Dear Reviewer,

we are very thankful for the positive feedback and the useful comments.

Please find below our replies as inserted blue text.

Kind regards,

Nena Griessinger, Jan Seibert, Jan Magnusson and Tobias Jonas

General Comments

The work of Griessinger et al. assesses the added value (or lack thereof) of a hierarchy of complexities in degree-day snow-models, possibly including SWE data assimilation. This type of models is frequently used in hydrological modelling. This manuscript is of high value for hydrological modellers in snow-dominated and snow-influenced catchments, and draws important conclusion as to the desirable level of complexity to be chosen depending on the type of catchment concerned (high snow/enduring-snow cover, low snow depth /ephemeral snow). The different model versions used here build a clever set-up to test the impact of different snow-melt parametrizations and of SWE data assimilation within a hydrological model.

However, a few important considerations are missing, which would strengthen the conclusions of the paper. These are listed below:

- First, the uncertainty associated with snow depth observation data is never mentioned. As I understand from the manuscript the collected snow depth data were rather punctual and to me, the mentioned 'flatness' of the terrain where they were collected does not guaranty their 'local' representativity. Elaboration on that, and precisions as to the snow depth measurement protocol, would be welcome. An ancillary aspect also regards the hydrological data, which are subject to quite high uncertainties in mountain catchments as a result of frequent shifts in the topography of the river beds. This aspect should at least be discussed.

We will include a discussion about the representativeness and uncertainty of the used punctual snow depth data. The stations used in this study were chosen carefully avoiding sites that are clearly influenced by strong wind drift of snow or frequent sensor failures. Similar datasets were already used for previous studies (Joerg-Hess et al. 2014, Magnusson et al, 2014). Since the dataset is used for operational monitoring of snow water resources, the data has been checked for erroneous or missing data. Measurement records taken at the same station at previous or following days and those taken at neighboring stations were used to appropriately replace or fill the measurement gaps. Regarding the runoff measurements, we rely on the quality of the data provided by the Federal Office of the Environment (FOEN). Nevertheless, we checked the data for missing values.

- Second, in most calibration and validation sets of simulations, M3 outperforms the upper-benchmark, which relies on a calibrated degree-day factorn whereas M3 relies on a constant degree-day factor for all catchments. To me this result is quite counter-intuitive and deserves an explanation.

Thank you for this remark, we agree that the finding M3 to outperform the upper benchmark may appear counterintuitive and should be discussed.

Dealing with liquid water content, refreezing, cold content dynamics and the partitioning of rain and snow are - among others - elements that influence the performance of temperature-index models. These elements differ between [M1, M2, M3] and HBV. In [M1, M2, M3] the representation of those processes have been particularly trained for optimal performance in the Swiss Alps.

Further, calibrating HBV for the melt season only – as done in our study – could result in a DDF that is too high during the snow accumulation period. The consequence might be an unbalanced performance with good snowmelt rates during the melt season at the price of too little accumulation earlier in the year with unwanted side effects on the snow depletion dynamics. M3 features a more moderate DDF of 2.5 mm°C^{-1}day^{-1} allowing for a more balanced performance over the entire snow season.

We will adapt the manuscript and provide this discussion to the reader.

- Finally, a distinct ‘discussion’ part could be inserted in the manuscript : Section 4.4 after line 11 could be part of it, as well as elements coming in response to point 2 mentioned above. Optionally, more elements as to the different, converging metrics used could be provided to the reader. The general decrease of (each) model performances with elevation could be commented and interpreted, in link with the quality of the interpolations (/extrapolation) of meteorological data and sometimes snow observations at these altitudes.

The setup of the manuscript was discussed with all authors in detail and we found the combination of results and discussion within one chapter as appropriate for this paper. Please also note that referee #1 particularly mentioned the good structure of the paper. We would like to thank for the interest in more interpretations which we could include in this chapter.

Minor Comments

- The last sentence of the abstract overlooks the fact that with altitude, not only the accurate estimation of snowmelt rate gains importance, but also the accurate estimation of SWE, which is one of the hypotheses tested by the paper’s set-up.

We will adapt the last sentence of the abstract to: “These findings suggest that with increasing elevation and correspondingly increased contribution of snowmelt to runoff, the accurate estimation of SWE and snowmelt rates gains importance.”

References:

Magnusson, J., Gustafsson, D., Hüsler, F., and Jonas, T.: Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. Water Resour. Res., 50(10), 7816-7835, doi: 10.1002/2014WR015302, 2014.

Jörg-Hess, S., Fundel, F., Jonas, T., and Zappa, M.: Homogenisation of a gridded snow water equivalent climatology for Alpine terrain: methodology and applications, The Cryosphere, 8, 471-485, doi:10.5194/tc-8-471-2014, 2014.