This is an interesting and generally well written paper, although it is unclear in parts. My comments are minor but numerous. The description of the simulation protocol needs to be improved; after reading the paper three times, I think there are three experiments – initial simulations, XP0 and XP1, also called “_new” – but I am rarely confident of which is being discussed at any time.

The introduction includes an extensive review of previous snow model intercomparisons, but these are all for seasonal snow with a focus on snowmelt and runoff, and are not very relevant for this paper. If going into this level of detail, however, uncertainty in model outputs due to uncertainty in meteorological driving data should also be mentioned https://hess.copernicus.org/articles/19/3153/2015/hess-19-3153-2015.html

The abstract should acknowledge that the intercomparison is for a single site on the Antarctic Plateau.

Can it be said that the surface temperature errors are consistent with the magnitude of sensible heat fluxes being too great both day and night?

This is a standard way to start a snow modelling paper, but snow is not a water resource at Dome C (and it isn't a key element of the landscape – it is the landscape!). The snow cover does not vary considerably in time and space.
“snow patterns” suggest spatial distribution of snow cover, which is not considered in Schlosser et al. (2000).

There are two sites in Etchevers et al. (2004).

“that that water may freeze”

The information in Table 1 is already provided in the text and the author list; it could be deleted.

All of the information in this sentence is repeated with more detail in the following sentences.

Rather than a list in the text, layer depths might be better presented in a table, which could then include the initial temperature and density profiles. How was initial temperature prescribed for single-layer models?

The sensitivity tests mentioned here are XP0 and XP1, not additional test deemed relevant by the participants?

No need for equations 1 and 2 to be bracketed with {

This is the only mention of calibration. It is important to know if some of the models have been calibrated for these simulations, and how.

The K&Z CM22 does not just measure visible radiation (and, in the Figure 1 caption, not just direct solar radiation).
Table 2
Wind direction is not used. A Vaisala HMP155 measures relative humidity, which requires temperature for conversion to specific humidity. Genthon et al. (2017) consider humicap measurements to be biased low for the conditions of Dome C. https://acp.copernicus.org/articles/17/691/2017/acp-17-691-2017.pdf

245
Could say when the temperature probes were installed.

Figure 2
Why are temperatures shown as not filled at four depths?

Table 3
The caption should explain the use of square brackets.

316
I assume this is not a linear profile between the surface and 10 m depth.

Figure 5
Adding air temperature to this figure would be an interesting comparison.

341
The PDFs in Figure 6 have too many extrema to be cubic functions. Were they, in fact, fitted with cubic splines?

389
What does “a root mean square error that varies from simple to double” mean?

391
The more sophisticated model do not have to represent the evolution of albedo in XP0, as I understand it.

400
Wind speed is always greater than or equal to zero, so taking its modulus does not add anything.
The proportionality constant is not simply the surface exchange coefficient; air density, and heat capacity for sensible heat flux, are also required.

Assumed overestimation of sensible heat fluxes in stable conditions is a longstanding feature of models, although it can prevent larger biases in surface temperature. https://journals.ametsoc.org/view/journals/clim/10/6/1520-0442_1997_010_1273_votseb_2.0.co_2.xml

Equation 5 is how Qh is measured and equation 6 is how it is modelled.

This is the bulk Richardson number, not the gradient Richardson number. I would have guessed that GDPS and CLM4 are singled out because they characterize stability by the Obukhov length, but that is the case for JULES also.

The vertical temperature profile is well initialized by construction.