

Reviewer 1

The authors sincerely thank the reviewer for the positive and encouraging evaluation of our work. We appreciate the acknowledgment of the novelty and significance of our study. Point-by-point responses are given below.

More detail on the modeling approach is needed to better understand how methodological choices may influence the analysis. Specifically, how is Nd diagnosed from CCN? Is saturation adjustment applied in the cloud microphysics scheme? How is Nd treated during mixing and entrainment processes? What is the vertical resolution near the cloud top? Do updraft velocities differ between the simulations as a result of aerosol perturbations?

Ans: Thank you very much for pointing this out. We have now revised the manuscript and added additional details on the modelling approach in the data and methodology section. Point-by-point response and the corresponding modified part of the text is given below.

The ICON-LES uses the Seifert & Beheng (2006) two-moment microphysical scheme, which predicts the number and mass of multiple hydrometeor classes, including cloud droplets. In the two-moment microphysical scheme of Seifert & Beheng (2006), the grid-scale cloud droplet nucleation rate is estimated as a function of CCN concentration and vertical velocity supersaturation (Equation 7). Following Kogan and Martin (1994), the Seifert & Beheng (2006) scheme applies the standard saturation adjustment technique to treat condensational growth of cloud droplets.

In the specific ICON-LES, entrainment is not parameterised; it arises from resolved advection and subgrid turbulent mixing.

In the selected case, the cloud top height is at ~ 1700 – 2000 m, with an effective vertical resolution of ~ 50 – 80 m.

The specific ICON-LES simulations used the same boundary conditions (May 2, 2013) for both 2013 and 1985 CCN simulations. Even with the identical large-scale forcing, the vertical velocity w would respond if the aerosol perturbation changes cloud-top cooling/evaporation and precipitation sufficiently. So, in this particular simulation, some dynamical differences (in vertical velocity) are expected in at least parts of the domain/time.

Line 35–36: “The numerous small droplets also lead to droplet sedimentation.” Isn’t the opposite expected, with smaller droplets having reduced sedimentation velocities?

Ans: Thank you very much for pointing this out. We have now revised it as:

Aerosol perturbations increase Nd, yielding smaller droplets that evaporate more efficiently during mixing and reduce droplet sedimentation, thereby redistributing liquid water toward the inversion and potentially strengthening cloud-top longwave radiative cooling. As radiative cooling occurs within a relatively thin layer near the cloud top, it also promotes the entrainment of warm, dry free-tropospheric air, further enhancing evaporation and potentially reducing LWP - a negative LWP adjustment (Williams and Igel, 2021; Wood, 2012; Bretherton et al., 2007; Ackerman et al., 2004; Wang et al., 2003).

Line 36: “Enhance radiative cooling at the cloud top.” This statement could be more precise, as radiative cooling occurs within a relatively thin layer near the cloud top.

Ans: Many thanks. Revised as suggested:

As radiative cooling occurs within a relatively thin layer near the cloud top, it also promotes the entrainment of warm, dry free-tropospheric air, further enhancing evaporation and potentially reducing LWP - a negative LWP adjustment (Williams and Igel, 2021; Wood, 2012; Bretherton et al., 2007; Ackerman et al., 2004; Wang et al., 2003).

Line 37: “Further, the entrainment of warm, dry air into the cloud leads to evaporation of the smaller droplets, resulting in decreases in LWP or negative LWP adjustment.” Entrainment leads to LWP reduction even if all droplets lose mass homogeneously, not only through preferential evaporation of smaller droplets.

Ans: Thank you very much for pointing this out. It is now revised as:

Notably, entrainment can reduce LWP even when droplet mass decreases approximately homogeneously, not just due to the preferential evaporation of smaller droplets.

Line 49: “The N_d –LWP relationship is non-linear, and the co-variability between LWP and N_d primarily drives it.” While this is true, is it the co-variability between aerosol and meteorological conditions that fundamentally drives both LWP and N_d ?

Ans: Many thanks for pointing this out, we have now revised it as:

The LWP adjustment is also regime-dependent (Glassmeier et al., 2021). In marine stratocumulus clouds, the N_d –LWP relationship is non-linear, and the apparent coupling largely reflects co-variability between aerosol loading and meteorological conditions, which fundamentally drives variations in both N_d and LWP (Goren et al., 2024).

Line 53: Please clarify that the modeling referred to here concerns global, low-resolution models.

