Investigating the sign of stratocumulus adjustments to aerosols in the global storm-resolving model ICON

Emilie Fons[1], Ann Kristin Naumann[2,3], David Neubauer[1], Theresa Lang[2,3], and Ulrike Lohmann[1]

[1]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
[2]Max Planck Institute for Meteorology, Hamburg, Germany
[3]Meteorological Institute, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany

Correspondence: Emilie Fons (emilie.fons@env.ethz.ch)

Abstract.

Since pre-industrial times, aerosol emissions have caused brightening of stratocumulus clouds, thereby cooling the climate. However, observational studies and climate models disagree on the magnitude of this cooling, in particular because of the liquid water path (LWP) response of stratocumulus clouds to increasing aerosols, with climate models predicting an increase in LWP, and satellites observing a weak decrease. With higher-resolution global climate models, there is hope to bridge this gap. In this

5 study, we present simulations conducted with the ICOsahedral Non-hydrostatic climate model (ICON) used as a global storm-resolving model (GSRM) with 5 km horizontal resolution. We compare the model outputs with geostationary satellite data, and we observe that, while ICON produces realistic low-cloud cover in the stratocumulus regions, these clouds look cumuliform and the sign of LWP adjustments to aerosols disagrees with satellite data. We evaluate this disagreement with a causal approach, which combines time series with knowledge of cloud processes in the form of a causal graph, allowing us to diagnose the

10 sources of discrepancies between satellite and model studies. We find that the positive LWP adjustment to increasing aerosols in ICON results from a superposition of processes, with an overestimated positive response due to precipitation suppression and cloud deepening under a weak inversion, despite small negative influences from cloud-top entrainment enhancement. Such analyses constitute a methodology that can guide modelers on how to modify model parameterizations and set-ups to reconcile

15 conflicting studies concerning the sign and magnitude of LWP adjustments across different data sources.

1 Introduction

Stratocumulus clouds cover 20 % of the Earth in the annual mean (Wood, 2012). Stratocumulus cloud cover is particularly large over the eastern part of the subtropical oceans, where they form the so-called "semi-permanent stratocumulus decks". Because these clouds are liquid low-level clouds located over dark oceanic surfaces, they efficiently reflect incoming shortwave

20 radiation but have small effects on outgoing longwave radiation. They are therefore crucial to cool the Earth. However, the radiative properties of stratocumulus clouds will be affected by climate change, not only due to increasing temperatures in the context of global warming, but also due to anthropogenic aerosol emission reductions in the context of air pollution mitigation (Forster et al., 2021). The IPCC predicts that subtropical marine low-level cloud cover and reflectivity will decrease in response

to future warmer temperatures, leading to a positive climate feedback of $0.2 \pm 0.16 \ \mathrm{W\,m^{-2}\,K^{-1}}$, i.e. an additional warming

25 effect. The clouds' response to globally changing aerosol emissions is more uncertain. Aerosols can modify liquid clouds by acting as cloud condensation nuclei, thereby increasing the cloud droplet number concentration N_d and reducing their effective radii r_eff, at an initially constant liquid water path (LWP). The increased total droplet surface area ($\propto N_\mathrm{d} r_\mathrm{eff}^2$) makes clouds instantaneously brighter (Twomey, 1977). The shift in cloud droplet size distribution can subsequently trigger precipitation suppression, leading to prolonged cloud lifetime and increased LWP and cloud cover (C) (Albrecht, 1989), and/or cloud-top

30 entrainment enhancement (Ackerman et al., 2004; Wang et al., 2003; Bretherton et al., 2007), leading to evaporation of cloud droplets and reduced LWP and C.

Observations (e.g. satellite data) and global climate models (GCMs) both agree that, since pre-industrial times, aerosol emissions have generated an overall cooling effect, with a total aerosol-cloud radiative forcing of -0.84 [-1.45, -0.25] $\mathrm{W\,m^{-2}}$ (Forster et al., 2021). Efforts to reduce future anthropogenic aerosol emissions will therefore generate an additional warming

35 due to reduced cloud brightness. However, there are concerns that the models might be right for the wrong reasons, as observations and models typically disagree on the sign and/or magnitude of LWP adjustments to aerosol perturbations (Quaas et al., 2009; Wang et al., 2012; Michibata et al., 2016; Neubauer et al., 2017; Malavelle et al., 2017; McCoy et al., 2020). Especially in stratocumulus regions, GCMs tend to predict LWP increases due to positive aerosol perturbations (e.g., Toll et al., 2017), implying a dominance of precipitation suppression, while satellite data analyses show a weaker, often slightly negative LWP

40 response, implying a dominance of cloud-top entrainment enhancement (e.g., Gryspeerdt et al., 2019; Possner et al., 2020).

Because of their coarse horizontal grid resolution (≈ 100 km), GCMs are not able to resolve the dynamics of stratocumulus clouds. The sub-grid processes that mediate aerosol effects on stratocumulus clouds have to be parameterized, leading to large intermodel variability and partly explaining why model estimates of aerosol-cloud interactions continue to disagree with observations. Thanks to progress in computing techniques, a new generation of kilometric-scale climate models, called global storm

45 resolving models (GSRMs), has emerged (Satoh et al., 2019). Inter-comparison experiments, like the DYAMOND (DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains) initiative by Stevens et al. (2019), have already been conducted to evaluate the representation of cloud processes in GSRMs. GSRMs still need to rely on parameterizations of sub-grid scale turbulence, cloud microphysics and radiation, but most of them can do without convective or cloud cover parameterizations (Hohenegger et al., 2023), making them particularly suited to study deep convection in tropical regions (e.g.,

50 Judt et al., 2021; Nugent et al., 2022; Lang et al., 2023). Although convection in stratocumulus regions is too shallow to be fully captured at kilometric resolutions, GSRMs start to bridge the gap between the physical process scale and the modeling scale for these regions as well (Stevens et al., 2020). Given the increasing popularity of GSRMs and their future use for years-long warming scenarios (e.g. as part of the "Next Generation Earth System Modelling Systems" or nextGEMS project[1]), it is crucial to study how low-level clouds are represented in these models. Heim et al. (2021) showed that the DYAMOND models were

55 able to produce realistic low cloud cover and characteristics of the marine boundary layer (e.g., inversion, subsidence) in the South-East Atlantic stratocumulus cloud deck. Nevertheless, the ensemble displayed large inter-model variability and systematic biases with respect to observations, in particular due to the turbulent mixing schemes. Furthermore, due to computational

---

[1]https://nextgems-h2020.eu/

cost considerations, the DYAMOND models all used one-moment cloud microphysics schemes, precluding an investigation of aerosol effects on low clouds as N_{d} is not parameterized.

60   In this study, we present the outputs of the global GSRM ICON (Zängl et al., 2015), run at 5km-resolution with Sapphire physics (Hohenegger et al., 2023), including a two-moment cloud microphysics scheme (Seifert and Beheng, 2006) (c.f. methods in Sec. A1-A2). We look at the stratocumulus microphysical properties and their interplay with the boundary layer and the cloud macrophysics over a 45-day period (27th June, 2021 to 10th August, 2021) in the four main stratocumulus regions (Fig. 1a-d): the North-East Pacific (NEP), the South-East Atlantic (SEA), the South-East Pacific (SEP) and the North-East Atlantic

65 (NEA). We run two simulations, one with 'low CCN' (fixed CCN concentration of 250 cm^{-3} in the lower troposphere), and one with 'high CCN' (1700 cm^{-3}). We also evaluate the model against geostationary satellite observations and reanalysis data (Benas et al., 2023; Walther and Straka, 2020; Huffman et al., 2023; Hersbach et al., 2018a, b) (c.f. methods in Sec. A3) in order to assess the realism of the stratocumulus cloud properties in the ICON GSRM (described in Sec. 2.1).

    It is challenging to do a direct comparison of aerosol-cloud interactions in the satellite data and in the model data as the

70 simulations are not nudged to observations (Sec. 2.2). One difficulty is the presence of potential confounders, i.e. environmental variables that change aerosol, cloud microphysics and boundary layer properties at the same time and therefore introduce spurious correlations in the statistical relationships. In the presence of confounding variables, it becomes difficult to disentangle superimposed processes that are expected to occur in aerosol-perturbed clouds, like precipitation suppression or entrainment enhancement. Several studies have used causal methods to remove spurious biases from aerosol-cloud interactions, either by

75 directly targeting environmental confounders (Gryspeerdt et al., 2016; Varble, 2018) or by using opportunistic experiments, like ship tracks or volcanic eruptions, to indiscriminately remove all environmental confounding (Manshausen et al., 2022; Chen et al., 2022). In Sec. 3, we use the causal methodology described in Fons et al. (2023) to directly target confounding induced by the precipitation and entrainment responses of stratocumulus clouds to aerosols. This methodology is based on a causal graph of LWP drivers, which was developed using prior domain knowledge of stratocumulus cloud processes ('expert'

80 causal graph), without using any causal discovery algorithms (Runge et al., 2023). This causal graph can be understood as a mass balance representation of the sources (condensation) and sinks (entrainment, precipitation) of LWP and can be used to detect the responses of these sources and sinks when stratocumulus clouds are perturbed by aerosols. We combine the physical knowledge contained in this graph with time series data, namely satellite data and GSRM model data from ICON. The high temporal resolution of both datasets ($\Delta t = 15$ min) allows us to use the precedence of cause with respect to effect and get

85 comparative causal insights into aerosol-cloud interactions. In that sense, the causal graph is used as a model evaluation tool (Nowack et al., 2020) to diagnose deviations of the model from observations.

2 Stratocumulus clouds in ICON-GSRM

2.1 How realistic are ICON's stratocumulus clouds?
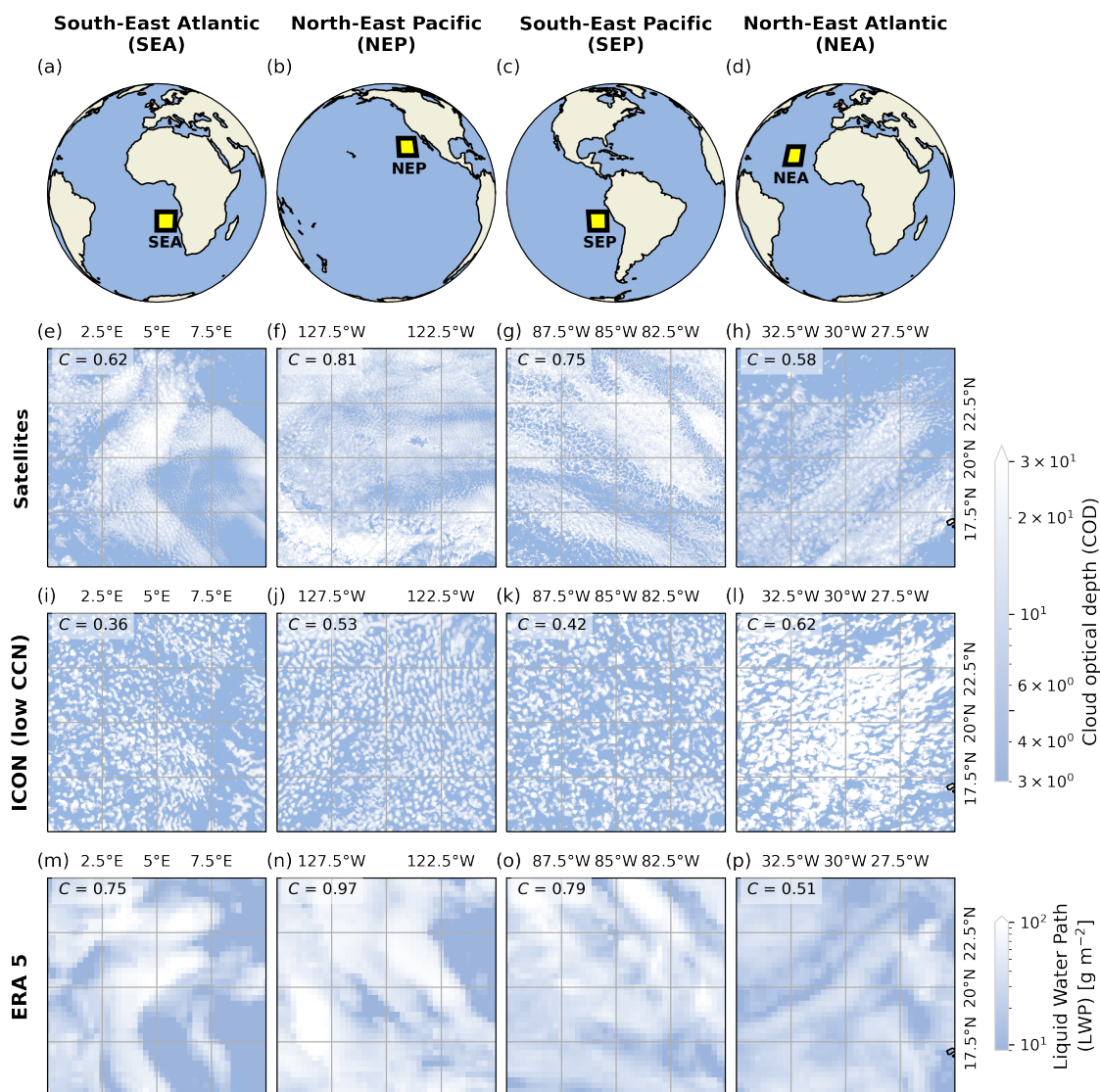
Figure 1. Low-level cloud cover C in stratocumulus regions. **(a-d)** show the locations of the four regions, e-h show bird eye's view snapshots (06/07/2021 15:00 local time) of stratocumulus clouds as seen by geostationary satellites, **(i-l)** and **(m-p)** show snapshots at the same timestep as simulated by ICON and ERA5, respectively. Satellite and ICON cloud covers are rendered using the cloud optical depth (COD) variable, while LWP is used for ERA5 data (COD cannot be calculated with ERA5's outputs). Lower thresholds of COD=4 and LWP=9 $\mathrm{g\,m^{-2}}$ are used for the sake of this figure. Note that the ICON GSRM is not nudged to the observations, so we do not expect to see the exact same cloud formations at the same timestep. Maps from cartopy (Met Office, 2010 - 2015).

