Supporting information for : "Model analysis of biases in satellite diagnosed aerosol effect on cloud liquid water path"

Harri Kokkola[1], Juha Tonttila[2], Silvia Calderón[1], Sami Romakkaniemi[1], Antti Lipponen[1], Aapo Peräkorpi[3], Tero Mielonen[1], Edward Gryspeerdt[4], Timo H. Virtanen[1], Pekka Kolmonen[1], and Antti Arola[1]

[1]Atmospheric Research Centre of Eastern Finland, Finnish Meteorological Institute, Kuopio, Finland
[2]CSC - IT Center for Science, Espoo, Finland
[3]Datrix S.p.A., Milan, Italy
[4]Grantham Institute - Climate Change and the Environment, Imperial College London, London, UK

Correspondence: Harri Kokkola (harri.kokkola@fmi.fi)

S1   Model initial settings

Simulations in UCLALES-SALSA require three main inputs, vertical profiles of atmospheric properties in the well-mixed layer at the beginning of the cloud event, aerosol composition and size distribution, and vertical profiles of background atmospheric properties to resolve the radiative transfer equation. Figure S1 summarizes sounding profiles for the cloud study derived from radiosonde measurements during flight RF-02 in the DYCOMS-II field campaign (Ackerman et al., 2009). Figure S2 shows size distributions used as initial aerosol loadings in our simulation scenarios, all of them share a common Aitken mode but have increasing number concentrations for the accumulation mode. In this way, we explore the effect of increasing CCN (i.e. in this study CCN are considered equivalent to particles in the accumulation mode) on cloud properties. For the purpose of comparison, aerosol number concentrations for particles with dry diameter above $0.1\,\mu m$ or $100\,nm$, N_{100} are $49.4\,cm^{-3}$, $114.05\,cm^{-3}$ and $228.1\,cm^{-3}$, respectively.

Background atmospheric properties include vertical profiles from surface to the top of the atmosphere for pressure, temperature and mixing ratio of water vapor and ozone. Simulations were performed assuming constant surface fluxes of sensible and latent heat equal to 16.0 and $94.0\,Wm^{-2}$ respectively; as well as constant divergence of the large- scale horizontal winds equal to $3.75{\times}10^{-6}\,s^{-1}$ as in previous cloud modelling studies for this case (Ackerman et al., 2009).

Horizontal and vertical resolutions were set to be $75\,m$ and $20\,m$, respectively. Equations were resolved with a maximum time step of $1\,s$ using an hourly period for spin-up without nudging. Each simulation scenario ran for $11\,h$ and model outputs were saved when the total simulation time reached $2\,h$, $6\,h$ and $10\,h$.
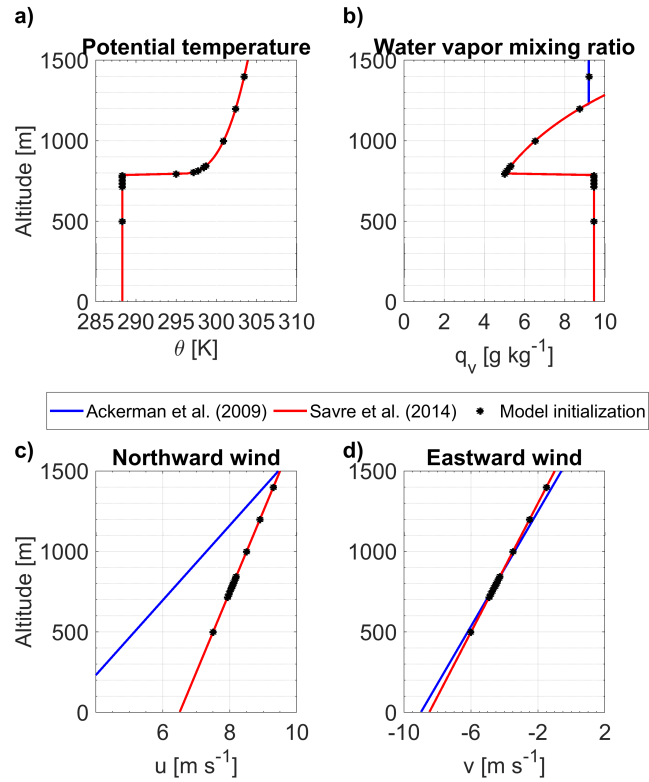
Figure S1. Vertical profiles of atmospheric properties used for model initialization in simulations of the DYCOMS-II RF02 cloud case (Ackerman et al., 2009)
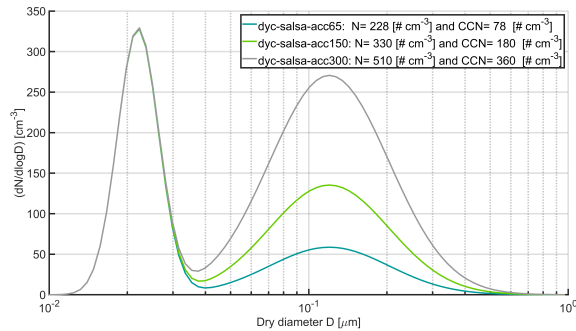
Figure S2. Aerosol size distributions used for model initialization in simulations of the DYCOMS-II RF02 cloud case. Values for total aerosol (N) and cloud condensation nuclei (CCN) number concentrations are shown to distinguish different simulation scenarios. Number concentrations of CCN were assumed to be equivalent to N_{100} values or number concentrations i.e., aerosol particles with dry size equal to or above 100 nm following common procedures in the cloud modelling community.

S2 Conditional sampling of modelled cloud properties

In this study, we follow a sampling methodology that mimics the so-called penetration depth bias (Grosvenor et al., 2018). We determine CER and CDNC values for a cloud upper region that we refer here in this study as the *extended cloud top* in which COT decreases by 3 units from its maximum value. Properties of the extended cloud top were calculated as extinction coefficient weighted average values. Expressions used to calculate CER and CDNC in the extended cloud top are

$$\mathrm{CER}_{\mathrm{model}} = \frac{\sum_{i=1}^{\mathrm{layers}} (b_{\mathrm{ext,i}} \mathrm{CER}_{\mathrm{i}})}{\sum_{i=1}^{\mathrm{layers}} b_{\mathrm{ext,i}}}, \tag{S.1}$$

and

$$\mathrm{CDNC}_{\mathrm{model}} = \frac{\sum_{i=1}^{\mathrm{layers}} (b_{\mathrm{ext,i}} \mathrm{CDNC}_{\mathrm{i}})}{\sum_{i=1}^{\mathrm{layers}} b_{\mathrm{ext,i}}}, \tag{S.2}$$

where the index i runs from 1 to layers to cover all layers in the extended cloud top.

