Reply to Anonymous Referee #1

We thank Anonymous Referee #1 for the thorough review. We give our reply (in italic) to the referee comments/ suggestions below.

This paper presents a study focusing on the modelling of snow accumulation and melting in an Himalayan catchment and the response of this catchment under different climate scenarios in terms of snow water equivalent (SWE) and melt runoff. This study addressed an interesting topic in a region where snow storage is crucial for water supply. The authors use data assimilation (Ensemble Kalman Filter, EnKF) of groundbased and remotely-sensed snow data to determine optimal parameters values in their modelling system. These optimal parameters are then used in climate sensitivity tests. My main comments about this study concern (i) the data assimilation method, especially the choice of variables to assimilate and the effects of these choices on final results and (ii) the limits of the climate sensitivity tests carried out with the optimized model. These questions need to be clarified prior to publication in TC. They are listed below (General comments) followed by more specific and technical comments.

General Comments

1) In the study, the EnKF is used to assimilate snow cover area per elevation band and snow depth at two locations. Four parameters are calibrated using the EnKF. My comments on this method concern (i) the choice and benefit of assimilating punctual snow depth measurements and (ii) the assimilation of MODIS snow cover. The assimilation of punctual snow depth is associated with high uncertainties due to the very limited representativeness of punctual snow depth measurement in mountainous terrain (e.g. Grünewald and Lehning, 2015).

We are aware of the high uncertainties related to the limited representativeness of punctual snow depth observations in complex terrain due to local influence of snow drift. We will add the reference, provided by the reviewer, to the revised manuscript for completeness. A key advantage of the EnKF is that it takes into account the uncertainty in the assimilated observations. Several observation uncertainties were tested (variance of 1cm, 16cm and 25cm) on how it influenced the posterior parameter distribution, prior to choosing the final observation uncertainty that is presented in this manuscript (variance of 25 cm). See the next reply for explanation on why we believe that a variance of 25 cm is representative for the uncertainty of punctual snow depth measurements in this case.
In addition we want to emphasize that the assimilation of snow depth observations is only used to calibrate the compaction parameter C_6 and not for calibration of the other three parameters.

For example, wind-induced snow transport can lead to erosion or accumulation of snow at the location of station. What would be the impact of such event when carrying out data assimilation with EnKF? Were the snow depth measurements assimilated in this paper impacted by such event?

Wind-induced snow transport is not included in the snow model presented in this study. If the snow depth measurements are affected by wind-induced snow transport there will definitely be a discrepancy between simulated and measured snow depth that cannot be correctly explained by the model. Using the EnKF for assimilation of measured snow depth to obtain optimal parameter values allows accounting for the uncertainty related to point snow depth measurements. Using a too small measurement uncertainty in the EnKF would result in model parameter values that are too strongly driven by the assimilated snow depth measurements, resulting in implausible values of the model parameters. We checked to which degree the parameter values were driven too strongly by the snow depth measurements as cause of a discrepancy between simulated and measured snow depth. This discrepancy *could (amongst other snow processes) result from wind-induced snow transport. It is difficult to assess whether the station data is influenced by this process based on only measurements of the snow depth. However, results from using different measurement uncertainties showed that a variance of 25 cm (that is used in this study) resulted a mild forcing of the model parameters and plausible values. Therefore, a variance of 25 cm for the measurement uncertainty was assumed to be representative for the uncertainty of punctual snow depth measurements in this case.*

The benefit of directly assimilating snow depth measurement is hard to identify throughout the paper. It would be interesting to have results obtained when only snow cover data are assimilated. In the present version of the manuscript, the advantage of simultaneous assimilation of snow cover and depth is not clear enough. Results in Section 3.3.1 and 3.3.2 could be presented (i) without assimilation, (ii) with assimilation of snow cover only and finally (iii) with simultaneous assimilation of snow cover and depth.

