Review of Staudinger et al. – Quantifying sensitivity to droughts – an experimental modeling approach – by A.F. Van Loon

In this paper the authors simulated the effects of two precipitation scenarios on streamflow, soil moisture, groundwater and drought indices for a number of catchments in Switzerland and analysed the sensitivity of the catchments to drought. They also investigated the main controls on this sensitivity and found that size, elevation and slope were important.

General comments:

My view is that this paper discusses an interesting topic, namely catchment sensitivity to drought and how catchment properties alter the drought signal. However, the research presented in this paper has some serious issues, both related to the methodology and results and the way these are presented. Here I will list the most important ones.

- The scenarios of 35 year of progressive drying are not realistic. The authors mention that in their discussion but do not make clear why these scenarios are still useful.
- The authors did 100 model calibrations and averaged the model results. They do this to avoid choosing the "best" parameter set, but introduce other issues. For example when the timing of peaks is different in the model runs, peaks are smoothed in the ensemble mean. Since this is a study on extremes (drought) this might strongly influence the results. I would suggest using the full ensemble of model results predicting the range of possible scenario effects, instead of taking the ensemble mean or choosing the "best" parameter set.
- The entire observation period is used for calibration, no validation of the model results is performed. This is not unusual, but for extrapolation outside the calibrated range (as is the case in this research with the two dry scenarios) validation of the model is needed to estimate its robustness in predicting values outside the calibrated range.
- Many of the results are not surprising and not new. For example the "SoYe scenario did not always result in lower streamflow values compared to the long-term mean, but had rather seasons with pronounced lower flows" and the "timing of the pronounced lower flows appeared to occur simultaneously" (p.7669). That a dry year has lower summer values than the long-term mean is to be expected. And the driest year from the reference run and the driest year from scenario SoYe are actually the same year, only with different initial conditions. It is logical that the low flows occur at the same time in the year and that differences between the scenario and the reference run are small because in Swiss catchments the multi-year memory of the hydrological system is low.
- On the other hand, the results that are surprising were not further investigated. The conclusion that small, high and steep catchments are less sensitive to drought than large, low-lying and flat catchments is counterintuitive, but it is not based on a thorough analysis. I found serious lacks in the statistical analysis of the results: i) only linear relationships have been tested while from the figure non-linear relationships are apparent, ii) the relationship between the tested variables is not investigated (how much is the added effect of the different variables size, elevation and slope, while high catchments are probably also small and steep?), iii) hydrogeology and land use are only investigated with one variable each although there are many possible factors related to geology and land use that might influence the sensitivity of catchments to drought. Some items really need further investigation, for example the negative effect of size and the influence of storage in snow. In the discussion the authors mention that size was also important in the study by Kroll et al. (2004), but it was not discussed whether size was positively or negatively related to low flows. In Kroll et al. (2004) the inclusion of hydrogeological indices was even more important emphasising the importance of hydrogeology in low streamflow prediction. Which parameters can represent the effect of hydrogeology on drought and low flow is still an important question in drought research, so the authors should focus on this analysis to present some novel results.
- The example of 2003 is not representative for other drought events as this event had a very wet winter and spring as initial conditions for a short intense summer drought (see Stoelzle et al. (2014). Therefore, more than one drought event should be investigated in this type of research.

- The paper could benefit from some rephrasing. Sometimes the author's reasoning is hard to follow and at some points I thought the opposite was meant (see some examples below).

Due to these issues I would advise to reject this paper, with possibility for resubmission after major reanalysis and rewriting.

Specific comments:

- P.7661, l.19-20: "additionally different occurrences" > what do you mean?
- P.7662, l.4: Thanks for citing me, but the correct reference is Van Loon & Van Lanen (2012), full reference: Van Loon, A. F., and H. A. J. Van Lanen (2012), A process-based typology of hydrological drought, Hydrol. Earth Syst. Sci., 16(7), 1915–1946, doi:10.5194/hess-16-1915-2012.
- P.7665, l.1: Since this study focusses on low flow and drought, why not use the Nash-Sutcliffe efficiency based on the logarithm of the discharge?
- P.7667, l.14-15: Why would you use the days exceeding Q90 if your study is on low flows?
- P.7667, l.15: "respective"?
- P.7667, l.17: Was the Q90 calculated from the reference simulation and then fixed for the scenario runs? If Q90 was recalculated the calculation of change is not correct.
- P.7668, l.14-16: Shouldn't this be opposite? I see the yellow lines closer to 1, so lower ΔQ in wet years.
- P.7668, l.19-20: This is only true for scenario SoMo. In SoYe the variability in ΔGW is similar to that in ΔQ.
- P.7668, l.22: "driest year of reference simulation" > driest in terms of P or Q? And which year is it? Is it the same year in the four example catchments?
- P.7668, l.22-23: Comparison with the long-term mean is not relevant as the driest year from the reference situation is also below the long-term mean.
- P.7668, l.24-26: Not surprising.
- P.7669, l.6: Not surprising.
- P.7669, l.6-8: I don't think they are so different, probably relative differences are minor. This does not prove differences in sensitivity to drying.
- P.7669, l.9-11: Not surprising.
- P.7669, l.10: Why preceding summer? The initial conditions in this scenario are determined by the preceding years.
- P.7669, l.14: "diminished"?
- P.7669, l.17-18: Again the relative differences are not so large.
- P.7669, l.24: "constant days"?
- P.7669, l.24 & p.7670, l.1: Lower elevation catchments are more vulnerable to drought? This requires more investigation.
- P.7670, l.15-17: Is that so? Or is it just hard to see because of the same scale of the y-axis?
- P.7670, l.19-21: Rephrase!
- P.7670, l.27-28: "IQR" > "Irel"?
- P.7671, l.1: Mention prod.no and %forest.
- P.7671, l.23-27: What do you mean to say here?
- P.7672, l.25-27: ?
- P.7673, l.20: Here you mention snow as a possible, but the effect of snow should be investigated in detail to make this study valuable. And what are the "greater storages"?
- Figure 2: Shouldn't the legend be the other way around (black = SoMo and grey = SoYe)? SoMo was more extreme, wasn't it?
- Figure 4: Put the grey line in front of the other lines or make the lines semi-transparent. Otherwise it is not visible when the blue line is higher than the grey line. Furthermore, the x-axis is not very clear. Indicate the years that were plotted.
- Figure 6 and 8: The correlation between some variables is significant, but the relationship is not always clear and linear. Refer to the txt for explanation of prod.no.

Reference:

Stoelzle, M., Stahl, K., Morhard, A., & Weiler, M. (2014). Streamflow sensitivity to drought scenarios in catchments with different geology. Geophysical Research Letters.