Estimating CCN number concentrations using aerosol optical properties:

Role of particle number size distribution and parameterization

Yicheng Shen1,2, Aki Virkkula2,1,1, Aijun Ding1, Krista Luoma2, Helmi Keskinen2,4, Pasi P. Aalto2,

Xuguang Chi1, Ximeng Qi1, Wei Nie1, Xin Huang1,

Tuukka Petäjä2,1, Markku Kulmala2, Veli-Matti Kerminen2

1Joint International Research Laboratory of Atmospheric Sciences, School of Atmospheric Sciences,
Nanjing University, Nanjing, 210023, China
3Institute for Atmospheric and Earth System Research, University of Helsinki, Helsinki, Finland
3Finnish Meteorological Institute, Helsinki, Finland
4Hyytiälä Forestry Field Station, Hyyttiäntie 124, Korkeakoski FI 35500, Finland

Keywords: CCN, scattering coefficient, backscatter fraction, Ångström exponent

Abstract

The concentration of cloud condensation nuclei (CCN) is an essential parameter affecting
aerosol-cloud interactions within warm clouds. Long-term CCN number concentration (N_{CCN})
data are scarce, there are a lot more data on aerosol optical properties (AOPs). It is therefore
valuable to derive parameterizations for estimating N_{CCN} from AOP measurements. Such
parameterizations have been made earlier, in the present work a new one is presented. The
relationships between AOPs, N_{CCN} and particle number size distributions were investigated
based on in-situ measurement data from six stations in very different environments around the
world. The parameterization derived here depends on the scattering Ångström exponent (SAE),
backscatter fraction (BSF) and total scattering coefficient ($\sigma_{sp}$) of PM10 particles. The analysis
showed that the dependence of N_{CCN} on supersaturation SS% is logarithmic:

$$N_{CCN} \approx ((287 \pm 45) \times SAE_{10} \ln(SS\%/(0.093 \pm 0.006))(BSF - BSF_{min}) + (5.2 \pm 3.3)) \sigma_{sp}.$$  

At the lowest supersaturations of each site (SS% \approx 0.1) the average bias, defined as the ratio of the
AOP-derived and measured N_{CCN} varied from ~0.7 to ~1.5 at most sites except at a Himalayan site
where bias was > 4. At SS% > 0.3 the average bias ranged from ~0.7 to ~1.3 at all sites. In other
words, at SS% > 0.3 \( N_{CCN} \) was estimated with an average uncertainty of approximately 30% by using nephelometer data. The squared correlation coefficients between the AOP-derived and measured \( N_{CCN} \) varied from ~0.5 to ~0.8. The coefficients of the parameterization derived for the different sites were linearly related to each other. To study the explanation of this, lognormal unimodal particle size distributions were generated and \( N_{CCN} \) and AOPs were calculated. The simulation yielded similar relationships between the coefficients as in the field data. It also showed that the relationships of the coefficients are affected by the geometric mean diameter and width of the size distribution and the activation diameter.

1. Introduction

Aerosol-cloud interactions (ACI) are the most significant sources of uncertainty in estimating the radiative forcing of the Earth’s climate system (e.g., Forster et al., 2007; Kerminen et al., 2012), which makes it more challenging to predict the future climate change (Schwartz et al., 2010). An essential parameter affecting ACI within warm clouds is cloud condensation nuclei (CCN) concentration, the number concentration of particles capable of initiating cloud droplet formation at a given supersaturation. Determining CCN concentrations and their temporal and spatial variations is one of the critical aspects to reduce such uncertainty.

CCN number concentrations (\( N_{CCN} \)) have been measured at different locations worldwide (e.g., Twomey, 1959; Hudson,1993; Kulmala et al., 1993; Hämeri et al., 2001; Sihto et al., 2011; Pöhlker et al., 2016; Ma et al., 2014). However, the accessible data especially for long-term measurement is still limited in the past and nowadays due to the relatively higher cost of instrumentation and the complexity of long-term operating. As an alternative to direct measurement, \( N_{CCN} \) can also be estimated from particle number size distributions and chemical composition using the Köhler equation. Several studies have investigated the relative importance of the chemical composition and particle number distributions (Dusek et al., 2006; Ervens et al., 2007; Hudson, 2007; Crosbie et al., 2015). For the best of our understanding, the particle number size distributions are more important in determining \( N_{CCN} \) than aerosol chemical composition. This makes particle number size distribution measurements capable of
serving as a supplementary of direct CCN measurements.

Considering the tremendous spatiotemporal heterogeneity of atmospheric aerosol, neither direct measurements of N\textsubscript{CCN} of the concentrations estimated from particle size distribution are adequate for climate research. In order to overcome the limitation of current measurements, many studies have attempted to estimate N\textsubscript{CCN} using aerosol optical properties (AOPs) (e.g., Ghan et al., 2006; Shinozuka et al., 2009; Andreae, 2009; Jefferson, 2010; Liu et al., 2014; Shinozuka et al., 2015; Tao et al., 2018). Most of these studies attempted to link N\textsubscript{CCN} with extensive AOPs, such as the aerosol extinction coefficient (\(\sigma_{\text{ext}}\)), aerosol scattering coefficient (\(\sigma_{\text{sp}}\)) and aerosol optical depth (AOD). Both N\textsubscript{CCN} and \(\sigma_{\text{sp}}\) are extensive properties that vary with a varying aerosol loading. The most straightforward approach to estimate CCN is to utilize the ratio between CCN and one of the extensive AOPs (e.g. AOD, \(\sigma_{\text{ext}}, \sigma_{\text{sp}}\)). However, the ratio is not a constant. Previous studies have also pointed out that the relationship between N\textsubscript{CCN} and extensive AOPs are nonlinear. On one hand, Andreae (2009) reported that the relationship between AOD and CCN number concentration at the supersaturation of 0.4% (CCN\textsubscript{0.4}) can be written as AOD\textsubscript{500}=0.0027\cdot(CCN\textsubscript{0.4})^{0.640}, which indicates AOT and CCN depend in a non-linear way on each other: for a larger AOD there are more CCN per-unit change in AOD. On the other hand, Shinozuka et al. (2015) indicated that the larger the extinction coefficient \(\sigma_{\text{ext}}\) was, the fewer CCN were per unit change of \(\sigma_{\text{ext}}\).

Some studies have also involved intensive aerosol optical properties, such as the scattering Ångström exponent (SAE), hemispheric backscattering fraction (BSF) and single-scattering albedo (SSA) to build up a bridge between the N\textsubscript{CCN} and AOPs. Jefferson (2010) used BSF and SSA to parameterize the coefficients \(C\) and \(k\) to present N\textsubscript{CCN}(SS\%) = \(C\times(SS\%)^k\), where SS\% is the supersaturation percent (Twomey, 1959). Liu and Li (2014) discussed how different aerosol properties affect the ratio of N\textsubscript{CCN} to \(\sigma_{\text{sp}}\), i.e., \(R_{\text{CCN}}/\sigma_{\text{sp}}\) based on in-situ and remote-sensing data. Shinozuka et al. (2015) used SAE and aerosol extinction coefficient to estimate N\textsubscript{CCN}. Tao et al. (2018) used a novel method to derive the ratio \(R_{\text{CCN}}/\sigma_{\text{sp}}\) which they named as AR\textsubscript{sp}, based on SAE and aerosol hygroscopicity using a humidified nephelometer. All the
studies mentioned above noted that the particle number size distribution (PNSD) plays an important role in estimating $N_{CCN}$ from aerosol optical properties.