The data assimilation is performed in a two-step approach (p8 l14-18). TT, Tlapse and precip were optimized by assimilation of snow extent, whereas C_6 was optimized by assimilation of snow depth. In section 3.3.1 distinction can only be made between the steps 'without assimilation' and 'with assimilation' of snow cover, respectively uncalibrated and calibrated in Figures 3 and 4 because C_6 does not influence the snow extent as it is an parameter that converts SWE in snow depth.
The simulated snow depth without assimilation ('uncalibrated') and after assimilation of both snow depth and snow extent ('calibrated') is already given in Figure 5. The simulated snow depth after assimilation of snow cover only will be included to this figure to show the advantage of assimilation of both snow extent and snow depth (see Figure 1 *below).*

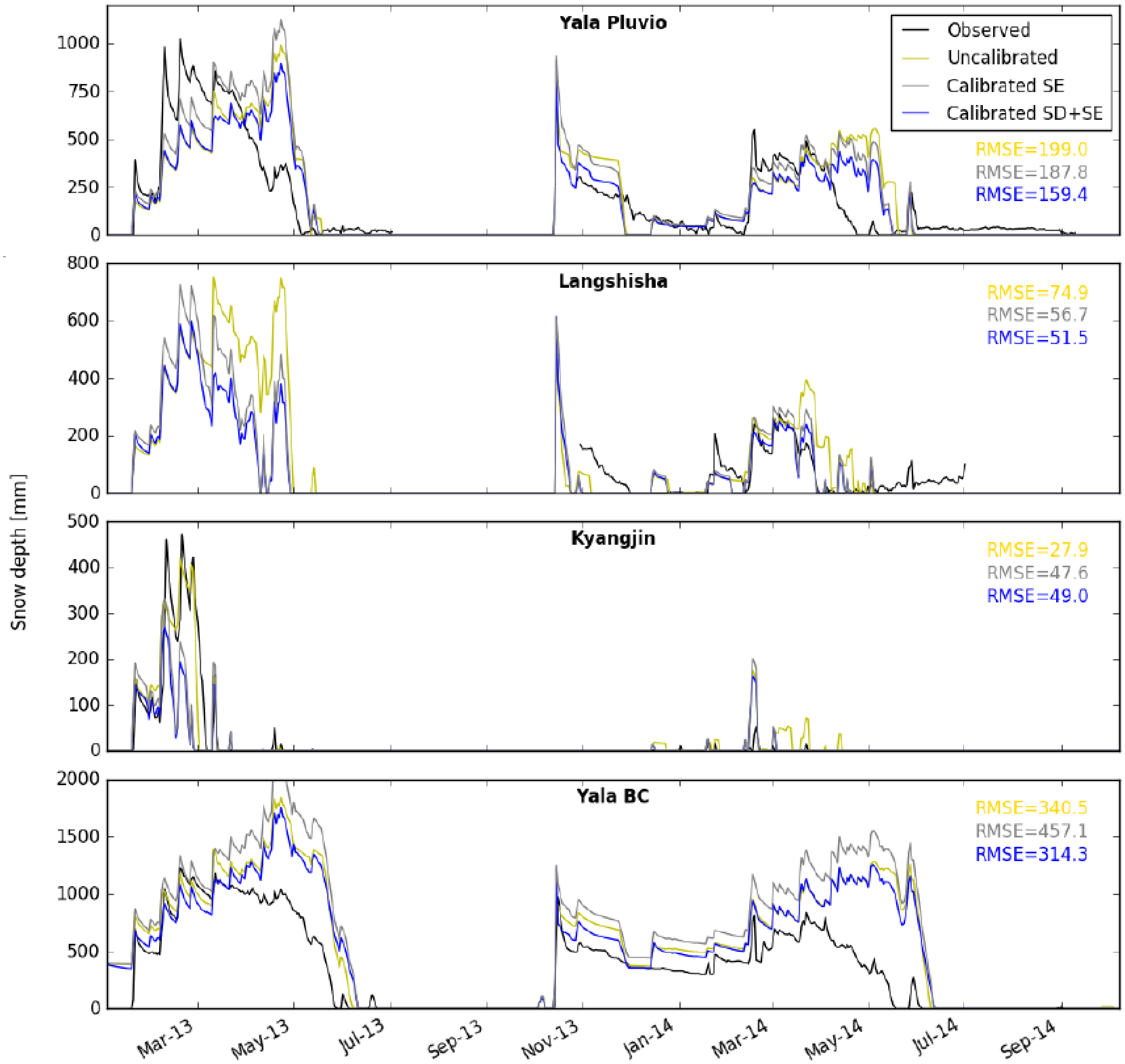

Figure 1 Observed snow depth and modelled snow depth before calibration, after assimilation of snow extent, and after assimilation of both snow depth and snow extent (ensemble mean) at four locations. The RMSE (mm) is given for the fit between modelled (before calibration, after assimilation of snow extent, and after assimilation of both snow depth and snow extent) and observed snow depth.

The assimilation of MODIS snow cover requires an observation operator to convert SeNorge output into simulated snow cover extent. Are the authors using a simple threshold value of SWE or snow depth to determine the presence or the absence of snow? Or are they using depletion curves?

The presence of snow in the model is based on a threshold value of 1 mm swe. This will be added to the manuscript.

MODIS snow cover are averaged per elevation band prior to assimilation. Can the author justify this choice? Indeed, averaging the information per elevation band reduce the information content brought by MODIS and remove the intra-band variability resulting from (i) the contrast between north-facing and south-facing slopes and (ii) the heterogeneous spatial distribution of precipitation.

The assimilation of snow cover would preferably be performed on a pixel to pixel basis to maintain all information on the spatial distribution of snow cover. However, the EnKF can only be used for continuous values and not for binary values (i.e. snow cover present or not). Therefore it is required to assimilate snow extent (continuous value) into the model. Rather than using the total snow extent for the entire catchment, we chose elevation bands to include more information on the spatial variability of snow cover. Elevation bands were chosen to capture the snow elevation line transition and therefore capture melt dynamics and spatial distribution of precipitation.

