In their study “Using MODIS estimates of fractional snow cover extent to improve streamflow forecasts in Interior Alaska” Bennett et al. investigate the value of two MODIS-derived snow cover area products (MOD10A1 and MODSCAG) to improve streamflow simulations in the interior of Alaska as compared to simulations where model-generated areal depletion curves are used. The authors conclude that there is only marginal improvement when evaluating the model performance with metrics such as NSE, RMSE etc., but argue that the MODIS derived snow covered area products might be valuable particularly in regions with sparse or poor quality observations. The methods and findings are sound and the article is generally written in a clear, concise and very structured way.

Thank you for your positive words about our study.

Nevertheless, I have some minor comments, which I would like to see addressed prior to publication:

General comments

• The basins with the sparsest streamflow observation had the greatest improvement in streamflow simulations. (P1L13,P14L 25-26) –> could indeed the sparse data be the main reason for the improvement rather than the product? This could be tested with a little model experiment adding some data gaps and see how the performance measure is sensitive to that maybe for the basin with the longest observations.

We are unsure if we could simply remove missing data from a ‘good’ record to test this. It is difficult to pull apart whether improvement is due to the missing streamflow observations, or if it is due to improving the snow observations in the catchment. However, it makes logical sense that improving observations (whether snow or streamflow) would be beneficial when before there are no observations. We talk about this in the discussion section of the paper. If this is a sticking point for the reviewer, we can put some thought into how to best test this, and perform the experiment for one watershed, and one year, as an example.

• Why do the authors include the Chatanika catchment, when it shows very poor performance both using the model generated and the MODIS product derived snow cover extend? Please clarify.

We were interested to show the results for the Chatanika as we feel that this basin is most representative of the poor-quality streamflow observations in Alaska. We think that even though it didn’t perform well, that this is likely the kind of issues modelers and observationalists will deal with when working with Alaskan hydrology data. This basin is also adjacent to the Caribou-Poker Creek instrumented watershed, and it represents a lowland site in comparison to the Salcha, Chena and Goodpaster systems which have more upland land cover. For this reason, we would like to keep the basin included in the paper, despite its poor performance.

• Move more detailed description of the derivation and differences between the MODIs products (interpolation, filtering, and smoothing) from the supplements in the main study.

We have now moved more of the details from the MODIS productions to the body of the paper. We originally took it out because we felt that the Methods section was too long, and we didn’t want to bog readers down. Based on this comment, and comments from Reviewer 1, we were too gregarious with these edits.

Minor/technical comments

• There are passages in the article where fill words like utmost, great, very are used abundantly. In my opinion, it would help to go through the article and check where these are really needed and where they could be dropped.

We have dropped these words throughout the text. Please see the track changed version of the manuscript.

• There is a mixed use of watershed and basin. Do the authors use it with equivalent meaning? If so, then use only one term throughout the paper, if not please clarify why the two terms are used.
We have removed the use of the word watershed and replaced it with basin throughout the paper. Thank you for noting this.

• Often model runs could be replaced for better clarity with SWE simulations (e.g. P10L30) or streamflow simulations, respectively.
We have replaced run(s) with simulation(s) throughout the paper.

• Please make the units consistent (sometimes there is sec, sometimes there is s, etc.)
We have made the units consistent throughout the text. Please see the track changed version of the manuscript.

P1L29/30 did both extent and duration decrease by the same percentage?
We have adjusted this sentence to read "Snowpack extents in Alaska have decreased over time by 18% (1966-2012) due to an earlier snow melt, while snowpack duration has also decreased (SWIPA, 2012)."

P1L35 Delete this sentence
We changed this sentence to read as follows "Rivers in Alaska have been observed to be changing as a result of an intensified or stronger hydrologic cycle that could lead to an increase in peak flows in the Northern American high latitudes (Cohen et al. 2012; Huntington, 2006; Rawlins et al., 2010)."

P2L3 Extremes –> Extreme
We corrected this typographic error. Thank you for noting it.

P3L14 in which P3L22/23 what does it mean that they perform better, better than in other regions or these models are better in these regions than the other models. Please clarify.
We meant that temperature index models are presumed to perform better than other models in highly variable landscapes with spare networks. We have added in "…than other models for regions…" to the sentence to clarify.

