Hydrology and Earth System Sciences

Open Access

Discussions

EGU

Interactive comment on "The predictability of reported drought events and impacts in the Ebro Basin using six different remote sensing data sets" *by* Clara Linés et al.

Clara Linés et al.

c.lines@unesco-ihe.org

Received and published: 4 May 2017

We thank the reviewer for providing detailed comments that helped us to further reflect on the analysis presented in the manuscript. Our replies to those comments are included below, together with indications on how we propose to improve the manuscript based on the suggestions of the reviewer.

Note: The numbering of the comments has been modified to fix some existing repetitions in the original numbering.

....

This study analyses drought indicators' relation and thus usefulness to predict drought impacts for a small case study region in Spain. I find the study a worthy addition to the drought literature, once its contribution has been worked better. In the current manuscript the reader 'gets lost' in the many correlations and individual results a bit and I think, some focus and highlight is needed to better appreciate the work and results. The manuscript also requires clarifications on a number of methodological details and better justification of some parts of the approach in order to assess and value the results.

Major comments:

1) A particular concern I have is related to the use of a very short time period for the 'benchmarking' by correlation analysis. In this study, a time series from 2001 to 2012 is used with 8 out of these 12 years representing drought events (Section 3.1., beginning) - hence 2/3 of the time. The common understanding is, however, that drought is defined as a rare extreme climatic/hydrologic event. This has a few implications for the chosen method and conclusions from the study:

Our aim in the analysis presented in the manuscript was to assess the usefulness of medium resolution global remote sensing products for drought management at basin scale. The length of the time series used in the analysis is therefore limited by the availability of those remote sensing products. We acknowledge that 12 years is a short period and that this limits the variability represented in the series. In fact our series includes three drought events, one of them having impacts over four hydrological

years. The short length of the data series available was one of the reasons to base the definition of drought we use to build the reference not on a frequency analysis, in which drought is defined as an extreme event with respect to the historical series, but on the occurrence of drought impacts. The other reason being that what managers need is to identify the conditions that may lead to drought impacts in order to take mitigation actions. For these two reasons we defined drought in the analysis as conditions of meteorological origin that may lead to impacts in sectors depending on water and we used impact data as benchmark information. This will be clarified in the manuscript in section 2.2.3 (Benchmarking datasets).

a.- To the international reader who doesn't know the Spanish climate history, it is not convincing whether simply 'dryness' or real 'drought' was analysed if 2/3 of the time were 'drought events' – how was this distinguished in the textual search? Does the Spanish language distinguish between the two? Some languages do, others don't.

The Spanish language does have different words for dryness (sequedad) and drought (sequía). The term 'sequía' was used for the textual search. As we commented in the discussion, the possible misuse of the word drought/sequía in the press was one of our concerns when reviewing the news. This was the reason to differentiate the records that are just mentions of drought, from others that are reporting the acknowledgement of drought from an official source. The classified records in figure 2 of the manuscript (upper part) show that only one of the records (the mention of drought recorded in 2003) is not backed-up by the acknowledgement of drought from official sources during the same period, and therefore may be considered a case of misuse of the word. We will include a specific comment in the discussion about the accuracy of the impact data (p.18, line 17).

The drought events identified by the textual search for a sector of the Ebro basin correspond with events observed at a larger scale. For example, Spinoni et al. (2015), use an indicator that combines three precipitation and potential evaporation-based indices to identify the drought events occurred in different regions of Europe during the period 1950-2012. Following that approach they identify three drought events for the Iberian Peninsula for the period 2000-2012 that match the ones obtained by the textual search, with the difference that the event starting in the hydrological year 2004-2005 has a shorter duration. This is caused by the different spatial scale of the analysis. While most of the basins in Spain received normal precipitation during the hydrological year 2006-2007, in the Ebro basin, and especially in the inner part of Cataluña, it was still low during that year (MMA, 2007). The reference to the work of Spinoni, corroborating our results in identifying three drought events during this short period will be also included in the discussion about the reliability of the impact data (after the second paragraph in section 4.1).

b.- To define drought based on monitored hydro-meteorological or remotely sensed anomalies, long time-series are necessary to obtain the full range of situations and hence define the average and moments of distribution of the variable to be used to index drought as an extreme. How does that influence the results? The time series and temporal resolution for each index need to be given (inconsistent in current methods section), and where applicable the reference time series for standardization/normalization.