2) The authors used the optimized version of their model to carry out climate sensitivity tests. They use the delta method and applied changes in temperature and precipitation for different climate scenarios (Table 3). The authors do not discuss the uncertainties associated with this method. Such discussion is really relevant in a paper dealing with climate sensitivity. The delta method assumes constant changes in space and time for temperature and precipitation. How relevant is this assumption for this region? - Are the changes on temperature and precipitation expected to depend on the season? What are the expected effects for the hydrological cycle in this region? - The authors use the monthly precipitation pattern of Collier and Immerzeel (2015) to spatially distribute precipitation, both in present and future climate. The authors should discuss the validity of this assumption of constant monthly spatial pattern under future climate.

The scope of this study is not to run long-term climate change runs as indeed the study period is too short and might not be representative. This short study period simply does not allow a full-fledged study on climate change scenarios. The focus is rather on the climate <u>sensitivity</u> of the SWE in this study area as a result of changes in temperature and precipitation (using the simple delta method). The study shows the combined effect of changes in air temperature and precipitation. We aim to show patterns that could potentially occur in future under changes in temperature and precipitation. The four RCP4.5 scenarios from Immerzeel et al. (2013) were mainly used to have realistic values for changes in precipitation and temperature.

The spatial distribution of precipitation was kept constant for simplicity. Changes in (spatial distribution of) precipitation are difficult to predict and simulate. In this study we aim to show the importance of knowing the spatial distribution of precipitation in future as we show that increased melt due to increased air temperature can be compensated by an increase in precipitation at high elevation.

So, by no means do we intend to present a climate change impact study, but merely a sensitivity study. In the revised manuscript we will make this clearer. The core of the study is to show how assimilation of snow depth and remotely sensed snow cover in a snow model can lead to a better understanding and quantification of snow water equivalent and snowmelt runoff in an inaccessible, data scarce environment. In particular the title may have given the reviewer the wrong impression about the focus of our study and we will modify the title: 'Assimilation of snow cover and snow depth in a snow model to estimate snow water equivalent and snowmelt runoff in a Himalayan catchment'.

