

The Cryosphere Discuss., author comment AC1
<https://doi.org/10.5194/tc-2022-69-AC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Matteo Guidicelli et al.

Author comment on "Spatio-temporal reconstruction of winter glacier mass balance in the Alps, Scandinavia, Central Asia and western Canada (1981–2019) using climate reanalyses and machine learning" by Matteo Guidicelli et al., The Cryosphere Discuss., <https://doi.org/10.5194/tc-2022-69-AC1>, 2022

We would like to acknowledge the reviewer for this thorough and critical review that has helped us to sharpen the focus of our study.

In the following, we report our responses (bold) to the reviewer's concerns (within quotation marks).

"The problems begin with the title, which overstates its importance. Only a tiny fraction, in fewer than half of the continents, of the world's glaciers are examined. The manuscript has too many figures and tables. The manuscript is supposed to be within 12 journal pages for TCD. The tables and figures alone, most of which occupy a full page, would take up this much space. The figures are bloated. For example, there is no need to illustrate "Tree 1" nor "Tree N", both of which are identical in Figure 3. The PCA section (4.1) doesn't tell the reader much more than the fact that elevation is the most important downscaling predictor.":

Title: We agree that the term "world's glaciers" can be misleading. In response to this comment we will change the title to: "Snow accumulation over glaciers in the Alps, Scandinavia, Central Asia and Western Canada (1981-2020) inferred from climate reanalysis and machine learning"

Number of figures and tables: We agree that some simplification is beneficial to the paper and we will accordingly perform major changes including a reduction of the number of Figures / Tables, as well as their content wherever possible and briefly described in the following:

Tables 2 and 3 will be moved to the Supplementary material. Fig. 2 will be moved to the Supplementary material as well. Fig. 5 could also be moved to the Supplementary material; even though it shows that other predictors than elevation are important to explain different biases between reanalysis' precipitation and snow accumulation on glaciers.

We also agree that Sec. 4.1 needs to be modified in order to better quantify the added value of each group of predictors on the model's performance. In the revised version of the paper we will show the changes in terms of overall model performance when suppressing the downscaled predictors (and/or other

predictors (e.g. topographical)). In fact, this might be a better evaluation of the predictors' importance than only showing the frequency of use of the main predictors (Fig. 4a and b) and their correlations (Fig. 4c and d).

Fig. 3 will be simplified and replaced by a smaller figure without the illustration of the "Trees".

"The leave one out validation is problematic as there is no independent validation dataset used, meaning that biases in precipitation are unlikely to be identified.":

Many thanks for this thought. However, we do not fully agree with this statement. For the "site-independent GBR", the model is always validated on a glacier that is independent from the model's training. Thus, as stated in the manuscript, the leave-one glacier-out cross-validation allows evaluating the generalization of the machine learning models for glaciers located in the same regions of the training data. Fig. 9 shows a more robust validation, where the performance of the machine learning models is also evaluated for completely independent regions (removing neighboring glaciers from the training data). Biases of reanalysis's precipitation against snow accumulation data (based on ground measurements and extrapolation techniques (see Sec. 2.2)) on the glaciers of the study are therefore identified (see Figs. 6 and 7).

Despite the glaciers used for validation being independent from the GBR model's training, it is true that they have an influence on the choice of the optimal hyperparameters of the GBR model, i.e.: the GBR model was optimized to perform well on the validation data. However, each single glacier (1 out of 95 glaciers) used for the validation has a very limited weight on the overall performance (mean squared error) and on the choice of the GBR's hyperparameters.

In order to make the proposed method even more robust, we will also define the hyperparameters independently from the test sites, i.e.: in turn, each glacier will be used to test the GBR model trained and validated (k-fold cross-validation for the selection of the hyperparameters) with the other glaciers.

"ERA-5 and MERRA-2 reanalyses are used without any mention of their potential large biases in the mountains. For example, Liu and Margulis (2019) report that MERRA-2 underestimates snowfall (which is based on the "PRECTTOLAND" variable used here) by 54% in High Mountain Asia.":

We are fully aware of the limitations of Reanalyses (because of missing and/or highly inaccurate in-situ observations) in high mountain region and specifically precipitation. In fact, our whole study is in principle motivated by this major challenge of improving the quantification of high altitude (solid) precipitation and SWE. In the current manuscript. Reanalysis biases in high-mountain regions are thus clearly mentioned including references in the introduction (lines 60-67). However, we agree that the biases observed in previous studies have not been described and quantified abundantly enough. In the revised paper we will better include them in the introduction thus enhancing the comprehensiveness of the manuscript. We will also add respective reference in the revised manuscript (e.g. Nitu et al, (2018), Zandler et al. (2019)).

References:

- Nitu, R., Roulet, Y. A., Wolff, M., Earle, M., Reverdin, A., Smith, C., ... & Yamashita, K.: WMO Solid Precipitation Intercomparison Experiment (SPICE). Tech. Rep., World Meteorological Organization, 2018. a., <http://hdl.handle.net/20.500.11765/10839>
- Zandler, H., Haag, I., and Samimi, C.: Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions, *Scientific Reports*, 9, 15 118, <https://doi.org/10.1038/s41598-019-51666-z>, 2019.

"It's not clear to me that the downscaling techniques presented here will correct that bias, as no independent evaluation of precipitation is presented.":

Reanalysis's precipitation is compared against snow accumulation data on glaciers. This data clearly is independent, and it is to our knowledge the only and thus best possible source of (cumulative) precipitation at very high elevation. The machine learning model is trained and validated against these snow accumulation data on glaciers. In general, from the results presented in the manuscript (e.g. Figs. 6 and 7) it is clear that, on average, the machine learning models can adjust the reanalysis' bias against snow accumulation on glaciers, which is among the main purposes of the study.

"Melt and sublimation are ignored in the "winter mass balance," which is then the wrong term.":

We do not fully agree with the reviewer here. The term "winter mass balance" refers to the snow water equivalent found on the glacier close to the maximum of snow depth, or the end of winter. Therefore, the winter mass balance – per definition – includes loss terms such as melt and sublimation, although they are not individually quantified. Furthermore, our periods of analysis are adjusted to optimally match the period where the components of melt and sublimation are small in comparison to accumulation by solid precipitation.

"After carefully searching through the text, I still cannot understand how precipitation phase was treated. It seems to have been ignored as SWE is used interchangeably with the downscaled precipitation on glaciers. But then, in Table B1 and B2 ERA-5/MERRA-2 snowfall variables are listed as predictors?":

Indeed, the precipitation phase was ignored. In the revised paper we will more clearly describe this choice, and also why we think that this simplification is justified, as briefly summarized in the following.

We adjusted the total precipitation variable of the reanalysis ("tp" for ERA-5 and "PRECTOTLND" for MERRA-2 (see Sec. 2.1.1 and Sec. 2.1.2)). We are aware that a different adjustment factor of precipitation might be needed depending on the precipitation phase. However, as we only adjust the total precipitation occurred during the accumulation season, the adjustment factors represent the "average" adjustment factor of all precipitation events. The snowfall variable was used as a predictor in order to give the chance to the GBR model to learn that a different

“average” adjustment factor should be applied depending on the proportion of snowfall against total precipitation (i.e. depending on the main precipitation phase during the accumulation season).