

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC2
<https://doi.org/10.5194/nhess-2021-223-AC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Edward E. Salakpi et al.

Author comment on "Forecasting vegetation condition with a Bayesian auto-regressive distributed lags (BARDL) model" by Edward E. Salakpi et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-223-AC2>, 2022

General comments

The paper presents a study aiming at forecasting vegetation conditions in arid and semi-arid environments, with up to 10 weeks lead-time in order to improve the management of droughts and anticipate their socio-economic impacts. The study uses different satellite product (NDVI, rainfall and soil moisture) to build statistical models for the forecasting of the vegetation index. The study builds on a previous paper by the same team (Barrett et al., Remote Sensing of Environment, 2020, <https://doi.org/10.1016/j.rse.2020.111886>), that used vegetation indices only and extends it using rainfall and soil moisture remote sensing data. Furthermore, the study uses a Bayesian framework for parameter estimation, that allows the determination of uncertainty on the forecast.

The paper is well written and well structured and the results are analyzed comprehensively. The conclusions appear well founded and show that the proposed methodology provides significant improvement as compared to Barrett et al. (2020), in particular in terms of lead time. Some elements in the paper could however be improved: results are compared to a benchmark model that should be better described. Some hypotheses of the work could lead to uncertainties in the results and should be discussed (use of one image in 2016 to identify grassland and shrublands, gap filling of data). Some details about methods for non-specialists of machine-learning techniques and on how figures were built are sometimes missing, precluding the good understanding of their meaning. Provided the authors address these minor comments, the paper will be suitable for publication in Natural Hazard and Earth System Sciences.

Specific comments

1/ P.2 lines 62-63: revise the sentence that is not correct.

Response: Comment accepted, has been fixed

2/ p. 5, line 91: you use sentinel 2 data from year 2016 to identify grasslands and shrublands pixels, but are you sure that this image is representative of the whole 2001-2018 study period? It is likely that land use changes over a 18 year period, so what would be the impact of errors on the grassland and shrubland pixels on the forecasting results?

Response: Comments accepted and has been added as a limitation in the discussion section. An example of an impact on the forecast in the case of a significant change in landcover over the period will be added as well.

3/ p.5 line 100-101: could you elaborate more about the gap filling method: how does it work and how the gap filling could impact the results of the forecasting model? What is the percentage of gap-filled data?

Response: Comment accepted, more details will be added on the gap-filling method. But as seen in the Barrett et al. 2019 paper the gap-filling method did not significantly impact forecasts.

4/ p.7 lines 122-123: could you elaborated a little more on the method used to assess the forecast probability distribution?

Response: Comment accepted and addressed.

5/ p. 8 line 154: incomplete sentence?

Response: Sentence fixed

6/ p.9 line 171-172. In order to assess the validity and robustness of a forecasting model, it is recommended to use different data for model calibration and evaluation. Could you explain more in details how you proceed with the model evaluation and if the data used for the evaluation are independent from the ones used to calibrate it.

Response: It was mentioned in line 172 that a rolling window approach was used. But to make this clearer, some more details or a walkthrough on how the split for training and test (held out) datasets was used for model training and evaluation will be added.

7/ p.9 lines 183: a figure explaining the computation of MPIW and PICP could be useful.

Response: A well-detailed figure of the MPIW and PICP can be found in the cited paper. However, a simple version will be added to this paper as suggested.

8/ p.9 lines 192-194: Could the authors provide more details about the AR benchmark model. I understood that it was built using only the vegetation index, but was a Bayesian framework also used for parameter estimation? If not, could the performance of the AR model be improved is a Bayesian framework was used for parameter estimation?

Response: A description of the AR method will be added to the paper. AR method used in the paper was not parameterized with the Bayesian approach, however, if a Bayesian method was used it might not significantly improve parameter estimation but will add probabilistic interpretation to the AR results. The improvement in model performance was mainly attributed to additional variables. Using the Bayesian approach was to give a straightforward way to assess forecast uncertainty.

9/ p.10 lines 203-205: for non-specialists (possibly in appendix or supplementary materials) explain how the Reliability diagram and Sharpness are built.

Response: Comment accepted and details have been added to the paper

10/ p.11 Fig. 3: explain in the figure caption how the figure is built: what are the ellipses on the figures?

Response: Comment accepted, more details on the ellipses, which are the joint distribution bins of the scatter plot between the forecasted and observed VCI3M values, will be added to the paper.

11/ p.15 Fig 6: what is AUC in the caption and on the figures? The lines for the two models have the same colors and types. It is not easy to understand which curve is related to what? Could you also explain how this curve was built?

Response: Curves in the figure are not the same types, AR curves are dotted and the BARDL curves are solid (explained in the caption). The line colours indicate the lead times (i.e. 6, 8, 10 & 12 weeks ahead) which are the same for both models, thus the two methods being compared in the plot cannot have different colours. More details on the AUC will be added.

12/ p.16 Fig. 7: same remark as for Fig. 6: explain how the figures are built and what is the information they carry out.

Response: Comment accepted, details will be added

12/ p.16 Lines 253-256: I do not understand these explanations.

Response: Comment accepted and fixed

13/ p. 17 lines 274-276: the result mentioned here was not shown before. It can be seen in the appendix but this should be mentioned.

Response: Comment accepted and well noted, it will be addressed

14/ p.18 line 280: figure ?? – please modify.

Response: Comment accepted and fixed

15/ p.17-18: a discussion of the limitations and impact of the gap filling and choice of the 2016 Sentinel 2 image to identify grassland and shrubland should be added.

Response: Comment accepted and limitation as has been acknowledged and added.

16/ Fig B1: Explain what is shown in these figures. They are not understandable with the current caption.

Response: Comment well noted. The caption has been updated