Review of "The Cold-Air Outbreaks in the Marine Boundary Layer Experiment model-observation intercomparison project (COMBLE-MIP), Part I: Model specification, observational constraints, and preliminary findings" by Juliano et al. [MS No.: egusphere-2026-1237]

This paper presents a very useful LES and SCM model intercomparison project that simulates a strong cold-air outbreak (CAO) event. The models participated the project run within a quasi-Lagrangian framework and are examined against a suite of ground observations, satellite data, and ERA5 reanalysis. The preliminary results are promising: LES models can broadly capture the evolution of observed cloud top height and liquid water path, along with the roll and open-cell structure of clouds in a km-scale domain. More insights into the role of ice microphysics are provided by comparing liquid-only simulations with mixed-phase simulations. The differences between LES and SCM runs and the source of modeling biases in longwave radiative fluxes at TOA and the surfaces are also investigated.

The paper is well written and organized, easy to follow. The initial conditions, large-scale forcings, and modeling setups are reasonably designed. One unique merit of this study is that they chose a very strong cold-advection event, manifested as more extreme surface fluxes forcing, so that cloud and radiative properties and their mesoscale organization along the CAO evolve significantly, providing a strong-signal test bed of CAOs for Earth system models, especially km-scale models. That said, the study could benefit from more discussions of the physical process involved, though the study focuses on presenting the preliminary modeling results. My comments and suggestions mainly concern some clarifications, and once they are addressed, I believe the paper would be suitable for publication in ACP.

Major comments:

1. The authors describe two methods for handling the large-scale wind forcings. I assume the presented simulations are based on the geostrophic adjustment approach, since wind speeds notably deviate from ERA5, as seen in Figure 4i. At 18:00 UTC on 03-13-2020, both ERA5 and the radiosondes show a wind speed peak between 1000 and 2000 m, indicating a strong wind shear in the PBL. This feature is not captured by the models, and the wind-shear driven turbulence may be missed as well, which could help explain why the simulated cloud-top heights are lower than those from ARSCL (Fig. 2f). Cloud-top heights also affect the LW radiation fluxes. In this regard, did the authors compare the simulations from the above two approaches regarding cloud-top heights and LW radiative fluxes (TOA and surface)?

2. As for the impact of ice microphysics on LW radiation fluxes, changes in cloud cover explain the LW radiation changes very well for the LES simulation, but not for SCM since both liquid-only and mix-phase SCM simulate overcast cloud decks, constraining the cloud cover ratio to near unity (Fig. 12c). In addition to cloud cover, the cloud-top height ratio might help explain the LW radiative flux changes, given the influence of cloud top on LW radiation and the notably divergent cloud-top height/temperature differences between liquid-only and mix-phase simulations (Figs. 3e-f or 3g-h).

3. The paper would benefit more if the authors could further clarify the physical processes involved: how including ice microphysical processes impact on cloud-top height and cloud cover. For the

first process, the authors argue that precipitation made the MBL more stratified, which hinders vertical mixing and thereby PBL deepening. However, comparing Fig. 2p with Fig. 2e, the surface precipitation does not increase and may even become weaker in the mixed-phase LES runs. This does not appear to support the authors' argument, unless there is evidence that sub-cloud evaporation becomes stronger in the mixed-phase run.

It would also help if the authors could elaborate on why ice microphysics leads to more cloud breakup or a smaller cloud fraction in LES models. If precipitation changes do not account for this, any other processes might be responsible? Relatedly, since cloud cover is an important variable during CAOs, can the authors also add cloud cover evaluations to in Figs. 2 and 3 if possible?

4. For open research, would the authors also consider sharing the simulation setup files? This would allow readers to reproduce the simulations if they are interested in specific cases and would facilitate the extension of the study with additional sensitivity tests.

Minor comments:

L68: Relative to "surface heterogeneities"? Please clarify.

L79: Revise "atmospheric scales" to "atmospheric motions"

L150: In Table 1, please also add the "SW radiation" setup

L172: The phrase "started from" seems to fit better here than "reached".

L174: Please clarify how the trajectory is computed, e.g., initial height, whether it is run on an isobaric level, and the time resolution of the trajectory.

L178: Since ERA5 data are used for the initial profiles, please note that ERA5 sometimes has lower boundary layer heights due to its coarse resolution, which could result in a shallower simulated boundary layer due to this initial bias. I suggest adding a caveat up front to give readers proper context.

L195: Are both nudging methods applied? Please clarify which one is used in the following simulations.

L324: Suggest mentioning the two modeling setups (liquid-only vs. mixed-phase) earlier, possibly in Section 2.2.4.

L355: Suggest narrowing the y-axis range in Fig. 3l for better visibility.

L395: "2e-d" should be "2e-f"?

L417: Revise "(Young et al., 2002, , their Fig. 4)" to "Young et al. (2002) (see their Fig. 4)"

L569: Please clarify the fixed numbers in the modeling setups, as well as the translation speed.