Reviewer #2:

In this paper the authors use the ACTIVATE field campaign to compare the E3SM SCM to different flavors of WRF (in both CRM and LES mode) with regards to the simulation of clouds and boundary layer turbulence observed during the campaign. The second part of the paper focuses on a set of E3SM-SCM experiments focused on the sensitivity to treatment of aerosols in the model.

While I found aspects of this paper to be interesting, it also felt like a hodge-podge of ideas/experiments that lacked a clear unifying focus as to what the authors hoped to accomplish/address. The two distinct sections of the paper feel a bit disjointed and I think the authors could do a better job tying them together a bit more. In addition, the second part of the paper focuses on just one case from ACTIVATE to draw some conclusions. I feel this needs to be addressed by testing robustness against more flights from the ACTIVATE campaign. In addition, there were several other sources of confusion in this document that need to be addressed (please see itemized list). Overall, I feel a major revision is necessary before this article is suitable for publication.

Overall, the paper is well written enough that I understand what the authors are saying; but there are frequent typos and grammar mistakes that are distracting and needs to be addressed upon resubmission.

We thank the referee for reviewing the manuscript and providing valuable comments. We have revised the manuscript accordingly in response to the general comments here as well as the major comments from Referee #1. The manuscript has also been checked through *Grammarly* and by native English-speaking coauthors thoroughly for grammar mistakes. Below please find our point-by-point responses to your specific comments.

In the conclusions of the paper the authors state (and elude to this on other sections of the manuscript): "A unique feature of this study is the multi-scale model intercomparison using SCM, CRM, and LES models, which provides a comprehensive process-level understanding of ACI in more details compared to individual models". I'm left very confused by this statement. The CRM and LES models were only used in the first half to compare the macroscopic aspects of the SCM simulation (clouds, turbulence, etc.); I do not see how they were used to help understand ACI directly other than being used as a validation tool.

We have now revised the title and the main text to emphasize the model evaluation and process-level understanding of ACI in E3SM-SCM using ACTIVATE observations as well as the CRM/LES results for validation/comparison. The new title is "Understanding Aerosol-Cloud Interactions Using a Single-Column Model for a Cold-Air Outbreak Case during the ACTIVATE Campaign", while the comparison with CRM and LES results is now part of the SCM evaluation.

I found the comparison of E3SM-SCM to WRF interesting but was confused why the authors felt it pertinent to include the SCM and LES runs with the CRM forcing. The

conclusions they draw of "proper combination of large-scale dynamics, sub-grid parameterizations, and model configurations is needed to obtain performance…" seems like a super obvious conclusion that I'm not sure why they felt needed detailed analysis. Unless I'm missing something I suggest that the authors remove these curves from the figures (which are too busy with these curves included) and perhaps state in a sentence or two that they explored the sensitivity to large-scale forcing. To me this analysis and section felt like a distracting tangent.

We agree that the SCM sensitivity to the large-scale forcing is expected (without an exception for this CAO case). As we follow the suggestion to de-emphasize SCM/CRM/LES intercomparison and focus more on the ACI in SCM, we have removed the results of SCM using CRM forcing from the main text.

Page 8, line 173 the authors state "…neither resolved nor parameterized at the sub-grid scale in E3SM-SCM". What exactly is "resolved" in an SCM? Isn't SCM just one column where all processes are parameterized? Or am I misunderstanding something about the E3SM SCM?

What we were trying to say is that this structure is smaller than the grid size represented by the SCM. To avoid such confusion, we have revised this sentence to

"However, this roll structure fails to be simulated in WRF-LES and is not parameterized in E3SM-SCM."

In the second half of the document, which explores E3SM-SCM to aerosol sensitivity the authors use one case and make the statement that their conclusions "warrant more cases" to test robustness. I completely agree… why not include more cases then?