The UCLALES-SALSA model version used in this study does not provide extinction coefficients neither cloud optical thickness as direct outputs (although used for internal calculations). For the purpose of validation, extinction coefficients were calculated off-line for cloudy layers with LWC is above $0.01\,\mathrm{gm^{-3}}$ using two approaches, a Mie-theory based approach and the parameterization of Savijärvi et al. (1997).

In the Mie-theory based approach, extinction coefficients b_{ext} were calculated at visible and near infrared bands for each model layer using wet diameters for cloud droplet and precipitation droplet bins as follows

$$b_{ext}(\lambda, t, z, x, y) = \sum^{\mathrm{cbins}} \frac{\pi}{4} (D_{\mathrm{wcba}}(t, x, y, z))^2 Q_{ext}(m, \xi) Ncba(t, z, x, y) +$$
$$\sum^{\mathrm{pbins}} \frac{\pi}{4} (D_{\mathrm{wpba}}(t, x, y, z))^2 Q_{ext}(m, \xi) Npba(t, z, x, y), \tag{S.3}$$

where D_{wcba} corresponds to the wet diameter of cloud droplets formed from aerosols and D_{wpba} corresponds to the wet diameter of precipitation droplets, while $Q_{ext}(m, \xi)$ is the dimensionless extinction efficiency that represents the ratio between the extinction cross section (optical shadow) to the geometric cross section (geometric shadow). This variable depends on the complex refractive index m and the dimensionless size parameter ξ that relates the ratio of the droplet diameter to the wavelength of light (Seinfeld and Pandis, 2016). Extinction efficiencies of water droplets were interpolated from a lookup table that matches droplet diameter with extinction efficiencies calculated according to the Mie theory using the algorithm of Wiscombe (1980) included as an auxiliary calculation routine "MIEV" in LibRadTran.

These extinction coefficients were used to determine COT as follows

$$\mathrm{COT}(\lambda, t, x, y) = \int_{\mathrm{CB}}^{\mathrm{CT}} b_{\mathrm{ext}}(\lambda, t, x, y, z)\, dz, \tag{S.4}$$

where b_{ext} is the extinction coefficient, λ is the wavelength, CB and CT correspond to altitude of the cloud top and cloud base; defined in this study as the minimum and maximum altitudes at which the liquid mixing ratio is above a threshold value of LWC>0.01 g m^{-3}.

Table S1. Model parameters to obtain cloud optical properties with the parameterization of Savijärvi et al. (1997)

Cloud property	Wavelength [μm]	Equation
Extinction coefficient b_{ext}	2.38-4.00	LWC $\left(\frac{1.50\mu\text{mm}^2\text{g}^{-1}}{r_{\text{eff}}} + \frac{2.0\mu\text{m}^2\text{m}^2\text{g}^{-1}}{r_{\text{eff}}^2} \right)$
Asymmetry parameter g	0.25-0.69	$0.841 + 2.08 \times 10^{-3}\mu\text{m}^{-1}r_{\text{eff}}$
Single-scattering coefficient ω	0.25-0.69	$1 - 5.58 \times 10^{-7} - 1.25 \times 10^{-7}\mu\text{m}^{-1}r_{\text{eff}}$
Cloud optical thickness (COT) ω	0.25-0.69	LWP $\left(\frac{1.50\mu\text{mm}^2\text{g}^{-1}}{r_{\text{eff}}} + \frac{0.50\mu\text{m}^2\text{m}^2\text{g}^{-1}}{r_{\text{eff}}^2} \right)$
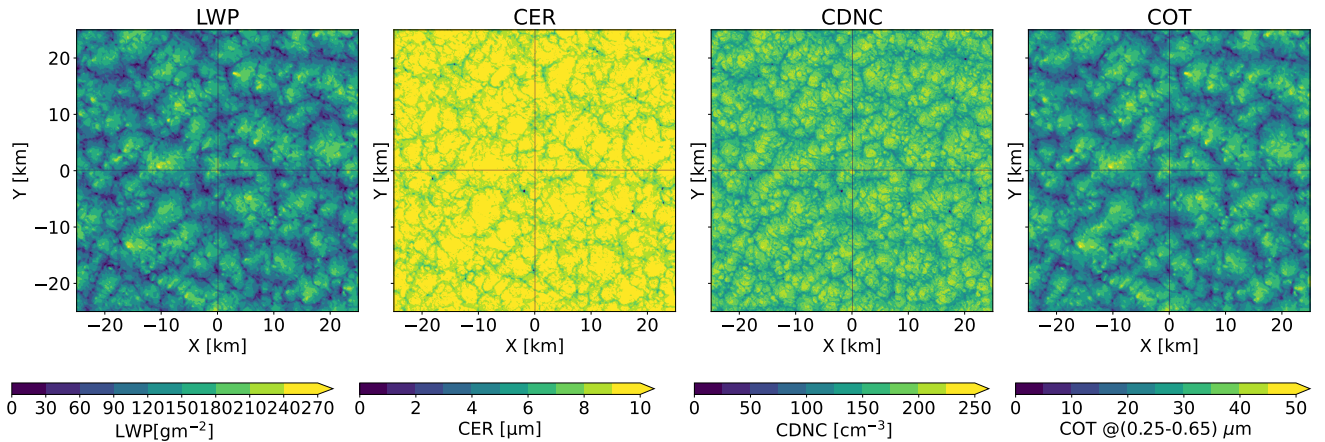
Figure S3. Modelled cloud properties for an initial CCN loading of $360\,\text{cm}^{-3}$ at the time instant of $10\,\text{h}$. Fields for cloud effective radius CER, Equation (S.1) and cloud droplet number concentration CDNC Equation (S.2) correspond to the extended cloud top region.

COT values in the visible wavelength were also obtained with the parameterization of Savijärvi et al. (1997) that is based on the two–stream approximation of the solar radiative transfer equation. Model expressions and parameters for cloud optical properties are summarized in Table S1. We did not find significant differences between COT values derived from droplet microphysics and those from the parametrization. Additional cloud optical properties such as the asymmetry parameter, single-scattering coefficient and cloud albedo were calculated and included in the model outputs for the purpose of completeness.

Figure S3 depicts property fields of LWP and COT in correlation with CDNC and CER in the extended cloud top region. Fields correspond to a late simulation time when the cloud structure is changing from closed to open cells introducing horizontal heterogeneity in cloud field properties.

S3 Surrogates of satellite-retrievals for CDNC and CER

Modelled CER and COT values from Equations (S.1) and (S.4) were used to calculate CDNC with two different equations commonly used in satellite data analysis. Both expressions are based on the pseudo-adiabatic cloud model whose main assumptions are that the liquid water content increases linearly with height being a constant fraction f_{ad} of the adiabatic liquid water content (i.e. negligible effects of temperature and pressure changes on the adiabatic condensational lapse rate) and, that there is homogeneous mixing and thus the droplet number concentration is vertically constant with an effective value of k and CDNC (e.g., Brenguier et al., 2000; Wood, 2006).

