Response to reviews of “Cloud water adjustments to aerosol perturbations are buffered by solar heating in non-precipitating marine stratocumuli” by J. Zhang et al.

We would like to thank the editor and the two anonymous reviewers for their insightful feedback and constructive comments and suggestions on our manuscript, which helped us improve the original manuscript.

Specific responses to each comment are contained below, with the reviewers’ comments provided in **blue** and our responses in **black**. Changes to the manuscript made in response to the reviewer are provided in ***red italics***. We have also made unsolicited changes to the manuscript to further polish the writing.

Reviewer 1

Summary: Zhang et al. describe a novel conditional Monte Carlo subsampling approach (cMC) to investigate the role of solar heating on marine stratocumulus cloud evolution. This approach allows them to artificially inflate the ensemble of large eddy simulations. In particular they investigate how solar heating changes the relationship between cloud droplet number concentration (Nd) and cloud water path (or liquid water path, LWP, in this case). They find that the Nd-LWP relation has a negative trend during the night time (in the absence of solar heating) and becomes less negative during the day (in the presence of solar heating), converging to a value around -0.2 in the late afternoon. This so-called "buffering effect" whereby the Nd-LWP relation is buffered back towards zero by the solar heating is attributed to the strong dependence of cloud absorbed SW radiation and cloud LWP; or simply, that thicker clouds absorb more strongly and thus thin at a faster rate than thinner clouds. They discuss the implications of these results for the time-dependent efficacy of aerosol injection, for example in the case of climate intervention via marine cloud brightening (MCB).

General comments:

Some limitations of this study, especially in it's relevance for MCB, include:

1. The simplistic and unrealistic assumption of the size and composition of the aerosol particles (ammonium sulfate, lognormally distributed with mean radius of 100nm). A more realistic MCB experiment would include seeding from larger, more hygroscopic sea salt spray.

2. The fixed (across the ensemble) prescribed SST and large-scale divergence which under-samples the relevant dynamical space these clouds occur in.

3. Of course, the mentioned restriction to the non-precipitating regime.

We thank the reviewer for raising these points and we agree that these are indeed the limitations on the implications for MCB based on this study. We covered the third point in the original manuscript and now added discussions around the other two points raised by the reviewer to the revised manuscript.

Overall, I think this paper is very through, well-written, and will be of interest to the aerosol-cloud interactions community. The cMC approach is also quite interesting and may be of broader interest outside the aerosol-cloud interactions community.

We thank the reviewer for the encouraging words and insightful comments! A point-by-point response is provided below.

Specific Comments:

- cMC sampling:

  - L144: Do you impose any threshold on the correlation coefficient for the regression when you build these subgroups, or just the slope?

No, we did not impose any threshold on the correlation coefficient for the cMC sampling, and this is done to ensure a pratical sampling efficiency. We actually tried imposing thresholds for correlation coefficient, and even for |r|>0.3, it takes forever to obtain 50 25-member random samplings that satisfy LWP-N_d slope conditions. This is because of the ~zero slope constraint we imposed between the 3 meteorological factors and N_d, as conditions for the cMC sampling. Fig. R1 below shows the distribution of r(ln(LWP), ln(N_d)) at the sampling time (i.e., sunrise), which indicates that the regression slope between ln(LWP) to ln(N_d) that we impose at sunrise is basically dictated by the correlation between them.

