Supplement of Atmos. Chem. Phys., 19, 15483–15502, 2019
https://doi.org/10.5194/acp-19-15483-2019-supplement

Atmospheric Chemistry and Physics
Open Access

Supplement of

Estimating cloud condensation nuclei number concentrations using aerosol optical properties: role of particle number size distribution and parameterization

Yicheng Shen et al.

Correspondence to: Aki Virkkula (aki.virkkula@fmi.fi) and Aijun Ding (dingaj@nju.edu.cn)

Supplement

The Supplement contains the following sections:

S1. Amount of data used for the analyses, fractions of accepted data and criteria for data removal

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;

There were 7587 available hourly-averaged-PNSD in MAO with 104 bins of each. A total of 5234 spikes were removed. This accounts for ~0.7% of the total number of bins. 423 out of 7587 (~5.6%) distributions had at least 1 bin(s) removed. A distribution with few missing bins are still usable if treated properly. Only 55 (~0.7%) distributions had more than 10 spikes removed.

Besides for MAO, other data sets rarely suffered from such spikes. 32 out of 11502 (~0.3%) distributions were removed for ASI. For SORPES and SMEAR2, less than 1% of distributions were removed. We didn't remove anything from PNSD of PVC and PNSD is not available in PGH.

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;

Besides from the QC flag within MAO dataset, additional 55, 112,120 and 123 data points were removed at SS=0.25%,0.4%, 0.6% and 0.8% respectively, which accounts for ~0.7%-1.6% of total available data. For SORPES and SMEAR2 ~1% of total available data were removed. For ASI, PVC and PGH, no further treatment was applied besides the original QC flag.

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.

51 successive hours of data from PVC were removed before analysis, which account for ~3% of the data we used in this study. 84 sparse data points were removed from the ASI data set, which account for ~0.7% of total available data. For SORPES and SMEAR2 less than 1% of data were removed.

In total, additional quality control removes ~2%, ~3%, ~1% and 0% of the total available data in MAO, PVC, ASI and PGH respectively. The exact number for SORPES and SMEAR2 is not applicable since those 3 criteria are within the original data process procedures. However, a rough estimation of fractional data removed by such criteria are 0.5%~2%.

The total number of available hourly-averaged data, accepted data and removed data and the fractions of these are presented in Table S1.

Table S1. Number of data and fractions of removed data from all stations

	Period (hours)		AOPs			CCN			Size distribution		
			from Dataset	Addtional QC	finalized data	from Dataset	Addtional QC	finalized data	from Dataset	Addtional QC	finalized data
SMEAR2	8784	N_total	--	--	8626	--	--	6973-6994	--	--	8461
		Percentage	--	--	98.2%	--	--	79.3-79.6%	--	--	96.3%
SORPES	8760	N_total	--	--	5266	--	--	4825~4906	--	--	5440
		Percentage	--	--	60.1%	--	--	55.1~56%	--	--	62.1%
ASI	12144	N_total	11851	84	11767	9894-10343	--	9894-10343	10931	32	10899
		Percentage	97.6%	0.7%	96.9%	81.5-85.2%	--	81.5-85.2%	90.0%	0.3%	89.7%
PVC	1800	N_total	1637	--	1637	1495	--	1495	1730	0	1730
		Percentage	90.9%	--	90.9%	83.1%	--	83.1%	96.1%	0.0%	96.1%
MAO	8160	N_total	7532	--	7532	7574-7653	55-123	7507-7541	7587	56	7541
		Percentage	92.3%	--	92.3%	92.8~93.8%	0.7-1.5%	92-92.4%	93.0%	0.7%	92.4%
PGH	3498	N_total	3453	--	3453	3380-3420	--	3380-3420	--	--	--
		Percentage	98.7%	--	98.7%	96.6-97.8%	--	96.6-97.8%	--	--	--

S2. Application of Reduced Major Axis (RMA) regression

The Matlab code of Trujillo-Ortiz and Hernandez-Walls (2010) was applied to calculate the reduced major axis (RMA) regressions of $R_{CCN/\sigma}$ vs. BSF to get the slope and offset (a and b, respectively) of $R_{CCN/\sigma} = a$ BSF + b at the supersaturations (SS) of the CCN counters at the six stations. The results are shown in Table S2. The values of Table S2 were plotted as a function of SS in Fig. S1 where also the fittings to the data are shown.

Table T2. Slopes (a) and offsets (b) of $R_{CCN/\sigma} = a$ BSF + b obtained with RMA. The unit of the coefficients is $[N_{CCN}]/[\sigma_{sp}]$ =cm^{-3}/Mm^{-1}.

