Reconciliation of comments on “Hydrologic landscape classification assesses streamflow vulnerability to climate change in Oregon, USA” by S. G. Leibowitz et al.

Two anonymous referees provided interactive comments on our manuscript. In addition, H. Gao also provided interactive comments on the manuscript. We thank these referees for providing useful reviews that improved our manuscript. Specific comments and our responses (*in italics*) are provided below.

Reviewer 1

The paper presents a novel way of assessing the impact of climate change on watersheds using hydrologic indices (climate and seasonality). The authors describe these changes using bias-corrected and downscaled GCM data from two GCMs for the entire state of Oregon and then in detail for smaller regions and watersheds. Overall, I find the paper to be extremely well written (albeit somewhat long) with a coherent structure. Just by reading it you can tell it was went through multiple iterations of review by multiple people. The discussion and conclusions are based on sound results. Therefore, I feel that my comments are only somewhat minor in nature. Mostly, my comments deal with the authors better explaining certain items. See comments:

[1] In the abstract, I don’t feel as if the authors do a good job at the beginning of explaining where they are doing this assessment. It takes a few sentences to mention Oregon.

We now state in the second sentence of the abstract that the analysis was conducted in the state of Oregon.

[2] The authors should describe the land use/cover in Section 2.1. Land use/cover can have a large control over hydrology so I believe it’s worth describing.

Land cover is now provided for western and eastern Oregon, based on the 2001 National Land Cover Database.

[3] Why did the authors use the Hamon method to estimate PET and not a more sophisticated method?

The Hamon approach is used because it could be applied to the entire state of Oregon at a 400 m resolution using PRISM temperature data and average day length.

[4] I’m a bit confused where PACK comes from. Is this observed, modeled, estimated?

We now explicitly state that PACK is a modeled snowpack value.

[5] The authors mention the sensitivity of the ECHAM and PCM models, stating that these models represented the highest and lowest global sensitivity. However, this is likely to not be true at the local scale. The authors should present the projections from these two models as percent changes in precipitation or increases in ave. monthly or annual temperature. This will give the readers an idea of just how extreme these GCMs are. It will also help out with the interpretation of the results. Authors can find this information at www.climatewizard.org

Maps depicting changes in monthly precipitation and temperature were included in the original Supplement, but were not called out when discussing the global vs. regional variation of the two GCMs. In response to the reviewer’s comment, we now mention these supplemental figures when discussing *global vs. regional variation in Section 2.3. However, we mapped absolute change in precipitation, rather than percent change.*

[6] Building on [5], most climate change impact studies use many GCMs to bracket the overall uncertainty and overall a more robust conclusion regarding the mean, max, and min impacts. Can the authors comment on why only two GCMs were chosen? I realize that these are supposed to be the max and min projections (see [5]), but using these two models doesn't give an overall idea of the mean projection.

As stated in the 'Introduction' under our objectives, the current analysis is intended as a proof-of-concept application of our approach. We recognize that we could have explored the model uncertainty space better, but we were limited in time and resources. Having said that, we think that we captured much of the basic behavior of the HL mapping using the scenarios we did select. However, we recognize that we did not return to this point in our summary and conclusions. In response, we now added several sentences to the part of the 'Summary and conclusions' section that discusses the limitations of our study, acknowledging that we used a relatively small number of future climate simulations because of the proof-of-concept approach, and stating that a much fuller representation of climate model-based uncertainty would be desirable in future work in order to span the widest possible range of potential future water resources outcomes. We now also mention the proof-of-concept idea when introducing the climate change realizations we selected in Section 2.3.

[7] There are instances in the paper where the authors use the phrase "observed changes". These aren't actually observed datasets but projections, right?

We thank the reviewer for catching this. Two usages of the word "observed" were changed to "simulated".

<u>Reviewer 2</u>

Overall comments: This is an interesting approach to try to depict modeled future climate change effects on hydrology in Oregon. However, the paper draws heavily on previous work and it is difficult to understand the current analyses without reference to these earlier papers.