In this study, we will introduce a new approach to estimate $N_{CCN}$, along with a brief discussion on how the ratio between $N_{CCN}$ and $\sigma_{sp}$ is related to BSF. The AOPs needed in our estimation are $\sigma_{sp}$, BSF and SAE obtained using a 3-wavelength nephelometer, either the TSI 3563 or Ecotech Aurora 3000. The main goal of this study is to provide a parameterization for calculating $N_{CCN}$ using AOPs, and to probe the physical explanations behind this parameterization. The method will be applied to six different sites worldwide.

2. Methods

2.1 Sites and measurements

In-situ measurements of AOPs, PNSDs, and $N_{CCN}$ were conducted at SMEAR II in Finland, SORPES in China, and 4 ARM Climate Research Facility (ACRF) sites (Mather and Voyles, 2013). The locations and measurement periods are listed in Table 1.

The Station for Measuring Forest Ecosystem-Atmosphere Relations (SMEAR II) is located at the Hyytiälä Forestry Field Station (61°51’ N, 24°17’ E, 181 m above sea level) of University of Helsinki, 60 km north-east from the nearest city. The station represents boreal coniferous forest, which covers ~8% of the Earth’s surface. Total scattering coefficient ($\sigma_{sp}$) and hemispheric backscattering coefficient ($\sigma_{bsp}$) of sub-1 µm and sub-10 µm particles are measured using a TSI-3563 3-wavelength integrating nephelometer at $\lambda = 450, 550,$ and 700 nm. The calibration, data processing, and calculation of AOPs followed the procedure described by Virkkula et al. (2011) and Luoma et al. (2018). $N_{CCN}$ was measured at supersaturations (SS%) of 0.1%, 0.2%, 0.3%, 0.5% and 1.0% using a DMT CCN-100 CCN counter, likewise in Schmale et al. (2017). A whole measurement cycle takes around 2 hours; data were interpolated to hourly time resolution to compare with other measurements. Particle number size distributions (PNSD) were measured with a custom-made Differential Mobility Particle Sizer (DMPS) system in size range 3–1000 nm (Aalto et al., 2001). A more detailed description of
CCN measurements and station operation can be found in Sihto et al. (2011) and Paramonov et al. (2015).

The Station for Observing Regional Processes of the Earth System (SORPES) is located at a suburb of Nanjing, a megacity in the Yangtze River Delta municipal aggregation (32°07’14” N, 118°57’10” E; ~40 m a.s.l.). $\sigma_{sp}$ and $\sigma_{bsp}$ of total suspended particles (TSP) are measured with an Ecotech Aurora-3000 3-wavelength integrating nephelometer at $\lambda = 450, 525, \text{ and } 635 \text{ nm}$ as described by Shen et al. (2018). $N_{CCN}$ is measured using a CCN-200 dual column CCN counter at 5 supersaturations: 0.1%, 0.2%, 0.4%, 0.6% and 0.8%. The two columns make the same cycle simultaneously to cross-check with each other. Each cycle takes 30 minutes. PNSD in the size range of 6 - 800 nm are measured with a DMPS built by University of Helsinki. More details of the measurements at SORPES are given by, e.g., Ding et al. (2013, 2016) and Qi et al. (2015).

The US Atmospheric Radiation Measurement (ARM) Mobile Facility (AMF) measures atmospheric aerosol and radiation properties all over the world. The first AMF (AMF1) was deployed in 2005 with both a CCN counter and a nephelometer. Between 2011 and 2018, AMF1 is operated at four locations: Ganges Valley (PGH) in the Himalayas, Cape Cod, Massachusetts (PVC) in a coastal area of U.S., Manacapuru (MAO) inside the Amazonian rain forest, and Ascension Island (ASI) on the South Atlantic Ocean downwind from Africa. Three of them are accompanied by a scanning mobility particle sizer (SMPS; Kuang, 2016). The SMPS is also part of the Aerosol Observing System (AOS) running side by side with AMF1 since 2012. Both PNSD and AOPs are available simultaneously at PVC, MAO, and ASI. $\sigma_{sp}$ and $\sigma_{bsp}$ of sub-1 $\mu$m and sub-10 $\mu$m particles are measured at all AMF1 locations by integrating nephelometers (Uin, 2016a). The size range of the SMPS is around 11 – 465 nm with slightly different ranges for different periods. $N_{CCN}$ is measured at different supersaturations, details are in Table 1. The supersaturations are typically calibrated before and after each campaign at an altitude similar to measurement site by instrument mentors according to CCN handbook (Uin, 2016b). Detailed information about each dataset and measurement site can be found on AOS handbook.
(Jefferson, 2011) or ARM web site (http://www.arm.gov/) and references thereby.

Ganges Valley (PGH) is located in one of the largest and most rapidly developing sections of the Indian subcontinent. The aerosols in this region have complex sources, including coal and fuel combustion; biomass burning; automobile emissions; and dust. In monsoon seasons, dust dominates the aerosol mass due to transportation (Dumka et al., 2017; Gogoi et al., 2015).

PVC refers to the on-shore data set for the ‘first column’ of the Two-Column Aerosol Project (TCAP) on Cape Cod, Massachusetts, USA. This is a marine site but still significantly affected by anthropogenic emissions (Berg et al., 2016).

MAO refers to Manacapuru in Amazonas, Brazil. Manaus pollution plume and biomass burning impact the background conditions alternately. During the period we selected for this study, no severe pollution episodes were observed. The $\sigma_p$ for PM10 never exceeded 250Mm$^{-1}$ in this study.

Ascension Island (ASI) locates in the southeast Atlantic where westward transport of southern Africa biomass-burning aerosols emphases heavy aerosol loading. Air mass at this site usually a mixture with aged biomass-burning plume and sea-salt aerosol. The aerosol loading can be very low without plume, in this case, there is substantial uncertainty on the backscatter fraction.

The primary purpose of this study is to use as basic and readily accessible measurement data as possible to estimate $N_{CCN}$. Aerosol optical properties are measured for different cutoff diameters, usually 1 μm, 2.5 μm, 10 μm or TSP. At several stations there are two sets of AOPs using two cutoff diameters. For this study we chose to use AOP data with the 10 μm cutoff or TSP that are more universally used than smaller cutoff diameters.

2.2 Data processing

Regardless of the time resolution of raw data, all the data in this study were adjusted into hourly
averages before further analyses. Suspicious data within the whole dataset were removed according to the following criteria:

1) for the size distribution data, all the data with unexplainable spikes were removed manually;
2) for CCN measurements, insufficient water supply may cause underestimation of CCN, especially at lower supersaturation ratios (DMT, 2009). N_{CCN} reading at lower SS% has a sudden drop a few hours before the similar sudden drop for higher SS% under such conditions, so data from such periods were removed;
3) if any obvious inconsistencies between the AOPs and PNSD or between the N_{CCN} and PNSD were found on closure study, all the data in the same hour were removed.

Special treatments were carried out for ASI dataset. There will inevitably be a considerable uncertainty in the backscattering fraction if zero point of either \( \sigma_{sp} \) or \( \sigma_{bsp} \) is inaccurate in very clean conditions. The measured \( \sigma_{sp} \) was in agreement with that calculated from the PNSD with the Mie model. However, in the data \( \sigma_{bsp} \) approaches 0.3 Mm\(^{-1}\) whenever \( \sigma_{sp} \) approaches 0. Thus, we subtracted from back scattering coefficients a constant 0.3 Mm\(^{-1}\) and no longer used any data points with \( \sigma_{sp} < 2 \) Mm\(^{-1}\) for this site to assure the data quality.