Ans: Many thanks for pointing this out. Here, modelling refers to high-resolution simulations, which have now been updated.

However, high-resolution modelling evidence has also shown negative LWP adjustment, in which the altered N_d leads to enhancement of entrainment mixing, thereby reducing LWP (Glassmeier et al., 2021; Ackerman et al., 2004).

Line 95: Could the averaging of data to a coarser grid resolution affect the results? The original LES output likely resolves cloud-scale structures, but aggregation to 1.2 km places the analysis in the “gray zone” for cloud dynamical processes. Was the analysis repeated over smaller spatial scales to confirm that averaging does not affect?

Ans: Thank you very much for pointing this out.

Yes, aggregating LES output to a coarser grid can affect diagnosed relationships by smoothing sub-kilometre variability and mixing cloudy with clear air within a grid cell (“grey-zone” effects). To minimise these impacts, our analysis is regime-conditioned and relies on cloud-only / cloud-core sampling (e.g., thresholds on LWC/ N_d and cloud-top detection), so the statistics are not dominated by diluted edge pixels. We also focus on distribution-based metrics (e.g., conditional means and joint relationships) that are less sensitive to small-scale positional noise than pointwise comparisons. Moreover, the 1.2 km coarse-grained HD(CP)² products have been used extensively in prior work, which indicates that the main aerosol–cloud sensitivities remain qualitatively robust after aggregation. In the HD(CP)² framework, however, only the coarsened products were retained due to the very large data volume of the native LES output, so we could not repeat the analysis at finer aggregation scales for a formal scale-sensitivity test. As a limitation, we acknowledge that some magnitude changes could occur at finer scales.

The amended text:

We have used the coarse-gridded data with a resolution of 1.2 km, a standard reduced-volume product that has been used in previous studies and evaluations (Costa-Surós et al. 2020, Dipu et al. 2022). Our results rely on regime-conditioned, cloud-only statistics that mitigate grey-zone smoothing. While coarse-gridding may influence quantitative values, the qualitative sensitivities remain robust.

Line 99–100: Why is this criterion used instead of one based on liquid water content? Such a layer may have very low LWC and therefore may not represent the actual cloud top, but rather a layer above it.

Ans: Thank you very much for pointing this out as well.

For filtering the liquid cloudy layer, we considered $LWC > 1 \times 10^{-8} \text{ kg kg}^{-1}$, in addition to cloud-top temperature $> 268 \text{ K}$. Further, for the analysis, the cloud-top Nd is filtered for cloud fractions equal to 1 (at the 1.2 km scale) and cloud optical thicknesses greater than 2.

Line 102–104: This assumption is unlikely to hold over a one-day simulation. However, I did not find an analysis in the paper that could be affected.

Ans: We agree that over a one-day LES run, individual cloud features will not remain co-located between the control and aerosol-perturbed simulations. However, our analysis does not rely on tracking identical cloud objects or pointwise cloud resemblance; comparisons between the two simulations at the same grid locations are used only to ensure identical large-scale forcing and to reduce broad spatial-gradient effects. The reported results are based on regime-conditioned, distribution-based statistics rather than instantaneous gridpoint-by-gridpoint differences, so cloud displacement mainly adds sampling noise rather than systematic bias.

This is already mentioned in the data methodology section.

Equation (2): Is the last term correct, given that vertical pressure velocity is used?

Ans: Thank you for the correction.

Equation 2 has been corrected now.

$$Q_2 = -L \frac{\partial q}{\partial t} - LV \cdot \nabla q - L\omega \frac{\partial q}{\partial p}$$

Equation (3): This equation appears to describe the mean volume diameter. If so, the subsequent calculation of the cloud dilution factor does not seem to be meaningful. Also the units appear incorrect.

Ans: Thank you very much.

In the revised manuscript, we have updated the cloud dilution ratio (equation 3) with adiabaticity (f_{ad}), which is defined as the ratio of LWP to adiabatic LWP (LWP_{ad}),

$$f_{ad} = \frac{LWP}{LWP_{ad}}$$

Here, LWP_{ad} is computed following the common adiabatic parcel approximation,

$$LWP_{ad} = \frac{1}{2} \gamma_{ad} H^2,$$

where γ_{ad} is the vertical gradient of adiabatic liquid water content and H is cloud depth.

Line 122: “Binned with a bin size of 1000.” I assume this refers to the number of bins rather than the bin width; please clarify.

Ans: Thank you for mentioning this.

