

Interactive comment on “Skill of large-scale seasonal drought impact forecasts” by Samuel J. Sutanto et al.

Anonymous Referee #2

Received and published: 7 April 2020

This study develops drought impact forecasting models based on Logistic Regression and Random Forest in Germany. The hydro-meteorological droughts were quantified using the Standardized Precipitation Index (SPI), the Standardized Precipitation Evaporation Index (SPEI) and the Standardized Runoff Index (SRI) for accumulation periods of 1, 3, 6, and 12 months. The skill of forecasted indices and impacts was evaluated using the Brier Skill Score (BSS) and Relative Operating Characteristic (ROC). The manuscript is well written and structured. The results are sound and discussions (incomplete though) support the results clearly. I have, however, some comments/suggestions that I would like the authors to address.

1- Based on McKee et al. (1993), the length of precipitation record for SPI calculation should be ideally a continuous period of at least 30 years. The same criterion is valid for

C1

SPEI. A short record of 28 years was however used to quantify drought hazards in this study. How would this short record length of the data affect the results? Specifically, how would the results be affected by natural climate variability (laying on oscillation high or low period)?

2- The Standardized Precipitation Index (SPI) and the Standardized Precipitation Evaporation Index (SPEI) were applied to quantify meteorological droughts and the Standardized Runoff Index (SRI) for hydrological droughts, all of them for accumulation periods of 1, 3, 6, and 12 months. While there exist many drought indices, the choice of the SPI, SPEI and SRI indices might be justified. An explanation might also be added on why the authors limited the accumulation periods to 12 months and didn't try longer periods (e.g., 24 months)?

3- There are many methods for calculation of potential evapotranspiration, ranging from simple temperature-based method to the standard Penman-Monteith method. It is not mentioned in the paper which method was used for the potential evapotranspiration calculation for the SPEI index and to simulate gridded runoff?

4- Simulated runoff was used as gridded observed runoff was not available. Is the bias of runoff simulations available to be added to the paper?

5- It was found that the SPEI index shows higher skill than the SPI for short accumulation periods. Can it be because of a longer memory of drought hazard calculated from SPEI which is based on water balance (precipitation minus potential evapotranspiration) compared to SPI which is based on only precipitation?

6- It was found that “drought indices with longer accumulation periods perform better than the ones with short accumulation periods”. Isn't it trivial as shorter drought indices with short (long) accumulation periods have more (less) fluctuations/noise and more difficult (easier) to forecast?

Comments on figures:

C2

- 1- Figures 2-4: The black dashed line in the figures indicates the threshold. It was explained in the caption of Figures 3 and 4, while it was defined as "Thes" in the legend of Figure 2. The same format might be used throughout the paper to keep consistency.
- 2- Figures 3 and 4: "ROC" can be added as y-axis label in Figures 3 and 4 and removed from the top.
- 3- Legends of Figures 3 and 4: The legend can be moved to the bottom of the plot with the explanation of each case in the front of it.
- 4- Caption of Figures 3 and 4: "The performance is measured using ROC" should be removed from the caption of Figure 3 as the plot and the other sentences in the caption clearly show it. Same comment goes for the caption of Figure 4. Revise the sentence in the caption of Figure 4 as "The boxes indicate the spread of ROC values for drought impact functions (15 ensemble members) for a) short lead-times (1-3 months), and b) long lead-times (4-7 months)."
- 5- Caption of Figure 5: In the statement "four different impact groups" - the word different is not needed. By stating that they are plural, i.e., four impact groups, the logical syntax of the statement means that they must be different.
- 6- Figure S2: Text on the plots is very small and not readable.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2020-67>, 2020.