Authors have performed an interesting study to assess the role of soil storage on climate-evapotranspiration (ET) interactions in three mountainous catchments using a distributed ecohydrologic model. In particular, the role of soil storage is considered by incorporating uncertainty of soil storage parameters in deriving precipitation, recharge and temperature relationships with ET. The manuscript is very well written and discussion of the results is very clear. However, the readers can benefit from a more focused conclusion summarizing main take home messages of the paper and its broader impact.

Reply: We thank the reviewer for the supportive comments and address the detailed comments below. Both referees suggest a more focused conclusion. We have edited the conclusion with this suggestion in mind -- some points have been moved to the discussion and other text has been removed. The edited conclusion now reads as follows:

We demonstrate how subsurface storage and drainage properties (AWC and parameters that control lateral redistribution) interact with climate-related drivers to influence ET in three western U.S. mountain watersheds with distinctive precipitation regimes. These watersheds reflect conditions found in many other western U.S. snow-dominated systems, where summer water availability is influenced by the magnitude of precipitation, timing of soil moisture recharge and spring temperature and its effect on snowmelt. We found that, for our three watersheds, estimates of longer-term average (15-year) watershed-scale ET vary across a range of physically realistic storage/drainage parameters. For all watersheds, the range in long term mean ET estimates across AWC estimates (e.g., mean ET at a high AWC versus mean ET at a low AWC) may be as large as inter-annual variation in ET, suggesting that the influence of AWC and drainage can be substantial.

Our results also point to the importance of lateral redistribution as a control on ET, particularly for CA-SIER. Only a few studies have emphasized the role of lateral redistribution in plot to watershed scale climate responses in the Western U.S. (Barnard et al., 2010; Tague and Peng, 2013). For the CA-SIER site, our model results suggest that there can also be interactions between AWC and hillslope to watershed scale redistribution as controls on ET. Lateral redistribution was less important for the CO-ROC, where summer precipitation was a more important contributor to annual ET values and the least important for the wetter OR-CAS site. Results emphasize that the role of subsurface properties, including both storage and drainage, will be different for different climate regimes.

These results have important implications both for predicting ET in basins where data is not available for calibration and for understanding and predicting the spatial variability of ET within a basin. AWC also affects the sensitivity of annual ET to climate drivers, particularly in the two more seasonally water-limited basins. Although the three watersheds show different responses of annual ET to these climate drivers, there are values of AWC that would eliminate these cross-basin differences. These sensitivities highlight the need for improved information on spatial patterns of subsurface properties to contribute to the development of science-based information on forest vulnerabilities to climate change. Improved accounting for plant accessibility to moisture has improved model-data ET comparisons in previous modeling studies at regional and global scales (Hwang et al., 2009; Tang et al., 2013; Thompson et al., 2011). With expected decreases in fractional precipitation received as snow with climate change (Diffenbaugh et al., 2013; Knowles et al., 2006), we might expect soil storage to play a more important role in providing water for forests in the future. Improved understanding of how climate and subsurface storage/drainage combine to control ET can enhance our understanding of forest water stress related to increased mortality (van Mantgem et al., 2009). Western U.S. forests show substantial vulnerability to drought, with declines in productivity and increases in mortality and disturbance in drought years (Allen et al., 2010; Hicke et al., 2012; Williams et al., 2013).Understanding these ecosystems' responses to primary climate drivers is of particular concern given recent warming trends (Sterl et al., 2008) and multi-year droughts (Cook et al., 2004; Dai et al., 2004). Identifying the physical conditions in which our ability to estimate ET is most sensitive or limited by knowledge of subsurface geologic properties helps to prioritize regional data acquisition agendas. Integrating results from recent advances in geophysical measurements and models such as those emerging from Critical Zone Observatories in the U.S. and elsewhere (Anderson et al., 2008) will be essential for analysis of climate ET interactions.

REVIEWER COMMENTS CONTINUE BELOW

Detailed comments:

1) Authors have used a spatially distributed model to perform simulations across three catchments but the final results are aggregated at the catchment scale. It will be very interesting to see how these climate sensitivities change across the catchment? Are they observing differences between uplands and lowland areas?

