Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-4-RC2, 2016

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Seasonal forecasting of hydrological drought in the Limpopo basin: A comparison of statistical methods." *by* M. Seibert et al.

Anonymous Referee #2

Received and published: 26 April 2016

This manuscript presents three methods for forecasting seasonal hydrological droughts in the Limpopo basin. All three are statistical forecasting methods, including multi linear regression, artificial neural networks, and random forest trees. The results presented are interesting. I think the development of the statistical methods themselves is of interest, though this has also been shown to be successful in several other publications. I do think that the finding of skill, albeit small, out to lead times of up to 12 months is quite remarkable. Also of scientific interest is the development of custom indicators in the Indian and Atlantic oceans, rather than the authors relying solely on the use of the standard set of indicators. Overall the manuscript is well written, though there is some room for improvements. This is particularly so for the conclusions where the writing

style reverts to very short staccato sentences. I would also like to ask the authors to carefully revise how they refer to the various indices, which is somewhat confusing at times. What confused me is that there are standard indices in both Atlantic and Indian Ocean, but also customized indices. The latter are then referred to as Atlantic and Indian Ocean only. I would propose that the authors revise this and always precede the customised indices with the word "Custom" or something like that. That would help clarify somewhat to my mind.

Overall the figures in the manuscript could be made a little larger to enhance visibility/interpretation. There are some very small figures, and at times the figures are not easy to read (e.g Figure 10 could be improved by plotting the thick black line differently).

My main comment on the paper is the influence of the dams within the catchment is not well explored. In some places the authors elude to the presence of dams, and also include details as to their total volume compared to the average annual volume that enters the dam. This shows that for some of the stations the anthropogenic influence is substantial. In many cases there is more storage than there is annual volume, such as is the case for Nauwpoort. And yet this is one of the two stations that are reported to have the highest skill (together with Hartbeeshoek, which has no upstream, dams). This is surprising. This is also linked to one of the findings of the authors that the predictability of the smaller catchments is better than for the larger catchments. This is an interesting conclusion because it is somewhat counterintuitive, because as the authors note, the lower skill in large catchments may be due to the anthropogenic influences. However, these catchments are really very small. This would mean that the skill found cannot be due to the persistence of the catchment initial conditions.

The last overall comment I have is on the selection of the customised indices. The authors note that these were selected over a large area. However, I can imagine that there is a trade-off between the large area and the ability to find significant differences/anomalies. I would expect that as the area gets larger, the detection of anomalies gets smaller. Perhaps the authors could comment on this. The manuscript

is oriented somewhat at hydroinfromatics techniques, and could be considered more suitable for e.g. The Journal of Hydroinformatics. I think to strengthen its appeal to the HESS community it would be good to include some hydrological reasoning why specific parameters are selected in the statistical models.

Specific comments: P1L4: assessed using statistical P1L9: as a proxy P1L15: warning, the models P2L5: which have severe P2L8: which may even P2L9: regarded as being highly affected P2L11: to studies that found P2L10-12: There is some discussion on the climate. I am not sure these comments are entirely relevant to this manuscript. P3L1: Atmospheric circulation processes have ... P3L6: it extends from the ocean P3L13: by the chaotic P3L26: These are particularly P3L33: The skill of the forecasting P3L33: The authors refer to the DJFMAM forecast. That is clear that this spans the wet season. But is this a single value, or is there a forecast for each month. Perhaps I missed it, but it may be good to clarify in the text what a forecast actually contains in terms of parameters and time steps. P5L3: There is some discussion on extracting the catchment areas. Why are these relevant other than to be included in the table describing the catchments. P5L12: event anomalies P6 Table 1: It may be useful to include the year in which the dam was built, or at least the main dam building period in the Limpopo. This can help interpret possible issues of stationarity in the time series. P7L20: with df = N-2 degrees of freedom P7L23: The region outlines P7L23: generously, so as to P8Table3: It is not so clear what the aggregation period of the streamflow indices is. Are these for one month? Or rather is the SSIDJMAM the aggregated streamflow index across the whole wet season. This should all be clarified a bit better. P9L2: linear regression is applied to estimate the values of parameters Bo to Bp. P9L10: until the addition or removal does not lead to an increase in model quality. P10L20: The hidunitj variable is somewhat long and should be avoided. Perhaps introduce something simpler, such as H, and explain it well. P11: I was not so clear how the forecast skill of the ANN is expressed, and if that is commensurate with how it is expressed for the criteria used to establish the MLM parameters. Please ensure that these are well defined, and that that if there are differences explain why the calibrated

models may then be compared. P11L22: The trees are trained P11L26: I am not sure what is meant by the final node size. P11: Overall the description of the Random Forecast Trees is difficult to follow for those not familiar. What are the 500 regression trees? What is the minimum final node size? I think the majority of the readers of HESS will not be familiar with this technique. ANN is more familiar I think. The authors use quiet a lot of jargon such as “bagging” etc. I would be very helpful if they provide a simple explanation of this technique and how a forecast is actually derived. P12L1: model over fitting (I changed this but please check the context) P1211: It is suggested that 2x2 contingency tables can be used only for probabilistic forecasts. I do not think that is correct as these can also be developed using deterministic forecasts. P12L17: has no skill, and is equivalent to a random forecast P13F4: The map is very small, making it difficult to read. Consider increasing its size. P13L16: In the proximity of southern Africa PL13: Chockwé is on the main river and therefore does not represent a sub-basin. P13L20: given the large sample size of 724 observations. What are these observations? Please explain. Are these months, or days? P14L16: Here some of the indicators are discussed before they are introduced. Perhaps add references to the tables here. P16Fig7: What is SRI_NOW? Is this the standardised runoff? I guess so – please clarify. Also clarify what is meant by interactions of selected predictors (grey). P16: It is not so clear what the differences are between ERSST and OISST. Please explain (briefly). These also achieve very different results. P16: In the discussion it is mentioned that the selection of the indicators is unexpectedly low in some cases, which is due to the low correlation. However, this may also be the case for the superiority of Darwin SLP over ENSO. Please try and generalise such findings.

P18L1-3: The results in the figure shows that the importance may vary quite dramatically at the same location during the year. This is not really explained (except that it is very changeable). Is this seasonality? P18L33: At several stations P20L3: also exhibit a strong P20L7: However, our study suggests P21L31: at a lower level P22L7: The discussion on if the errors are small then there is skill seems somewhat trivial. But maybe there is something missing? P23L9: here it is suggested to explicitly consider

the human influence. I cannot agree more. However, I am not sure what is meant by: with the scope on the role of..please clarify P24Fig14: This figure is small and difficult to read. P25: The conclusions can be improved, primarily in writing style. The current style is very staccato and does not flow well. Try and make a bit more of an essay./stroryline. .

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-4, 2016.