Fig. 1 shows snapshots of the low-level cloud cover C in the four stratocumulus regions. Even though these images are snapshots, they are quite representative of the typical cloud organization over the time period of this study. ICON produces realistic C in the four stratocumulus regions (Fig. 1i-l). The mean C in the different regions ranges from 55 to 90 % in the satellite data and 55 to 80 % in the ICON data (Fig. 2a). However, even when the simulated C is similar to the observed C (e.g. Fig. 1h, l), the mesoscale cloud organisation looks very different in the model and in the observations. The satellite snapshots show how stratocumulus clouds are highly aggregated in closely connected cells, whether they are closed cells (Fig. 1e, f), open cells (Fig. 1g) or a mix of closed and disorganized cells (Fig. 1h) (Wood and Hartmann, 2006). ICON simulates more individual, less connected cells, i.e. clouds that seem to have cumuliform features. Increasing the CCN concentrations did not significantly change the overall C (Fig. 2a) or cloud organization in the 4 stratocumulus regions (not shown). Fig. 1m-p show the ERA5 C on the same day and same time step as Fig. 1e-h. Even though similar cloud structures appear in the observations and the reanalysis data, the coarser ERA5 C does not agree with the observed C on a pixel-by-pixel basis. This has implications for the co-location of cloud variables from the satellite and reanalysis datasets (see methods in Sec. A3).

Fig. 2 b-j shows mean values of cloud and boundary layer properties for the ICON simulations and the geostationary satellite data for 10 different variables: C, N_{d}, r_{eff}, cloud depth (H), cloud-top entrainment rate (w_{e}), surface precipitation (RR), cloud-base vertical velocity (w_{CB}), LWP, boundary layer height (BLH) and estimated inversion strength (EIS) (see methods for calculations). While C, r_{eff}, H, w_{e}, LWP and BLH are of the same order of magnitude between the model and the observations, there is a significant underestimation of N_{d} and a significant overestimation of RR by ICON compared to the satellite data. Due to the horizontal grid discretization, both ICON and ERA5 underestimate the variability of in-situ vertical velocities (Fig. 2g), with maximal w_{CB} much smaller than real updraft speeds at stratocumulus cloud base (≈ 0.5 m s^{-1}; Wood, 2012). The underestimation and low variability of N_{d} in ICON is directly explainable by the low vertical velocities, as the CCN activation is parameterized as a function of the (fixed) CCN concentration and the vertical velocities at cloud base (Segal and Khain, 2006). The underestimated N_{d} leads to slightly overestimated r_{eff} and strongly overestimated precipitation rates, both at cloud-base and at the surface. In the high CCN experiment, N_{d} and r_{eff} are shifted to more realistic values compared to observations, although it should be noted that the high CCN concentration amounts to 1700 cm^{-3}, which is unrealistically high for marine regions. Even though cloud-base precipitation decreases by a factor of ≈ 2 in the high CCN experiment, precipitation still remains 1 to 2 orders of magnitude higher than the satellite-measured precipitation, indicating a systematic high bias for stratocumulus precipitation in this ICON model configuration with respect to satellite-derived precipitation estimates.

Additionally, there is a slight underestimation of EIS in ICON compared to ERA5. It can be noted that ERA5 itself might underestimate EIS in stratocumulus regions: Although model developments led to an increasingly better representation of boundary-layer stratocumulus clouds in ERA reanalysis (Köhler et al., 2011), biases remain (Ahlgrimm et al., 2018), and the sharpness of the inversion at the top of the boundary layer is often underestimated (Kalmus et al., 2015; Zheng and Miller, 2022).

Fig. 3 shows the vertical structure of the clouds in the low CCN simulation, at the same timestep as in Fig. 1, while Fig. 4 shows average vertical profiles for selected variables. ICON's clouds are well confined in the lower part of the troposphere, at the top of the boundary layer (Fig. 3 and Fig. 4e). The top of the boundary layer is marked by a temperature inversion of a few
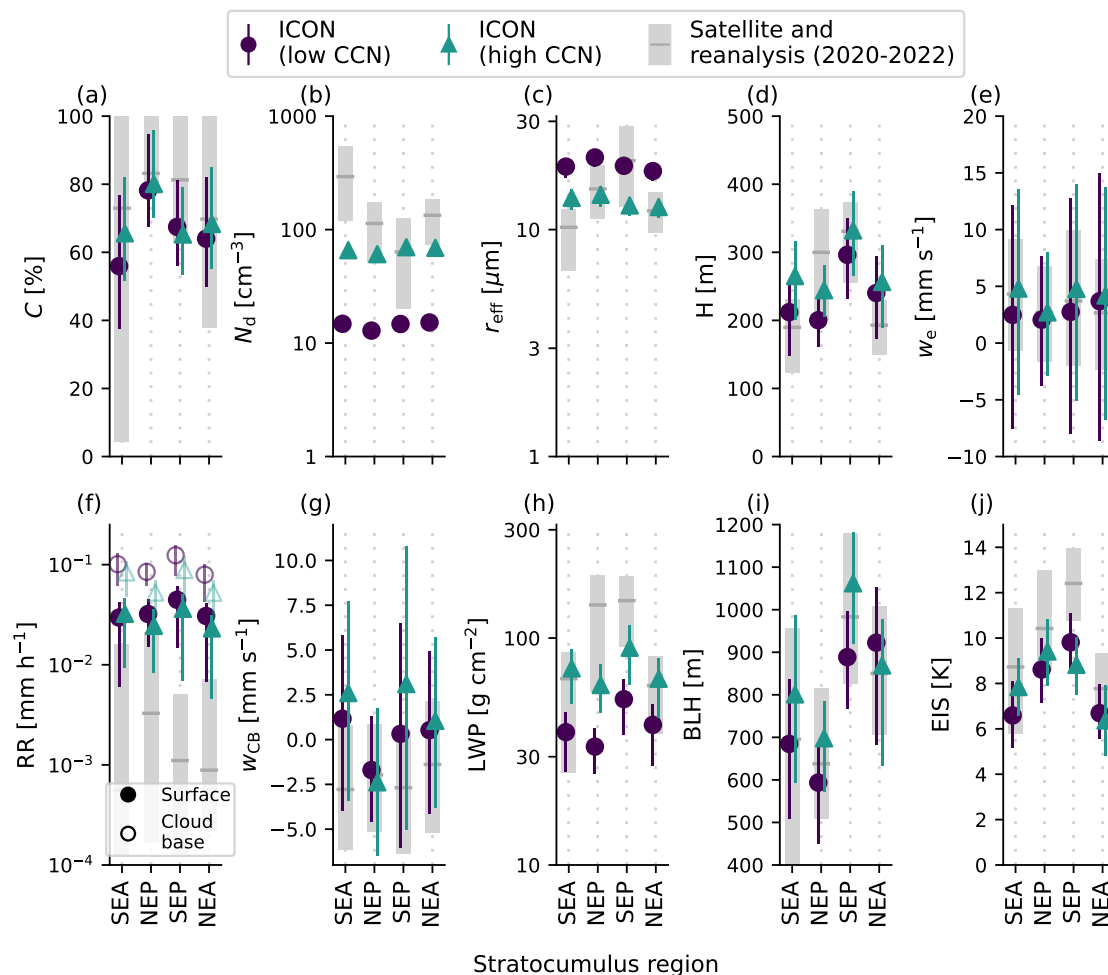
Figure 2. Cloud and boundary layer properties for the four stratocumulus regions. The markers show the average value for ten different variables: cloud cover **(a)**, cloud droplet number concentrations **(b)**, cloud droplet effective radii **(c)**, cloud depth **(d)**, cloud-top entrainment rate **(e)**, rain rate **(f)**, cloud-base updraft **(g)**, liquid water path **(h)**, boundary layer height **(i)**, estimated inversion strength **(j)**. The vertical bars (purple, green and thick grey) show the spread between the 25th and the 75th percentile of the instantaneous grid points (for the low ICON CCN, high ICON CCN and satellite/reanalysis data, respectively). The grey bar for satellite data covers three years (from 01/07 to 09/07 for 2020, 2021, 2022 vs. 2021 only for ICON) to allow for a comparison of the model values to a climatological spread. For the grey variables, w_e, BLH and EIS are derived from ERA5, while the other variables are derived from satellite products (see methods in Sec. A3).

Kelvin (Fig. 4a), which is slightly weaker than the inversion in the reanalysis data. The free troposphere is characterized by large-scale subsidence and low relative humidity (Fig. 4b,c). The ICON clouds seem to have cumuliform features: Even though, on average, the cloud height is similar to the observed height (Fig. 2d and Fig. 4e), the depth of individual cloud cells is highly variable, with small and deep cells that seem rather disaggregated. Fig. 3a-d and Fig. 4e show that the liquid water content
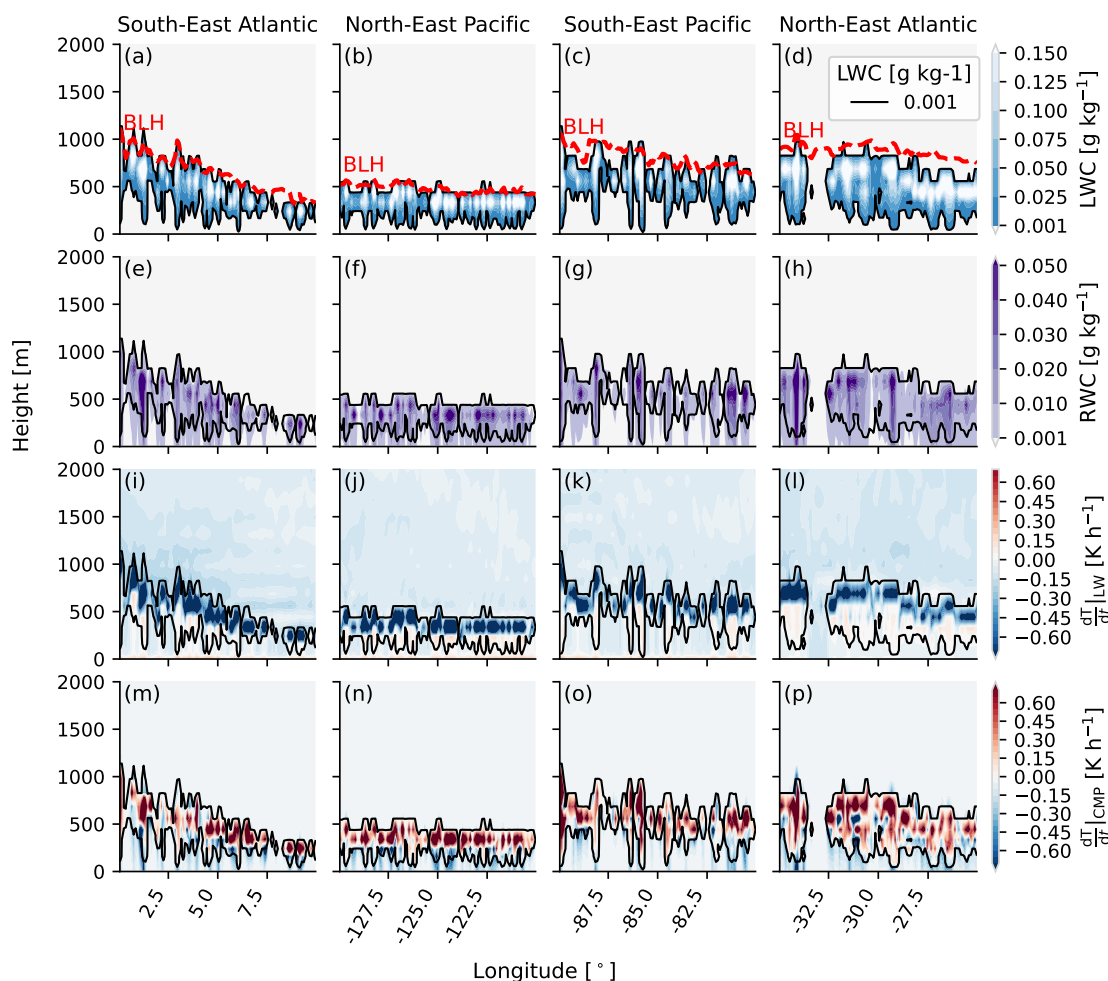
Figure 3. Cross section snapshots of cloud properties simulated by ICON. The cross-sections are taken along the middle latitude of each region, and at the same timestep as Fig. 1. The first row **(a-d)** shows the cloud liquid water content (LWC) in all four regions, the second row **(e-h)** shows the rain water content (RWC), the third row **(i-l)** shows the temperature tendency due to longwave (LW) radiation and the fourth row **(m-p)** shows the temperature tendency due to phase changes caused by the cloud microphysics (CMP) scheme. On the first row, the red dotted line indicates the boundary layer height (BLH).

