Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-284-RC3, 2018

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "Statistical approaches for assessment of climate change impacts on low flows: temporal aspects" *by* Anne Fangmann and Uwe Haberlandt

M. Hanel (Referee)

hanel@fzp.czu.cz

Received and published: 3 September 2018

The paper deals with statistical modelling of drought indices for current and future climate. The authors fit various linear models to predict runoff drought indices for 28 stations in Lower Saxony. The subset of stations is also used for comparison of the results obtained with statistical models with the results of conceptual rainfall-runoff model. In addition, the best fitting models are then applied to project the indices for past and future climate as simulated by CORDEX RCMs.

Although the paper is well within the scope of the journal, I find it not always clear with respect to used methods but also with respect to the motivation of the study. Therefore

Printer-friendly version

Discussion paper

I suggest to consider the paper for publication only after major revision. The details are given bellow.

General comments:

1. Motivation of the study.

In introduction the authors write: "Even though the application of process-based hydrological models is supposedly the most accurate means for analysis of climatic impacts, model set up and application may be difficult and time consuming, especially for detailed regional analyses and if large numbers of climate change scenarios and climate models are to be considered. Also, data scarcity may pose an issue for model calibration."

While I agree that statistical models are valuable for drought assessment I find the motivation for their use in this study somewhat unconvincing. The authors use many meteorological indices requiring at-site/catchment data as well as indices calculated from runoff time series. I.e. all information needed for calibration of hydrological model is in principle available. Then, in my opinion, the process-based hydrological models should be always preferred, especially for climate change studies, since it might be not clear what happens if the statistical model (calibrated on current conditions) is applied for climate that is warmer e.g. by 4°C. For instance the snow water storage regime may change dramatically etc., introducing non-linearity and making the extrapolation uncertain.

In addition, the statistical framework presented by authors is rather complex. One may ask, whether the application of simple process-based hydrological model is not easier even for large number of catchments, especially when lot of hydrological models are available for easy scripting e.g. in R.

Therefore I see the potential in the presented framework rather in assessment of drought drivers or understanding the key processes responsible for drought generation in current climate (through regression analysis, PCA, finding the most informative variables etc.) than in drought prediction. In this perspective, I think the paper would benefit from focusing more on the information we can get from the statistical models.

2. Study design

Authors calibrate the statistical models using observed data for present climate, find the best model and apply it on bias corrected RCMs for present and future climate. Then the bias for present conditions and the changes between control and future periods are assessed.

I find this rather unfortunate. There are many reasons why the observation-based statistical model cannot work with the RCM simulations. First of all, as stated by authors, the bias from the climate models is removed by monthly linear scaling. It is well known, that the bias in climate models has multi-scale nature and correction at one scale does not warrant correction at different scales (see e.g. Hanel et al., 2017). Therefore the daily meteorological indices are likely biased impacting the results of the whole analysis. Second, there is spatial scale mismatch between the climate model and observed data, as also noted by authors. Especially the indices related to precipitation extremes are then smoothed and also characteristics of temperature related to altitude are likely biased.

More natural approach would be to fit the statistical model considering the climate model derived meteorological indices as input and the observed drought indices as output. This would then include the correction into the statistical model.

3. Clarity

Although the paper is in general well written, some methods are not presented clearly. See specific comments.

Specific and minor comments:

1. p. 3, l.10

"A hydrological model could only be set up for 7 of these stations" - why?

2. p. 6, l.5

"The climate model data is available for a smaller domain than the observed climate data" - actually the RCM data are available for the whole Europe ...

3. p.7, l.9

change "to be e relevant" to "to be a relevant"

4. p.7, l.18

"X" should be changed to Chi

5. p.9, l.25-33 - The procedure for detection of non-stationarity - not very clear to me

Authors use bootstrap to examine the nature of the relation between variables. For a candidate model, they resample with replacement the time series. Then, it is not clear whether they refit the model or just check if the regression fitted for original series is still valid for the resampled series (they say "the regression is significant").

Any case, they further write that "if the regression is significant for all subsamples, the relationship is considered continuously stationary. If, on the other hand, the slope of the regression line is time dependent, random samples taken from different periods will not yield a significant slope coefficient."

I have several concerns regarding the procedure:

- It would be more natural to allow for some fraction of insignificant samples. The requirement that ALL subsamples yield significant regression seems to be quite strong

- If the models were refitted for each subsample then in principle even if the regression is significant, the coefficient can be oposite, or very different... The regression may be also insignificant due to natural variability not as a result of non-stationarity.

- If the models were not refitted for each subsample then the procedure is kind of cross-validation

- In summary, the procedure needs to be described in more detail, consider the concerns above and be preferably based on published literature

6. p.10, l.5-10 - Calibration/validation procedure

"... the calibration consistently precedes the validation period. This set-up is chosen to test the ability of models fitted to a past period of time to predict 'future'". I agree that it is useful to assess the performance of model fitted in calibration period in the validation period. I would, however, be careful with presenting this as a test for the ability of models to predict future. Especially in the climate change context, the differences between the calibration and projection period are much larger, especially for temperature.

7. Model fitting and evaluation section

- it is not clear which combinations of input variables were tested (where they all possible combinations considering the restrictions?). How many models it was?

8. p.10, l.5-10

It is not clear, what you mean by "time series are reversed" and "the inversion is applied to preserve the continuity of the time series". Are the calibration and validation periods just swapped?

9. p.10, l.23

"MLR" not defined

10. Model performance section (p. 10 and further)

In Figure 4, the authors present "the performance of all tested model configurations ...". It seems that in the boxplots also the models that are not appropriate are presented. I find this strange, especially for situations when the same predictors are used e.g in one model assuming independence of residuals and the other model allowing for autocorrelation. Then only one of the two models should be considered (based on the likelihood ratio test), since the other is not appropriate.

In general, it would be more interesting to focus on the information we can gain from such regression exercise.

11. p.13, l.19-21

"Maximum likelihood fitting of models with linear time dependence of the coefficients (validated via likelihood ratio tests) was tried ... but could not yield the expected results for the small calibration period."

- could you be more specific on this, please?

- what you mean by expected results?

12. Figure 6 and related discussion

It seems that the mean error increases with the temporal distance from the calibration period. Does this also apply when you swap the calibration and validation periods? Could this be a sign of trend in temperature? Would the figure look same if you omit models relying on temperature (and ETP)?

13. Prognosis section - see general comments

14. How was the ETP calculated for climate model data?

References:

Hanel, M., Kožín, R., HeřmanovskÃ¡, M. and Roub, R., 2017. An R package for assessment of statistical downscaling methods for hydrological climate change impact studies. Environmental Modelling & Software, 95, pp.22-28.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-284, 2018.