Station	SS(%)	a	(a_{LOW} - a_{HIGH})	b	(b_{LOW} - b_{HIGH})
SMEAR II	0.1	175	(170 - 181)	-15.0	(-15.8 - -14.3)
	0.2	511	(502 - 521)	-49.8	(-51.2 - -48.5)
	0.5	1031	(1011 - 1050)	-110.1	(-112.9 - -107.3)
	1	1492	(1459 - 1525)	-164.4	(-169.1 - -159.7)
SORPES	0.1	121	(117 - 125)	-9.1	(-9.5 - -8.7)
	0.2	333	(326 - 341)	-25.8	(-26.6 - -25.0)
	0.4	657	(643 - 671)	-53.0	(-54.6 - -51.5)
	0.8	926	(905 - 946)	-76.6	(-78.9 - -74.4)
PGH	0.12	-53	(-54.6 - -51)	5.1	(5.0 - 5.2)
	0.22	161	(156 - 167)	-6.9	(-7.3 - -6.5)
	0.48	712	(689 - 734)	-37.6	(-39.2 - -36.0)
	0.78	849	(823 - 876)	-44.1	(-46.0 - -42.3)
PVC	0.15	517	(500 - 534)	-42.4	(-44.5 - -40.3)
	0.25	989	(956 - 1023)	-85.8	(-89.9 - -81.7)
	0.4	1465	(1416 - 1514)	-130.7	(-136.7 - -124.7)
	1	2452	(2369 - 2536)	-223.5	(-233.7 - -213.3)
MAO	0.25	472	(462 - 481)	-46.7	(-48.1 - -45.4)
	0.4	833	(817 - 849)	-83.4	(-85.6 - -81.1)
	0.6	1188	(1163 - 1213)	-122.1	(-125.6 - -118.7)
	1.1	2128	(2065 - 2190)	-226.5	(-234.9 - -218.2)
ASI	0.1	150	(147 - 153)	-15.9	(-16.3 - -15.4)
	0.2	319	(312 - 325)	-34.0	(-34.9 - -33.1)
	0.4	372	(365 - 380)	-39.8	(-40.9 - -38.7)
	0.8	406	(397 - 414)	-42.4	(-43.6 - -41.1)

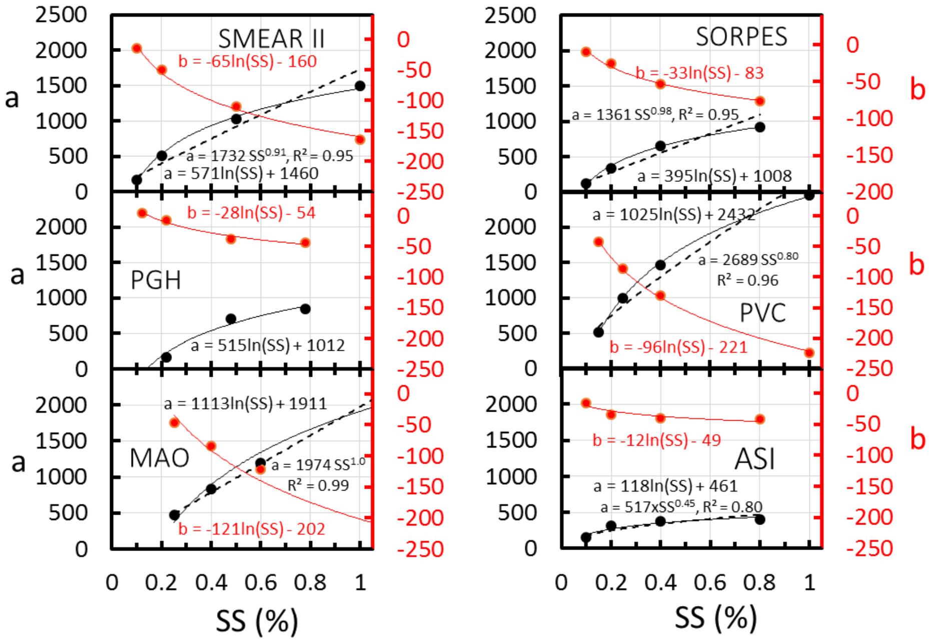

Figure S1. The RMA-derived coefficients a and b of each station (Table S2) as a function of supersaturation. Two types of functions, a logarithmic and a power fuction were fitted to the coefficient a, to coefficient b only a logaritmic function. The squared correlation coefficients R^2 are shown only for the power function fittings, for the logarithmic fittings they were all > 0.99. The unit of the coefficients is $[N_{CCN}]/[\sigma_{sp}]$ =cm^{-3}/Mm^{-1}.