2.3 Light scattering calculated from the particle number size distributions

Light scattering coefficients were calculated using the Mie code similar to Bohren and Huffman (1983) for SMEAR II. The refractive index was set to the average value of 1.517+0.019i reported for SMEAR II by Virkkula et al. (2011). The wavelength for Mie modeling was set to 550 nm, which is same as in the measurements. The whole size range of the DMPS or the SMPS, depending on the station, was used. The total scattering coefficient (\( \sigma_{sp} \)) and hemispheric backscattering coefficient (\( \sigma_{bsp} \)) represent the scattering phase function integrated over the scattering angles of 0-180° and 90-180°, respectively. The backscatter fraction (BSF) is the ratio between \( \sigma_{bsp} \) and \( \sigma_{sp} \).
2.4 CCN number concentration calculated from the particle number size distribution

Under the assumption of fully internally mixed particles, the CCN number concentration calculated from the particle number size distributions ($N_{CCN}(PNSD)$) is obtained by integrating the PNSD of particles larger than the critical dry particle diameter ($D_m$):

$$N_{CCN}(PNSD) = \int_{D_m}^{\infty} n \log D \, d \log D$$  \hspace{1cm} (1)

At a given SS, $D_m$ is a diameter above which all particles can act as CCN. For a selected dry diameter of a particle having given hygroscopicity is computed from the maximum of: The Critical Diameter $D_m$ is the minimum dry diameter ($D_d$) that ensure the $\kappa$-Köhler curve (Petters and Kreidenweis, 2007) to have one real solution:

$$S(D) = \frac{D^3 - D_d^3}{D^3 - D_d^3 (1 - \kappa)} \exp \left( \frac{4\sigma_{sa} M_w}{RT \rho_w D} \right).$$  \hspace{1cm} (2)

Here $D_d$ is the dry diameter, $\sigma_{sa}$ is the surface tension of the solution/air interface, $R$ is the universal gas constant, $T$ is temperature, and $D$ is the diameter of the droplet, $\rho_w$ is the density of water, $M_w$ is the molecular weight of water, and $\kappa$ is the hygroscopicity parameter. $S(D)$ in this particular case is set to the same supersaturation ratio as CCN being measured (e.g., 0.1%, 0.2%, 0.3%, 0.5% and 1.0% for SMEAR II).

The accuracy of $N_{CCN}(PNSD)$ is affected by the treatment of $\kappa$. In this study, we are not trying to achieve an accurate value of $\kappa$ but instead want to illustrate that even an arbitrary setting of $\kappa$ can yield reasonable CCN concentrations. This approach is named as ‘unknown chemical approach’ in (Kammermann et al., 2010) and as ‘Prediction of $N_{CCN}$ from the constant $\kappa$’ in Meng et al., (2014). Both of them give a detailed discussion of how this approach performs.

Arbitrary $\kappa$ is not performing as good as a proper $\kappa$ when calculating $N_{CCN}$, yet we believe that it is good enough to be an alternative to measuring CCN in the empirical estimation of this study. Wang et al. (2010) also claimed that $N_{CCN}(PNSD)$ may be successfully obtained by assuming an internal mixture and using bulk composition few hours after emissions. For SORPES, ASI and PVC, we simply set a global-average value of 0.27 for $\kappa$ (Pringle et al., 2010; Kerminen et al., 2012). For the forest sites, SMEAR II and MAO, we set $\kappa = 0.12$, which is
close to the value of $\kappa$ for Aitken mode particles reported previously by studies at forest sites (Sihto et al., 2011; Hong et al., 2014).

2.5 Aerosol optical properties and CCN concentrations of simulated size distributions

For studying the relationships of particle size, $N_{\text{CCN}}$ and AOPs we generated unimodal particle number size distributions $n_{\text{num}}(\text{GMD}, \text{GSD})$ with varying the geometric mean diameter (GMD) and geometric standard deviation (GSD). For them we calculated the same AOPs with the Mie model as were obtained from the real measurements from the stations $\sigma_{\text{sp}}$ and $\sigma_{\text{bsp}}$ and from these the BSF at the wavelengths $\lambda = 450, 550$ and $700$ nm. $N_{\text{CCN}}$ was calculated simply by integrating number concentrations of particles larger than a critical diameter of 80 nm, 90 nm, 100 nm, and 110 nm.

3. Overview of measured properties

At SMEAR II the average values of $\sigma_{\text{sp}}$ at $\lambda = 550$ nm in PM$_1$ and PM$_{10}$ were 11.2 and 13.7 Mm$^{-1}$ and corresponding backscattering fractions were 0.162 and 0.155, respectively, during our study period. These values are consistent with those reported by Virkkula et al. (2011). The average ± standard deviation $N_{\text{CCN}}$ was $129 \pm 99, 303 \pm 228, 391 \pm 303, 512 \pm 384$ and $736 \pm 492$ cm$^{-3}$ at SS% of 0.1, 0.2, 0.3, 0.5 and 1.0%, respectively. The CCN spectrum here did not quite follow the traditional fitting $N_{\text{CCN}}(\text{CSK}) = C \times \text{SS\%}^k$ (Jefferson 2010, Twomey 1959). One year average $N_{\text{CCN}}(\text{CSK})$ at SMEAR II for SS=0.1% is 197/cm$^3$, 53% higher than $N_{\text{CCN}}(\text{mea})$ for the same period. Also, $R^2$ of the linear regression between $N_{\text{CCN}}(\text{CSK})$ and $N_{\text{CCN}}(\text{mea})$ is 0.78 at SS=1.0%, which means that Jefferson's method performs approximately as well for SMEAR II as at the other sites presented by Jefferson (2010). However, our motivation is to develop a method that needs no absorption data.

3.1 AOPs and CCN calculated from particle size distributions

Aerosol optical properties calculated from particle number size distributions matched well with the measured scattering coefficients in PM$_1$. For $\sigma_{\text{sp}}$ larger than about 40 Mm$^{-1}$, the calculated values were slightly lower than the measured ones. The measured and calculated BSF also
matched well with $r^2=0.93$ for the data with $\sigma_{sp}>10$. Another quality check of the CCN data is that the $N_{CCN}(\text{PNSD})$ calculated from Eq. (1) was consistent with the measured CCN number concentration $N_{CCN}(\text{meas})$: for the linear regression $r^2$ was 0.80, 0.91, 0.94 and 0.92 for SS=0.1%, 0.2%, 0.5% and 1.0%, respectively, and the corresponding slopes varied between 0.85 and ~1.2 depending on the value of SS%. The correlation between $N_{CCN}(\text{PNSD})$ and $N_{CCN}(\text{meas})$ was the weakest for the lowest set of supersaturation (0.1%), most probably because the measurement uncertainty is much higher at lower values of SS% compared with higher SS% for DMT CCN counter (Rose et al., 2008).

3.2 Relationships between AOPs and CCN

The correlation between $N_{CCN}$ and $\sigma_{sp}$ was weak at SMEAR II, especially for higher supersaturations (Fig 1). In spite of this, when color-coded with respect to BSF, the relationship between $N_{CCN}$ and $\sigma_{sp}$ becomes clear: the scatter plot points of $N_{CCN}$ grows almost linearly as a function of $\sigma_{sp}$ for a narrow range of values of BSF. This indicates BSF can serve as a good proxy for describing the ratio between $N_{CCN}$ and $\sigma_{sp}$ at SMEAR II.

