

We are grateful to the reviewer for the thoughtful and detailed comments. Here we present our response to the comments and plans of revision.

#### Reviewer comment

This paper has the potential to be an interesting contribution to the literature, as the changing roles of temperature and precipitation under a changing climate are of both scientific and practical importance. I did find a number of places within the text that additional information is needed to properly assess this work.

1. I believe this paper quite nicely follows the paper by McCabe and Wolock [2011] that is cited throughout. Specifically, the last sentence of that 2011 paper states, "If temperatures continue to increase, as projected by climate models, the effects of temperature on runoff may become more apparent." I think it is important to add to the introduction that this manuscript appears to be the first such contribution that looks at the relative contributions of P and T under future climate scenarios. It should also be worth noting in the introduction that the joint roles of P and T are examined here and not treated as independent forcings, as was done by McCabe and Wolock [2011]. Adding these points would help to draw context to this work and provide a further case for its relevance in the literature.

#### Author reply

We agree. This study was indeed inspired by McCabe and Wolock (2011) as cited in our paper. We will highlight the novelty of this important literature and explain the contributions of the present study.

#### Reviewer comment

2. p. 4, line 22: The statement is made "although the possible underestimation of the influence of temperature: : has been discussed in recent researches" but only one citation is offered with an "e.g." - either there is only one paper that has discussed this or there are more. If there are more, they should also be cited. If there is only this one paper, the "e.g." should be removed. If this was a recent discussion in the literature, then there should not be a reason to only cite a sample of papers - simply cite them all. This also lends credence to the manuscript.

#### Author reply

We agree. As far as we know, there have been numerous studies addressing the sensitivity of runoff to climate change, but only a few have compared the relative role of P and T. We will try to cite all the relevant papers.

#### Reviewer comment

3. p. 6, lines 10-12: Here it is noted that 10 different land-cover classes are used to account for water demands. Later the authors note that changing land conditions are not considered. To avoid confusion, I would add this limitation here in addition to later in the text. Something like, "Whereas land use is included in the assessment of water demand, the effects of changing land use are not." I know this point is noted on line 19 but perhaps moving that up to line 12 would make this point clearer.

#### Author reply:

In Section 2.1, we intended to describe the model structure, setup, and validation in separate paragraphs. Although the future change in land-cover was not the focus in this study, the model itself is flexible enough to use different land-cover distributions when such datasets are readily available.

Reviewer comment

4. p. 6, line 20: Please explain how the monthly climate inputs were scaled to the HUC-12 watersheds.

Author reply:

The original 4x4 km climate data were scaled to the watershed-level by area-weighted averages. We will clarify this in the revised manuscript.

Reviewer comment

5. p. 7, line 6: The ability to estimate ET well seems an important component to the validity of the results. Are these parameters static in the model, as the manuscript notes that are estimated empirically? It would seem that the ET parameters would adjust under future climate conditions. If they are static, could this bias the estimates of the relative effects of precipitation and temperature under future conditions?

Author reply:

Monthly ET rates by land cover type were modeled as a function of precipitation, PET, Leaf Area Index, and soil water availability. The land-cover distribution, LAI, and the parameters were assumed to be static in the future. We fully agree that the future changes in other environmental factors and the uncertainty in model parameters may affect the results to various degrees, but we believe that the roles of these factors are not likely to be as substantial as precipitation and temperature. It is out of the scope of this study to examine how LAI will change and how the change will affect the relative roles of P and T on watershed hydrology under a changing climate. However, we expect the change in LAI is not likely to alter the key conclusions from this study.

Reviewer comment

6. p. 7, lines 20-21: The R-squared values for only the highest values are reported and not for the all results. This is a bit misleading; the reader cannot assess the performance across the entire CONUS and WRR scales. When looking at figure 2, the captions need to also show the number of data points in each plot. It is unclear why the WRR plots have a different number of points than the CONUS plots and this is not well-explained in the text. There is no mention of the number of gauges used on the validation process. This is a concerning oversight that must be addressed and further reviewed.