In the first approach, satellite retrievals of CDNC are calculated using Equation (1) in the main paper with explicit consideration of the adiabatic factor, the water condensational lapse rate and the breadth of the droplet distribution besides the dependencies to COT and CER.

Following common practices in satellite data analysis, we assumed $Q_{ext} = 2$ and $\rho_w = 1000\,\mathrm{kg m^{-3}}$. Values of k and c_w used in this equation were calculated as extinction coefficient weighted average values in a similar way as it was shown for CER in Equation (S.1).

Condensational lapse rate values c_w were calculated at every model layer of the selected upper cloud region using properties of the moist air as

$$c_w = \left((\epsilon + r) \frac{r \gamma_a L_v}{R_d T P} - \frac{r P g}{(P - p) R_d T P} \right) \rho_a, \tag{S.5}$$

where r is water vapor mass mixing ratio, p is vapor pressure, P is total pressure, T is temperature, R_d is the gas constant for dry air, γ_a is the moist adiabatic lapse rate (Stull, 2017), ϵ is the ratio between the gas constant for dry air and pure water vapor, g is the gravitational acceleration and L_v is the latent heat of water vaporization. We included temperature effects on water vapor pressure and latent heat of vaporization.

The adiabatic factor was calculated as the ratio between the modelled adiabatic liquid water path and the adiabatic liquid water path $\mathrm{LWP_{adiab}}$ calculated as

$$f_{ad} = \frac{\mathrm{LWP}}{0.5 c_{w,model} H^2 + \mathrm{LWC_{model,CB}} H}, \tag{S.6}$$

where $c_{w,model}$ is the water condensational lapse rate in the extended cloud top region, H is the cloud geometrical thickness or the difference between the maximum and minimum altitudes at which the liquid mixing ratio LWC>0.01 g m^{-3} (Brenguier et al., 2000). In this study, we introduced a correction factor to account for our definition of the cloud base (i.e. at the cloud base altitude, the liquid water content $\mathrm{LWC_{model,CB}}$ was assumed to be 0.01 gm^{-3}).

Biases between modelled values (X as CDNC or LWP) and emulated satellite retrievals were calculated as relative deviations from the model

$$\Delta(X) = \frac{X_{model} - X_{satellite}}{X_{model}} \times 100. \tag{S.7}$$

Figure S4 depicts variable fields for CER and adiabatic fraction, main variables used in Equation (1) and Equation (2) in the main text as well as biases in CDNC values. Satellite-based biases in CDNC values are always positive despite the fact
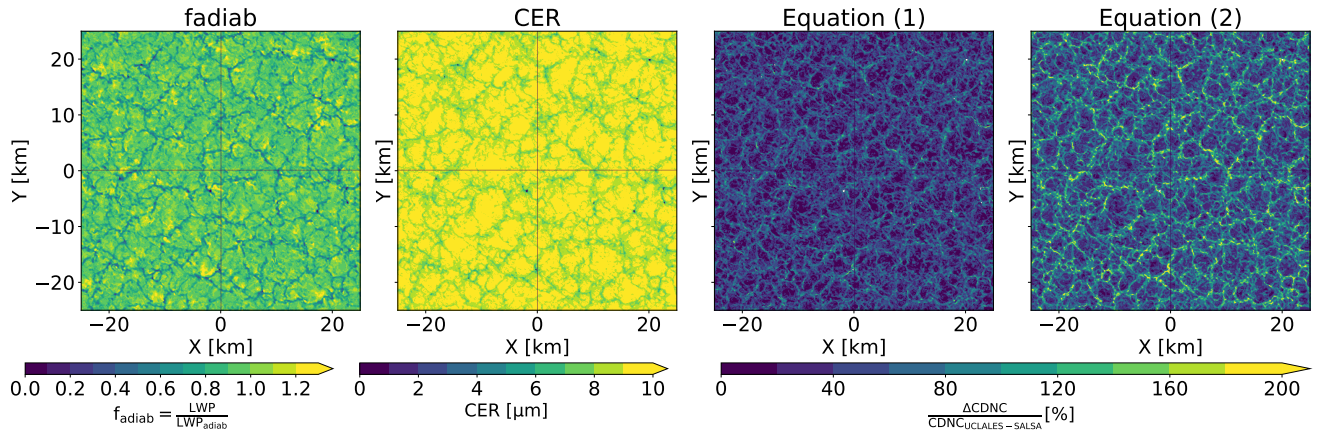
Figure S4. Biases between modelled and surrogate satellite-retrievals of CDNC using Equation (1) and Equation (2) in a simulation initialized with a CCN loading of $360\,\mathrm{cm}^{-3}$ at the time instant of $10\,\mathrm{h}$.
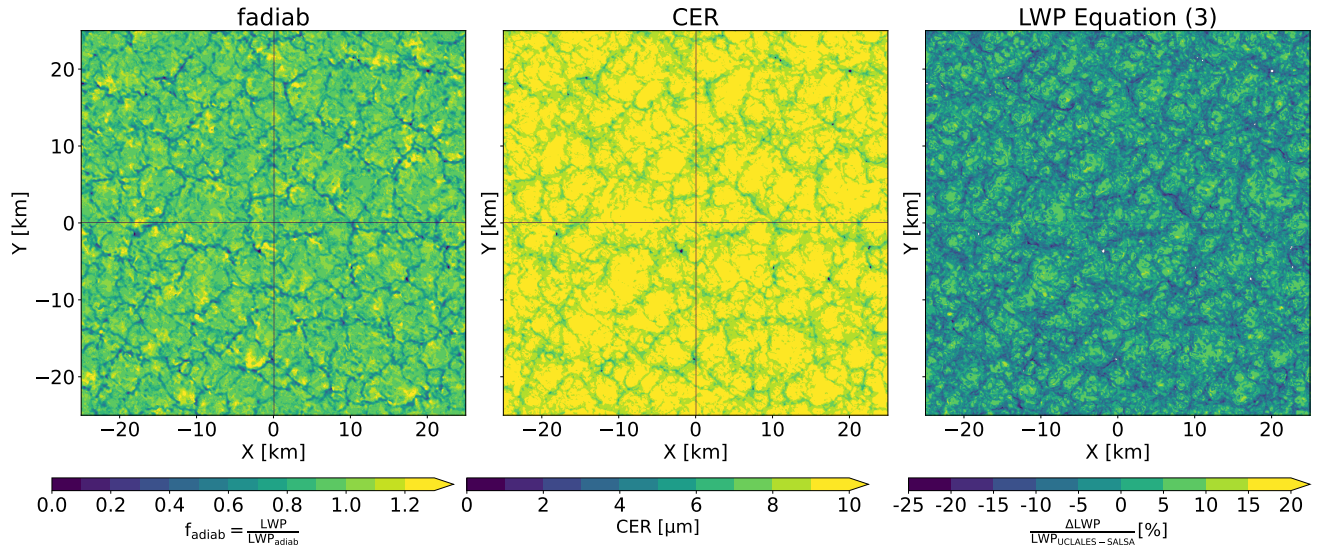
Figure S5. Biases between modelled and surrogate satellite-retrievals of LWP with Equation (3) in a simulation initialized with a CCN loading of $360\,\mathrm{cm}^{-3}$ at the time instant of $10\,\mathrm{h}$.