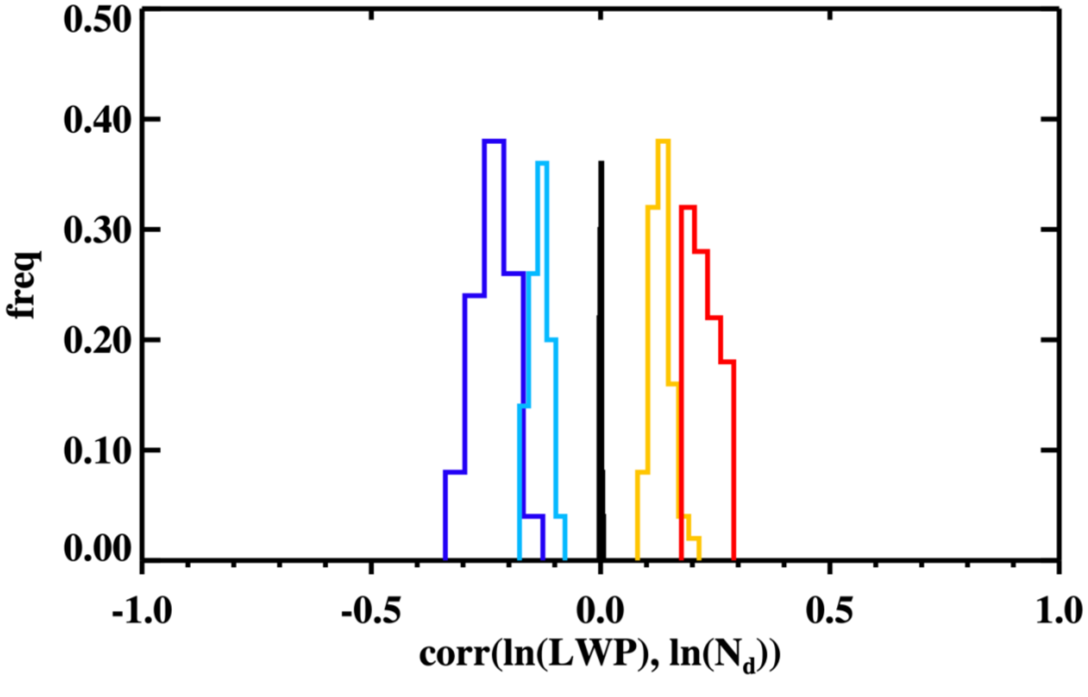

Figure R 1. Correlation distribution for 5 LWP-N_d slope groups at sunrise.

- Fig 1. I'm curious if it's possible to give an indication of the strength of the correlation in this figure. How robust is the slope over time? Is the r-value similar across these sub-ensembles? across time? Is the r-value always fairly large? If not, what does that indicate? And can that be shown in the plot? Maybe when r < 0.5 (or some other value state) you could make the lines more transparent?

We thank the reviewer for raising this critical point, as mentioned in the response to the previous comment, we did not impose any threshold on the correlation coefficient when performing cMC sampling. This means that the evolution in correlation coefficient across sub-ensembles closely tracks that of the regression slope (see Fig. R2 below). In other words, the regression slopes (S) that we show in this study are predominately controlled by the correlation between LWP and N_d, given $S(A, B) = corr(A, B) \times \frac{\sigma(B)}{\sigma(A)}$. This indicates that our cMC approach cannot (or, isn't designed to, given practical sampling efficiency and the non-linear nature between LWP and N_d) select a narrow, linear band of points in ln(LWP)-ln(N_d) space (i.e., with high r-value), but rather relies on the randomness of sub-sampling a large ensemble of simulations with given correlations between LWP and N_d to infer the relationship between LWP and N_d. We don't think this dominating role of the r-value in depicting the regression slope will affect our conclusion, given the relatively large sub-sample size (25 simulations per sub-ensemble), just that the points appear scattered, rather than as a narrow band in the ln(LWP)-ln(N_d) space.

In light of the reviewer's comment #1 & 2, we add Figure R2 to the supplemental material and mention it in the main text of the revised manuscript, as "*The diurnal evolution in the Nd–LWP slope (and correlation coefficient) of the five subgroups is shown in Figure 1 (and Fig. S1) …*" We also revised the statements around L144 to clarify how the correlation coefficient is handled, which now reads "*In order to maintain practical sampling efficiency of the cMC approach while approximating desired regression slopes, we impose arbitrary bounding values (or thresholds) around the desired slopes without any threshold on the correlation coefficient between N_d and LWP. We note that our approach is not designed to select a narrow, linear band of points in ln(LWP)–ln(N_d) space but rather relies on the correlation between N_d–LWP to infer the relationship between them, given the relatively large number of samples in each sub-ensemble*".