P3L33 Date missing in reference
We have added the date to the reference. Thank you.

P4L33 delete above the Steese Highway (I do not see the relevance of this information)
We have deleted this part of the sentence.

P5L23 delete at the Steese Highway site
We have deleted this part of the sentence.

P5L35 add eq. 1 in brackets
We have added Equation 1 to the first equation in the paper (SCA_{fadj}), and Equation 2 for this equation.

P6L1 does still really need to be expressed in feet?
Because of the way that the equation was developed, you cannot obtain the same answer by converting 1000 ft to meters. Thus, I believe that you must enter the value of the elevation in feet. It is confusing because I used meters in the example. I have now added the meters to feet conversion to make it clearer, and added the unit value after elevation in brackets.

P6L17 how sensitive are streamflow simulations to this lapse rate? What motivates the assumption that the fixed lapse rate of 0.6_C/100m holds?
This lapse rate is a default in SNOW17 to represent the saturated adiabatic lapse rate, and is used to calculate the percentage of the watershed where precipitation falls as snow. However, the value can be changed for each basin and sub-basin, if warranted by the input temperature data, and also different methods can be applied to separate rain versus snow. The lapse rate is used to find the air temperature threshold value, and this value is used to relate to an elevation, for which the basin area snow fall can be calculated. It is important to note that this is different from other uses of lapse rates in the model.
Although we could not find previous studies that account for the sensitivity of this parameter, there are six main parameters in SNOW17 that have been identified as the most sensitive parameters for SNOW17, SCF, MFMAX, MFMIN, SI and UADJ (Anderson 2002, Tang et al. 2007). The use of the single lapse rate value for these calculations is widely applied in studies across the globe (e.g. Clark et al. 2011). We have added more detail and these references to this section of the manuscript.

P6L30 delete "and is set to . . . "already mentioned above
We have deleted this part of the sentence.

P6L33 mm/mb/6hr is that the unit for rain on snow, then move to melt from rain of snow earlier in the sentence
We have moved mm/mb/6hr to come after UADJ events in the revised sentence.

P7L12-15 It might be helpful to see a sketch of how the ADC works.
While we agree it might be useful, we have many figures in this paper already. Thus, we have added a reference to the images that depict the ADC relationships in Anderson 2002's paper (Anderson 2002, Figure 7.4.3, 7.4.4).

P7L16 add a reference to this look up table
We have added in the reference for the ADC look up table.

P7L29 change "produces streamflow simulates" to "simulates streamflow"
We have changed this sentence as suggested.

P8L12 delete study
We have deleted study.

P8L30 to P910 all of these statistics are well known, I think it would be sufficient to just add some reference for each, else, I would summarize them in a separate table. In case the authors want to keep the equations, check the equations carefully: MAE has to additionally be divided by the number of observations, Spearman correlation coefficient could simply be written in a simpler form.
We agree with this comment and have opted to remove the section and add references for a few of the statistics.

P9L10 make units consistent
We are unsure which units you are referring to here, but we have reviewed all units in the manuscript to ensure they are consistent. Please let us know if we have missed anything.

P10L1 delete "for reasons that are discussed in the following section"
We have deleted this part of the sentence.

P10L5 why it is the maximum recommended value and who recommended it? Maybe refer again to the table with parameter ranges.
Anderson 2002 recommends the ranges for these parameters. We have changed the sentence to include the reference and refer to the table.

P10L14 What does a more rigorous calibration mean here?
We have deleted this sentence.

P11L10 Why for May 15th 2001?
The May 15th date for this region in Alaska represents a time when snow is melting, and we should be partway through the snow recession. For this reason, some snow will be represented at higher elevations and likely less at lower elevations. The year, 2001, was selected somewhat arbitrarily, it is a moderate snow year, which we thought would show these relationships and differences across MODIS data more clearly.

P12L6 1-R
We have corrected this typographical error. Thank you.

Figures

- Most figures are difficult to read when printed in black and white. This could be improved easily by adapting the color palette.
  We changed the color palettes in most of the figures.
- Figure1: not all elevation classes are used in the map. Units are missing for the elevation zones. Drop last sentence in caption.
  Although hard to see, the upper two elevation classes (green shades) are found in the Salcha and Goodpaster basins. However, we have revised the color palatte and color ramp on the figure.We added elevation units and dropped the last sentence as suggested.