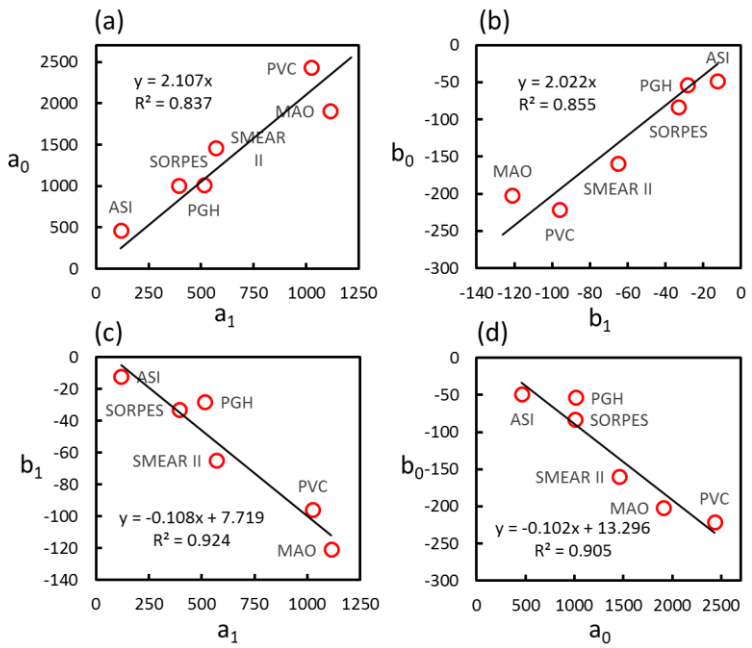

Figure S2. Relationship between the coefficients a_0, a_1, b_0 and b_1 shown in Fig. S1 that were obtained from the fitting of $a = a_1\ln(SS) + a_0$ and $b = b_1\ln(SS) + b_0$ with the data in Table S2. a) a_0 vs. a_1, b) b_0 vs. b_1, c) b_1 vs. a_1, d) b_0 vs. a_0. The unit of the coefficients is $[N_{CCN}]/[\sigma_{sp}]$ = cm^{-3}/Mm^{-1}.

When using RMA-derived slopes and offsets of $R_{CCN/\sigma} = a\ BSF + b$ the relationship between the factor a_1 and SAE became $a_1 \approx 391 \cdot SAE_{10}$ (Fig. S3).

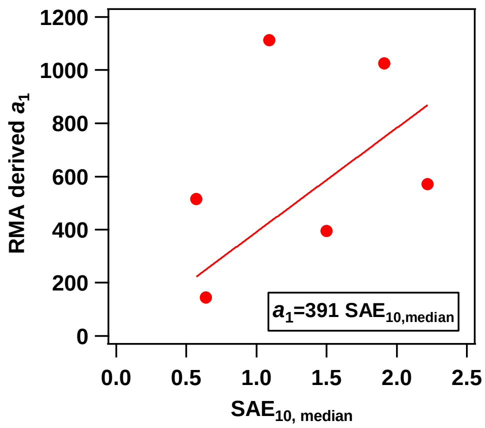

Figure S3: Relationship between RMA-derived a_1 and $SAE_{10,median}$.

This was further used to estimate CCN number concentration in the formula

$$N_{CCN}(RMA) \approx \left(\ln\left(\frac{SS}{0.12 \pm 0.02} \right) a_1 (BSF - BSF_{\min}) + R_{\min} \right) \sigma_{sp} \tag{S1}$$

The derivation of Eq. (S1) is presented in supplement S4. N_{CCN}(RMA) is in general in agreement with the $N_{CCN}(AOP_2)$ and N_{CCN}(meas). However for SS~0.1% the performance of RMA method is poor. At SS~0.1%, R^2 between N_{CCN}(RMA) and N_{CCN}(meas) is much lower than between $N_{CCN}(AOP_2)$ and N_{CCN}(meas) which indicates using RMA gives very uncertain results ast lowest SS. Nevertheless, for SS>0.15%, OLS-derived $N_{CCN}(AOP_2)$ and RMA-derived N_{CCN}(RMA) agree well. Figure S4 shows the scatter plots for N_{CCN}(RMA) vs. N_{CCN}(meas) and R^2 and bias. The R^2 are between 0.5~0.85 and bias are within 0.5~2 when SS>0.15% for $N_{CCN}(AOP_2)$.