Our drought reference is based on the occurrence of impacts, rather than on the identification of extremes in the time series. This is explained in more detail above, in the first paragraph of the reply to this comment, and will be also clarified in the

manuscript in section 2.2.3 (Benchmarking datasets).

The length of the period of analysis doesn't have an influence in the identification of drought events based on impact records. However, having a longer series, and therefore potentially a larger number of drought events, would provide more robust results in the correlation analysis. Ideally the results should be updated as the period of record of remote sensing data grows. The implications of the length of the remote sensing series will be added to the discussion in section 4.2.

We have verified the information provided in the manuscript regarding the temporal resolution of the datasets and have realised that the reference to the monthly GPP and PsnNet product used in the analysis, downloaded from the Numerical Terradynamic Group (NTSG), is missing. This will be added to the text (p.6, line 16). The temporal resolution of all the selected products is shown in Table 1 in the column 'Original time interval'. No additional data has been used as a reference for standardization/normalization.

c.- Correlation depends on variability, but if two thirds of the time period are mostly dry, I would expect that this will have a considerable effect on the results of a statistic that relies on variability in the data. Some analysis regarding the sensitivity of the index series (and range of variability) on the results is therefore necessary to make an assessment of the uncertainties.

The correlation analysis is performed with a monthly time step. According to the news review, during the period of analysis we have drought conditions occurring 63 months out of 136 (46%) and impacts 58 months out of 136 (42%). These records

Printer-friendly version

Discussion paper

correspond to three drought events. One of them is a multiyear event of extreme severity (2004-2008). In fact, the hydrological year 2004-2005 was characterised as one of the most intense droughts of the record in the Iberian Peninsula (García-Herrera et al., 2007). The other two events affect a single hydrological year at a smaller scale. However, despite the shortness of the period (12 years), it encompasses a wide range of different conditions, including also one of the wettest hydrological years of the country's record, the hydrological year 2003-2004 (MMA, 2005). The variability of the period will be mentioned in the description of the study area.

We acknowledge that a sensitivity test would contribute to reinforce the validity of the results. However, a longer series than the one available would be necessary to ensure an adequate estimation of the sensitivity. A detailed discussion on how the significance of the results and the possible issues of the remote sensing series were considered in the analysis is provided in the reply to comment 9.

2) The objectives and used methods need to be better harmonized, in particular the relation of the use of statistical correlation analysis and the aimed for 'predictability' of impacts needs to be developed more clearly. An assessment of predictability of impacts would require a predictive application (some validation experiment). For hazards, it is also common to consider false alarm rates for an assessment of predictablity. Overall, the use of statistics and their interpretation in this study requires more clarification and precision (see specific comments below).

The analysis presented in the manuscript can be considered a preliminary study on drought prediction. Indeed, we agree that the development of an operational detection method or tool would require further analysis and validation. At this initial stage,

however, our aim is to identify the datasets that can be useful for operational drought detection at the basin scale. Drought detection in this case is closely related to the predictability of impacts, as the conditions that need to be detected are those that may lead to impacts. However, these impacts do not necessarily occur immediately; their occurrence can be delayed as the effects of drought propagate through the different components of the hydrological cycle. To identify the remote sensing parameters that represent conditions that anticipate the occurrence of drought impacts, and therefore have potential to support the prediction of drought, we explore the correlation between the remote sensing data and the drought impacts at different time lags. While using correlation in this way may say less about the long term correlation of two time series, it does provide insight in the relationship between correlation and lag. We will include a clarification of the aim of the analysis in section 2.3 (description of the correlation analysis).

3) The discussion section is rather vague in its comparison with other studies. In particular, there are many other studies that have correlated agriculture yields with drought indices in many countries (e.g. see table of the review by Bachmair et al., 2016 in WIRES Water). How do the findings for best-correlated indices and time scales compare to other studies? If this case study wants to contribute to the international literature these need to be better compared in the discussion. The niche of this study within the wider range of studies needs to be worked out more specifically to appreciate this small scale case study's contribution to the field.

We agree that the correlation of remote sensing data, especially SPI and NDVI, to agriculture yield data has been widely researched and applied. Rather than aiming to add something new to that field, the purpose of including that part was to provide

a comparison of the results obtained in the correlation to text based impact data and results obtained with the most commonly used type of impact data, crop yields, and discuss the advantages and limitations of one with respect to the other. An explanation of this purpose and references to the wide use of correlation between crop yields and drought indices will be added to section 2.3 (description of the correlation analysis). Additional references and comparison with other studies will be also added as mentioned in the replies to comment 1a and to the reference suggestion at the end of this review.