Hereafter, we will use the term $R_{CCN}=N_{CCN}/\sigma_{sp}$ to describe the relationship between CCN concentration light scattering and similar to Liu and Li (2014). Note that this same ratio was defined as $AR_{scat}$ in Tao et al. (2018). $R_{CCN}$ varies over a wide range of values, so a proper parameterization to describe it is of significance.

4. Development of the parameterization

4.1 Site-dependent parameterization for each measured supersaturation, $N_{CCN}(\text{AOP}_{i})$

The first step in the development of the parameterization was to calculate linear regressions of $R_{CCN}$ vs BSF. $R_{CCN}$ depends clearly on BSF (Fig. 2) as

$$R_{CCN}= a \ BSF + b$$ (3)

At SMEAR II the correlation between BSF and $R_{CCN}$ is strong when $\sigma_{sp} > 10 \ \text{Mm}^{-1}$. At $\sigma_{sp} < 10 \ \text{Mm}^{-1}$ the uncertainty of the nephelometer is higher which may at least partly explain the lower correlation. For each dataset and individual supersaturation, $a$ and $b$ the slope and offset
of the linear regressions has a different value as presented in Table 2. The parameterization gives the formula for calculating $N_{CCN}(AOP)$, i.e., $N_{CCN}$ calculated from measurements of AOPs:

$$N_{CCN}(AOP) = (a_{SS\%} \cdot BSF + b_{SS\%}) \cdot \sigma_{sp}$$  \hspace{1cm} (4)

The subscript 1 for $AOP_1$ indicates the first set of parameterization.

Scatter plots of $N_{CCN}(AOP_1)$ vs $N_{CCN}(meas)$ are presented for the supersaturations used at the SMEAR II CCN counter in Fig 3 and for the highest and the lowest SS% used at the other stations in Fig 4. At SMEAR II this approach yields $R^2$ of 0.70, 0.86, 0.75 and 0.55 for SS=0.1%, 0.2%, 0.5% and 1.0%, respectively, and the slopes (and intercepts) are 0.95(13), 0.92(28), 0.86(52) and 0.76(87), respectively. All slopes are slightly less than 1 and the intercept are slightly over 0. One explanation is that when both $x$ and $y$ have uncertainties, the least-squares method in the linear regression trend to underestimate the slope (Cantrell, 2008). $N_{CCN}(AOP_1)$ overestimates (or underestimates) $N_{CCN}(meas)$ by 4.8%, 1.2%, -4.2% and -12.5% at the above specified supersaturations. For the overall dataset regardless of supersaturations, $R^2$, slope(intercept) and difference between $N_{CCN}(AOP_1)$ and $N_{CCN}(meas)$ are 0.73, 0.81(56) and $-5.1\%$ respectively.

$R^2$ between $N_{CCN}(AOP_1)$ and $N_{CCN}(meas)$ is higher at lower supersaturations than at higher supersaturations in most of the scatter plots shown in Figures 3 and 4. The reasonable explanation is that the higher the supersaturation is the smaller are the particles that can act as CCN. The smaller are the particles the less do they contribute to both total scattering and backscattering and the higher is the relative uncertainty of both of them and thus also the uncertainty of $N_{CCN}(AOP_1)$.

### 4.2 General combined parameterization $N_{CCN}(AOP_2)$

In the next step, the slopes and offsets obtained from the linear regression (Table 2) were plotted as a function of SS% (Fig 5). The data obviously depend logarithmically on SS% so that (4) becomes

$$N_{CCN}(AOP) = \left( a_{SS\%} \cdot BSF + b_{SS\%} \right) \sigma_{sp} = \left( (a_1 \ln(SS\%) + a_0) \cdot BSF + b_1 \ln(SS\%) + b_0 \right) \sigma_{sp}$$  \hspace{1cm} (5)

The coefficients $a_0$, $a_1$, $b_0$ and $b_1$ obtained from the regression of $a_{SS\%} = a_1 \ln(SS\%) + a_0$ and $b_{SS\%} =$
b_1 ln(SS%) + b_0 for each station are presented in Table 3. The relationships of the coefficients can be used for a combined, more general parameterization. Obviously the a_0 vs. a_1, b_0 vs. b_1, a_1 vs. b_1 and b_0 vs. b_1 pairs from all stations follow very accurately same lines (Fig 6). This suggests that there is some underlying reason for it. Linear regressions yielding a_0 = (2.38 ± 0.06)a_1, b_0 = (2.33 ± 0.03)b_1, and b_1 = (-0.097 ± 0.013)a_1 + (6.4 ± 5.9) were used, after the simple algebra in the supplement, to get

\[ N_{CCN}^{(AOP)} = \ln \left( \frac{SS\%}{0.093 \pm 0.006} \right) \left( BSF - (0.097 \pm 0.013) \right) + (6.4 \pm 5.9) \sigma_sp \]

(6)

where both the coefficient a_1 and the constant 6.4 ± 5.9 have units of [N_{CCN}]/[\sigma_sp] = cm^{-3}/Mm^{-1}. This is the general formula for the parameterization. In both (5) and (6) the only unquantified coefficient is now a_1. However, we can find some ways to quantify also it.

For a given station, if there are simultaneous data of N_{CCN}(meas) and \sigma_sp for some reasonably long period, (6) can be adjusted. Instead of subtracting (0.097 ± 0.013) from BSF the minimum BSF = BSF_{min} in the data set will be used. Further, when BSF = BSF_{min} the factor a_1(BSF - BSF_{min}) = 0 and N_{CCN}^{(AOP)} ≈ R_{min}\sigma_sp where R_{min} is the minimum R_{CCN} in the data set. It follows that

\[ N_{CCN}^{(AOP)} = a_1 \ln \left( \frac{SS\%}{0.093 \pm 0.006} \right) \left( BSF - BSF_{min} \right) + R_{min} \sigma_sp \]

(7)

The derivation of (7) is shown in the supplement. In the data processing the 1st percentiles of both BSF and R_{CCN} are used as BSF_{min} and R_{min}, respectively. Here the free parameters are a_1, BSF_{min} and R_{min}. The coefficient a_1 is positively correlated with SAE. The linear regressions of a_1 and the average and median scattering Ångström exponent of PM_{10} particles (SAE_{10}) (Table 3) at the 6 sites in the analyzed periods yield a_1 = (298 ± 51)SAE_{10} cm^{-3}/Mm^{-1} and a_1 = (287 ± 45)SAE_{10} cm^{-3}/Mm^{-1}, respectively (Fig. 7). The uncertainties are large but, the main point is that the correlations show that a_1 and thus N_{CCN}^{(AOP)} is the higher the higher SAE_{10} is. R_{min} was estimated by calculating the 1st percentile of R_{CCN} at each site at each SS%. The average and standard deviation of R_{min} was 5.2 ± 3.3 cm^{-3}/Mm^{-1}. Consequently the parameterization becomes

\[ N_{CCN}^{(AOP)} = \left( 287 ± 45 \right) SAE_{10} \ln \left( \frac{SS\%}{0.093 \pm 0.006} \right) \left( BSF - BSF_{min} \right) + (5.2 ± 3.3) \sigma_sp \]

(8)
5. Results and discussion

5.1 Comparison of $N_{CCN}$ from the AOP parameterization and measurements

The parameterization in Eq (8) was applied to the data of the 6 stations and $N_{CCN}(AOP_2)$ was compared with the $N_{CCN}(meas)$ at the supersaturations used in the respective CCN counters. The results are presented as scatter plots of $N_{CCN}(AOP_2)$ vs. $N_{CCN}(meas)$ (Fig 8a and 8b), the bias of the parameterization calculated as $N_{CCN}(AOP_2)/N_{CCN}(meas)$ (Fig 8c) and the squared correlation coefficient $R^2$ of the linear regression of $N_{CCN}(AOP_2)$ vs. $N_{CCN}(meas)$ (Fig 8d).