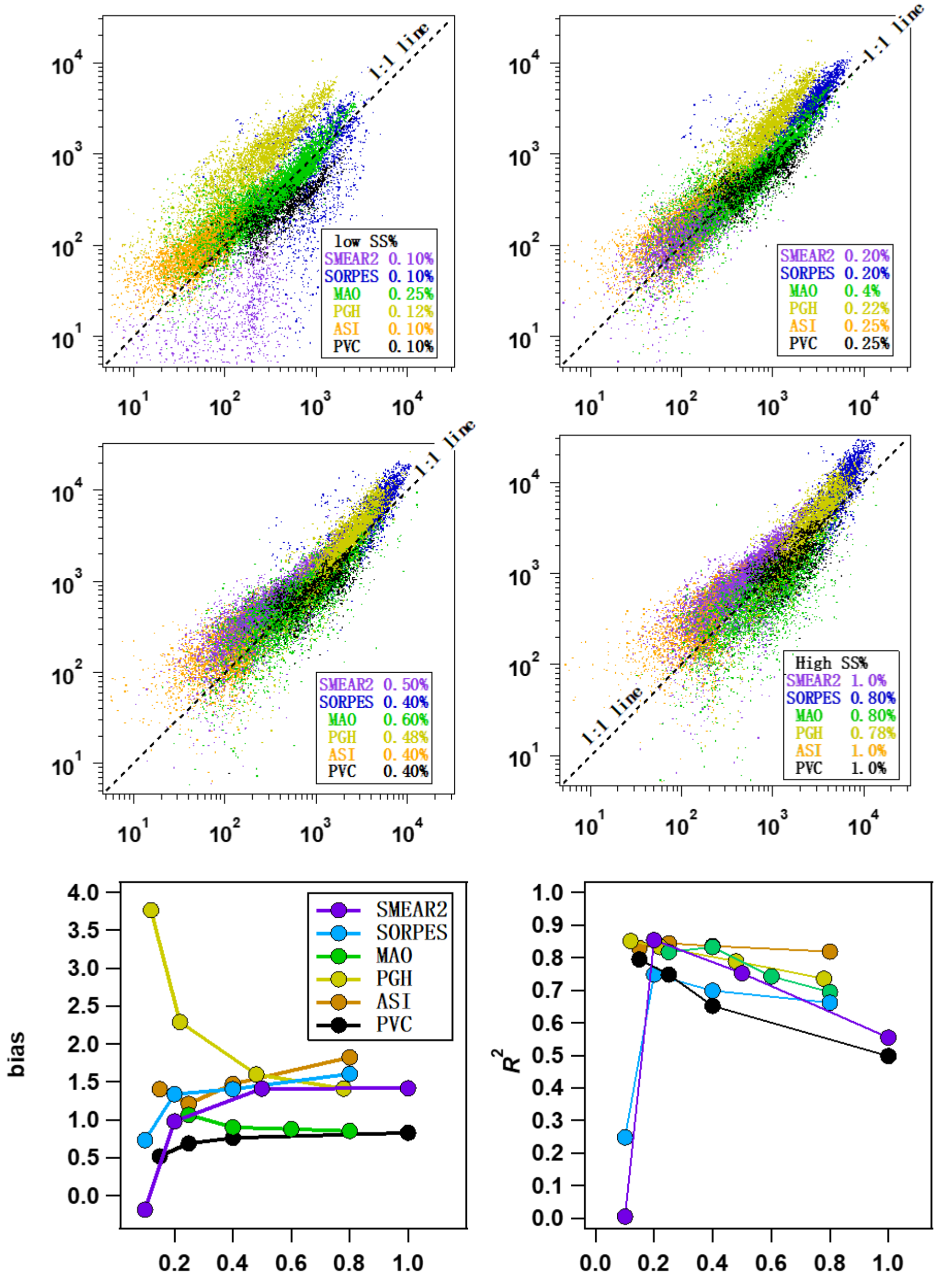

Figure S4. Statistics of N_{CCN}(RMA) from parameterization in Eq. (ES1). N_{CCN}(RMA) vs. N_{CCN}(meas) at different sites at different supersaturations, bias = N_{CCN}(RMA)/ N_{CCN} (meas) at different sites and supersaturations, and R^2 of the linear regression of N_{CCN}(RMA) vs. N_{CCN} (meas) at different sites and supersaturations. same as Figure 8, but for N_{CCN}(RMA).

The choice between OLS and RMA

Many studies use the reduced major axis (RMA) method instead of ordinary least squares (OLS) method to define a line of best fit for a bivariate relationship when variable represented on the X-axis contains measurement error. Smith (2009) point out that the major difference RMA and OLS is not in the difference in the assumption made about the distribution of error, but in their symmetry/asymmetry property. The reduced major axis regression is to describe the symmetric relationship between two variables and not for predictive use of the variable x with respect to y or y with respect to x (Smith, 2009). For predictive use OLS is preferred.

References

Smith, R. J.: Use and Misuse of the Reduced Major Axis for Line-Fitting, Am. J. Phys. Anthropol., 140, 476–486, doi:10.1002/ajpa.21090, 2009

Trujillo-Ortiz, A. and Hernandez-Walls, R.: gmregress: Geometric Mean Regression (Reduced Major Axis Regression), a MATLAB file available at: http://www.mathworks.com/matlabcentral/fileexchange/27918-gmregress, 2010.

S3. Derivation of Equation (8)

a) Using slopes and offsets from ordinary linear regressions

$$N_{CCN}(AOP) = \left(a_{SS}BSF + b_{ss}\right)\sigma_{sp} = \left(\left(a_1\ln(SS) + a_0\right)BSF + b_1\ln(SS) + b_0\right)\sigma_{sp}$$

$$R_{CCN/\sigma} = \frac{N_{CCN}(AOP)}{\sigma_{sp}} = a_{SS}BSF + b_{ss} = \left(a_1\ln(SS) + a_0\right)BSF + b_1\ln(SS) + b_0$$