Thank you for the comments. We agree that one single case is limited to generate a robust or general conclusion on ACI. We did attempt to consider other "process study" cases during the ACTIVATE field campaign. However, in practice, there are a few issues preventing us adding more cases in this ACI study of CAO over the western North Atlantic:

1. Aerosol-cloud interactions over the western North Atlantic are highly dependent on synoptic conditions. This study is mainly focused on cold-air outbreak conditions during the ACTIVATE campaign. For the limited number of process-study cases sampled during the ACTIVATE campaign, most of them are for summer cumulus clouds (e.g., Sorooshian et al., 2023; Li et al., 2024). Therefore, we don't have many cases to choose from if we want to evaluate the SCM results against extensive ACTIVATE observations and compare with fine-resolution process-model results.
2. The E3SM SCM fails to reproduce the cloud evolution of the other ACTIVATE CAO case. We conducted E3SM-SCM simulations for the CAO case on February 28, 2020. However, the SCM fails to reproduce the observed persistent MBL clouds and the cloud growth later in the day (see the plot below). As ACI metrics and sensitivity to aerosols are dependent on cloud and thermodynamic conditions, the failure of reproducing cloud evolution makes it questionable to study ACI in SCM for this CAO case.

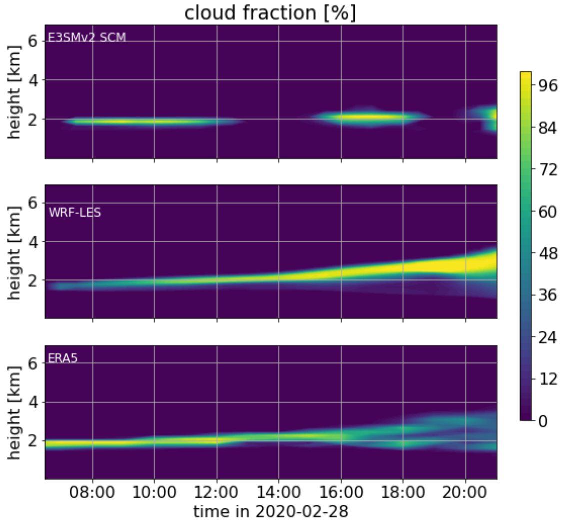

Figure R1: cloud fraction evolution from E3SM-SCM, WRF-LES, and ERA5 for the 28 February case.

Part of the reason for choosing the 1 March 2020 case is the well-represented CAO synoptic conditions (and forcings from reanalysis) and the good performance of MBL clouds in other models. As seen in the previous studies with WRF-CRM and WRF-LES models (Chen et al., 2022; Li et al., 2022, 2023), a single well-simulated case may still provide useful information on understanding the dynamics, thermodynamics, and aerosol-cloud-meteorology interactions over this region. We have also revised the structure of this paper and added a few more sensitivity tests and analyses to further understand the behavior of model physics in E3SM-SCM. On the other hand, we do have a plan of running long-term SCM simulations for statistical analysis over the ACTIVATE domain. However, including long-term statistical analysis in this CAO case study would defeat the purpose of evaluating the SCM against field observations and intercomparing SCM with CRM/LES results, so we decided to keep this paper as a single CAO case study and focus on the long-term analysis in a follow-up study.

In the second half of the paper the authors state "…since Nd is underestimated it is difficult to demonstrate the value of adding aerosol vertical variation" which is blamed on weak vertical velocity updraft coming from the model. Why not do sensitivity experiments where the input vertical velocity to the aerosol activation is boasted by a certain factor to test the sensitivity? This is the type of experiment that the SCM is ideal for. In the conclusions the authors state "the evaluation of SCM simulations against the ACTIVATE measurements is helpful for understanding and improving turbulence representation over this region". I don't think the authors have currently done this, but experiments that show the possible improvements/sensitivity with better turbulence linked to aerosol activation could provide justification for such a statement to be retained.

Thank you for the great suggestion. We have now conducted a sensitivity test by directly enhancing the vertical velocity variance by a factor of 2, which makes the simulated turbulence close to the aircraft observation. As a result, the simulated cloud droplet number concentration and effective radius are closer to the aircraft observation (on top of the effect of using observed aerosol number concentration). This result is now added to Section 4.1 and the plot is included as Fig. 8 (shown below).

