Response to Anonymous Referee #1

The authors have done a commendable job responding to the reviewer comments. I think the addition of the ERA5 family reanalyses, as well as the more in-depth analysis of the various product resolutions and their impacts, helps understand the added-value (or not) of the global datasets. The text is also more nuanced regarding strengths and weaknesses of various products. Overall, I am satisfied with the quality of this new and improved version of the paper.

One point still bothers me somewhat, however. The authors acknowledge that the observed reference precipitation can be affected by undercatch, and a correction factor is applied. However, this method is pretty coarse and brings lots of uncertainty. Then, in the text around figure 3, the reference dataset is taken as the "truth" to which the other products are compared. I think it would be wise to restate that the results are relative and that it is not known (at this point in the paper) which ones are closer to the reality.

I think this can be easily resolved with a sentence or two in the text, which is why I recommend minor revision which can be handled at the editorial board level next.

Two additional sentences in ll. 283-284: “Although we consider the measured data to be the reference, it is not sure at this point of the paper that it performs best in modelling the snow cover. Therefore, the presented results are relative.”

Good work by the authors and an overall very interesting paper.

Authors’ answer: We once again thank the reviewer for his/her constructive feedback and very much appreciate the effort he/she took.

Regarding his/her comment: We agree, and added the sentences as suggested.

Response to Anonymous Referee #2

General comments:

The manuscript is much improved from the initial review and easier to follow. Besides the specific comments to follow, I would recommend focusing on key findings in the results, perhaps generalizing by category, and then discussing the significance of the findings with respect to model choices. While I realize that a full discussion of consideration of basin size, roughness, elevation, etc relative to available datasets in beyond the scope of this paper, the community would benefit from an organized discussion of these factors. This is hinted at when discussing the results of other research but could be made much more explicit. For example, rather than start with "Study A" exhibited better results, consider starting with the theme "basin size and elevation matter" and then demonstrate this with case studies including results from this work. This kind of organized discussion would make this a much stronger and more interesting paper.

I suggest that the authors place their findings into a framework of considerations others may find useful when selecting meteorological or topographic datasets. This would be one of the more valuable contributions of this paper.

Authors' answer: We once again thank the reviewer for his/her constructive feedback and very much appreciate the effort he/she took.

In our opinion, it is currently not possible to implement an encompassing profound framework in the discussion part including findings of other studies, as similar studies explicitly for high-alpine areas in various other regions are still lacking in literature. However, we definitely agree with the reviewer that such a framework would be highly valuable and should be established as soon as more studies on comparing various global products explicitly in different high-alpine regions come up. This goes also in line with our general remarks and suggestions for next steps in Section 4.3, where we suggest to set up such comparisons at further similar study sites, e.g. of the INARCH network. We added several points in the discussion in this regard - please consider our answers to your last three specific comments for more details.

Specific comments:

Table 1. Are all measured parameters at DWD really available since 1900? Qsi and Qli were probably added more recently?

Authors' answer: You are right, radiation data was recorded since 2009. We changed that in the Table.

Figure 1. What is UFS? Adding locations of glaciers would help readers see where snow depth is measurement on and off glacial surfaces.

Authors' answer: We added the explanation of UFS, which is the Environmental Research Station Schneefernerhaus to the figure caption. We also added the glacier surfaces as suggested to the map.

Figure 2. What is the large blank area? This seems significant as this appears to be where one of the snow depth measurements is taken.

Authors' answer: The white area is HRU10, as it is explained in the legend of this figure. HRU10 represents the glaciated areas on which DWD snow depth is measured. The added glaciated area in Figure 1 should help to clarify this as well now.

Line 196. What does it mean to run the model in "gauged basin mode"? Does this mean that parameters are adjusted to match outflow? There is mention of a stream gauge, but it is not clear how or if these data are used. Furthermore, how well does the model work? Summarize findings of Weber et al 2020.

Authors' answer: "Gauged basin mode" refers to the model run with in situ measured data. It has not been adjusted to match the outflow. In terms of IAHS, the term "gauged" does not only refer to a stream gauge but to any measurement device in a basin such as temperature sensors, snow gauges, and precipitation gauges etc. As written in section 3.4, we do not use measured runoff in our study due to data gaps and the general poor data quality. We added a sentence in ll. 199-200 to prove the achieved good model accuracy of Weber et al. 2020: "Measured snow depth could be modelled with an accuracy of NSE > 0.7 (Nash-Sutcliffe-Efficiency, (Nash and Sutcliffe, 1970))."

Line 205: Similarly, what does “ungauged basin mode” mean?