- Figure2: what happens to MODIS SASC North between May 10 and 17?
  It looks like the MODIS SASC recorded that snow cover extent increased to 80% of the Upper Chena River basin on May 13th, 2010 at 6:00 am. After this point, at the next recorded interval the snow had all melted. We correlated this with the SNOTEL gauge across the Upper Chena river basin, which are Upper Chena, Teuchet Creek, and Monument Creek. Although it looks like some precipitation fell on the 11th, no snow fell at all around this time. Thus, we believe this is an error in the MODIS data, potentially where clouds were interpreted as snow cover. Also, comparing directly with the SNOTEL gages indicates that all snow cover extents should also have been at zero at this point, however all model results indicate that there was still residual amounts of snow (0.1 fractional SASC) in the catchment. However, the plot is meant to show the differences between the SASC in the SNOW17 when different remote sensing tools are applied. Therefore, to not distract the reader, we opted to remove this outlier data point as quality control. We have not changed the text as we feel that this level of detail is not warranted, but if you think we should explain the removal of the data point, we will add in a sentence.

- Figure3: It seems there is a difference in the fractional snow cover extend seasonal development for the years that are used for calibration and the years that are used for validation. Is this also the case for each individual catchment, or is one catchment causing this difference? Is it ok to shift the year in validation, calibration period in the Chatanika catchment compared to the other catchments?
  This figure shows the average snow cover extent across all the catchments based on elevation bands, and we tried to capture high and low streamflow years in the calibration and validation periods. Although, the figure shows the variability in snow cover extent across the years, it also shows that there are high and low years where there was variability across elevation zones in the melt trajectories, some years where there was a larger range of snow melt out dates (2000, 2006), while other years there was more consistency in the melt out (2002, 2007). We shifted the calibration and validation period in the Chatanika due to the improvement in quality of the data for the last 5 years of the record, and we think this is ok. To show the variations in the streamflow data, we show all the years for the Upper Chena (Figure 1) and the Chatanika basins (Figure 2) are given below.