Major issues:

1. Title and abstract are misleading, because paper does not clearly define or use vulnerability assessments, and there is a strong focus on salmonids not reflected in the title.

Although there are studies that specifically define a vulnerability assessment as an analysis that addresses exposure, sensitivity, and adaptation, this is not universally the case; a quick scan of Google Scholar found many papers that use this term less formally. Regardless, we now use the terms 'evaluate' and 'evaluation' in the title and throughout the manuscript when discussing vulnerability, instead of 'assess' and 'assessment', in order to avoid any confusion. Regarding salmonids, we respectfully disagree that this focus should be reflected in the title. While we do think effects of climate change on salmonids is an important component of the manuscript, the manuscript is predominately about hydrologic effects, and the quantitative analysis we perform is on hydrologic variables. We deal with salmonids interpretively (i.e., with respect to how they might respond to the expected hydrologic changes), and only address these effects in the three case studies; no statewide analysis of salmonid vulnerability is included. Given that, we believe that including salmonids in the title would be misleading.

2. Paper argues that the classification approach is distinctive from analysis of historical data or use of models, but does not make the argument convincingly.

See response to reviewer's comment on p. 2879, l. 16.

3. The outcomes depend heavily on a calculated S value which appears to represent water available for runoff, yet this value does not match observed runoff and this issue is not addressed in the paper.

See response to reviewer's comment on p. 2905, l. 24 and 27.

4. The main conclusions from the three case studies refer to aspects of groundwater, geology, and soils but these aspects of the classification system were not described or defined in this paper.

We added descriptions of the aquifer permeability, terrain, and soil permeability indices to the end of Section 2.2.

5. The choice of only 2 GCM models is not justified in terms relevant to the study question.

See response to comment 6 of Reviewer 1.

6. Results are pretty much what we already know from GCMs directly - what are we learning?

We agree that some of the specific findings of our analysis – such as the loss of snow-dominated area – are already known. However, through the use of our three case studies we demonstrate the effect of geology on these results, e.g., loss of seasonal snowpack and subsequent snowmelt will be mitigated to some degree in the Sandy River basin, compared to the Middle Fork John Day basin, due to high aquifer permeability. Further, we have demonstrated that a classification-based approach, such as provided by the Wigington et al. (2013) hydrologic landscapes, can allow hydrologic vulnerability to be mapped and evaluated.

7. Tables and figures need more explanation.

Additional detail was added per the reviewer's last 11 comments.

Detailed comments

The title of the paper is misleading. "Hydrologic landscape classification assesses streamflow vulnerability to climate change in Oregon, USA" - A vulnerability assessment strictly speaking is not conducted, and the contents of the manuscript focus on three case study basins as well as

As stated in response to the reviewer's major comment #1, we now use the terms 'evaluate' and 'evaluation' when discussing vulnerability, instead of 'assess' and 'assessment', in the title and throughout the manuscript. Regarding the three case study basins, we do not believe that omitting this from the title is misleading, since the title does not state the scale or extent of the evaluation. We omitted this information because it would make the title too long and cumbersome.

p. 2879, l. 10-23. Vulnerability is mentioned here, but not defined. This study does not seem to qualify as a vulnerability assessment, because it does not address exposure, sensitivity, and adaptation. The manuscript contains a lot of discussion about salmonids. Maybe salmonid vulnerability is what the authors were concentrating on? But this does not come across clearly. Needs tightening.

