Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-467-RC1, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Reliability of meteorological drought indices for predicting soil moisture droughts" *by* D. Halwatura et al.

Anonymous Referee #1

Received and published: 7 October 2016

General comment The paper tests the capability of simple meteorological drought indices to detect drought events, as defined by simulated soil moisture time-series. The topic is of interest for practical applications in drought monitoring, since simulations are often hard to be performed over some areas.

My opinion is that the overall quality of the paper is negatively affected by some basic assumption made by the authors during the analysis, which are not clearly presented and sometime poorly described. Often the reported results seems off, due to errors or unclear explanations. Hence, I suggest to the authors to carefully reread the paper before to proceed with a full evaluation of the paper.

Specific comments

First of all, they compared the 3-month SPI and RDI against a time-series of monthly

minimum pF. If I have understood correctly, this means that this time series is obtained by choosing the minimum pF value (out of roughly 30 values) for each month in the simulation period. If this is the case, I'm really surprised to see the really good correspondence between SPI and pF as shown in Fig. 4 (and 5 as well). Since SPI (as well as RDI) is a standardized variables, its "random" behavior in Fig. 4 is justified, but the same cannot be said for minimum pF which should retain a sort of seasonality depending on the climate of the area. I'm not familiar with the climate of the specific study region, so it is possible that this behavior is due to the peculiar climate of the region, but in general it is not advisable to perform a correlation analysis between a standardized variable (SPI) and a non-standardized one. This should lead to unfocussed readers to assume that this approach is valid also over other regions.

Also, the authors do not clarify if the 75% threshold is computed separately for each month or for the whole dataset. I assume is the first case (based on the data in Fig. B), but this is never clearly stated. Following this topic, in the same figure it seems that the 75% threshold corresponds to an SPI value around 1.2. This means either that: 1) both tails of the distribution are accounted in this computation, but the correct approach would be to consider just one tail since drought event (i.e., extreme dry conditions) are analyzed here, or 2) the fitting of you distribution is poor since the theory suggests that only about 11% of the data should be < -1.2 according to the normal distribution (about 3 values). It is fundamental that this issue is clarified and eventually fixed.

The analysis on extreme values is really misleading, and it also needs to be extended by including other indices. The authors say that FR and FAR are identical in all the cases, but this shouldn't be the case. FR is equal to FAR only if a and c are the same, but this is really unlikely to happen in real cases. For instance, in your example in Fig. B (which I assume is from one of your cases): FR = 5/8=62.5% whereas FAR = 4/7 = 57.1%. Please recheck your calculation of those indices. Also, FR and FAR are not the only indices relevant in this case, e.g., what about the skill of the SPI? Is it better than the climatology or the random case?

Finally, the results of the sensitivity analysis are surprising and need some clarifications. In almost all the case you have FR/FAR values between 30 and 50% higher than in the case of SPI. This means that FR/FAR values for the perturbed simulation are in the order of 65-70% in all cases, included several cases where only a 10% error in 1 parameter is added/subtracted (ie., Bourke 5 cm, Cairns 5 cm, Melbourne 5 cm). I'm really surprised by this result, since in my experience, even for a very sensitive parameter, a 10% change can rarely leads to have 2/3 of the previously detected extremes not detected anymore. It would be useful to have a figure with the reference and perturbed simulations (only the maximum and median ones), as well as the corresponding threshold values, in order to better understand how these changes affect the results. Also, judging from Fig. B it seems that the same 75% threshold is used for both the reference and the perturbed simulation. I assumes that this is not the case, and it is just a coincidence, but I suggest to clarify in the text that the 75% threshold is adapted for each simulation accordingly to the simulated values.

Minor comments

P1, L6. Replace evapotranspiration with potential evapotranspiration.

P1, L6. Rephrase as "used as proxy of severity and duration. . .."

P1, L15. ". . .the frequency with which the simulated. . . below threshold". Actually, you do not want to estimate the frequency, since the frequency is already known as soon as the threshold is defined. Please rephrase.

P2, L7-8. ". . .water is controlling. . .. (e.g., water cycle)". Please rephrase.

P3, L6. I would rephrase as something like "The analysis . . .." Since you are not actually strictly describing a "method".

P3, L 27. How many years were used for the fitting? Which period (the full period?). Please clarify. Also, you should say something about the quality of the fittings (the same is true for RDI).

P4, L21. Appendix A is just a table. Do you really need an appendix for a table? Same for the other appendices.

P4, L22. Remove the parenthesis before "Australian" and move it before "2011".

P5, L1. Please clarify if minimum means minimum among the 30sh daily values in a specific month. Also, please include a standardization of the variable for the successive comparisons with SPI, RDI.

P5, L8-16. Please add at least a skill score.

P5, L17-24. This part on the definition of the threshold is unclear. Please clarify if the threshold is calibrated or not, since you contradict yourself successively in the text. Also, is the threshold computed for each month separately (e.g., 12 thresholds) or for the whole year? The first would be definitely better for the pF.

P5, L20. In Arnold et al. (2014) is reported that there is still seeding also at the wilting point, which does not means that there is no stress. The capability to germinate is clearly reduced compared to optimal water conditions. You should check Cammalleri et al. (2016) "A novel soil moisture based drought severity index (DSI) combining water deficit magnitude and frequency" where a combination of water stress and frequency is used to define drought from simulated soil moisture. Your definition based only on frequency can lead to erroneous estimates over wet areas.

P5, L21. "all values below zero. . .". It seems that this is not what was done since the 75% threshold is not at zero.

P5, L22-24. This statement is true only in theory (see Fig. B) and it also highlights how it does not make sense to test two indices if is known a-priori that they would be the same.

P6, L9. How do you define the "most extreme droughts"?

P6, L10. How was the interval -0.5, 0.5 chosen? Also, please report that 10% steps were adopted.

P6, L9-14. To compare extreme pF values in different sites does not make much sense, since one site can be "naturally" drier than another which is not related to the occurrence of a drought event (e.g., pF in a dry area after a rainy period can be higher that pF during a drought in Sweden). This is the reason why standardized SPI is used.

P6, L28-29. This contradict what stated in the methodology.

P8, L28-30. The FR alone cannot fully explain the performance of SPI. For instance, is this better that randomly guessing drought events? E.g, How skillful is this index?

P9, L20. Is more representative compared to what? (I assume to PDSI considering the reference). I would rephrase as "it represent well. . .".

P10, L9. "rather well. . . then. . .". Please rephrase.

P10, 18-19. This sentence is not clear; also, Fig. 5 seems not relevant to this discussion.

P11, L11-12. This is not necessarily the case. If a significant trend in soil moisture is observed on the site, the use of a longer time-series could negatively affect the analysis (without proper de-trending, etc.).

Fig. 4. Please report the starting/ending dates, as well as a time scale that is multiple of 1 year (e.g., 12-24, 48. . .) to make the figure more readable.

Fig. 5. "The plots represent the highest correlation". What does it means? Please clarify.

Fig. 6. Please re-arrange this figure to make it clearer. E.g., order for site and depth, etc.. Also, the acronym C30, M5, etc. are not defined. Please clarify.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-467, 2016.