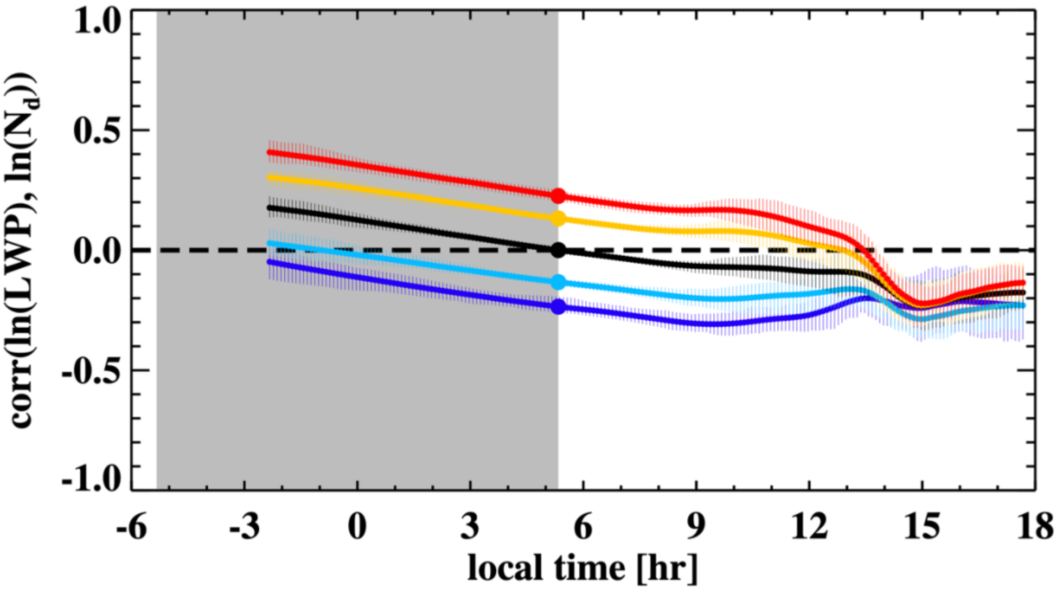

Figure R 2. Same as in Fig. 1 of the manuscript, but for the correlation between LWP and N$_d$.

- L173: Can you quantify the entrainment velocity from your output? How does entrainment velocity quantitatively depend on droplet radius in these simulations?

Yes, we can. Entrainment velocity (w_e) can be written as $w_e = \frac{d(z_i)}{d(t)} - w_s(z_i)$, which can be easily diagnosed from the motion of inversion height (z_i) and the prescribed large-scale subsidence profile ($w_s(z_i)$), directly from the simulations. Below, I show w_e's dependence on N$_d$ (Fig. R3) and r$_e$ (Fig. R4) for 5 different LWP-N$_d$ slope groups as in Fig. 1 of the manuscript. As expected, entrainment velocity positively (negatively) correlates with N$_d$ (r$_e$), indicating an increasing number of smaller drops enhances entrainment at cloud tops, with exceptions in the sub-ensemble groups where LWP-N$_d$ relationship is negative (cyan and blue). In addition to the drop-size dependence, entrainment velocity is also sensitive to the LWP of the cloud (as cloud top radiative cooling scales with LWP). This explains what we see in the cyan and blue groups where high-N$_d$ is associated with low-LWP, leading to a smaller entrainment velocity that offsets the enhancement due to smaller drops, resulting in negative (positive) slopes between w$_e$ and N$_d$ (r$_e$). Note this high-N$_d$ – low w$_e$ relationship does not contradict our finding of the persistent negative trend in the LWP-Nd slope at night, because we take into account the contribution from the radiation-process to the LWP tendency (recall Fig. 2c and 2d in the manuscript).

