

Summary

The authors describe a new modelling framework to quantify mass balance and frontal ablation with a joint calibration method of parameters. The model is applied to 71 glaciers in Svalbard, and the model output is compared to decadal observational data. The results show good accordance with observations with frontal ablation being slightly overestimated. I commend the authors on their thorough description of the methodology, an overall well written paper, and the development of a very useful tool for the community. However, I do have one major and some minor comments that should be addressed.

Major Comments

- In the limitations you state that the majority of the glaciers in your study are surge-type glaciers, yet your model is unable to capture surge-type behavior. This makes me question why surge-type glaciers are included in the study and how this inclusion impacts the results on a regional scale. I suggest to either exclude surge-type glaciers from the study and only use the remaining 35 glaciers or to differentiate surge-type/non-surge-type glaciers in the results.

Minor Comments

- Line 73: Reads slightly odd. Rephrase to “PyGEM simulates the climatic mass balance monthly for every 10 m of elevation...”.
- Line 73: Remove space at the end of the sentence.
- Lines 126 – 128: Maybe a workflow graph would be helpful here so that the reader can easily understand the integration of PyGEM and OGGM rather than having to follow this in the text.
- Line 134: Replace “the 4 x1 vector” with “a 4 x 1 vector”
- Line 203: You use ERA5 temperature data – Do you correct for the elevation difference between ERA5 grid cell height and glacier surface height? This might not change temperatures much, but it should be considered.
- Line 218: Remove space between parenthesis and 1.5%
- Lines 222 – 224: This paragraph should be integrated in section 4.1. It is directly applicable to section 4.1. but not to section 4.2.
- Line 226: Remove space between Svalbard and parenthesis
- Line 226: If the previous paragraph is integrated into this section, it is not necessary to state that this is an illustrative case and it should therefore be removed.
- Lines 228 – 229: I understand that the authors want to show the best case scenario, but I think it would be good to also show results that are affected by heavy data infilling (at least in the supplementary).

- Figure 1d: It is not quite clear why this panel is relevant as it is not referred to in the text. Either make clear why it is important or remove.
- Figure 2: Here would be a good place to compare the results of a single glacier with observational data as is stated in Lines 210–211.
- Line 257: Interesting. Can you explain what drives the per-glacier mismatches? Is there a common factor e.g. glacier geometry that causes the mismatches?
- Lines 257 – 258: I understand why you chose to only show one period (2010-2019), but it would be good to show both in the main manuscript.
- Figure 4/ Figure 5: I was a bit surprised to see that total length change is shown for each individual glacier, but frontal ablation is only shown cumulative for all glaciers. I think it would be good to show total length change, mass balance and frontal ablation for all glaciers individually as well as cumulatively.
- Lines 299 – 300: Given that SERMeQ only outputs terminus positions along the flow line, it also means that you do not account for changes in glacier front geometry i.e. if the glacier retreats faster on the lateral boundaries. Could you implement a different methodology in SERMeQ e.g. the curvilinear box method? Do you think this makes a difference in your frontal ablation estimates?
- Line 313: Do you exclude these three glaciers in addition to the 9 glaciers that did not converge? Please clarify.
- Lines 319 – 320: Dashes appear quite long. Replace with normal dash (–).
- Line 322: Again, is that glacier removed in addition to the nine glaciers that did not converge? Please clarify.