REPLY: We agree that the spatial patterns of these climate sensitivities would be interesting to observe across the catchment, however a thorough analysis is beyond the scope of this paper and we will explore more spatial patterns in future work. It would afford less room in the manuscript length to address how these climate sensitivities interact with subsurface properties, which we believe to be the novel contribution of our work.

2) Does the sensitivity of ET change for different land cover types in a given catchment?

REPLY: We expect that ET estimates would vary with different land cover types in each catchment. Two of our catchments, CA-SIER and OR-CAS, are uniformly covered in conifers. CO-ROC, which is significantly larger than the other two catchments, is comprised of other land types including meadows and rock. We expect that these land cover types are responding to climate drivers differently than the conifers. We have added text to acknowledge this important point:

> **We also note that CO-ROC is considerably larger than our other two study sites and, as such, includes significant fractions of other land cover including rock**

and meadow. We expect the different vegetation types to confound the response of ET to climate.

3) What about sensitivities of plant transpiration and NPP to AWC and precipitation?

REPLY: We explored the response of transpiration and NPP to the climate drivers presented in our study and found similar patterns in response across watersheds. We chose to focus on ET because we were able to validate our model estimates of annual values to field based observations in each catchment, which we believe strengthen our paper's results and discussion. We were unable to find similar transpiration observations at an appropriate scale/temporal resolution for model validation. We used annual NPP estimates to validate our carbon cycle, but chose to focus our presentation on the sensitivity of our hydrologic metric (ET) because it is more directly related to availability of soil moisture (AWC). Adding discussions of NPP/plant transpiration would also substantially lengthen an already long paper but we will consider this in future work.

4) Can authors specify which of the soil parameters generate most of the scatter in their results like in Figure 5 or 6? In other words, what is the most sensitive parameter? Is the most sensitive parameter different among the catchments?

REPLY:
We examined the influence of individual soil parameters to the sensitivity of ET estimates (not shown) and found that the sensitivity often varied with combinations for parameters rather than a single parameter value. For all catchments, streamflow estimates were most sensitivity to the 'm' parameter that controls the decay of conductivity with depth and defines an effective soil depth.

5) Can authors specify which metric they used for annual NPP during calibration (page 7899)?

REPLY: We used estimates of annual NPP that we found in peer-reviewed literature to define a minimum and maximum range of NPP values then selected parameters that fell within this range. That range of values is provided in Table 2. For calibration we selected parameters that fell within this NPP range and also provided reasonable estimates of streamflow based on the NSE and the daily bias.

6) Since R75 is not the actual recharge, I suggest authors rephrase it to timing of potential recharge.

REPLY: We agree that timing of potential recharge is more appropriate and have rephrased as follows:

> **To assess the impact of timing of potential recharge (as influenced either by year to year variation in precipitation timing, snowmelt or rain-snow partitioning) we calculate R_{75}, the day of water year by which 75% of the total potential annual recharge has occurred.**

7) Can authors briefly describe patch elements in RHESSys (page 7898)?

REPLY: Patch units are not necessarily grid shaped, but instead are delineated based on landscape characteristics including elevation, land cover classification, and aspect. Average patch sizes range from 90 to 8100m2 with average patch size of 3600m2. Soil, vegetation and climate processes are calculated at the scale of the patch.

8) A brief description of snow module will be helpful. How the results are impacted by the snow parameters?

REPLY:
We have added the following text to the methods section that describes the RHESSys model:

RHESSys partitions rain to snow at a daily timestep based on each patch's air temperature. Snowmelt is estimated using a combination of an energy budget approach for radiation-driven melt and a temperature index-based approach for latent heat-drive melt processes.

We agree with the reviewer that results are potentially sensitive to snow parameters that control the rate of accumulation and melt. However we assume that in order to achieve reasonable rates of model performance relative to daily streamflow observations, the snow parameters used are reasonable and provide a basis for assessing the sensitivity to subsurface characteristics, which is the central focus of this paper.