than modelled values should reflect the penetration depth bias. Biases are larger in the edges of cloud cells where cloud get

thinner showing smaller adiabatic fraction values (i.e. smaller LWP values compared to adiabatic) and CER values. Biases are significantly larger for CDNC retrievals with Equation (2) that considers a very simplified pseudo-adiabatic model. However, despite the larger degree of sophistication of Equation (1), positive biases in CDNC still remain and show an average value of 38.6 % compared to an average value of 59 % for Equation (2).

Nonetheless, biases in LWP between the model and Equation (3) are significantly lower with larger values at cloud edges as can be seen in Figure S5. All datasets at different simulation times show similar trends suggesting that biases are caused by processes at cloud edges related to stratocumulus dissipation (e.g. evaporative cooling during cloud top mixing or lateral mixing) which are not considered in the pseudo-adiabatic cloud model from which satellite equations are derived.

S4 LWP susceptibility to changes in CDNC

We used joint histograms to study the LWP susceptibility to changes in CDNC using model-derived values and surrogates of satellite retrievals. The histograms were built using bin sizes determined with the Freedman Diaconis Estimator, a more robust method recommended for non-Gaussian distributions with extreme values.

Figures S6, S7, and S8 show the LWP susceptibility to CDNC variations at three different time instances, 2, 6, and 10 hours. Simulations with different aerosol loadings are colour coded in both, joint and marginal histograms; and mean values are represented with continuous lines. Marginal histograms of LWP for both, model-derived and satellite-surrogate values do not show significant deviations when CCN concentrations increase. However, there are positive biases between modelled and satellite-derived CDNC distributions. The histograms of satellite-surrogate CDNC are shifted to larger values and show less skewness and kurtosis compared to the modelled CDNC.

Joint histograms in panels a), b) and c) of S7 indicate an increase in the LWP susceptibility to changes in CDNC when CCN increases. Although, LWP and CDNC values correlate linearly on a logarithmic scale in the majority of the CDNC range, the variation in LWP with CDNC decreases at lower CDNC values, especially for the simulation scenario with the lowest CCN. This suggest that droplet growth and coalescence during drizzle formation dominate the LWP behaviour.

Joint histograms for both satellite-equations give an "inverted v" shape for the LWP variation as a function of CDNC variation regardless of the time instance or the CCN number concentration used in the simulation scenario. The negative branch appears at LWP/CDNC pair values with very low probability of occurrence in our data set. This suggests that the decreasing trend of LWP with increasing CDNC is a numerical artifact caused by averaging operations in CDNC concentrations that have a very small amount of data and may not reflect realistic cloud conditions due to extreme positive bias in satellite retrievals of CDNC at very small CER values.

Time instance of 2h
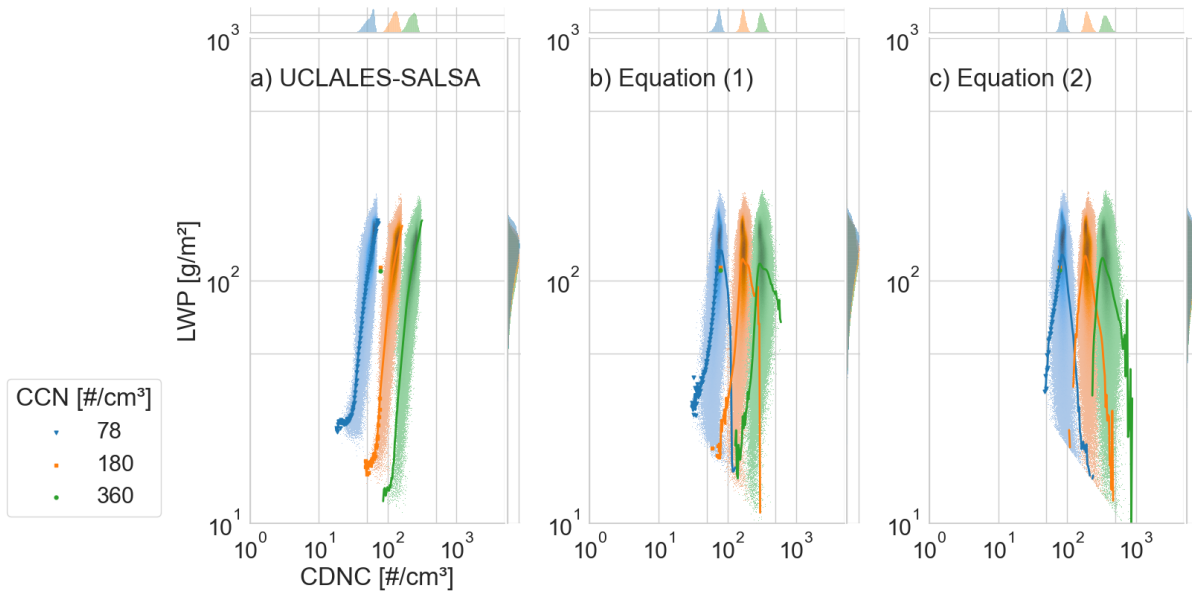
Figure S6. Joint and marginal histograms for LWP and CDNC values using a) UCLALES-SALSA b) Equation (1) and c) Equation (2) at a time instance of 2 hours. Simulations are colour coded according to CCN concentrations used in the model initialization. The intensity of colour in joint histograms increases when the probability increases. The probability is represented as a density function calculated as counts/sum(counts)/bin area. Continuous lines indicate the arithmetic mean.