The study period (Jan. 2013 to Sep. 2014) should be compared to the present climatology of the catchment for temperature and precipitation. Is this period considered as cold or warm and wet or dry? Is it representative of the averaged current climate conditions in the Langtang catchment? The author apply the delta method to a short time period (from a climate perspective) and this short time period must be better characterized.

The study period will be compared to climatology of the catchment. An additional figure (see Figure 2 *below) will be added comparing the study period to the 1988-2009 climatology. However, we want to emphasize again that we do not intend to present a climate change impact study, but merely a sensitivity study.*

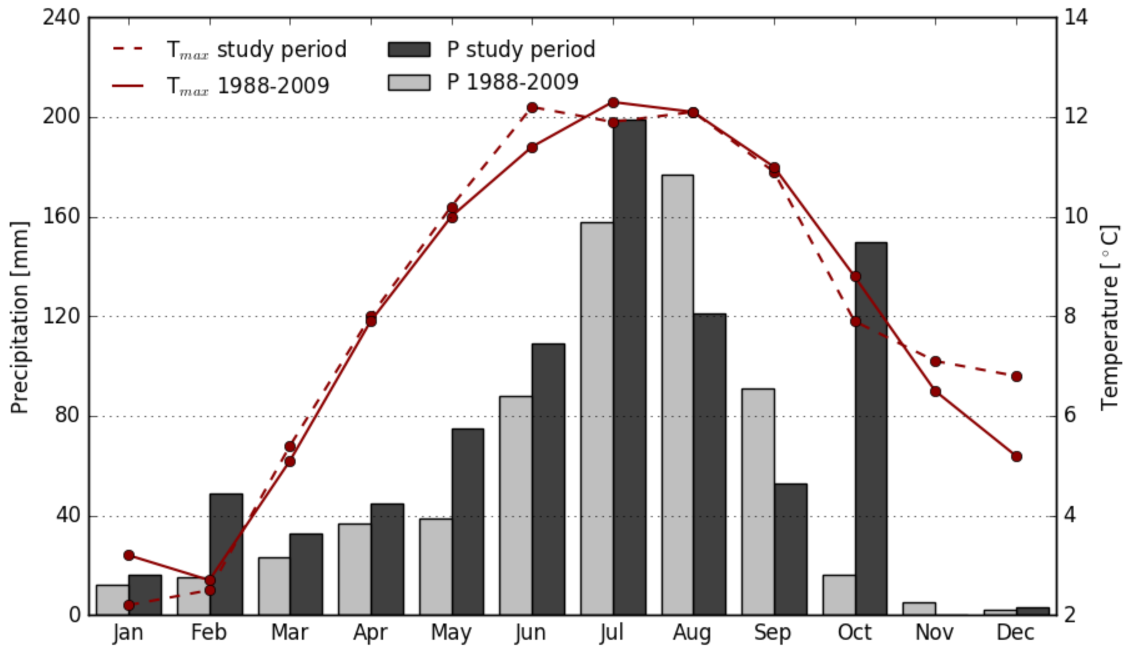

Figure 2 Comparison of maximum temperature (T_{max}) and cumulative monthly precipitation (P) for the study period and the 1988-2009 time series (based on measurements in Kyangjin). The average yearly cumulative precipitation is 853mm and 663mm for the study period and the 1988-2009 time series respectively.

In section 3.5 at P 13 L1, L 13-14 and L 17-18, they authors discuss how the SWE and changes in SWE depend on elevation. This discussion is supported by Figures 7 and 8 that provide maps of SWE for the study period and change of SWE in the different climate sensitivity tests. I recommend the authors to provide complementary figures showing these variables as a function of elevation. It would help the reader to clearly identify the influence of elevation.