Yes, it refers to the number of bins rather than the bin width, and it is updated in the revised manuscript.

For joint histogram analysis, the variables are binned with a bin size (number of bins) of 1000.

Line 143: The results are qualitatively consistent but far from being quantitatively consistent.

Ans: Yes, the results are qualitatively consistent, but far from being quantitatively consistent, because the effects, magnitudes and the Nd threshold of the transition are expected to vary with

background meteorology, sampling scale and aerosol loading.

The revised section is given below.

For both simulations, the $\overline{\text{LWP}}$ shows a non-linear relationship consistent with prior studies (Gryspeerdt et al., 2019; Dipu et al., 2022), but the magnitude and the N_d threshold of the transition differ due to differences in meteorological regime, sampling/aggregation scale, and aerosol loading.

Line 148: “Which sustained droplet activation.” The term sustained does not seem appropriate in this context.

Ans: Thank you for your feedback. Revised as suggested.

This shift toward higher N_d reflects higher CCN concentrations in 1985, which enhanced droplet activation during cloud formation, increasing N_d and delaying LWP depletion.

Line 164: “Initially, both simulations exhibit positive LWP sensitivity.” What is the reason for this behavior? Is it related to assumptions about the initial conditions? Are the initial thermodynamic profiles identical, differing only in aerosol loading?

Ans: Initially, both simulations show positive LWP sensitivity because the aerosol increase elevates N_d , reduces droplet size, and suppresses/delays warm-rain formation, allowing LWP to accumulate (Albrecht effect).

The simulations are initialised with identical thermodynamic profiles and forcing. They differ only in aerosol/CCN loading (droplet activation), so the early positive response reflects rapid microphysical adjustment rather than differing initial conditions.

The following text is included in the revised manuscript.

Initially, both simulations exhibit positive LWP sensitivity because CCN acts to elevate N_d , reduce droplet size, and suppress or delay warm-rain formation, allowing LWP to accumulate (Albrecht, 1989). Both simulations use identical initial conditions and forcing and differ only in the prescribed CCN, so this positive LWP sensitivity reflects rapid microphysical adjustment rather than differences in initialisation.

Line 175: How are the different critical values determined? Are they derived across different spatial domains or over different time intervals?

Ans: Thank you very much for pointing this out.

From the temporal evolution of LWP sensitivity, the critical N_d and Reff values are estimated. The critical values are specified only for time steps at which the domain-wide bin mean LWP exhibits a non-linear dependence on N_d , with a clear maximum (i.e., a transition from positive to negative LWP sensitivity). From the time the temporal evolution of the LWP sensitivity (Fig. 4), if a non-linear relationship is present, the mean N_d/Reff of the bin at which LWP reaches its maximum (sign change in sensitivity) is taken as the “critical” value for that time step. If no clear non-linear relationship exists, no critical values are assigned. They are derived across the same spatial domain and time interval.

The above text has been amended in the revised manuscript.

Line 188: Related to the earlier comment on the definition of cloud top, how representative is N_d at cloud top of N_d throughout the cloud volume?

Ans: Thank you for this comment.

Cloud-top N_d is not expected to be identical to N_d everywhere in the cloud, because N_d can

vary vertically due to activation near cloud base, entrainment-driven dilution near cloud top, and microphysical processes. In this study, N_d is primarily intended as a consistent metric for comparing aerosol perturbations and linking to cloud-top processes (radiative cooling and entrainment) that control LWP adjustments. Therefore, the cloud top N_d is not interpreted as a volume-mean value but as an indicator of the droplet population near the region most relevant for the LWP adjustment mechanisms.

The above text has been amended in the revised manuscript (Data and methodology section).

Line 192: This statement holds only if saturation adjustment is not used and condensation is explicitly calculated.

Ans: Thank you for this comment.

Yes, it is correct that in the saturation adjustment scheme, supersaturation is not explicitly predicted, and condensation is diagnosed to restore saturation within each time step. However, this does not mean evaporation is absent. Mixing or entrainment that produces subsaturation still leads to evaporation and cooling, while ascent and condensation result in heating. Therefore, our interpretation does not rely on explicitly resolved supersaturation but on the microphysical-dynamical balance between condensation and evaporation tendencies, as represented by the saturation-adjustment closure.

So, we have revised the text to:

As the cloud develops, collision-coalescence shifts the droplet spectrum toward larger Re_{eff} and promotes precipitation processing. As a result, the relative contribution of evaporative cooling associated with mixing/entrainment decreases, and Q_1 becomes less negative with increasing N_d .

