

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/nhess-2021-110-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on nhess-2021-110

Anonymous Referee #2

Referee comment on "Applying machine learning for drought prediction in a perfect model framework using data from a large ensemble of climate simulations" by Elizaveta Felsche and Ralf Ludwig, Nat. Hazards Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/nhess-2021-110-RC2>, 2021

This study present a methodology for drought prediction at seasonal scale using machine learning algorithms. The study is treating a highly relevant subject with the usage of a novel methodology based on machine learning for droughts predictions. However, the issue with machine learning and climate is the length of the climate records, which does not allows to build AI models. Therefore, this study proposes a methodology fully based on a down scaled ESM. The study is interesting however, I have major comments listed bellow, in particular, at that stage it is very hard to evaluate properly the manuscript since the data/method is not clear enough:

1) the method (if I understood it correctly) is fully based on model data, therefor it is not really a study about prediction but according to me it is only potential predictability, since this study does not demonstrate any skill in predicting observed past climate in the two regions of interest, but only the ability to forecast the model climate. The paper should be much clearer about this, for example the title and the abstract should use the term "perfect model framework" and/or potential predictability.

2) The method description is very unclear about the prediction aspects. What are the target month analyzed? From which start date? For example, it is really confusing to me to predict SPI1 with one month lead time for different seasons. What do you mean here? Do you mix all together the start date of March (to predict the SPI1 of April), April (to predict the SPI1 of May) and May (to predict the SPI1 of June)? Or do you predict SPI1 integrated over MAM, but in this case, to my understanding it is not SPI1 but SPI3. In any case the methodology should be much clearer about this point, at this stage I cannot evaluate properly the manuscript without this clarification.

3) "The data from the years 1957 - 1999 was used as training data, the years 2000-2005 were used for the testing purpose." Do you mean that the score calculation is performed only for 6 years from 2000 to 2005? This is a far too short period for any skill assessment. Usually, in seasonal prediction the skill is assessed over the whole hindcast period (1957-2005), using cross validation to construct the prediction.

4) The discussion does not mention at all the main limitation of this study according to me: at that stage the authors have demonstrated some ability to predict a model using AI, but we don't know how to use such method for real prediction. Would it be possible to apply your model on observation and then verify its skill? If yes, it should be included in the study and if not this should be clearly mentioned.

Typos:

This study uses the monthly sea level pressure (pr)

The strong influence of ps/psl and NAO shows the influence of the atmospheric pressure