Specific comments

Methods section:

4) The twofold use and steps of impact report analyses, (1)construct a narrative of the events, (2) use as binary indicator of impact in correlation analysis, should also be introduced as such in the methods section. The narrative in the first part of the results otherwise comes really unexpected, whereas the reader expects only correlation results.

We agree and will mention the use of the impact records to reconstruct the onset and evolution of drought conditions during the period of analysis in section 2.2.3 (Benchmark datasets).

5) P1 Last paragraph: another reason that RS data needs benchmarking are the short time series - compared to precipitation and hydrometric records

a definition of drought as statistical extreme is not possible. This requires elaboration on assumptions made and limitations for the analysis.

As mentioned in the reply to comment 1b, the analysis was not based on a definition of drought as a statistical extreme, but as the occurrence of certain conditions of meteorological origin that lead to impacts in sectors depending on water. This will be clarified in the last paragraph of the introduction and the implications for the results described in the reply to comment 1b (second paragraph) will be included in the discussion.

6) P8 lines 3-5. It is necessary to add some justification for the categorizations of sectors and of type of information. Sectors: e.g. readers may wonder why only rainfed 'cereal' - are there no other rainfed crops? Is all irrigated agriculture similar regarding seasonality/demand? Type: why were these distinguished and e.g. what would be different whether a drought is retrospectively reported or "mentioned". I couldn't find how this classification was used in the correlation analyses - if it wasn't used in the analyses, it is not relevant to mention in the methods section.

We agree that 'rainfed agriculture' is a better label for that sector than 'rainfed cereal' and we will substitute it in the text and figure 2.

The main types of irrigated crops in the area are fruit orchards, alfalfa and maize and they indeed have different seasonality and water demand. The irrigation campaign typically runs from March to October. Most of the records classified as impacts to irrigated agriculture in the analysis refer to insufficient water available for irrigation

and curtailments to the irrigation quota. This is a situation that will affect all crops that require water during the period of the curtailment, although certain crops may be more affected depending on the growing stage at the time of the curtailment.

As mentioned in the reply to comment 1a, the reason to make a distinction between the records that are just mention of drought from others that are reporting the acknowledgement of drought from an official source is the different level of reliability attributed to the two types of records. This distinction was not used in the correlation analysis, but was considered in the interpretation of the results.

7) p.8 line 14 ff. From that section it is unclear, which variables were correlated with which exactly. I suggest to name all variables in the data section and then here add a clear list/matrix of what to expect in the results section.

The correlation is performed between each of the remote sensing parameters and both the timeline that aggregates all types of drought events records and the timeline that aggregates all types of drought impacts (figure 2 in the manuscript). We will rephrase the first sentence of the section 2.3 to make it clearer.

8) P8 line 15 Isn't it the other way round: predictability may be (! But not necessarily) related to the strength of the relation ..... Please elaborate more precisely the link between a correlation analysis and it's potential for prediction (of what exactly?) - this is not necessarily methodologically straightforward.

With the sentence 'The strength of the relationship is related to the predictability of

the occurrence of drought and drought impacts provided by the remote sensing time series' we wanted to point out that the two things are connected. If the remote sensing product does provide information on the occurrence of drought or drought impacts the correlation observed should be stronger. This occurrence of drought impacts is what we are most interested in predicting, since the managers require this information to apply measures to mitigate those possible impacts. When analysing the results of the correlation we need to take into account the limitations of the test before we can establish causal relations between the correlation values obtained and the predictive capability of the product. We will substitute 'is related to' with 'is a function of' in that sentence to avoid the ambiguity.

9) p.8 line 28. How is binary information used in the calculation of correlations, which normally require ordinally scaled data. How is significance tested? How is auto-correlation corrected/considered (Spain always emphasizes that they have multi-year droughts). Citing an R function cannot replace a complete introduction of the statistical methods used, incl. all variable transformations and an introduction of all the measures used later in Results and Discussion sections.

For the correlation analysis between the remote sensing parameters (continuous variable) and the data of drought (impact) occurrence or non-occurrence (dichotomous categorical variable) Pearson correlations were calculated by assigning a numeric binary code to the categories of the dichotomous variable. In this case 1 was assigned to drought (impact) occurrence and 0 to the non-occurrence. The correlation between a continuous and a dichotomous variable is sometimes called point-biserial correlation, but the formula is mathematically equivalent to that of Pearson correlation (Cohen, Cohen, West, & Aiken, 2003).