Linear regressions of the coefficients in Table 2 yield

$$a_0 \approx (2.38 \pm 0.06)a_1, b_0 \approx (2.33 \pm 0.03)b_1, b_1 \approx -(0.096 \pm 0.013)a_1 + (6.0 \pm 5.9)$$

$$\Rightarrow$$

$$a_1\ln(SS) + a_0 \approx a_1\ln(SS) + (2.38 \pm 0.06)a_1 \approx a_1(\ln(SS) + (2.38 \pm 0.06))$$

$$b_1\ln(SS) + b_0 \approx b_1\ln(SS) + (2.33 \pm 0.03)b_1 = b_1(\ln(SS) + (2.33 \pm 0.03))$$

$$\approx \left(-(0.096 \pm 0.013)a_1 + (6.0 \pm 5.9)\right)(\ln(SS) + (2.33 \pm 0.04))$$

$$\Rightarrow$$

$$R_{CCN/\sigma} = \left(a_1\ln(SS) + a_0\right)BSF + b_1\ln(SS) + b_0$$

$$\approx a_1(\ln(SS) + (2.38 \pm 0.06))BSF + \left(-(0.096 \pm 0.013)a_1 + (6.0 \pm 5.9)\right)(\ln(SS) + (2.33 \pm 0.03))$$

Approximation, since (2.33±0.03) $\approx$ (2.38±0.06)

$$\Rightarrow$$

$$R_{CCN/\sigma} \approx a_1\left(\ln(SS) + (2.38 \pm 0.06)\right)BSF - (0.096 \pm 0.013)a_1\left(\ln(SS) + (2.38 \pm 0.07)\right) + (6.0 \pm 5.09)\left(\ln(SS) + (2.38 \pm 0.06)\right)$$

$$\approx a_1\left(\ln(SS) + (2.38 \pm 0.06)\right)(BSF - (0.096 \pm 0.013)) + (6.0 \pm 5.9)\left(\ln(SS) + (2.38 \pm 0.06)\right)$$

$$\approx \left(\ln(SS) + (2.38 \pm 0.06)\right)\left(a_1(BSF - (0.096 \pm 0.013)) + (6.0 \pm 5.9)\right)$$

$$\approx \left(\ln(SS) - \ln(0.093 \pm 0.006)\right)\left(a_1(BSF - (0.097 \pm 0.013)) + (6.0 \pm 5.9)\right)$$

$$\approx \ln\left(\frac{SS}{0.093 \pm 0.006}\right)\left(a_1(BSF - (0.096 \pm 0.013)) + (6.0 \pm 5.9)\right)$$

b) Using slopes and offsets from reduced major axis regressions

$$N_{CCN}(RMA) = \left(a_{SS}BSF + b_{ss}\right)\sigma_{sp} = \left(\left(a_1\ln(SS) + a_0\right)BSF + b_1\ln(SS) + b_0\right)\sigma_{sp}$$

$$R_{CCN/\sigma} = \frac{N_{CCN}(RMA)}{\sigma_{sp}} = a_{SS}BSF + b_{ss} = \left(a_1\ln(SS) + a_0\right)BSF + b_1\ln(SS) + b_0$$

RMA regressions $\Rightarrow a_0 \approx (2.11 \pm 0.16)a_1, b_0 \approx (2.02 \pm 0.16)b_1, b_1 \approx -(0.108 \pm 0.016)a_1 + (7.7 \pm 11.0)$

The same steps as above in (a) $\Rightarrow$

$$R_{CCN/\sigma} \approx \left(\ln(SS) + (2.11 \pm 0.16)\right)\left(a_1(BSF - (0.108 \pm 0.016)) + (7.7 \pm 11.0)\right)$$

$$\approx \left(\ln(SS) - \ln(0.12 \pm 0.02)\right)\left(a_1(BSF - (0.108 \pm 0.016)) + (7.7 \pm 11.0)\right)$$

$$\approx \ln\left(\frac{SS}{0.12 \pm 0.02}\right)\left(a_1(BSF - (0.11 \pm 0.02)) + (8 \pm 11)\right)$$

S4. Derivation of Equation (9)

If the original slopes and offsets were calculated using ordinary linear regressions

$$N_{CCN}(AOP) \approx \ln\left(\frac{SS}{0.093 \pm 0.006}\right)\left(a_1(BSF - BSF_{\min}) + \mathrm{C}\right)\sigma_{sp},$$

where C is an unknown constant.

If $BSF = BSF_{\min}$

$$\Rightarrow a_1(BSF - BSF_{\min}) = 0$$

$$\Rightarrow N_{CCN}(AOP, BSF_{\min}) \approx \ln\left(\frac{SS}{0.093 \pm 0.006}\right)\mathrm{C}\cdot\sigma_{sp}$$

$$\Leftrightarrow \mathrm{C} \approx \frac{1}{\ln\left(\frac{SS}{0.093 \pm 0.006}\right)}\frac{N_{CCN}(AOP, BSF_{\min})}{\sigma_{sp}} \approx \frac{1}{\ln\left(\frac{SS}{0.093 \pm 0.006}\right)}R_{\min}$$

$\Rightarrow$

$$N_{CCN}(AOP) \approx \ln\left(\frac{SS}{0.093 \pm 0.006}\right)\left(a_1(BSF - BSF_{\min}) + \frac{1}{\ln\left(\frac{SS}{0.093 \pm 0.006}\right)}R_{\min}\right)\sigma_{sp}$$

$$\approx \left(\ln\left(\frac{SS}{0.093 \pm 0.006}\right)a_1(BSF - BSF_{\min}) + R_{\min}\right)\sigma_{sp}$$

If the original slopes and offsets were calculated using reduced major axis regressions

$$N_{CCN}(RMA) \approx \left(\ln\left(\frac{SS}{0.12 \pm 0.02}\right)a_1(BSF - BSF_{\min}) + R_{\min}\right)\sigma_{sp}$$

S5. Analysis of the uncertainty related to the number of samples

The following procedure was used for testing how different values would be change if the number of samples decrease.