of the clouds increases with height in the cloud, in line with cloud adiabaticity assumptions (Lohmann et al., 2016), even though it decreases again close to cloud top, suggesting cloud-top entrainment. Fig. 3e-h and 4d show the rain water content of the clouds. In the 4 regions, most cloud cells precipitate, with strong precipitation coming from the deepest cloud cells. Even though some precipitation evaporates (blue color in Fig. 3m-p), there is surface precipitation across most of the domain.
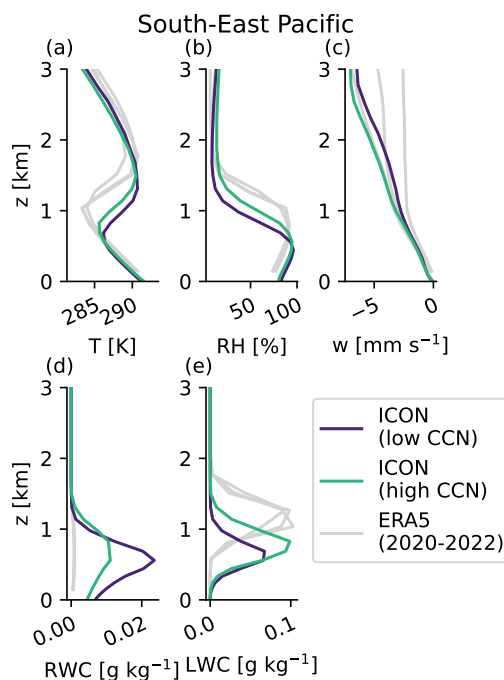
Figure 4. Vertical profiles of cloud and boundary layer properties. Temperature T **(a)**, relative humidity RH **(b)**, vertical velocity w **(c)**, RWC **(d)** and cloud liquid water content LWC **(e)** are shown for the SEP region. Results are similar for the three other regions, and available in Suppl. Fig. 1

This confirms the high precipitation bias from stratocumulus regions in this ICON configuration, both compared to satellite estimates (Fig. 2f) and to reanalysis data (Fig. 4d). As expected, the clouds produce longwave (LW) cooling at cloud-top (negative temperature tendencies in Fig. 3i-l) due to efficient emission of LW radiation to space through the dry and cloud-free

135 free troposphere. In stratocumulus clouds, this cloud-top cooling is an important driver of convective and turbulent mixing (Wood, 2012). Fig. 3m-p shows the temperature tendency due to phase changes caused by the cloud microphysics scheme, with a positive tendency indicating net condensation, and a negative tendency indicating net evaporation. Net condensation is visible inside clouds, while net evaporation of rain water occurs below cloud base. Net evaporation of cloud water is also visible in-cloud, especially when the cloud cells are deep (e.g. Fig. 3m), indicating that evaporation of cloud/rain water also

140 occurs inside the cloud due to entrainment.

Overall, the cloud characteristics shown in Fig. 1-4 indicate that ICON simulates rather realistic boundary layer properties and low-level cloud characteristics in the stratocumulus regions. In particular, specific features like extensive cloud cover at the top of the boundary layer and cloud-top radiative cooling are well represented by ICON. Additionally, ICON-GSRM seems to be able to generate reasonable values of cloud-top entrainment (Fig 2e), which is a key process for the water budget and

145 evolution of stratocumulus clouds. However, there are systematic biases with respect to observations, namely too low cloud

droplet number concentrations due to low vertical velocities at cloud base, and very high rain rates. These biases are bound to affect the radiative budget of the stratocumulus regions and should be carefully studied before using the GSRM in warming scenarios. In the next sections, we look at how these biases affect the LWP adjustments of stratocumulus clouds to aerosols.

2.2 LWP adjustment to aerosols

150 LWP adjustments to aerosols are typically quantified by the derivative of LWP with respect to an aerosol proxy A (e.g. $A =$ Aerosol Optical Depth): $\frac{d\,\text{LWP}}{d\,A} = \frac{\partial\text{LWP}}{\partial N_d}\frac{\partial N_d}{d\,A}$. Given the difficulty of obtaining a satellite-derived aerosol proxy A that correlates well with CCN concentrations (e.g., Stier, 2016), the two partial derivatives on the right-hand side of this equation are often evaluated separately. Here, we focus on the first term, $\frac{\partial\text{LWP}}{\partial N_d}$, and how it is impacted by aerosol-driven entrainment and precipitation responses.
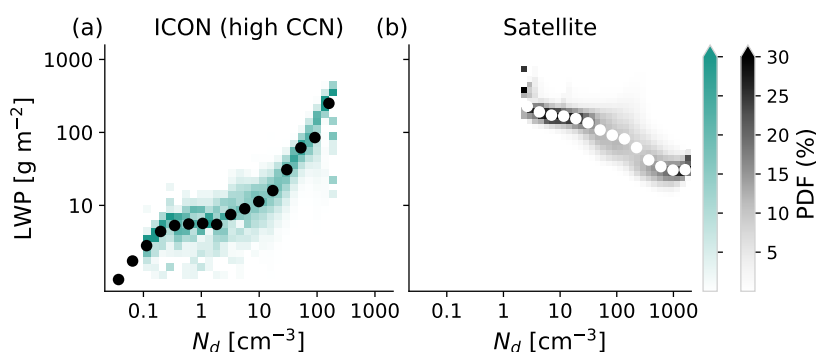
Figure 5. 2D histograms of LWP vs. N_d. The subplots show the relationship in the satellite data **(a)**, and the model data **(b)**. All data from the four regions are aggregated here. See Suppl. Fig. 2 for histograms disaggregated by stratocumulus region. The color shading shows the probability density function (PDF), after binning the data points in narrow intervals of N_d values.

155 Fig. 5 shows that LWP values are positively correlated with N_d values in ICON data but negatively correlated in satellite data. This suggests that, like in GCMs, and contrary to observations, ICON simulates LWP increases in response to aerosol increases. However, correlation does not mean causation. For more insight into actual aerosol effects on the LWP, we can also look at the comparison between the low CCN and the high CCN experiments conducted with ICON. Fig. 2h confirms that the LWP responds positively to increases in CCN and N_d in the model, increasing confidence in the fact that this is indeed a causal

160 response. Fig. 4e shows that this increase in LWP can be divided into an increase in LWC and an increase in cloud depth. At this point, it is important to note that Fig. 4 shows all-sky vertical averages over space and time, so that differences between the low and high CCN cases could be (partly) due to modified cloud fractions/lifetimes in the high CCN case. However, Jiang et al. (2006) showed no cloud lifetime effects for cumulus clouds. Histograms of instantaneous grid point properties are also provided in Suppl. Fig. 3 and confirm that the differences between the low and high CCN simulations are still present when a

165 potential lifetime effect is taken out.

It has been postulated that the positive LWP adjustment in GCMs is due to the coarse resolution and lack of explicit parameterization of size-dependent entrainment processes in GCMs, while the rain autoconversion process is explicitly parameterized to be dependent on N_d, thereby driving a strong positive response of LWP to aerosol perturbations. It is clear from the results presented above (Fig. 2f, 3e-h, 4d) that clouds precipitate too readily in ICON, with rain rates one or two orders of magnitude

170   larger than the ones recorded by the satellite precipitation product or estimated by ERA5. In this highly precipitating regime, aerosol-induced shifts in the cloud droplet distribution can cause reductions in precipitation rates and result in strong absolute increases in LWP. The comparison of Fig. 4d and Fig. 6a,b shows how the precipitation profile shifts concomitantly with the droplet number and droplet size profiles when the CCN concentration is increased.
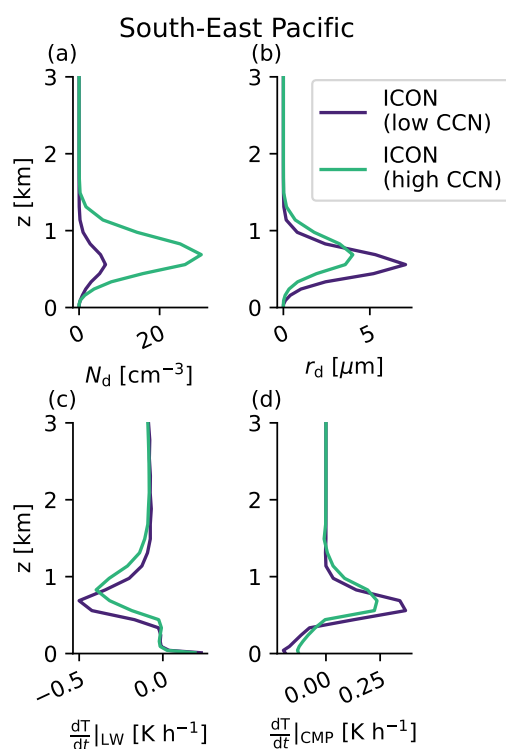
Figure 6. Vertical profiles of cloud microphysics and radiative properties. The subplots show: the cloud droplet number concentration **(a)**, the cloud droplet radius r_d **(b)**, the temperature tendency due to LW radiation **(c)** and the temperature tendency due to phase changes caused by the cloud microphysics scheme **(d)**. Only the SEP region is shown here, results are similar for the 3 other regions (Suppl. Fig. 1).

But what happens to cloud-top entrainment in ICON-GSRM? Contrary to GCMs, sub-kilometric Large-Eddy Models

175   (LEMs) have been able to simulate negative entrainment influences on the LWP (Ackerman et al., 2004; Wang et al., 2011). With a resolution that is much better than GCMs, but still coarser than LEMs, GSRMs could be expected to simulate entrainment processes via the turbulence scheme, thereby driving aerosol-induced LWP reductions. The net LWP increase that is observed in the model does not necessarily mean that there is no entrainment response. Fig. 2e does show that cloud-top

entrainment rates are slightly enhanced in the high CCN simulation. Fig. 6c,d give more insight into potential mechanisms
for entrainment enhancement in ICON with the vertical profiles of the temperature tendencies related to LW radiation and
CMP phase changes. Given that there is no cloud droplet sedimentation in this model set-up, the only possible mechanisms for
cloud-top entrainment enhancement are: increased cloud-top LW cooling and/or enhanced evaporative cooling, both resulting
in increased cloud-top turbulence (Bretherton et al., 2007). Fig. 6c shows that, despite the increased LWP in the high CCN
experiment, cloud-top LW cooling (negative $\frac{dT}{dt}|_{LW}$) is not enhanced: it is even reduced. That can be explained by the higher
cloud tops in the high CCN simulation (Fig.4e), resulting in colder LW emission temperatures. This eliminates the LW cooling
hypothesis. Fig. 6d shows how net cloud water condensation leads to latent heat release (positive $\frac{dT}{dt}|_{CMP}$) within the cloud.
In the high CCN case, in-cloud $\frac{dT}{dt}|_{CMP}$ is reduced compared to the low CCN case. This behavior can also be seen in the
instantaneous temperature tendencies in Suppl. Fig. 3. A possible explanation is larger latent heat consumption from enhanced
evaporation in the high CCN case. This means that the evaporation hypothesis is a plausible mechanism for entrainment en-
hancement in ICON. Interestingly, because of saturation adjustments in the model, there is no possible mechanism for droplet
size-dependent evaporative rates. Therefore, the enhanced evaporation is purely due to the increase in LWP caused by pre-
cipitation suppression, i.e. there is more water left in the cloud to evaporate. This can generate more in-cloud turbulence and
enhance the cloud-top entrainment. This means that the real mechanisms for entrainment enhancement and the mechanisms at
play in the model might be different, and this explains why we will use two different causal graphs (one for satellite data and
the other for model data) in the next section.

Because cloud-top entrainment is not explicitly parameterized in the ICON model, it is not as straightforward to evaluate
entrainment influences as it is for precipitation influences via the autoconversion parameterization. For example, it is possible
that, as expected for aerosol-perturbed stratocumulus clouds, there is entrainment enhancement and resulting cloud evaporation,
but that this process is weaker than precipitation suppression, resulting in net LWP increases when CCN are increased. It is
also possible that entrainment is enhanced but that it drives cloud deepening by raising the cloud top (Wood, 2007), resulting
in an enhanced positive LWP response. Another difficulty is that the ICON simulations are global, and CCN concentrations
were increased globally from the low to high CCN case. This means that the boundary conditions (subsidence, inversion) of the
four stratocumulus regions might change between the low and high CCN experiments. For instance, the entrainment response
to CCN increases might be driven by the coupling of stratocumulus cloud processes with the large-scale circulation (e.g., the
Walker circulation, Dagan et al., 2023) and not directly driven by local CCN increases.