An additional figure will be added with boxplots of SWE per elevation zone for the reference run and the four climate sensitivity tests (see Figure 3 *below).*

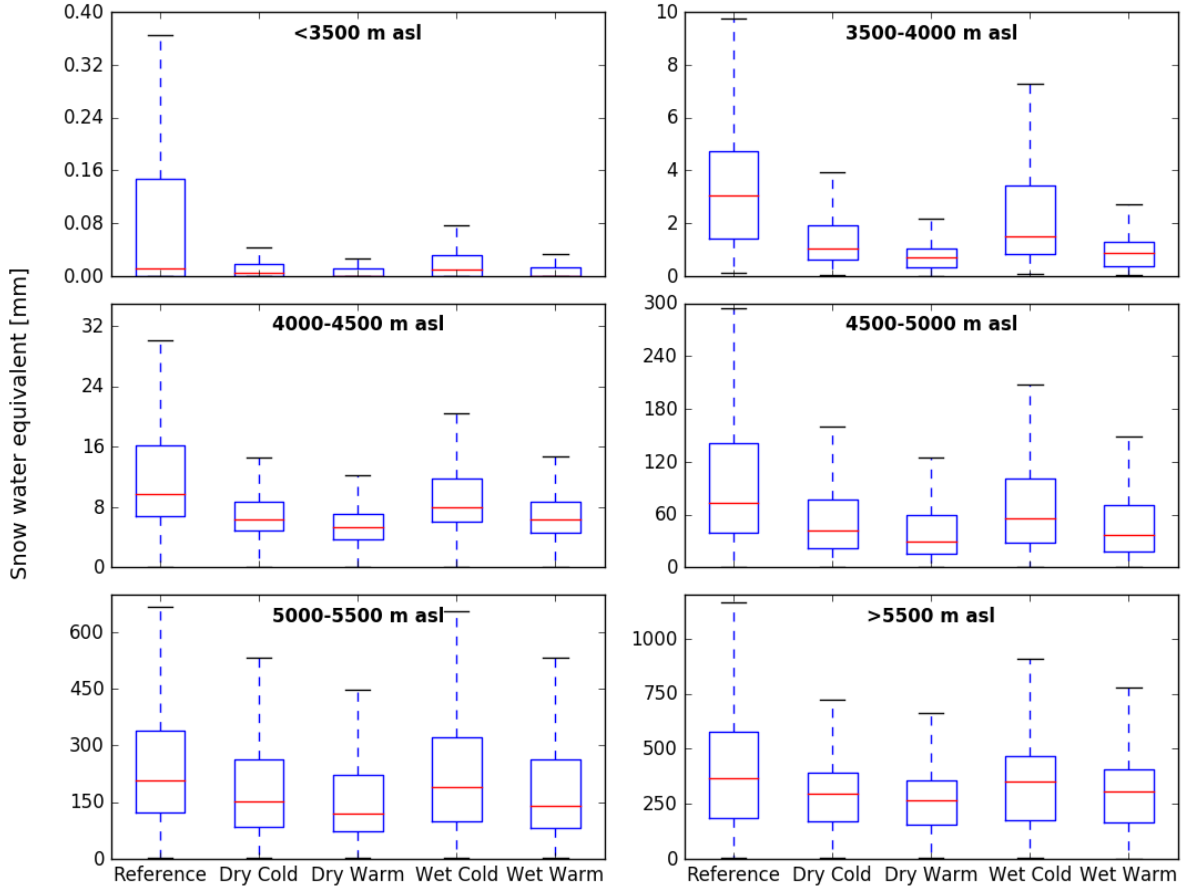

Figure 3 Boxplots of SWE per elevation zone averaged over the simulation period and all ensemble members for the reference run and the four climate sensitivity tests.

Specific comments

Introduction: the introduction is rather short and only presents earlier studies carried out in the Himalayan region. I recommend the authors to write more general paragraphs on (i) data assimilation of ground-based and remotely-sensed snow data in snowpack model and (ii) distributed snowpack modelling applied in mountainous region to simulate the cryospheric and hydrological response of mountain catchments under present and future climate. They should present in this introduction how techniques developed in other mountainous regions could be applied to an Himalayan catchment.