Authors’ answer: Analogues to the previous point, in the “ungauged basin mode” no in situ measured data are available for modelling. Therefore, other forcing data have been used, as explained in this sentence. We refer to the first sentence in the introduction in ll. 29-30, where we explain the meaning of ungauged. For clarification we added the subordinate clause to l. 209: “in which we assume to have no measured model forcing data”.

Section 2.4: Are the different DEM’s used with the reference simulation climatological data? If so, what parameters are adjusted as a result. Section 3.2 states that the reference simulation was used and that it was explained in section 2.3. This however is not the case.

Authors’ answer: All different DEMs are used with the reference meteorological data as explained in Section 3.2 (ll. 330-332). Also as explained in Section 3.2, as well as in Section 2.4, the adjusted parameters are altitude, slope and aspect, which in turn influence the meteorological data. Section 3.2 states that the reference meteorological data, which means the in situ measured data, are used and that these data are adjusted to the DEM specific HRU altitude, aspect and slope as explained in Section 2.3. In Section 2.3 it is described how the meteorological data is adjusted to the HRUs. We think this should be clear enough to the reader.

Line 324-326. There is no explanation of how reference climate data are adjusted using ALOS, SRTM, and GTOPO30 DEMs. I believe the referenced section (2.3) is incorrect. It should be section 2.4. Please clarify.

Authors’ answer: Section 2.3 is the correct reference in this case. In this Section, we explain how the meteorological data is adjusted to the HRUs and thus the DEM specific values of altitude, aspect and slope angle. The same method is also applied for the non-reference meteorological and DEM input data which is presented in the following.

Please consider ll.203-206: “The corrected data were then transferred to the HRUs following the method of Liston and Elder (2006) as well. This method uses monthly variable lapse rates for temperature (Kunkel, 1989) and a monthly variable precipitation adjustment factor (Thornton et al., 1997) for the altitude-dependent adjustment of temperature and precipitation. Relative humidity was adjusted via dew point temperature which has a relatively linear dependence on elevation (Liston and Elder, 2006).”

The radiation data is treated by CRHM internally which is explained in Section 2.2.

Figure 5. What is “(d) setup”? Is this the reference simulation and Lidar topography? Please clarify.

Authors’ answer: We changed the figure caption so this is clear now.

Discussion: Many of the findings with respect to climate products are not too surprising given their coarse resolution. This has been demonstrated several times in mountainous terrain. What would make this discussion much more interesting is if the authors examined factors that limited their use by topic such as basin size, basin homogeneity, basin elevation, area climatological variability, etc and then supported statements with their findings as well as those of other researchers.

Authors’ answer: We agree that several studies applied such global products; however, we are not aware of other studies, which explicitly compare numerous global products specifically for high-alpine regions as we did. In the discussion, we tried to mention existing studies, which point somehow in this direction, but which might not be directly comparable as we discussed. The majority of these studies consider alpine regions including their forelands and not explicitly only high alpine sites, e.g. above the tree line and highly characterized by complex terrain.

In our opinion, it is currently not possible to compare and categorize the findings of our study with others in general for a sound framework for high alpine catchments. As this might have been unclear in the discussion, we added the following in ll. 515-518: “According to the knowledge of the authors, studies comparing different global meteorological input data explicitly for high-alpine regions are very sparse and not available in the extent of this study. Therefore, a framework-based comparison of different studies in this regard is not possible at current stage. In the following, however, we discuss further studies, which point to a certain extent in this direction.”

Moreover, at the beginning we included that we focused exemplarily on the RCZ (see ll. 505-506 "In the following, we discuss the potential applicability of global products for snow hydrological modelling in high alpine regions as exemplarily demonstrated for the RCZ in this study.").

As already stated concerning the general comment of the reviewer, we see great need for, e.g., a follow-up study including several high-alpine catchment of similar category than the RCZ to be able to make such statements. This would of course be helpful, as it would provide a clear guidance to action for the use of model forcing data in ungauged basins. Therefore, we added in ll.596-601: "This would enable to create a comprehensive framework of how to use globally available data for model forcing in ungauged high alpine basins. Regarding globally available meteorological setups, a framework categorization could include statements on catchment size and homogeneity, climatological variability as well as its topographic characteristics."

Line 545. How will one know if transferring data from a catchment within 100km will provide the best results? Again, if this were addressed in a framework mentioned in the previous comment, it would be more helpful.

Authors' answer: We agree that 100 km might have been a too speculative number although it was the case for the transfer of the meteorological data of Mt. Wendelstein. We deleted this number, but kept the statement that this is true for stations in the closer vicinity in the same mountain range with similar climatology.

Line 559. Again using a framework to consider the use of alternate topographic products would be helpful. For example, use of GTOPO30 might be ok if the area under consideration is well above the current snowline, but would be problematic in basins where the much of the snow accumulation area lies close to the freezing line (where small errors in Ta result in the wrong precipitation phase).