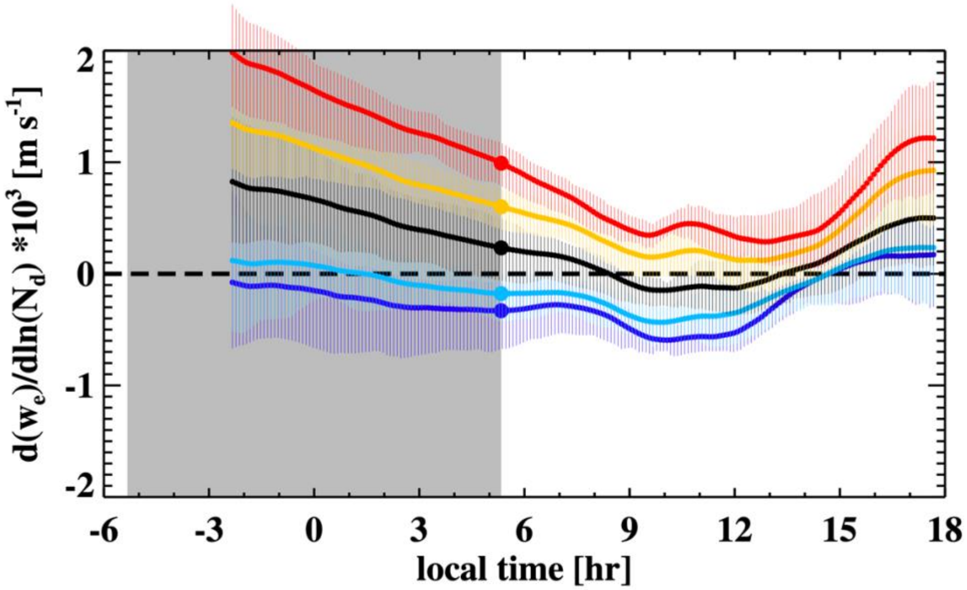

Figure R 3. The slope between entrainment velocity (w_e) and N_d, for 5 LWP-N_d slope groups.

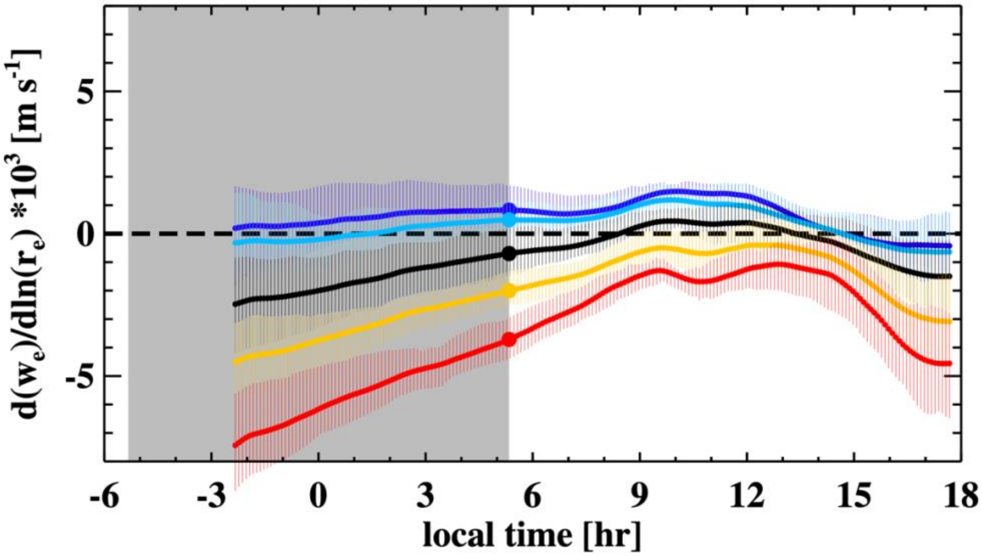

Figure R 4. The slope between entrainment velocity (w_e) and r_e, for 5 LWP-Nd slope groups.

- L204: If phi_ENT tendency is calculated as the residual between the total phi tendency and the phi_RAD tendency, then what is the "residual" referred to here?