Moreover, the mechanism for cloud deepening (Fig. 2d, 4e) is not immediately obvious. Cloud deepening could be driven
by large-scale meteorological changes, or it could be locally driven by entrainment under a weak inversion. Because the clouds
have cumuliform features in ICON, it is also possible that other processes, such as cloud deepening from delayed precipitation
(Seifert et al., 2015), or aerosol-induced warm cloud invigoration (Douglas and L'Ecuyer, 2021), are driving cloud deepening
and LWP increases on top of the precipitation response.

The above considerations show how difficult it is to draw causal conclusions from simple correlations or experiment com-
parisons. In the next section, we apply a causal graph for LWP drivers (Fons et al., 2023) to disentangle the LWP response to
increases in aerosols and cloud droplet number concentrations. This will allow us to answer the following questions concerning

the model data: (1) Are cloud responses driven by local aerosol perturbations or non-local processes? (2) Does an entrainment

215 response to aerosols exist along the precipitation response? (3) What is the sign of this entrainment response? (4) Do additional

processes, like cloud deepening (driven by entrainment, warm cloud invigoration or rain delays), also have an importance?

3 Causal diagnosis of LWP drivers

3.1 Precipitation and entrainment
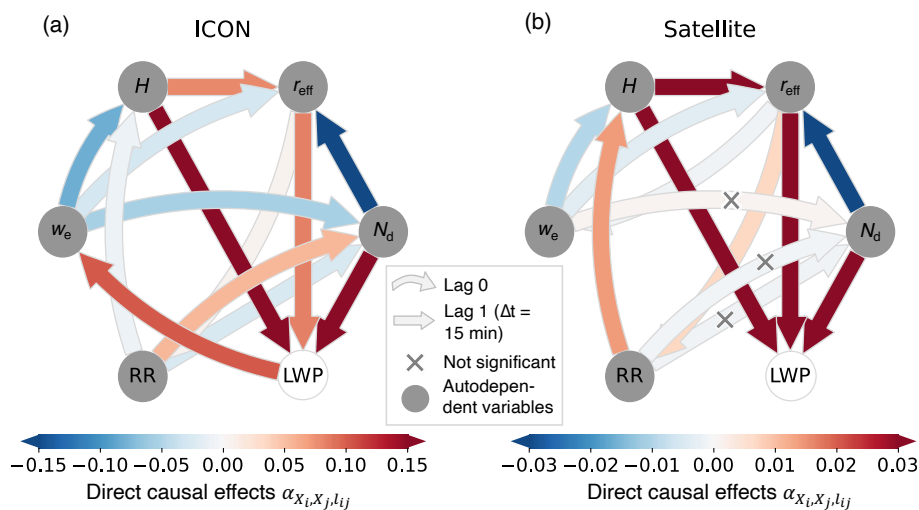
Figure 7. Direct causal effects for LWP drivers in stratocumulus clouds. **(a)** is the causal graph applied to the ICON data set, and **(b-c)** are the causal graph applied to the satellite data set, where H is calculated from the satellite data in **(b)** and from ERA5 in **(c)**. In **(b-c)**, one arrow has been moved compared to **(a)** due to the different mechanisms possible for entrainment enhancement in reality and in the model set-up. The straight arrows indicate contemporaneous (lag 0) effects, while the curved arrows indicate lagged (lag 1 = Δt = 15 min) effects. The color shades of the arrows indicate the sign and magnitude of the causal effects (note the different scales for a) and b-c)). The small grey crosses indicate the arrows that are not significantly different from 0 based on a 90 % bootstrap confidence interval.

Fig. 7 shows the direct causal effects $\alpha_{X_i, X_j, l_{ij}}$ calculated from the LWP 'expert' causal graph. The graphs for the model

220 data and for the satellite data are different (see arrow pointing to w_e) due to the different mechanisms that are possible for

entrainment enhancement in reality and in the model set-up. Both graphs describe how contemporaneous (lag 0, straight

arrows) and lagged (lag 1 = Δt = 15 min, curved arrows) responses in precipitation and entrainment processes can lead to LWP

adjustments following aerosol perturbations. Each individual arrow describes a given physical process, as indicated in Table 1.

The causal graphs presented here contain one additional arrow (from RR to H) compared to the graph presented in Fons et al.

225 (2023), due to the consideration of another physical process, cloud deepening via rain delays.

Table 1. Association of the direct causal effects $\alpha_{X_i,X_j,l_{ij}}$ with the corresponding cloud physical processes. Almost all arrows are correctly detected, i.e. their sign agrees well with the direction of the underlying physical process.

Arrow (Fig. 7)	Physical description	Expected sign	Correctly detected in satellite?	Correctly detected in model?
$\alpha_{N_d,\text{LWP},0}$, $\alpha_{r_\text{eff},\text{LWP},0}$ and $\alpha_{H,\text{LWP},0}$	Definition of LWP	+	Yes	Yes
$\alpha_{N_d,r_\text{eff},0}$	Due to mass balance for a given LWP	-	Yes	Yes
$\alpha_{H,r_\text{eff},0}$	Condensational growth with height	+	Yes	Yes
$\alpha_{r_\text{eff},w_e,1}$	Entrainment suppression by droplet sedimentation	-	Yes	N/A
$\alpha_{\text{LWP},w_e,1}$	Entrainment enhancement by evaporation of cloud droplets	+	N/A	Yes
$\alpha_{w_e,N_d,1}$, $\alpha_{w_e,r_\text{eff},1}$ and $\alpha_{w_e,H,1}$	Evaporation due to cloud-top entrainment	-	Yes (homogeneous)[a]	Yes (mix)[a]
$\alpha_{r_\text{eff},\text{RR},1}$	Rain enhancement	+	Yes	Yes
$\alpha_{\text{RR},N_d,0}$	Cloud water removal	-	Not significant	Yes
$\alpha_{\text{RR},N_d,1}$	Wet scavenging/Dynamical adjustments	0/-	Not significant	No
$\alpha_{\text{RR},H,1}$	Cloud deepening due to rain delay	-	Inconclusive[b]	Yes

[a]: 'Homogeneous' and 'mix' refer to the entrainment mixing regimes described by Hill et al. (2009). [b]: Results are very sensitive to the proxies used for H and RR. See text for more details.

Each $\alpha_{X_i,X_j,l_{ij}}$ in Fig. 7 corresponds to the regression coefficient of X_i on X_j along the l_{ij}-lagged arrow after any confounding indicated by the causal graph has been removed. See Fig. A1 in the methods section for the example calculation of the $\alpha_{N_d,r_\text{eff},0}$ direct causal effect. A negative (positive) direct causal effect means that an increase in X_i leads to a decrease (increase) in X_j after a lag l_{ij}. Here, the negative $\alpha_{N_d,r_\text{eff},0}$ describes how, when a cloud is perturbed by aerosols and N_d

230 increases, the droplet size r_eff decreases to satisfy the water mass balance at an initially constant LWP. The direct causal effects can be multiplied according to Wright's path rule (Wright, 1921) to compute total causal effects $\beta_{X_i,X_j,l}$ between two variables that are not directly linked by an arrow in the graph. For example, the effect of N_d on RR (mediated by r_eff) is computed with: $\beta_{N_d,\text{RR},1} = \alpha_{N_d,r_\text{eff},0} \times \alpha_{r_\text{eff},\text{RR},1} < 0$. The negative sign of $\beta_{N_d,\text{RR},1}$ agrees well with the precipitation suppression mechanism that we expect to see in aerosol-perturbed clouds (Albrecht, 1989).

235 Almost all direct causal effects $\alpha_{X_i,X_j,l_{ij}}$ have the same sign in the satellite and the model data, although magnitudes are different (Fig. 7, Table 1). We suspect that the magnitudes in the model graph might be higher because the model data set is more self-contained than the co-located satellite/reanalysis dataset, where the variables are derived from different sources. For this reason, we mostly comment on the sign of the physical processes. The signs of the satellite causal effects calculated here also agree well with the calculations from Fons et al. (2023), which were conducted on 2-year, coarser-resolution timeseries of

240 the SEA region only.

In Fons et al. (2023), the lag 0 arrow from RR to N_{d} was found to be negative, indicating cloud water removal by precipitation, while the lag 1 arrow was close to zero. We hypothesized that the lag 1 arrow was the superposition of a negative process (CCN scavenging by precipitation; Grandey et al., 2014) and a positive process (speculatively linked to dynamical adjustments of updraft speeds below cloud base when cold pools are present; Terai and Wood, 2013). In this study, both the lag-0 and lag-1

245   arrows are found to be insignificant, i.e. the confidence interval includes 0. This might indicate that the noise-to-signal ratio is not favorable with this satellite precipitation proxy. In the model data, the lag-0 arrow from RR to N_{d} is accurately detected to be negative (cloud water removal), but the lag-1 arrow is detected to be positive. This might be because there is no CCN scavenging in the model, as CCN concentrations are constant in space and time, and other dynamical adjustments might cause the lag-1 arrow to become positive. Interestingly, when we replace cloud-base precipitation with surface precipitation rates in

250   the ICON data, the lag-0 arrow becomes positive and the lag-1 arrow becomes negative (Suppl. Fig. 4a). This seems to indicate that the quantification of precipitation processes is very sensitive to the chosen precipitation product and to the choice of causal time lags. Because the satellite proxy quantifies surface precipitation, and surface precipitation from stratocumulus is hard to retrieve from space (e.g., Zhu et al., 2022), precipitation processes might be inaccurately estimated. Additionally, although 15 min is a good temporal resolution to capture stratocumulus cloud updrafts (see methods), it might be too coarse to resolve faster

255   precipitation influences, potentially resulting in spurious positive precipitation influences in the model data (Runge, 2018). It should also be noted that the causal method used here is based on a linearity assumption which might be broken for non-linear precipitation processes. Finally, hidden confounding variables (i.e. not yet included in the causal graph), like relative humidity (Grandey et al., 2014), might also further bias the arrows from RR to N_{d} and will need to be evaluated in future research. We focus on the other arrows of the causal graph in the rest of the discussion.

260   Aerosol-induced increases in N_{d} cause entrainment enhancement both in the satellite data and in the model. In the satellite graph, this is described mathematically by $\beta_{N_{\mathrm{d}},w_{\mathrm{e}},1} = \alpha_{N_{\mathrm{d}},r_{\mathrm{eff}},0} \times \alpha_{r_{\mathrm{eff}},w_{\mathrm{e}},1} > 0$. This describes how, following aerosol increases, reductions in r_{eff} will slow down the sedimentation of cloud droplets at the cloud top, where more efficient longwave and evaporative cooling of the smaller droplets will generate turbulence and enhance entrainment of free tropospheric air into the boundary layer. In the model data, entrainment enhancement is directly described by $\alpha_{\mathrm{LWP},w_{\mathrm{e}},1} > 0$. This is different from

265   the satellite graph, as explained in 2.2, and is due to the lack of cloud droplet sedimentation and to the saturation adjustment step in the CMP scheme. This means that there is no size-dependence of cloud droplet sedimentation, LW cooling and evaporation in the model. Instead, the arrow from LWP to w_{e} indicates that, in the model, enhanced entrainment is purely due to the increase in LWP (caused by precipitation suppression), i.e. there is more water left in the cloud to evaporate from a water budget perspective. This will cause evaporative cooling, generate more in-cloud turbulence and enhance cloud-top entrainment.

270   Entrainment of warm and dry free-tropospheric air into the cloud has the same effect on cloud droplets in the satellite data and in the model (see blue arrows from w_{e} to N_{d} and r_{eff}): evaporation. In the model, both N_{d} and r_{eff} are reduced, indicating a mix of homogeneous and inhomogeneous entrainment regimes (Hill et al., 2009). In this study, the satellite data display a more homogeneous regime with a reduction in r_{eff} but not N_{d}, indicating that the mixing timescale is longer than the evaporation timescale, allowing the entrained air to mix with the cloud and evaporate all cloud droplets homogeneously. Fons et al. (2023)

275   found a mix of homogeneous and inhomogeneous entrainment regimes. The difference might arise due to the different regions

and seasons considered in the studies. After the entrained air has led to the evaporation of cloud droplets, the latent heat consumption by evaporation generates cooling and turbulence, leading to further entrainment enhancement. Mathematically, this feedback loop can be confirmed in the satellite graph by following the path from w_e, through H and r_{eff} and then back to w_e: $\beta_{w_e,w_e,2} - \alpha_{w_e,w_e,1} \times \alpha_{w_e,w_e,1} = \alpha_{w_e,r_{eff},1} \times \alpha_{r_{eff},w_e,1} + \alpha_{w_e,H,1} \times \alpha_{H,r_{eff},0} \times \alpha_{r_{eff},w_e,1} > 0$ ($\alpha_{w_e,w_e,1} \times \alpha_{w_e,w_e,1}$ is the causal autodependency component of the total causal effect). It can be noted that the formula is more complicated for the model graph because of the mediation by RR which delays the effect. For a complete description of the other arrows in the causal graph, see Fons et al. (2023).