Authors' answer: We agree with the reviewer that also a framework regarding global DEM products for high-alpine catchments would be of great interest. However, similar to the global meteorological products, we think that for a comprehensive and profound framework regarding the topographic products it is not sufficient to rely only on the results obtained only in RCZ and as findings in other sites are still missing or very sparse. We included in Section 4.2, ll. 563-570 some literature findings for other than explicitly high alpine regions on different DEM resolutions in hydrological modelling as well as one study on glacier melt. The added references are the following:

- Hopkinson, C., Chasmer, L., Munro, S., and Demuth, M. N.: The influence of DEM resolution on simulated solar radiation-induced glacier melt, 24, 775–788, https://doi.org/10.1002/hyp.7531, 2010.
- Nagaveni, C., Kumar, K. P., and Ravibabu, M. V.: Evaluation of TanDEMx and SRTM DEM on watershed simulated runoff estimation, Journal of Earth System Science, 128, 73, https://doi.org/10.1007/s12040-018-1035-z, 2019.
- Sørensen, R. and Seibert, J.: Effects of DEM resolution on the calculation of topographical indices: TWI and its components, Journal of Hydrology, 347, 79–89, https://doi.org/10.1016/j.jhydrol.2007.09.001, 2007.
- Vaze, J., Teng, J., and Spencer, G.: Impact of DEM accuracy and resolution on topographic indices, Environmental Modelling & Software, 25, 1086–1098, https://doi.org/10.1016/j.envsoft.2010.03.014, 2010.

For a proper comparison in high alpine regions, it would require, similar to the meteorological products, a much broader choice of high alpine test sites with different topographic characteristics. Such studies should especially consider the degree of topographic complexity. Therefore, we added the following suggestion regarding the DEM products in ll. 599-601: "Regarding the usage of global DEM products an investigation especially on the degree of topographic complexity and especially the steepness of terrain would be most interesting."

Response to Anonymous Referee #3

We once again thank the reviewer for his/her constructive feedback and very much appreciate the effort he/she took. In the following, we give an individual answer to each comment.

1) Need thorough improvements in writing, especially to fix grammatical errors.
Authors' answer: We gave the manuscript to a native speaker for proofreading of this revised version and mentioned explicitly to have also an eye on grammar issues.

2) Line 4 in Abstract: Specify the 10-year period (Sept 2000 – Aug 2010)
Authors' answer: We added the time period (l.4).

3) Line 126: convert 'ha' into 'square kilometers'. I'm not sure if 'ha' is an MKS unit.
Authors' answer: We changed it to km^2 (l.126).

4) Line 178: 'specifically' may be better than 'especially'.
Authors' answer: We changed it to 'specifically' (l.179).

5) Line 308: 'data' may be better than 'curves'.
Authors' answer: We changed it to 'data' (l.306).

6) Line 320: What is 'internal' variance? Please explain 'internal'.
Authors' answer: We rewrote the sentence (ll.326-327): "The variance of slope and aspect angles in the coarse GTOPO30 DEM setup is very low."

7) Lines 367-403: This is a huge paragraph that contains discussions on two separate subtopics, the sensitivity to the met data (367-395) and the sensitivity to the topo data (395-403). Dividing the paragraph into two paragraphs (for the met data and for the topo data) will it make easier to capture the key statements in this paragraph.
Authors' answer: We agree and we divided it into two paragraphs according to the suggestion (ll.373-410).

8) Related to (7), Figure 6 may also be split into two like; Fig. 6a to present the met data sensitivity and Fig. 6b to present the topo data sensitivity.
Authors' answer: We split Figure 6 into a) and b) regarding meteorological and DEM data, respectively, as the lines were obviously too close together. We also added a) and b) in this regard in the Figure capture.

9) L149:151: Figure 9 may also be split into two like; Fig. 9a for the met data sensitivity and Fig. 9b for the topo data sensitivity.
Authors' answer: For Figure 9, however, we do not see the necessity to split it, since the signature of the bar plots clearly separates the met and topo data.

10) L495: Comparison of the topo data sensitivity to the met data sensitivity is largely irrelevant. May change 'all other setups' to 'all other DEM setups' to clarify that this statement refers to the topo data sensitivity experiment.
Authors' answer: We changed it to 'all other DEM setups' (l.502).

11) Section 5. Conclusion is another huge paragraph with mixed subtopics. I suggest to split it into 4 paragraphs; Lines 593-605, 605-617, 617-631, and 631-641.
Authors' answer: We agree that the readability of the conclusion improves and split this paragraph according to the suggestions of the reviewer (ll.620-669).