Reviewer 2

This paper is focused on the extremeness of recent drought events in Switzerland by looking at different types of drought, including meteorological, hydrological, agricultural, and groundwater drought. The paper is a new research study and is generally well-written as it explains the methodology, the mathematical framework and the assumptions used. However, the application research part needs minor improvements to verify the novelties of the method employed in the study area.

Based on this general comment the following points should be addressed and clarified.

Section 3.1.2 (B) Simulations. Please provide information about hydrological model application showing briefly calibration and validation statistics for the 140 catchments of the PREVAH model. Please also show application model efficiency for the selected six (6) catchments. A comparison table with the observed and simulated drought characteristics (deficit and deficit duration, and estimated observed and simulated return periods) should be provided to demonstrate the successful application of the method.
Reply: *We added some information on the calibration and validation statistics obtained by* Köplin et al., 2010 *for the 140 basins used in the calibration procedure: 'The model was calibrated on runoff time series from 140 mesoscale catchments covering the different runoff regimes in Switzerland resulting in Nash-Sutcliffe efficiency values >0.7 for most catchments both in the calibration and validation periods* (Köplin et al., 2010)*'. The observed and simulated drought characteristics are distributions (see Figure 1 for an example catchment below). As suggested, we added a Table to the manuscript listing median of the observed and simulated drought characteristics i) discharge deficit duration, ii) discharge deficit, iii) precipitation deficit duration, and iv) precipitation deficit. Figure 2 included in the response to the reviewers displays the estimated return periods derived from observations and*

simulations. Summarizing and including these results in the manuscript would make it too long.

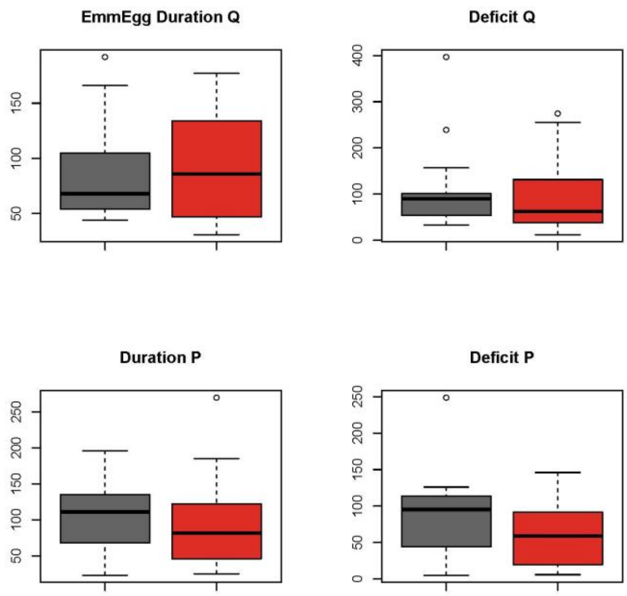

Figure 1: Comparison of the observed distribution (grey boxplot) and the simulated distributions (red boxplots) of 1) discharge deficit duration (upper left), discharge deficit (upper right), precipitation deficit duration (lower left), precipitation deficit (lower right) for the station Emme-Eggiwil.

