Comment on nhess-2021-223

Anonymous Referee #2

Referee 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-RC2, 2022

This paper proposes a new model for forecasting the vegetation condition index (VCI) based on a Bayesian autoregressive distributed lag (BARDL) model. The new model can provide the probability distribution of VCI instead of a deterministic value. In a forecasting framework, it is clear that the BARDL model can improve the current methods, as supplying a probability distribution is crucial for decision making. The BARDL model is applied to a set of counties in Kenya with arid and semi-arid conditions. VCI is forecasted from the available information about precipitation and soil moisture content, considering the current information about drought conditions. The new BARDL model is compared with the results obtained by using a deterministic AR model. The comparison is based on a set of measures that quantify both accuracy and precision. The paper offers a new method that can overcome some limitations of the current models to forecast droughts. However, the paper needs to address the comments included below before accepting it for publication.

General comments

- The paper uses the vegetation condition index (VCI) to forecast droughts in Kenya. However, other indices are available like SPI, SPEI, PDSI, multi-variate standardised dry index (MSDI), the temperature condition index (TCI), the vegetation temperature condition index (VTCI), and the temperature vegetation dryness index (TVDI), among others. A discussion could be included in the paper to support the selection of VCI in the paper.
- The Introduction Section focuses on three existing techniques to forecast VCI: Auto-

Regression, Gaussian Processes and Artificial Neural Networks. A longer revision of the techniques used in last years to develop EWS for droughts could be included in this section, as well as other papers that develop similar tools. For example, stochastic algorithms based on different types of Markov Chains, autoregressive moving-average (ARMA), autoregressive integrated moving average (ARIMA) techniques, support vector machines, Kalman filters, multiple regression tree techniques, among others, have been used in last years to forecast droughts.

- While the BARDL algorithm supplies a probability distribution, the AR model supplies a deterministic value. Therefore, the comparison between the two models is not straightforward. In the paper, a confidence interval for the AR model is estimated from RMSE and z-score. However, this is a simplified way to estimate the prediction uncertainty, supplying a constant confidence interval regardless the magnitude of both VCI and the explanatory variables. This step is very important to compare BARDL results with AR results in a proper way. In addition, the methodology to compare both models should be clarified in the paper, as it is not clear how most of measures used to quantify accuracy and precision have been applied to the probabilistic forecast supplied by BARDL.
- The Discussion Section should be rewritten, as in its current form it is mostly a mixture of conclusions with some additional results considering seasonality.
- The Conclusions Section could be extended to summarise the main findings of the study.

Specific comments:

- Abstract: Some sentences could be included in the abstract about the case study used in the paper.
- 14: The acronym AR has not been introduced in the paper at this point yet.
- 30: The acronym USAID is not introduced in the paper and could be explained at this point.
- 46: The ARDA model has been applied to assess droughts previously, such as Zhu et al. (2018). References to previous studies in which the ARDA technique is applied to droughts should be included in the paper.
- 51: The paper proposes the use of a Bayesian framework in the ARDA model to incorporate the prior knowledge about model parameters in the analysis, obtaining a probability distribution for VCI results. Bayesian networks have been also applied to develop a long-term drought forecast (Shin et al., 2019), supplying probabilistic results that can assess forecast uncertainties. A discussion could be included in the paper, stating the benefits of a BARDL model compared to Bayesian networks.
- Section 2.1: Some information about the number of counties considered in the study could be included in this section, as well as the number of counties that are arid and semi-arid. In addition, some information about the area in km2 that is considered in the study could be useful for the reader.
- 70: 'estimates' should be changed to 'estimate.
- 98-99: The description of NVIi and NDVIi variables should be included in this paragraph too.
- 103-104: 'long term' should be changed to 'long-term'.
- 111: The acronym AR has been introduced in the paper above.

