, JFJ, MEL, MHD, NOT, SEO, SMR, and VAV are already given for STP conditions, while ATT data is normalized to STP using Eq. (1). Observed Nccn from ASI, AWR, COR, ENA, FKB, GRW, MAO, NIM, NSA, OLI, PGH, PVC, PYE, SGP are given for ATP conditions. Accordingly, CAMS-derived Nccn are compared to observations under the same conditions.

In the files from the ARM sites, two types of supersaturations are given: One setpoint supersaturation and one calculated supersaturation, which can differ significantly, as described in Andrews et al. (2025) and Lance et al. (2006). Here, the calculated supersaturation was chosen as it resembles the actual supersaturation in the flow chamber.

Since CAMS-derived Nccn are available only once per day at 00:00 UTC, we sub-sampled the corresponding observed data and computed medians and interquartile ranges of the hourly measured Nccn from 22:00 UTC to 02:00 UTC. We will refer to these medians as 00:00 UTC daily medians.

As mentioned, $\mathrm{CCN_C}$ are given at 6 specific supersaturations (s = 0.1 %, 0.2 %, 0.4 %, 0.6 %, 0.8 %, and 1.0 %). Observed Nccn are usually not measured at the same supersaturations. Therefore, $\mathrm{CCN_O}$ are directly taken (or averaged if several exist) within a range of ±0.05 %. In case there is no data in the ±0.05 % range, we interpolated linearly in the narrow range of ±0.1 %. The resulting uncertainties from this interpolation are small in comparison to the measurement uncertainties. However, not all supersaturations are covered at each measurement site, not even in a ±0.1 % range. To be able to interpolate across a wider range of supersaturations, we preferred a power law to a linear function. Nevertheless, we limit the inter-/extrapolation interval $[s_1, s_2]$ to a maximum range of ±0.4 % around the desired value s_x, to reduce the inaccuracies, which are introduced by using a constant slope k across a range of supersaturations (see discussion in Sect. 2.4).

Thus, an Nccn value that is missing at a supersaturation s_x, but available at s_1 and s_2, is estimated by applying a power law to a given Nccn value at $s_i \in [s_1, s_2]$:

$$\mathrm{Nccn}(s_x) = \mathrm{Nccn}(s_i) \left(\frac{s_x}{s_i} \right)^k, \tag{2}$$

with the exponent k given from:

$$k = \frac{\ln\left(\frac{\mathrm{Nccn}(s_1)}{\mathrm{Nccn}(s_2)}\right)}{\ln\left(\frac{s_1}{s_2}\right)}, \tag{3}$$

as it was derived e.g. in Block et al. (2024) and Désalmand (1985). Equation (2) is a generalized form of the power law proposed by Twomey (1959), which has been widely used for the approximation of CCN activation spectra (e.g., Martins et al., 2009; Andreae, 2009; Hudson, 1980; Cohard et al.,

Table 1. Information on ground-based measurement sites that are used in this study. The given time frame indicates year and month of the first and last data point of the respective dataset of observations. The table contains a list of the supersaturations that are present in the given data files of each observation. For ARM stations, supersaturation setpoints are given. For SGP, four lists of supersaturations are provided, one for each available data product (AOSCCN, AOSCCNAVG, AOSCCN1COLAVG, AOSCCN2COLAAVG).

ID	Station name	Time frame	Available supersaturations [%]
ASI	Ascension Island (UK Overseas)	April 2016–October 2017	0.10, 0.20, 0.40, 0.60, 0.80, 1.00
ATT	Amazon Tall Tower Observatory, Brazil	March 2014–February 2015	0.11, 0.15, 0.20, 0.24, 0.29, 0.47, 0.61, 0.74, 0.92, 1.10
AWR	McMurdo Station, Antarctica	May 2016–December 2016	0.10, 0.20, 0.27, 0.37, 0.50, 0.61, 0.80, 0.92, 1.00
BRW	Utqiaġvik (formerly Barrow), Alaska US	July 2007–June 2008	0.20, 0.30, 0.50, 0.60, 1.00, 1.20, 1.45
CES	Cabauw, Netherlands	January 2012–December 2014	0.10, 0.20, 0.30, 0.50, 1.00
COR	Córdoba, Argentina	September 2018–April 2019	0.10, 0.20, 0.40, 0.60, 0.80, 1.00
ENA	Graciosa Island, Azores (Portugal)	January 2015–October 2020	0.10, 0.20, 0.50, 0.80, 1.00
FIK	Finokalia, Greece	January 2014–December 2015	0.20, 0.40, 0.60, 0.80, 1.00
FKB	Black Forest, Germany	March 2007–January 2008	0.08, 0.20, 0.34, 0.49, 0.78, 1.07, 1.37
GRW	Graciosa Island, Azores (Portugal)	May 2009–December 2010	0.10, 0.33, 0.55, 0.78, 1.00, 1.12
JFJ	Jungfraujoch, Switzerland	January 2012–December 2014	0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50, 0.70, 1.00
MAO	Manacapuru, Brazil	July 2014–September 2014	0.15, 0.25, 0.40, 0.60, 0.80, 1.10
MEL	Melpitz, Germany	January 2012–December 2014	0.10, 0.20, 0.30, 0.50, 0.70
MHD	Mace Head, Ireland	January 2011–December 2012	0.10, 0.25, 0.35, 0.50, 0.75, 1.00
NIM	Niamey, Niger	December 2005–January 2007	0.20, 0.40, 0.60, 0.80, 1.00
NOT	Noto, Japan	May 2014–March 2015	0.10, 0.20, 0.50, 0.80
NSA	Northern Slope of Alaska, US	March 2007–August 2011	0.10, 0.23, 0.40, 0.56, 0.89, 1.23, 1.56
OLI	Oliktok Point, Alaska (US)	April 2017–September 2019	0.10, 0.20, 0.40, 0.60, 0.80, 1.00
PGH	Nainital, Uttarakhand (India)	June 2011–March 2012	0.06, 0.17, 0.31, 0.46, 0.60, 0.75
PVC	Cape Cod, Massachusetts (US)	July 2012–March 2013	0.15, 0.18, 0.22, 0.29, 0.36, 0.46, 0.55, 0.66, 0.74, 0.86, 0.94, 1.00
PYE	Point Reyes, California (US)	March 2005–September 2005	0.30, 0.50, 0.70, 0.90
SEO	Seoul, South Korea	January 2006–December 2010	0.20, 0.40, 0.60, 0.80
SGP	Southern Great Plains, Oklahoma (US)	September 2006–December 2021	0.15, 0.20, 0.40, 0.60, 0.80, 1.00, 1.15; 0.22, 0.27, 0.48, 0.68, 0.89, 1.10, 1.25; 0.00, 0.10, 0.20, 0.40, 0.80, 1.00; 0.08, 0.19, 0.34, 0.49, 0.79, 1.08, 1.38
SMR	Hyytiälä (SMEAR II), Finland	January 2012–December 2014	0.10, 0.20, 0.30, 0.50, 1.00
VAV	Vavihill, Sweden	December 2012–November 2014	0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50, 0.70, 1.00, 1.40

1998). Commonly, s_i is assumed to be 1.0 % and the power law is written as

$$\mathrm{Nccn}(s_x) = \mathrm{Nccn}\,(1\,\%) s_x^k\,, \tag{4}$$

with a constant slope parameter k, that is independent of the desired supersaturation s_x.

2.4 Limitations to Twomey's power law

Several experimental and theoretical studies have pointed out that the simplicity of Twomey's power law comes with its limitations (e.g. Pöhlker et al., 2016; Deng et al., 2013; Ji and Shaw, 1998; Khvorostyanov and Curry, 2006; Cohard et al., 1998). Most importantly, it does not account for an aerosol-limited regime (e.g. Reutter et al., 2009) at high supersaturations, potentially leading to overestimated Nccn at high supersaturations. One reason is that the slope parameter k in Eq. (4) is assumed to be constant over the entire range of supersaturations; an assumption that has been contradicted by multiple observational studies (e.g. Kim et al., 2011; Khvorostyanov and Curry, 2006; Yum and Hudson, 2001). Instead, k has been shown to decrease with increasing supersaturation (Khvorostyanov and Curry, 2006; Co-

hard et al., 1998). The curvature of the activation spectrum depends on the size distribution of aerosol particles and thus, the environment of the respective air mass. Marine air masses with hygroscopic and bigger aerosol tend to have lower k values than continental environments, where small, hydrophobic particles are present (Jayachandran et al., 2020; Pruppacher and Klett, 1997). Cohard et al. (1998) stated that a power law with a constant k is not able to reproduce this curvature.

Several studies attempted to expand the original power law, e.g. using logarithmic (Wang et al., 2021) or exponential fits (Ji and Shaw, 1998). Khvorostyanov and Curry (1999) proposed a superposition of single-k power laws for different ranges of supersaturation and aerosol sizes, letting k decrease with increasing supersaturation and decreasing aerosol size.

In the following, we will show that using generalized s-dependent k parameters retrieved from observations give a very accurate representation of measured CCN spectra (Sect. 3.2) and are suited to be applied to simulated Nccn at one single supersaturation in order to create a full CCN spectrum.

2.5 Statistical approaches and analyzed parameters

The correlation between simulated and observed Nccn, as well as the quality of fit, are assessed using the bias B, the Pearson correlation coefficient R and, as a combined metric, the Kling–Gupta efficiency (KGE).

The relative bias B is computed from the medians of CAMS-derived and observed Nccn:

$$B = \frac{\mathrm{med}\,(\mathrm{CCN_C}) - \mathrm{med}\,(\mathrm{CCN_O})}{\mathrm{med}\,(\mathrm{CCN_O})} \cdot 100\,\%\,. \tag{5}$$

A bias B of 0 corresponds to equal medians. Unless stated otherwise, the sign of the bias is kept and indicates overestimation (+) or underestimation (−). This means especially that an increasing B is not necessarily equal to a larger deviation of simulated to observed Nccn. It might as well indicate a transition from under- to overestimation. When averaging biases from a group of stations, the sign will be disregarded, which gives the absolute value $|B|$.

Following its standard definition, the Pearson R is

$$R = \frac{\sum_{i=1}^{n}(\mathrm{CCN_{O,i}} - \overline{\mathrm{CCN_O}})(\mathrm{CCN_{C,i}} - \overline{\mathrm{CCN_C}})}{\sqrt{\sum_{i=1}^{n}(\mathrm{CCN_{O,i}} - \overline{\mathrm{CCN_O}})^2 \sum_{i=1}^{n}(\mathrm{CCN_{C,i}} - \overline{\mathrm{CCN_C}})^2}}\,, \tag{6}$$

with n being the number of points in the observation period.

The Kling–Gupta efficiency (KGE; Gupta et al., 2009) is a metric used to quantify the quality of a model's fit to observations and is widely used in hydrology. It is employed here to jointly account for bias (quantified by B) and explained variability (represented by R) and is defined as

$$\mathrm{KGE} = 1 - \sqrt{(R-1)^2 + (\alpha-1)^2 + (\beta-1)^2}\,. \tag{7}$$

R is the Pearson correlation coefficient, $\alpha = \frac{\mathrm{std_C}}{\mathrm{std_O}}$ is the ratio of the standard deviations of $\mathrm{CCN_C}$ and $\mathrm{CCN_O}$, $\beta = \frac{\overline{\mathrm{CCN_C}}}{\overline{\mathrm{CCN_O}}}$ is the ratio of the means of $\mathrm{CCN_C}$ and $\mathrm{CCN_O}$. The higher a KGE value is, the better the fit is between observation and simulation. A maximum KGE of 1 is obtained for two perfectly matching datasets. Negative KGEs are possible but do not necessarily correspond to anti-correlating data or a certain sign of the bias. In this study, the KGE of each station is mainly used as a comparison metric and will always be related to other stations. An absolute statement, whether a certain number is considered high or low, will not be made.

Total median Nccn have been examined using the introduced metrics B, R, and KGE. These primary metrics are further distinguished for their spatial variability (e.g. accounting for environmental setting and pollution regime) and for their variability in time (e.g. accounting for seasonal differences).

The assessed spatial variability is based on a classification defining the environment at each station as either marine or continental. Sites that are located south of 70° S or north of 70° N (outside the CAMS assimilation range) are classified as a separate group of polar stations (Fig. 1).