1. For each site 2%,3%,5%,10%,20%,30%,50% and 100% of samples were taken from the whole period.
2. The slope and offset a, b, BSF_{min} (calculated as the 1st percentile of the BSF data) and SAE_{10},median were calculated from the randomly chose subsets.
3. The a, b, BSF_{min} and SAE_{10},median should be slightly different if the sub-set is different. Therefore the random sampling was repeated 100 times resulting in 100 different results
4. The averages and standard deviations of the 100 results were calculated and plotted below for all the sites. The average are the reds circles and the stds the error bars in the plots.

Results of the analysis

The averages of a,b, BSF_{min} and SAE_{10},median don't have clear dependence on the number of samples. However, the uncertainty is very large at low number of samples and decreases with increasing number of samples. The uncertainties depend on parameter and site. The plots suggest that if the number of samples is larger than 1000 the uncertainty is low enough. For example, the std of BSF_{min} is ~0.0005-0.005 and the std of SAE_{10},median is ~0.01-0.02. For a and b, std is ~10% of the a average value.

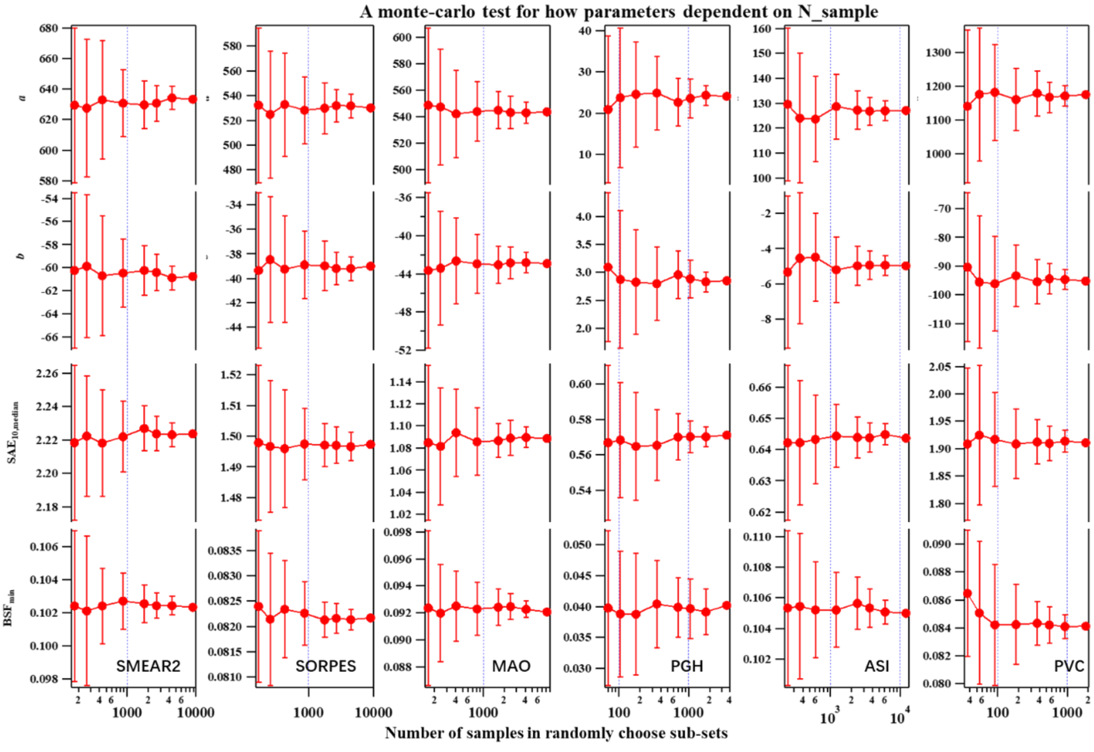

Figure S5. A monte-carlo test on the depencence of the parameters a, b, $SAE_{10,median}$ and BSF_{min} on the number of hourly-averaged samples. The average are the reds circles and the stds the error bars.

