Title: Multivariate regression trees as an 'explainable machine learning' approach to exploring relationships between hydroclimatic characteristics and agricultural and hydrological drought severity. Case of study Cesar River basin.

Authors: Paez-Trujilo et al.

Recommendation: minor revision

Assessment

I appreciate the authors' careful consideration of my main concerns in their revised manuscript. The changes made have significantly improved the clarity and overall quality of the paper. The chapter 2 is now well structured compared to previous version. For the general comments, I disagree with authors response on SSMI although this is irrelevant with the manuscript goal. I also notice that the revised manuscript was not written carefully. I identified numerous mismatches between figures and their corresponding explanation text and figure references (chapter 3.4). I kindly request the authors to carefully review the revised manuscript. I look forward to seeing the final version of the paper.

General Comments

1. Thanks for addressing my comments. However, I would like to raise a concern regarding the explanation provided on the use of the Standardized Soil Moisture Index (SSMI or SSI and ESMI in some papers) that is not correct. The authors stated that 1) SSMI is an agricultural drought index derived from daily satellite data, 2) the index is developed for short-term drought monitoring, and 3) there is no previous assessment of the index performance using simulated SM as the input parameter. All these three arguments are misleading. First, the SSMI is a standardized drought index like SPI (for precipitation), SSI (for streamflow), SGI (for groundwater). It uses monthly data instead of daily data (e.g., Ndehedehe et al., 2016; Carrão et al., 2016; Das et al., 2022). These publications used monthly soil moisture data to derive SSMI. Second, SSMI has been employed not only for drought monitoring but also for drought forecasting in some studies (e.g., AghaKouchak, 2014; Xu et al., 2018). Last, some of the aforementioned publications have indeed utilized models to simulate soil moisture variable used in drought identification (SSMI). The second paragraph is accepted.

2. Thanks for your explanation. The authors may consider to move the PCA analysis in the appendix or supplementary material instead of delete it. PCA analysis is still useful to indicate the variance.
3. OK

Line by line comments

L refers to line and P refers to page.

I acknowledge the responses from the authors for line-by-line comments. However, the new colors used in the figures for the revised manuscript can be improved. My remarks and suggestions are underlined below.

P5L117: Figure 2. Thanks for changing the colors in the legend. I suggest to further improve the distinction between different land cover types. The colors for pasture, crops, and shrubs have bluish colors. The authors may use color dark green for forest, light gray for pasture, color light green for cropland, and brow for shrubs as commonly used in many land use map, or just see some examples from the published land use map.

P13L292: Figure 3. Same suggestion. The use of gray and blue colors in the graph is not contrast. The authors may use blue color for observed and red color for simulated. The vertical line to divide calibration and validation periods can be black.

New line by line comments

P1L1: Suggestion for title: "Multivariate regression trees as an 'explainable machine learning' approach to explore relationships between hydroclimatic characteristics and agricultural and hydrological drought severity: Case of study Cesar River basin"

P1L17: The authors may write "(SWAT)" here.

P1L18: Suggested text revision: ….the drought indices namely Soil Moisture….

P2L59-60: Rephrase this sentence. It is unclear.

P3L83: Mentioned -> The aforementioned research

P4L98: Between indices and Soil Moisture Deficit Index, the author my write either "which are" or "i.e.,"

P10L231: Instead of "(the drought indices give categories)" -> "(moderate, severe, and extreme)"

P10L234: What do the authors mean with four technique attributes are relevant to this study?

P10L238: The authors may remove "The drought indicators give these three categories to represent the drought severity". It is redundant.

P12L274: What is SS?

P13L310: Please give a low flow definition here. How do the authors identify low flow? Is it using a threshold method?

P15L324: Figure 4. It is annual average right? Also, readers need an explanation about soil type A, B, C, and D. What are those?

P15L328: What drought category is represented in Table 6? Is it severe drought, extreme, or moderate drought?

Starting from here, please read carefully and do comprehensive check.

P17L357: Here the authors say: potential evapotranspiration (1,679 mm). However, I cannot see this number in Figure 7.

P18L369: I think it is lower and not above.

P18L372-373: It is not Figure 8b and also please check your statement about "highest median of months in the severe drought category"

P18L375,376: It is not above but lower.

P18377: Here the authors stated that moderate drought category was above 20 months and severe category was above 10 motnhs -> but this is not for Figure 8h.

P20L404: Check the number 1362 mm, cannot see this number.

P20L421: I think it is higher and not lower.

P20L426: Check if it is Figure 9f?

P20L428: Check if it is 37?

P21L442: Why figure 8e? 8d is higher

P21L444: Clusters h, i, and j are seen in Figure 8i, j, and k?

P22L475: It think it is not groups I and j but h and i.

P23L479: This sentence is confusing. Subbasins grouped at leaf i showed in Figure 10b and c?

P25L544: were located

P26L591: software -> model

Reference

Carrão et al.: An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data, https://doi.org/10.1016/j.jag.2015.06.011, 2016.

Ndehedehe et al.: On the potential of multiple climate variables in assessing the spatio-temporal characteristics of hydrological droughts over the Volta Basin, https://doi.org/10.1016/j.scitotenv.2016.03.004, 2016.

Das et al.: A non-stationary based approach to understand the propagation of meteorological to agricultural droughts, https://doi.org/10.1007/s11269-022-03297-9, 2022.

AghaKoucak, A.: A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought, doi:10.5194/hess-18-2485-2014, 2014.

Xu et al.: Standardized soil moisture index for drought monitoring based on soil moisture active passive observations and 36 years of north American land data assimilation system data: A case study in the Southeast United States, https://doi.org/10.3390/rs10020301, 2018.