Commonly, aerosol and CCN concentrations are also distinguished for different regimes. For example, Varghese et al. (2016) used a classification into clean ($< 1000\,\mathrm{cm^{-3}}$), polluted and highly polluted ($> 2000\,\mathrm{cm^{-3}}$), based on the Nccn at a supersaturation of 1 %. Andreae (2009) proposed a threshold of $600\,\mathrm{cm^{-3}}$ at 0.4 % supersaturation for polluted marine regions and Hudson and Yum (2002) defined groups of clean ($< 200\,\mathrm{cm^{-3}}$) and polluted air masses ($> 1000\,\mathrm{cm^{-3}}$) for 1 % supersaturation, with a transition regime in between. In this study, most analyses are based on 0.4 % supersaturation as a mid-range supersaturation level. At this supersaturation, a classification into a clean ($\mathrm{CCN_O} < 500\,\mathrm{cm^{-3}}$) and polluted ($500\,\mathrm{cm^{-3}} \leq \mathrm{CCN_O}$) regime has been constructed, which is similar to the mentioned classifications by Andreae (2009) and Hudson and Yum (2002), but with lower thresholds than in Varghese et al. (2016).

Since one aim of this study is to validate the spectral behavior of CAMS-derived Nccn, the sensitivity of the Nccn to a change in supersaturation will be used as a secondary metric. These Nccn–s spectra are approximated by applying Twomey's power law. The slope of the spectrum is then described by the slope parameter k, using Eq. (3) for $s_1 = 0.1\,\%$ and $s_2 = 1.0\,\%$. The larger k the steeper the activation spectrum is. Large k are usually found for continental or polluted air masses (Martins et al., 2009; Pruppacher and Klett, 1997).

3 Results and discussion

3.1 The performance in time and space

Different environments and regimes certainly affect the order of magnitude of Nccn and thus the validity of $\mathrm{CCN_C}$ presented here. In order to compare different stations with

Table 2. Cluster (as defined by observed pollution regime and environment), Kling–Gupta efficiency KGE, bias B, correlation coefficients R and number of points # as derived from daily CAMS CCN data at 00:00 UTC and observational data from 22:00 UTC to 02:00 UTC for all 25 stations, sorted by KGE from highest to lowest. CCN are taken at $s = 0.4\,\%$, except for ATT where $s = 0.4\,\%$ was not available as such and instead $s = 0.47\,\%$ was applied here – which eventually might lead to a slightly decreased bias. R values in parentheses indicate not significant correlations (95 % confidence interval). The cluster names are abbreviated: clean continental as "cle-con", polluted continental as "pol-con", clean marine as "cle-mar", polluted marine as "pol-mar". In comparison, correlations were also computed for 11-d running medians and monthly medians, both based on or around 00:00 UTC data only.

		Daily				11-d running medians		Monthly medians	
ID	Cluster	KGE	B [%]	R	#	R	#	R	#
MHD	cle-mar	0.505	+14.09	0.669	188	0.814	222	0.736	8
ATT	pol-con	0.493	+11.15	0.543	172	0.854	287	0.744	12
PVC	pol-mar	0.406	−11.36	0.455	142	0.285	167	(0.491)	7
FKB	pol-con	0.373	+25.19	0.506	266	0.659	280	0.742	11
SMR	pol-con	0.279	−38.47	0.321	746	0.748	878	0.849	30
ASI	cle-mar	0.164	+63.80	0.818	420	0.833	461	0.740	16
NOT	pol-mar	0.155	−5.30	0.214	233	0.412	292	0.717	10
VAV	pol-con	0.147	−27.97	0.173	388	0.267	450	(0.178)	21
FIK	pol-mar	0.146	−13.42	0.172	271	0.539	323	0.775	11
SGP	pol-con	0.084	−38.46	0.174	3847	0.184	4208	(0.147)	146
CES	pol-con	0.075	−40.60	0.287	560	0.432	583	(0.446)	19
JFJ	cle-con	0.066	−42.31	0.385	548	0.666	696	0.868	25
AWR	cle-mar	0.023	+201.19	0.487	198	0.821	205	0.911	8
GRW	cle-mar	−0.016	+22.90	0.121	563	(0.020)	583	(−0.098)	21
NIM	pol-con	−0.042	+99.45	0.243	262	0.364	283	(0.388)	12
SEO	pol-con	−0.097	+69.15	0.183	586	0.455	786	0.558	31
OLI	cle-mar	−0.126	+190.46	(0.019)	256	−0.122	317	(−0.032)	13
MEL	pol-con	−0.141	−45.02	0.129	263	0.202	579	(−0.162)	25
PYE	cle-mar	−0.175	+93.13	0.282	178	(0.092)	186	(−0.504)	7
PGH	pol-con	−0.201	+75.59	0.609	217	0.404	271	(−0.123)	10
ENA	cle-mar	−0.207	+37.82	0.293	682	0.548	744	0.664	26
NSA	cle-mar	−0.230	−38.95	0.146	1329	(0.037)	1491	(0.099)	50
COR	pol-con	−0.240	−68.71	(−0.061)	184	−0.309	207	(−0.590)	8
MAO	pol-con	−0.339	+48.74	0.441	83	0.760	81	(0.376)	3
BRW	cle-mar	−1.558	−17.85	0.225	129	(0.074)	264	(0.196)	11

each other we make use of the primary statistical quantities B, R, and KGE from daily data at or around 00:00 UTC and distinguish the stations by observed pollution regime and environmental setting. A summary is given in Table 2.

3.1.1 Evaluating the performance in time

Using the advantage of the CAMS-derived Nccn dataset as a continuous multi-year time series, a qualitative assessment of its ability to reproduce temporal variability is performed. Table 2 compares R values for each station distinguishing daily medians, running medians with windows of 11 d and monthly medians, all based on 00:00-UTC time periods.

Using daily 00:00-UTC medians, only 5 out of 23 stations with significant correlation coefficients show R values above 0.5 indicating at least moderate correlation of day-to-day fluctuations. This is a poor result considering that the CCN data are derived from a reanalysis in which the underlying aerosol fields are constrained by observations. Diurnal variations which might impact bias and correlation coefficient cannot fully be evaluated here only using 00:00-UTC CAMS data. Hence, applying 3-hourly CAMS CCN data could be very beneficial to account for these effects. However, in the current dataset by Block (2023), this temporal resolution is not available yet. A more detailed analysis on short-term variability of selected sites is given in Anders et al. (2025).

Applying an 11-d running median on both datasets largely removes variability shorter than the synoptic scale, thereby improving temporal clarity of the time series at most stations. This filtering increases the correlation coefficient of 17 stations, compared to the daily medians, and thus, enhances the confidence that the CAMS-derived Nccn captures variability on synoptic time scales. Still, only 10 of 21 stations yielding significant result show R values larger than 0.5, which is about half of all stations. Correlation coefficients increase again for 8 stations when using monthly medians, especially for NOT, FIK and JFJ the increase is quite drastic. However, only for 11 of the 25 stations significant correlation coeffi-

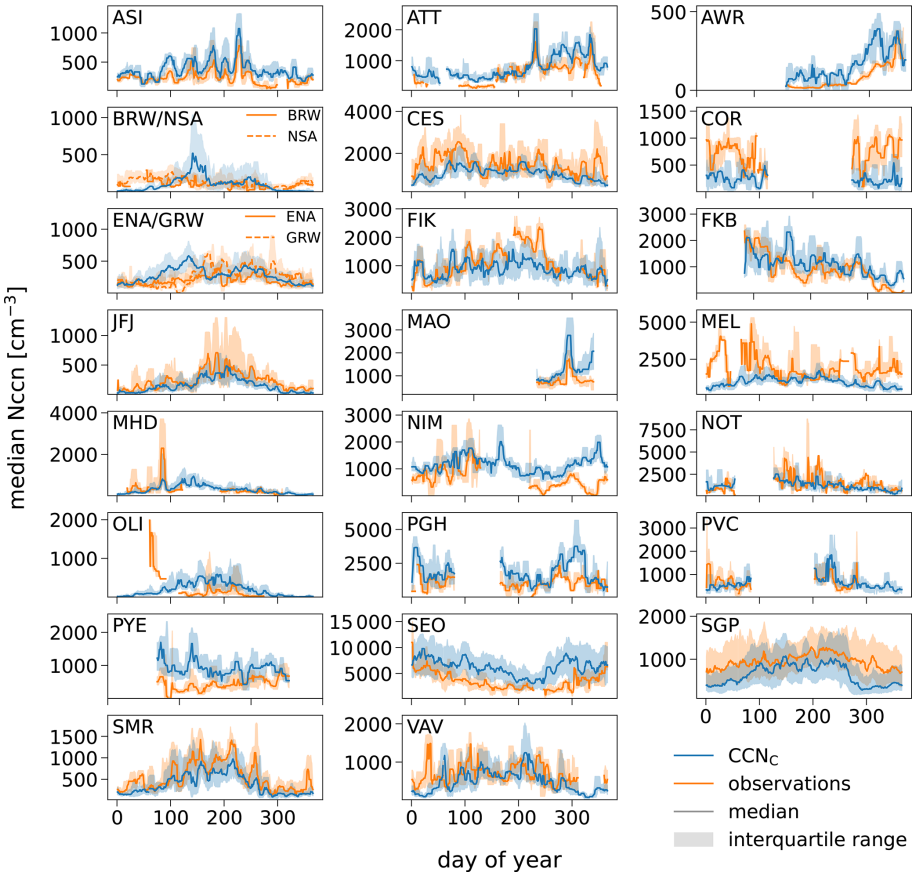


Figure 2. Climatologies for all stations at a supersaturation of 0.4 %. CCN_C in blue and CCN_O in orange, both with a running median and interquartile range in a window of 11 d. Please note: Data from 29 February has been removed and the co-located data from BRW and NSA as well as ENA and GRW are shown in the same tile.

cients are found. For those, R values range from +0.558 to +0.911 within all remaining 11 stations across all environments and pollution regimes.

The results of Table 2 are further illustrated in Fig. 2 showing multi-year 11-d running medians, and Fig. A2 showing multi-year monthly medians of Nccn ($s = 0.4\,\%$) for all stations. Even though a correlation analysis of seasonal and annual medians does not produce significant results, as the available measurement periods are not long enough, we can visually judge the simulated temporal variability over several months indicating a yearly cycle from Figs. 2 and A2. From those 18 stations with at least 10 months of continuous observations available, 13 agree with CAMS in the yearly distribution of CCN, featuring e.g. correct differences in summer/wintertime CCN (SEO, SMR, SGP or JFJ), spring/autumn CCN (ATT, CES or FKB) or the double summer peak for VAV. Furthermore, at stations featuring wet and dry seasons, such as ATT, FIK, NIM and SEO (Pöhlker et al., 2016; Kouvarakis et al., 2000; Galy-Lacaux and Modi, 1998; Schmale et al., 2018), the increase in observed dry season Nccn is well captured by CAMS-derived Nccn (Fig. A3).

From this we conclude that CAMS-derived Nccn mostly reproduces the observed variability on monthly and seasonal time scales and therefore is suitable to be used for climatological applications and studies focusing on these timescales.

3.1.2 Dependencies on environment and pollution regime

There are no systematic differences in KGE at one fixed supersaturation (here 0.4 %), when comparing the four clusters. Averaging between all clean marine stations results in a KGE of −0.181, for polluted marine a KGE of 0.236, for clean continental 0.066 and for polluted continental stations 0.033. These differences are small compared to the spread between stations of the same cluster (Table 2). Furthermore, analyzing the KGE in dependence of the stations' geographical position or separately their environment or pollution regime has given no clear result either.

Block et al. (2024) made the comment, based on PGH data only, that very polluted sites might be overestimated by CAMS CCN. To further evaluate this feature we examined the bias of the most polluted sites available in our observa-

tions. Observations at SEO, MEL, CES, NOT, FIK and PGH all show median $CCN_O > 1000\,cm^{-3}$ for $s = 0.4\,\%$, outranging all other observations. But only SEO and PGH are overestimated by CAMS, while the other 4 stations are underestimated. Thus, we cannot confirm that heavily polluted sites are high biased in general. However, examining data from other megacities, such as in Asia and Africa, might be necessary, as our list of stations only included one (SEO).

For most stations, the pollution regime as identified by CCN_C ($s = 0.4\,\%$) matches the regime it would have if regimes were defined by the observational CCN_O ($s = 0.4\,\%$). A comparison of both definitions is part of Table A2. Only 4 stations deviate as a result of their bias, and thus would fall into a different category. Due to their negative bias, SMR, VAV and COR are classified in CCN_C as clean regimes instead of polluted ones. Meanwhile, PYE is classified as polluted instead of clean due its positive bias. A decrease in absolute bias would lead to a correction of pollution regime identification.

