The Referee's comment below are in italics, our answer in plain font

The paper by Goose et al. applies the OGGM model to test glacier length changes using simulated climatology and compares the results to geologic observations. I have provided my technical comments as a .pdf. Overall, I find the paper easy to read, the figures to be mostly clear, and the general conclusions interesting enough to be published in Climate of the Past as it will be of interests to a broad readership to data generators and modelers alike. There are still some open questions that I've provided to the authors, but overall I think this paper will make a nice addition to the literature. I enjoyed reading it and as a data collector I find what they are modeling and testing to be of interest.

We would like to thank the Referee for the positive evaluation of our manuscript and for the suggestions. The technical corrections will be implemented in the revised version. We will also include all the suggested references in the revised version, except from the suggested citation on page 8: 'Gibbons et al. 1984 Geology'. We do not plan to include the later one as we were not able to find a corresponding reference that would match with our text.

Our answers to the general points included in the supplement are:

Page 8: It would be useful to include some of the Solomina et al. 2016 interpretations onto one of the figures. At the moment the connection between this part of the text and the figures is hard to understand. Also, Solomina et al. is a general review and I would suggest citing some original literature for the Alps from which they base their conclusions.

Fig. 5 of Solomina et al. (2016) clearly shows the periods of global glacier advances but, when focusing on a region like the Alps, the signal becomes very noisy (see figure 4 of Solomina et al. 2016) because of the smaller number of records, the discontinuities in those records and the difficulty to interpret some of them in a quantitative way. We have tried to include the information provided in Fig. 5 of Solomina et al. (2016) on our Figure 5 but with very limited success. Consequently, we decided to not include it in the revised version information directly on the figures but to modify the text to make the interpretation more straightforward. We will also cite some of the original literature as suggested.

Page 9: I agree that the investigation period is shorter, but the temperature changes is much larger. I was very surprised to see that the relationship between the summer temperature and the glacier length breaks down over the last 100 years (i.e. Figure 6a). I think a bit more explanation is warranted here since there appears to be no strong relationship between the warming, which is the largest of the last 1000 years, and the glacier lengths.

Do you really think it is because of the short period of investigation given that glacier retreat across the world is well documented (e.g. Oerlemans, 2005)? I guess I'm generally confused why this is the case but perhaps I'm missing something. This seems like a major point to address if you are going to conclude that there exists a dominate temperature control on glacier length.

The changes between 1700-1850 and 1970-2000 are larger on average by a nearly a factor 2 than the changes between 1970-2000 and 1900-1930. In addition to this larger signal, the longer averaging period allows to reduce the contribution of internal climate variability and the impact of the long response time of the glaciers. The variability of precipitation is also much higher for short periods than for longer ones. This explains why the link between temperature and glacier length appears relatively clearly for the longer period while it seems weaker for the shorter ones.

Nevertheless, the influence of temperature is still strong over the 20th century. The large majority of the glaciers are retreating over this period in the simulations driven by all the climate models (Sup. Fig 1), except in some members of CESM that have a very low warming. Consequently, our results are compatible with the observed glacier retreat across the world. It is just that the link between temperature and glacier length is more direct for longer timescales than when comparing two 30 year periods over the 20th century. This will be explained in more details in the revised version of the manuscript, adding a specific paragraph devoted to this point:

‘The warming over the 20th century has a clear impact on glacier length, inducing a simulated retreat of nearly all the glaciers in agreement with observations, except in some experiments driven by CESM members that display a weak temperature increase over the Alps (Fig. 3 and Sup. Fig. 1). Nevertheless, the contemporaneous temperature does not appear to be the only variable driving the glacier length changes when comparing two 30-year period at the beginning and the end of the 20th century (Fig. 6a). The contribution of temperature is present but the response time of the glacier as well as the influence of precipitation variability, for instance, can still obscure the link between temperature and glacier length for those relative short periods.’

The panels in Figure 2 are difficult to see in both the printed and .pdf versions. I would suggest finding a way to expand them to help the reader see them better.

To clarify in Figure 2 caption - the shaded region is the range and not the standard deviation about the mean?

Also, it a little disconcerting that the CESM ensemble are so far off the others during the last 100 years. The reason is obvious when looking at the temperature reconstructions (Figure 1) where the ensemble produces a late response to the abrupt climate warming. This makes me wonder about the utility of the CESM ensemble in general.

For the revised version, the panel of Figure 2 will be expanded.

Yes, as mentioned in the caption, the shaded area is the range of the ensemble.

We had no a priori in the selection of the models except the data availability and the absence of spurious trends or discontinuity in 1850 (see the method section). The difference between the results of OGGM driven by CESM results and observations is for us an interesting conclusion and including CESM results shows the interest of the proposed approach to evaluate the skills of climate models. We should insist that the performance of CESM is good for many characteristics of the climate of the last millennium (Otto-Bliesner et al. 2016) and there were no reason to suspect this behavior. For the ensemble itself, having an ensemble for all the models would be the ideal solution but it is unfortunately not the case. Nevertheless, different processes may control the amplitude of the internal variability, as measured by the range of the ensemble, and the forced response governing the trends over the past two centuries. The range of CESM provides thus one estimate of the role of internal variability but this is only one estimate and it must be taken with caution as discussed in section 3 (Page 7, line 16-20 of the submitted manuscript).

Figure 5 - I don't find this figure particularly helpful and overall it is a very confusing figure.

We agree that Figure 5 is not as clear as the other figures but it is helpful in order to make the comparison with equivalent figures produced from observations (as in Solomina et al. 2016). It also shows that the timings of advances and retreat are strongly influenced by internal variability so we propose to keep it in the revised version.