At the site-specific lowest SS% the scatter plots of $N_{CCN}(AOP_2)$ vs. $N_{CCN}(meas)$ of data from most stations get clustered along the 1:1 line, but for the Himalayan site PGH the parameterization yields significantly higher concentrations (Fig 8a). It was mentioned above that we applied also the Jefferson (2010) parameterization to SMEAR II data. At SS=0.1% the average $N_{CCN}(CSK)$ was 53% higher than $N_{CCN}(meas)$ and $R^2$ of the linear regression was 0.78 at SS=1.0%. The bias of our method at SS% = 0.1 was ~0.66 so it underestimated measurements by 34% and $R^2 \approx 0.65$ (Fig 8), lower than that of Jefferson (2010). At SS% > 0.3 the bias varied from 1.1 to 1.3. At the highest SS% the deviations from the 1:1 line are smaller also for PGH (Fig 8b). At PGH at the lowest SS% the bias is > 4 but decreases to ~1.1-1.2 at SS% = 0.4% and even closer to 1 at higher SS%. At SS% > 0.4% the AOP-derived $N_{CCN}$ is higher than the measured concentration at four sites with their bias varying between ~1.1 and ~1.3. For the US coastal site PVC the parameterization constantly underestimates the CCN concentrations by about 30%. For the Amazonian site MAO the bias is close to 1 at the lowest SS% but for the higher SS% it varies from 0.68 to 0.79.

5.2 Evaluation of the effect of particle size distribution to the parameterization

The linear relationships of the coefficients of Eq. (5) are so clear (Fig. 6) that there should be some common underlying reason. To study this we generated lognormal unimodal size distributions as explained in section 2.5. GMD was given logarithmically evenly-spaced values from 50 nm to 1600 nm and GSD was given two values: 1.5 representing a relatively narrow size distribution and 2.0 a wide size distribution. We then calculated AOPs, $N_{CCN}$ and $R_{CCN/\sigma}$.
for these size distributions.

The reasoning for the approach of estimating $N_{CCN}$ from $\sigma_{sp}$ and BSF can easily be explained by the similar variations of $R_{CCN0}$ and BSF (Fig. 9). $R_{CCN0}$ is the highest for the smallest particles, i.e. for GMD = 50 nm and it decreases with the growing GMD as also BSF. Note that the width of the size distribution has very strong effects on $R_{CCN0}$: for the wide size distribution it is approximately an order of magnitude lower than for the narrow size distribution.

Note also that the rates of decrease of $R_{CCN0}$ and BSF. We used this information for estimating particle sizes with a stepwise linear regression. An example is given by the linear regressions of $R_{CCN0}$ vs. BSF calculated for 5 consecutive size distributions, first for those that have their GMDs from 50 nm to 100 nm and the second for those that have their GMDs from 100 nm to 200 nm (Fig. 10). Note that it is obvious that linear regressions are applicable for short intervals but do not well for the whole size range. The absolute values of the slopes and offsets are clearly lower for the larger particle size range. The particle size that is used for describing the size range of each regression we define here as the equivalent geometric mean diameter GMD$_e$, the geometric mean of the range of the GMDs of the unimodal size distributions used for each regression. It will be shown below that GMD$_e$ is a mathematical concept helping in explaining the observed relationships, not an actual GMD of the particle size distribution at the sites.

For the wide size distributions the slopes and offsets of the regressions of $R_{CCN0}$ vs. BSF decrease and increase, respectively, monotonically with an increasing GMD$_e$ in the whole size range studied here (Fig. 11). For the narrow size distribution the slope decreases to GMD$_e$ = 300 nm and then increases which means there is no unambiguous relationship between them.

Note also that the ranges of the absolute values of the slopes and offsets of the wide and narrow size distributions are very different. However, they decrease and increase simultaneously. This is the link to the observations from the field stations. We plotted the offset vs slope of the unimodal size distributions and those obtained from the linear regressions of the field data at...
the supersaturations presented in Table 2 and below it the GMD$_e$ vs. the slopes of the regressions of the unimodal size distributions (Fig 12). In Fig. 12 also the effect of the choice of the activation diameters of 80 nm, 90 nm, and 110 nm is shown.

Several observations can be made of Fig. 12. First, for the simulated wide size distributions the relationship of the offset and slope is unambiguous but not for the narrow size distributions at sizes GMD$_e$ > ~200 nm (Fig 12b). Secondly, the field data points obviously follow the lines of the simulations. This supports the approach for the interpretation of the relationships presented above (Fig. 6) for the coefficients in Eq. (5). Especially, note the similar ranges of $b_0$ vs $a_0$ in Fig 6d and the ranges of $b$ vs $a$ in Fig 12a. This is the link to Eq. (5): if we set SS% = 1, the equation reduces to $N_{CCN} = (a_0 BSF + b_0)\sigma_{sp}$. Together with the clear linear relationships between the coefficients this suggests that the coefficients of Eq. (5) depend on the GMD and GSD of the particle size distributions.

Most field data agree well with the $b$ vs $a$ line of the unimodal wide size distribution with the lowest activation diameter of 80 nm. For instance, the PVC data point corresponding to the highest supersaturation has the highest slope (1970 cm$^{-3}$Mm$^{-3}$, Table 2) and it is close to the above-mentioned line (Fig. 12a). The corresponding GMD$_e$ of the unimodal size distribution is also ~80 nm (Fig 12b). The SMEAR II high SS% offset vs. slope fits best with the corresponding lines of the narrow unimodal size distributions with all activation diameters and the corresponding GMD$_e \approx 150 – 180$ nm.

At the lowest SS% the offset vs. slope points of all stations agree with the lines derived from the lines derived from the unimodal modes. This is interesting considering the high uncertainties involved in the regressions at the lowest SS% (Fig. 2). For ASI the slopes and offsets of the lowest and highest SS% are especially close to each other, closer than at any other station (Fig. 12a), and the corresponding GMD$_e \approx 750$ nm and 400 nm, respectively, when the GMD$_e$ vs. a relationship of any of the wide distributions is used (Fig. 12b). For PGH at the lowest SS% the slope is actually negative which is not obtained from the simulations at all so
no GMD, cannot be given for it.

5.3 Aerosol size characteristics for all site

As it was shown above, particle size distributions affect the coefficients of the parameterization. It is therefore discussed here how the size distributions vary at the six sites of the study and whether they support the interpretations presented above. The size distributions are discussed using the particle number size distribution and the ratios of $\sigma_{10}$ of PM$_1$ and PM$_{10}$ size ranges data from those stations where they are available.

5.3.1 Diurnal variation of particle number size distribution

Fig. 13a shows the averaged diurnal cycle of PNSD at the sites where either a DMPS or SMPS is available. New particle formation (NPF) events is a significant source of uncertainty in the prediction of $N_{\text{CCN}}$ (Kerminen et al., 2012; Ma et al., 2016). Complete NPF events start from a burst of sub 10 nm particles and continuous growing up to a few hundred nanometers. As a result, the size distribution varies significantly. NPF is one possible explanation of the poor $N_{\text{CCN}}-\sigma_{10}$ correlation.