3.1.3 Polar sites

Four sites (AWR, BRW, NSA and OLI) lie in polar regions and therefore outside of the global domain of the data assimilation scheme of ECMWF's IFS between 70° S–70° N. Due to possible irregularities caused by the lack of observational constraints in this area, CAMS-derived Nccn data need to be used carefully. In particular, the co-located stations BRW and NSA exhibit simulation features that did not appear at any other station. There are partially sharp peaks, exceeding the observations by a factor of up to 4, and extremely low concentrations below $1\,cm^{-3}$ in the winter months. These extremes result in very low correlation coefficients at both BRW and NSA. The same holds for the nearby station OLI. The Antarctic station AWR, however, exhibits a value of R that is comparable to stations of the mid-latitudes but it produces a very large bias (Table 2) even though, both CCN_O and CCN_C are very low. The presence of wrong peaks and large biases illustrates very clearly that the benefits of the assimilation of satellite observations are missing for the polar regions.

3.2 The CCN-supersaturation spectrum

3.2.1 Evaluating the CCN spectrum

As depicted from Fig. 2, CCN are overestimated by CAMS at some stations (e.g. AWR, NIM or ASI), while underestimation is present at others (e.g. MEL, JFJ or SGP). These offsets change depending on the chosen supersaturation (Anders et al., 2025), as different fractions of CCN species get activated and thus shape the CCN spectrum.

Figure 3 compares observed and CAMS-derived Nccn-supersaturation spectra, which are computed from daily median Nccn for all six supersaturations, for which CAMS-derived total CCN are available. Both, CCN_C and CCN_O increase with supersaturation as expected. However, while CCN_O show shallower slopes for higher supersaturations, CCN_C grow almost linear with supersaturation and do not reach an aerosol-limited regime as the observations. Consequently, this leads to both an overestimation (positive bias B) at high supersaturations and an underestimation for low supersaturations. Table A2 summarizes nicely how biases evolve with supersaturation. While for $s = 0.1\,\%$ only 8 stations were positively biased but 17 stations had a negative bias, this reversed for $s = 1.0\,\%$, where 19 stations had a positive bias but only 6 were negatively biased. For 20 out of 25 stations, B at a supersaturation of 1.0 % is larger (more positive or less negative) compared to its value at 0.1 %. When assessing how large the biases are in comparison, we need to look at their absolute values. 11 stations (of which 7 are marine) have larger $|B|$ at higher than at lower supersaturations, while 14 stations (of which 9 are continental) depict larger $|B|$ at low supersaturations. In the median of all stations, the absolute value $|B|$ is 65.89 % for $s = 0.1\,\%$ and 53.13 % for $s = 1.0\,\%$. Hence, the underestimation at low supersaturations (which is reached more often in natural cloudy conditions) is an at least equally important issue as the overestimation at high supersaturations.

As found by Anders et al. (2025), the best-fitting supersaturation s_{best}, for which B is closest to zero, differs between stations. As noted before and depicted in Fig. 4, marine stations are least biased at low supersaturations, while continental stations are least biased at higher supersaturations.

All three remote marine stations have their s_{best} at 0.2 % and, together with MHD, they create a "clean marine" category. The "polluted marine" category is dominated by coastal stations with mostly marine but also continental influence. These stations show a larger variability of s_{best}, ranging from 0.1 to 0.6, with a median value of $s = 0.3\,\%$. Stations of the "polluted continental" category mostly occupy the larger supersaturation ranges around 0.6 %. A separate "clean continental" category emerges because CAMS-derived Nccn underestimates the pollution level at several inland stations (COR, VAV, and SMR). Without these classification mismatches, this category would contain only JFJ. This class has a median s_{best} at 0.9 %.

Some stations do not fit perfectly into this framework. For example, COR exhibits nearly identical biases for 1.0 % as for 0.1 % (Table A2) and is just by minimal deviations an outlier here. Because of a regime mismatch, similar holds for PYE, which would fit very well to the according cluster if it would actually be clean-marine as observed. However, because our objective is to correct the CAMS-derived dataset itself, all classifications are based on the aerosol regimes represented by CAMS rather than on observational categories. Finally, we find no systematic relationship between s_{best} and pollution regime alone. Instead, both environmental setting and pollution regime are required to explain the observed variability in s_{best}.

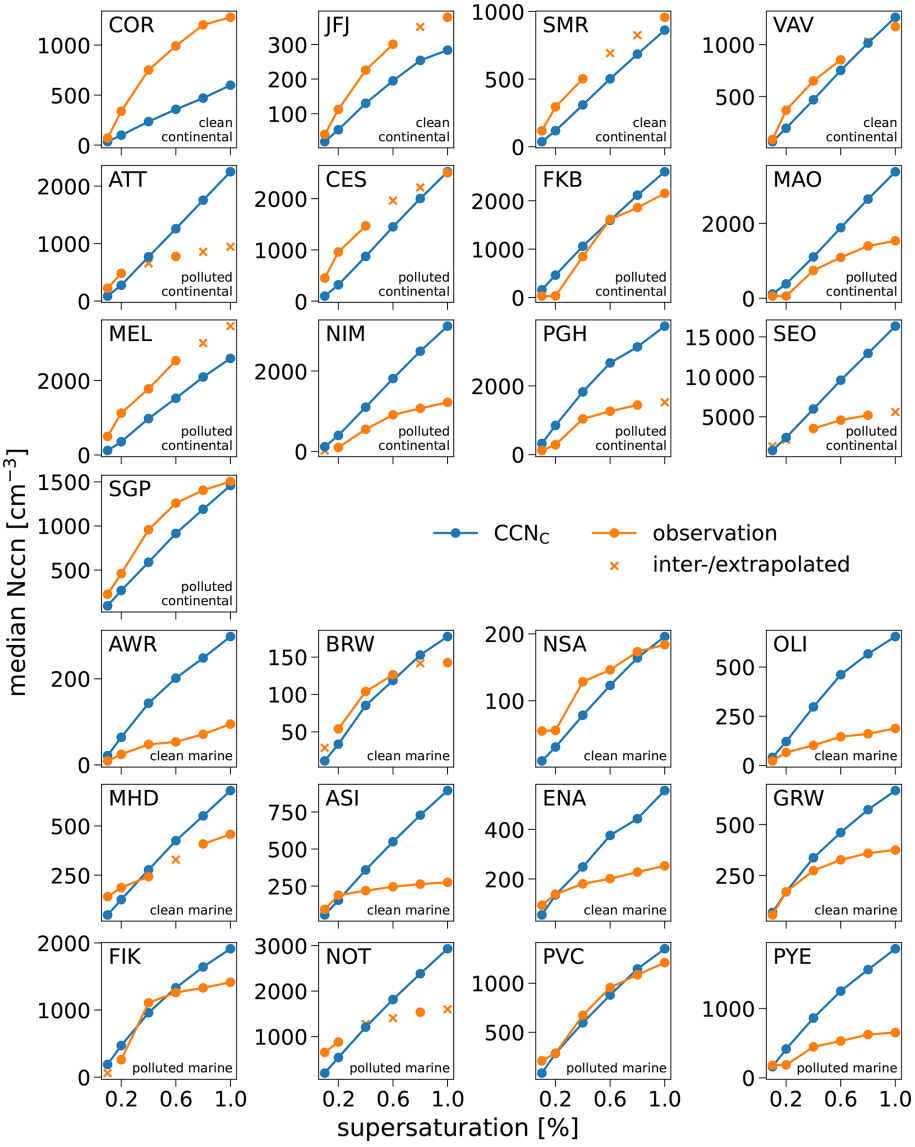


Figure 3. Nccn–s spectra showing the medians of every analyzed ground-based measurement campaign over the corresponding supersaturation. CCN_C in blue and CCN_O in orange. Observation values that are indicated by an "x" are medians of an inter- or extrapolated set of 00:00-UTC daily Nccn, using Twomey's power law (Eq. 2). Please note, the y axis is different in every plot. The x axis is shared.

The curvature of a CCN spectrum depicted in Fig. 3 and thus the bias for each supersaturation might be dependent on the involved aerosol species. As already depicted in Block et al. (2024), Nccn derived from CAMS are mostly dominated by organic matter (OM) and sulfate (SO_4), while black carbon (BC) fractions are mostly small and the fractions of the three sea salt (SS) modes are negligible, even under remote marine conditions. With higher supersaturation, smaller and less hygroscopic aerosols get activated. Hence, the OM and BC fractions increase with supersaturation, while the portions of SO_4 and SS are reduced. Looking at the ratio of the OM and SO_4 fractions, we found a very clear relationship to the above-mentioned slope parameter k (Fig. 5): The larger the OM-to-SO_4 ratio, the larger is k. A logarithmic fit was found with an R^2 of 0.865. Large OM-to-SO_4 ratios above 1.5 at 0.2 % supersaturation are found at MAO, NIM and ATT. These three stations indeed exhibit similar shapes in Fig. 3, having a strongly increasing B with increasing supersaturation. On the other hand, stations with low OM-to-SO_4 ratios (FIK, AWR and BRW), below 0.3 at 0.2 % supersaturation, tend to have a weaker increase in B. It seems that the stations with lower OM fractions are likely to have a smaller absolute value $|B|$, but tend to produce a lower correlation coefficient R. As a result of these two contradictory effects, a systematic relation of the OM-to-SO_4 ratio on the KGE could not be determined. Examining the sea salt and

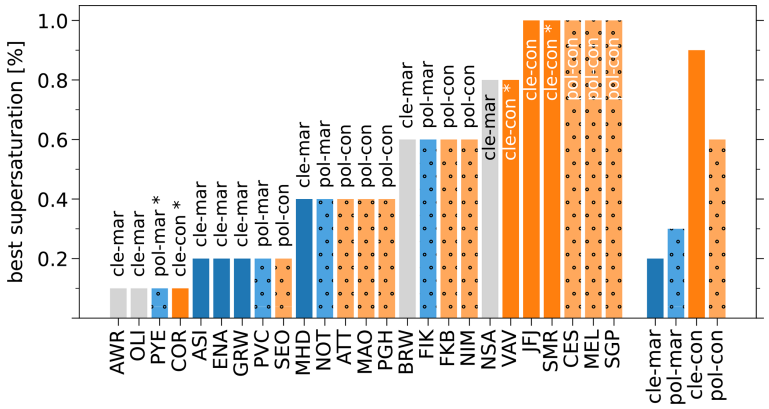


Figure 4. Supersaturation where CCN_C and CCN_O show best fit, for each station respectively. Distinguished into clean marine ("cle-mar", blue), polluted marine ("pol-mar", purple), clean continental ("cle-con", green) and polluted continental ("pol-con", orange) clusters, as defined by CCN_C. Cluster names with an asterisk indicate stations where the regime classifications of CCN_C and CCN_O do not match (see also Table A2). On the right, the respective median best supersaturation of each cluster. Polar stations are shown in grey for completeness but are excluded from the cluster median.

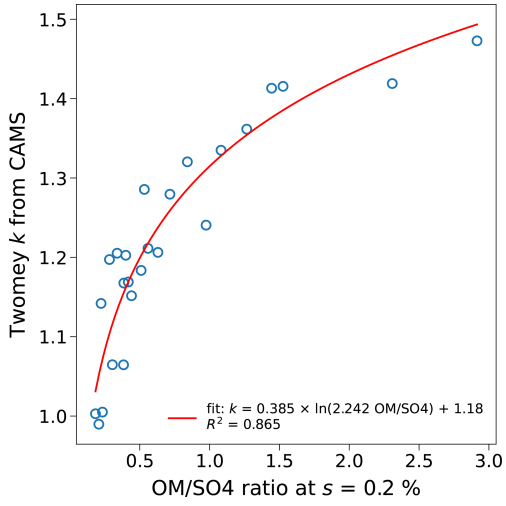


Figure 5. Slope parameter k from CCN_C, as defined by Eq. (3) with $s_1 = 0.1\,\%$ and $s_2 = 1.0\,\%$, against the ratio of the median organic matter (OM) and sulfate (SO_4) CCN_C fractions. Each data point (blue circle) represents one station. A logarithmic fit was performed (red solid line).

BC fractions, no significant correlation to the KGE is obtained either.

This specific feature suggests a possible cause of the biased CCN spectrum. The discrepancy could stem from uncertainties in the emission fractions of CCN-relevant aerosol species (as these are not constrained in CAMS), from limitations in the assumed aerosol size distributions and composition used in the CCN calculations (e.g. fixed global modes and kappa values, assumption of external mixtures), or from a combination of both. Nevertheless, the overall Nccn spectrum may be improved without modifying aerosol size distributions or species concentrations. Instead, this approach relies solely on the observed CCN spectrum as demonstrated in the following section.

3.2.2 Deriving a spectral $k(s)$ from observations

As described in the previous section, we found large biases in CAMS-derived Nccn at both low and high supersaturations. To reduce these biases, we present an approach tailored to the unique characteristics of the CAMS-derived Nccn dataset. The idea is to fit the simulated Nccn–s spectrum to the observations. For that, due to its simplicity and its prevalent application in CCN activation spectra (e.g. Martins et al., 2009; Cohard et al., 1998), we use Twomey's power law (Eq. 4). However, its original form must be modified to serve the targeted purpose, reducing biases in CCN_C at both low and high s.