As described in the response to the reviewer's major comment #1, we now use the terms 'evaluate' and 'evaluation' when discussing vulnerability. In addition, we now include in the Introduction our definition of a climate change vulnerability evaluation; i.e., an evaluation of how a system will likely be altered by climate change. Most of the manuscript deals with changes in climate and seasonality class or changes in the Feddema Moisture Index or surplus, at statewide or case study basin scales. These serve as evaluations of vulnerability, since they evaluate alterations in hydrologic properties by climate change. While vulnerability to salmonids is also included, this is not the major focus of the paper.

p. 2879, l. 16 onward. This section makes an argument that classification (as defined by their approach) can represent an improvement over what the authors call "diagnostic" (analysis of historical data) and "prognostic" (analysis of model results) approaches. However, their method is a form of prognostic analysis, because they simply use projected climate (P and T) to drive calculation of a simple index of water availability (a highly simplified hydrologic model, seemingly). So this argument seems specious.

We respectfully disagree. While the climate index is, as the reviewer suggests, a simplified hydrologic model that represents water availability, we use this to provide a static measure of average annual water availability that is used to identify units having similar climate, rather than to provide the kind of time series analyses that are used in diagnostic and prognostic approaches. Further, this index is combined with four other indices (seasonality, aquifer permeability, terrain, and soil permeability) to develop a five element classification system that Wigington et al. (2013) used to classify all 5,660 assessment units across the state of Oregon.

p. 2880, l. 15 "the HL classification to show how results from six potential future climate realizations may impact water resources." Vulnerability is mentioned but not defined.

See response to reviewer's comment on p. 2879, l. 10-23.

p. 2882, l. 5. "accumulation of a seasonal snowpack in the Cascades, with annual average depths ranging from 7620 to 13 970mm (Ruffner, 1985)" Express in SWE, use Snotel?

To the best of our knowledge, there are not any good SWE estimates for the Cascades as a whole. Calculating an average for the Cascades using SNOTEL data would require considerable effort, since data must be downloaded individually by station and by year. Given this level of effort, we believe the value that we previously cited is suitable for an overview of the study area.

p. 2883, l. 14 - Is S' (as implied) based on mean monthly data over the reference periods (i.e., mean monthly values for 1971-2000 and 2041-2070, or a total of 24 values)? Or is S' recomputed for each year of the simulation based on modeled values of P and T? Please clarify.

We clarified that S' is calculated as a 30 year monthly normal.

p. 2884, l. 23 to p. 2885, l. 13: why only two models? Why only use predicted GHG emissions as the criterion for model selection? Given the high variability among models (demonstrated by differences even in the two models shown here) there is great uncertainty about future hydrology. Why not pick models that differed from one another in terms of the drivers of hydrology?

See response to comment 6 of Reviewer 1. Regarding why we focused on predicted GHG emissions as the criterion for model selection, we believe that climate – both precipitation and temperature – is the main driver of hydrology, and that picking the two models with the largest range in response to CO_2

would best bound our proof-of-concept analysis. However, we agree that a fuller representation of climate model-based uncertainty would be desirable in future work, as we now state in Section 4.

p. 2885, l. 14-15: so 1/8 degree data were interpolated to 400-m data? Please clarify. How was the interpolation validated? Were any attempts made to relate the interpolated values to any measured data?

As is explained in the text, the 1/8th degree data were resampled to a 400 m grid using a bilinear interpolation. This essentially smoothes the data over a higher resolution surface. No attempts were made to validate the interpolated values – which are climate predictions – against measured data. As noted on the NCAR/UCAR Climate Data Guide website (https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/regridding-overview), bilinear interpolations are one of the most commonly used methods of climate grid interpolation when regridding to a common grid, and are adequate when used for smoothly varying variables.

p. 2886, l. 19-20. S_ represents the area-weighted monthly watershed positive surplus from each 20 of the n assessment units in a basin, and Ai is the area of assessment unit i . three case study basins (Fig. 1): the Siletz and Sandy, in western Oregon, and the Middle Fork John Day in eastern Oregon.