Testing the significance of cross-correlation results is indeed a delicate issue. The bounds provided by the function in R are just rough guidelines and must be interpreted carefully. The bounds provided in this case are below 0.2 and above -0.2 and for this reason values between -0.2 and 0.2 were not considered for the plots. In view of the difficulty to calculate exact boundaries to test if the cross-correlation results are significantly different from 0, Chatfield (2004) suggests leaning on graphic or numeric tools to support the results.

The time, autocorrelation and partial autocorrelation plots for the series were explored to detect possible issues with the data. The autocorrelation plots show that the auto-correlation diminishes quickly with increasing lag, with the exception of the reservoir indices series. Most of the autocorrelation plots for the reservoir level series present a small peak of autocorrelation at a lag of 12 months, and one of them (management unit 132) presents autocorrelation values declining more slowly (significant values until lag 20).

For the remaining products, autocorrelation for ET, LST, GPP, PsnNet, SM and SPI-3 dissipates mostly at a lag of 2 months. For SPI-1 it goes quicker and is non-existent in some cases. NDVI takes 3-4 months and for SPIs with longer accumulation periods (SPI-6, 9 and 12) the correlation dissipates slower (4, 6 and 8 months respectively), which is inherent to the product. This can indeed have an influence on the correlation of the impacts timeline to NDVI and SPI-6 (and longer), showing stronger correlations. This will be emphasised in the discussion section.

The R function used performs just the cross-correlation function, which we consider to be a well-known function documented in many reference statistic texts. One of these

references was cited in the manuscript. We will rephrase the sentence to have the reference text citation before the mention to the R function. We will also add to this section the explanation regarding the purpose of the correlation test between crop yields and drought indices as described in the reply to comment 3.

10) P9 line 13-end belongs into the methods section. Now it is partly a 1:1 repetition (unnecessary) and more explanation on the method than in the methods section - should be the other way round.

We agree and will remove that paragraph from the results section and complete the explanation in the methods section.

Results section

11) Similar to the impact reports, some time series of the longer series standardized indices (and RS variables?) need to be shown to support that we are indeed looking at drought as an extreme event (e.g. also in 3.1).

We are basing the definition of drought for the analysis on impact occurrence rather than on a frequency analysis, in which drought is defined as an extreme event with respect to the historical series. This is explained in more detail in the reply to comment 1. To put the variability of the test period in the context of the historic record we will include additional references in the section that describes the basin area as mentioned in the reply to comment 3.

12) p.11 line6ff - a description of figure legend in the text is not necessary. This info should be clear from methods section/variable definition and figures' legends and captions. Instead, a more detailed description of correlation patterns shown by the figures needs to be given to warrant having them all in the paper.

Other reviewers, prior to the submission of the manuscript, considered it useful to have a written description of those plots since they found them uncommon. But we will remove the sentence about the x-axis (line 5) as it is already mentioned in the caption of the figure.

We will also add more detail about the correlation patterns as suggested. Especially regarding the differences between rainfed and irrigated areas that appear for some of the products.

13) Quite a bit of discussion is already included in the results section, which makes it a bit difficult for the reader to see the main outcomes, i.e. the strengths of correlations found, the differences and similarities, etc... without immediately being biased towards a possible background explanation. I suggest to moved all sentences with references into the discussion section and here purely describe own results first (see also introductory comment on better focus on unique results).

We agree and will move the explanations regarding the possible reason for weaker correlations for soil moisture (p.11, line 20-22) and GPP-PsnNet (p.11, line 33 - p12)

and of stronger correlations for ET and NDVI (p. 15, line 2 - p. 16, line 1) to the discussion section.

The mention to local press reporting drought impacts of the hydrological year 2006-2007 that were not reflected in the regional press (p. 17, lines 5-7), will be also moved to the discussion section as an example of scale issues (p.18, line 24).

Figures

14) Fig. 3 (and similar) The heading is misleading. Many readers would expect scatter-plots of indicators vs impacts. The legend needs a header relating to a clear variable from the methods section (maximum of cross-correlation what - coefficient? What exactly is shown?)

The headers of figures 3 and 4 will be removed, since the captions already define the plots as the cross-correlation of drought indicators and drought events or impacts.