S6. N_{CCN}(AOP) calculated by using the site-specific median SAE

The general combined parameterization was presented in the main test as Eq.10:

$$N_{CCN}(AOP_2) \approx \left(a_1 \ln\left(\frac{SS}{0.093 \pm 0.006}\right)(BSF - BSF_{\min}) + R_{\min} \right)\sigma_{sp}$$
$$\approx \left((286 \pm 46)\mathrm{SAE} \cdot \ln\left(\frac{SS}{0.093 \pm 0.006}\right)(BSF - BSF_{\min}) + (5.2 \pm 3.3) \right)\sigma_{sp}$$

In the main text, we used SAE of hourly-averaged σ_{sp} to estimate $N_{CCN}(AOP_2)$. Here we give another alternative for using this formula by using the site-specific median SAE values (Table 4 in the main text). The N_{CCN}(AOP) calculated by using the site-specific median SAE is compared with N_{CCN}(meas) in Figure S6. When compared with N_{CCN}(AOP) calculated by using the hourly-varying SAE (Fig. 8 in the main text), it is obvious that the two approaches are competitive with each other. A comparison of the biases and correlation coefficients is presented in Table S3 below. For some combinations of SS and sites, the site-specific median SAE gives a smaller R^2 and a higher bias than the hourly SAE especially for ASI.

However, site-specific median SAE is very probably always positive, while the hourly SAE is sometimes negative which may yield negative N_{CCN}(AOP). For the 6 sites of this study, the fraction of negative SAE of all hourly data varied between 0-6%. To estimate N_{CCN} for a site with a large fraction of negative SAE, we recommend to use site-specific median SAE.

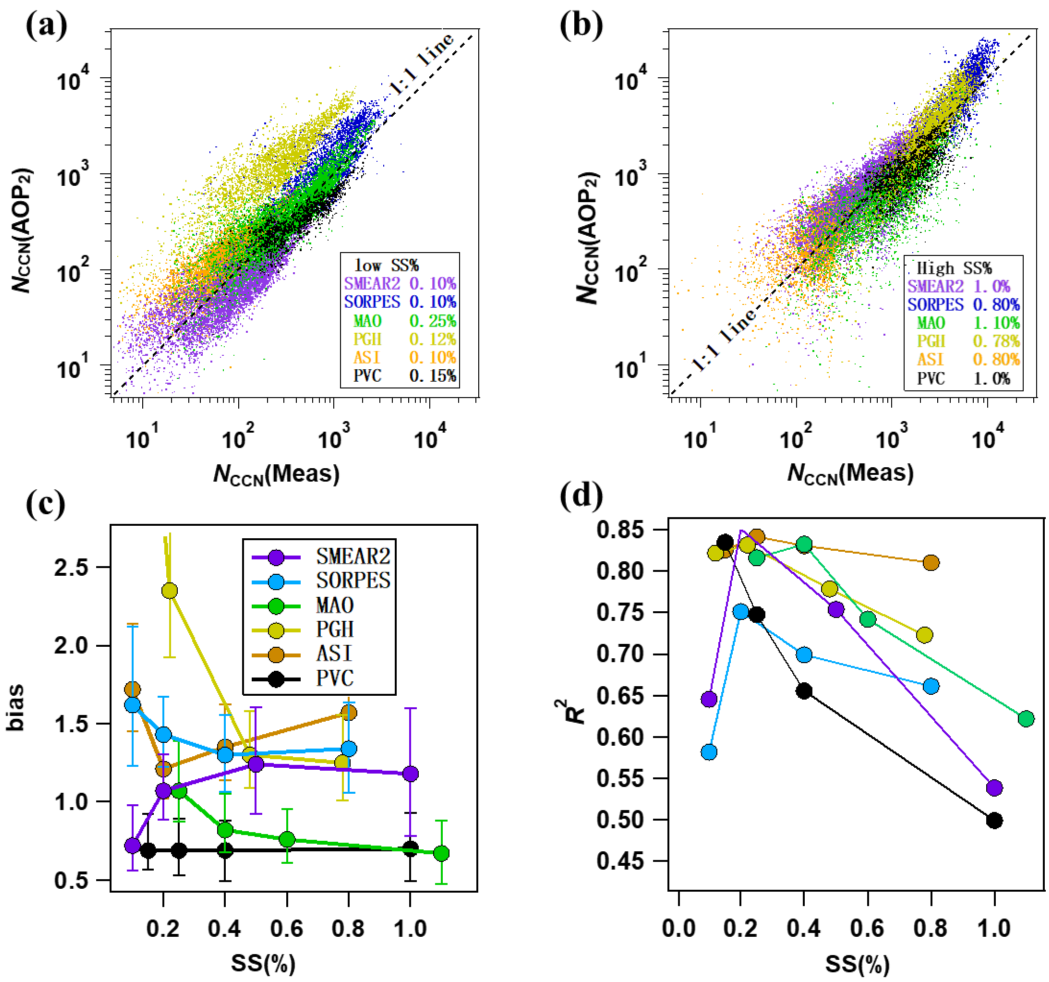

Figure S6. Same as Figure 8 in the main text, but N_{CCN}(AOP) calculated by using the site-specific median SAE. For details see the caption of Fig. 8 in the main text

Table S3. Performance of the general combined parametrization using SAE of hourly-averaged scattering coefficients and site-specific median SAE at the supersaturations of the CCN counters of each station.