SMEAR II and SORPES are reported to have an appreciable frequency of NPF (Kulmala et al., 2004; Dal Maso et al., 2005; Sihto et al., 2006; Qi et al., 2015). Continuous growth in particle size at SORPES can usually last for several days after NPF (Shen et al. 2018). Similar growth patterns have also been observed in the Two-Column Aerosol Project (TCAP; http://campaign.arm.gov/tcap; refers as PVC in this study) according to Kassianov et al. (2014). NPF is rarely observed in the Amazon forest where MAO is located. However, it does take place also at MAO as is shown in the diurnal cycle of PNSD. At ASI, there no evidence of NPF according to the PNSD diurnal cycle.

These observations of the NPF are compared with the bias and correlation coefficients of the parameterization discussed in section 5.1 (Fig. 8). The correlation coefficient of $N_{\text{CCN}}(\text{AOP}_2)$ vs. $N_{\text{CCN}}(\text{meas})$ is the highest, $R^2 \approx 0.8$ at all SS% at ASI where no NPF takes place and clearly lower
at the other sites (Fig 8d). For the bias NPF appears not to have a clear influence: for both SMEAR II and SORPES bias varies from ~1.1 to ~1.4 at SS% > 0.1%.

5.3.2 Distribution of geometric mean diameter

Figure 13b presents the normalized distribution of the geometric mean diameter at SMEAR II, SORPES, PVC, MAO and ASI. It varies from 20 nm to 200 nm at all sites, with the most frequent GMD between ~70 nm and ~120 nm depending on the site. This shows clearly that the above-presented equivalent geometric mean diameter GMD, calculated assuming a unimodal size distribution is not a quantitative GMD of the size distribution, it is a mathematical concept that explains partially the relationships of R_{CCN_{opt}} and BSF.

The frequency distribution of GMD at SMEAR II is the widest among five sites with PNSD data available, followed by SORPES and PVC. At MAO the frequency distribution of GMD has two peaks in this study, different from that at ATTO in Amazonas (Schmale et al., 2018). The lower peak is possibly due to the burst of sub-20 nm particles since they have little chance to grow to sizes that can serve as CCN. The second peak around 100 nm possibly represents the GMD without the burst of sub-20 nm particles and it is distinctly narrower than at SMEAR II, SORPES and PVC.

A comparison of the correlation coefficients of N_{CCN}(AOP_{2}) vs. N_{CCN}(meas) (Fig. 8) and the widths of the GMD frequency distributions (Fig. 13b) does not show any clear relationships between them, other than that of ASI. The frequency distribution of GMD is the narrowest at ASI indicating that the average particle size does not change much throughout the whole period. This is in line with the low variation of the slope and offset of the R_{CCN} vs BSF of ASI (Fig 12a). At ASI also the correlation coefficient of N_{CCN}(AOP_{2}) vs. N_{CCN}(meas) is the highest, R^2 ≈ 0.8 at all SS%.

5.3.3 Contribution of light scattering by sub-μm particles

There is one more measure related to particle size distribution, the ratio between \( \sigma_p \) of sub-1 μm and sub-10 μm aerosol (\( \sigma_p(\text{PM}_1)/\sigma_p(\text{PM}_{10}) \)). At SMEAR II, the contribution of submicron particles usually varies within range 0.8-0.9 and it is the highest among all sites in this study. PVC has two
peaks in the $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ distribution, the peak around 0.2 corresponds to air masses from the sea, with a very low scattering coefficient and $N_{CCN}$. By ignoring the cleanest air masses ($\sigma_{sp}<5$ Mm$^{-1}$), the fraction of $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ is usually around 0.8, which is just slightly lower than at SMEAR II. At PGH and MAO, the distribution of the ratio is wider, and the peak position is around 0.65. The overall contribution of sub-μm particle light scattering at PGH is moderate among the sites in this study. At ASI $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ is the lowest among all sites in this study, indicating that particles larger than 1 μm contribute a considerable fraction of light scattering. For SORPES $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ is not available.

Among those five sites, when $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ decreases, the correlation between BSF and $R_{CCN0}$ decreases. At some sites (e.g., ASI) the BSF of PM$_{10}$ is often be even larger than that of PM$_1$ which is most probably an error in the measurements but it may also be due to non-spherical particles like sea salt and dust, which will blur the correlation between BSF and $R_{CCN0}$. In such a case the increase of the amount of large particles leads to an increase of BSF and a decrease of $R_{CCN0}$ which is opposite to the usual positive correlation between BSF and $R_{CCN0}$ in this study. Thus, the lower $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ may in principle result in a poor performance of our method. However, a comparison of the correlation coefficients and the $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ frequency distributions of each site shows the opposite. At the highest SS% of each site the $R^2$ in a decreasing order is ASI, PGH, MAO, SORPES, SMEAR II, and PVC (Fig. 8d). The peaks of the frequency distribution of $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$ are, in a growing order, ASI: 0.375, PGH: 0.625, MAO: 0.65, PVC: 0.825, SMEAR II: 0.875. Note that at SORPES there is only one size range measured. Of these only PVC and SMEAR II are not in the same order. On the other hand, the bias at the highest SS% has no clear relationship with $\sigma_{sp}(\text{PM}_1)/\sigma_{sp}(\text{PM}_{10})$: for MAO our parameterization underestimates the $N_{CCN}$ the most (bias ≈ 0.68) and for ASI it overestimates the most (bias ≈ 1.28).

6. Conclusions

The relationships between aerosol optical properties, CCN number concentrations ($N_{CCN}$) and particle number size distributions were investigated based on in-situ measurement data from six stations in very different environments around the world. The goal of the work was to find
a parametrization to obtain $N_{\text{CCN}}$ from sites where AOPs are measured but no CCN counter is available.

There are many previous parameterizations for doing just the same. As a starting point we used the parameterization presented by Jefferson (2010). That one needs also absorption measurements since it includes single-scattering albedo. We instead studied how the parameterization would look like if only total scattering and backscattering data were available.

The basic idea for the parameterization is that $N_{\text{CCN}}$ is proportional to $\sigma_p$ and a function of the backscatter fraction (BSF), as is also in the parameterization of Jefferson (2010). One clear difference is that our data analysis showed that the dependence on supersaturation is logarithmic, different from that of Jefferson (2010). Actually this result is qualitatively in line with the relationship between AOD and CCN reported by Andreae (2010).

The coefficients of the parameterization derived for the different sites showed that they appear to be linearly related to each other. A simulation with unimodal size distributions showed that the relationships are affected by the size and width of the size distribution and the activation diameter.

We were able to derive a parameterization that describes all sites. The parameterization not only depends on BSF but also on the wavelength dependency of scattering, i.e. the scattering Ångström exponent SAE. Further studies need to be done to compare different parameterizations for data from various different sites.

Author contributions
YS carried out measurements at SORPES in China, analyzed and visualized data of all sites, and wrote the original draft. AV contributed to data analysis and visualization, writing and editing the original draft, and supervised the work of YS in Finland. AD provided funding for the measurements and research at SORPES in China, acquired funding for YS in China, and supervised the work of
YS. KL, HK and PA carried out measurements, data collection and maintenance of measurement data of SMEAR II in Finland. YS, XC, XQ, WN and XH carried out measurements, data collection and maintenance of measurement data of SORPES in China. MK and TP provided the funding for YS in Finland. MK provided funding for the measurements and research at SMEAR II in Finland. TP and VMK formulated the goals of the research and supervised it.