Time instance of 6h
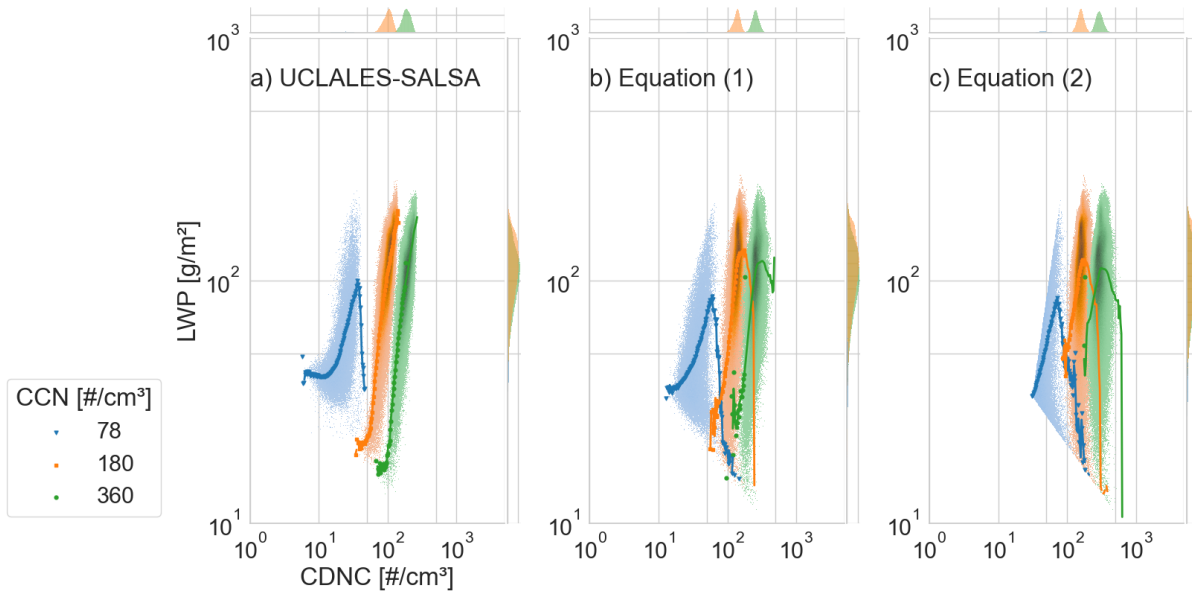
Figure S7. The same as Figure S6 for the time instance of 6 h.
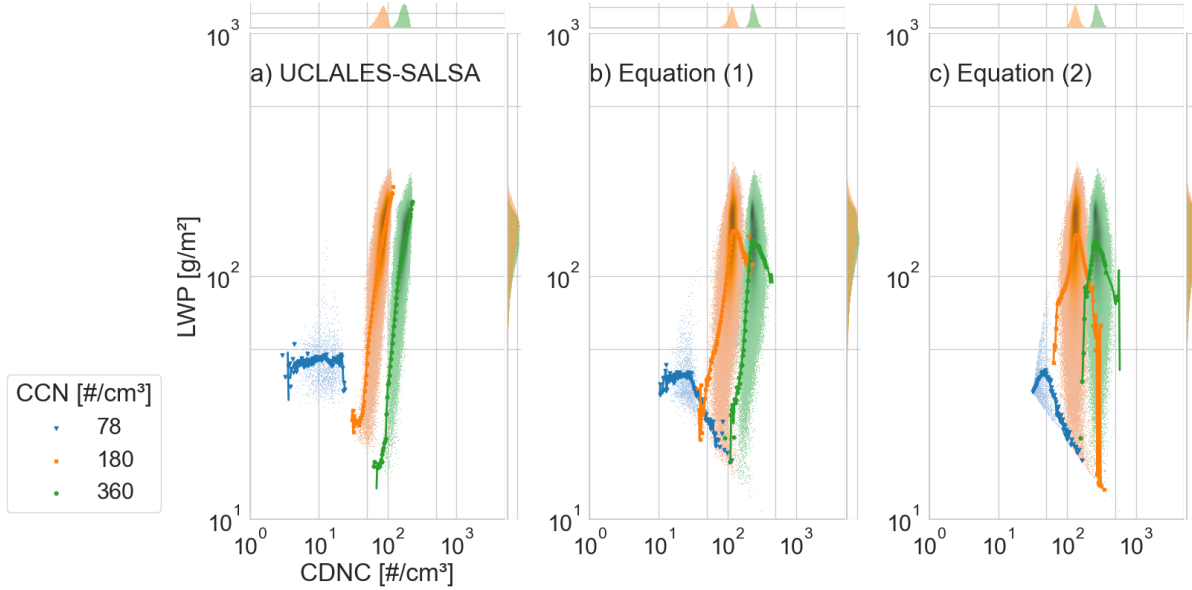
Figure S8. The same as Figure S6 for the time instance of 10 h.

S5 Spatial aggregation of cloud properties

To assess possible effects of cloud heterogeneity on the LWP susceptibility to changes in CDNC, we calculated average cloud properties in subdomains of 1.425 km by 1.425 km (19 model vertical columns each 75 m by 75 m) whose size mimics the spatial resolution of satellite observations (e.g. cloud properties in the MODIS L2 product have either 1 km or 5 km pixel resolution). Since radiances are directly proportional to cloud optical thickness, we use COT values in cloudy columns as a weighting factor to perform horizontal averaging operations along subdomains. For a cloud property X, its average value in the subdomain is calculated as

$$\bar{X}_{\mathrm{model}} = \frac{\sum_{i=1}^{n}(\mathrm{COT}_i X_i)}{\sum_{i=1}^{n}\mathrm{COT}_i}. \tag{S.8}$$

Spatially aggregated values for relevant cloud properties (e.g., CER, CDNC, k, f_{ad}, c_w) were used to emulate satellite retrievals of LWP and CDNC with equations 3, 2 and 1, respectively.

