Reviewer #1

Thank you for the feedback and suggestions for improvements. We agree with all the suggestions, including the suggestion for additional analysis related to basin-level peak SWE/Q comparisons. All suggestions can be incorporated in the revised manuscript. Specific responses are provided in detail below in red font.

Summary

This manuscript examines the implications on simulated snowpack conditions associated with changing a precipitation partitioning model (partitioning precipitation into rain and snow fractions), from an approach considering only contemporary air temperature to an approach using both contemporary air temperature and humidity. The authors carry out a suite of numerical experiments in the Pacific Northwest region of the United States using the two alternative partitioning schemes and compare the approaches to each other and to observational data at USDA SNOTEL observation sites. Simulations and observations are compared on the basis of peak snow water equivalent (SWE), the timing of peak SWE, snow phenology (that is, the onset, conclusion, and duration of seasonal snow cover), and the fraction of runoff from snow at SNOTEL sites within the Columbia River Basin. Overall, the switch to a bivariate precipitation partitioning approach yields improvements in simulated snow conditions with respect to observations. Where there is degradation, the authors effectively argue that the degradations are likely the influence of factors beyond those that would be influenced by the change in partitioning approach. The paper is of interest to the readership of HESS and makes important, albeit primarily methodological, contributions to snowpack and hydrologic modeling in snow-dominated regions. I believe that it can be published with only minor revisions.

Major Comments:

a) I can surmise why VIC-CropSyst is being used in this study, instead of VIC without the CropSyst model coupled. However, some readers might be left wondering why VIC-CropSyst is being used when, for example, value additive aspects of the crop model (e.g., dynamic crop yields, etc.) are not being examined and when the comparisons with SNOTEL observations are primarily in locations where there is little or no cultivation. As such, I would strongly encourage the authors to provide some additional context for using VIC-CropSyst for readers. For example, if the work summarized in the manuscript is part of larger and/or ongoing efforts to develop more sophisticated regional projections of how climate change might affect agriculture in the region, it would be helpful for readers to know.

The use of VIC-CropSyst in this study, rather than the standalone VIC model, was indeed a deliberate choice. We agree that this aspect may be unclear to some readers, and we appreciate the suggestion. Reviewer 2 raised a similar point as well. We can revise Section 2.3 of the manuscript (VIC-CropSyst Model and Calibration) to clarify why we chose the VIC-CropSyst model rather than the standalone VIC model with text similar to what we have following two paragraphs.

This study is part of a larger, ongoing effort to understand the interplay between water supply, agricultural water availability, and the impacts of water shortages on agricultural productivity. The VIC-CropSyst model combines the VIC hydrology engine with the dynamic crop growth engine from the CropSyst model, enabling it to address both supply and demand sides of water usage in an integrated manner. This allows the model to capture how changes to water supply influence agricultural demand, and vice versa.

While the specific modifications related to the rain-snow partitioning scheme will be the same whether implemented in the VIC model or the VIC-CropSyst model, and the simulated impacts on snow and streamflow will be identical if the CropSyst crop growth engine is not invoked, implementing our changes in the VIC-CropSyst model offers significant advantages. It will enable us to address both standalone water supply applications and those related to the complex interaction between water supply and demand in future studies. Accurate snow simulations are crucial for the coupled VIC-CropSyst model, as snowmelt estimates directly affect soil moisture, water availability, and, consequently, crop growth and water demand.

We can update the manuscript to include this clarification and to better explain the rationale for using the VIC-CropSyst model.

b) From a water balance perspective, the analysis of snowmelt contribution to streamflow is interesting. However, given that there is not really meaningful observational constraint at SNOTEL sites, we are left only with model-to-model comparisons. While the authors are clear that this is the case, it is also possible that they could select a few watersheds within the region and compare simulated SWE/Q to the USDA's Basin Snow Water Equivalent estimates of SWE, normalized by runoff volume for the same watersheds. This represents some additional analysis, but it might provide a much more insightful comparison of the degree to which the change in precipitation partitioning influences whole-basin estimates of water supply.