Line 195 onward: The effects of radiative cooling and heating are not discussed, despite being speculated as key mechanisms driving the negative N_d -LWP relationship. Why are these processes not analyzed in more detail?

Ans: Thank you very much for pointing this out.

An explanation is already included at the end of the discussion sections, which is also provided below.

It should be emphasised, however, that evaporation and advection are not the only processes associated with negative LWP sensitivity. While they appear to be the dominant mechanisms in the above simulations, additional processes such as cloud-top radiative cooling, droplet sedimentation, and turbulence entrainment feedbacks may also play important roles in driving cloud depletion at high N_d . Future studies should aim to disentangle and quantify the relative contributions of these pathways, for example, by combining targeted LES experiments with process-level diagnostics and Lagrangian cloud tracking.

Line 198: When discussing high-resolution simulations and processes occurring at the cloud edge, is advection the appropriate term, or would mixing be more accurate? More generally, modeling and also the analysis details may significantly affect the results, as numerical diffusion is unavoidable when Lagrangian processes are simulated in Eulerian framework. This may influence the diagnosed apparent heating and moisture sink terms.

Ans: Thank you for this comment.

We agree that in high-resolution simulations and processes occurring at the cloud edge, mixing or entrainment is a more precise term. Rather than pure mean advection alone, it is the combined mixing/entrainment that dilutes cloudy air with warmer, drier environmental air, leading to droplet evaporation. We therefore revise the wording to “entrainment/mixing of warm, dry air” instead of “warm-air advection”.

Yes, we acknowledge that numerical diffusion can affect local evaporation rates and, in turn, the diagnosed apparent heating/moisture tendencies. However, our analysis is based on domain and regime-conditioned statistics, not pointwise tendencies/comparison at a single cloud edge, which reduces sensitivity to numerical diffusion. We nevertheless acknowledge that numerical diffusion may influence the magnitudes of the diagnosed apparent heating and moisture sink terms near cloud edges and add this as a limitation. The following text is added at the end of the discussion section.

We also note that numerical diffusion can affect the magnitudes of the diagnosed apparent heating and moisture sink near cloud edges, which is a limitation of the present analysis and should be considered when interpreting the results.

Figure 1: Mean values could be added to both panels.

Ans: Thank you for this comment. Revised as suggested.

Figure 2: “the 1985 simulation uses the pre-industrial (1985) CCN”. This is a new definition for pre-industrial era.

Ans: Thank you very much for pointing this out. Revised as suggested.

Figure 6: left axis is cut of from the figure

Ans: Thank you very much for pointing this out. Revised as suggested.

Figure 7: ylabel “aparent”

Ans: Thank you very much for pointing this out. Revised as suggested.

Figure A1: Check the units.

Ans: Thank you very much for pointing this out. Revised as suggested.

Reviewer 2

The authors sincerely thank the reviewer for the positive and encouraging evaluation of our work. We appreciate the acknowledgment of the novelty and significance of our study. Point-by-point responses are given below.

General Comment

The research topic is very relevant and interesting to the scientific community. However, the paper has several concerns in data methods and lacking proper justification/reference. More clarification is also required on the simulation setup, especially how aerosol spatial distribution is initialized and perturbation is imposed. Another concern arises from the fact that in a real LES simulation where environmental conditions are not identical at all points (as the domain is very large), whether the LWP sensitivity of aerosol can be achieved?. There can be other environmental factors also controlling this sensitivity (Nd – LWP relation). In a sensitivity experiment, in general we keep all other parameters such as environmental conditions same for all the forcing runs to ensure impact of aerosol forcing only. This is more suitable for an ideal LES run.

Considering all these aspects, I feel that the paper is not suitable for publication in its present form and a significant revision is required.

Ans: In this paper, we have used the available large-domain ICON-LES simulation from the HD(CP)² project. In the ICON-LES simulations, the aerosol perturbation is imposed only through prescribed CCN fields, while the meteorology is kept identical across the paired simulations. Following Costa-Surós et al. (2020), the CCN concentrations are provided as temporally and spatially varying offline fields (time-varying 4D distributions) generated with the regional coupled chemistry–aerosol model system COSMO-MUSCAT, and then read into ICON-LEM. The control run uses a CCN representative of 2 May 2013, while the perturbed run uses a CCN representative of peak European pollution around 1985.