First of all, we cannot use 1.0 % as the reference supersaturation s_i, since this would preserve the large biases at this supersaturation. Instead, we use a cluster-specific reference, which accounts for the different s_{best} we found in the previous section for the different clusters (Fig. 4). Based on these findings, we defined four cluster-specific s_i: 0.2 % for clean marine, 0.4 % for polluted marine, 0.6 % for polluted continental sites and 1.0 % for clean continental stations.

Generalizing Twomey's power law by applying it to other supersaturations than 1.0 % has been done already (e.g. Block et al., 2024; Andreae, 2009) and in our case, it is a promising modification to avoid high-supersaturation biases and to target each environment specifically. It needs to be noted that this flexibility in particular is the main advantage of the power law over other parameterizations. The exponential function by Ji and Shaw (1998), the error function by Pöhlker et al. (2016) and the logarithmic formulation by Wang et al. (2021) all have Nccn (1.0 %) as one of their inputs. We do not want to evaluate the capabilities of all these

parameterizations for fitting observed Nccn–s spectra here as they are not suitable for our purpose.

Furthermore, we found for our data that keeping a constant exponent k in the power law leads to increased biases at low supersaturations. Similar results have been shown e.g. by Kim et al. (2011) and Wang et al. (2021). Furthermore, k derived from observations decrease with supersaturation (e.g. Khvorostyanov and Curry, 2006; Cohard et al., 1998; Yum and Hudson, 2001). This is also the case for the majority of observations we use in this study: Fig. 6 depicts k values of each station, as calculated from Eq. (3). For this and the following other k parameterizations, we use Nccn medians of the full daily cycle that is available in the observations (not limited to the period between 22:00–02:00 UTC as for the analysis above). We do this to increase the robustness and confidence in our improvements of CAMS-derived Nccn.

We propose to use these observation-derived, spectral $k(s)$ instead of a constant k in the power law:

$$\mathrm{Nccn}(s_x) = \mathrm{Nccn}(s_i)\left(\frac{s_x}{s_i}\right)^{k(s)}. \tag{8}$$

Similar notations of a power law with a supersaturation-dependent exponent have been used by e.g. Khvorostyanov and Curry (2006, 2008), and Liu et al. (2011). For the purpose of reducing low supersaturation biases in $\mathrm{CCN_C}$ and ultimately, simulating a more accurate CCN activation spectrum, this approach is suitable to modify the yet too steep slope of the $\mathrm{CCN_C}$ spectrum.

As a first step, we average the calculated station-specific $k(s)$, to obtain a global spectrum. Since the majority of stations we use are continental, calculating the median across all stations would result in a bias towards the continental stations. Thus, we group the observation-derived $k(s)$ spectra of the stations by environment and compute the median of both groups. To account for the uneven global distribution of continental (30 %) and marine environment (70 %), we perform a weighted average of these environment medians for each supersaturation, to obtain a global spectrum of observation-derived $k(s)$. At the six CAMS supersaturations, the values are 0.838, 0.727, 0.669, 0.464, 0.455, and 0.436, respectively. For comparison, we calculated the same global, environment-weighted mean $k(s)$ from $\mathrm{CCN_C}$ which gives 1.312, 1.162, 1.099, 1.006, 0.952, and 0.868, respectively. It becomes clear that the observation-derived $k(s)$ differ widely from an analogously CAMS-retrieved $k(s)$, revealing the root of the biased CAMS-derived CCN spectrum.

Lastly, to reduce the risk of overfitting to our set of observations and to extrapolate $k(s)$ to 0.1 % and 1.0 %, we applied a power law fit for supersaturations starting at 0.2 % and a linear fit extrapolating to 0.1 %. The power law appears to be the best-fitting function for the given points of $k(s)$. The second linear fit is necessary due to the nature of the power law, which would grow to infinity at low values of s. Similarly, Khvorostyanov and Curry (2006) have found observed k spectra to decrease more slowly in the range below 0.12 % supersaturations compared to higher s. Wang et al. (2021) used a piecewise function that fits Nccn below 0.1 % linearly, which gives better results than the Twomey parameterization.

The global $k(s)$ parameterization can be written as:

$$k(s) = \left\{ \begin{array}{ll} -1.041286\,s + 1.069804 & \text{for } s < 0.2\,\% \\ 0.370341\,s^{-0.524598} & \text{for } s \geq 0.2\,\% \end{array} \right\}. \tag{9}$$

We also explored other possible derivations of a k–s spectrum, by treating each regime/environment/cluster individually instead of a global mean. In Fig. A4, we show a comparison of these approaches at SMR, SGP, ENA and FIK, which are the stations with the highest number of points in their respective cluster.

For the constant k parameterizations, we apply Eq. (3) with $s_1 = 0.1\,\%$ and $s_2 = 1.0\,\%$ for each individual station, resulting in one single-value k per station. For applying a correction with a constant global k, we group these k values by environment and perform a weighted average of the median of all continental stations (30 % weight) and median of all marine stations (70 % weight). Thus, we use one single-value k globally. The correction with a constant cluster k is done similarly but with grouping by cluster before applying the median single-value k in the respective cluster.

Additionally, bias corrections are performed with four different $k(s)$ parameterizations in comparison. First, an s-dependent spectrum of $k(s)$ is calculated for each station. For applying a correction using a global mean $k(s)$ on all stations, we group these station spectra by environment and perform a weighted average of the median of all continental stations (30 % weight) and median of all marine stations (70 % weight). The resulting global average $k(s)$-spectrum is fitted, as described above. Thus, we use one global $k(s)$-spectrum for each station. The parameterization with a regime dependent $k(s)$ is done by grouping the station spectra by regime and fit the regime-median $k(s)$. This fitted spectrum is used for each station of the respective regime. Parameterizations using the environment dependent $k(s)$ and the cluster dependent $k(s)$ are done similarly, but by grouping first by environment or by cluster, respectively.

All Nccn parameterizations are then retrieved from Eq. (8) with the respective k or $k(s)$ and the cluster-specific reference supersaturation s_i (clean continental: 1.0 %, polluted continental: 0.6 %, clean marine: 0.2 %, polluted marine: 0.4 %.)

For most stations the differences between group-specific (regime, environment, cluster) and global $k(s)$ are not very large (as at ENA and FIK in Fig. A4), thus it would be unnecessary to distinguish between clusters (or environments/regimes), as we did for the definition of s_i. This simplifies applications on a global scale in any future developments or use cases. However, another reason why we prefer the parameterization with a globally weighted mean $k(s)$ is that the bias correction produces the most balanced biases for low and high supersaturations, as can also be seen in Fig. A4 for SMR and SGP. Please note that all 25 stations behave

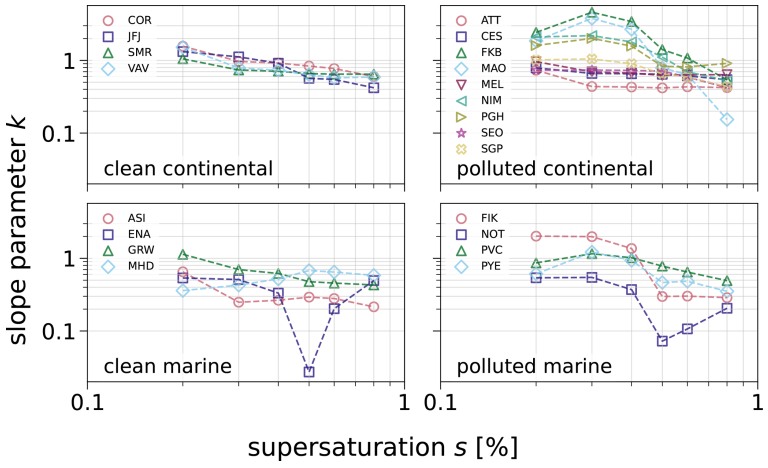


Figure 6. The slope parameter $k(s)$, derived from the observations (Eq. 3) and grouped by clusters identified by CAMS-derived Nccn. Polar stations are excluded.

differently, which is why the chosen parameterization is not necessarily the best fit at each station. It is rather a trade-off between the individual stations and the applicability on a global scale. For completeness, all group-specific $k(s)$ parameterizations are listed in Table A3.

3.2.3 Towards an improved CCN spectrum

Applying Eq. (9) to the six supersaturations CAMS-derived Nccn are available at, we receive a global spectrum of $k(s)$ values, which are independent of the environment. Using the new global set of s-dependent k values and the environment-specific s_i, we applied Eq. (8) to the original daily $CCN_C(s_i)$. The resulting medians are depicted in Fig. 7, which illustrates the improvement this new approach achieved. The curvature of the activation spectra at most stations is simulated more accurately. Especially, at the clean marine stations ASI, ENA and GRW as well as polluted marine FIK and NOT, this is the case. In addition, ATT, MAO, NIM, PGH and SEO are examples from the polluted continental cluster where high supersaturation biases are reduced significantly. However, Nccn which the original CCN_C have already underestimated (e.g. for CES, COR, JFJ, MEL, SGP) do not benefit from the adjustments and negative biases at continental stations and high supersaturations are enhanced. To account for these very region-specific biases, a modification of the underlying size distribution for CAMS-derived Nccn or a redistribution of CCN-relevant aerosol masses might still be necessary.

Despite these remaining challenges, we could reduce the biases for the majority of all stations (Fig. A5 in the Appendix). Especially at the high supersaturations 0.8 % and 1.0 %, the absolute values of the bias are decreased substantially. Also at low supersaturations, absolute biases $|B|$ are reduced in both environments. Specifically, the medians of the biases of all non-polar stations at 0.1 %, 0.2 %, 0.4 %, 0.6 %, 0.8 %, and 1.0 % supersaturation are reduced by 30.1 %, 29.5 %, 41.4 %, 34.5 %, 51.7 %, and 43.3 %, respectively. The improved median absolute values $|B|$ are 42.9 %, 36.8 %, 22.6 %, 26.1 %, 26.6 %, 30.1 %, respectively. Unlike Fig. A5, these absolute values do not indicate a direction of the bias (under-/overestimation) but show that we could reduce the magnitude of the biases at all supersaturations substantially.

4 Summary and conclusions

The recently published 3D-dataset of CCN number concentrations (Nccn) derived from CAMS reanalysis aerosol mass mixing ratios (Block et al., 2024) is evaluated against direct observations from 25 ground-based measurement sites. The evaluation focuses on three aspects of the CAMS-derived total Nccn: (1) its ability to reproduce observed temporal variability, from day-to-day fluctuations over synoptic-scale variations to monthly timescales; (2) its suitability across different environments (marine, continental) and pollution regimes (clean and polluted); and (3) its sensitivity to supersaturation.