The application of the causal graph shows that low-level clouds in ICON respond as expected to aerosol perturbations, with precipitation suppression and cloud-top entrainment enhancement. This answers questions (1), (2), and (3) as enumerated at the end of the previous section: (1) we do detect local aerosol effects on the ICON clouds; (2) entrainment is enhanced even without cloud droplet sedimentation; (3) entrainment enhancement leads to LWP reductions due to evaporation, even though these reductions are masked in the positive net LWP adjustment (Fig. 2h).

3.2 Cloud deepening

The causal processes described above do not explain why cloud depth seems to increase with aerosols in ICON. In fact, we could even predict cloud thinning due to entrainment-related evaporation, as indicated by the blue arrow from w_e to H (in both the satellite and the model data). Cloud thinning by entrainment can happen under certain thermodynamic conditions (Wood, 2007) and has been simulated in limited area models (Ackerman et al., 2004; Bretherton et al., 2007). However, cloud thinning by entrainment can be compensated by other processes, potentially causing a net deepening effect. This invites us to think of new mechanisms to explain cloud deepening: warm cloud invigoration (Douglas and L'Ecuyer, 2021) or aerosol-induced rain delays and subsequent cloud layer growth (Seifert et al., 2015; Vogel et al., 2016).

Warm cloud invigoration occurs because condensational growth on numerous and smaller droplets is more efficient than condensational growth on larger droplets, thereby generating additional latent heat and turbulence, which can drive increases in updraft speeds. This can drive cloud deepening and increase N_d, due to the dependence of CCN activation on supersaturation (Segal and Khain, 2006). However, such a mechanism is not possible in this model set-up due to saturation adjustments in the CMP scheme. In case of supersaturation, the saturation adjustment scheme condenses all the excess water vapor, independent of the cloud droplet number concentration or radius. Instead, we look at how aerosol-induced rain delays might affect the cloud depth: as aerosols reduce droplet sizes and precipitation is suppressed, clouds deepen to produce rain and maintain a so-called subsiding radiative-convective equilibrium (SCRE) that is imposed by the balance of the large-scale subsidence and evaporative and convective tendencies (Seifert et al., 2015; Rosenfeld et al., 2019). LEM experiments by Vogel et al. (2016) show that, as precipitation is reduced (or even suppressed), convective mixing moistens the inversion if the subsidence is weak, resulting in cloud vertical growth and boundary layer growth (seen in three regions in Fig. 2i). Considering the low subsidence in most of the study regions (Fig. 4c and Suppl. Fig. 1) and the weak inversion strength (Fig. 2j), the rain-delay hypothesis is a likely process to explain the cloud deepening as seen in ICON.

The arrow from RR to H allows us to test the rain-delay hypothesis (Fig. 7). In the model data (Fig. 7a), the lag-1 arrow from
310 RR to H is found to be negative, describing how early precipitation onset prevents further cloud deepening. In other terms, when
precipitation is suppressed due to aerosol increases, smaller droplets can be lifted higher as they grow in the updraft region,
resulting in cloud deepening under a weak inversion. One immediate consequence is LWP build-up, as $\alpha_{N_{\mathrm{d}},r_{\mathrm{eff}},0} \times \alpha_{r_{\mathrm{eff}},\mathrm{RR},1} \times$
$\alpha_{\mathrm{RR},H,1} \times \alpha_{H,\mathrm{LWP},0} > 0$. Eventually, the cloud deepening can allow the cloud to form precipitation ($\alpha_{H,r_{\mathrm{eff}},0} \times \alpha_{r_{\mathrm{eff}},\mathrm{RR},1} > 0$)
and satisfy the SCRE imposed by large-scale forcing.

315 The arrow from RR to H is found to be strongly positive in the satellite data (Fig. 7b). A positive effect of precipitation
on cloud depth seems physically unlikely (Wood, 2007), therefore we hypothesize that this is due to the violation of the
adiabaticity assumption used to compute H from satellite retrievals when clouds are precipitating. The arrow is found to be
weakly negative when H is taken from ERA5 (Suppl. Fig. 4b), in agreement with the ICON datam but weakly negative when
both H and RR are estimated using ERA5. Christensen and Stephens (2011) used satellite data to evaluate aerosol effects
320 on stratocumulus cloud deepening, and found no significant cloud deepening in closed cell regimes, while it could occur for
open cell regimes in unstable boundary layers. Given the prevalence of closed cell morphologies in the stratocumulus regions
considered here (e.g., Wood and Hartmann, 2006, or Suppl. Fig. 5), one could have expected a null effect of aerosols on cloud
depth in the satellite data. However, given the dependence on the choice of cloud depth and precipitation proxies, the effect of
RR on H in the satellite data is marked as 'Inconclusive' in Table 1.

325 The analysis of the RR to H arrow answers question (4), i.e. are additional processes, like cloud deepening, also important
for LWP adjustments in ICON? We confirm that aerosol increases can result in cloud deepening due to rain delays under
a weak inversion in the model, contributing to positive LWP adjustments. In the satellite data, it is not as clear whether
cloud deepening contributes to LWP build-up due to the uncertainty surrounding both precipitation retrievals and cloud depth
estimations. Analysis of other data sources, e.g. in situ field data might help to answer this question.

330 ## 3.3 Temporal developments

Fig. 8 shows how the precipitation and entrainment responses develop with time. These temporal developments $\beta_{X_i,X_j,l}$ were
computed using Wright's path rule, with time lags from $l = 0$ to $l = 24\,\mathrm{h}$. The temporal developments are consistent with
the direct effects discussed above: $\beta_{N_{\mathrm{d}},\mathrm{RR},l}$ is consistently negative, in both the satellite data and the model data, indicating
precipitation suppression (Fig. 8a). Interestingly, $\beta_{N_{\mathrm{d}},RR}$ is weaker in the model than in the satellite data. This too-weak
335 dependence of RR on N_{d} could explain why precipitation rates do not decrease much when increasing CCN to 1700 cm^{-3} in
the high CCN experiment (Fig. 2f).

In the satellite data, $\beta_{N_{\mathrm{d}},w_{\mathrm{e}},l}$ is consistently positive, indicating long-lasting cloud-top entrainment enhancement (Fig. 8b).
In the model data, the initial positive $\beta_{N_{\mathrm{d}},w_{\mathrm{e}},l}$ becomes slightly negative after three hours, indicating that cloud-top entrainment
first increases, then decreases slightly instead of continuing to increase. This different model behavior is probably explained
340 by the limited mechanisms for entrainment enhancement that are possible in the model.

Fig. 8c, d show the fraction of the total effect of N_{d} on LWP that is mediated by RR and w_{e}, respectively. Note that these
effects are not direct and immediate but mediated and delayed. The mediated effect of RR is computed along the path that starts
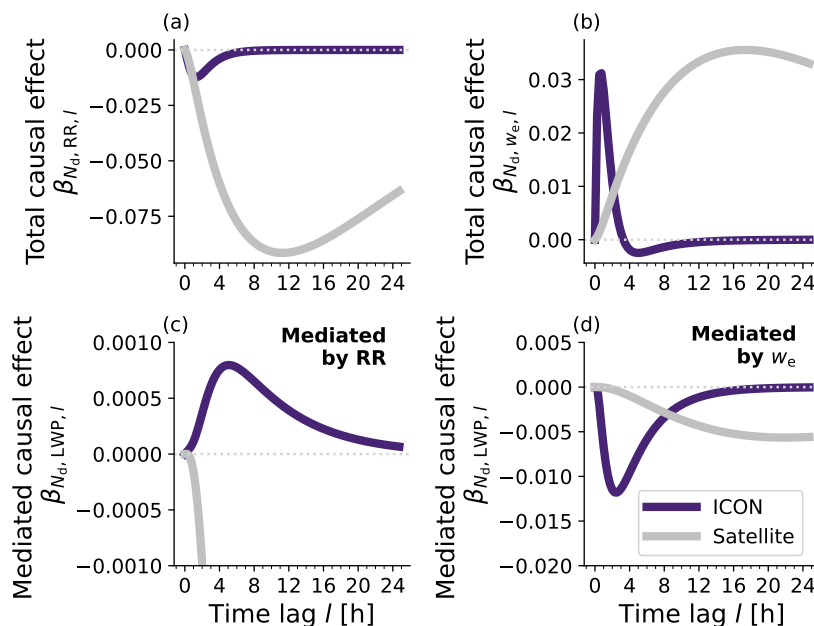
Figure 8. Temporal developments of causal effects. The subplots show the temporal developments of **(a)** the total causal effect of N_d on RR, **(b)** the total causal effect of N_d on w_e, **(c)** the RR-mediated causal effect of N_d on LWP, and **(d)** the w_e-mediated causal effect of N_d on LWP, after an initial positive perturbation in N_d at $l = 0$. In each subplot, the two lines correspond to the two causal graphs presented in Fig. 7a, b. Note that the vertical axis of subplot **(c)** is cropped at -0.001 and therefore cuts the grey line. This is to ensure that the other line is not squished.

from N_d, goes through r_eff and RR (precipitation suppression), and then back to N_d and H (cloud water removal and cloud deepening) before integrating to LWP. For the satellite data, the mediated effect of w_e is calculated along the path that starts at
345   N_d, goes through r_eff then w_e (entrainment enhancement), and then back to N_d, r_eff and H (evaporation) before integrating to LWP (for the model, the path is longer as it goes through RR, then LWP, and then finally w_e).

Fig. 8c shows that the effect of aerosol-induced precipitation suppression on the LWP is positive (albeit weak) in the model data (purple line), in line with the expected LWP build-up from precipitation suppression. For the satellite data, the RR-mediated LWP response (grey line) gets becomes very negative very quickly (cropped in Fig. 8c) due to the spurious positive
350   arrow from RR to H in Fig. 7b. Negative precipitation-mediated influences on the LWP, as seen in the satellite data, are deemed to be non-physical and probably stemming from the violation of the adiabaticity assumption when clouds are precipitating and/or limitations of remote-sensing precipitation products for stratocumulus clouds (e.g., Zhu et al., 2022). The positive precipitation-mediated response of stratocumulus clouds to aerosols should mostly result from reduced cloud water removal, and to some extent from cloud deepening if the inversion is weak. However, the causal graph actually detects that the positive
355   response mostly results from cloud deepening in the model (compare Fig. 8c to Suppl. Fig. 6). This indicates that the causal method applied here might be limited to accurately detect all precipitation processes, as already discussed in Sec 3.1.

Fig. 8d shows that the effect of aerosol-induced entrainment enhancement on the LWP is negative, denoting continued evaporation due to mixing of the warm and dry free tropospheric air into the cloud. By comparing the scales of Fig. 8c-d, it would seem that the entrainment-mediated reductions in LWP should prevail compared to the precipitation-mediated build-up in the model. However, this is not possible since the net aerosol effect on LWP is positive in the model (Fig. 2h): we hypothesize that the issues associated with RR in the causal graph, as identified above, might be responsible for this too-weak mediated effect of RR.

The timescales of precipitation suppression and entrainment enhancement are much shorter for the ICON clouds, with model peak responses occurring within a couple of hours and satellite peak responses occurring after 12 hours. It would make sense that these processes occur faster in the model than in reality, as ICON simulates clouds with cumuliform features rather than stratiform features and cumulus clouds have notably shorter lifetimes (< 4-5 h, Seifert et al., 2015) than stratocumulus clouds (< 24 h, Christensen et al., 2020). However, it is possible that the longer timescales in the satellite data are due to the mismatch of temporal resolution between the co-located satellite and reanalysis products, introducing an artificial delay in the computed causal effects.

The temporal developments confirm the results of the analyses of the direct causal effects: aerosols cause precipitation suppression and associated LWP build-ups, as well as entrainment enhancement and associated LWP reductions in both the observations and in the model. In particular, cloud deepening enhances the positive influence of RR on LWP in the model. The temporal developments also provide additional information concerning timescales, and seem to confirm that ICON clouds are shorter-lived than real stratocumulus clouds, with faster precipitation and entrainment responses to aerosols than observations.

4 Conclusions

The ICON model is able to simulate realistic low-level clouds in stratocumulus regions, even though there are systematic biases concerning cloud droplet number concentrations, precipitation rates and cloud morphologies with respect to satellite observations. We also observe that, similarly to typical GCMs, ICON simulates positive stratocumulus LWP adjustments to aerosol increases. This suggests that positive effects from aerosol-induced precipitation suppression prevail compared to negative influences of cloud-top entrainment enhancement. We used a causal graph as a diagnosis tool to evaluate observation-model differences. We found that the positive response of LWP to aerosols in the model results from a superposition of different processes.

First, precipitation suppression leads to LWP build-up, as expected from Albrecht (1989). As in Fons et al. (2023) this is only partially detected by the causal approach, potentially due to uncertainties in precipitation estimates (e.g., Zhu et al., 2022)), non-linearities of precipitation processes, hidden confounding (e.g., by relative humidity Grandey et al., 2014), or a temporal resolution that is coarser than the process timescale of precipitation processes (Runge, 2018). However, it is clear that, with average rain rates one to two orders of magnitudes higher than observed rain rates, aerosol effects will cause precipitation suppression and LWP build-up due to the explicit parameterization of rain autoconversion as a function of cloud droplet number concentration in the cloud microphysics parameterization scheme (Seifert and Beheng, 2006).

