Overview & Summary: This manuscript evaluated cloud radiative biases in the ACCESS-AM2 model and the CERES satellite product against ground-based observations during the MICRE campaign at Macquarie Island. The study used a suite of instrumentation to establish a positive bias in surface shortwave radiation in both the model and CERES, which is consistent with numerous past radiative flux modeling studies over the Southern Ocean. The authors also employed a new lidar simulator to constrain model output with observational instrument specifications, therefore allowing an evaluation of cloud attribution to the radiative biases. They found that excessive absorbed surface shortwave radiation is associated with too low of a cloud fraction. Overall, the manuscript is well-written, descriptive where needed and brief where appropriate, and the experiment is well-designed. The largest concern is the lack of discussion of the potential role that cloud microphysics have in the discussed biases. While I understand this is not a microphysics study, there are long-standing microphysical biases in GCMs that are suggested to be relevant to simulating Southern Ocean clouds (e.g., prevalence and maintenance of supercooled liquid) that could be discussed in more detail as potential caveats. Otherwise, I have provided some general comments and line-specific comments below, none of which should require too much extra effort to delay publication. I suggest this manuscript be accepted for publication after these minor revisions are addressed.

General Comments:

- In the second paragraph of the Introduction when you introduce the ALCF, I think some expansion is needed for unfamiliar readers of how this framework operates. Typically, simulators are thought of as being applied to model data, not on observational data, so this can be confusing and deserves a little more explanation. Based on your description in Section 2.6, it seems ALCF is used on the observational ceilometer dataset as more of a means of calibration rather than "simulation".
- In the fourth paragraph of the introduction, I think something should be said about satellite limitations in observing low-level clouds over the SO, which is a very strong motivation for evaluating ground-based observations. As one example, Tansey et al. (2022; https://doi.org/10.1029/2021JD035370) looked at surface precipitation measurements during MICRE in comparison with CloudSat, with a few notable differences based on satellite instrument sensitivities and algorithm structure.
- You discuss how observed cloud fraction is defined by the all-sky cloud camera (Section 2.2.3) and then state it is used to evaluate CF biases on line 337 in Section 4, but it's not entirely clear how CF is being defined by ACCESS-AM2. Is it the prognostic value? Is it the CF computed by ALCF? It would be good to mention this at the beginning of Section 4.
- Section 2.1 and other parts of manuscript—I wouldn't call this in-situ observations, since many of what is included are more commonly thought of as passive remote sensing instruments. Perhaps change to "ground-based" observations, or convince me otherwise.

- Figure 1 caption—what does the blue color scale represent in Fig. 1a? I'd briefly mention it in the caption.

- Data availability: There was no data source given for the University of Canterbury ceilometer (CL51) data. Please include it as follows: https://doi.org/10.26179/5d91835e2ccc3 . In addition, no data sources were given for the radiometers or the all-sky cloud camera. The ARM data availability statement suffices for the ARM instruments, but if a DOI exists for these AAD instruments, they should be listed too.

- Perhaps a *little* more could be said about cloud phase in the last 2 paragraphs of Section 6 and how more or less supercooled liquid in the model relative to observations can impact your results. In general, the discussion of cloud microphysics in the manuscript is rather weak, and providing some speculative pathways for explanation would be very helpful.

Line-specific Comments:

Line 108: "was" should be "were"
Line 110: The two clauses should be joined by a conjunction, not a comma.
Line 136: I think you have to be careful when saying that supercooled clouds are typically not visible in the backscatter profile. Liquid-based supercooled clouds can exist at rather low altitudes (< 1-2 km AGL) over Macquarie Island, and in the absence of an underlying layer, will show a sharp gradient in attenuated backscatter consistent with a liquid cloud base identification. I would also mention that fog can frequently be observed in the backscatter profiles.
Line 216: Please give some reference to what type of CBH algorithm is used, since these are not trivial nor converged methods. I assume it is what is described in Section 5.3 of Kuma et al. (2021), and if so, I'd list the thresholds for attenuated volume backscattering coefficient that were employed here.
Lines 270-271: This statement is very confusing to me. I'm not sure what you mean by large spreads in the LW_{cs} distributions, as the violin plot in Fig. 3d doesn't really show spread that is larger in model/satellite relative to ERA5. Also not sure what you mean by "paying more attention to the LW_{cs} models than the SW_{cs} models". Are you suggesting that LW_{cs} biases are more important than SW_{cs} biases, or appear to have a higher sensitivity? Please clear this up.
Line 275: Careful with this statement. "Too few, too bright" refers to compensating errors. All else being equal, "too bright" should *reflect* more surface SW radiation, so this component can't really be linked to an overestimation of absorbed surface SW.
Lines 289-291: Aside from the absolute values being ~ 2 W/m^2 off, the values you listed do seem consistent with Hinkelman and Marchand (2020)---SW bias of +8 for your study and +10 for their study, and LW bias of -12 for your study and -10 for their study. Understanding that 2 W/m^2 is not a small amount, I'd suggest rewording because "consistency" usually refers to how biases operate in sign, even if magnitudes are different.
Line 299: It's not necessary for this study, but perhaps the frequent soundings released at Macquarie Island can give you an idea of potential humidity and temperature biases in ERA5.

Line 302: It's not clear to me what you mean by three algorithms used for SW_cs biases. Are you including the ERA5 SW_cs calculations in this statement?

Line 304: Again, I wouldn't use "in-situ" here, as I think much of the community thinks of these as passive remote sensing instruments.

Lines 301-305: This paragraph seems a little out of place as you start discussing CREs. I expected this to be a transition to the next section, but that also wasn't clear because the paragraph ends on a "future guidance" type of statement. Suggest making the transition more clear or moving this paragraph to the next section.

Line 316: You haven't yet used the term "downward CRE" and it's not used anywhere else in the manuscript, so I'd avoid it here to avoid confusion.

Line 323: You say larger negative values of CRE_LW, but CRE_LW values are strictly positive using your convention.

Line 355: Fig. 6e does not show an overestimation in CF in spring as stated here, but rather the same mean CF for ACCESS-AM2 and observations.

Line 360: Is there a reason for suggesting cloud microphysics are a lesser control on the CRE compared to other properties? I would add some references to back this up if so. Droplet radius and size distributions are inherently linked to the cloud's optical thickness, after all.

Line 382: "results" should be singular "result"

Line 385: Suggest using "While much of that is…" instead of "While a lot of that is…"

Line 473: Saying "other" CFs is rather vague. I would provide a little more detail here and give the range where the model produced lower CF (between 0.2 and 0.6), or just say "lower" CFs instead.

Line 484: I think you mean under positive SW *bias* conditions. Also, "result" should be "results".

Line 488: Suggest making it clear that you are referring to a greater dependence on low-level CFO for LW biases *compared to* SW biases, if that is indeed what you mean.

Line 489: I wouldn't say "inappropriate" cloud representation. It's as appropriate as can be given scale separations, but must inevitably be parameterized. I'd suggest using "parameterized" instead of "inappropriate".

Line 499: Again, wouldn't use "in-situ"

Line 523: Should be either "These analyses suggest" or "This analysis suggests".

References:

Tansey, E., Marchand, R., Protat, A., Alexander, S. P., and Ding, S.: Southern Ocean Precipitation Characteristics Observed From CloudSat and Ground Instrumentation During the Macquarie Island Cloud & Radiation Experiment (MICRE): April 2016 to March 2017, Journal of Geophysical Research: Atmospheres, 127, e2021JD035 370, https://doi.org/10.1029/2021JD035370, 2022.