Regarding temporal variability, performance varies substantially between stations, even within the same environment or pollution regime. In particular, the representation of day-to-day variability differs considerably across sites, with significant Pearson correlation coefficients ranging from +0.121 to +0.818. Only 5 of the 23 stations with statistically significant correlations at the 95 % confidence level exhibit correlation coefficients exceeding 0.5, indicating limited skill in reproducing daily variability despite the underlying aerosol fields being constrained by reanalysis data. However, model performance improves markedly as the temporal aggregation period increases. The highest correlation coefficients are obtained for monthly median values, with R ranging from +0.558 to +0.911, within all 11 stations still yielding statistically significant correlations. In addition,

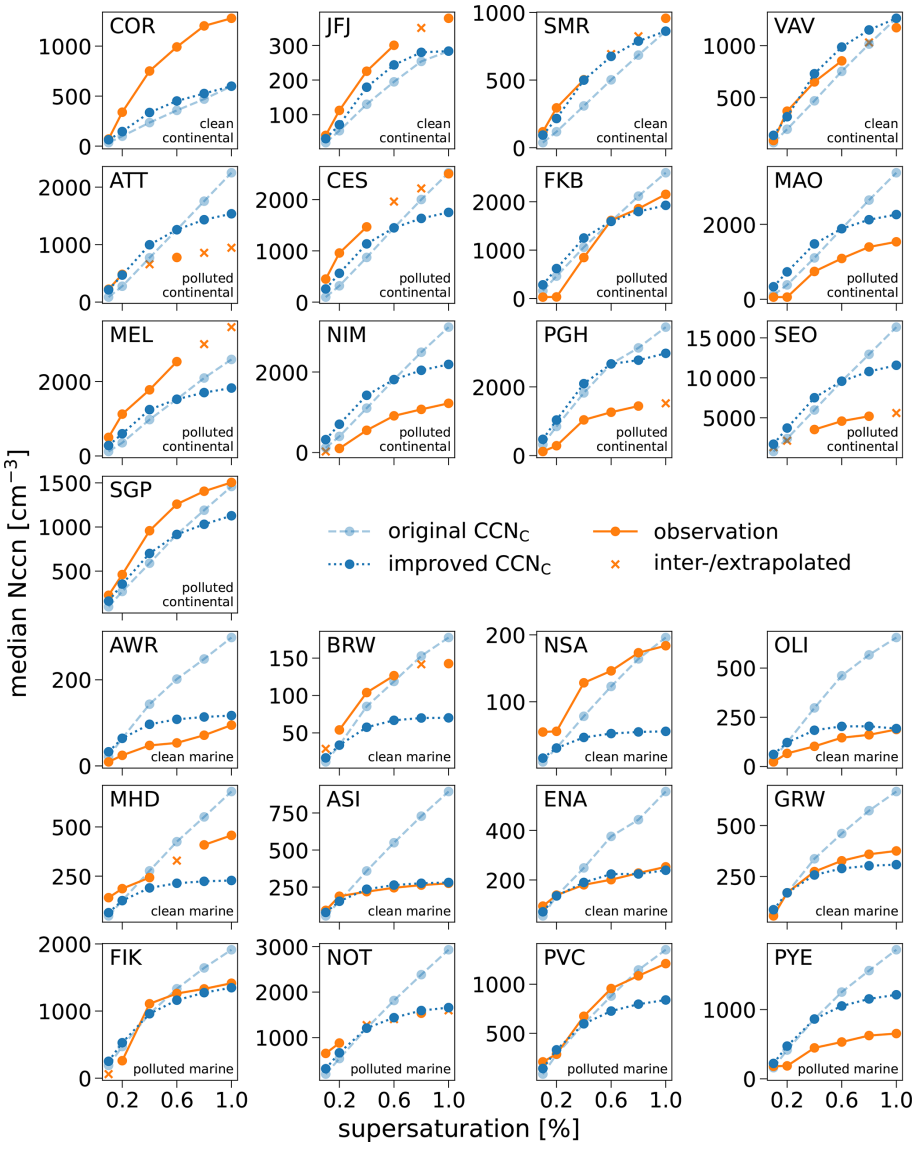


Figure 7. Same as Fig. 3, but with the addition of improved CCN_C (dark blue).

the CAMS-derived Nccn generally reproduces the observed seasonal cycles, including the occurrence of dry and wet seasons, at most stations yielding long enough continuous measurements. These results suggest that the dataset is well suited for climatological applications and studies focusing on monthly to seasonal timescales, while caution is warranted when interpreting day-to-day variability.

Considering the suitability of CAMS-derived Nccn in different environments and regimes, we evaluated Nccn for a mid-range supersaturation of $s = 0.4\,\%$ and made use of Kling–Gupta efficiency (KGE) as an estimate of quality of fit, taking into account correlation as well as bias. The resulting values show a large spread between stations and no clear dependence of the model's performance to environmental settings nor to pollution regimes. Investigations of the most polluted sites could not confirm the statement in Block et al. (2024) that there might be a tendency for overestimation at heavily polluted sites, even though we find that PGH and SEO are high biased. However, apart from SEO, we did not investigate data from other megacities. It should further be noted that CAMS-derived CCN data are indeed problematic in Arctic and Antarctic regions as large biases and low correlation coefficients are obtained at some of the investigated polar stations. This could be due to the lack of constrainment by the CAMS data assimilation scheme which is only active between 70° S–70° N. As already stated in Block et al. (2024), one should be careful when using the CAMS-derived CCN dataset in these regions.

The examination of the sensitivity of Nccn to different levels of supersaturation revealed an increasing bias with

increasing supersaturations. At several stations, this results in a transition from underestimation at low supersaturations to overestimation at higher supersaturations. The bias shift could be described using the slope parameter k of Twomey's power law, which was applied onto the CCN-supersaturation spectrum. For almost all stations, CAMS-derived Nccn exhibit a larger k than observed Nccn, which is related to an overly steep activation spectrum. Further analysis suggests that these discrepancies are linked to the ratio of organic matter (OM) to sulfate (SO_4) of the related CCN species. Stations characterized by lower OM-to-SO_4 ratios generally exhibit smaller k values, implying a weaker supersaturation dependence. This points to potential deficiencies in the assumed aerosol composition and/or size distributions used in the derivation of the dataset. Consequently, refinements of the underlying size distribution may further improve the agreement with observations. While such modifications were beyond the scope of the present study and remain a topic for future dataset development, we instead developed an observation-based parameterization to address this issue. We demonstrate that its application to CAMS-derived total Nccn improves agreement with observations without modifying aerosol size distributions or species concentrations. These improvements further enhance the suitability of the CAMS CCN dataset for global aerosol-cloud-interaction studies.

A significant reduction of the biases in CAMS Nccn could be achieved by applying a modified power law with an s-dependent k spectrum, which has been derived from the observations. For the application on CAMS data we grouped the data into classes, distinguishing environmental settings (marine/continental) and pollution regimes (clean/polluted) for which we computed the least biased level of supersaturation. Thus, we find that e.g. marine stations are least biased at low supersaturations, while continental stations are least biased at higher supersaturations.

At a supersaturation of 1.0 %, the median bias decreased by 43.3 % to an absolute value of 30.1 %, indicating a reduction of the overestimation. For lower supersaturation of 0.1 %, it decreased by 30.1 %, to an absolute value of 42.9 %, which corresponds to a reduction of the underestimation at this supersaturation. The majority of stations exhibited more accurate re-presentations of their CCN-supersaturation spectra, especially those in marine environment. However, biases at high supersaturations persist at some stations.

Despite our efforts of exploring and evaluating CAMS-derived Nccn, further investigations should also focus on the validation of the vertical distribution, as some problems have already been addressed by Kulkarni et al. (2025) and Anders et al. (2025). It is also of importance to investigate the quality of CAMS-derived Nccn over yet neglected regions in Africa, Asia and Australasia and above the Southern ocean, since this study could only partially contribute to the discussion started by Choudhury et al. (2025). Additionally, the comparison to more stations of long-term CCN measurements over multiple years will help in assessing the representation of not only day-to-day but also annual variability and eventually multi-decadal trends.

Appendix A

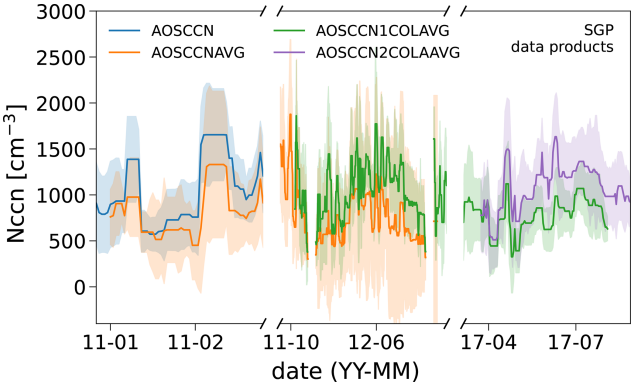


Figure A1. Periods of overlap of four data products of SGP that were available for our analysis: AOSCCN is the Aerosol Observing System Cloud Condensation Nuclei product, AOSCCNAVG is the AOSCCN Average product, AOSCCN1COLAVG is the AOSCCN Single Column Average product and AOSCCN2COLAAVG is the AOSCCN Dual Column Average product. We have combined these data products to one dataset, which we have used for this study. For overlapping time periods, we have selected that one with the higher reported level of data quality. Full-day Nccn are shown at 0.4 % supersaturation and with a 11-d running median and interquartile range.

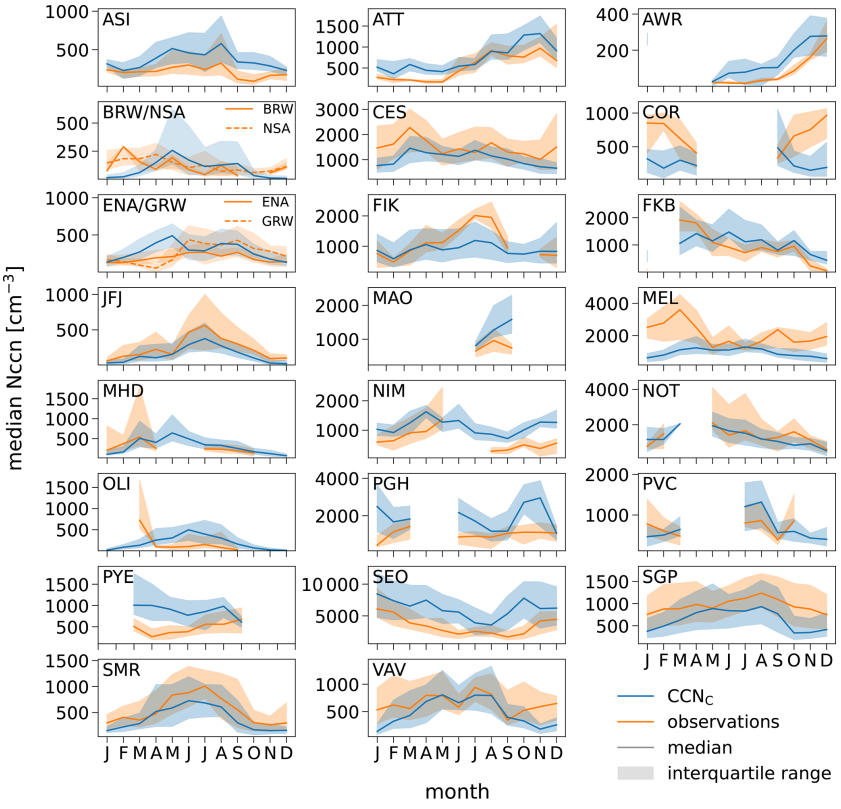


Figure A2. Climatologies for all stations at a supersaturation of 0.4 %. CCN_C in blue and CCN_O in orange, both showing monthly medians and interquartile ranges of 00:00-UTC daily values. Please note: The co-located data from BRW and NSA as well as ENA and GRW are shown in the same tile.

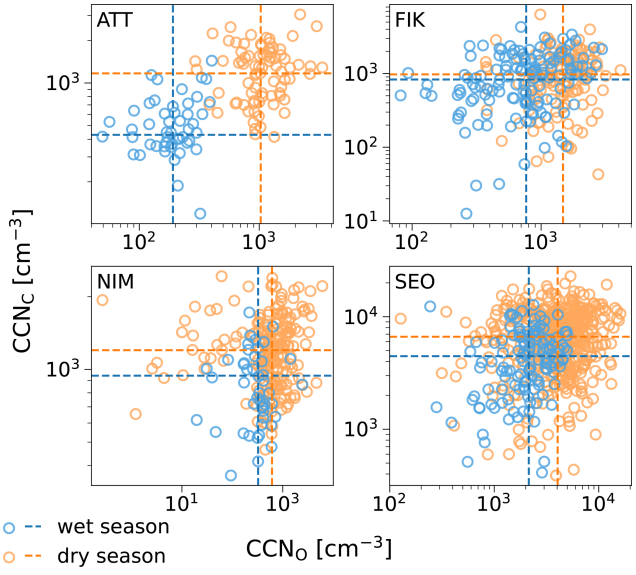


Figure A3. 00:00-UTC median Nccn at ATT, FIK, NIM and SEO. Days of wet season periods in blue, dry season in orange. Season medians shown as dashed lines. Please note, all tiles have different axis ranges.

