Author response to Comment RC3 on gi-2021-16

We thank the reviewer for her/his time to assess the manuscript and for the valuable feedback suggestions. Below, we respond to each suggestion and comment one by one. The reviewer comments are highlighted in blue while our responses are kept in black.

In their manuscript “Evaluating methods for reconstructing large gaps in historic snow depth time series”, the authors compare different methods for filling large gaps in measured snow depth time series in Switzerland.

The manuscript is very well written and of high technical and scientific quality. The comparison of the presented methods is in general of interest for the respective snow hydrological community, however, in my opinion only for a very small number of real use cases. The authors show that already a very simple snow model approach using measured temperature and precipitation as input can yield more or less the same results. There are much more temperature and precipitation measurements available than snow depth observations, especially in data sparse region. For that reason, I don’t see very much applicability of the results.

[Answer]: We only partly agree. Yes, we cover quite a specific use case with our study, which is the reconstruction of long-term gaps in historic snow depth time series. We indeed believe that for this and similar use cases our method comparison is of large value to the scientific community. For example, the homogenization community is often confronted with the problem of such gaps in the step of the break-detection (see the newly added paragraph at the end of the discussion section). Solely the conclusion that a simple temperature-index snow model is able to represent a decent amount of variability when it comes to reproduce snow climate indicators such as HSavg, HSmax or dHS1 is in our opinion of value. Moreover, there is an increasing number of cases in high alpine environments, where either precipitation is not measured or the measured precipitation amount is strongly limited due to under-catch, but on the other hand a large number of neighboring snow stations is available due to specific needs like avalanche warning.

As the authors state, HS is a good-natured variable for gap-filling. This holds true for measurements in terrain where the presented stations are usually located and for continuous snow coverage and typical seasonal, continuous accumulation and ablation dynamics. Therefore, it is quite obvious that good results can be obtained using statistical interpolation methods (more or less regardless of type) using neighboring stations of similar elevation. Much more interesting would be an extension of the analysis to terrain characteristics (lateral snow redistribution, steep terrain, slope, aspect, i.e. small scale heterogeneity in mountainous terrain). This could be tackled by connecting the presented methods to stations clustered not only by elevation and distance, but also slope, aspect, etc. However, I see that this is probably not possible due to the stations located at “representative”, flat, unobstructed terrain locations.

[Answer]: This is a very good point because with the extension of the methods to consider terrain characteristics, it would be also possible to interpolate snow to areas where we do not have station information. Unluckily, as you already write, we lack the necessary data to train any method of that kind since the stations are located at sites that do not differ too much from each other regarding slope or aspect. However, we ultimately are interested in getting continuous snow depth time series at a station location. Therefore, despite being interesting, these questions should maybe rather addressed in another study and are out of scope of our study.

Regarding the results of dHS1, it would be interesting to see the same analysis for dHS10 or dHS5, i.e. a threshold for a snow day of 10 or 5 cm snow depth, as 1 cm is within a range of errors/uncertainties of all measuring and modeling methods. Probably the results will be much clearer using a slightly higher threshold.

[Answer]: We calculated the number of snow days for thresholds of 1, 2, 5, 10 and 30 cm (dHS1, dHS2, dHS5, dHS10, dHS30). We added a paragraph to the discussion regarding the effect of different snow cover day thresholds and included a table (Table A1) in the appendix.

As pointed out by reviewer 2, the study would highly benefit from an additional comparison to derivates or direct model values from, e.g. reanalysis products, which are readily available globally.

[Answer]: Please see our answer to RC2 regarding the use of reanalysis products. An assessment of downscaling methods for snow depth or temperature and precipitation from different reanalysis products would get to broad and would be out of scope of this paper. Nevertheless, we added a few sentences to the discussion section where we discuss the potential of using reanalysis data but refrain from including it to the results.

I support the idea raised in the other comments of including a table with particular strengths and weaknesses of the methods depending on the application and data availability.

[Answer]: We added an overview table (Table 3) to the discussion which summarizes the suitability of different methods in different situations (dense and sparse station networks, gaps in neighboring stations) and for different applications.

Apart from the – in my opinion – rather low applicability of the presented results in other scientific use cases, the article presents a technically well performed study. The findings and conclusion are presented in a very clear and concise way.