Author reply

The R-squared values here were reported for all the results, i.e., the aggregated annual runoff in the entire CONUS and in each WRR. There are 50 points (50 years) in the CONUS plots, and 900 points (50 years multiple 18 WRR regions) in the WRR plots. We agree that comparisons at CONUS and WRR scales are not enough to validate a model. The idea was to verify the model performance at large scales to complement to our previous validations in gauged sites (Caldwell et al., 2012; Sun et al., 2015), and also to show readers the general water balance in the CONUS. The R-squared values evaluated at large scales

are usually better than that at site level, because the overestimations and underestimations at different sites may offset each other.

References:

- Caldwell, P., Sun, G., McNulty, S., Cohen, E., Moore Myers, J., 2012. Impacts of impervious cover, water withdrawals, and climate change on river flows in the conterminous US. *Hydrology and Earth System Sciences*, 16(8): 2839-2857.
- Sun, S. et al., 2015. Drought impacts on ecosystem functions of the US National Forests and Grasslands: Part I evaluation of a water and carbon balance model. *Forest Ecology and Management*, 353: 260-268.

Reviewer comment

Equation 1: Consider writing in terms of R so that the reader can look at the P and T terms on one side of the equation.

Author reply

We agree. We will rephrase the equation as:  $R = P - E_T + dS_M/dt + dS_P/dt$  (runoff equals precipitation minus ET and changes in soil moisture and snowpack).

Reviewer comment

p. 8, lines 5-7: I was left wondering if the ET is oversimplified to the point that it does not consider the non-linearity and feedbacks in the climate system. Could the authors comment on this?

Author reply

The Equation (1) here was only used to describe the basic idea of water balance. ET is estimated by two steps: (1) a maximum ET is calculated as a function of precipitation, PET, Leaf Area Index; (2) the actual ET is simulated at monthly basis using Sacramento soil water accounting model, which considers water availability in different soil layers. Thus, ET and runoff indeed respond to climate change nonlinearly in the WaSSI model.

Reviewer comment

p. 11, Section 3.1: Cite the corresponding evidence (figures) for these observations.

Author reply

The corresponding figures were cited in the first sentences of each paragraph.

Reviewer comment

p. 12, lines 17-19: What are these observations based on? I think the figures are created based on equation 2 but please add this detail here.

Author reply

Agreed. We will cite Eq. 2 in this sentence.

Reviewer comment

p. 13, line 1: GCMs do not treat changes in temperature and precipitation separately, correct? So why is the word independent used here. Please clarify.

### Author reply

We divided the effects of changing P and T into three components to quantify their relative roles: (1) the independent effects of P, i.e., the changes in R caused by the changes in P, while T is assumed to be static; (2) the independent effects of T, i.e., the changes in R caused by the changes in T, while P is assumed to be static; (3) the effects of interactions between P and T, i.e., the changes in R that cannot be explained by the above two independent effects due to the nonlinearity of hydrologic processes. These independent effects are estimated by hydrologic modeling experiments using different P and T data (i.e., different combinations of data in the baseline and future).

We understand the reviewer's concerns that P and T are related to each other in the climate system, and their changes should occur simultaneously in "reality" in the future. However, we believe that such sensitivity-test methods still make good sense statistically when it comes to quantifying the relative roles.

### Reviewer comment

p. 14, Use of the Wilcoxon signed rank test: I believe more detail is needed here. What are the sample sizes used in the test - is it N=20 for the number of GCMs per region? More details about the test and the assumptions are needed. This statistical test assumes that the data come from the same distribution, which may not be the case if the 20 data points come from 20 different GCM results. Furthermore, on p. 14, line 22, are these differences significant? An increase in T from 60 to 63 percent does not seem that would be significant. Lastly, the Wilcoxon test does not account for the substantial variability inherent in each of these estimated changes - variability from the GCM output, the downscaling, etc. I would think there would be much less of change to find any significant differences if the actual variability of the estimated changes were included.