We agree that comparing simulated SWE/Q to the USDA's basin estimates from observations, would provide additional insight, and can integrate this into a revised manuscript. This comparison would help to better assess the degree to which changes in precipitation partitioning influence observation-simulation matches of whole-basin water supply estimates. However, to ensure that observed flows are comparable to the VIC-CropSyst model outputs, which are "natural" flows without human influence, we focus on six watersheds (Boise River near Twin Springs (BOTWI), Clearwater River at Orofino (CLEAR), Coeur d'Alene above Shoshone creek near Prichard (COEUR), St. Joe River at Calder (JOECA), South Fork Payette River at Lowman (PAYLO), and Stehekin River at Stehekin (STEHE)) with minimal human influence as noted by USGS. For each basin, we follow the same methodology used by USDA in terms of the stream gauge and SNOTEL stations considered.

The analysis results for all simulation years across the six selected watersheds are provided below. Interestingly, we found that while dynamic partition decreases relative bias as compared to static partitioning and is almost always the best performing model for peak SWE (Figure R1 a) and annual Q (Figure R1 b), that performance improvement does not translate to a better match with observations for the SWE/Q ratio metric in all those cases. While dynamic partitioning is still the best method for the majority of cases for the ratio metric, it is the best model in fewer cases than for peak SWE and annual Q (the length of the vertical black lines in Figure R1).

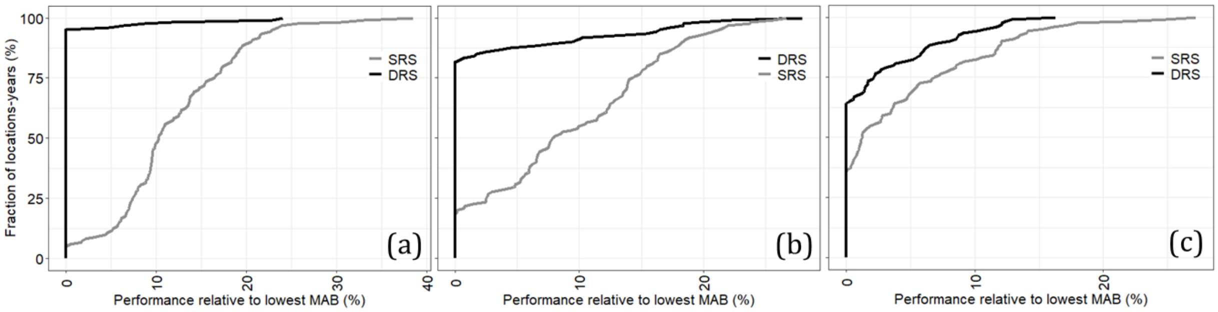

Figure R1. Relative model performance (RMP) chart for (a) peak SWE (SWE), (b) streamflow (Q), and (c) SWE/Q ratio, comparing the SRS (static rain-snow) and DRS (dynamic rain-snow) partitioning methods. The Y-axis is the fraction of stations for which a particular RMP is achieved, and X-axis is the difference between each model's mean absolute bias (MAB) and the best-performing model's MAB. The closer a model's curve is to the Y-axis and for longer, the better. The length of a model's curve exactly on the Y-axis indicates how frequently the model is best performing, and the distance of the curve from the Y-axis indicates how much worse a model's performance is relative to the best model. If the model under consideration is the best-performing one, the RMP value will be zero. If not, the RMP provides an indication of how far the model's performance is from the best-performing model.

While this may seem counter intuitive, it is possible. We take the example of 2008 in the CLEAR watershed to understand why.

Observations:
```
peak SWE = 1047mm, annual Q = 9,640cfs, SWE/Q = 0.11
```
Static Partitioning:
```
peak SWE = 845mm, annual Q = 7,890cfs, SWE/Q = 0.11
Absolute Relative Bias
peak SWE = 20%, annual Q = 18%, SWE/Q (mm/cfs) = 0%
```
Dynamic Partitioning:
```
peak SWE = 891mm, annual Q = 9,437cfs, SWE/Q = 0.09
Absolute Relative Bias
peak SWE = 15%, annual Q = ~0%, SWE/Q (mm/cfs) = 18%
```

In this case, with the dynamic partitioning, both peak SWE and annual Q got closer to the respective observations. However, the SWE/Q ratio got worse. This is because, in this case, annual Q improved a lot more with dynamic partitioning (18% absolute relative bias reduced to ~0%) in relative terms as compared to peak SWE (20% absolute relative bias reduced to 15%). Therefore, the ratio of the two worsened (0% absolute relative bias increased to 18%) as compared to the observed ratio. With the static partitioning, while there were larger errors in both peak SWE and annual Q than dynamic partitioning, the values were in the same ratio as observations with a better match of the ratio. This is a case where the original static method gave the right results for the wrong reason.

