Dear Enrico Mattea and colleagues,

Your manuscript received positive comments, and after the new round of revisions the reviewer who had the most substantial comments (Adrien Gilbert) indicated that "The authors did a very good job in addressing my comments and concerns, I don't have anything to add.".

I went through the last version of the manuscript, and found the story to be very clear and well presented. The figures nicely support your findings – thank you also for incorporating the initial suggestions that I made, e.g. on Figure 2 – and I appreciate the 'honest' way in which your results are presented. The model capabilities are highlighted, but you are not 'hiding' what does not work and give an elaborate explanation concerning the possible discrepancies. I am convinced that some of your main findings (e.g. the role of micro-melt events) will be of large relevance to scientists working on firn modelling and those interested in deriving palaeoclimatic information from ice cores on alpine glaciers.

At this stage, I have formulated a list of final remarks and suggestions that I would like you to address when uploading your (most likely final) version of your manuscript. These are mainly easy-to-incorporate changes, although some may require a small amount of work:

- l.2: "...need to improve understanding and further develop" → "...need to improve our understanding and to further develop"
- l.37: order the references in chronological order
- l.102: "...the coupled model in..." → "...the coupled model is described in..."
- l.114-116: here you describe how other station data is used to perform quality checks, to fill gaps in time series and to determine parameters that were not measured at CM. This seems like quite a lot of work, which must not have been trivial. A few questions here:
    - Do I understand it correctly that you have performed this work yourself? If so, would be good to state this more explicitly
    - Can you say something about the "quality checks"? Potentially give some numbers
    - The new dataset you have created through your approach seems to be quite unique – it is definitely than what is directly available from the data providers. As this dataset would be quite valuable for other researchers, could you make it directly available? So far you mention "the meteorological time series...should be requested from the respective providers" in the "code and data availability section": but what you provide here goes "beyond" this.
- Table 1: what is CAE? Please define
- l.129: "...within the model input", maybe rephrase to "...as model input"?
- l.149: "...18 boreholes, some locations having been measured..." → "18 boreholes, where some locations have been measured..."
- l.184-186: "...corrected for...calibration parameters.". I got a bit lost in this sentence and could not understand 'what is what'. Could you rephrase this, potentially by splitting the explanation in two different sentences?
- Table 4: very nice and useful for the reader! In the caption, you mention "Additional parameters not listed here were kept at the default value". It would probably be useful for the reader to also have this information directly at hand, without having to dive into the three studies that you mention (and having to look for 'what is where'). Could you add this information in a separate table (e.g. in suppl. mat.)?
- l.313: you initialize by running the model over the time period 2004-2011 for 8 times. Two questions here:
    - Why did you decide to go for 8 times, and for instance not 4 times, or 10 times? Would be good if you could give a hint.
    - How sensitive are you results to running this time period 8 times? Could you include a short analysis on this, for instance in Appendix B?
- l.315: you mention '20 m / 1h' and '100 m / 3h' as spatio-temporal resolution. I found this quite confusing at first, as this made me think that you change two things at once: the spatial resolution and the temporal resolution: how can you then discern the effect of both on your results? Subsequently, I kind of understood that you probably change the time resolution to ensure numerical stability (correct?). My question here: would the results differ if you would run with '100 m / 1h' vs. '100 m / 3h' (where the former is obviously computationally more expensive). I would appreciate it if

you could provide some information on this, and potentially reconsider reformulating this to 20 m vs. 100 m, and only mention the time resolution separately, indicating that this is changed to ensure numerical stability (if this would the case of course, I'd gladly be corrected here, but I am trying to take away any possible source of confusion ☺).

- l.355-356: "in every month sublimation is a more effective energy sink than melt": ok, nice. Probably not that surprising of a finding for specialists in the field, but was for me at when reading the sentence at first. Maybe also consider mentioning this explicitly in your conclusion?
- l.360: "with only minor amounts...": could you quantify this statement? Would be useful, as is difficult to visually derive from figure 9
- l.364: "... it becomes statistically significant over the rest of the period": maybe good to be more specific here directly – to have info without having to refer to figure explicitly. i.e. "... it becomes statistically significant over the rest of the period (p<0.05 for 2004-2018)"
- l.370: "The majority of melt happens under clear sky conditions...". This contrasts with the previous sentence, from which I had derived that the cloudiness / sky conditions do not play a big role ("...unlike cloud cover which appears to have almost no effect"). Maybe consider slightly rewording? Could potentially remove the clear sky info and just focus on the "slightly positive temperatures"?
- l. 375: "...mean melt rates, slightly decreasing the likelihood of melt under high winds": could you add a sentence on why the likelihood is decreasing in this case?
- l.392-400: you explain how the density does not increase as long as melt-refreezing occurs at same location, and how your model is not accounting for this. Could you provide a hint somewhere about how this could be solved? Not suggesting that this needs to be changed, but would be good to provide a possible solution (like you do for some of the other limitations that you nicely put forward!)
- l.415: maybe reword to: "...be affected by the lack of SW radiation reflected from the surrounding terrain in the modeled SEB"
- l.442-443: "temperatures were initialized with repeated model runs over 2004-2011": see related question earlier. If you would have repeated this more (or less) than 8 times, how would this have affected your firn warming?
- "Conclusions and outlook": nice overview of your study! It would maybe be useful, for someone who has a quick look at the paper and directly looks at your conclusions, to state here which data is used for calibration/tuning, and which one for evaluation: e.g. for external reader it is not clear whether the fact that the firn temperatures are reproduced reasonably well is a result from the fact that you tune to this (i.e. calibration) or that this is just an outcome of your model without specific tuning to this (i.e. evaluation). Would be useful to shortly say something about this.
- l.555: 'code and data availability': would really be nice if you could also provide the processed series for GC directly, where the data gaps are filled and non-measured variables are derived from other station data (see comment on l.114-116)
- Please acknowledge the three reviewers (Adrien Gilbert, Vincent Verjans and one anonymous) in the 'Acknowledgments' section

Thank you for going through this final series of minor comments. Once these are addressed, we should be able to proceed to the acceptance of your manuscript.

Best regards,
Harry