- 118: A discussion could be included about the selection of the OLS method for estimating parameters of ARDL. Some other methods are also available.
- 131 – Eq. 3: The variable subscripts should be revised in Eq. 3. D_{t-q} seems to be the drought indicator in a constant time step $t-q$, which seems to be constant in the first summation regardless the value of i. Similarly, P_{t-p} and S_{t-p} seem to be constant values in the summations. In addition, the regression coefficients are also constant values in the summation, though they could change in terms of i. A discussion should be included about the use of constant values in summations.
- 137 – Eq. 4: How does X_{t-i} represent several variables? How can i vary from 0 to i?
- 143-146: The variable theta should be explained to readers in this paragraph.
- 145-146: The term $P(X_t)$ is ignored because it is difficult to compute. This is not a proper statement for ignoring a variable in a research paper.
- 152-154: An analysis should be done to fix the distribution function that best characterises the regression parameters. Why mu is set to 0 and sigma to 0.5?
- 153-154: Something is missing in this sentence.
- 164: This is not the standard form of AIC.
- 161-163: Some figures could be included in the paper to show how a time lag of 6 weeks obtains the best AIC and R2 results.
- 168: What is i? What is y hat?
- 176: The R2 measure of Eq. 9 is not a good measure to quantify accuracy of forecasts.
- 188-189: What is m?
- 196: 'inputs' should be changed to 'input'.
- 213: How r, R2 and RMSE are calculated for the BARDL model? The BARDL model supplies a probability distribution, but observations are deterministic.
- 214: R2 is not a good measure of forecast accuracy. RMSE is more adequate than R2. Therefore, the gain in performance metrics could be assessed with RMSE. However, the BARDL model supplies a probability distribution of VCI. How do you obtain a RMSE value from the comparison between probability distributions and deterministic values of observations?
- Figure 3: What do the coloured lines mean?
- 222-224: The R2 values do not correspond with the values shown in Table 2.
- Table 2: This table could be summarised in a figure.
- 229-231: The table in Appendix A could be summarised in a figure and included in the main text of the paper, in order to analyse the comparison between the two models. The results included in Table 1 show that PICP values are smaller for AR than for BARDL, meaning that a greater number of observations are out of the confidence intervals for BARDL. This result should be discussed in the paper. In addition, most of PCIP values for the BARDL model are smaller than 94-96 %, in contrast to the statement of line 229.
- Figure 5: Please use the same y-axis scale in each row to compare the AR and BARDL results. The dashed line of the left column differs from the dashed line of the right column, though observations do not change. The green line represents the forecast. What is such a forecast for the BARDL model given that it supplies a probability distribution?
- 235: A drought is forecasted when VCI3M values are smaller than 35. This is straightforward for the AR model, as it is deterministic. However, how do you apply this criterion to the BARDA outputs considering probability distributions?
- 251-253: The BARDL lines lie above the main diagonal of the reliability diagram. This means that the probabilities supplied by the BARDL model tend to underestimate droughts. A comment about this point should be included in the paper.
- 253-256: The sharpness diagrams are mostly flat for 10 and 12 weeks. The low values close to 1 means that the BARDL model is not able to forecast droughts. Therefore, the BARDL model is useful to forecast droughts with 6 weeks ahead but it is not for 10 and 12 weeks. A comment about this point should be included in the paper.
- Figure 7: The sharpness diagram should plot percentages in the y axis.
- 266-276: These two paragraphs could be moved to the Conclusions Section.

- 273-274: The Authors state that the BARDL model gains 2 weeks based on the results of R2. However, more measures should be taken into account to conclude such a statement.
- 275-276: This statement is not clear from the results included in Section 4.
- 280: The number of the figure is missing.
- 284-295: This paragraph with figures of Appendixes C to F could be extended to form a new section 4.6 devoted to the seasonality analysis.

References

Zhu N., Xu J., Li W., Li K., Zhou C. (2018) A Comprehensive Approach to Assess the Hydrological Drought of Inland River Basin in Northwest China. Atmosphere, 9(10): 370. https://doi.org/10.3390/atmos9100370

Shin, JY, Kwon, H-H, Lee, J-H, Kim, T-W. Probabilistic long-term hydrological drought forecast using Bayesian networks and drought propagation. Meteorol Appl. 2020; 27:e1827. https://doi.org/10.1002/met.1827