Thanks for raising this point, we realize our statements here could cause confusion. Because of the way <Ø>_ENT is calculated (i.e., as the residual after accounting for all processes other than ENT), there is no residual in the <Ø> (i.e., theta_l and q_t) budget. The 'residual' we referred to in the text is the residual between the LWP tendency (L') diagnosed directly from the simulation and those calculations based on mixed-layer theory (i.e., L' in equation (1), referred to as LWP tendency budget analysis), following Chen et al. 2024, ACPD. Even though there is no residual for <Ø>, there can still be a residual in the LWP tendency between mixed-layer theory and the actual tendency in LWP derived from the difference between two time steps in the simulations.

We revised these statements, which now reads *"We caution that during the late afternoon the difference between the L' from the budget analysis (i.e., Eqn. 1) and the L' diagnosed directly from the simulations increases, …"*

- Fig 2. I would just recommend using a different color palette in panel a) to distinguish these RAD and ENT components from the coloring of the sub-ensembles used in the other panels.

Great suggestion! and done!

- Fig 3. Can you use the same colors (shade of blue) as the sub-ensembles that you are referring to? It says in the caption that blue refers to blue, but since it's a different color blue this is a little confusing.

Done, and thanks for the suggestion!

- L260: This should be caveated with the assumption that all other conditions are unchanging over time, besides Nd.

We agree, revised text now reads *"Given the persistent decreasing trend in dln(LWP)/dln(Nd) during the night (Fig. 1), assuming unchanged large-scale meteorological conditions throughout the day, one can relate the sunrise value of dln(LWP)/dln(Nd) to the elapsed time since the perturbation in* N_d *was introduced."*

- L268: Can you clarify the details of your regression two-sided t-test to determine the near-zero slopes? What is the p-value?

Sorry for the confusion here, the way we select near-zero slopes is exactly the same as we introduced in Section 2.2 (L145-146), that is to draw 25-simulations that satisfy $-0.005 \leq \mathrm{dln(LWP)/dln}(N_d) \leq 0.005$ and $-0.05 \leq \mathrm{d}f_c/\mathrm{dln}(N_d) \leq 0.05$, in addition to the flat N_d-MET slopes that we already imposed. We did not use a t-test to determine whether the slope is insignificantly different from zero.

We modified the text to clarify this, which now reads "*We use the cMC method to subsample conditions where a 25-member subset of the LES ensemble has near-zero* N_d*–LWP and* N_d*–fc slopes, to mimic flat slopes between cloud micro- and macro- physical properties, in addition to the constraint on* N_d*–MET covariations. (See Sec. 2.2 for the threshold values used to impose these constraints.)*"

- L272: Re: my comment of limitation #1, I suggest adding a comment here in the text to clarify that this implication for MCB is limited by the opportunistic sampling strategy. Because you do not simulate actual injection, the "injected particles" necessarily come from the same underlying distribution as the background particles. However, in a realistic MCB simulation you would probably seed with larger, more hygroscopic particles to resemble sea salt. The distinction here may be subtle, but so is the prospect of MCB efficacy. The "aerosol perturbation" then referenced is really more similar to a perturbation in the background aerosol, some co-variability between meteorology and aerosol, than a deliberate MCB seeding.

We totally agree with this excellent point and we thank the reviewer for raising it. This statement is now revised to read as "*Although our opportunistic sampling strategy based on background aerosol conditions does not fully represent deliberate aerosol seeding, such as MCB, which will likely inject larger and more hygroscopic particles than we assumed in these simulations, it does provide insights into the qualitative relationship between MCB efficacy and seeding time.*"

- Fig 6. Is the "aerosol perturbation" time the same as local time? Can you add back the grey shading that you have on all the other figures to indicate night from daytime?

Yes, indeed, the x-axis is intended to indicate the 'local time' at which "aerosol perturbation" occurs. We have changed the x-axis label to "*local time since midnight at which 'aerosol perturbation' occurs [hr]*". Since all "aerosol perturbations" occur (by design) during nighttime, we don't think it's necessary to show the full 24-hr on the x-axis and therefore that there is no need to add the grey shading to indicate nighttime.

- Fig 8. How is the cloud aspect ratio defined? Where does the scaling come from? What are the assumptions that go into this scaling? Please give more explanation and a citation, if one exists.