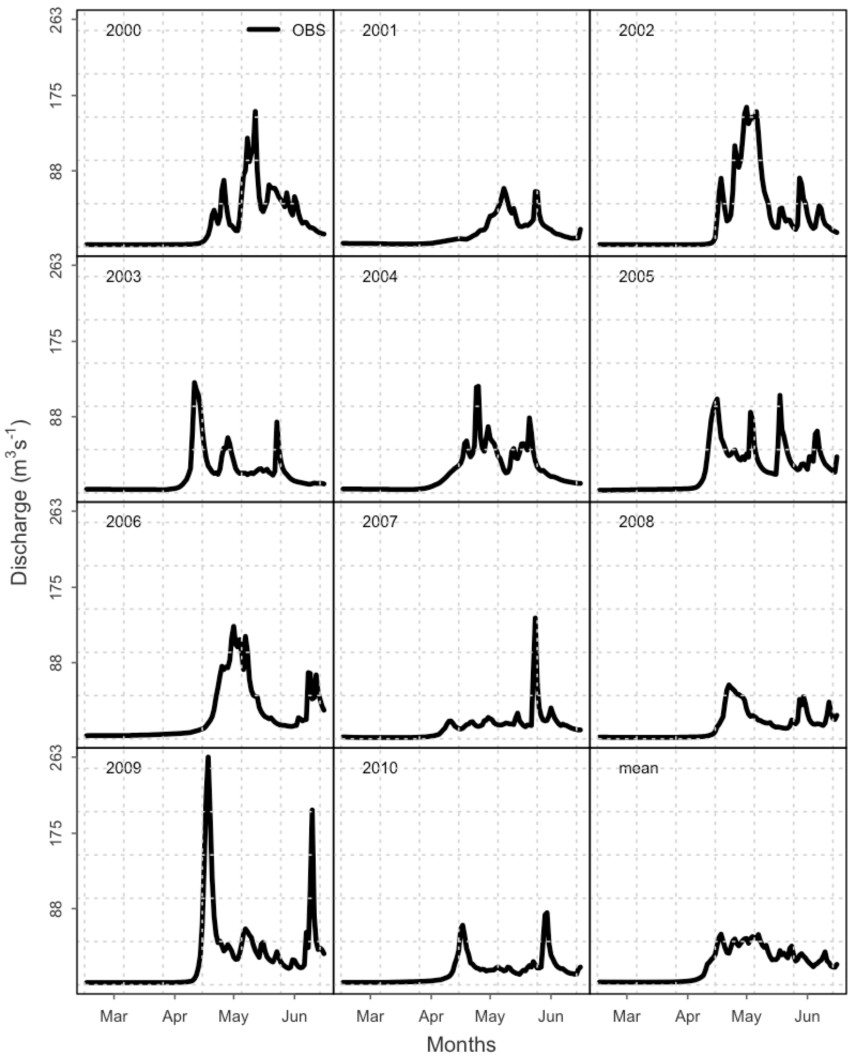

Figure 1. The Upper Chena River basin observed streamflow for the calibration and validation years.

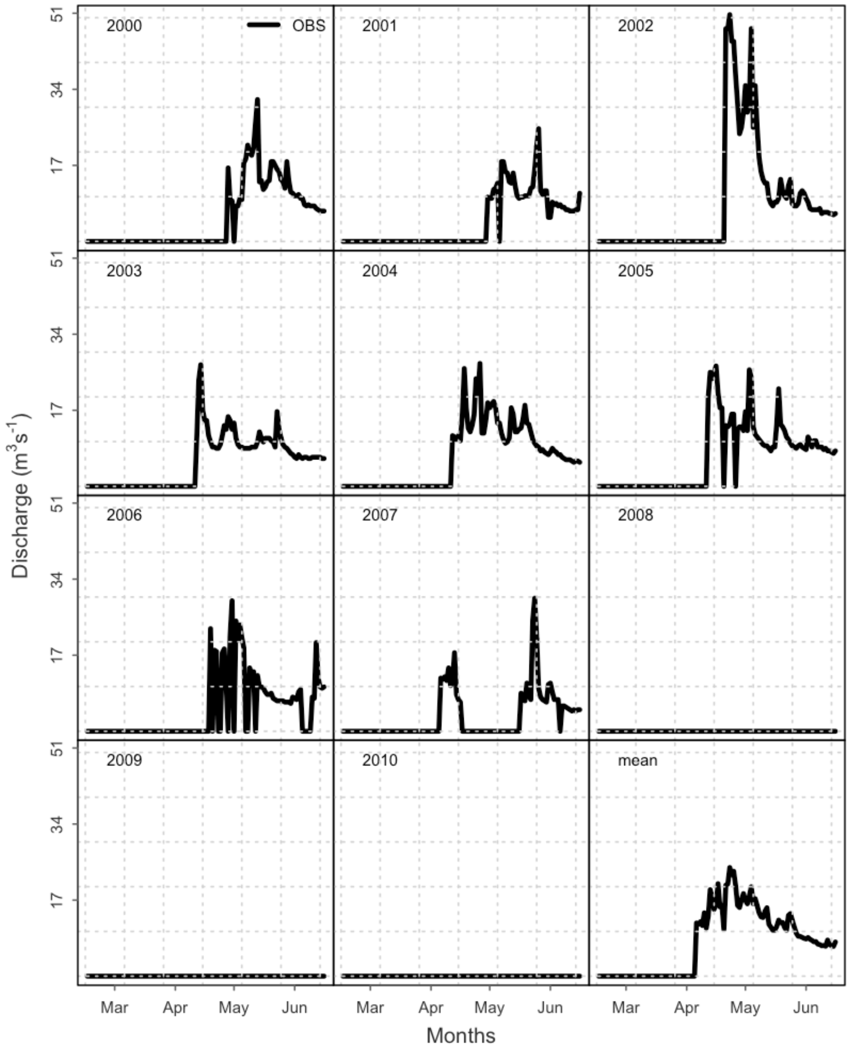

Figure 2 The Chatanika River basin observed streamflow for the calibration and validation years.

To show the variations in the catchments we generated the year 2000 based on the upper basin areas (Figure 3). We do not think there is a lot of variation across the basins, and hence we feel that the variability observed in each panel in Figure 3 is due to the elevation differences, and the year-to-year variations in climate, which occur across all basins.