Subdomains with a cloud fraction less than 0.3 were excluded from calculations of the LWP susceptibility to changes in CDNC (Grosvenor et al., 2018). The cloud fraction was calculated with Equation (S.9) where i is the index of subdomain. Cloudy columns were identified using the same criteria applied to high resolution model outputs: columns whose COT is
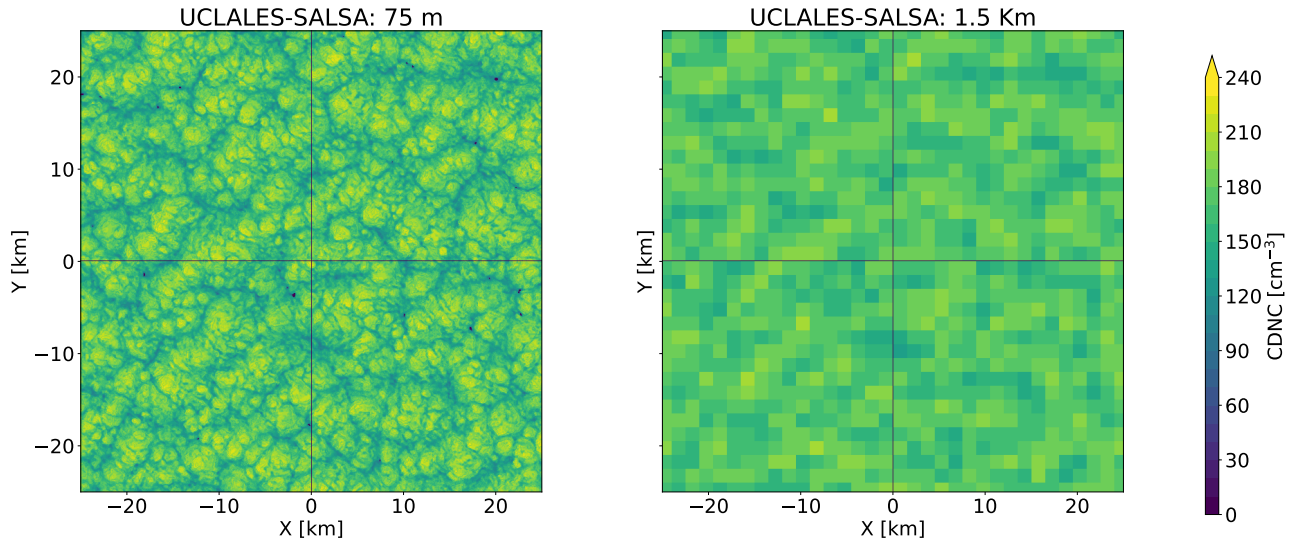
Figure S9. Variation in modelled CDNC fields after spatial aggregation to a coarser resolution. CDNC values correspond to the simulation initialized with a CCN loading of $360\,\mathrm{cm}^{-3}$ at the time instance of $10\,\mathrm{h}$.
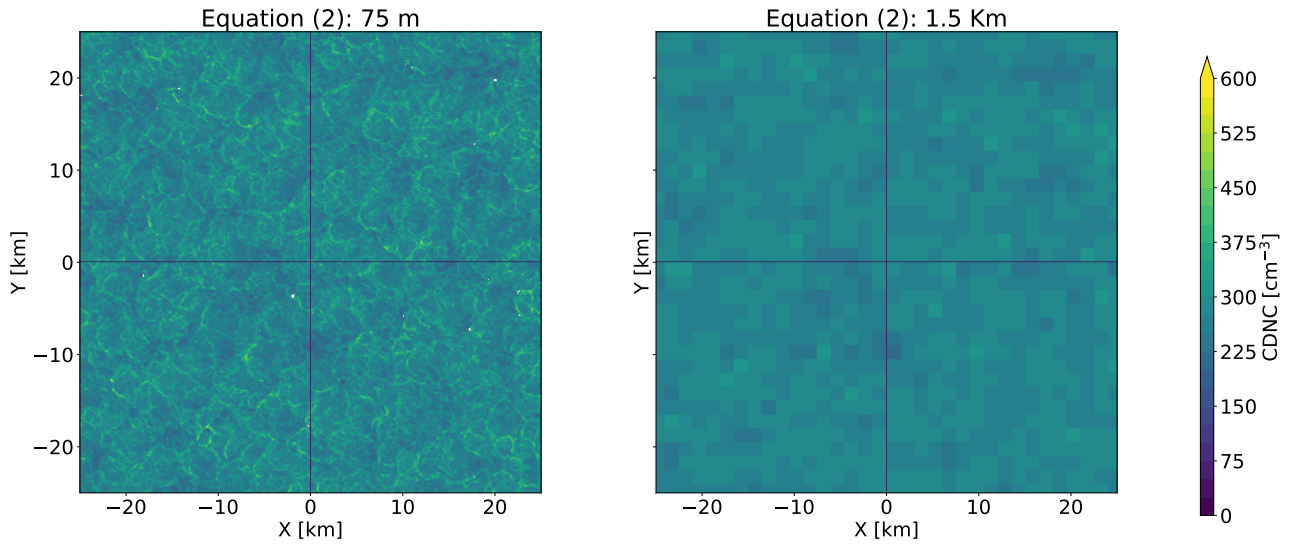
Figure S10. Variation in satellite retrievals of CDNC fields calculated with Equation (2) after spatial aggregation to a coarser resolution. CDNC values correspond to a simulation initialized with a CCN loading of $360\,\mathrm{cm}^{-3}$ at the time instant of $10\,\mathrm{h}$.

above 4.0 and CER larger than $4\,\mathrm{\mu m}$ are considered as cloudy (Gryspeerdt et al., 2019; Arola et al., 2022), CER values larger than $15\,\mathrm{\mu m}$ were used to identify possible precipitation effects on the LWP susceptibility (e.g., Arola et al., 2022).

$$cf = \frac{\sum_{i=1}^{n} \left(\mathrm{COT}_i > 4.0, \mathrm{CER}_i > 4\,\mu\mathrm{m} \right)}{n}. \tag{S.9}$$

The effects of spatial aggregation on CDNC and LWP fields can be observed in Figures S9 and S10. The usage of COT weighting in the aggregation flattens the signal from the cloud edges because optically thicker columns dominate the average sub-domain property. Figure S11 depicts how this results in a reduction of the skewness and kurtosis of the probability density function of CDNC biases.
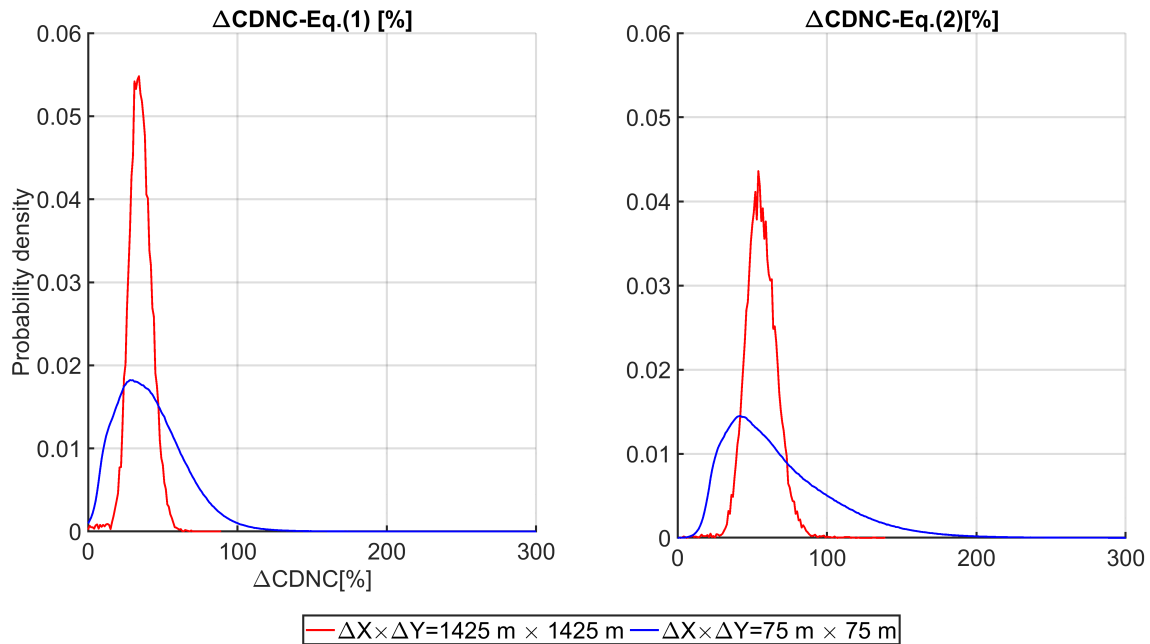
Figure S11. Probability density function of biases between modelled and satellite-surrogate CDNC determined in fine resolution (75 m) and coarse resolution (1.5 km simulation scenarios. Each continuous line depicts the distribution shape of CDNC biases for all simulation scenarios and time instances. The left and right panels describe the CDNC biases with respect to Equations (1) and (2), respectively.