We agree there is a lack of information on the use of 'cloud aspect ratio'. The original attempt was to roughly indicate the ratio between the vertical and horizontal extent of the clouds (i.e., vertical-to-horizontal aspect ratio) to illustrate the transition from more stratiform clouds to cumuliform clouds as part of the diurnal evolution. We now find this added layer of information not critical and does not add to the point that we are making, thus, we have decided to remove it from Fig. 8.

- L364: You discuss how advection will change the large-scale forcing (SST, subsidence). But again, re: my earlier point about the limitations, you also should add to the discussion how the variability in initial large-scale conditions may alter these results. Some of these limitations may be introduced earlier in the paper.

Yes, we agree that the fixed SST and subsidence profile leads to an under-sampling of the dynamical space of initial large-scale conditions that real-world clouds occur in. Since the focus of our study is the evolution of the LWP-N_d relationship, the use of this conditional sub-sampling approach, by design (selecting sub-ensemble that do not have an apparent MET-N_d relationship, or ‘controlling meteorology’), alleviates the dependence of our conclusion (a buffered evolution in LWP-N_d relationship attributed to solar heating) on the initial large-scale conditions that the LES-ensemble encompasses. In other words, the randomness of the repeated sub-sampling (50 times for each group) is designed to capture various combinations of initial MET conditions that lead to the same LWP-N_d slope.

That said, we do acknowledge that large-scale conditions, such as SST and subsidence, affect the rate of thermodynamical and dynamical processes, which may change the exact evolution in the LWP-N_d slope, but not the fact that it appears buffered, because the underlying relationship between cloud properties (LWP, N_d) and processes that govern the slope evolution, i.e., entrainment and radiation, remains unchanged.

To prove this point and test the robustness of our conclusion, we re-ran some of our simulations with different SSTs (still fixed in time). The result (Fig. R5) suggests a robust feature of the buffered evolution in LWP-N_d slope with subtle differences in the timing of convergence and perhaps the strength of the buffering.

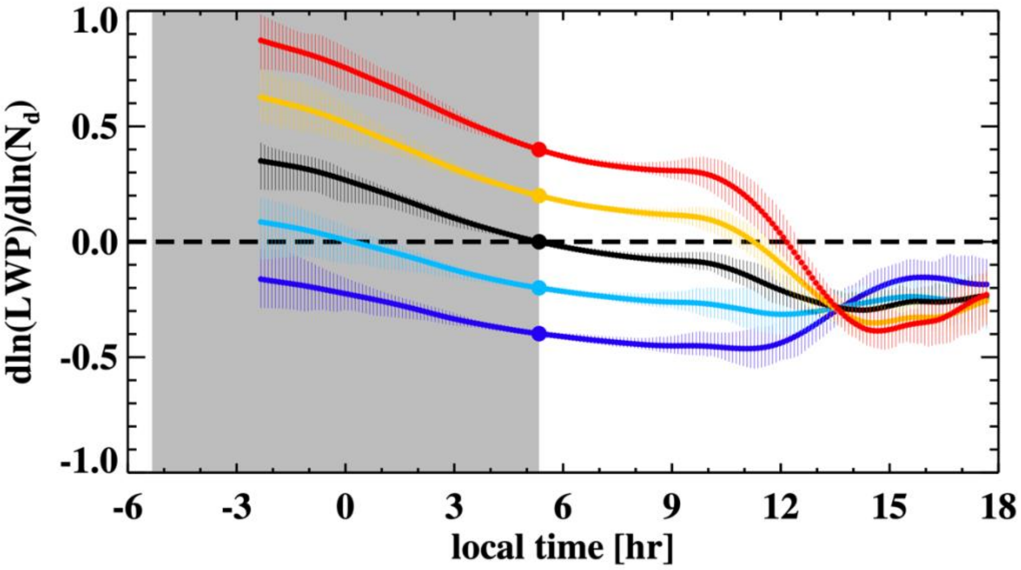

Figure R 5. As in Fig. 1 of the manuscript, but for simulations with variant SST conditions, everything else being the same as the original ensemble.