A more extensive introduction will be given, including the points given above.

P 4 L 16-17: the description of the location of the snow depth measurements is confusing. Are the 4 sites measuring snow depth located along the 2 transects? Figure 1 suggests that this is not the case. The authors should clarify this point.

There are 2 transects of surface temperature measurements. Only two out of four snow depth measurements are located along the 2 transects. This will be adjusted.

P 4 L 28: which uncertainties are taken into account with the correction factor precip? Does it include: - uncertainties in solid precipitation measurements at the station due to wind undercatch? - spatial and temporal representativeness across the catchment of the precipitation measured at the station?

Correction factor precip accounts for the uncertainties related to undercatch. This factor does not account for the uncertainties in the representativeness of the chosen monthly spatial distribution of precipitation. We believe that quantifying the uncertainties in spatial distribution is beyond the scope of this study. The fact that we use the monthly spatial precipitation distribution based on a high resolution weather model based Collier and Immerzeel (2015) is already a great improvement over previous studies where fixed lapse rates were used for the entire catchment to regionalize precipitation based on observations of a single station.

P 6 L 3-4: please mention that in Brock et al. (2000) the snow albedo remains constant when the maximum air temperature is below 0 _C.

This will be added.

P 6 L 22: the sentence "Separate transport ... this study" should be reformulated. It suggests than when wet snow avalanches occur the ice and liquid phases are transported separately. This is not the case in the nature. It seems that the authors mentions this point only because seNorge treats separately the solid and the liquid phase in the snowpack.

This was indeed mentioned because seNorge separately treats the solid and liquid phase in the snowpack. The sentence will be reformulated.

P 7 L 8: the runs used for the sensitivity analysis are not clearly described. For each run, are the authors using the model to simulate the evolution of snow cover and SWE over the whole study period (January 2013- September 2014) and the whole catchment? Or are they using different time period and sub-domains?

The sensitivity runs simulate the evolution of snow cover and SWE over the whole study period and the whole catchment. This will be better described.

P 7 L 10: how are computed the mean snow cover extent and snow depth? Are they averaged over the whole period and the whole domain? This point is similar to my previous point regarding the characteristics of the simulations used in the sensitivity analysis.

They are indeed averaged over the whole study area and study period. The description will be improved.

P 7 L 10 (and in the rest of the paper): the author should precise how they compute the snow cover extent from the output of seNorge. Cf my general comments about the observation operator.

See the answer on the general comment.

P 8 L 22-23: how is modified the maximum air temperature in the climate sensitivity tests?

This is a good point. In the previous version of the manuscript we have only perturbed the mean temperature forcing, but for the revised manuscript we have now imposed the same delta change on the maximum temperatures. The impact is however minimal and only for the dry warm case, a difference of a few percent in SWE is simulated compared to the previous version. This is because the maximum temperature is only used in the albedo decay algorithm, so when Tmax is higher, then the albedo decay will start a bit earlier, however the impact is minimal in these sensitivity tests. Figures 8 and 9 will be updated in the revised manuscript with the climate sensitivity tests including perturbation of Tmax.

P 10 L 19-25: this paragraph should also discuss model results in the elevations zones above 5000 m. For example, could the author discuss the differences between summer 2013 and 2014 in terms of snow extent in the elevation zones 5000-5000 m and >5500 m? What can explain the underestimation of SCE in these zones for summer 2013 whereas better results are achieved in summer 2014?

We will add a paragraph to the results and discussion section, where we discuss the potential reasons for those high altitude differences.

P 10 L 29: differences in classification accuracy with and without calibration are hard to identify on Figure 4. A map of differences of classification accuracy could help the reader to better identity the regions where large differences are found between the two simulations.