We also agree that in a large LES domain, environmental heterogeneity and aerosol–meteorology co-variability can influence the apparent Nd–LWP relationship. However, the experimental design compares the control and the perturbed simulation with identical meteorological forcing, differing only in the imposed CCN fields. Costa-Surós et al. (2020) showed that, even in a large LES domain with substantial spatial variability, the aerosol signal can be quantified using domain-wide, regime-conditioned statistics, rather than pointwise cloud matching. In our analysis, we have used a similar methodology in which regime-conditioned statistics and distribution-based metrics mitigate confounding meteorological factors while retaining the aerosol signals imposed by the perturbation.

Specific comments:

L12-14: “stronger cloud-top heating and moisture sinks are simulated during negative LWP sensitivity phases, particularly for high Nd in 2013, consistent with enhanced evaporation and entrainment.”

Seems opposite statement as enhanced entrainment and evaporation leads to stronger cooling by latent heat absorption.

Ans: Thank you very much for pointing this out.

We agree that the wording was unclear. Yes, entrainment-driven evaporation is associated with latent cooling, not heating. However, in our diagnostics, Q1 is an apparent heating that includes contributions from advection and vertical motion in addition to microphysical latent heating/cooling. Therefore, during negative LWP sensitivity at high Nd (especially in the 2013 case), we mainly interpret it through the apparent moisture sink (Q2). The positive Q2 at high Nd indicates dry-air entrainment and mixing, along with associated evaporation/dilution,

consistent with the concurrent decrease in LWC/LWP. Any positive (or less negative) Q1 in this regime is attributed to warming terms (e.g., warm-air advection/subsidence) offsetting evaporative cooling, rather than evaporation causing heating.

The revised text is given below.

Specifically, during negative LWP sensitivity phases, stronger cloud-top drying (moisture sinks) is simulated, particularly at high Nd in 2013, consistent with enhanced entrainment/mixing and evaporation-driven cloud dilution.

L67-70: In this study, we aim to investigate the LWP sensitivity in a non-precipitating continental cloud regime using high-resolution LES model simulation in numerical weather prediction mode, with initial and boundary conditions for a real weather situation, and with interactive land surface (Heinze et al., 2017; Costa-Surós et al., 2020).

Does it mean that two aerosol runs have different initial and boundary conditions?.

Ans: Again, thank you very much for pointing this out.

No. Both simulations use the same initial and boundary conditions for the real weather case (including the interactive land surface and the same large-scale/lateral forcing). The paired simulations differ only in the prescribed CCN fields used for droplet activation (aerosol perturbation), while the meteorological setup is otherwise identical.

Amended manuscript text:

In this study, we investigate LWP sensitivity in a non-precipitating continental cloud regime using high-resolution LES simulations in numerical weather prediction mode, with initial and boundary conditions from a real weather situation and an interactive land surface (Heinze et al., 2017; Costa-Surós et al., 2020). Aerosol–cloud interaction effects are quantified using control and aerosol-perturbed simulations that differ only in the prescribed CCN fields for droplet activation, while the meteorology is kept identical across the simulations. The following methodology section describes the model setup and aerosol perturbation.

L85-87: The CCN concentrations in the model are prescribed as a spatially and temporally varying distribution. The control simulation uses CCN concentrations as estimated for 02 May 2013 (Costa-Surós et al., 2020)., and for the perturbed simulation, CCN concentrations valid for the year approximately 1985 were selected, in which the pollution level in Europe was at its peak (Smith et al., 2011).

How did you consider aerosol perturbation for 1985?. Is it based on observations?

Ans: Thank you for this comment.

The 1985 CCN fields are not taken directly from observations; instead, they are obtained by scaling the 2013 CCN fields using species-dependent scaling factors derived from emission ratios following Genz et al. (2019). It is derived from emissions-based scaling rather than direct observations.

The following text is now amended in the manuscript.

The 2013 CCN concentrations are generated from 2013 emissions using a regional coupled model system (Wolke et al., 2004, 2012). The 1985 CCN concentrations are obtained by scaling the 2013 CCN concentrations with species-dependent factors derived from emission ratios following Genz et al. (2019). A detailed description is provided in Costa-Surós et al. (2020).

L88 & L89 : Experiment date mismatch.

Ans: Thanks. Revised as suggested.

L95: “we have used the coarse-gridded data with a resolution of 1.2 km”

What is the impact of coarse-gridding in the analysis results?

