Review of "*Supercooled Liquid Water Clouds observed and analysed at the Top of the Planetary Boundary Layer above Dome C, Antarctica*" by Ricaud et al.

Ricaud et al. present a very nice study of two cases of supercooled liquid water cloud layers measured using a suite of remote sensing instruments at Dome C, Antarctica in the summer of 2018. Exemplar cases of a "typical" and a perturbed boundary layer are detailed to show what cloud and boundary layer properties may be expected in the region and how these properties can be affected by warm moist oceanic air masses. The authors show that these perturbed boundary layers can greatly change the radiative properties of the clouds which form within them and thus affect the surface energy balance.

Comparisons with the ARPEGE-SH model show that the model fails to capture key observed characteristics of the clouds and boundary layer structure in both scenarios; specifically, the model severely underpredicts cloud supercooled liquid water and exhibits almost systematic biases in temperature and water vapour with respect to the observations. The failure of the model to capture cloud presence and phase distribution the both cases is very important to highlight to the community. The difference between observed and modelled net surface radiation in the perturbed case (up to 50 W/m^2) is particularly striking.

The study is well explained and provides clear figures to support the conclusions drawn from the observation-model and inter-case comparisons. It will provide an excellent resource as a reference study for future observation-model comparisons focusing on boundary layers on the Antarctic Plateau. The authors operate with transparency by providing access to all data used within the study, which is fantastic to see. I recommend publication subject to some minor comments and restructuring.

General comments:

I appreciate that statistical analyses of PBL properties measured/modelled at Dome C is within the future work remit of this study; however, it would be useful to provide the reader with an estimation of how representative each of these two cases were for e.g. the summer of 2018/19. It would be useful to know how typical is "typical" with perhaps an estimation of occurrence percentage. Also, it would be helpful to have some synoptic overview of the two cases presented, perhaps as a supplementary to the manuscript, with some reference to the mean synoptic state over e.g. the summer of 2018/19.

The authors mention that model 24-h model forecasts and meteorological analyses were provided, where 4D-VAR data assimilation of the latter took place every 6 h. It is stated that most of the model-observation comparison uses the forecasts, while the analyses are used at 0000 UTC and 1200 UTC (understandably due to radiosonde ingestion improving the comparison and providing a “best guess” at these times). What is unclear to me is the combination of these two datasets, and this should be made clearer in the manuscript. When are the forecasts initialised? Are they initialised every 12-h (0000 UTC or 1200 UTC) or once per day? Are they 24-h in total, or are 24 hours of each forecast (which may be for longer) used for comparison with the observations? Or, are they 24-h in total, initialised twice daily at 0000 UTC or 1200 UTC, providing the latter 12 hours for comparison with the observations (to avoid any spin up issues)?

Indeed, if 12-h subsets of the 24-h forecasts are used, with each beginning at either 0000 UTC or 1200 UTC, then the sharp transitions at 1200 UTC would be somewhat expected from this re-initialisation as the model is effectively brought from maximum divergence back to the “best guess” of the atmospheric state. The authors mention the poor agreement in cloud properties at 1100 UTC and 2300 UTC in Section 3.1, so I am inclined to believe this is how the model is being operated. One would expect improved agreement with the observations at these re-initialisation times (albeit there may still be some discrepancies with the observations which should be emphasised). If I have misunderstood, please accept my apology; however, I feel that the model use and implications of using it in forecasting mode should be discussed in greater detail with a focus on how this re-initialisation (if conducted) may be affecting any of the conclusions drawn with respect to the transitions between wet/dry conditions. Additionally, if this is the case, there is scope for more discussion on model deficiencies: if the model diverges so strongly within the forecast comparison window, it may suggest severe parametrization deficiencies within the model.

Following from this last point, the study would benefit from more discussion on why the model fails to capture the SLW layers, perhaps with specific reference to which cloud microphysical parametrizations are used within the model. Did the authors look into the process parametrizations and identify which may need to be changed to remedy the poor model-observations comparison? E.g. are the deposition freezing / Wegener-Bergeron-Findeisen mechanisms too efficient? To what extent can the SLW deficiency be caused by the poor agreement in temperature / water vapour?

The study is well written; however, it could benefit from a distinct Discussion section for the case and literature comparisons. For example, the 1st paragraph and point (1) of the last paragraph of Section 3.2, 4th and 5th paragraphs of Section 3.3 etc. read like a discussion and should be presented separately from the main study results.