We are unclear here as to what the reviewer's concern is. Lines 19 and 20 present S^*, as defined by Wigington et al. (2013). We then stated on lines 25-26 that S^* was calculated for three case study basins: the Siletz, Sandy, and Middle Fork John Day. Perhaps the issue is with the "20" in the phrase "from each 20 of the n assessment units" that the reviewer refers to. However, the "20" was not part of the text, but was the line number.

p. 2887, l. 3-8. Because groundwater turns out to be a big deal in mediating results (at least according to the discussion), the way the groundwater index is determined should be presented here, and the indices for the case study basins should be presented in the methods. "Wigington et al. (2013) use the Q/S_ ratio to assess whether the river experiences groundwater losses or gains: a Q/S_ > 1 indicates that runoff is greater than available 5 surplus, and thus suggests groundwater imports (changes in storage are assumed to be zero since 30 year normals are used). Conversely, a Q/S_ < 1 suggests groundwater exports, since runoff is less than available surplus."

A description of the aquifer permeability index was added to the end of Section 2.2. We also mistakenly omitted the source of the data for discharge (Q), which has now been added to Section 2.4.

p. 2887, l. 10-13. First mention of the fact that each unit was assigned an initial class - maybe add a sentence to the introduction summarizing the distribution of classes in Oregon, and include a map showing how basins were classified. "The percentage of assessment units that change class ranges from 4.4% for the ECHAM_B1 realization to 18.3% for PCM_A1b, with a mean of 10% over all six realizations"

The initial class was first mentioned in Section 2.2 of the Methods, which describes the calculation of the initial (1971-2000) hydrologic landscape maps. That text included reference to Fig. S3 in the Supplement, which is a map depicting the initial distribution. We also stated at the end of Section 2.3 that we calculated changes in the various variables. We now clarify that the changes were between 1971-2000 and 2041-2070.

p. 2904, l. 5. "The specific effects of these changes in timing and delivery of S_ are mediated by the geology of the basin. We discuss in detail results from three case study basins to demonstrate how the HL approach can be useful for understanding climate change impacts in diverse hydroclimatic and geologic settings, and to illustrate how the approach could support management." Given that these factors aren't mentioned in the introduction and methods, this comes as a big surprise. The methods section does not contain any information about how parts 3, 4, and 5 ((3) aquifer permeability, (4) terrain, and (5) soil permeability) of the classification system are estimated, nor does the introduction anticipate how these factors might influence runoff response. If this is the big story, it should be mentioned in the introduction, the methods for these parts of the classification system should be presented, and the logic for selecting the case study basins should be clarified.

We added descriptions of the aquifer permeability, terrain, and soil permeability indices to the end of Section 2.2. In addition, we now include geology as one of the factors we mention in the Introduction that can cause variation in the impacts of climate change. Finally, the paragraph describing the Oregon Hydrologic Landscape (HL) characterization in the Introduction now states that the HL classification includes information on geology, which can influence hydrologic response to climate change, and we include a citation for this.

p. 2905, l. 8. "we have demonstrated that the Wigington et al. (2013) HL approach can provide a method for mapping and interpreting vulnerability to climate change." Many methods could be used to map and interpret climate change - but we would like to know whether this method is accurate. The discussion does not address this.

We added a paragraph to the 'Summary and conclusions' addressing the accuracy of our assessment, noting that this is dependent on both the accuracy of the ECHAM and PCM models and the accuracy of the initial HL classification.

p. 2905, l. 24 and l. 27. "the relationship between modified surplus and runoff is not quantified," and "Use of a model that estimates Q from S_ would allow for more objective conclusions." Please elaborate. I don't think it's a matter of objectivity; it seems more a matter of accuracy. It looks like S* deviates significantly from observed Q in the case studies. This suggests that S* does not accurately predict runoff (Q).

The reviewer correctly observes that S does not accurately predict runoff, at least for some basins. We neglected to fully describe the purpose of the S* variable in the Methods, and its intended use. We now explain in Section 2.4 that S* represents two major components of stream runoff – rainfall and snowmelt (minus PET) – but that it does not include changes in groundwater storage of lags from groundwater movement. We state that, knowing the geology of an area, S* and Q can be used to interpret how watershed positive surplus and groundwater contribute to basin runoff, which is what we do in the three case studies. Thus, the very deviations that the reviewer observes are used to interpret basin hydrodynamics. For example, lags observed between S* and Q in the Sandy River basin, along with its high permeability aquifer, suggest that these lags are due to aquifer recharge.*

p. 2916 Table 2. These changes are based on the change in precip or temperature between the reference time period (1971-2000) and the target future time period (2041-2070) in the GCMs? Please expand table caption to clarify.