S6   Spatial aggregation effects on the LWP susceptibility to changes in CDNC

LWP and CDNC values in the dataset with coarser resolution (i.e. spatially aggregated data) correlate linearly and joint histograms for both modelled and satellite retrievals, do not show of the "inverted v" shape as it is shown in Figures S12, S13 and S14. This factor together with the fact that the signal from cloud edges is flatten out after spatial aggregation of cloud properties (Figure S9), support the hypothesis that the inverted-V shape in satellite-based studies is caused by extreme positive biases in satellite retrievals of CDNC at very small CER values. This happens at cloud edges where the pseudo-adiabatic cloud model

does not fully represent cloud dynamics due to cloud top and lateral mixing. Satellite-equations suggest a shift onto negative LWP adjustments regardless of the aerosol loading or time instance in the simulation. Negative adjustments are expected in clouds affected by evaporation–entrainment or sedimentation-entrainment (Zhang et al., 2022). However, our model was not

145   able to reproduce them, except in a precipitating-cloud scenario (i.e. low aerosol loading after long simulation time) when it is expected to have a reduction of the LWP. The negative LWP adjustment also appeared after temporal aggregation of simulation with different aerosol loadings reflecting different stages in the cloud lifetime.
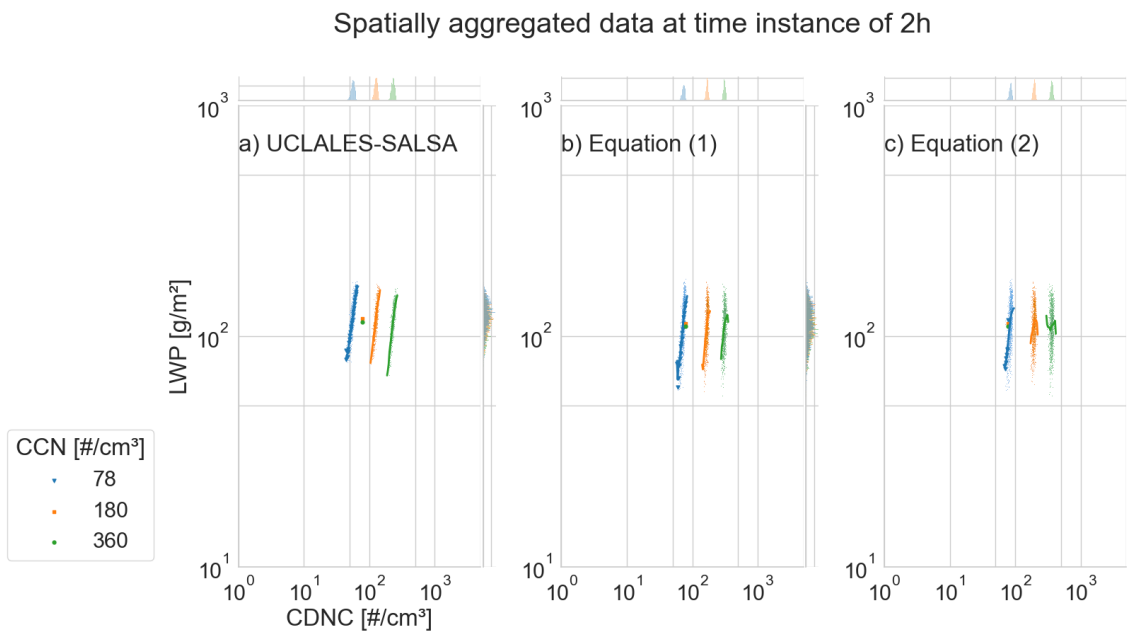
Figure S12. Joint and marginal histograms for spatially aggregated LWP and CDNC values using a) UCLALES-SALSA b) Equation (1) and c) Equation (2) at a time instance of 2 hours. Simulations are colour coded according to CCN concentrations in the aerosol loading used for model initialization. The intensity of colour in joint histograms increases when the probability increases. Continuous lines indicate the arithmetic mean.
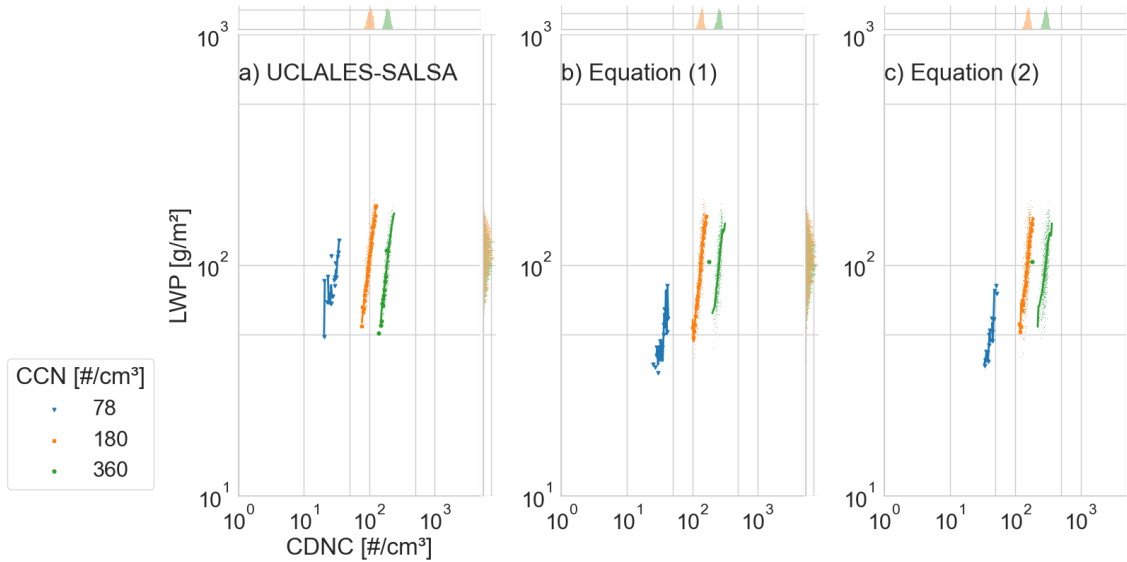
Figure S13. The same as Figure S12 for time instance of 6 h
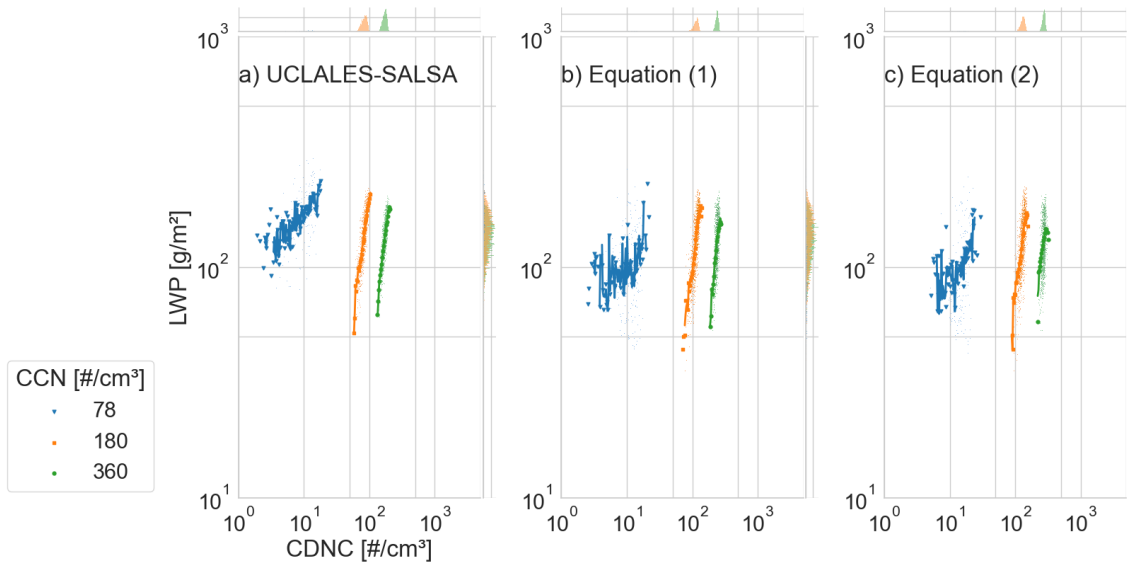
Figure S14. Joint and marginal histograms for spatially aggregated LWP and CDNC values using a) UCLALES-SALSA b) Equation (1) and c) Equation (2) at a time instance of 10 hours. For this particular time instance, we included model columns with CER > 15 μm to explore precipitation effects. Simulations are colour coded according to CCN concentrations in the aerosol loading used for model initialization. The intensity of colour in joint histograms increases when the probability increases. Continuous lines indicate the arithmetic mean.

