Response to reviewer 2

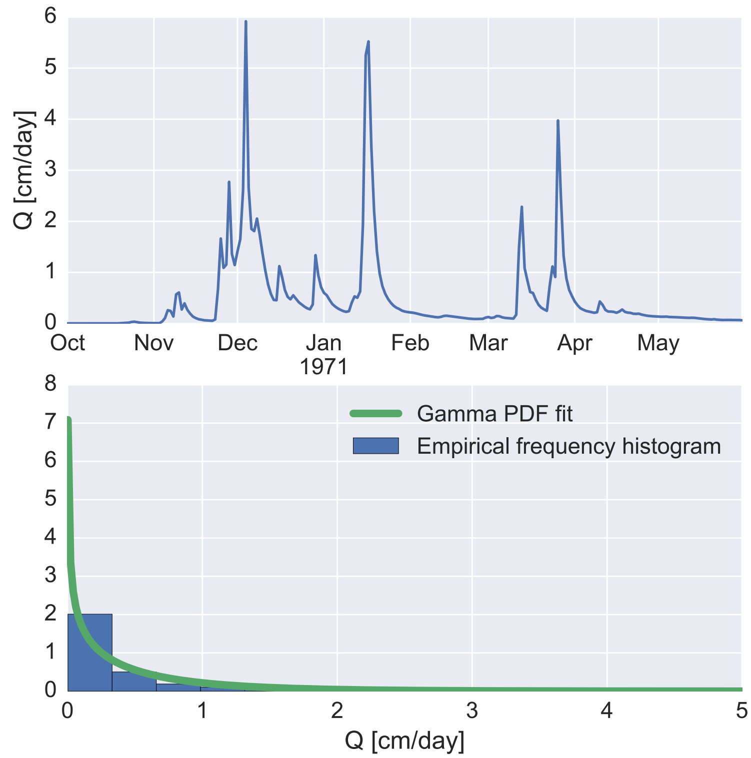
We thank the reviewer for an extremely thorough review and for numerous constructive suggestions. In the following, we have addressed the reviewer's primary issues, which relate to the contextualization of the manuscript objectives and findings, manuscript ordering and organization, and visualization of data.

Response to major comments

M1: The reviewer makes two important points in this first comment:

- **Method choices:** We completely agree. While the relevance of each choice is detailed in the section *Section 2.2: Overview of the methods varied across recession analyses*, the introduction would benefit from a brief overview of these choices and their prevalence in the literature. We will include this in the revised manuscript.
- **What defines a good catchment?** We agree with the reviewer that it is important to describe the characteristics of a catchment that would be relevant to our study. We will include wording clarifying that the study focuses on forested, relatively steep, rainfall dominated catchments, without significant snowfall and without significant regional groundwater systems.

M2: We completely agree with the reviewer that erratic streamflow regimes would be best for this type of study. Fortunately, these study catchments, from coastal California and Oregon, exhibit highly seasonal [Fatichi et al, 2012] Mediterranean climates. Consequently, the watersheds exhibit high inter and intra-annual variability in streamflow. Quantitatively, *Botter et. al.* [2013] define erratic vs. persistent streamflow regimes using a gamma distribution fit to the streamflow empirical frequency histogram. Erratic regimes are those for which the probability distribution function (PDF) is monotonically decreasing in Q (mode at zero flow, but with a heavy tail), and persistent regimes are those for which the PDF is humped at some value of Q greater than zero. For the Eel River watershed (one of the watersheds featured in our study), we have performed many such probabilistic analyses [e.g. Dralle et. al., 2015], and most pacific coast Mediterranean watersheds are classified as erratic by this metric. To demonstrate, we present a typical year taken from the USGS gage on the Eel River watershed, at the Scotia, CA, along with the corresponding period-of-record streamflow PDF derived *only* from wet season months (Nov – April):



Even without including the dry season period (May – Oct) in the above PDF, the best fit gamma distribution is clearly monotonic, indicating an erratic streamflow regime.

The reviewer's comment, however, indicates that we did not clearly describe the features of the flow regimes of the study catchments, which is critical for understanding the relevance of the results. We will follow the reviewer's suggestion to include more information on the catchments' streamflow regimes, climatic features of the region, and a plot similar to the one above demonstrating the highly variable nature of the flow time series.

M3: We thank the reviewer for this aesthetic suggestion. The edited manuscript will include improved labels (especially in Figures 3 and 4) to make the individual method choices more clear. For more details on figure changes, see **m16.**

M4: We agree with the reviewer; the peak filtering approach could be described more clearly. The edited manuscript will include a new figure illustrating the peak selection algorithm, along with a sketch demonstrating the method for determining recession end.

M5: The reviewer makes a good point. The edited manuscript will outline these research questions more clearly in the introduction, preparing the reader for the more complete description found in Section 2.3, which cannot reasonably be presented prior to outlining both the method combinations (MSCL) and the recession measures (a, b, T_R).

M6: This is a good question. Instead of single values for *a, b,* and *Tr*, our analyses provide populations of these variables for each catchment. However, to rank catchments, we needed single number descriptors of the population. Obvious choices could include the mean and median for measures of central tendency, and standard deviation or the inter-quartile range for variability. We did not want the occasional erroneous fit confounding our rankings, and so we chose to use the median and inter-quartile range, which are robust against biasing effect of outlier fits. We will add language in the manuscript explaining this choice.