We added the time frames to the caption, and explain that the 2041-2070 climate and seasonality classes are based on simulated changes in precipitation and temperature.

p. 2917 Table 3. all climates get drier. Makes sense: precip does not change, temp increases. Except for scenario B1 in which some climate classes get wetter. Since scenario B1 has greenhouse gas forcing, it is hard to understand how climate would get wetter.

The B1 emissions scenario has the smallest temperature increases, since it has the least greenhouse gas forcing. Therefore, any local increases in precipitation don't have to compete with as big an increase in ET, making it easier to get circumstances where some climate classes get wetter.

p. 2920. Table 4. All 7 summer-peaking water surplus units move to spring peaking; and 35% of spring peaking water surplus units move to winter peaking.

We did not add this to the table caption, since we feel that the caption should be limited to explaining the table. However, this information was added to the text. Note that the reviewer misread the table; the 35% value is actually the percent of assessment units that remained in the spring seasonality class. The actual change to winter seasonality was much more dramatic (57-68%, excluding PCM_B1).

p. 2922. Table 5. HL class definitions should be provided in the caption; readers should not be expected to look these up in another publication.

Done.

p. 2924. Figure 2. Why are only a subset of basins in Oregon highlighted in these maps? If only some parts of Oregon have been classified by previous work, can those original basins and their classifications be shown? Overall, it appears that the GCM simulations do not adequately capture the very strong spatial gradient of climate, because the changes in climate attributable to model projections show a strong gradient.

The caption for Fig. 2 now states that areas depicted in white did not experience changes in climate class (a similar statement is now included with Fig. 4). Regarding the reviewer's comment about climate gradients, the monthly changes in precipitation and temperature were included as Figs. 24-S9 and S10-S15 the Supplement, respectively. Although some strong gradients exist for individual months, patterns are not consistent over time, and would be averaged out over the annual period that the change in climate class represents.

p. 2925. Figure 3. Again, please clarify why Figure 2 covers only selected portions of Oregon while Figure 3 covers the whole area.

All assessment units experienced some change in the FMI index (Fig. 3). However, as was originally noted in the text (Section 3.1.1), the patchiness observed in Fig. 2 is due to the fact that the magnitude of change in the FMI was small relative to the range of each climate class bin.

P. 2930. Figure 4. Why does seasonality change only in the high-elevation portions of Oregon? Is this simply a reflection of the spatial distribution of snowpack, i.e., it is orographically enhanced?

The fact that seasonality change is limited to the high-elevation areas is because spring and summer seasonality is limited to these areas. As the reviewer suggests, this is reflective of orographic effects on precipitation. Most of the state (4931 out of 5660 assessment units, as was described in Section 3.1.2) has winter seasonality – due to lower elevations – which does not change under these simulated climate conditions.

p. 2931. Figure 5. Part (a) shows the expected pattern from future climate change: increased water available for runoff in winter and declines in summer, due to declining snowpack/shifts to rain in winter, and the expectation that this will enhance summer drought. However, part (b) does not make sense: if absolute changes are negative for all models in the month of July, how can percent changes be positive? Something seems wrong here.

It was previously stated in Section 3.1.2 that a negative departure can have a positively valued percent departure if the denominator is negative, i.e., if conditions during that month represent a deficit. This is now repeated in the figure caption.

p. 2932. Figure 6. Map shows expected changes: biggest losses of water surplus in the mountains for May, June, and July, because models lead to expected losses of snowpack and associated water storage.