Before calibration the snow model already shows high performance in simulating snow cover (Figure 4a). After calibration there is only a modest improvement in accuracy in most regions, whereas there is a slight decrease in performance in few other regions. The small changes in performance are therefore indeed hard to identify. The only pronounced improvement is in the lower area on the northern slope and this improvement also shows from Figure 3. A difference map was tested, but resulted in rather chaotic patterns of small increases and decreases in accuracy in steep terrain and therefore does not help visualization and interpretation. Given the little information content, the fact that elevation dependent improvements are already shown in Figure 3 and since we already add a new figure (boxplots) we have decided not to include the proposed figure in the revised manuscript.

P 10 L 30-31: the authors associate the low classification accuracies in the northern part of the catchment with model errors due the avalanching parametrization. However, it seems that this difference can also arise from errors in the meteorological forcing used to drive seNorge. For example: (i) errors in precipitation phase and amount, (ii) errors in the spatial distribution of precipitation. Indeed, the spatial distribution of precipitation is based on monthly precipitation patterns derived from Collier and Immerzeel (2015). For a given precipitation event, the spatial distribution of precipitation can vary from the monthly pattern from Collier and Immerzeel (2015) and strongly affect the snow cover. Please add a discussion about the different potential sources of error.

This is correct and a more complete description of reasons for low classification errors will be added to the discussion.

P 11 L 23-24: please consider reformulating the last sentence of this paragraph. Indeed, the improvement for Kyangjin in 2014 is not really clear.

This will be reformulated.

P 11 L 25: the authors point out the lack of independent stations for the evaluation of snow depth and SWE. Are glacier mass balance data available for a glacier in this catchment to bring complementary values for evaluation? For example, winter mass balance data can provide interesting evaluation on the cumulated precipitation during the winter.

The yearly mass balance is available for Yala glacier, which is a clean-ice glacier positioned in the study area. However, the yearly mass balance is negative. As only snowmelt is simulated, and no glacier melt, it is impossible to simulate a negative mass balance with the snow model. Therefore no comparison was made between output of the snow model and available mass balance data.

Though, an extra data set of snow depth measurements will be used (Yala BC; see Figure 1 *above) to improve the validation of the simulated snow depth. This will be added to the revised manuscript.*

P 11 L32: the absence of underestimation or overestimation concerns snow depth and not SWE.

This will be adapted.

P 12 L 5-30, Section 3.4: This section does not contain new and original results and only presents the effect of well-established parametrizations introduced in seNorge to improve the snowpack dynamics without comparison with measurements. I recommend the authors to remove the discussion concerning the snow compaction and the snow albedo since it does not bring additional value to their paper.

This section will be removed from the manuscript.

Concerning the avalanche parametrization, the discussion at lines 7-10 (P 12) suggests that avalanching strongly affects the simulation results. It would be really interesting if the authors could illustrate how the avalanching parametrization improves the representation of the snow depth distribution in the model. Figure 7 shows that, in the simulations, snow accumulates at the bottom of the steep slopes of the catchment. Are these zones of additional snow accumulation identified on the LandSat images at 30-m resolution? Such discussion on avalanche processes and a comparison with remotely-sensed observation would substantially improve the quality of this section on snow processes. Otherwise, I recommend to remove this section from the paper.

The snow accumulation zones are not clearly visible in Landsat 8 imagery. This is caused by i) the shadow in the steep areas, ii) a too small extent of the accumulation zones, and iii) potentially a wrong timing of the acquisition date. Therefore a comparison of simulated and remotely sensed snow cover in the deposition zones is impossible. Based on extensive field experience we know that avalanches regularly occur at these steep walls and that snow is deposited downslope these steep walls. The walls are too steep to have a substantial snow depth. Including avalanching in the model improves the snow redistribution as there can be no unrealistic accumulation of snow in these steep zones. Although, there is no possibility to support this with Landsat 8 imagery, we respectfully disagree with the recommendation of removing this section, as we believe it is important and improves the distribution of snow in the model.

Technical comments
Text

P 16 L 25: modify the reference to Immerzeel et al. (2014)

This will be modified.

References (not included in the submitted manuscript)

Grünewald, T., Lehning, M. (2015). Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements

This will be included.