

COMMENTS 2

Authors: We are very grateful for the comments and suggestions to improve the study.

- In line 110, there is extra text. "Anexo 1- poner algo mas?" . You need to eliminate it

Authors: Done.

- In Figure 2, it is mentioned that it is used for the model in Equation 12 to estimate parameters. I believe it should say Equation 3.

Authors: Done.

- What is the motivation for reducing the value of w to 0-1 by replacing values of w above 1? If later a index (WCI that will range between 0 and 1) is used?

Authors: Although W and WCI range between 0-1, the meaning differs. WCI uses the expression:

$$WCI = \frac{W - W_{min}}{W_{max} - W_{min}}$$

where W is each value of the time series and W_{min} and W_{max} are, respectively, their multiyear minimum and maximum for every 10 days. The maximum and minimum values of the denominator reflect the best and worst conditions of surface soil moisture, respectively, and the difference between them somewhat reflects the condition of the local soil moisture.

- Where does Equation 3 come from (there is no reference, although it seems that it could be found in Sadaghi 2017 from Figure 2)? I think either a reference should be provided, or the foundation and interpretation of this index should be explained, or both.

Authors: We have improved the explanation of this index in the manuscript.

- In my opinion, the explanation of Equation 6 is unnecessary; I believe conditional probability should be well-known to everyone.

Authors: Done.

- In line 173, it is indicated how the phases are generated. To impose the limits of these phases, it seems like the median is used, although it is not explicitly stated. In my view, representing it with a Box-Cox diagram would allow a discussion of these points since the third quartile would suggest that phase 2 begins some weeks later, especially in Los Velez. Wouldn't it be clearer to exclusively represent a line connecting the medians of each ten-day period for greater clarity?

Authors: We added that the VCI median was used to delimit the phases and modified the figure 3 to include lines connecting the VCI and WCI medians.

- Figure 4 is not clear, especially regarding rainfall. It is difficult to distinguish those "abundant rains" in phase 3 (at least when comparing them with phase 2 where the extreme points are higher). There is also a contradiction stated in lines 188-189 regarding the explanation of Figure 4, as it indicates that there is more rain in phase 2.

Authors: Now it says: “Phases 3 and 4 in Bajo Aragón are very similar to Los Vélez with the same duration and time limits. The only difference is that the VCI maximum in Los Vélez is higher than in Bajo Aragón, probably, the vegetation of the area had a better condition to grow at Los Vélez in phase 2 (see Fig. 6A)”.

- I suggest using more distinct colors to better distinguish the series in Figure 5 (for example, red and blue).

Authors: Done.

- In line 192, it is mentioned that there is a smoother profile of the ZVCI series; perhaps a cause explaining this should be suggested.

Authors: We have included the following explanation:

“This behaviour is revealing that vegetation response to environmental changes is slower than soil moisture response.”

- Although providing probabilities in percentages may not be the most appropriate, it is understandable given the calculation method. However, I am more in favor of indicating the value between 0 and 1 when discussing probabilities.

Authors: In the methodology section, we have added an explanation about the calculation of the estimations of probabilities using percentual relative frequencies.

- Mentioning expected probabilities on standardized data could be considered. For example, if the indices followed normal distributions (which may not be the case), for ZWCI and ZVCI, the probabilities would be $P(\text{Index} < -0.5) = 0.309$, $P(\text{Index} < -0.7) = 0.242$, and $P(\text{Index} < -0.1) = 0.159$.

Authors: We have added the following paragraph:

” In addition to this, probabilities of Z_{WCi} and Z_{VCI} for the different thresholds can be compared with expected probabilities on standardized data. If the indices, Z_{WCi} and Z_{VCI} , followed normal distributions, the probabilities would be $P(\text{Index} < -0.5) = 0.309$, $P(\text{Index} < -0.7) = 0.242$, and $P(\text{Index} < -1.0) = 0.159$, which is not the case for the majority for the 10-day periods.

- The comments about the probability of an index being below the threshold of -0.7 being lower than being below -0.5 seem unnecessary. This is obvious: $P(Z < -0.5) = P(Z < -0.7) + P(-0.7 \leq Z < -0.5)$ (the same for $Z < -1$). In any case, the greater the difference, the higher the probability that the index is between $[-0.7, -0.5]$.

Authors: Thank you for your comment, you are right. However, thinking of potential readers, not specialists in statistics, we prefer to keep it.

- In line 215, why is there now a condition for the threshold of -0.3? I don't understand if there is justification or at least it is not explained.

Authors: The conditional probability we are obtaining is defined as:

$$- P(Z_{VCI} < th \text{ at the period } i \mid Z_{WCi} < -0.3 \text{ at the same period } i):$$

The probability of an anomaly occurring in VCI at the period i (10-day) under the condition of an anomaly occurring in WCI ($Z_{WCI} < -0.3$) at the same period. " th " are the three different thresholds: -0.5, -0.7 and -1.0.

The reason to use the value -0.3 in Z_{WCI} is to consider a great range of anomalies in the condition.

- In Figures 8, 9, 10, 11, $P(VCI < -0.7)$ is stated. Obviously, this probability is always 0 because the VCI value is in the interval (0,1). It should be $P(ZVCI < -0.7)$.

Authors: Thank you very much. You are right, there is a mistake in the legend of the graphs.

- In the paragraph between lines 224 and 230, there is no reference to the possibility that the correlation could be negative in the summer months (from late May to September). Additionally, the term "significantly" is used, but based on what criterion or test is it significant?

Authors: We have eliminated the term "significantly", the paragraph now says:

"Los Vélez shows high values in the lag-4 conditional probability with minor anomalies (-0.5 and -0.7), often reaching from 50% to 65% of probability from November to January (compared to an average of 20-30% of base probability). These probabilities decrease when threshold -1.0 is used, reaching a maximum of 50% in the middle of January."

- In lines 238 and 239, it is said, "as with Los Velez, the lag-4 conditional probabilities for the threshold -1.0 remain above the base probability." It seems incorrect to me; in Los Velez, these probabilities for the summer months were below.

Authors: Thank you very much. You are right, we have eliminated this paragraph.

- In line 241, it is said that the probability with a 4-period delay is higher than the base probability, but this also occurs for the probability without a delay.

Authors: You are right, there are some periods where lag-0 conditional probability is higher than lag-4 conditional probability, but in the discussion section we are focusing on the capability of predicting VCI anomalies, so we focused on lag-4 conditional probability.

- I wonder why, with two time series, a transfer function model is not carried out, identifying cross-correlations to determine the delay that best explains one series in terms of the other. It may have been done, but it is not indicated in the article. Why have you chosen 4 lag?

Authors: Previously, we did the cross-correlation to select optimal lag, but it is not included in the manuscript. If you consider it necessary, we can include it in an annex.