We added to the text in Section 3.1.2 that the high losses of surplus in the mountain areas are due to simulated losses of snowpack and associated water storage.

p. 2933. Figure 7. In part (b), it appears that S* anticipates Q by about a month, and is lower than Q, except in Oct and Nov. This implies that the water budget used to estimate S* overestimates actual ET.

Our interpretation was that this difference between S and Q was possibly due to some minimal groundwater imports. However, we now state that this could also be due to potential evapotranspiration overestimating actual evapotranspiration.*

p. 2934-5. Figures 8 and 9. In part (b), that S* is quite different from Q - S* is higher than Q in fall and spring and lower in winter. What accounts for this difference?

The relationships between Q and S for the Sandy (Fig. 8b) and Middle Fork John Day (Fig. 9b) basins were previously described in Sections 3.22 and 3.23, respectively.*

<u>H. Gao</u>

The Hydrologic Landscape (HL) model is interesting and offers a promising venue for modelling. However, it is not completely new. The presented model includes information about climate, seasonality, aquifer permeability, terrain, and soil permeability. For topography, the authors classified into three classes: mountain, transitional, and flat. The classification is based on relief (maximum elevation minus minimum elevation) and total percentage of flatland (slope <1%). After the landscapes classification, the authors evaluate the hydrological change in 2041-2070. We agree with the authors' statement: "A major strength of the HL approach is that results can be applied to similarly classified, ungauged basins." To our knowledge, this approach is reasonable and useful in practice. However, we would like to raise several issues:

1. The model may be over-parameterized. Since too much information (5 categories, including 5 climate classes, 3 seasonality classes, 3 aquifer permeability classes, 3 terrain classes, 3 soil permeability classes) has been considered. In principle, there could be 486 landscapes classes. This makes the model too complicated, and hard to calibrate.

The hydrologic landscape (HL) approach is not a model in the sense that a set of equations are defined for some variable and then parameterized using some optimization approach. Rather, the five quantities we use in the HL maps are indices that are conceptually defined and then evaluated using existing data sets. For example, the six climate classes (very wet, wet, moist, dry, semi-arid, and arid) are based on Feddama Moisture Index thresholds using Eq. 1 and calculated using precipitation and potential evapotranspiration (PET) data. The precipitation data are based on 400 m PRISM coverages, as are the temperature data used to calculate PET. Similarly, the aquifer permeability metric (high, moderate, and low) is based on binning of a statewide aquifer permeability map as is now described in Section 2.2. Thus the different classes are not the result of a single, parameterized model. The manuscript explicitly referred to these as indices (Section 2.2).

2. Land cover information has not been included in the HL model, which is odd. We think this information is essential, as it affects hydrological processes directly.

We agree that land cover is not included, and explicitly stated in the 'Summary and conclusions' that the HLs "do not deal with the influences of vegetation, land use, or other human activities – all of which could influence vulnerability to climate change." We also acknowledged in that same section that these factors could exacerbate or mitigate against climate impacts. The HL map was designed to represent the major geoclimatic factors influencing streamflow. Land use effects would be especially important to consider in heavily urbanized areas. However, these are fairly limited in Oregon.

3. The authors apparently missed relevant literature on landscape-based hydrological modelling. We think the HL model is based on the same idea as FLEX-Topo (Savenije, 2010;Gharari et al., 2013;Gao et al., 2013), but with a different classification approach, since the HL model divided the catchment by climatic, seasonal, geologic, topographic and soil information. FLEX-Topo uses essentially topographic information and land cover for hydrological landscape classification. We suggest reference is made to the mentioned publications, as well as to Winter, (2001).

We thank H. Gao for making us aware of the FLEX-Topo papers. We have included reference to Savenije (2010) and Gao et al. (2014) in the portion of the Introduction on classification. The Winter (2001) paper is the conceptual basis for our approach, and is cited as such in Wigington et al. (2013). We have now explicitly included a reference to Winter (2001) in Section 2.2 of the revised paper.