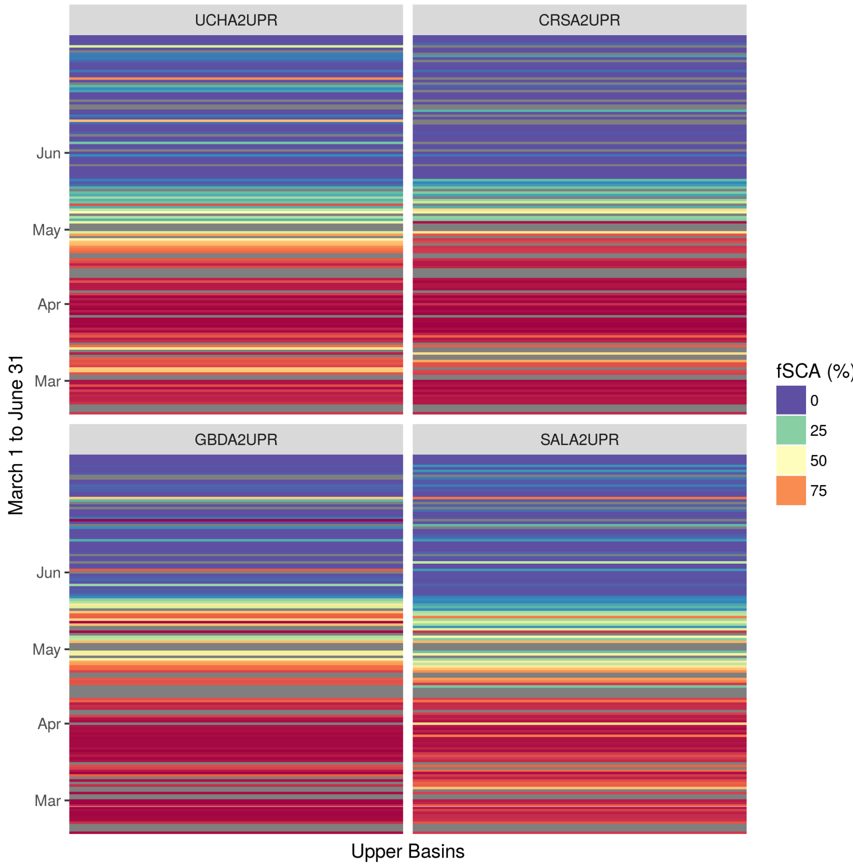

Figure 3. Four upper basins and their SCA (%) values.

- Figure9: could the time be extended spanning April to September as also used in Table 1? Would it be possible to add quantiles of the streamflow to get an idea about the range per season? Unit should be m3/s, I do not understand what the average of all basins can tell me.
  We decided to change the table to align with Figure 9. So, the table now shows precipitation values for Nov-Dec-Jan-Feb (winter), and Mar-Apr-May-June (spring). Because we calibrated the models only for the snow melt season, the rainy season is not illustrated in the plots. Also, we corrected the units in the plots. The average of all basins was included to show the regional streamflow magnitude and hydrograph shape for the average of all basins. We have opted to retain it in the plots but if you feel strongly about this we can remove it.

Tables

- Table1: units: elevation m a.s.l.; Q m3/s?
  We have corrected the Table, and changed the period of P to align with Figure 9.
- Table2: maybe replace current with the last year included.
  We have replaced current with last year as suggested.
- Table3: SCF Max values seem messed up
  We are unsure what you mean by this comment. The SCF values ranges from 0.65 to 0.95 across the catchments.
- Table4: mention that period of record is not the same for each catchment!
  We have added this to the table caption as suggested.

References:

Anderson, E.A., 2002. Calibration of conceptual hydrologic models for use in river forecasting. Office of Hydrologic Development, US National Weather Service, Silver Spring, MD. http://140.90.113.200/oh/hrl/modelcalibration/1.%20Calibration%20Process/1_Anderson_CalbManual.pdf Accessed August 17th, 2018. pp 372.

Clark, M.P., Hendrikx, J., Slater, A.G., Kavetski, D., Anderson, B., Cullen, N.J., Kerr, T., Hreinsson, E.Ö. and Woods, R.A., 2011. Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, *47*(7).

Tang, T., P. Reed, T. Wagener, K. Van Werkhoven. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrology and Earth System Sciences Discussions, European Geosciences Union, 2007, 11 (2), pp.793-817.