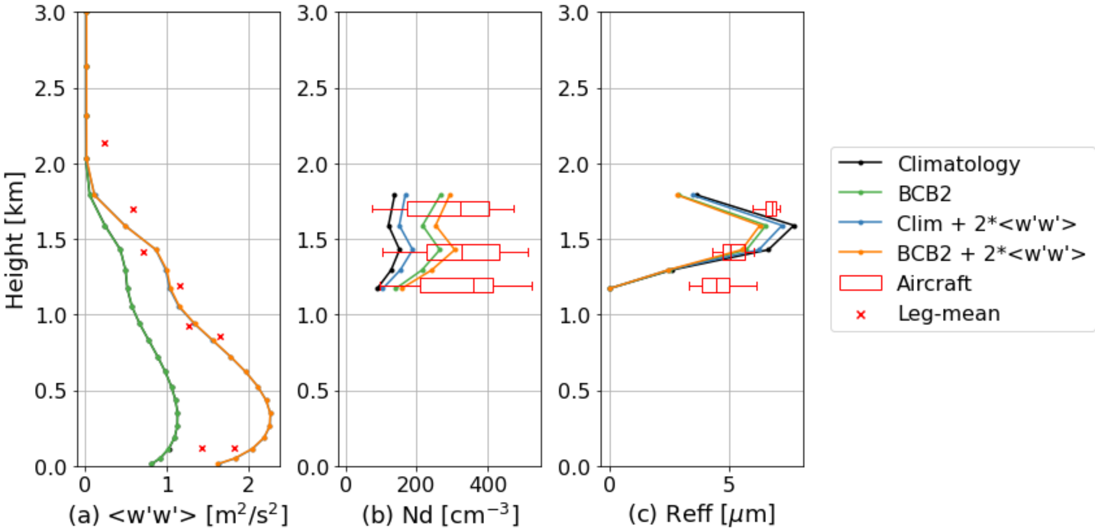

Figure 8: (a) Vertical velocity variance <w'w'>, (b) cloud droplet number concentration N_d and (c) cloud droplet effective radius R_{eff} averaged between 15:00 and 16:00 UTC, when the aircraft measurements (shown in red crosses and boxes) were made. In the figure legend, "Climatology" is the original SCM simulation with prescribed aerosol concentration; "BCB2" is the SCM simulation with aerosol number concentration from the aircraft measurement at the BCB2 leg; and "2*<w'w'>" means the vertical velocity variance is enhanced by the factor of 2 in the SCM aerosol activation scheme.

Section 4.3 ends with a statement that "E3SM SCM cannot provide information on sensitivity of aerosol to vertical distribution"… then why present results in this section at all? I feel adding a couple sentences or a paragraph to the conclusions/summary section stating that the authors attempted to address this in the SCM but couldn't because of x, y, and z would be good and sufficient… Which could help to motivate development of a validated aerosol model in the SCM.

We have now removed this part from the main text, with only a brief statement in the Summary and Discussion section:

"In the current SCM framework using observed aerosols, usually only one set of aerosol parameters, characterizing the spatially mean properties (i.e., particle number size distribution and composition), is fed into the model regardless of the aerosol vertical distribution (Liu et al., 2011; Liu et al., 2007; Klein et al., 2009; Lebassi-Habtezion and Caldwell, 2015; Li et al., 2023). The prescribed aerosol information based on observations is usually taken from in-situ measurements below the cloud base (e.g., Liu et al., 2011; Li et al., 2023), assuming that hygroscopic aerosol particles are readily activated into cloud droplets in the saturated air driven by updrafts. However, as aerosol concentration usually decreases with height in the lower atmosphere, regional aerosol vertical distribution may be changed

by in-cloud scavenging, horizontal transport, and vertical mixing, which can further affect cloud microphysical properties by secondary activation above cloud base (Wang et al., 2013; Wang et al., 2020). We conducted a sensitivity experiment with a specified aerosol vertical distribution (Fig. S5), but the configuration of prescribed aerosols in SCM only shows the response of clouds to aerosols given at the level of cloud formation. A more comprehensive consideration of complete aerosol processes (e.g., vertical transport, scavenging, deposition, etc.) is needed to include the cloud and dynamical feedback on aerosols and better understand the aerosol-cloud interactions."