Station	Fraction of hourly SAE < 0	SS	N_{CCN}(AOP) calculated using hourly-varying SAE R^2	bias	median SAE R^2	bias
SMEAR II	0.0%	0.10%	0.675	0.72	0.657	0.72
		0.20%	0.832	1.09	0.850	1.07
		0.50%	0.719	1.26	0.754	1.24
		1.00%	0.504	1.20	0.554	1.18
SORPES	0.0%	0.10%	0.595	1.61	0.587	1.62
		0.20%	0.773	1.36	0.751	1.43
		0.40%	0.650	1.22	0.699	1.30
		0.80%	0.636	1.27	0.687	1.34
MAO	6.0%	0.25%	0.840	1.24	0.816	1.07
		0.40%	0.834	0.97	0.832	0.82
		0.60%	0.725	0.91	0.742	0.76
		1.10%	0.583	0.71	0.622	0.67
PGH	4.4%	0.12%	0.852	4.53	0.821	4.71
		0.22%	0.871	2.13	0.832	2.35
		0.48%	0.784	1.13	0.779	1.30
		0.78%	0.703	1.07	0.723	1.25
ASI	0.04%	0.10%	0.872	1.92	0.828	1.72
		0.20%	0.923	1.41	0.844	1.21
		0.40%	0.900	1.61	0.836	1.35
		0.80%	0.857	1.90	0.818	1.57
PVC	0.3%	0.15%	0.880	0.71	0.835	0.69
		0.25%	0.780	0.70	0.747	0.69
		0.40%	0.687	0.71	0.655	0.69
		1.00%	0.519	0.71	0.499	0.70

The biases of $N_{CCN}(AOP_2)$ presented in Table S3 were calculated from the ratio $N_{CCN}(AOP_2)/N_{CCN}(meas)$. Since $N_{CCN}(AOP_2) \approx (a_1 \ln(SS/0.093)(BSF - BSF_{min}) + R_{min})\sigma_{sp}$ it is obvious that biases of a_1 affect the bias of $N_{CCN}(AOP_2)$. If we consider the a_1 values in the main text Table 4 as the accurate station-specific values then the fitted line $a_1 = 286 \cdot SAE$ overestimates or underestimates a_1 by +37%, +30%, -20%, -32%, -20% and +251% for SMEAR II, SORPES, PGH, PVC, MAO and ASI, respectively. These values were calculated from $100\%(286 \cdot SAE - a_1)/a_1$. The biases of a_1 calculated from $286 \cdot SAE/a_1$ are therefore 1.373, 1.295,0.796 ,0.675, 0.792, 3.509 for the respective stations. The average biases of $N_{CCN}(AOP_2)$ at all supersaturations of each station presented in Table TS3 are compared with the biases of a_1 in Figure S7. For each station two values are shown: the average bias of $N_{CCN}(AOP_2)$ calculated by using the median SAE of each station and the average bias of $N_{CCN}(AOP_2)$ calculated by using the hourly-varying SAE. For PGH the average bias of $N_{CCN}(AOP_2)$ at all supersaturations and at SS> 0.3% are shown because the biases at the lowest supersaturations are anomalously high. The plot shows that for most stations the bias of $N_{CCN}(AOP_2)$ can be explained by the bias of a_1: when a_1 is underestimated so is $N_{CCN}(AOP_2)$ and when a_1 is overestimated so is $N_{CCN}(AOP_2)$. PGH is the only exception to this, especially at the lowest two supersaturations (SS = 0.12% and 0.22%) and we cannot explain why. For ASI the bias of $N_{CCN}(AOP_2)$ is clearly smaller than the bias of a_1. This would happen when in the formula $N_{CCN}(AOP_2) \approx (a_1 \cdot \ln(SS/0.093)(BSF - BSF_{min}) + R_{min})\sigma_{sp}$ both SAE and BSF are very small and especially when BSF is close to BSF_{min}. Both of these would take place when aerosol is dominated by large aerosols. This is true especially for ASI, a site dominated by marine aerosols.

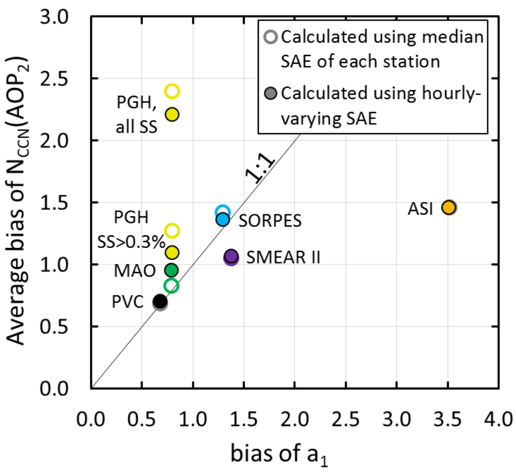

Figure S7. Biases of $N_{CCN}(AOP_2)$ vs the bias of a_1 calculated from $a_1 = 286 \cdot SAE$. The biases of $N_{CCN}(AOP_2)$ are the averages of biases at all supersaturations presented in Table S3. For each station two values are shown: the average bias of $N_{CCN}(AOP_2)$ calculated by using the median SAE of each station (open circles) and the hourly-varying SAE (filled circles). For PGH the average bias of $N_{CCN}(AOP_2)$ at all supersaturations and at SS> 0.3.