References

Ackerman, A. S., vanZanten, M. C., Stevens, B., Savic-Jovcic, V., Bretherton, C. S., Chlond, A., Golaz, J.-C., Jiang, H., Khairoutdinov, M., Krueger, S. K., Lewellen, D. C., Lock, A., Moeng, C.-H., Nakamura, K., Petters, M. D., Snider, J. R., Weinbrecht, S., and Zulauf, M.: Large-Eddy Simulations of a Drizzling, Stratocumulus-Topped Marine Boundary Layer, Monthly Weather Review, 137, 1083–1110, https://doi.org/10.1175/2008MWR2582.1, 2009.

Arola, A., Lipponen, A., Kolmonen, P., Virtanen, T. H., Bellouin, N., Grosvenor, D. P., Gryspeerdt, E., Quaas, J., and Kokkola, H.: Aerosol effects on clouds are concealed by natural cloud heterogeneity and satellite retrieval errors, Nature Communications, 13, 7357, 2022.

Brenguier, J.-L., Pawlowska, H., Schüller, L., Preusker, R., Fischer, J., and Fouquart, Y.: Radiative Properties of Boundary Layer Clouds: Droplet Effective Radius versus Number Concentration, Journal of the Atmospheric Sciences, 57, 803–821, https://doi.org/10.1175/1520-0469(2000)057<0803:RPOBLC>2.0.CO;2, 2000.

Grosvenor, D. P., Sourdeval, O., and Wood, R.: Parameterizing cloud top effective radii from satellite retrieved values, accounting for vertical photon transport: quantification and correction of the resulting bias in droplet concentration and liquid water path retrievals, Atmospheric Measurement Techniques, 11, 4273–4289, https://doi.org/10.5194/amt-11-4273-2018, 2018.

Gryspeerdt, E., Goren, T., Sourdeval, O., Quaas, J., Mülmenstädt, J., Dipu, S., Unglaub, C., Gettelman, A., and Christensen, M.: Constraining the aerosol influence on cloud liquid water path, Atmospheric Chemistry and Physics, 19, 5331–5347, https://doi.org/10.5194/acp-19-5331-2019, 2019.

Savijärvi, H., Arola, A., and Räisänen, P.: Short-wave optical properties of precipitating water clouds, Quarterly Journal of the Royal Meteorological Society, 123, 883–899, https://doi.org/https://doi.org/10.1002/qj.49712354005, 1997.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics : from air pollution to climate change, vol. 40, John Wiley and Sons, Inc., 3 edn., https://doi.org/10.1080/00139157.1999.10544295, 2016.

Stull, R.: Practical Meteorology: An Algebra-based survey of Atmospheric Science, The University of British Columbia, 1.02b edn., 2017.

Wiscombe, W. J.: Improved Mie scattering algorithms, Appl. Opt., 19, 1505–1509, https://doi.org/10.1364/AO.19.001505, 1980.

Wood, R.: Relationships between optical depth, liquid water path, droplet concentration and effective radius in an adiabatic layer cloud, https://atmos.uw.edu/~robwood/papers/chilean_plume/optical_depth_relations.pdf, 2006.

Zhang, J., Zhou, X., Goren, T., and Feingold, G.: Albedo susceptibility of northeastern Pacific stratocumulus: the role of covarying meteorological conditions, Atmospheric Chemistry and Physics, 22, 861–880, https://doi.org/10.5194/acp-22-861-2022, 2022.