Acknowledgments

This work was supported by National Natural Science Foundation of China (41725020), Academy of Finland via Center of Excellence in Atmospheric Sciences (project no. 272041), the Centre for International Mobility CIMO of the Finnish Government Scholarship Pool programme, and the Collaborative Innovation Center of Climate Change supported by the Jiangsu 2011 Program. Data were also obtained from the Atmospheric Radiation Measurement (ARM) User Facility, a U.S. Department of Energy (DOE) Office of Science user facility managed by the Office of Biological and Environmental Research. For the SMEAR II data we thank the SMEAR II technical team. Y. Shen was supported by the CSC-Finnish joint PhD studentship.

References


Ding, A. J., Nie, W., Huang, X., Chi, X., Sun, J., Kerminen, V. M., Xu, Z., Guo, W., Petaja, T.,


Schmale, J., Henning, S., Henzing, B., Keskinen, H., Ovadnevaite, K., Bougiatioti,


Tao, J., Zhao, C., Kuang, Y., Zhao, G., Shen, C., Yu, Y., Bian, Y., and Xu, W.: A new method for calculating number concentrations of cloud condensation nuclei based on measurements of


**Table 1. Site and data description**

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Period</th>
<th>Location</th>
<th>CCN</th>
<th>Size distribution</th>
<th>AOPs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Instrument</td>
<td>SS%</td>
<td>Instrument size range</td>
</tr>
<tr>
<td>SMEAR II</td>
<td>2016.1.1-2016.12.31</td>
<td>43° 51’ N, 24° 17’ E</td>
<td>CCN-100</td>
<td>0.1%, 0.2%, 0.5% and 1.0%</td>
<td>custom-made DMPS</td>
</tr>
<tr>
<td>SORPES</td>
<td>2016.06.01-2017.05.31</td>
<td>32° 07’ N, 118° 56’ E</td>
<td>CCN-200</td>
<td>0.1%, 0.2%, 0.4%, and 0.6%</td>
<td>custom-made DMPS</td>
</tr>
<tr>
<td>PGH</td>
<td>2011.11.01-2013.03.25</td>
<td>29° 22’ N, 79° 27’ E</td>
<td>CCN-100</td>
<td>0.12%, 0.22%, 0.48% and 0.78%</td>
<td>NA</td>
</tr>
<tr>
<td>PGX&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2012.07.16-2012.09.30</td>
<td>42° 2’ N, 30° 3’ W</td>
<td>CCN-100</td>
<td>0.15%, 0.25%, 0.4% and 1.0%</td>
<td>SMPS TSI-3936</td>
</tr>
<tr>
<td>MAO&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2014.01.29-2014.12.31</td>
<td>3° 13’ S, 60° 36’ W</td>
<td>CCN-100</td>
<td>0.25%, 0.4%, 0.6% and 1.1%</td>
<td>SMPS TSI-3936</td>
</tr>
<tr>
<td>ASI&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2016.06.01-2017.10.19</td>
<td>7° 58’ S, 14° 21’ W</td>
<td>CCN-100</td>
<td>0.15%, 0.25%, 0.4%, and 0.8%</td>
<td>SMPS TSI-3936</td>
</tr>
</tbody>
</table>

<sup>a</sup> products used: aiavglogrenM1.c1., and aosccnavgM1.c2.

<sup>b</sup> products used: aiavglogrenM1.s1., noaaaosccn100M1.b1., and aossmpsS1.a1.

<sup>c</sup> products used: aip1ogrenM1.c1., aosccn2colaavgM1.b1., and aossmpsM1.a1.

<sup>d</sup> products used: asinephdryM1.b1., aosccnaosccnM1.b1., and aossmpsM1.a1.

<sup>*</sup> may vary slightly
Table 2. The slopes and offsets of the linear regressions of \( R_{CCN/\sigma} \) vs. BSF at the different supersaturation SS\% at the studied sites. s.e.: standard error of the respective coefficient obtained from the linear regressions.

<table>
<thead>
<tr>
<th>SITE</th>
<th>SS%</th>
<th>( a \pm \text{s.e.} )</th>
<th>( b \pm \text{s.e.} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMEAR II</td>
<td>0.10%</td>
<td>91 ± 3</td>
<td>-2.9 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>0.20%</td>
<td>433 ± 5</td>
<td>-39 ± 0.7</td>
</tr>
<tr>
<td></td>
<td>0.50%</td>
<td>867 ± 10</td>
<td>-86 ± 1.5</td>
</tr>
<tr>
<td></td>
<td>1.00%</td>
<td>1155 ± 17</td>
<td>-116 ± 2.5</td>
</tr>
<tr>
<td>SORPES</td>
<td>0.10%</td>
<td>62 ± 2</td>
<td>-2.6 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>0.20%</td>
<td>266 ± 4</td>
<td>-18 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>0.40%</td>
<td>531 ± 7</td>
<td>-39 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>0.80%</td>
<td>738 ± 11</td>
<td>-56 ± 1.2</td>
</tr>
<tr>
<td>PGH</td>
<td>0.12%</td>
<td>-18 ± 1</td>
<td>2.6 ± 0.1</td>
</tr>
<tr>
<td></td>
<td>0.22%</td>
<td>24 ± 3</td>
<td>2.8 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>0.48%</td>
<td>244 ± 12</td>
<td>-4.4 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>0.78%</td>
<td>344 ± 14</td>
<td>-8.3 ± 1.0</td>
</tr>
<tr>
<td>PVC</td>
<td>0.15%</td>
<td>417 ± 9</td>
<td>-30 ± 1.1</td>
</tr>
<tr>
<td></td>
<td>0.25%</td>
<td>793 ± 17</td>
<td>-62 ± 2.1</td>
</tr>
<tr>
<td></td>
<td>0.40%</td>
<td>1176 ± 25</td>
<td>-95 ± 3.1</td>
</tr>
<tr>
<td></td>
<td>1.00%</td>
<td>1945 ± 43</td>
<td>-161 ± 5.3</td>
</tr>
<tr>
<td>MAO</td>
<td>0.25%</td>
<td>273 ± 5</td>
<td>-19 ± 0.7</td>
</tr>
<tr>
<td></td>
<td>0.40%</td>
<td>544 ± 8</td>
<td>-43 ± 1.2</td>
</tr>
<tr>
<td></td>
<td>0.60%</td>
<td>678 ± 13</td>
<td>-51 ± 1.8</td>
</tr>
<tr>
<td></td>
<td>1.10%</td>
<td>868 ± 32</td>
<td>-58 ± 4.3</td>
</tr>
<tr>
<td>ASI</td>
<td>0.15%</td>
<td>22 ± 2</td>
<td>2.2 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>0.25%</td>
<td>105 ± 3</td>
<td>-3.6 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>0.40%</td>
<td>127 ± 4</td>
<td>-5.0 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>0.80%</td>
<td>136 ± 4</td>
<td>-4.0 ± 0.6</td>
</tr>
</tbody>
</table>
Table 3. The coefficients $a_0$, $a_1$, $b_0$ and $b_1$ obtained from the fitting of $a = a_1 \ln(\text{SS}%) + a_0$ and $b = b_1 \ln(\text{SS}%) + b_0$ with the data in Table 2. The unit of the coefficients is $[N_{\text{CCN}}] / [\sigma_{sp}] = \text{cm}^{-3}/\text{Mm}^{-1}$.

s.e.: standard error of the respective coefficient obtained from the regressions.

