Peer review of manuscript "Modelling Cold Firn Evolution at Colle Gnifetti, Swiss/Italian Alps" by Marcus Gastaldello, Enrico Mattea, Martin Hoelzle, and Horst Machguth, submitted for publication in *The Cryosphere*.

General comments

This manuscript presents simulations of firn conditions at a high-elevation site in the Swiss/Italian Alps using the distributed snow model, COSIPY. The evolution of firn temperature is interrogated over the period 2003-2024, focussing on conditions in the top 30 m. Spatial patterns of firn conditions are simulated using distributed meteorological data. A continuation of previously identified trends in englacial temperature and simulated surface melt. The role of surface melt in englacial warming is discussed.

In general, the manuscript reads well, the figures are well presented, and the methods are generally well described. However, the manuscript needs to clearly articulate the additional contribution to knowledge over previous work at the same site. E.g. beyond a continuation of the warming trend observed, what new insights are provided? How does the current modelling effort differ from previous work? Do we have additional statistical confidence in the trends occurring? To this end, the manuscript may benefit from a hypothesis (or series of hypotheses) to focus the manuscript on research questions of wider significance.

I also have some concerns about the model validation that limit confidence in the model results and need addressing:

- The model tuning is somewhat ad-hoc and the validation largely occurs after the interpretation of some key model results (e.g. temperature trends, spatial patterns). A thorough validation of the model using additional data should occur before the interpretation. This could include additional comparison to near-surface and deep temperature and density profiles. A more thorough sensitivity analysis of key variables would strengthen the interpretation of key results and help justify the parameter choices.
- At present the validation of deep subsurface temperatures mixes subsurface (percolation, heat conduction) and surface (new snow density, incoming longwave and albedo) processes as well as meteorological interpolation (e.g. lapse rate). Independent validation of surface energetics (e.g. through observed surface temperature or near-surface temperature) would help give confidence in the model choices. Figure 8 indicates there is systematic variation in subsurface temperature bias with elevation and slope/aspect, yet the effect of this on the spatial patterns of trends is not discussed.
- The effect of model spin-up on the simulations results needs to be treated in a more systematic way, given the weight given to the interpretation of the large model trends at lower elevations in the manuscript. Model sensitivity to key parameters needs quantified, perhaps through a table of sensitivity runs showing model performance and changes in key results (e.g. trends) for different model configurations.
- More detail is also needed on surface energetics, specifically incoming longwave and turbulent flux calculations, especially given their connection to the development of cooling during periods of clear-skies and low wind speed that are vital to the maintenance of cold firn.

Line comments

4 – please clarify if melt water has infiltrated the archive already or whether it could.

9 – a 'prolongation' – is unclear what is meant here – is the previously melt rate increasing or just continuing at the same rate?

13 – 'rotation' please explain what is meant in more detail here.

13 'in the uppermost 30 metres' – the results shown in the manuscript cover the top 15-30 metres, please revise.

20-24 – either introduce the definition of polythermal in the first sentence, or define polythermal with respect to 'cold' and 'temperate' ice in the second sentence.

100 – the method used to calculate lapse rates needs introduced here or at line 113.

106 – how representative are the relative humidity measurements given the elevation difference between he sites and the target area is 600-1000 m? How sensitive is the model to potential errors in this assumption given it will control the surface energetics through portioning turbulent flux between sublimation and sensible heat.

Table 2 – would be useful to include parameterisation of incoming short and longwave radiation parameterisations in the table.

122 – are the thermistor chain data used in the analysis? If not, please remove.

159 – how was a significant snowfall event defined?

175 – what values of $c_{emission}$ were chosen?

188 – were any limits placed on the Richardson stability parameter? If not, this could introduce unrealistic cooling during periods of low wind speed. Please discuss.

Table 3 – the values for cloud transmissivity coefficient (a,b) have a large bearing on modelled LW but aren't tuned to local observations. Please discuss

Table 3 – useful to include parameters h_{max} and z_interp1 and 2 (line 414) in the table as they are discussed later.

240 – Please indicate where the results of the tuning of z_{lim} are shown? If this the same results as in Figure 10? If so, the top 5 metres are not shown so is hard to assess the statement. Please discuss.

259 – please refer the observations of references that indicate the overestimation of thermal conductivity.

284 – given the importance of the model spin up, further detail about the method used to construct the spin up timeseries is need, including an assessment of uncertainty.

321 – what is the justification for excluding 2003 from the analysis? how extreme was this year compared to others at other sites with longer measurement series?

Figure 5 – the shading a transparency make it hard to clearly see the changes in melt as the color scale includes a graduation in shade. Please revise.

335 – ‘particularly evident at the ZS in Fig 6a’ – would be useful to give years referred to as is ambiguous in figure.

Figure 7 – a map of initial temperature in 2003 would be insightful here.

Figure 7 – please add the col-temperate transition line referred to at line 454.

356 – the negative model bias at lower elevations shown in Figure 8 indicates that the early period may be too cold – how might this impact the trends shown?

359 – ‘depth averaged’ what depths are averaged here?

381 ‘depth of seasonal variation (~15m)’. Figure 6 shows the depth as between 15 and >20 m. please revise.

Figure 10 – several profiles show a poor fit to the model results (year 2000 and 2021) – perhaps this is an artefact of the seasonal variation in near-surface profile. Plots of the simulated and observed temperature profiles would give more confidence in the model results.

Also, the much flatter profile below 20 m depth indicates that model is warming too much at depth compared to observations. Integrated over the depth, this is a large difference in energy. Please discuss.

421-3 – no validation of the internal temperatures profile is made except at SP, so care must be taken when extrapolating the model confidence to unvalidated areas of different slope and aspect. Please revise.

426 – the high-resolution thermistor results should be shown here to be relevant.

454 – the transition zone is not shown in Figure 7a. please add.

476 – no statistics are shown for the correlation of the MAAT and melt – please revise.

522 – a comment on the overly cold temperature at low elevation after spin-up is needed here.

533 – ‘a deviation in excess of 10 C’ – not shown on figure A1i. please revise.

536 – “The deficiency for the deep firn to retain thermal energy” do the authors mean the inability of the default model scheme to simulate the warming within the firn?

Editorial comments (*suggested changes in italics*)

210 ‘Monte Moro’ do you mean ‘Passo del Moro’

356 ‘area are as of result’ -> ‘*area are a result’*

371 ‘CG82-01/81’ -> *‘CG82-01/82’*

526 – ‘ thermal energy’ – ‘cold content’ or ‘thermal energy deficit’ are perhaps better terms to use here.