Second, we detect that delays in the onset of precipitation can cause cloud deepening in ICON (Seifert et al., 2015). In the satellite data, such a response is not as clearly detected, potentially due to the difficulty in reliably estimating cloud depths from satellite or reanalysis data. However, in a stable stratocumulus-topped boundary layer, we might not expect any vertical growth if the inversion strength is high enough, as observed by Christensen and Stephens (2011).

Third, we observe that cloud-top entrainment enhancement does occur in ICON, even without a parameterization of entrainment as a function of droplet sizes. Entrainment enhancement was not a priori expected as the ICON set-up used here does not parameterize cloud droplet sedimentation, which has been identified as a key process to initiate the entrainment enhancement feedback loop (Ackerman et al., 2004; Bretherton et al., 2007). In the model, entrainment enhancement is purely driven from increased evaporation of accumulated LWP due to precipitation suppression. This is different from droplet-size dependent entrainment mechanisms that occur in observations (radiative and evaporative feedbacks) due to parameterization choices that limit the realm of what is possible in the model. Like in observations, entrainment enhancement causes evaporation of the cloud droplets, with a negative effect on the LWP. However, this negative effect is weaker than the positive effects due to the superposition of precipitation suppression and cloud deepening in ICON.

Finally, we observe that cloud processes occur faster in the model compared to observations, which is consistent with how the clouds look: In ICON, clouds are more disagreggated and look more cumuliform, while in reality, stratocumulus clouds are stratiform and more aggregated with a longer lifetime than cumulus clouds (Jiang et al., 2006; Christensen et al., 2020).

In summary, a causal graph can be used as a diagnosis tool to evaluate cloud processes across different data sources by removing confounding from the large-scale environment and decomposing a net confounded effect into its physical causal components. Of course, the conclusions are dependent upon the validity of the causal graph and the reliability of the given data sets. In Fons et al. (2023), we tested several other causal graphs to evaluate the robustness of the results, allowing us to select the causal graph with the highest physical plausibility. This physical graph is adjusted for the present study to account model-reality differences, and to add a physical process (cloud deepening due to rain delays). Despite differences in the data sets (regions, seasons, resolution), the signs of the causal effects are consistent between the previous study and this study, increasing our confidence in the methodology. However, the causal methodology seems unable to reliably estimate absolute magnitudes of causal effects, e.g. making it difficult to assess whether the entrainment response or the precipitation response dominates (compare Fig. 8c and d). This could be related to the linearity assumption used in the present causal approach, and calls for an investigation of non-linear methods in future research (Runge, 2018).

Even though the absolute magnitudes of the causal effects are considered unreliable, signs of causal effects can be used to diagnose the existence (and direction) of physical processes in the model, hinting at model modifications that might yield a better model-observation agreement. For example, a strong inversion is a key feature of stratocumulus regions (Wood, 2012) and might be achieved in the model by increasing the vertical resolution and/or perturbing the turbulence scheme (Possner et al., 2014, e.g.), possibly preventing cloud deepening and a shift from a stratiform to a cumuliform regime. This cloud regime shift could also reduce the high precipitation bias, and a stronger humidity gradient at the inversion could enhance the negative effects of entrainment on LWP. Should it still be necessary, other modifications could be implemented, e.g.: activating cloud droplet sedimentation, as previous studies (Ackerman et al., 2004; Bretherton et al., 2007) have highlighted the key role of

425 this process in cloud-top entrainment enhancement and resulting negative LWP adjustments; or reducing the efficiency of rain autoconversion. Implementing such changes in the model could bring model estimates of aerosol-cloud radiative forcing closer to observations, and eventually increase our confidence in climate projections made with GSRMs.

Code and data availability. Code for the data processing and analysis is provided on Zenodo (10.5281/zenodo.10580241).

The ICON model code is available online (https://code.mpimet.mpg.de/projects/iconpublic/wiki/Instructions_to_obtain_the_ICON_model_

430 code_with_a_personal_non-commercial_research_license). The simulation runscripts are available with the rest of the code on zenodo.

The satellite timeseries used for the analyses were generated from co-located SEVIRI (Copyright (c) (2020) EUMETSAT), GOES-R (NOAA), GPM and ERA5 data (generated using Copernicus Climate Change Service information [2022]). MODIS Level 3 data (used for comparison purposes) were downloaded from https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/61/MOD08_D3. MIDAS morphology data (used for a supplementary figure) were provided by I. McCoy in a personal communication. GOES data were downloaded from

435 NOAA's AWS API. SEVIRI data are freely available from https://wui.cmsaf.eu/safira, GPM data from https://disc.gsfc.nasa.gov/datasets and ERA5 data from https://cds.climate.copernicus.eu. The processed timeseries and analyses outputs are provided on Zenodo (10.5281/zenodo.10580438).

Appendix A: Methods

A1   ICON set-up

440 We use the ICOsahedral Nonhydrostatic model (ICON, Zängl et al., 2015) as a GSRM in the Sapphire configuration (Hohenegger et al., 2023), with prescribed sea surface temperatures and initial conditions taken from the European Centre of Medium Range Weather Forecast (ECMWF) analysis. ICON is run with the same set up as in Lang et al. (2023), i.e. in a global set-up, for 45 days, from the 27th of June, 2021 to the 9th of August, 2021. The first 4 days are considered model spin-up and left out of the analysis. The model time step is 40 s, and the outputs are instantaneous values with a an output frequency of 15 min

445 to enable the comparison with the satellite data. The model is run at approximately 5 km horizontal resolution. While this can seem too coarse to accurately capture stratocumulus cloud processes, Heim et al. (2021) noted that there was no significant improvement in low cloud representation when increasing the horizontal resolution from 4 km to 500 m. The vertical grid is made of 110 hybrid sigma levels between the surface and a height of 75 km. In the boundary layer, the vertical resolution increases progressively from 25 m at the surface to about 400 m above the inversion.

450 In order to evaluate aerosol effects on the stratocumulus cloud decks, we used the two-moment cloud microphysics scheme by Seifert and Beheng (2006). We ran two simulations: one with moderate CCN concentrations (250 cm^{-3}) and one with high CCN concentrations (1700 cm^{-3}), all other set up parameters being equal otherwise. Note that the CCN concentration is not uniform over the whole atmospheric column, but instead is set to the fixed value in the lower troposphere and decays exponentially with altitude. For the sake of simplicity, these 2 simulations are referred to as 'ICON (low CCN)' and 'ICON

455 (high CCN)' in this article. The first simulation was run for 45 days (minus 4 spin-up days), while the second simulation was

only run for 12 days (minus 4 spin-up days) due to the high computing time requirements of global storm-resolving runs with a two-moment cloud microphysics. The common analysis period for the 2 simulations (01.07.21 to 09.07.21) is 8-day long, while the common analysis period for the satellite data vs. the low CCN experiment is 41-day long. The shorter common period is implicitly used throughout the article whenever the high CCN experiment is analyzed (Fig. 2, 4, 5), while the longer period
460   is used when only the low CCN experiment is used (all of Sec. 3).

Other parameterization choices include: the 3D turbulence mixing scheme from Smagorinsky (1963) with the modification by Lilly (1962) (as implemented by Dipankar et al., 2015; Lee et al., 2022), the RTE-RRTMGP scheme by Pincus et al. (2019) for radiative transfer, and the JSBACH land model from Raddatz et al. (2007). The shallow and deep convection schemes are switched off, and we use an all-or-nothing cloud scheme, i.e. the cloud fraction is set to 0 or 1 depending on a threshold for
465   cloud water and cloud ice content.

A2   ICON time series

The standard outputs of the model were processed to make up time series for the 6 variables of interest in this study: low-level cloud droplet number concentrations N_{d}, cloud droplet radii (called r_{eff} for consistency with the satellite denomination), cloud depths H, cloud-top entrainment rates w_{e}, rain rates RR and liquid water paths LWP. Stratocumulus clouds were identified as
470   the uppermost cloud layer within the boundary layer. Their H was computed as the sum of the cloudy model level heights. Their N_{d} is averaged over the cloud height, while their r_{eff} is computed as the maximum over the cloud height. We made this choice to try to be consistent with the adiabatic assumption that is used for satellite retrievals of cloud microphysical properties. In adiabatic clouds, the N_{d} profile is constant with height, while the r_{eff} profile increases with height above cloud base, meaning that satellites will observe cloud-top droplet radii that are close to the maximum (Lohmann et al., 2016). The
475   cloud's LWP is obtained by summing the product of the LWC times the layer depth for each cloud level. LWP is defined as the cloud liquid water path only, and does not include rain water. To identify the uppermost cloud level in the boundary layer, we had to diagnose the boundary layer depth BLH. BLH was estimated as the height at which the relative humidity drops below 50 % (as in Bretherton et al., 2013). As shown in Fig. 3, this corresponds well to the cloud top. w_{e} was diagnosed using a boundary layer mass balance equation from Stull (1988):

480
$$\frac{\mathrm{d\,BLH}}{\mathrm{d}t} = \frac{\partial\,\mathrm{BLH}}{\partial t} + \boldsymbol{v} \cdot \nabla\mathrm{BLH} = w_{\mathrm{e}} + w_{\mathrm{subs}},$$

where $\boldsymbol{v}$ is the horizontal wind vector and w_{subs} is the large-scale subsidence rate, which we estimate as the vertical velocity at the inversion level. Like in Heim et al. (2021), we make the approximation that the instantaneous term of the derivative is negligible compared to the advective term, and we thus calculate w_{e} as $\boldsymbol{v} \cdot \nabla\mathrm{BLH} - w_{\mathrm{subs}}$. EIS was diagnosed following Wood and Bretherton (2006). For RR, only surface precipitation is a standard output of the model. Therefore, we diagnosed cloud-
485   base RR a posteriori, as a function of the rain water content and the rain drop number concentrations following the parametric equations from Seifert and Beheng (2006).

A3 Satellite time series

The simulation outputs are compared to geostationary satellite cloud retrievals obtained from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) aboard the 11th Meteosat satellite, and the Advanced Baseline Imager ABI aboard the 16th and
490   17th Geostationary Operational Environmental Satellites (GOES 16 and 17). With a central longitude of 0°, Meteosat 11 can see the South-East and North-East Atlantic, including the Namibian and Canarian Sc decks, as defined by Klein and Hartmann (1993). GOES 17 sees the Californian deck in the North East Pacific (central longitude of -137.2°), while GOES 16 sees the Peruvian deck in the South East Pacific (central longitude of -75.2°). We match the time period of the satellite data to the time period of the GSRM runs (1st of July- 9th August, 2021). For SEVIRI, we downloaded the CPP and CTX products of
495   the CLoud property dAtAset using SEVIRI (CLAAS) 3.0 (Benas et al., 2023). For GOES, we downloaded the cloud fraction, cloud optical depth, cloud droplet effective radius, cloud top temperature and pressure (ABI-L2-ACMF, ABI-L2-CODF, ABI-L2-CPSF, ABI-L2-ACHTF and ABI-L2-CTPF) (Walther and Straka, 2020). Low clouds were filtered by selecting only liquid clouds in SEVIRI and by selecting cloud top pressures larger than 680 hPa in GOES. We calculated the cloud droplet number concentration, the cloud depth and the liquid water path using the cloud adiabaticity assumption, following Brenguier et al.
500   (2000) and Quaas et al. (2006). For further details concerning these calculations, see Fons et al. (2023). Both GOES-ABI and Meteosat-SEVIRI data have been validated (Walther and Straka, 2020; Benas et al., 2023) against more commonly used polar-orbiting satellite instruments, like the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Aqua and Terra satellites (Platnick et al., 2015). However, few studies specifically validate N_d derivations from geostationary satellites against those from polar-orbiting satellites, so we include in the supplementary material a comparison of GOES/SEVIRI N_d to
505   MODIS N_d (Suppl. Fig. 7). The geostationary N_d agrees pretty well with the MODIS N_d, except for a constant positive bias in the SEA region (unproblematic given the data standardization), and a non-constant positive bias in the NEA region.

     We colocated the satellite data with precipitation data from the Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals Version 7 (iMERG V07) (Huffman et al., 2023). We then added reanalysis data from ERA5 (Hersbach et al., 2018a, b): BLH, EIS, w_e, H. EIS was diagnosed from pressure level ERA5 data following Wood and Bretherton (2006).
510   w_e was calculated following the equation from Stull (1988) as described above. Although we already had an estimate for H from the satellite data, we also diagnosed it from ERA5 as a comparison, by summing cloudy level heights in the boundary layer.

     All data were colocated to the 0.25° grid of ERA5 and linearly interpolated to the temporal resolution of SEVIRI (15 min). Because of the mismatch between the low cloud cover between ERA5 and the satellites (e.g. Fig. 1 e, m), we average the
515   co-located data to a coarser 0.5° resolution. We assume that, at these coarser resolutions, the cloud top entrainment derived from ERA5 will be approximately co-located with the right cloud structures in the satellite data, even if the cloud structures were not strictly matching on a pixel-by-pixel basis in the high resolution data. To allow for a fair comparison between the satellite and the model data, the model time series (obtained as described above) were also averaged to a coarser 0.5° grid. Note that Fons et al. (2023) carried out the same type of causal analyses as the ones carried out here, but using coarser satellite
520   data (10°).

