Dear editor and referee #1,

We thank referee #1 for his/her very constructive and helpful comments. We appreciate that the referee has understood our aim of assessing the predictive power of our snow evolution model using downscaled bias-corrected climate forcing over the instrumental period to get an overview of forcing-specific issues when it comes to future projections of glacier's surface mass balance (SMB) in **glacier change impact studies**.

When comparing with the comments of other two referees, we have realized that some issues raised by the other reviews and even some conflicting ideas between this review and the other two may have originated from very different scientific cultures in our native field - glaciology – and other fields, e.g., hydrology, climatology, etc. To accommodate these "cultural" differences, we have decided to substantially re-frame the manuscript to focus on an in-depth present-day evaluation of different bias-corrected climate forcings dynamically downscaled by the Regional Climate Model (RCM) for glacier change impact studies.

In this way, we are not simply presenting the results from the benchmark and CORDEX-driven simulations as it was in the initial version of the manuscript but will zoom into the differences in the SMB model reconstructions, their intrinsic drivers and what they mean for the performance of bias-correction methods and downscaled climate model products. We have therefore decided to modify the title of the manuscript to 'Synopsis of the uncertainties introduced by bias-corrected climate forcings in regional glacier surface mass balance evolution studies - A case study using a CORDEX chain envelope in western Norway' to re-define the focus of the article.

In response to the referee's comments and those of other referees, we will provide significantly more details on the methodology used, including the validation/evaluation of the benchmark model results and assessment of the results of the simulations driven by CORDEX. Below we are listing specific responses to the major questions of referee #1 (written in light blue and *italic* font):

1. I am not sure how they actually modelled the surface mass balance of the glaciers. As I read the methodology presented in section 2 it seems to me that SMB is not explained here as (glacial) Surface Mass Balance, but it seems more to be described as the Snow Mass Balance. From section 2 onwards to the final sections this was what I interpreted this work to be. I kept forgetting that SMB was glacial mass balance, and I had to hop back the introduction to remind myself what this study was about. The methodology is a very clear description of how the snow balance is treated, but it seems like a subsection of how the glacier surface mass balance, or the climatic mass balance (van Pelt et al, 2019) was calculated. A clarification of this would be a necessary addition to the manuscript. There are some physics in addition to the snow model described in section 2, that usually is applied when calculating the CMB of glaciers (Hock, 2005). I would recommend expanding section 2 with new a sub-section that describe the physics used to calculate the CMB and reference the used methodology. Another possible reference may be to look into Huss et al (2008): Modelling runoff from highly glacierized alpine drainage basins in a changing climate. Hydrol. Process., 22(19), 3888–3902 (doi: 10.1002/ hyp.7055).

We agree with the referee that our methods section, including the description of our approach to calculating the glacier's SMB (defined as a difference between the precipitation that has accumulated on the glacier surface and what has been lost due to melt and eventual runoff and sublimation), needs further clarification and expansion. We are now providing a subsection in Sect. 2 where we detail different components entering our calculations of snow water depth, glacier ice melt and SMB. Also, in response to the other reviews, we have further expanded the methodology section where the forcing procedure and forcing datasets are presented in a greater detail.

2. Another question is how useful the CORDEX data was in this study. The comparison of the NORA10 and the different CORDEX datasets was interesting, showing NORA10 seem to beat the CORDEX data on most of the parameters tested for. I am not sure how meaningful the continued use of CORDEX is after seeing the results in Figure 3, with their large RSMEs. Or is the use of CORDEX *in interest for driving the model in future scenario mode? If not the CORDEX output is well argued to be important here, the space and number of figures can be substantially shortened.*

After having read and discussed the reviews, we have fully realized that the aims and objectives of our study could be significantly strengthened by refocusing on the potential of the present-day products of different bias-corrected climate forcing dynamically downscaled by RCMs to produce realistic results in glacier surface mass loss impact studies. As part of this analysis, we argue that the remaining high bias in CORDEX outputs after bias-corrections doesn't mean that one should not use these products, but rather account for the errors that come with such biases and strive to introduce regional improvements in the climate forcings that are used to model future changes in glaciers and their impacts in Norway. These do not only include general trends in the surface energy/precipitation changes, but also the direct impacts of partially or entirely missing weather patterns (as pointed out by referee# 2) in raw climate model outputs.

Existing articles on bias-corrections of RCM outputs have pointed out that all bias-correction methods have their limitations (e.g. Maraun, 2016; Holthuijzen et al., 2021). In this study, we aim to address these limitations and quantify the uncertainties they might bring to impact studies. We agree that the way we present the results should be improved in the revised manuscript by reframing and deepening the analysis in Sect. 3 and 4 to not simply present the results from the benchmark and CORDEX-driven simulations as it is in the current manuscript but elaborate on potential mechanisms, drivers, and long-reaching consequences of inaccuracies in climate forcings.

3. A third issue is the validation of the model output with the seNorge data in section 3.1. As I understand the seNorge data is model data, and is calibrated with, or have assimilated observational data in the model input. Although I would guess the hindcasted NORA10 as well as the CORDEX data, both of the products from HIRLAM, may use assimilated observational data in the hindcast mode. With this I see a question with validating modelled data from NORA/CORDEX (yi, ...yi+x), with modelled data from seNorge (x). Maybe the observational data in seNorge has a larger weight than in they have in the NORA10/CORDEX simulations, but that needs to be stated. One way to manage this is to use the observational data, or the nodes in seNorge that are anchored to observational data to manage a cross correlation check. That is, using only the pixels / nodes where seNorge has observations, and where the observation bias should be weighted highest in the seNorge output. Although I think it is now possible to download the observational data from the seNorge webpage, if the raw data of the observations is wanted for a correlation test.