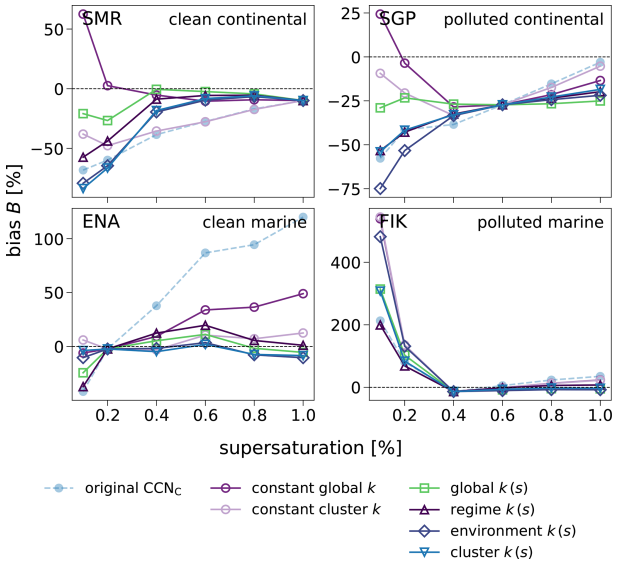


Figure A4. B–s spectrum at the SMR, SGP. ENA and FIK, showing the bias over the corresponding supersaturation for different parameterizations using the observation-derived slope parameter k. Please note, the y axis is different in each tile, the x axis is shared. We excluded the polar stations from the derivation of k and $k(s)$. For the constant k-parameterizations, we apply Eq. (3) with $s_1 = 0.1\,\%$ and $s_2 = 1.0\,\%$ for each individual station, resulting in one single-value k per station. (1) Constant global k is a weighted mean of environment dependent k values. (2) Constant cluster k represents the median k derived from each cluster and applied to the respective cluster. Then for each station, we use the median of the single-value k in their respective cluster. For the s-dependant parameterizations, we calculate $k(s)$ for each station. (3) Global $k(s)$ is a weighted mean of environment dependent $k(s)$ values. (4) The regime $k(s)$ represents the fitted median $k(s)$ derived for each pollution regime and applied to all stations of that regime. (5) Similarly, the environment $k(s)$ represents the fitted median $k(s)$ fore each environment, respectively, while (6) the cluster $k(s)$ represents the fitted median $k(s)$ for each cluster, respectively. All Nccn parameterizations are then retrieved from Eq. (8) with the respective $k(s)$ and the reference supersaturation s_i (clean continental: 1.0 %, polluted continental: 0.6 %, clean marine: 0.2 %, polluted marine: 0.4 %).

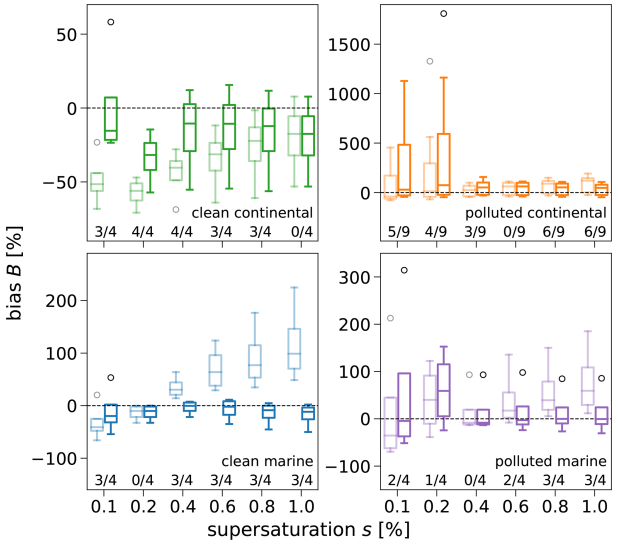


Figure A5. Statistics of the bias B at each supersaturation, depicted for clean continental (green), polluted continental (orange), clean marine (blue) and polluted marine (purple) cluster. Polar stations are excluded. CAMS-derived Nccn in light hue and the improved new CCN_C in dark hue. Each box plot includes the biases of all stations of the respective cluster and supersaturation. The numbers below show how many stations of the total in this category exhibit smaller absolute values of their bias after the improvement. Please note, the y axis is different in each tile, the x axis is shared and outliers are indicated by black circles.

Table A1. References to observation data and removed files during data checking. Please note: This table does not list missing files, but only the removed files during the process of data checking.

ID	Reference	Removed dates
ASI	Koontz et al. (2016b)	25 April 2016–31 May 2016, 16 June 2016, 20 August 2016, 16 September 2016–16 November 2016, 19 January 2017–27 January 2017, 21 August 2017
AWR	Koontz et al. (2016a)	2 June 2016, 5 October 2016, 28 November 2016, 2 January 2017–3 January 2017
COR	Koontz et al. (2018)	18 October 2018–22 October 2018, 11 February 2019–12 February 2019, 10 March 2019–13 March 2019, 31 March 2019–9 April 2019
ENA	Koontz et al. (2015)	28 September 2016–29 September 2016, 11 November 2016, 28 July 2017–29 July 2017, 21 March 2018–23 March 2018, 8 December 2018–2 April 2019, 22 September 2019–26 October 2020
NIM	Enekwizu et al. (2005a)	22 March 2006–23 March 2006, 4 May 2006–11 August 2006, 4 September 2006–7 September 2006
NSA	Andrews et al. (2007)	24 February 2008, 7 May 2008, 12 May 2008, 25 May 2008, 28 February 2009, 7 May 2009, 27 May 2009, 22 January 2010, 8 February 2010, 5 April 2011, 8 April 2011, 11 April 2011, 6 May 2011
OLI	Koontz et al. (2017a)	7 June 2019–25 June 2019, 4 August 2019, 6 August 2019
SGP	Enekwizu et al. (2006)	22 May 2007–13 October 2007, 18 July 2008, 5 January 2009, 6 July 2009–10 July 2009, 15 May 2009–3 August 2009
	Flynn et al. (2011)	30 November 2012–11 June 2013
	Koontz et al. (2011)	30 November 2012, 1 December 2012, 2 March 2013, 12 June 2013, 19 June 2013, 3 July 2013, 17 January 2016, 4 March 2016–10 May 2016, 7 December 2016
	Koontz et al. (2017b)	22 August 2017, 27 November 2017, 16 April 2018, 1 June 2018–7 November 2018, 18 March 2019–21 March 2019, 28 March 2019–1 April 2019, 21 June 2019–1 July 2019, 12 July 2019, 24 September 2019, 16 November 2019, 17 November 2020–25 November 2020, 8 June 2021–10 June 2021, 12 July 2021, 21 September 2021–24 September 2021, 30 September 2021–4 October 2021, 18 October 2021, 23 October 2021–28 October 2021, 8 December 2021

No dates needed to be removed for: ATT (Pöhlker et al., 2016) and BRW, CES, FIK, JFJ, MEL, MHD, NOT, SEO, SMR, VAV (Petäjä et al., 2017), and FKB (Enekwizu et al., 2007), PYE (Enekwizu et al., 2005b), GRW (Shilling and Kulkarni, 2009), MAO (Flynn et al., 2014), PGH (Shilling and Kulkarni, 2011), PVC (Shilling and Kulkarni, 2012).

Table A2. Regime, environment and for all 25 stations. The regime is derived from the 0.4-%-supersaturation median of (a) CCN_O and (b) CCN_C. Check marks indicate matching regimes in CCN_O and CCN_C. Furthermore, biases B [%] at at all six supersaturations s are shown, sorted by the best supersaturation at each station. The respective best bias is printed in bold font.

ID	CCN_O regime	CCN_C regime	Environment	$s = 0.1\,\%$	$s = 0.2\,\%$	$s = 0.4\,\%$	$s = 0.6\,\%$	$s = 0.8\,\%$	$s = 1.0\,\%$
AWR	clean	✓	marine	**+130.16**	+160.82	+201.19	+277.39	+248.46	+214.74
OLI	clean	✓	marine	**+75.42**	+82.79	+190.46	+215.50	+252.34	+248.53
PYE	clean	polluted	marine	**−11.14**	+122.25	+93.13	+135.59	+150.05	+185.06
COR	polluted	clean	continental	**−50.99**	−70.81	−68.71	−63.97	−60.93	−53.13
ASI	clean	✓	marine	−40.25	**−18.01**	+63.80	+123.84	+176.66	+224.82
ENA	clean	✓	marine	−41.77	**−2.27**	+37.82	+86.85	+94.45	+120.06
GRW	clean	✓	marine	+20.35	**−0.05**	+22.90	+40.87	+59.55	+77.63
PVC	polluted	✓	marine	−59.70	**−1.04**	−11.36	−7.99	+5.49	+11.68
SEO	polluted	✓	continental	−40.56	**+12.82**	+69.15	+109.02	+149.16	+191.96
MHD	clean	✓	marine	−65.89	−32.61	**+14.09**	+29.24	+34.77	+48.70
NOT	polluted	✓	marine	−69.62	−38.68	**−5.30**	+29.05	+55.14	+83.21
ATT	polluted	✓	continental	−61.38	−42.89	**+17.09**	+62.17	+104.84	+138.07
MAO	polluted	✓	continental	+98.53	+561.65	**+48.74**	+73.07	+89.73	+120.24
PGH	polluted	✓	continental	+171.28	+195.70	**+75.59**	+111.49	+117.43	+145.67
BRW	clean	✓	marine	−60.76	−37.68	−17.85	**−6.13**	+7.81	+24.49
FIK	polluted	✓	marine	+212.84	+81.04	−13.42	**+5.60**	+23.47	+35.25
FKB	polluted	✓	continental	+454.33	+1327.13	+25.19	**−1.24**	+13.96	+20.73
NIM	polluted	✓	continental	+353.68	+295.20	+99.45	**+98.76**	+132.38	+154.77
NSA	clean	✓	marine	−81.47	−44.78	−38.95	−15.92	**−5.29**	−6.66
VAV	polluted	clean	continental	−23.21	−47.04	−27.97	−11.82	**−1.56**	+7.71
JFJ	clean	✓	continental	−52.08	−52.27	−42.31	−35.15	−27.57	**−25.06**
SMR	polluted	clean	continental	−68.21	−59.84	−38.47	−27.50	−16.91	**−9.94**
CES	polluted	✓	continental	−78.32	−66.79	−40.60	−26.13	−9.77	**+0.78**
MEL	polluted	✓	continental	−75.98	−68.44	−45.02	−39.88	−30.28	**−25.07**
SGP	polluted	✓	continental	−57.71	−41.51	−38.46	−27.26	−15.32	**−3.04**

Table A3. Coefficients of the parameterizations we explored to improve the CAMS-derived Nccn spectrum. Cluster refers to the group of stations the respective parameterization should be applied to. A constant k is only given for parameterizations which rely on a single k value. The coefficients m and n are parameters of the linear function to fit the supersaturation range below 0.2 %: $k(s) = m\,s + n$. The coefficients a and b are parameters of the power function to fit the supersaturation range at and above 0.2 %: $k(s) = a\,s^b$. Thus, m, n, a, and b are only given for spectral $k(s)$-parameterizations.

Parameterization	Cluster	Constant k	m	n	a	b
constant global k	global	0.652950	–	–	–	–
constant cluster k	cle-con	1.071864	–	–	–	–
	pol-con	0.829455	–	–	–	–
	cle-mar	0.478487	–	–	–	–
	pol-mar	0.646918	–	–	–	–
global $k(s)$	global	–	−1.041286	1.069804	0.370341	−0.524598
regime $k(s)$	clean	–	−2.074001	1.442125	0.412154	−0.567478
	polluted	–	−0.707944	1.270485	0.503336	−0.501873
environment $k(s)$	continental	–	−2.337264	1.780602	0.452349	−0.662178
	marine	–	−0.501462	0.770894	0.336834	−0.427843
cluster $k(s)$	cle-con	–	−3.038368	1.955403	0.417673	−0.727884
	pol-con	–	−0.969338	1.305138	0.534918	−0.454285
	cle-mar	–	−0.667750	0.685041	0.350527	−0.281582
	pol-mar	–	0.262107	0.951143	0.392417	−0.583421

Data availability. The CAMS-derived CCN data (https://doi.org/10.26050/WDCC/QUAERERE_CCNCAMS_v1) from Block (2023) that is being evaluated here, is archived at the World Data Center for Climate (WDCC) at DKRZ and can be downloaded from https://www.wdc-climate.de/ui/entry?acronym=QUAERERE_CCNCAMS_v1 (last access: 23 June 2026).

An overview of the 25 evaluation datasets can be found in Tables 1 and A1, with references therein.

10 of the 25 CCN datasets providing CCN data at various supersaturations – Cabauw (CES), Finokalia (FIK), Jungfraujoch (JFJ), Melpitz (MEL), Mace Head (MHD), Hyytiälä (SMR), Vavihill (VAV), Utqiaġvik – formerly Barrow (BRW), Noto (NOT), Seoul (SEO) – were obtained from https://doi.org/10.6084/M9.FIGSHARE.C.3471585.V1 (Petäjä et al., 2017).

CCN data for 14 of our 25 stations – Ascension Island (ASI), McMurdo Station (Antarctica, AWR), Cordoba (COR), Graciosa Island (ENA and W, covering two time periods), Black Forest (FKB), Manacapuru (MAO), Niamey (NIM), North Slope of Alaska (NSA, same location as BRW), Oliktok Point (OLI), Nainital (PGH), Cape Cod (PVC), Point Reyes (PYE) and Southern Great Plains (SGP) – were provided by the United States' Department of Energy's Atmospheric Radiation Measurement (ARM, https://www.arm.gov, last access: 23 June 2026) user facility.

The data for ATTO Observatory (ATT) was directly provided by Mira Pöhlker (Pöhlker et al., 2016). The Supplement related to this article is available online at https://doi.org/10.5194/acp-16-15709-2016-supplement.