This is just one example, but there are other cases that lead to the static partitioning resulting in a better match with observations in spite of peak SWE and annual Q both showing a reduction in bias (blue color points in the lower darker quadrant of Figure R2) . But in most cases, the dynamic partitioning is still the best and when it is not, its performance is not that much lower than that of static partitioning (Figure R1 c).

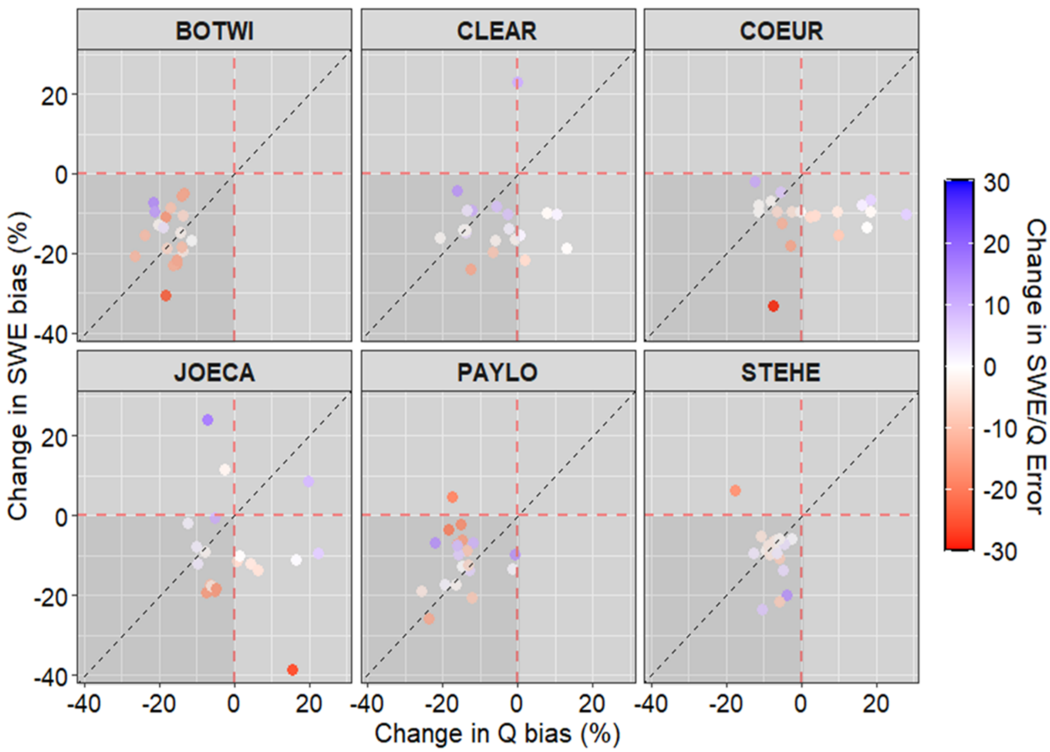

Figure R2: Change in bias for peak SWE, annual average Q and the ratio of the two. Each point corresponds to a year of simulation. Change in bias for the ratio of peal SWE to annual Q is indicated by the color of the point. Each panel corresponds to a watershed. The darker lower left quadrant of each panel, with negative values of change in bias on both the X- and Y-axis indicate years with reductions in both streamflow and SWE bias (better simulation-observation match) after the implementation of dynamic rain-snow partitioning. Red color of the points indicate reduction in bias (improvements in simulation-observation match) for the SWE/Q ratio, while the blue color represents years where the SWE/Q ratio simulation-observation match has degraded after applying the dynamic partitioning method.

We will integrate this analysis and discussion into the manuscript.

Minor Comments:

Line 84: I would quibble slightly that the Columbia River Basin is entirely snow dominated, given that some significantly large areas within the basin are, in fact, rain dominated. I'd suggest slightly rewording this to "The CRB encompasses multiple states and portions of Canada in North America, and snow comprises a substantial fraction of annual precipitation in much of the watershed, particularly high mountain areas."

Yes, we agree. We can revise section 2.1 to add the suggested sentence.