Also, the following additional references could be of benefit to the study:

- O'Shea, S. J., Choularton, T. W., Flynn, M., Bower, K. N., Gallagher, M., Crosier, J., Williams, P., Crawford, I., Fleming, Z. L., Listowski, C., Kirchgaessner, A., Ladkin, R. S., and Lachlan-Cope, T.: In situ measurements of cloud microphysics and aerosol over coastal Antarctica during the MAC campaign, Atmos. Chem. Phys., 17, 13049–13070, doi:10.5194/acp-17-13049-2017, 2017.
- Grosvenor, D. P., Choularton, T. W., Lachlan-Cope, T., Gallagher, M. W., Crosier, J., Bower, K. N., Ladkin, R. S., and Dorsey, J. R.: In-situ aircraft observations of ice concentrations within clouds over the Antarctic Peninsula and Larsen Ice Shelf, Atmos. Chem. Phys., 12, 11275-11294, doi:10.5194/acp-12-11275-2012, 2012.
- Young, G., Lachlan-Cope, T., O'Shea, S. J., Dearden, C., Listowski, C., Bower, K. N., et al. Radiative effects of secondary ice enhancement in coastal Antarctic clouds. Geophysical Research Letters, 46. doi:10.1029/2018GL080551, 2019.
- King, J. C., Gadian, A., Kirchgaessner, A., Kuipers Munneke, P., Lachlan-Cope, T. A., Orr, A., Reijmer, C., Broeke, M. R., van Wessem, J. M., and Weeks, M.: Validation of the summertime surface energy budget of Larsen C Ice Shelf (Antarctica) as represented in three high-resolution atmospheric models, J. Geophys. Res. Atmos., 120, 1335–1347, doi:10.1002/2014JD022604, 2015.

Specific comments:

Page 2, line 29: Please define ARPEGE-SH as an acronym at first point of use.

Page 5, line 89: Please define YOPP SOP as SOP-SH to avoid confusion with SOPs 1-3 in the northern hemisphere. This is described in more detail in Section 2.6, but it would be beneficial to move these specific dates up to this point (or just repeat).

Page 5, line 95: Here, it is not clear what the authors mean by the "adjustable time resolution" statement, as it is not clear whether 7 mins was chosen as the time resolution or whether it is the limit of what is achievable by the instrument. However, this becomes clear within the Methods section. Please rephrase for clarity or remove.

Page 6, line 117: typo (synthesizes)

Section 2.1: can the authors comment of whether you would expect instrument functionality to be affected by the altitude difference between Pic du Midi and Dome C?

Section 2.2 (and throughout): The authors often refer to measurements with respect to mean sea level; however, given the high altitude of Dome C, this is misleading to the reader. Please could measurements made at Dome C be rephrased to "above ground level"? This would avoid any confusion.

Section 2.6: How many model levels were within the PBL? Can the authors comment on whether the vertical resolution of the model may limit its ability to capture the ~100m thick SLW clouds observed? Additionally, the spatial resolution is quite coarse: Young et al., 2019 (full reference above) found that resolution can affect cloud modelling skill, can the authors comment on whether this may be affecting their comparisons?

Page 11, lines 222-223: Is it not only the SLW from the lidar that is shown in Figure 2?

Page 11, lines 232 – 234: Small values are presented from the HAMSTRAD, what is the measurement accuracy of this instrument?

Page 12, line 260: Second pass of CALIOP/CALIPSO not shown – please consider adding figure in supplementary material to the manuscript.

Page 12, line 266: Please include reference edition number for Pruppacher and Klett as figure numbering may change between editions.

Page 14, lines 313 – 316: mention that this positive gradient indicates a stable layer, as previously explained. As currently written, the authors are leaving it to the reader to make this conclusion and should be explicitly emphasised.

Figures:

- The description of the model BL calculation is included for the first time in the introduction to Figure 5 on page 13. As this BL height is used for the first time in Figure 1, it should be introduced at this first point of use (page 10).
- I would suggest using a different colourmap between observations and modelling results to make it clear to the reader which data are being presented.
- Figure 4: Could benefit from (a) larger tick labels / different labels to explain phase masks rather than providing the number allocation; (b) indicating the altitude of Dome C to illustrate cloud layer altitude relative to ground level.

- For anomaly figures (5 and 13), please consider using a diverging colourmap with white at 0 (e.g. blue-white-red) to ease readability. Changing the colormap may make the sub-figure headings clearer, which would also be useful.
- Figure 7: it may be useful to adapt the scale to make the subtle maxima easier to see.
- Figure 14: it's quite hard to see the measured SLW cloud layer (red) on top of the high values of delta Theta (red), perhaps changing the colour (maybe white?) of the measurements would make this easier to distinguish.
- Side note: Figures 9 and 16 are fantastic, the webcam images do a great job at illustrating the different cloud conditions between the two cases. The radiation values alone don't truly convey how different cloud distribution can be, so these images are invaluable to emphasise this.