A4 Causal inference

The causal method used here consists in applying a causal graph of LWP adjustments to time series data of cloud properties. The graph was drawn by the authors, i.e. it was derived from domain knowledge and not obtained by causal discovery (Runge et al., 2023). The time series used here have a temporal resolution of Δt = 15 min, which is close to the average process

525 timescale of air parcel movements from base to top in a stratocumulus cloud. This is ideal as it allows to resolve feedback loops and witness how changes in cloud properties (e.g. N_d) propagate in time, using the precedence of cause with respect to effect (Fons et al., 2023).
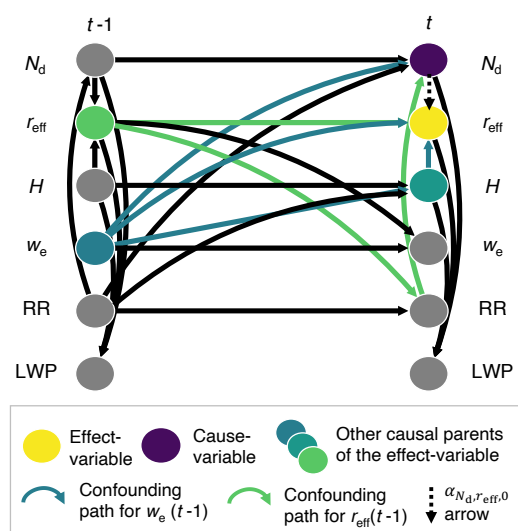
Figure A1. Calculations of direct causal effects using Wright's approach. This figure demonstrates how to calculate $\alpha_{N_d, r_{eff}, 0}$ in the framework of Wright's linear approach, by carrying out a multiple linear regression of $r_{eff}(t)$ on all its causal parents, including $N_d(t)$. The other causal parents are: $r_{eff}(t-1)$ and $w_e(t-1)$, which are both confounders (see teal green and lime green arrows), as well as $H(t)$. Note that $H(t)$ itself is not a confounder but because it is a parent of the effect-variable, controlling for its effect by including it in the regression helps to reduce the variance of the results (Runge, 2021).

Using Wright's linear path approach (Wright, 1921; Runge et al., 2015), the causal graph is used in combination with the data to detect and remove confounding influences and compute causal effects of aerosols on the LWP, as opposed to traditional

530 linear regression coefficients. Direct causal effects refer to the causal effects between two variables that are directly connected by a directed arrow in the causal graph. Direct causal effects are calculated as partial linear regression coefficients of the effect-variable X_j on the cause-variable X_i in the multiple linear regression of X_j on all its causal parents, i.e. those variables with arrows pointing directly towards X_j in the causal graph (see an example in Fig. A1). Total causal effects designate causal effects between two variables that are not directly connected by an arrow in the causal graph. Total causal effects are calculated

535 from the direct causal effects using Wright's sum of product rule. Mediated causal effects correspond to the fraction of a total causal effect that is mediated by a given variable. For further details, please see the methods section in Fons et al. (2023),

which uses the same exact methodology. All computations and graph plots are made using the Tigramite package in Python (https://github.com/jakobrunge/tigramite).

Here, the LWP causal graph is applied to the 41 days of the satellite and ICON time series. Note that all variables X in the time series were corrected for the diurnal cycle prior to the causal effect computations, as the diurnal cycle can constitute a source of confounding.

$$X_{\text{corrected}}(t) = \frac{X(t) - \overline{X(t)}}{\sigma_{X(t)}},$$

where $\overline{X(t)}$ and $\sigma_{X(t)}$ are the diurnal instantaneous average and standard deviations, i.e. the average/standard deviation of X at a given time of day (e.g. 10:15 am), computed over the whole time series. The time series were not corrected for the seasonal cycle given their short duration. The causal effect computations are run for the four regions all together (i.e. aggregated), assuming the low-level clouds in these four regions obey to the same physical processes.

Confidence intervals were computed using a bootstrapping method with $n = 100$ members. Direct causal effects are considered significantly positive or negative when the bootstrap confidence interval does not include 0.

Author contributions. E.F. developed the concept of the study together with U.L. and D.N. T.L. and E.F. ran the ICON model simulations under the guidance of A.K.N. E.F. wrote the code for the data processing, for the causal workflow and for the data post-processing. E.F., A.K.N., D.N. and U.L. worked on to the interpretations of the results. E.F. drafted the manuscript with contributions from all other co-authors.

Competing interests. The authors declare no competing interests.

Acknowledgements. This work was supported by the European Union's Horizon 2020 research and innovation program under Marie Sklodowska-Curie grant agreement No. 860100 (iMIRACLI). A.K.N. and T.L. were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2037 'CLICCS—Climate, Climatic Change, and Society'—Project Number 390683824.

References

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014–1017, https://doi.org/10.1038/nature03174, 2004.

Ahlgrimm, M., Forbes, R. M., Hogan, R. J., and Sandu, I.: Understanding Global Model Systematic Shortwave Radiation Errors in Subtropical Marine Boundary Layer Cloud Regimes, Journal of Advances in Modeling Earth Systems, 10, 2042–2060, https://doi.org/10.1029/2018MS001346, 2018.

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

Benas, N., Solodovnik, I., Stengel, M., Hüser, I., Karlsson, K.-G., Håkansson, N., Johansson, E., Eliasson, S., Schröder, M., Hollmann, R., and Meirink, J. F.: CLAAS-3: The Third Edition of the CM SAF Cloud Data Record Based on SEVIRI Observations, Earth System Science Data Discussions, pp. 1–38, https://doi.org/10.5194/essd-2023-79, 2023.

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.

Bretherton, C. S., Blossey, P. N., and Uchida, J.: Cloud Droplet Sedimentation, Entrainment Efficiency, and Subtropical Stratocumulus Albedo, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL027648, 2007.

Bretherton, C. S., Blossey, P. N., and Jones, C. R.: Mechanisms of Marine Low Cloud Sensitivity to Idealized Climate Perturbations: A Single-LES Exploration Extending the CGILS Cases, Journal of Advances in Modeling Earth Systems, 5, 316–337, https://doi.org/10.1002/jame.20019, 2013.

Chen, Y., Haywood, J., Wang, Y., Malavelle, F., Jordan, G., Partridge, D., Fieldsend, J., De Leeuw, J., Schmidt, A., Cho, N., Oreopoulos, L., Platnick, S., Grosvenor, D., Field, P., and Lohmann, U.: Machine Learning Reveals Climate Forcing from Aerosols Is Dominated by Increased Cloud Cover, Nature Geoscience, 15, 609–614, https://doi.org/10.1038/s41561-022-00991-6, 2022.

Christensen, M., Wk, J., and P, S.: Aerosols Enhance Cloud Lifetime and Brightness along the Stratus-to-Cumulus Transition, Proceedings of the National Academy of Sciences of the United States of America, 117, 17 591–17 598, https://doi.org/10.1073/pnas.1921231117, 2020.

Christensen, M. W. and Stephens, G. L.: Microphysical and Macrophysical Responses of Marine Stratocumulus Polluted by Underlying Ships: Evidence of Cloud Deepening, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD014638, 2011.

Dagan, G., Yeheskel, N., and Williams, A. I. L.: Radiative Forcing from Aerosol–Cloud Interactions Enhanced by Large-Scale Circulation Adjustments, Nature Geoscience, 16, 1092–1098, https://doi.org/10.1038/s41561-023-01319-8, 2023.

Dipankar, A., Stevens, B., Heinze, R., Moseley, C., Zängl, G., Giorgetta, M., and Brdar, S.: Large Eddy Simulation Using the General Circulation Model ICON, Journal of Advances in Modeling Earth Systems, 7, 963–986, https://doi.org/10.1002/2015MS000431, 2015.

Douglas, A. and L'Ecuyer, T.: Global Evidence of Aerosol-Induced Invigoration in Marine Cumulus Clouds, Atmospheric Chemistry and Physics, 21, 15 103–15 114, https://doi.org/10.5194/acp-21-15103-2021, 2021.

Fons, E., Runge, J., Neubauer, D., and Lohmann, U.: Stratocumulus Adjustments to Aerosol Perturbations Disentangled with a Causal Approach, npj Climate and Atmospheric Science, 6, 1–10, https://doi.org/10.1038/s41612-023-00452-w, 2023.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 923–1054,

[Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)] Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/9781009157896.009, 2021.

Grandey, B. S., Gururaj, A., Stier, P., and Wagner, T. M.: Rainfall-Aerosol Relationships Explained by Wet Scavenging and Humidity, Geophysical Research Letters, 41, 5678–5684, https://doi.org/10.1002/2014GL060958, 2014.

Gryspeerdt, E., Quaas, J., and Bellouin, N.: Constraining the Aerosol Influence on Cloud Fraction, Journal of Geophysical Research: Atmospheres, 121, 3566–3583, https://doi.org/10.1002/2015JD023744, 2016.

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.

Heim, C., Hentgen, L., Ban, N., and Schär, C.: Inter-model Variability in Convection-Resolving Simulations of Subtropical Marine Low Clouds, Journal of the Meteorological Society of Japan. Ser. II, 99, 1271–1295, https://doi.org/10.2151/jmsj.2021-062, 2021.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thébaut, J.-N.: ERA5 Hourly Data on Single Levels from 1959 to Present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.adbb2d47, 2018a.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thébaut, J.-N.: ERA5 Hourly Data on Pressure Levels from 1959 to Present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.bd0915c6, 2018b.

Hill, A. A., Feingold, G., and Jiang, H.: The Influence of Entrainment and Mixing Assumption on Aerosol–Cloud Interactions in Marine Stratocumulus, Journal of the Atmospheric Sciences, 66, 1450–1464, https://doi.org/10.1175/2008JAS2909.1, 2009.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales, Geoscientific Model Development, 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, publisher: Copernicus GmbH, 2023.

Huffman, G., Stocker, E., Bolvin, D., Nelkin, E., and Tan, J.: GPM IMERG Final Precipitation L3 Half Hourly 0.1 Degree x 0.1 Degree V07, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/GPM/IMERG/3B-HH/07, 2023.

Jiang, H., Xue, H., Teller, A., Feingold, G., and Levin, Z.: Aerosol Effects on the Lifetime of Shallow Cumulus, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026024, 2006.

Judt, F., Klocke, D., Rios-Berrios, R., Vanniere, B., Ziemen, F., Auger, L., Biercamp, J., Bretherton, C., Chen, X., Düben, P., Hohenegger, C., Khairoutdinov, M., Kodama, C., Kornblueh, L., Lin, S.-J., Nakano, M., Neumann, P., Putman, W., Röber, N., Roberts, M., Satoh, M., Shibuya, R., Stevens, B., Vidale, P. L., Wedi, N., and Zhou, L.: Tropical Cyclones in Global Storm-Resolving Models, Journal of the Meteorological Society of Japan. Ser. II, 99, 579–602, https://doi.org/10.2151/jmsj.2021-029, 2021.

Kalmus, P., Wong, S., and Teixeira, J.: The Pacific Subtropical Cloud Transition: A MAGIC Assessment of AIRS and ECMWF Thermodynamic Structure, IEEE Geoscience and Remote Sensing Letters, 12, 1586–1590, https://doi.org/10.1109/LGRS.2015.2413771, 2015.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587–1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

635 Köhler, M., Ahlgrimm, M., and Beljaars, A.: Unified Treatment of Dry Convective and Stratocumulus-Topped Boundary Layers in the ECMWF Model, Quarterly Journal of the Royal Meteorological Society, 137, 43–57, https://doi.org/10.1002/qj.713, 2011.

Lang, T., Naumann, A. K., Buehler, S. A., Stevens, B., Schmidt, H., and Aemisegger, F.: Sources of Uncertainty in Mid-Tropospheric Tropical Humidity in Global Storm-Resolving Simulations, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 443, https://doi.org/10.1029/2022MS003443, 2023.

640 Lee, J., Hohenegger, C., Chlond, A., and Schnur, R.: The Climatic Role of Interactive Leaf Phenology in the Vegetation-Atmosphere System of Radiative-Convective Equilibrium Storm-Resolving Simulations, Tellus, Series B - Chemical and Physical Meteorology, 74, 164–175, https://doi.org/10.16993/tellusb.26, 2022.

Lilly, D. K.: On the Numerical Simulation of Buoyant Convection, Tellus, 14, 148–172, https://doi.org/10.1111/j.2153-3490.1962.tb00128.x, 1962.

645 Lohmann, U., Lüönd, F., and Mahrt, F.: Microphysical Processes in Warm Clouds., in: An Introduction to Clouds: From the Microscale to Climate, pp. 186–217, Cambridge University Press, Cambridge, 2016.

