Interactive comment on "Evapotranspiration feedbacks shift annual precipitation-runoff relationships during multi-year droughts in a Mediterranean mixed rain-snow climate" *by* Francesco Avanzi et al.

Anonymous Referee #1

Received and published: 25 September 2019

The manuscript by Avanzi et al. is generally well-written and well-referenced. They utilize a hydrologic modeling approach to quantify how a mountain watershed responds to short duration (sub-decadal) drought episodes, and find that how simulated evapotranspiration responds to drought conditions is a major issue with regards to runoff estimation. While unsurprising, this finding is important to motivate improvements in ET in hydrologic models in mountain regions. Overall, the modelling approach is acceptable, but will need some additional information and analysis (noted below), especially with regards to the ET estimation approach and some additional longer simulations.

I am also a bit skeptical of the fact that a groundwater model was not used to show whether or not such a model is necessary to accurately capture hydrologic responses in volcanic, subsurface flow-dominated basins. I liked the shift identification approach for changes in precipitation-runoff relationships. I also thought it was a valuable addition to include a range-wide assessment to highlight not only the application of the approach but also to show how basin response to drought varies as a function of elevation. Some general and specific comments follow below.

General comments: 1. The recent 2012-2015 drought period considered by the authors is inconsistent (P2 L33) with the official declaration period of drought on the website provided (P3 L25), which gives 2012-2016. The modelling exercises will need to be repeated to include water year 2016 if the authors would like to stick to the official declaration of drought. As they do note that the results are sensitive to the duration of drought episodes, it would be worth including (and comparing) the 2012-2015 (what could be called 'peak drought' conditions) and 2012-2016 (the 'official drought') periods. I think it would strengthen the paper to include some additional modeling studies of longer droughts, even if these are hypothetical. Many water agencies (and model experiments) use 'design' droughts based on a few known drought episodes lumped together or informed by past extended droughts. While I see equation 3 was applied by lumping results together on Page 6, I would like to see how a few iterations of a design drought of decadal (or longer) scale change (or do not change) the results.

2. PRMS is a fine modeling approach, but in regions with known groundwater/surface water interactions, such as the volcanic Almanor sub-basin, why was a groundwater model not also developed and coupled to PRMS (i.e., GSFLOW?). I understand this may not improve results, but without showing that it does not, it leaves the reader wondering if such an addition would improve results.

3. P8 L5: The authors misconstrue the definition of 'warm snow drought', perhaps because the definition can be misleading if no thresholds in precipitation anomalies are specified. Though the 2012-2015 drought may have had greater precipitation than

the 1977 drought, many of these years in both drought periods satisfied the dry snow drought conditions demonstrated in Hatchett and McEvoy (2018). Warm snow droughts should only be defined by years with near to above-normal precipitation and below average snowpack.

4. P8 L25: Snow/rain ratios should not be based upon a single point in time value of accumulated precipitation and snowpack at that time (April 1), as this neglects numerous factors that may be controlling the state of the snowpack on this date. For example, if a third of an above-average snowpack was lost during a warm, humid period in late March, the value on April 1 would not correctly represent the fact that otherwise a year had experienced above normal snow and perhaps above normal rain/snow ratios. Just a hypothetical example. I suggest calculating snow fraction over the course of the year and using the total precipitation estimated as snow divided by total precipitation (ending on Apr 1) as a more robust metric.

5. The procedure used to estimate ET is very interesting, but as this method has not yet been accepted by the scientific community (as noted by a submitted paper on P5 L4 and L6), I need to see comparisons of these results with some standard, easily implementable methods to calculate ET.

6. The concept of climate-driven ET elasticity is very cool. I would recommend a schematic figure be added to highlight this concept. The additional model runs using longer drought episodes (warm and dry versus cool and dry at decadal time scales) could play into making this figure more robust by helping constrain the temporal and climate condition sensitivity of the elasticity.

7. The map figures (Figure 1) are not up to publication standards. They need to be projected with latitude and longitude coordinates. The bins of precipitation are far too large and substantial precipitation variability is lost. The inset map needs to be projected and should show the entire west coast of North America (or at least the western United States). I suggest simply binning precipitation by 50 mm increments to better

highlight topographic gradients. The geologic mapping appears to be hand drawn (and is of very poor quality) and needs to be markedly improved. The map should also show the locations of the stations used in the study since they are referenced in the main text.

Specific Comments: P3 L11: Please change all instances of Oroville Lake to the correct title, "Lake Oroville"

P3 L26: When referring to climate, one must not neglect temperature if discussing precipitation. Please change to "dry, hot summers and wet, mild winters".

P3 L28: It would help to add a figure demonstrating the basin hypsometry to quantify 'most'. Does "most" mean 59% or 92%?

P4 L4: Add 'anomalous' before 'low'

P4 L13: Should be "Cascade Range".

P4 L14: Suggest to add more geologic context, specifically on soils.

P4 L22: Please provide the temporal resolutions used. Perhaps a table could be used?

P4 L22: What are the resolutions of the spatially distributed indices?

P4 L29: What types of precipitation gauges were used? Were gauges heated? Are there concerns for undercatch?

P5 L1: Which PRISM products were used, 4 km or 800 m? Monthly or daily?

P5 L25: What alpha level was significance assessed at?

P8 L1: I'm a bit confused here, is the paper referring to flow below Oroville, or total inflow to Oroville? If the former, these numbers need to be prefaced by discussion of water deliveries that may have been subject to changed allocations in response to the drought. Similarly, the storage value in the reservoir doesn't really add much, given that a reservoir's operational goals might be to completely drain the reservoir by the end of

the water year. If additional context is provided, then this number becomes meaningful. Last, please correct 'norm' to the proper spelling 'normal'.

P8 L3: For consistency with the discussion on the 1987-1992 drought on P8 L8, please include the temperature anomaly for the 2012-2016 (note I used the official period there, not 2012-2015).

P8 L9: Section should be singular.

P8 L9: The start of this paragraph is a bit strange, i.e., is there any reason to suspect that runoff seasonality should not be preserved? I think this sentence could be removed or replaced with some more insightful findings. Perhaps discuss any temporal shifts?

P10 L11: Can you add some additional clarity about these winter precipitation events? I would expect these to be extreme precipitation rain-on-snow events to produce peak flows, but as it is written, any precip event could generate a peak flow event.

P10 L30: I am balking at the use of 'observed' water balance, as that implies that the water balance has been completely observed, when in reality it is merely an estimate based upon models for precip (PRISM) and ET (NDVI-GAM approach). "Estimated" might be a better word but could confuse readers against "modeled". I leave this to the authors to ponder if a better descriptor could be used.

P11 L8: What is "soft data"?

P12 L25: Suggest to add a citation for the second and third sources of uncertainty.

P13 L6: I like the concept of tree mortality (despite the myriad complexities controlling tree mortality in addition to ET like pests, disease, etc), but I feel like this sentence detracts from the previous, powerful statement defining climate elasticity of ET. Ending the paragraph with this definition and instead another elucidating sentence about the value of this metric would be a strong way to close out this subsection.

Table 1: Please add the period of record means (or medians) for each variable. These

statistics are not provided in Figure 2 as the caption implies.