Response to Reviewer #1

We thank Juraj Parajka for taking his time to carefully read the paper and providing critical remarks.

General comments

This study presents an approach for calibrating different degree-day snowmelt approaches by using MODIS snow cover data. The second aim is to examine different degree-day variants for snowmelt simulations and calibrate or validate them using satellite snow cover data. The approach is tested in two regions (Baden-Württemberg in Germany and Switzerland). The results indicate a slight increase in overall NSE runoff performance and a better NSE performance during the winter period.

I read with interest the manuscript because we did numerous similar experiments in the past (and recently). I have to say that the manuscript presents some interesting and novel experiments, but as a whole, it is not ready for publication in its current form. The main reasons for such assessment are:

The Introduction section needs to be improved. In its current form, it is not specifically presenting which approaches are already available, what the research gaps are and how this research goes beyond existing studies? There are numerous studies (for example, please see some references below, and references cited in these studies) investigating and comparing different degree-day snowmelt models and studies investigating calibration of conceptual hydrologic models (their snow part) to MODIS snow cover data. The introduction needs to clearly present the research done so far and to formulate what the novel scientific contribution of this study is. In my opinion, a comparison of existing degree-day models is not novel. Nor a general use of MODIS snow cover data in hydrological modelling. Still, I think the study presents some interesting approaches which can be turned into novel research objectives, such as how many and which MODIS images are needed for robust calibration of conceptual snowmelt models.

The introduction of the paper is certainly not complete, but this is not a review paper. In our understanding, the purpose of a non-review paper is to provide new ideas and not to give a complete picture of the state of the art. If the paper contains facts that are already known and not referenced, then this should be pointed out by the reviewers. Recently publications have extremely long review like introductions and number of cited papers increased significantly in the past years. In the opinion of the second author, this did not improve the quality of the publications. However, with the comments from the reviewer, we can definitely extend the introduction if required.

The structure of the document/story is not easy to understand, and the clarity of the presentation can be improved. If the study's main aim is

to propose some novel approach/method, then I would suggest presenting it first and describing the study region and data later. This will allow the reader to understand the novelty and eventually to apply the general approach to other regions/models. I would also suggest presenting a general strategy at the beginning clearly. This will create a storyline and improves the clarity of the presentation. In its current form, there are many subsections and the order reads more like a summary of all technical works done but does not present clearly what the novel scientific contribution/research hypothesis is.

Thank you for this suggestion. We'll try to reconstruct a better narrative. If one is too close to the topic, the presentation may focus on things which are not well understandable.

The study needs to be more focused on the novel contribution. I'm not sure how the interpolation and its cross-validation contributes to the novel scientific findings in the field of using satellite data form model calibration? Perhaps the crossvalidation can be presented only in a supplement. The more interesting point is to analyse which MODIS images are needed for robust model calibration. I do not understand why not use all available images, particularly for model validation? How sensitive are presented results to the selection of dates of MODIS images? There should be a more detailed analysis and evaluation for supporting the results and interpretations made. It is also not clear why not to use the concept of the HBV for simulation of snowmelt accumulation and melt. Why is it needed to separate the degree-day part and then link it back with the hydrological modules instead of using it together (i.e. to calibrate only the snow module first and then apply the complete model)?

The cross validation of temperature and precipitation is not a novel finding. It was only included to partly quantify the possible error of the input data. We intend to shorten this part of the paper. In our opinion, the separation of the calibration and validation of the snow model from the hydrological model makes sense because of many reasons. Some of them are: (i) different hydrological models may use the same snow model (ii) an independently calibrated snow model reduces the uncertainty of the model parameters as no compensation of the model errors is possible through model parameters (reduced equifinality) (iii) snow models may be used individually for the estimation of available resources.

The discussion of the results is not comprehensive. It will be interesting to link the findings with similar studies calibrating the hydrologic models by using MODIS or comparing different variants of degree-day models.

This is an interesting suggestion. We will definitely look further into this.

I believe the manuscript presents an interesting topic and can be an interesting contribution, but it needs a very substantial revision and extension.

We will do our best to improve the paper.

Specific comments

Which MODIS version is applied?

Version V6 for both MODIS Terra (MOD10A1.006) and MODIS Aqua (MYD10A1.006) were used in the study. This will be added in the revision.

Kriging. Was the spatial correlation model (semi-variogram) fitted separately for each day?

Yes, the semi-variogram models were fitted for each day.

Radiation based model: how was the Linke coefficient estimated. Does it vary seasonally?

The Linke coefficient was set at a constant value of 3.0 (close to the annual mean for rural-city areas) for the model. Seasonal variation was not accounted for in the study but it could be an interesting addition. Instead, a diffusion factor (0.2 for clear sky to 0.8 for overcast conditions) was used to account for the diffusion.

Cross-validation of interpolation. Leave-one –out crossvalidation is typically used to camper different interpolation methods. Were the residuals smaller than obtained by some other interpolation method? How do the resulting maps compare with existing gridded (precipitation, air temperature) products provided by MeteoSwiss or DWD?

The goal of presenting the cross validation results was to give some information on input parameter uncertainty. The presented results suggest the Residual Kriging worked better than other methods for precipitation and External drift Kriging worked better for temperature data.

We have not compared the Kriged surfaces with the DWD or Meteoswiss gridded products yet. This can be a part of the revision as an annex if required.