### Author reply

(1) Yes, N=20. We agree that the assumption of coming from a same distribution may not always hold true due to the limited sample size and the high regional diversity. However, the other assumptions of the test seem to fit our case well, such as coming from two related and matched samples, and each pair being chosen randomly and independently. We believe that the test of significance can be a good indicator of the agreement among the different GCMs. We plan to revise this part.

(2) The reviewer is right. We will revise this sentence to avoid misleading. The difference of T contributions between S1 (RCP4.5/2030s) and S2 (RCP4.5/2080s) scenarios is not significant at the CONUS level, while the difference between S3 (RCP8.5/2030s) and S4 (RCP8.5/2080s) is much larger. This can be explained by the different climate projections under RCP4.5 and RCP8.5. RCP4.5 is a relatively conservative emission scenario, under which the increase in T from 2030s to 2080s is not as significant as that under RCP8.5.

(3) The Wilcoxon test is only used for the relative contributions of P and T, which are estimated from 20 GCMs, one downscaling model, and one hydrologic model. We did not specifically quantify the uncertainties from downscaling model and hydrologic model, but the general trends derived from different downscaling or hydrologic model might be similar (although larger uncertainty is expected at regional scale). After all, the projected changes in P and T by different GCMs are probably the most important uncertainty source.

### Reviewer comment

The authors should be commended for the discussion section, which helps to address the limitations of their analysis and place appropriate qualifiers on the interpretation of the results. Nicely done.

### Editorial comments:

There are numerous places where there is awkward wording or pronouns missing. I was able to catch some of them in my reading but more still exist in the manuscript. Please do a thorough read-through to ensure the manuscript addresses this.

p. 3, line 11: Change “the” to “a”

p. 3, line 21: Change “results” to “result” and “decrease” to “decreases”

p. 4, line 4: The authors use climate change throughout. Change “global warming” to “climate change”

p. 4, line 12: Change to read “under a warming climate”

p. 4, line 15: Change to “CONUS has reached”

p. 4, line 22: Change to “has been discussed recently”

p. 5, line 6: Delete “the state of the art”

p. 5, lines 6-8: Delete the sentence that starts “In another word: : :” This is confusing and does not add anything.

p. 5, line 9: Should this read, “quantify the relative contributions of: : :”

p. 5, line 12: After “spatial pattern of runoff change” add over what period you are examining the change.

p. 6, line 7 (and throughout): HESS guidelines are very clear about how to abbreviate evapotranspiration and how to reference other variables. The authors should review these practices and adjust accordingly. See more information below.

Specifically (from [http://www.hydrology-and-earth-systemsciences.net/for\\_authors/manuscript\\_preparation.html](http://www.hydrology-and-earth-systemsciences.net/for_authors/manuscript_preparation.html)):

“Multi-letter variables should be avoided. Instead use single-letter variables with subscript (e.g. ERMS instead of RMSE, or ET instead of ET). Single-letter variables or parameters and user-defined function symbols should be italic (e.g.  $x$ ,  $Y$ ,  $_{}$ ,  $f(x)$ ). Multiletter variables, if they cannot be avoided, should be roman.”

p. 14, line 15: What is “it” in this sentence?

p. 18, line 2: Precipitation, temperature and runoff should be abbreviated to be consistent with your use of these terms elsewhere in the discussion section and throughout the manuscript.

Tables and figures should stand alone. In the tables, only S1-S4 are referred to as scenarios with no other explanation. These scenarios should be explained on each table. Abbreviations should be spelled out (see examples in figure 1, CONUS as an example where this is not done; figure 2, R and P are used in the figures but not explained in the caption.)

### Author reply

We will check through the manuscript more carefully and address all the issues raised by the reviewer. We will also rephrase the captions of all the figures and tables.