Based on this comment, we have realized that different forcing datasets used in this study and why we chose them for this analysis should be described in more details in Sect. 2.

Strictly speaking, seNorge is not based on data assimilation; It contains grided meteorological data statistically interpolated from measurements of all the weather stations in Norway. To reconcile the differences in the opposing views of referees from glaciology, hydrology and climatology mentioned above, we are now zooming into the analyses of these datasets from the interdisciplinary perspectives. It has been stated by referee #2 that only overall statistics can be compared between downscaled CORDEX data and the observation. Thus, we don't think direct comparison of downscaled CORDEX data on specific dates should be performed against automatic weather station measurements at each anchored point, even though this is traditionally practiced in glaciological studies (e.g. van Pelt et al., 2012, 2019). To reconcile these differences of opinion and to better reach the readers from other research communities, such as from the fields of hydrology and climatology, we should rather look at some statistical quantities at a glacier-wide or drainage-basin wide scale, even though it is still our intention to quantify to which degree (if any) future projections capture weather/sub-seasonal patterns relevant to the calculations of glacier surface mass balance and what a potential lack of sub-seasonal signals may mean for the reliability of future glacier projections. For the above purposes, it is more convenient and straightforward to use a grided dataset (seNorge) instead of in situ observations at individual weather stations.

Minor comments

Li 111. Precipitation into (?) our model.

We will change the sentence from ‘MicroMet is used to interpolate coarser-resolution RCM and reanalysis outputs … precipitation onto our model grid…’ to ‘MicroMet is used to interpolate coarser-resolution RCM and reanalysis outputs … precipitation into the model…’.

Li 276. Is Table S1 provided?

Yes. It is in the ‘Supplement’ section.

Section 4.1./Fig. 10. I do not follow the discussion with reference to the correlation matrix in Figure 10, probably because I am not sure what this matrix show. Is all melt from SMB calculated as runoff? What about (temporal) storage, evapotranspiration etc? The two lowest arrays roffw and roffa should they not be same as SMB runoff? I guess a few lines of text describing this figure would help to motivate this part of the results.

We have realized that we did not explain the composition of the runoff well enough. It consists of melt water from snow and glacial ice and rainfall. Evapotranspiration is not included here. But only the runoff results on the glacier covered the region will be included in the revised manuscript. Thus, Sect. 4.1 will be drastically changed, and Fig. 10 will be removed.

Li 410. Maybe add the reference here again of where you got the data of glacier cover, to repeat this to the reader or call the delineation you refer to in section 2 as glacier cover.

Agreed. We will add the reference there.

Li 484. Bondhuselva?

Thanks for correcting. It will be changed to Bondhuselva.

The response to the comment of referee #1 on the figures are followed. **Due to the substantial re-framing of the manuscript some of these comments might not apply to the new figures in the revised manuscript**:

Figure 1. Please add lines for each zoom-in picture that join the frames of the area in the bigger map.

Figure 1. Would it be possible to make the hydrography clearer in the zoom-in maps? You could add a blue streamline following the hydrography pointed out in each of the zoom out maps, and number them to follow the legend of the streams. That would make the zoom-out maps more clear and will make it easier to navigate in them.

Figure 1. The upper right zoom-out. Grå should be Grås?

We will improve Fig. 1 according to the three suggestions given by the referee.

Figure 4. Perhaps name the panels a-d. As now it is hard to follow the caption as what of the matrices are linked to what part of the description in the caption.

We will improve Fig. 4 according to the suggestion.

Figure 4. Do the two matrices in b) indicate reverse signs of the SMB between observed and some of the modelled data? That would be remarkable. I am not sure what these matrices show. Make this clearer, or it may be a source of confusion on the reader side.

We admit that panel (b) is confusing. The first row is the observation and rest are the difference between the modelled results and the observation. We will present this information in a different way in the revised manuscript.

Figure 5. The left side panels in this figure should me made with more contrast, and perhaps larger. As now it is hard to see what they contain.

We will improve the figure.

Figure 10. See comments above on Section 4.1.

We have realized that we did not explain the composition of the runoff well enough. It consists of melt water from snow and glacial ice and rainfall. Evapotranspiration is not included here. But only the runoff results on the glacier covered the region will be included in the revised manuscript. Thus, Sect. 4.1 will be drastically changed, and Fig. 10 will be removed.

Figure 11. Make it clearer in the captions that the 18 different catchments are ordered with respect to the glacier cover, and add their number 1 to 18 in at least one of the point distributions to make it more transparent where each of the catchments are representing which point.

Fig. 11 will very likely be removed from the manuscript.

Reference

Holthuijzen, M. F., Beckage, B., Clemins, P. J., Higdon, D., and Winter, J. M.: Constructing High-Resolution, Bias-Corrected Climate Products: A Comparison of Methods, Journal of Applied Meteorology and Climatology, 60, 455–475, https://doi.org/10.1175/JAMC-D-20-0252.1, 2021.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

van Pelt, W., Oerlemans, J., Reijmer, C., Pohjola, V., Pettersson, R., and van Angelen, J.: Simulating melt, runoff and refreezing on Nordenskiöldbreen, Svalbard, using a coupled snow and energy balance model, The Cryosphere, 6, 641–659, https://doi.org/10.5194/tc-6-641-2012, 2012.

van Pelt, W., Pohjola, V., Pettersson, R., Marchenko, S., Kohler, J., Luks, B., Hagen, J. O., Schuler, T., Dunse, T., Noël, B., and Reijmer, C.: A long-term dataset of climatic mass balance, snow conditions, and runoff in Svalbard (1957–2018), 13, 2259–2280, https://doi.org/10.5194/tc-13-2259-2019, 2019.