We added discussion on the limitation on our MCB-related implications, due to the fact that SST and subsidence being the same among simulations, *“Although many aspects of the boundary layer thermodynamic structure are varied to construct the large ensemble, two large-scale conditions, namely SST and free-troposphere subsidence, are fixed among ensemble members. The cMC approach is designed to effectively limit the role that the variability in these large-scale conditions can play in driving the evolution in the Nd-LWP relationship, by sub-sampling simulations with flat slopes between Nd and other cloud controlling factors at the beginning of the simulations. Although such a variability in the prescribed large-scale conditions can cause* *subtle differences in the exact timing and strength of the "buffered" feature, the finding of the feature itself remains robust based on a sensitivity test with variable SSTs simulations (not shown). Once again, the concept of using a large ensemble with cMC sampling is not to provide a reference value for the Nd-LWP relationship, which may still be weakly dependent on the prescribed SST and subsidence even after applying cMC, but to explore features of the Sc system that are robust even in the context of (co-)varying large-scale conditions, e.g., in the real world."*

We also added a sentence at the end of the Introduction to raise these limitations on the implications for MCB earlier in the paper, *"This has implications for the optimal timing of deliberate aerosol perturbations in the context of Marine Cloud Brightening (MCB), one of the proposed climate intervention approaches (National Academies of Sciences, Engineering, and Medicine (NASEM) report, 2021; Latham and Smith, 1990; Latham et al., 2012), to the extent that they are constrained by the duration and the prescribed, time-invariant large-scale conditions of these simulations."*

References

Chen, Y.-S., Zhang, J., Hoffmann, F., Yamaguchi, T., Glassmeier, F., Zhou, X., and Feingold, G.: Diurnal evolution of non-precipitating marine stratocumuli in an LES ensemble, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-1033, 2024.

Reviewer 2

This paper uses a conditional Monte Carlo subsampling approach to analyze the diurnal response in the liquid water path (LWP) adjustment to solar heating in non-precipitating stratocumulus using large-eddy simulations, finding that LWP has a strong dependence on shortwave heating which act to modulate the overall adjustment. Overall, I think this is a strong paper that is well written and only have a few critiques that I want the authors to address prior to publication.

We thank the reviewer for the encouraging words and insightful comments! A point-by-point response is provided below.

General Comments:

1. L80-81:
    1. Assuming you chose 24-hour simulations because you want to investigate the diurnal variability in the LWP adjustment, would there be any benefit in running longer simulations (e.g. 36, 48, or 72 hrs.)?

    There is benefit in extending the same set of simulations to a second and/or a third day, given the aerosol lifetime in marine boundary layer (a few days), but only to a certain extent, as, ideally, one would want to have Lagrangian simulations of extended duration (e.g., 72 hours) to capture the temporally-integrated cloud responses under varying large-scale conditions to represent the full extent of cloud evolution in the real world.

Our group is currently running Lagrangian multi-day simulations under observationally informed large-scale conditions in a follow-up work.

An imperfect attempt to artificially extend our conclusions to 3 diurnal cycles (without needing to actually run 72-hour simulations) is provided in Fig. S4 in the supplement material, where we "recycled" the 24-hour simulations and re-subsampled simulations (using cMC) after spin-up based on the LWP-Nd slope values at the end of the first diurnal evolution. This exercise suggests the "buffered" feature observed in the first diurnal cycle is robust.

2. All your simulations start at 18:40 local, do your results depend on when the simulations start (i.e. the overall trends in figure 1)?

We don't think our result is sensitive to when the simulations start, as long as there is enough turbulence spin-up time before sunrise (to examine the impact of solar heating). The results shown in Fig. 1 are based on a series of LWP-N_d relationships conditioned at sunrise, by which time turbulence has sufficiently spun up, given a 18:40 start time. Moreover, Fig. S4 in the supplement material shows evidence of a robust "buffered evolution" feature if the 24-hour simulations were reused for a second and a third day (i.e., if the simulations had started one or two days earlier).