Author contributions. YEA and KB downloaded, prepared and processed observational as well as model data. MP provided measurement data for ATT. KB and JQ designed this study and provided scientific supervision. YEA conducted the statistical analysis, created plots and tables and drafted the initial manuscript. YEA and KB prepared, revised and submitted the final manuscript with contributions from all co-authors. All authors edited the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. The authors acknowledge data provision and data quality reporting from the Atmospheric Radiation Measurement (ARM; https://www.arm.gov, last access: 23 June 2026) User Facility, a US Department of Energy (DOE) Office of Science user facility managed by the Biological and Environmental Research Program. The authors also wish to acknowledge the ACTRIS (Aerosol, Clouds, and Trace Gases Research Infrastructure; http://www.actris.eu/, last access: 23 June 2026) for its essential role in providing access to high-quality data, facilities, and tech-

nical support that were integral to the execution of this research. We further would like to acknowledge the World Data Centre for Climate (WDCC) for providing the CAMS-derived CCN data (https://doi.org/10.26050/WDCC/QUAERERE_CCNCAMS_v1).

The authors would like to thank the reviewers Christina Williamson and Ying Chen for their valuable comments and suggestions.

Financial support. Karoline Block and Johannes Quaas acknowledge funding by the "WarmWorld" project (FKZ 01LK2202G) funded by the German Federal Ministry of Research, Technology and Space. Johannes Quaas further acknowledges funding by the European Commision Horizon Europe programme grant "CleanCloud" (grant no. 101137639).

Review statement. This paper was edited by Fiona O'Connor and reviewed by Christina Williamson and Ying Chen.

References

Anders, Y., Block, K., and Quaas, J.: Validation of CAMS derived CCN variability with in-situ observations, Meteorologische Arbeiten (XXX) und Jahresbericht 2024 des Instituts für Meteorologie der Universität Leipzig, 11–26, https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-1001574 (last access: 23 June 2026), 2025.

Andreae, M. O.: Correlation between cloud condensation nuclei concentration and aerosol optical thickness in remote and polluted regions, Atmos. Chem. Phys., 9, 543–556, https://doi.org/10.5194/acp-9-543-2009, 2009.

Andrews, E., Enekwizu, O., Laskin, A., Hayes, C., and Uin, J.: Cloud Condensation Nuclei Particle Counter (AOSCCN), 2007-03-02 to 2011-08-30, North Slope Alaska (NSA), External Data (satellites and others) (X1), ARM [data set], https://doi.org/10.5439/1988402, 2007.

Andrews, E., Zabala, I., Carrillo-Cardenas, G., Titos, G., Casquero-Vera, J. A., and Hallar, A. G.: Harmonized aerosol size distribution, cloud condensation nuclei, chemistry and optical properties at 10 sites, Scientific Data, 12, 937, https://doi.org/10.1038/s41597-025-04931-y, 2025.

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.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res.-Atmos., 114, https://doi.org/10.1029/2008JD011115, 2009.

Block, K.: Cloud condensation nuclei (CCN) numbers derived from CAMS reanalysis EAC4 (Version 1), World Data Center for Climate (WDCC) at DKRZ [data set], https://doi.org/10.26050/WDCC/QUAERERE_CCNCAMS_V1, 2023.

Block, K., Haghighatnasab, M., Partridge, D. G., Stier, P., and Quaas, J.: Cloud condensation nuclei concentrations derived from the CAMS reanalysis, Earth Syst. Sci. Data, 16, 443–470, https://doi.org/10.5194/essd-16-443-2024, 2024.

Che, H., Stier, P., Watson-Parris, D., Gordon, H., and Deaconu, L.: Source attribution of cloud condensation nuclei and their impact on stratocumulus clouds and radiation in the south-eastern Atlantic, Atmos. Chem. Phys., 22, 10789–10807, https://doi.org/10.5194/acp-22-10789-2022, 2022.

Choudhury, G. and Tesche, M.: A first global height-resolved cloud condensation nuclei data set derived from spaceborne lidar measurements, Earth Syst. Sci. Data, 15, 3747–3760, https://doi.org/10.5194/essd-15-3747-2023, 2023.

Choudhury, G., Block, K., Haghighatnasab, M., Quaas, J., Goren, T., and Tesche, M.: Pristine oceans are a significant source of uncertainty in quantifying global cloud condensation nuclei, Atmos. Chem. Phys., 25, 3841–3856, https://doi.org/10.5194/acp-25-3841-2025, 2025.

Cohard, J.-M., Pinty, J.-P., and Bedos, C.: Extending Twomey's analytical estimate of nucleated cloud droplet concentrations from CCN spectra, J. Atmos. Sci., 55, 3348–3357, https://doi.org/10.1175/1520-0469(1998)055<3348:ETSAEO>2.0.CO;2, 1998.

Deng, Z. Z., Zhao, C. S., Ma, N., Ran, L., Zhou, G. Q., Lu, D. R., and Zhou, X. J.: An examination of parameterizations for the CCN number concentration based on in situ measurements of aerosol activation properties in the North China Plain, Atmos. Chem. Phys., 13, 6227–6237, https://doi.org/10.5194/acp-13-6227-2013, 2013.

Désalmand, F.: An attempt to characterize a relationship between supersaturation spectrum, size spectrum and solubility of CCN observed over a tropical rain forest, J. Atmos. Sci., 42, 472–477, https://doi.org/10.1175/1520-0469(1985)042<0472:AATCAR>2.0.CO;2, 1985.

Enekwizu, O., Hayes, C., and Uin, J.: Cloud Condensation Nuclei Particle Counter (AOSCCN), 2005-12-21 to 2007-01-07, ARM Mobile Facility (NIM), Niamey, Niger (M1), ARM [data set], https://doi.org/10.5439/1988402, 2005a.

Enekwizu, O., Hayes, C., and Uin, J.: Cloud Condensation Nuclei Particle Counter (AOSCCN), 2005-03-04 to 2005-09-15, ARM Mobile Facility (PYE), Point Reyes, CA (M1), ARM [data set], https://doi.org/10.5439/1988402, 2005b.

Enekwizu, O., Laskin, A., Hayes, C., and Uin, J.: Cloud Condensation Nuclei Particle Counter (AOSCCN), 2006-09-13 to 2011-03-09, Southern Great Plains (SGP), Central Facility, Lamont, OK (C1), ARM [data set], https://doi.org/10.5439/1988402, 2006.

Enekwizu, O., Hayes, C., and Uin, J.: Cloud Condensation Nuclei Particle Counter (AOSCCN), 2007-03-19 to 2008-01-03, ARM Mobile Facility (FKB), Black Forest, Germany (M1), ARM [data set], https://doi.org/10.5439/1988402, 2007.

Flynn, C., Shilling, J., and Kulkarni, G.: Aerosol Observing System (AOS): cloud condensation nuclei data, averaged (AOSCCNAVG), 2011-01-01 to 2014-09-29, Southern Great Plains (SGP), Central Facility, Lamont, OK (C1), ARM [data set], https://doi.org/10.5439/1985586, 2011.

Flynn, C., Shilling, J., and Kulkarni, G.: Aerosol Observing System (AOS): cloud condensation nuclei data, averaged (AOSCCNAVG), 2014-07-01 to 2014-09-29, ARM Mobile Facility (MAO), Manacapuru, Amazonas, Brazil; AMF1 (M1), ARM [data set], https://doi.org/10.5439/1985586, 2014.

Galy-Lacaux, C. and Modi, A. I.: Precipitation Chemistry in the Sahelian Savanna of Niger, Africa, J. Atmos. Chem., 30, 319–343, https://doi.org/10.1023/A:1006027730377, 1998.

Gryspeerdt, E. and Stier, P.: Regime-based analysis of aerosol-cloud interactions, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL053221, 2012.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Holanda, B. A., Pöhlker, M. L., Walter, D., Saturno, J., Sörgel, M., Ditas, J., Ditas, F., Schulz, C., Franco, M. A., Wang, Q., Donth, T., Artaxo, P., Barbosa, H. M. J., Borrmann, S., Braga, R., Brito, J., Cheng, Y., Dollner, M., Kaiser, J. W., Klimach, T., Knote, C., Krüger, O. O., Fütterer, D., Lavrič, J. V., Ma, N., Machado, L. A. T., Ming, J., Morais, F. G., Paulsen, H., Sauer, D., Schlager, H., Schneider, J., Su, H., Weinzierl, B., Walser, A., Wendisch, M., Ziereis, H., Zöger, M., Pöschl, U., Andreae, M. O., and Pöhlker, C.: Influx of African biomass burning aerosol during the Amazonian dry season through layered transatlantic transport of black carbon-rich smoke, Atmos. Chem. Phys., 20, 4757–4785, https://doi.org/10.5194/acp-20-4757-2020, 2020.

Hudson, J. G.: Relationship between fog condensation nuclei and fog microstructure, J. Atmos. Sci., 37, 1854–1867, https://doi.org/10.1175/1520-0469(1980)037<1854:RBFCNA>2.0.CO;2, 1980.

Hudson, J. G. and Yum, S. S.: Cloud condensation nuclei spectra and polluted and clean clouds over the Indian Ocean, J. Geophys. Res.-Atmos., 107, https://doi.org/10.1029/2001JD000829, 2002.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

IPCC – Intergovernmental Panel on Climate Change: The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity, Cambridge University Press, 923–1054, https://doi.org/10.1017/9781009157896.009, 2023.

Jayachandran, V. N., Varghese, M., Murugavel, P., Todekar, K. S., Bankar, S. P., Malap, N., Dinesh, G., Safai, P. D., Rao, J., Konwar, M., Dixit, S., and Prabha, T. V.: Cloud condensation nuclei characteristics during the Indian summer monsoon over a rain-shadow region, Atmos. Chem. Phys., 20, 7307–7334, https://doi.org/10.5194/acp-20-7307-2020, 2020.

Ji, Q. and Shaw, G. E.: On supersaturation spectrum and size distributions of cloud condensation nuclei, Geophys. Res. Lett., 25, 1903–1906, https://doi.org/10.1029/98GL01404, 1998.

Jia, H., Ma, X., Yu, F., and Quaas, J.: Significant underestimation of radiative forcing by aerosol–cloud interactions derived from satellite-based methods, Nat. Commun., 12, 3649, https://doi.org/10.1038/s41467-021-23888-1, 2021.

Khvorostyanov, V. I. and Curry, J. A.: A simple analytical model of aerosol properties with account for hygroscopic growth: 1. Equilibrium size spectra and cloud condensation nuclei activity spectra, J. Geophys. Res.-Atmos., 104, 2175–2184, https://doi.org/10.1029/98JD02673, 1999.

Khvorostyanov, V. I. and Curry, J. A.: Aerosol size spectra and CCN activity spectra: Reconciling the lognormal, algebraic, and power laws, J. Geophys. Res.-Atmos., 111, 2005JD006532, https://doi.org/10.1029/2005JD006532, 2006.

Khvorostyanov, V. I. and Curry, J. A.: Kinetics of cloud drop formation and its parameterization for cloud and climate models, J. Atmos. Sci., 65, 2784–2802, https://doi.org/10.1175/2008JAS2606.1, 2008.

Kim, J. H., Yum, S. S., Shim, S., Yoon, S.-C., Hudson, J. G., Park, J., and Lee, S.-J.: On aerosol hygroscopicity, cloud condensation nuclei (CCN) spectra and critical supersaturation measured at two remote islands of Korea between 2006 and 2009, Atmos. Chem. Phys., 11, 12627–12645, https://doi.org/10.5194/acp-11-12627-2011, 2011.

Koontz, A., Flynn, C., Andrews, E., Uin, J., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN1COLAVG), 2011-08-10 to 2017-08-16, Southern Great Plains (SGP), Central Facility, Lamont, OK (C1), ARM [data set], https://doi.org/10.5439/1255094, 2011.

Koontz, A., Flynn, C., Andrews, E., Uin, J., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN1COLAVG), 2015-01-01 to 2020-10-28, Eastern North Atlantic (ENA), Graciosa Island, Azores, Portugal (C1), ARM [data set], https://doi.org/10.5439/1255094, 2015.

Koontz, A., Flynn, C., Andrews, E., Uin, J., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN1COLAVG), 2016-05-29 to 2017-01-03, ARM Mobile Facility (AWR), McMurdo Station Ross Ice Shelf, Antarctica; AMF2 (M1), ARM [data set], https://doi.org/10.5439/1255094, 2016a.

Koontz, A., Uin, J., Andrews, E., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN2COLAAVG), 2016-04-25 to 2017-10-30, ARM Mobile Facility (ASI), Ascension Island, South Atlantic Ocean; AMF1 (M1), ARM [data set], https://doi.org/10.5439/1323894, 2016b.