Ans: Thank you for this comment.

Yes, aggregating LES output to a coarser grid can affect diagnosed relationships by smoothing sub-kilometre variability. To minimise these impacts, our analysis is regime-conditioned and relies on cloud-only / cloud-core sampling (e.g., thresholds on LWC/Nd and cloud-top detection), so the statistics are not dominated by diluted edge pixels. We also focus on distribution-based metrics (e.g., conditional means and joint relationships) that are less sensitive to small-scale positional noise than pointwise comparisons. Moreover, the 1.2 km coarse-grained HD(CP)2 products have been used extensively in prior work, which indicates that the main aerosol-cloud sensitivities remain qualitatively robust after aggregation. In the HD(CP)2 framework, however, only the coarsened products were retained due to the very large data volume of the native LES output, so we could not repeat the analysis at finer aggregation scales for a formal scale-sensitivity test. As a limitation, we acknowledge that some magnitude changes could occur at finer scales.

The amended text:

We have used the coarse-gridded data with a resolution of 1.2 km, a standard reduced-volume product that has been used in previous studies and evaluations (Costa-Surós et al. 2020, Dipu et al. 2022). Our results rely on regime-conditioned, cloud-only statistics that mitigate grey-zone smoothing. While coarse-gridding may influence quantitative values, the qualitative sensitivities remain robust.

L99: The cloud top is defined as the topmost level with $N_d > 2 \text{ cm}^{-3}$, which is further filtered for cloud fractions equal to 1 (at the 1.2 km scale) and cloud optical thicknesses greater than 2. Please give justification or references to support this cloud top detection method. I think $N_d > 2 \text{ cm}^{-3}$ is very low value to consider as cloud. Several studies reported cloud detection threshold by $LWC > 0.001 \text{ g/m}^3$ or $N_d > 10 \text{ cm}^{-3}$. In this case, please check if $N_d > 2 \text{ cm}^{-3}$ satisfies to the LWC criteria.

Ans: Thank you for this comment.

We agree that $N_d > 2 \text{ cm}^{-3}$ alone could be too low, so the cloud top is not defined from N_d only. In the revised manuscript, we clarify that the cloud top is identified as the highest model level satisfying liquid-cloud conditions ($LWC > 1 \times 10^{-8} \text{ kg kg}^{-1}$ and $T > 268 \text{ K}$), and that $N_d > 2 \text{ cm}^{-3}$ is used only as an additional indicator to avoid selecting levels with vanishing droplet number. For the analysis, we further restrict to fully cloudy coarse-grid columns (cloud fraction = 1 at 1.2 km) and optically thick clouds ($COT > 2$), which reduces cloud-edge and cloud contamination.

The revised text is given below.

The analysis is restricted to single-layered liquid clouds by excluding the clouds with a cloud-top temperature below 268 K. Cloud top is diagnosed as the uppermost model level with liquid cloud water present (liquid water content, $LWC > 1 \times 10^{-8} \text{ kg kg}^{-1}$). We additionally constrained $N_d > 2 \text{ cm}^{-3}$ and restricted the analysis to overcast and optically detectable clouds (cloud fraction = 1 at 1.2 km cloud optical thicknesses, $COT > 2$) to minimise cloud-edge contamination.

L115: The cloud dilution ratio, which serves as a proxy for entrainment, is defined as the ratio of the cloud effective radius to the adiabatic radius.

Provide justification or references to support this statement. As cloud dilution is generally represented by the ratio of LWC to adiabatic LWC.

Now, for this case if $LWC/LWC_{ad} = R_{eff}/R_{effad}$

This simply means that mixing in the clouds is extremely inhomogeneous where evaporation is rapidly adjusted to constant R_{eff} . Further analysis to support this process is not provided by using time scale analysis. Fig. 3 clearly shows that mixing is not extremely inhomogeneous as indicated by rapidly decreasing R_{eff} with decreasing N_d .

Ans: Thank you for this comment.