Malavelle, F. F., Haywood, J. M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., Allan, R. P., Karset, I. H. H., Kristjánsson, J. E., Oreopoulos, L., Cho, N., Lee, D., Bellouin, N., Boucher, O., Grosvenor, D. P., Carslaw, K. S., Dhomse, S., Mann, G. W., Schmidt, A., Coe, H., Hartley, M. E., Dalvi, M., Hill, A. A., Johnson, B. T., Johnson, C. E., Knight, J. R., O'Connor, F. M., Partridge, D. G., Stier,
650 P., Myhre, G., Platnick, S., Stephens, G. L., Takahashi, H., and Thordarson, T.: Strong Constraints on Aerosol-Cloud Interactions from Volcanic Eruptions, Nature, 546, 485–491, https://doi.org/10.1038/nature22974, 2017.

Manshausen, P., Watson-Parris, D., Christensen, M. W., Jalkanen, J.-P., and Stier, P.: Invisible Ship Tracks Show Large Cloud Sensitivity to Aerosol, Nature, 610, 101–106, https://doi.org/10.1038/s41586-022-05122-0, 2022.

McCoy, D. T., Field, P., Gordon, H., Elsaesser, G. S., and Grosvenor, D. P.: Untangling Causality in Midlatitude Aerosol–Cloud Adjustments,
655 Atmospheric Chemistry and Physics, 20, 4085–4103, https://doi.org/10.5194/acp-20-4085-2020, 2020.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface, Exeter, Devon, https://scitools.org.uk/cartopy, 2010 - 2015.

Michibata, T., Suzuki, K., Sato, Y., and Takemura, T.: The source of discrepancies in aerosol–cloud–precipitation interactions between GCM and A-Train retrievals, Atmospheric Chemistry and Physics, 16, 15 413–15 424, https://doi.org/10.5194/acp-16-15413-2016, 2016.

Neubauer, D., Christensen, M. W., Poulsen, C. A., and Lohmann, U.: Unveiling Aerosol–Cloud Interactions – Part 2: Minimising the Effects
660 of Aerosol Swelling and Wet Scavenging in ECHAM6-HAM2 for Comparison to Satellite Data, Atmospheric Chemistry and Physics, 17, 13 165–13 185, https://doi.org/10.5194/acp-17-13165-2017, 2017.

Nowack, P., Runge, J., Eyring, V., and Haigh, J. D.: Causal Networks for Climate Model Evaluation and Constrained Projections, Nature Communications, 11, 1415, https://doi.org/10.1038/s41467-020-15195-y, 2020.

Nugent, J. M., Turbeville, S. M., Bretherton, C. S., Blossey, P. N., and Ackerman, T. P.: Tropical Cirrus in Global Storm-Resolving Models:
665 1. Role of Deep Convection, Earth and Space Science, 9, e2021EA001 965, https://doi.org/10.1029/2021EA001965, 2022.

Pincus, R., Mlawer, E. J., and Delamere, J. S.: Balancing Accuracy, Efficiency, and Flexibility in Radiation Calculations for Dynamical Models, Journal of Advances in Modeling Earth Systems, 11, 3074–3089, https://doi.org/10.1029/2019MS001621, 2019.

Platnick, S., King, M., and Hubanks, P.: MODIS Atmosphere L3 Daily Product. NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA., https://doi.org/http://dx.doi.org/10.5067/MODIS/MOD08_D3.006, 2015.

670 Possner, A., Zubler, E., Fuhrer, O., Lohmann, U., and Schär, C.: A Case Study in Modeling Low-Lying Inversions and Stratocumulus Cloud Cover in the Bay of Biscay, Weather and Forecasting, 29, 289–304, https://doi.org/10.1175/WAF-D-13-00039.1, 2014.

Possner, A., Eastman, R., Bender, F., and Glassmeier, F.: Deconvolution of Boundary Layer Depth and Aerosol Constraints on Cloud Water Path in Subtropical Stratocumulus Decks, Atmospheric Chemistry and Physics, 20, 3609–3621, https://doi.org/10.5194/acp-20-3609-2020, 2020.

675 Quaas, J., Boucher, O., and Lohmann, U.: Constraining the Total Aerosol Indirect Effect in the LMDZ and ECHAM4 GCMs Using MODIS Satellite Data, Atmospheric Chemistry and Physics, 6, 947–955, https://doi.org/10.5194/acp-6-947-2006, 2006.

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

Raddatz, T. J., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the Tropical Land Biosphere Dominate the Climate–Carbon Cycle Feedback during the Twenty-First Century?, Climate Dynamics, 29, 565–574,
685 https://doi.org/10.1007/s00382-007-0247-8, 2007.

Rosenfeld, D., Zhu, Y., Wang, M., Zheng, Y., Goren, T., and Yu, S.: Aerosol-Driven Droplet Concentrations Dominate Coverage and Water of Oceanic Low-Level Clouds, Science, 363, eaav0566, https://doi.org/10.1126/science.aav0566, 2019.

Runge, J.: Causal Network Reconstruction from Time Series: From Theoretical Assumptions to Practical Estimation, Chaos: An Interdisciplinary Journal of Nonlinear Science, 28, 075 310, https://doi.org/10.1063/1.5025050, 2018.

690 Runge, J.: Necessary and Sufficient Graphical Conditions for Optimal Adjustment Sets in Causal Graphical Models with Hidden Variables, in: Advances in Neural Information Processing Systems, edited by Ranzato, M. et al., vol. 34, pp. 15 762–15 773, Curran Associates, Inc., 2021.

Runge, J., Petoukhov, V., Donges, J. F., Hlinka, J., Jajcay, N., Vejmelka, M., Hartman, D., Marwan, N., Paluš, M., and Kurths, J.: Identifying Causal Gateways and Mediators in Complex Spatio-Temporal Systems, Nature Communications, 6, 8502,
695 https://doi.org/10.1038/ncomms9502, 2015.

Runge, J., Gerhardus, A., Varando, G., Eyring, V., and Camps-Valls, G.: Causal Inference for Time Series, Nature Reviews Earth & Environment, 4, 487–505, https://doi.org/10.1038/s43017-023-00431-y, 2023.

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., and Düben, P.: Global Cloud-Resolving Models, Current Climate Change Reports, 5, 172–184, https://doi.org/10.1007/s40641-019-00131-0, 2019.

700 Segal and Khain: Dependence of Droplet Concentration on Aerosol Conditions in Different Cloud Types: Application to Droplet Concentration Parameterization of Aerosol Conditions, Journal of Geophysical Research: Atmospheres Wiley Online Library, 111, D15, https://doi.org/https://doi.org/10.1029/2005JD006561, 2006.

Seifert, A. and Beheng, K. D.: A Two-Moment Cloud Microphysics Parameterization for Mixed-Phase Clouds. Part 1: Model Description, Meteorology and Atmospheric Physics, 92, 45–66, https://doi.org/10.1007/s00703-005-0112-4, 2006.

705 Seifert, A., Heus, T., Pincus, R., and Stevens, B.: Large-Eddy Simulation of the Transient and near-Equilibrium Behavior of Precipitating Shallow Convection, Journal of Advances in Modeling Earth Systems, 7, 1918–1937, https://doi.org/10.1002/2015MS000489, 2015.

Smagorinsky, J.: General Circulation Experiments with the Primitive Equations: I. The Basic Experiment, Monthly Weather Review, 91, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2, 1963.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama,
710 C., Kornblueh, L., Lin, S.-J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., Wedi, N., and Zhou, L.: DYAMOND: The DYnamics of the Atmospheric General Circulation Modeled On Non-hydrostatic Domains, Progress in Earth and Planetary Science, 6, 61, https://doi.org/10.1186/s40645-019-0304-z, 2019.

Stevens, B., Acquistaoace, C., Hansen, A., Heinze, R., Klinger, C., Klocke, D., Rybka, H., Schubotz, W., Windmiller, J., Adamidis, P., Arka, I., Barlakas, V., Biercamp, J., Brueck, M., Brune, S., Buehler, S. A., Burkhardt, U., Cioni, G., Costa-Suros, M., Crewell, S., Crüger, T.,
715 Deneke, H., Friedrichs, P., Henken, C. C., Hohenegger, C., Jacob, M., Jakub, F., Kalthoff, N., Köhler, M., Laar, T. W. v., Li, P., Löhnert, U., Macke, A., Madenach, N., Mayer, B., Nam, C., Naumann, A. K., Peters, K., Poll, S., Quaas, J., Röber, N., Rochetin, N., Scheck, L., Schemann, V., Schnitt, S., Seifert, A., Senf, F., Shapkalijevski, M., Simmer, C., Singh, S., Sourdeval, O., Spickermann, D., Strandgren, J., Tessiot, O., Vercauteren, N., Vial, J., Voigt, A., and Zängl, G.: The Added Value of Large-Eddy and Storm-Resolving Models for Simulating Clouds and Precipitation, Journal of the Meteorological Society of Japan, 98, 395, https://doi.org/10.2151/jmsj.2020-021,
720 2020.

Stier, P.: Limitations of Passive Remote Sensing to Constrain Global Cloud Condensation Nuclei, Atmospheric Chemistry and Physics, 16, 6595–6607, https://doi.org/10.5194/acp-16-6595-2016, 2016.

Stull, R. B., ed.: An Introduction to Boundary Layer Meteorology, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-009-3027-8, 1988.

725 Terai, C. R. and Wood, R.: Aircraft Observations of Cold Pools under Marine Stratocumulus, Atmospheric Chemistry and Physics, 13, 9899–9914, https://doi.org/10.5194/acp-13-9899-2013, 2013.

Toll, V., Christensen, M., Gassó, S., and Bellouin, N.: Volcano and Ship Tracks Indicate Excessive Aerosol-Induced Cloud Water Increases in a Climate Model, Geophysical Research Letters, 44, 492–500, https://doi.org/10.1002/2017GL075280, 2017.

Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, Journal of the Atmospheric Sciences, 34, 1149–1152,
730 https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.

Varble, A.: Erroneous Attribution of Deep Convective Invigoration to Aerosol Concentration, Journal of the Atmospheric Sciences, 75, 1351–1368, https://doi.org/10.1175/JAS-D-17-0217.1, 2018.

Vogel, R., Nuijens, L., and Stevens, B.: The Role of Precipitation and Spatial Organization in the Response of Trade-Wind Clouds to Warming, Journal of Advances in Modeling Earth Systems, 8, 843–862, https://doi.org/10.1002/2015MS000568, 2016.

735 Walther, A. and Straka, W.: Algorithm Theoretical Basis Document For Daytime Cloud Optical and Microphysical Properties (DCOMP), NOAA NESDIS Center for Satellite Applications and Research, 2020.

Wang, H., Rasch, P. J., and Feingold, G.: Manipulating marine stratocumulus cloud amount and albedo: a process-modelling study of aerosol-cloud-precipitation interactions in response to injection of cloud condensation nuclei, Atmospheric Chemistry and Physics, 11, 4237–4249, https://doi.org/10.5194/acp-11-4237-2011, publisher: Copernicus GmbH, 2011.

740 Wang, M., Ghan, S., Liu, X., L'Ecuyer, T. S., Zhang, K., Morrison, H., Ovchinnikov, M., Easter, R., Marchand, R., Chand, D., Qian, Y., and Penner, J. E.: Constraining cloud lifetime effects of aerosols using A-Train satellite observations, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL052204, 2012.

Wang, S., Wang, Q., and Feingold, G.: Turbulence, Condensation, and Liquid Water Transport in Numerically Simulated Nonprecipitating Stratocumulus Clouds, Journal of the Atmospheric Sciences, 60, 262–278, https://doi.org/10.1175/1520-0469(2003)060<0262:TCALWT>2.0.CO;2, 2003.

Wood, R.: Cancellation of Aerosol Indirect Effects in Marine Stratocumulus through Cloud Thinning, Journal of the Atmospheric Sciences, 64, 2657–2669, https://doi.org/10.1175/JAS3942.1, 2007.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373–2423, https://doi.org/10.1175/MWR-D-11-00121.1, 2012.

Wood, R. and Bretherton, C. S.: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability, Journal of Climate, 19, 6425–6432, https://doi.org/10.1175/JCLI3988.1, 2006.

Wood, R. and Hartmann, D. L.: Spatial Variability of Liquid Water Path in Marine Low Cloud: The Importance of Mesoscale Cellular Convection, Journal of Climate, 19, 1748–1764, https://doi.org/10.1175/JCLI3702.1, 2006.

Wright, S.: Correlation and Causation, Journal of Agricultural Research, 20, 557–585, 1921.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) Modelling Framework of DWD and MPI-M: Description of the Non-Hydrostatic Dynamical Core, Quarterly Journal of the Royal Meteorological Society, 141, 563–579, https://doi.org/10.1002/qj.2378, 2015.

Zheng, Q. and Miller, M. A.: Summertime Marine Boundary Layer Cloud, Thermodynamic, and Drizzle Morphology over the Eastern North Atlantic: A Four-Year Study, Journal of Climate, 35, 4805–4825, https://doi.org/10.1175/JCLI-D-21-0568.1, 2022.

Zhu, Z., Kollias, P., Luke, E., and Yang, F.: New Insights on the Prevalence of Drizzle in Marine Stratocumulus Clouds Based on a Machine Learning Algorithm Applied to Radar Doppler Spectra, Atmospheric Chemistry and Physics, 22, 7405–7416, https://doi.org/10.5194/acp-22-7405-2022, 2022.