<table>
<thead>
<tr>
<th>Location</th>
<th>$a_1 \pm \text{s.e.}$</th>
<th>$a_0 \pm \text{s.e.}$</th>
<th>$b_1 \pm \text{s.e.}$</th>
<th>$b_0 \pm \text{s.e.}$</th>
<th>average $\pm \text{std}$</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMEAR II</td>
<td>464 $\pm$ 11</td>
<td>1170 $\pm$ 16</td>
<td>-49 $\pm$ 1.5</td>
<td>-118 $\pm$ 2.1</td>
<td>2.11 $\pm$ 0.67</td>
<td>2.22</td>
</tr>
<tr>
<td>SORPES</td>
<td>331 $\pm$ 12</td>
<td>817 $\pm$ 18</td>
<td>-26 $\pm$ 0.9</td>
<td>-62 $\pm$ 1.4</td>
<td>1.45 $\pm$ 0.33</td>
<td>1.50</td>
</tr>
<tr>
<td>PGH</td>
<td>205 $\pm$ 30</td>
<td>385 $\pm$ 41</td>
<td>-6.3 $\pm$ 1.5</td>
<td>-9.1 $\pm$ 2.0</td>
<td>0.53 $\pm$ 0.30</td>
<td>0.57</td>
</tr>
<tr>
<td>PVC</td>
<td>810 $\pm$ 31</td>
<td>1933 $\pm$ 21</td>
<td>-70 $\pm$ 1.7</td>
<td>-160 $\pm$ 2.1</td>
<td>1.79 $\pm$ 0.52</td>
<td>1.91</td>
</tr>
<tr>
<td>MAO</td>
<td>393 $\pm$ 45</td>
<td>858 $\pm$ 40</td>
<td>-25 $\pm$ 6.6</td>
<td>-60 $\pm$ 5.8</td>
<td>1.00 $\pm$ 0.55</td>
<td>1.09</td>
</tr>
<tr>
<td>ASI</td>
<td>64 $\pm$ 25</td>
<td>168 $\pm$ 31</td>
<td>-3 $\pm$ 2.2</td>
<td>-6 $\pm$ 2.7</td>
<td>0.73 $\pm$ 0.41</td>
<td>0.64</td>
</tr>
</tbody>
</table>

$R_{CCN/\sigma} = (a_1 \ln(\text{SS}%) + a_0)BSF + b_1 \ln(\text{SS}%) + b_0$ SAE$_{10}$
FIGURES

Figure 1. Measured CCN number concentration $N_{CCN}^{(meas)}$ vs. PM$_{10}$ scattering coefficient $\sigma_{sp}$ at $\lambda = 550$ nm at SMEAR II at four supersaturations (SS%): a) 0.1 %, b) 0.2 %, c) 0.5 % and d) 1.0 %. Colorcoding: backscattering fraction (BSF) at $\lambda = 550$ nm.

Figure 2. Relationship between $R_{CCN/\sigma}$ ($= N_{CCN}^{(meas)}/\sigma_{sp}$) and BSF at SMEAR II at four supersaturations (SS%): a) 0.1 %, b) 0.2 %, c) 0.5 % and d) 1.0 %. Grey symbols: all data, red symbols: data at $\sigma_{sp} > 10$ M$m^{-1}$. Both $\sigma_{sp}$ and BSF were measured at $\lambda = 550$ nm.
Figure 3. Comparison between $N_{CCN}(\text{AOP})$ and $N_{CCN}(\text{meas})$ at SMEAR II. $N_{CCN}(\text{AOP})$ was calculated by using the constants $a$ and $b$ in Table 2 for each supersaturation.

Figure 4. $N_{CCN}(\text{AOP})$ vs. $N_{CCN}(\text{meas})$ at a) SORPES, b) MAO, c) PVC, d) ASI and e) PGH. $N_{CCN}(\text{AOP})$ was calculated by using the constants $a$ and $b$ in Table 2 for each supersaturation.
Figure 5. The coefficients $a$ and $b$ of each station (Table 2) as a function of supersaturation.

Figure 6. Relationship between the coefficients of Equation (5) presented in Table 2 for the 6 stations. a) $a_0$ vs. $a_1$, b) $b_0$ vs. $b_1$, c) $b_1$ vs. $a_1$, d) $b_0$ vs. $a_0$. 
Figure 7. Relationship of the $a_1$ coefficient in Equation (6) with the average a) geometric mean diameter of the PNSD data size ranges of the sites, b) volume mean diameter of the same size range, and c) PM$_{10}$ scattering Ångström exponent (SAE10).

Figure 8. Statistics of $N_{CCN}(\text{AOP}_2)$ from parameterization 2. $N_{CCN}(\text{AOP}_2)$ vs. $N_{CCN}(\text{meas})$ at different sites at relatively (a) low and (b) high supersaturations, (c) bias = $N_{CCN}(\text{AOP}_2)/N_{CCN}$ (meas) at different sites and supersaturations, and (d) $R^2$ of the linear regression of $N_{CCN}(\text{AOP}_2)$ vs. $N_{CCN}$ (meas) at different sites and supersaturations.
Figure 9. Size distribution of a) $R_{CCN/\sigma}$ and b) backscatter fraction BSF ($\lambda = 550$ nm) of simulated narrow (GSD = 1.5) and wide (GSD = 2.0) unimodal size distributions. GMD: geometric mean diameter, GSD: geometric standard deviation.

Figure 10. Linear regressions of $R_{CCN/\sigma}$ vs backscatter fraction BSF ($\lambda = 550$ nm) of simulated unimodal a) narrow (GSD = 1.5) and b) wide (GSD = 2.0) size distributions. The regressions were calculated assuming that the data consist of size distributions with GMD ranging from 50 nm to 100 nm and 100 to 200 nm.
Figure 11. Size distributions of the coefficients of the linear regressions of $R_{CCN}/\sigma$ ($\lambda = 550$ nm) vs backscatter fraction BSF ($\lambda = 550$ nm) of narrow and wide size distributions. a) slopes of $R_{CCN}/\sigma$ vs. BSF, b) offsets of $R_{CCN}/\sigma$. $R_{CCN}/\sigma$ was calculated assuming particles larger than 90 nm get activated. The regressions were calculated for 5 consecutive size distributions. GMDe is the geometric mean of the range of the unimodal size distributions used for the regressions.
Figure 12. a) Relationships of the slopes and offsets of the linear regressions $R_{CCN}/\sigma = aBSF + b$ of the simulated unimodal narrow (GSD = 1.5) and wide (GSD = 2.0) size distributions and those obtained from the similar regressions of the station data at low and high supersaturations (Table 2). b) Equivalent geometric mean diameter ($GMD_e$) of the unimodal modes used for the linear regression vs. the slope of the linear regression of $R_{CCN}/\sigma$ vs. BSF. The vertical error bars show the ranges of the GMDs of the unimodal size distributions used in the respective linear regressions.
Figure 13. Analyses of particle size distributions at the six sites. a) Average diurnal cycle of PNSD and b) normalized size distribution of GMD at SMEAR II, SORPES, PVC, and ASI, c) normalized frequency distribution of $\sigma_{pm}(PM_{1})/\sigma_{pm}(PM_{10})$ at SMEAR II, PVC, MAO, PGH and ASI.