We agree that “cloud dilution” (as a proxy for entrainment/mixing) is most consistently quantified using thermodynamic dilution of condensate, i.e. the ratio of in-cloud liquid water to its adiabatic reference, rather than using a microphysical size ratio such as $R_{eff}/R_{eff,ad}$. In particular, R_{eff} responds not only to dilution/entrainment but also to collision-coalescence, sedimentation, activation/CCN changes, and precipitation removal, so it is not a unique proxy for entrainment. Therefore, in the revised manuscript, we removed the definition of cloud dilution based on $R_{eff}/R_{eff,ad}$ and replaced Eq. (3) with a standard adiabaticity (sub-adiabatic fraction) based on the liquid water path:

$$f_{ad} = \frac{LWP}{LWP_{ad}}$$

Here, LWP_{ad} is computed following the common adiabatic parcel approximation,

$$LWP_{ad} = \frac{1}{2} \gamma_{ad} H^2,$$

where γ_{ad} is the vertical gradient of adiabatic liquid water content and H is cloud depth.

L119: LWC unit should be g/m³

Ans: Thanks for pointing this out. Revised as suggested.

L130: “However, the LWP distribution shows relatively small shifts towards higher LWP in the 1985 simulation compared to 2013 (Fig. 1b)”

From the figure it seems an insignificant change in the LWP PDFs. This result clearly indicates that aerosol perturbation has no clear impact (positive) on LWP sensitivity.

The analysis related to Nd – LWP joint probability distribution is valid for both low and high CCN cases and produces a similar relation with slight shift. This relation does not produce a LWP sensitivity to aerosol perturbation as similar relation is valid for low and high CCN cases. This relation arises due to spatial variability in aerosol and other environmental conditions. Therefore, it is difficult to understand how aerosol alone impacts LWP in this case.

Ans: Thank you for this comment.

We agree that Fig. 1b shows only a small shift in the LWP PDF in the 1985 simulation compared to 2013, implying that the domain-mean LWP response is modest in this case. This does not imply that the aerosol perturbation is insignificant, because the domain includes multiple cloud states, competing sink processes, and dynamical adjustments can offset microphysical LWP increases, leaving only a small net shift in the bulk distribution. The aerosol signal is most evident in the microphysical state, as the higher prescribed CCN in 1985 causes a pronounced rightward shift in droplet number, with Nd increasing from 100 cm⁻³ (2013) to 300 cm⁻³ (1985). We therefore quantify aerosol effects primarily using regime conditioned Nd–LWP statistics and their temporal evolution, rather than relying on the LWP PDF. Moreover, the two simulations have identical meteorological forcing and initial conditions, and differ only in the prescribed CCN, so the diagnosed differences in Nd and conditional Nd–LWP behaviour reflect the aerosol perturbation rather than differing environmental conditions.

The revised text is given below.

The contrasting response of the Nd and LWP to the aerosol perturbation suggests that higher droplet activation alone does not directly translate into proportional increases in bulk water content, highlighting the importance of compensating microphysical and dynamical processes (e.g., entrainment/mixing and cloud dilution) that can offset LWP increases.

L145: It is not clear how the transition point of regime shift (positive to negative) is moved to higher side in 1985 case. It is only the aerosol change or the temperature change of global warming is also associated.

Ans: Thank you for this comment.

The shift of the transition point to higher Nd in 1985 is not due to global warming. The two

simulations have identical meteorological forcing and initial conditions; only the prescribed CCN field differs. Therefore, the shift in the transition point reflects the aerosol-driven shift in the droplet activation. The transition from positive to negative LWP sensitivity occurs when the cloud dilution exceeds the precipitation suppression. In the 1985 case, the aerosol perturbation shifts cloud droplets to higher Nd, so the transition point of the negative-sensitivity is also shifted to higher Nd, indicating an aerosol-driven activation shift rather than a temperature-driven climate change effect.

The following text is added for clarification.

Since the two simulations have identical meteorological forcing and initial conditions, and differ only in CCN, the shift in the transition point is attributed to the aerosol perturbation rather than thermodynamic warming.

L151: “The cloud dilution ratio, which is defined as the ratio of R_{eff} to the adiabatic Rad serves as a proxy for entrainment mixing. Values close to 1 indicate adiabatic clouds, while lower values suggest dilution.”

This statement is not correct. To detect adiabatic clouds, liquid water dilution ratio (LWC/LWC_{ad}) is best suited. The dilution ratio shown presented in this study depends on mixing type (homogeneous vs. inhomogeneous). For example, in case of extreme inhomogeneous mixing, this ratio (R_{eff}/R_{ad}) will be close to 1, even if strong dilution mixing occurred.

Ans: Thanks for pointing this out.