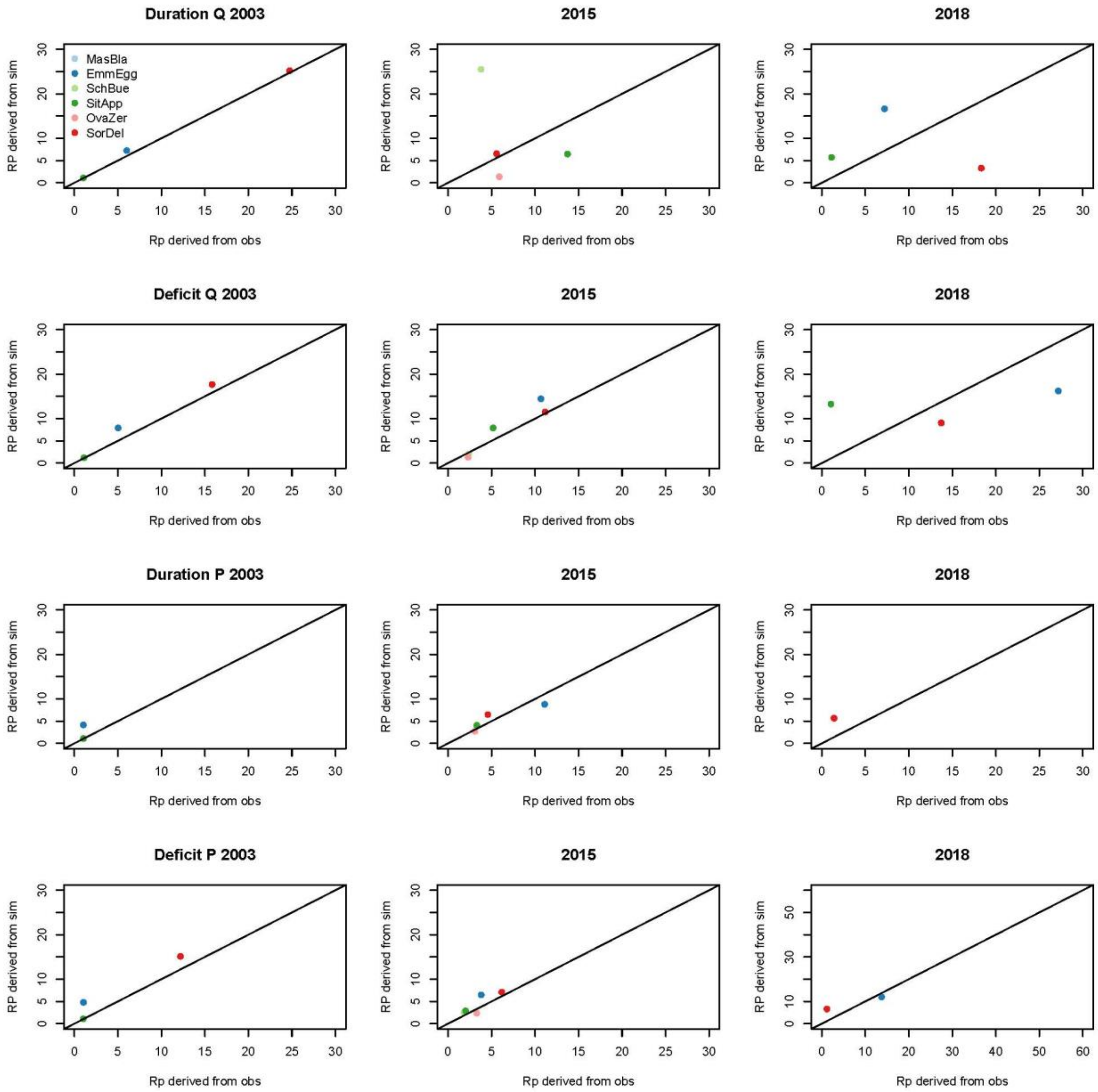

Figure 2: Estimated return periods for the events 2003 (left), 2005 (middle), and 2018 (right) derived from the observations (x-axes) and simulations (y-axes) for the 6 example catchments. Not all events were identified in all catchments.

Modification: p:6, l:10-20 and Table 1.

2. Section 3.2 Event identification. How important is the threshold selection on the final general results? Why the authors select a fixed threshold and why a variable threshold method is not selected for this study (e.g. Van Loon, 2015)? I would expect from the authors to use at least a monthly varying threshold for this type of presented analysis. How different would be the presented results if monthly or daily thresholds are used. Please justify this issue on the revised manuscript.

Reply: *We selected a fixed over a variable threshold because we were interested in the droughts occurring in the main drought season (here summer) as compared to deviations from a seasonal threshold. Choosing a variable threshold would be an option if one was interested in anomalies rather than the drought events happening in the main drought* *season. A fixed threshold is therefore an appropriate choice with respect to the research question. This reason for the threshold choice is provided in the manuscript.*

Furthermore, please discuss the selection of 50% percentile for all the study variables. A preliminary sensitivity analysis could be useful to justify the selection of the selected percentile and the smoothing window of 60 days.

Reply: *We chose a threshold at the 0.5 percentile because a lower threshold (at the 0.6, 0.7, 0.8, or 0.9 percentile) resulted in too few events for a statistical analysis (see Figure 3 in this answer to the reviewers). This explanation is provided in the text. A higher threshold would result in the selection of minor events which is not desirable either. A smoothing window of 60 days was found suitable for ensuring independent events.*

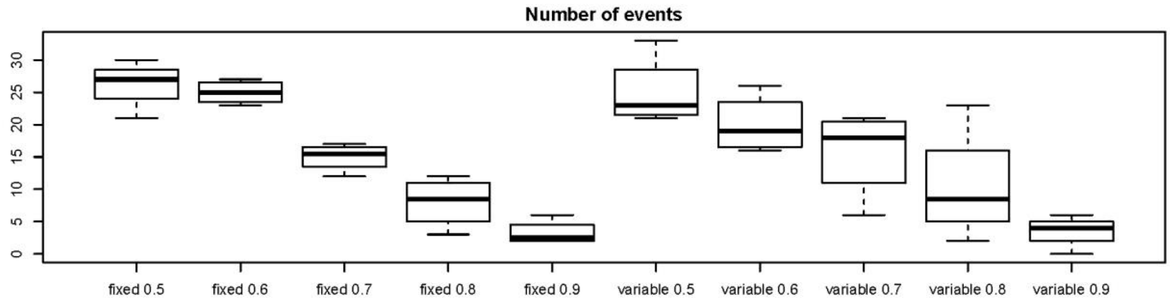

Figure 3: Number of events chosen across the stations (boxplots) for fixed and variable thresholds, where the threshold was chosen as the 0.5, 0.6, 0.7, 0.8, and 0.9 percentile.

For the motivations listed above, the paper in its present form needs revisions in order to evaluate the innovative character of the proposed method. The paper is of general interest for international audience and merits publication in NHESS journal when the major revisions and comments are addressed. Addressing these comments will improve the quality of the paper and help the general reader of the paper.

References used in the answers to the reviewers

Köplin, N., Viviroli, D., Schädler, B., Weingartner, R., 2010. How does climate change affect mesoscale catchments in Switzerland? - A framework for a comprehensive assessment. Adv. Geosci. 27, 111–119. https://doi.org/10.5194/adgeo-27-111-2010