Some of the figures (Figure 2, Figure 3) have some issues with the frames of the figure areas and seem misaligned with subfigures. I'd encourage the authors to address these minor formatting issues.

Yes, we can address this. Thanks for catching it.

For figure 9, the bin labels on the x-axis are slightly confusing. Is bin 4 really 10-20% bias? Or is it 0-20% bias? If the former, were there no results in the 0-10% bias range? Or is this what is meant by "Only bins with at least 8 stations are displayed." If so, in the text, please indicate that "Fewer than 8 SNOTEL stations fall within the range of a 0-10% change in bias, therefore we exclude this range in bias change from these box plots." For figure 9(d) the label of the plot could be more clear. I believe this is the fraction of precipitation to SWE, so perhaps P/SWE (-) would be clearer.

Yes, bin 4 corresponds to 10-20% bias and yes, it is excluded because there are fewer than 8 stations. We can modify the caption to clarify this. And yes, for figure 9(d), we agree that P/SWE (-) would be clearer, and we can make this change.

A modified Figure 9 from the original manuscript including these changes is pasted below.

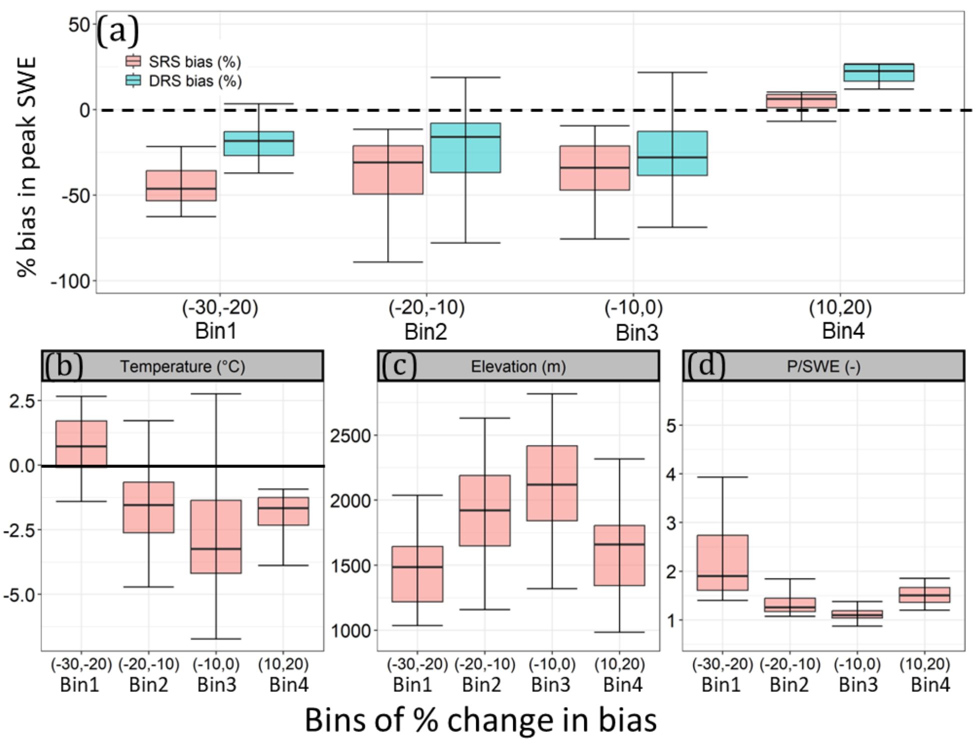

Figure 9. Percent bias in peak SWE and average temperature, elevation, and fraction of accumulative precipitation to peak SWE for bins of % change in bias in increments of 10% change in bias. Only bins with at least 8 stations are displayed and this corresponds to 84% of the stations. Bins 1, 2, and 3 with negative changes in bias correspond to performance improvements with DRS. Bin 4, with a positive change in bias corresponds to degradation in performance with DRS. Note that fewer than 8 SNOTEL stations fall within the range of a 0-10% change in bias, therefore we exclude this range in bias change from these box plots.

For each bin, the distribution of four aspects are provided across the stations within each bin. These include (a) biases of SRS and DRS methods, (b) average daily temperatures on wet days (precipitation > 1mm) during the snow season (November-April), (c) elevation, and (d) the ratio of cumulative precipitation from October 1st until peak SWE has been attained and the peak SWE value from SNOTEL observations. This is to get a sense of potential precipitation undercatch issues.