In the revised manuscript, we have updated the cloud dilution ratio (equation 3) with adiabaticity, which is defined as the ratio of LWP to adiabatic LWP (LWP_{ad})

$$f_{ad} = \frac{LWP}{LWP_{ad}}$$

Where $LWP_{ad} = \frac{1}{2} \gamma_{ad} H^2$, where γ_{ad} is the vertical gradient of adiabatic LWC, and H is the cloud depth

L171: “Notably, the 2013 simulation exhibits a stronger negative slope than 1985, highlighting more rapid LWP depletion”. . . . If I understood correctly, stronger entrainment–evaporation feedback should occur in high aerosol loading case. But here, the response is opposite.

Ans: Thank you for this comment.

Enhanced entrainment–evaporation feedback under higher aerosol loading is not expected to be monotonic. Also, the magnitude of negative LWP sensitivity depends strongly on the thermodynamic state and on whether the cloud regime is dilution-dominated. In the 1985 simulation, the LWP transition point is shifted to higher Nd and delays the onset of the depletion regime. While the diagnosed negative Nd–LWP slope is smaller because the net drying/depletion is weaker on average (consistent with greater moisture retention and later occurrence of positive Q2; explained later in the text). Thus, the 2013 simulation exhibits a steeper negative slope because the cloud regime enters drying/entrainment-driven depletion more readily, leading to more rapid LWP loss.

The following is the revised text with clarifications.

Notably, the 2013 simulation exhibits a steeper negative slope than the 1985 simulation, indicating more rapid LWP depletion. Although the transition to negative LWP sensitivity shifts to higher Nd in 1985, the magnitude of the negative slope is smaller than in 2013. The steeper negative slope in 2013 suggests that the cloud field enters drying/entrainment-driven depletion more readily, leading to more rapid LWP depletion. In contrast, the 1985 perturbation primarily shifts the transition to higher Nd and delays the onset of the depletion regime. Importantly, higher aerosol loading does not necessarily imply a more negative Nd–LWP slope. The strength

of the negative LWP sensitivity is state-dependent and depends on whether the clouds are in a dilution-dominated (entrainment–evaporation) state.

L245: “The analysis clearly indicates that aerosol perturbations and aerosol levels have a significant impact on the LWP sensitivity and associated processes.” This is not really reflected in the presented results. The aerosol perturbation impact on LWP sensitivity is not significant, However, more microphysical and thermodynamical response is found.

Ans: Thank you very much for pointing this out.

We agree that the domain-wide LWP response to the aerosol perturbation is modest in this case and that the most pronounced signals appear in the microphysical and thermodynamic diagnostics. We therefore revise the text, and it is given below.

The analysis shows that the aerosol perturbation induces a clear microphysical and thermodynamic response and modulates the regime-conditioned Nd–LWP sensitivity, while the change in the bulk LWP distribution remains modest.

L247: “The observed positive tendency in Q1 for negative LWP sensitivity could be due to warm entrainment, which leads to evaporation of cloud droplets”. Entrainment is generally associated with downward motion in cloud top. This should bring colder air inside the cloud?

Ans: Again many thanks pointing this out.

Entrainment at the cloud top does not imply that cold air is brought into the cloud. In stratocumulus and shallow convective clouds, entrainment primarily involves the mixing of free-tropospheric air above the inversion into the cloudy layer. That air is typically warmer and drier than the cloudy boundary-layer air. The entrained air is incorporated mainly through turbulent mixing, and the resulting mixture becomes subsaturated, which promotes cloud droplet evaporation and LWP depletion. This can be further explained by the apparent heating Q1 and the moisture sink Q2. The apparent heating Q1 includes contributions from advection and vertical motion in addition to microphysical latent heating/cooling. Therefore, during negative LWP sensitivity at high Nd (especially in the 2013 case), we mainly interpret it through the apparent moisture sink (Q2). The positive Q2 at high Nd indicates dry-air entrainment and mixing, along with associated evaporation/dilution, consistent with the concurrent decrease in LWC/LWP. Any positive (or less negative) Q1 in this regime should be attributed to warming terms (e.g., warm-air advection/subsidence) offsetting evaporative cooling, rather than evaporation causing heating.

The following text is added for clarification.

Here, entrainment refers to turbulent mixing of relatively warm, dry free-tropospheric air across the inversion into the cloud-top layer, which promotes subsaturation and evaporation. Since Q₁ is an apparent heating residual that includes advection and vertical motion, its sign reflects the net balance of warming and evaporative cooling rather than evaporation alone.