We revised the discussion around this point in Section 4 in light of this comment, *"Extending the analysis to three diurnal cycles by re-using the 24-hour simulations for cMC subsampling results in similar conclusions with respect to the persistent nighttime negative trend in the N_d–LWP slope and the daytime buffering due to SW absorption, which essentially makes the N_d–LWP slope oscillate between -0.1 and -0.4 after convergence during the first afternoon (Fig. S3)".*

2. L141: Why did you choose the thresholds on cloud-top height, surface sensible heat flux, and 800 hPa relative humidity listed here?

These arbitrary choices are intended to minimize the regression slope between N_d and meteorological factors (that control cloud evolution) while maintaining practical sampling efficiency. In other words, a trade-off between approximating a zero-slope and time needed for cMC to obtain enough samples. (recognizing that the smaller are the bounds around zero, the longer it takes to obtain the desired number of "random" samples)

We added *"In order to maintain practical sampling efficiency of the cMC approach while approximating desired regression slopes, we impose arbitrary bounding values (or thresholds) around the desired slopes without any threshold on the correlation coefficient between N_d and LWP. We note that our approach is not designed to select a narrow, linear band of points in ln(LWP)–ln(N_d) space but rather relies on the correlation between N_d–LWP to infer the relationship between them, given the relatively large number of samples in each sub-ensemble"* to make this clearer in the main text.

3. L352-360: How frequent are non-precipitating stratocumulus and, given recent observational studies demonstrating the diurnal impact of cloud-top entrainment on the LWP adjustment may also be modulated by precipitation (e.g. Smalley et al. 2024), how representative are your simulations of the real world? On a side note, are there any plans in the future to do similar analyses of precipitating stratocumulus cases?

Precipitation process is definitely an important aspect of aerosol-cloud interactions and MCB in particular. While acknowledging the limitations on MCB-related implications, we focus on non-precipitating stratocumulus clouds only in this study. And, yes, our group is currently analyzing the precipitating cases in this large ensemble.

How frequent are non-precipitating stratocumulus? It depends strongly on the geographical location (or large-scale meteorological conditions) and how one defines precipitation. For the conditions that these simulations are initialized with (the heart of the NE Pacific Sc deck), about 20% of the 316 simulations have precipitation (defined as cloud base rain rate exceeding 0.5 mm/day). This fraction compares well with a satellite-based study that finds the frequency of precipitating clouds is about 22% over the northeastern Pacific stratocumulus deck (using an effective radius threshold of 15 micron, see Zhang et al. 2022, ACP).

How representative are your simulations of the real world? The design of Latin-Hypercube sampling is to cover the range of conditions in the real world with the least number of samples. Being constrained by the range of conditions in the real world, based on ERA5 climatology, the goal of the large ensemble is to simulate stratocumulus clouds under conditions that are representative of the real world. The fact that these initial conditions lead to a similar fraction of Sc that precipitates, compared to that observed by the satellite, indicates the suitability of the ensemble approach.

We added mention of Smalley et al. 2024's finding along the line of these discussions *"For one, we focus only on non-precipitating Sc systems, whereas studies have shown that precipitation can modulate the impact of cloud-top entrainment on the LWP adjustment (Smalley et al., 2024; Stevens et al., 1998). Furthermore, suppressing or even preventing precipitation in Sc systems can potentially generate larger radiative impacts, compared to brightening non-precipitating systems (e.g.,Wang and Feingold, 2009; Prabhakaran et al., 2023, 2024; Chun et al., 2023)."*

Minor Comments:

1. L20: "lead to more, smaller" sounds awkward. Maybe change it to "leads to an increase in smaller"

Thanks for the suggestion. Text revised accordingly.

2. L35: "Making the quantification of LWP adjustment" should be "making the quantification of the LWP adjustment"

Corrected, thanks!

3. Figure 1: For ease of interpretation, could you move the threshold values of dln(LWP)/dln(N_d) listed in the caption to a plot legend instead?

Done! Thanks for the suggestion.

References

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