Koontz, A., Uin, J., Andrews, E., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN2COLAAVG), 2017-04-20 to 2019-09-06, Oliktok Point, AK, USA; Long-term Mobile Facility (OLI), Oliktok Point, Alaska; AMF3 (M1), ARM [data set], https://doi.org/10.5439/1323894, 2017a.

Koontz, A., Uin, J., Andrews, E., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN2COLAAVG), 2017-04-12 to 2026-01-21, Southern Great Plains (SGP), Lamont, OK (Extended and Co-located with C1) (E13), ARM [data set], https://doi.org/10.5439/1323894, 2017b.

Koontz, A., Uin, J., Andrews, E., Enekwizu, O., Hayes, C., and Salwen, C.: Cloud Condensation Nuclei Particle Counter (AOSCCN2COLAAVG), 2018-09-24 to 2019-04-30, ARM Mobile Facility (COR), Cordoba, Argentina; AMF1 (main site for CACTI) (M1), ARM [data set], https://doi.org/10.5439/1323894, 2018.

Kouvarakis, G., Tsigaridis, K., Kanakidou, M., and Mihalopoulos, N.: Temporal variations of surface regional background ozone over Crete Island in the southeast Mediterranean, J. Geophys. Res.-Atmos., 105, 4399–4407, https://doi.org/10.1029/1999JD900984, 2000.

Kulkarni, G., Mei, F., Sivaraman, C., Wang, J., Shilling, J. E., Newsom, R. K., Christensen, M. W., Berg, L. K., and Fast, J. D.: Assessment of extinction-, satellite-, and model-based vertical cloud condensation nuclei (CCN) retrieval methods using airborne CCN measurements over the southern Great Plains, J. Geophys. Res.-Atmos., 130, e2024JD042565, https://doi.org/10.1029/2024JD042565, 2025.

Lance, S., Nenes, A., Medina, J., and Smith, J. N.: Mapping the Operation of the DMT Continuous Flow CCN Counter, Aerosol Sci. Tech., 40, 242–254, https://doi.org/10.1080/02786820500543290, 2006.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989–3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Liu, J., Zheng, Y., Li, Z., and Cribb, M.: Analysis of cloud condensation nuclei properties at a polluted site in southeastern China during the AMF-China Campaign, J. Geophys. Res.-Atmos., 116, https://doi.org/10.1029/2011JD016395, 2011.

Martins, J. A., Gonçalves, F. L. T., Morales, C. A., Fisch, G. F., Pinheiro, F. G. M., Leal Júnior, J. B. V., Oliveira, C. J., Silva, E. M., Oliveira, J. C. P., Costa, A. A., and Silva Dias, M. A. F.: Cloud condensation nuclei from biomass burning during the Amazonian dry-to-wet transition season, Meteorol. Atmos. Phys., 104, 83–93, https://doi.org/10.1007/s00703-009-0019-6, 2009.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, J. Geophys. Res.-Atmos., 114, https://doi.org/10.1029/2008JD011235, 2009.

Petäjä, T., Kulmala, M., Ovadnevaite, J., O'Dowd, C. D., Matsuki, A., Baltensperger, U., Herrmann, E., Bukowiecki, N., Kiendler-Scharr, A., Ogren, J., Prévôt, A. S. H., Slowik, J. G., Kanakidou, M., Wiedensohler, A., Nenes, A., Kristensson, A., Iwamoto, Y., Ehn, M., Liine Heikkinen, Schmale, J., Henning, S., Henzing, J. S., Keskinen, H., Sellegri, K., Bougiatioti, A., Kalivitis, N., Stavroulas, I., Jefferson, A., Minsu Park, Schlag, P., Pringle, K., Reddington, C., Aalto, P., Äijälä, M., Birmili, W., Paramonov, M., Fjæraa, A. M., Fiebig, M., Myhre, C. L., Frank, G., Swietlicki, E., Svennigsson, B., Cerina Wittbom, Fröhlich, R., Hammer, E., Gysel, M., Frumau, A., Kos, G., Kinochi, K., Ruper Holzinger, Mihalopoulos Nikolaos, Piccard, D., Poulain, L., Sonntag, A., Stratmann, F., Yum, S. S., Carslaw, K., Furuya, M., Hyono, H., and Tsurumaru, H.: Data from collocated observations of cloud condensation nuclei, particle size distributions, and chemical composition, figshare [data set], https://doi.org/10.6084/M9.FIGSHARE.C.3471585.V1, 2017.

Pöhlker, M. L., Pöhlker, C., Ditas, F., Klimach, T., Hrabe de Angelis, I., Araújo, A., Brito, J., Carbone, S., Cheng, Y., Chi, X., Ditz, R., Gunthe, S. S., Kesselmeier, J., Könemann, T., Lavrič, J. V., Martin, S. T., Mikhailov, E., Moran-Zuloaga, D., Rose, D., Saturno, J., Su, H., Thalman, R., Walter, D., Wang, J., Wolff, S., Barbosa, H. M. J., Artaxo, P., Andreae, M. O., and Pöschl, U.: Long-term observations of cloud condensation nuclei in the Amazon rain forest – Part 1: Aerosol size distribution, hygroscopicity, and new model parametrizations for CCN prediction, Atmos. Chem. Phys., 16, 15709–15740, https://doi.org/10.5194/acp-16-15709-2016, 2016.

Popp, T., De Leeuw, G., Bingen, C., Brühl, C., Capelle, V., Chedin, A., Clarisse, L., Dubovik, O., Grainger, R., Griesfeller, J., Heckel, A., Kinne, S., Klüser, L., Kosmale, M., Kolmonen, P., Lelli, L., Litvinov, P., Mei, L., North, P., Pinnock, S., Povey, A., Robert, C., Schulz, M., Sogacheva, L., Stebel, K., Stein Zweers, D., Thomas, G., Tilstra, L., Vandenbussche, S., Veefkind, P., Vountas, M., and Xue, Y.: Development, production and evaluation of aerosol climate data records from European satellite observations (Aerosol_cci), Remote Sens.-Basel, 8, 421, https://doi.org/10.3390/rs8050421, 2016.

Pruppacher, H. R. and Klett, J. D.: Microphysics of clouds and precipitation, no. 18 in Atmospheric and oceanographic sciences library, in: 2nd rev. and enl. edn., Kluwer Academic Publishers, ISBN 9780792342113, 1997.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristjánsson, J. E., Liu, X., Balkanski, Y., Donner, L. J., Ginoux, P. A., Stier, P., Grandey, B., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevåg, A., Iversen, T., Seland, Ø., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J.-F., Iacono, M. J., Kinne, S., and Schulz, M.: Aerosol indirect effects – general circulation model intercomparison and evaluation with satellite data, Atmos. Chem. Phys., 9, 8697–8717, https://doi.org/10.5194/acp-9-8697-2009, 2009.

Quaas, J., Arola, A., Cairns, B., Christensen, M., Deneke, H., Ekman, A. M. L., Feingold, G., Fridlind, A., Gryspeerdt, 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.

Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M., Bozzo, A., Ades, M., Huijnen, V., Benedetti, A., Engelen, R., Peuch, V.-H., and Morcrette, J.-J.: Description and evaluation of the tropospheric aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS-AER, cycle 45R1), Geosci. Model Dev., 12, 4627–4659, https://doi.org/10.5194/gmd-12-4627-2019, 2019.

Reutter, P., Su, H., Trentmann, J., Simmel, M., Rose, D., Gunthe, S. S., Wernli, H., Andreae, M. O., and Pöschl, U.: Aerosol- and updraft-limited regimes of cloud droplet formation: influence of

particle number, size and hygroscopicity on the activation of cloud condensation nuclei (CCN), Atmos. Chem. Phys., 9, 7067–7080, https://doi.org/10.5194/acp-9-7067-2009, 2009.

Roberts, G. C. and Nenes, A.: A continuous-flow streamwise thermal-gradient CCN chamber for atmospheric measurements, Aerosol Sci. Tech., 39, 206–221, https://doi.org/10.1080/027868290913988, 2005.

Rose, D., Gunthe, S. S., Mikhailov, E., Frank, G. P., Dusek, U., Andreae, M. O., and Pöschl, U.: Calibration and measurement uncertainties of a continuous-flow cloud condensation nuclei counter (DMT-CCNC): CCN activation of ammonium sulfate and sodium chloride aerosol particles in theory and experiment, Atmos. Chem. Phys., 8, 1153–1179, https://doi.org/10.5194/acp-8-1153-2008, 2008.

Schmale, J., Henning, S., Henzing, B., Keskinen, H., Sellegri, K., Ovadnevaite, J., Bougiatioti, A., Kalivitis, N., Stavroulas, I., Jefferson, A., Park, M., Schlag, P., Kristensson, A., Iwamoto, Y., Pringle, K., Reddington, C., Aalto, P., Äijälä, M., Baltensperger, U., Bialek, J., Birmili, W., Bukowiecki, N., Ehn, M., Fjæraa, A. M., Fiebig, M., Frank, G., Fröhlich, R., Frumau, A., Furuya, M., Hammer, E., Heikkinen, L., Herrmann, E., Holzinger, R., Hyono, H., Kanakidou, M., Kiendler-Scharr, A., Kinouchi, K., Kos, G., Kulmala, M., Mihalopoulos, N., Motos, G., Nenes, A., O'Dowd, C., Paramonov, M., Petäjä, T., Picard, D., Poulain, L., Prévôt, A. S. H., Slowik, J., Sonntag, A., Swietlicki, E., Svenningsson, B., Tsurumaru, H., Wiedensohler, A., Wittbom, C., Ogren, J. A., Matsuki, A., Yum, S. S., Myhre, C. L., Carslaw, K., Stratmann, F., and Gysel, M.: Collocated observations of cloud condensation nuclei, particle size distributions, and chemical composition, Scientific Data, 4, 170003, https://doi.org/10.1038/sdata.2017.3, 2017.

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.

Shilling, J. and Kulkarni, G.: Aerosol Observing System (AOS): cloud condensation nuclei data, averaged (AOSCCNAVG), 2009-04-16 to 2010-12-30, ARM Mobile Facility (GRW), Graciosa Island, Azores, Portugal; AMF1 (M1), ARM [data set], https://doi.org/10.5439/1985586, 2009.

Shilling, J. and Kulkarni, G.: Aerosol Observing System (AOS): cloud condensation nuclei data, averaged (AOSCCNAVG), 2011-06-09 to 2012-03-26, ARM Mobile Facility (PGH), ARIES Observatory, Nainital, Uttarkhand, India; AMF1 (M1), ARM [data set], https://doi.org/10.5439/1985586, 2011.

Shilling, J. and Kulkarni, G.: Aerosol Observing System (AOS): cloud condensation nuclei data, averaged (AOSCCNAVG), 2012-07-16 to 2013-03-30, ARM Mobile Facility (PVC), Highland Center, Cape Cod MA; AMF1 (M1), ARM [data set], https://doi.org/10.5439/1985586, 2012.

Spracklen, D. V., Carslaw, K. S., Pöschl, U., Rap, A., and Forster, P. M.: Global cloud condensation nuclei influenced by carbonaceous combustion aerosol, Atmos. Chem. Phys., 11, 9067–9087, https://doi.org/10.5194/acp-11-9067-2011, 2011.

Stier, P.: Limitations of passive remote sensing to constrain global cloud condensation nuclei, Atmos. Chem. Phys., 16, 6595–6607, https://doi.org/10.5194/acp-16-6595-2016, 2016.

Twomey, S.: The nuclei of natural cloud formation part II: The supersaturation in natural clouds and the variation of cloud droplet concentration, Geofisica pura e applicata, 43, 243–249, https://doi.org/10.1007/BF01993560, 1959.

Varghese, M., Prabha, T. V., Malap, N., Resmi, E., Murugavel, P., Safai, P., Axisa, D., Pandithurai, G., and Dani, K.: Airborne and ground based CCN spectral characteristics: inferences from CAIPEEX – 2011, Atmos. Environ., 125, 324–336, https://doi.org/10.1016/j.atmosenv.2015.06.041, 2016.

Wang, Y., Niu, S., Lu, C., Fan, S., Lv, J., Xu, X., Jin, Y., and Sun, W.: A new CCN activation parameterization and its potential influences on aerosol indirect effects, Atmos. Res., 253, 105491, https://doi.org/10.1016/j.atmosres.2021.105491, 2021.

Yum, S. S. and Hudson, J. G.: Vertical distributions of cloud condensation nuclei spectra over the springtime Arctic Ocean, J. Geophys. Res.-Atmos., 106, 15045–15052, https://doi.org/10.1029/2000JD900357, 2001.