M7: We agree with the reviewer; the paper could benefit from some sort of summary figure detailing the steps of analysis. We will add a decision tree to the edited manuscript, detailing the ordering of extraction and fitting, the variables derived from these procedures, and the subsequent analyses performed on the populations of recession measures.

M8: We thank the reviewer for this observation. We will take time during the first revision to separate any discussion points from the results section, and vice versa.

M9: We agree with the reviewer, this information should and will be included in the table. We will also include discussion on the potential effect of different record lengths on the results.

M10: We agree that the numerous box plots may be somewhat un-inspired. For full details on numerous figure changes, see **m16**.

<u>Response to minor comments</u>

m1: The edited abstract will include discussion of these points.

m2: This will be changed.

m3: We thank the reviewer for mentioning this important citation; it will be added.

m4: Thank you, this would be useful to include. We will add a few sentences mentioning the various motivations for event-scale analysis.

m5: There are few studies prior to the mid-2000's that extract individual recession events. Some, such as *Wittenberg* (1999) extract individual events for non-linear fitting, but provide no methodological information concerning the minimum recession length. In the context of bulk recession analysis, the method introduced by *Brutsaert and Nieber* (1977) removed the need to identify the start/end of a recession event, and so "minimum length" is not typically discussed. Still, some studies prior to 1977 (e.g. *Howe* 1966) do extract individual events, and some mention minimum length requirements. We will cite these papers, along with some bulk recession analysis papers that include minimum length requirements.

m6: The fitting procedures will be more thoroughly described.

m7: The scale correction procedure will be more thoroughly described.

m8: We will clarify this (e.g. variance of *b* is a higher order moment of the distribution of *b*), thank you.

m9: We thank the reviewer for catching this. We will explicitly cite the Figure number, and will also clearly delineate which figures present Elder Creek results, and which present results for all watersheds.

m10: We agree with the reviewer and will remove all but the last two sentences of the first paragraph in Section 3. The description of the Spearman rank will be relegated to the methods section.

m11: This may be related to the fact that these figures were saved as .png files. The revised manuscript will include higher resolution versions of these figure images, which should fix the blurring issue.

m12: We thank the reviewer and will implement this suggestion (relevant changes described in **m9**).

m13: We agree that this is somewhat unclear and will clarify this reference.

m14: We only reference Figure 4 to give the reader a sense of the range over which the recession measures (*a, b,* and *Tr*) typically vary. The important results here are that some measures (e.g. *a*) were found to be considerably more robust with respect to ranked analysis than others (e.g. *Tr*). This has implications for comparative recession analyses, where the relative values of recession measures are used to classify or contrast catchments. We will make this clearer in our discussion.

m15: We will transfer this analysis to the methods section and present the plot in the results section.

m16: We agree with the reviewer, the plot could be re-arranged to better facilitate and match the discussion section. In order to address the reviewer's general concerns about figure quality, the following specific changes to figures 3, 4, and 5 will be implemented:

- Figure 3 will lump **00, **01, **10, **11 to make 4 plots, rather than 16 separate boxplots -- the message being that concavity and linearity choices are the primary drivers of fit quality.
- Figure 4 will display on median(b) vs. (**00, **01, **10, **11) and IQR(b) vs. (00**, 01**, 10**, 11**), since the primary finding is that median(b) increases along the sequence **00, **01, **10, **11 and IQR(b) decreases along the sequence (00**, 01**, 10**, 11**). This strategy will be used for *a* and *Tr* as well.
- Similar to figure 4 changes, figure 5 will plot only distributions of medians and distributions of IQRs for 0***, 1***.

In all figures, we will add a code legend, and will implement the reviewer's suggestion to use more intuitive labels for the methods (i.e. 0000 → mscl, or 0101 → mScL).

m17: We believe the reviewer misunderstood the purpose of this plot. These are exactly as the reviewer suggested: plots of *a, b,* and *Tr* across all methods for a single catchment, Elder Creek. As the reviewer suggested earlier, however, this confusion could be avoided with more clear labeling when the results are relevant to Elder Creek, or to all catchments. We intend to more clearly label plots in this way.

m18: We will encode this information into a decision tree and pair this with a diagram showing the steps of analysis (also see **M7**).

<u>Bibliography</u>

Fatichi, S., V. Yu Ivanov, and E. Caporali. "Investigating interannual variability of precipitation at the global scale: Is there a connection with seasonality?." *Journal of climate* 25.16 (2012): 5512-5523.

Botter, Gianluca, et al. "Resilience of river flow regimes." *Proceedings of the National Academy of Sciences* 110.32 (2013): 12925-12930.

Dralle, David N., Nathaniel J. Karst, and Sally E. Thompson. "Dry season streamflow persistence in seasonal climates." *Water Resources Research*

(2016).

Wittenberg, Hartmut. "Baseflow recession and recharge as nonlinear storage processes." *Hydrological Processes* 13.5 (1999): 715-726.

Brutsaert, Wilfried, and John L. Nieber. "Regionalized drought flow hydrographs from a mature glaciated plateau." *Water Resour. Res* 13.3 (1977): 637-643.

Howe, J. W. "Recession characteristics of Iowa streams: Part I – Temporal and areal distribution of recession constants." (1966).