The symbol for the cross-correlation (defined in equation 4) will be added in the legend. We have found that the formulas for the cross-covariance and correlation included in the manuscript correspond to the theoretical cross-covariance and correlation, rather than the sample cross-covariance and correlation. The former will be substituted by the latter, which is identical but with Latin instead of Greek letters. The header for the legend on figures 3 and 4 will then be r_{xy}.

15) Figure 5 (and similar) is very hard to see/read as labels are too small on A4 print. Also captions and legend labels need to be more precise (p-Values for which test - methods section?) shown are circles scaled to p-values (range not one value), but not "significances" (better to define previously).

The font size in figures 5 and 6 has been increased and the p-values in the legend have been substituted by intervals. The p-values in the figures correspond to the correlation test and refer to the probability to obtain those (or more extreme) results if the variables were not correlated. We will change the word 'significance' to 'reliability' in the caption to avoid the ambiguity.

Discussion section

16) The discussion misses a critical evaluation of the statistics used with respect to assumptions on methods and data (besides the impact reports). In particular, more insight into the RS data and what they inherently 'see' or not linked to drought impacts would help advance the selection of future usage of these as indices.

This is related with the issues discussed in the reply to comment 9 in relation to testing the significance of the results and, most notably, with the implications of the autocorrelation length of the different parameters. As discussed above the latter can have an influence in the strength of the correlation obtained for NDVI, especially for SPI-6 and longer. We will include this in the discussion section, together with the discussion of the results that was originally included in the results section (see reply to comment 13).

17) p. 18 line 17ff. The list of caveats is good, but will be more useful if it were expanded by how these will actually each affect the results.

We agree that this is a useful addition to the discussion. The effects on the results observed for the different aspects included in the list of caveats are the following:

– Accuracy (related to the misuse of the word drought in the newspaper): As previously mentioned in the reply to comment 1a, this issue was the reason to classify the records of drought occurrence according to the source of the information to make a distinction between official sources such as mandated authorities, managers and scientists; and non-official sources such as journalists or water users. This second type of source is the one that is most susceptible to accuracy issues. Particularly for the case of the mandated authorities there are clear procedures with which drought is officially acknowledged, which are defined in the drought management plan. In the records reviewed, only the mention of drought conditions recorded in 2003 is not backed-up by mention of drought from official sources during the same period, and may therefore be regarded as a misuse of the word. Thus, we consider accuracy issues to have little impact on results.

– Completeness: Certain types of impacts received a wider and more detailed coverage. In Fig. 2 of the manuscript some unlikely situations can be identified. For example impacts to livestock in May and July 2006, but not in June. The effect of the incompleteness is reduced by aggregating all reports of impacts across different sectors.

– Scale: The results of the test with crop yield data show values for the hydrological year 2006-2007 for which no drought impacts were identified in the reviewed

regional newspaper that are similar to three other hydrological years for which drought impacts were recorded. Local press for the specific area of the test (Alto Aragón), however, reported a lack of rain from October to March, aggravated with high temperatures, in Monegros and Bajo Cinca that had an impact on rainfed cereals and pastures.

– Bias (related to the over or understatement of drought caused by political or public interests): These do not have an influence in our analysis since we are only considering binary data of occurrence or non-occurrence, but this issue could have a significant impact on the reliability if the records were used to estimate the severity of the event.

These effects will be added to the list of caveats.

Suggestion: For consideration of seasonal aspects and time trends in impact occurrence modelling, perhaps the discussion by Stagge et al. 2015 may provide some further ideas.

Thank you for the relevant reference. It does provide additional insights in the topic and the methodological approach based on logistic regression and Generalized Additive Models (GAMs) is worth considering for further steps. We will reference it in the introduction section as an example of use of impact data to assess indicators for drought detection (p. 3, line 14). It is also a good reference to illustrate in the discussion how the most informative indicators of drought occurrence may vary depending on specific characteristics of the country or basin, such as management practices or dominant water uses.

References

Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Smith, K., Wall, N., Fuchs, B., Crossman, N., Overton, I. (2016) Drought indicators revisited: the need for a wider consideration of environment and society Wiley Interdisciplinary Reviews: Water, 3: 516-536. DOI: 10.1002/wat2.1154

Stagge, J.H., Kohn, I., Tallaksen, L.M., Stahl, K. (2015) Modeling drought impact occurrence based on meteorological drought indices in Europe Journal of Hydrology 530, 37-50 doi.:10.1016/j.jhydrol.2015.09.039

Chatfield, C. (2004). Time-series forecasting. New York: Chapman Hall/CRC.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioural sciences (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

García-Herrera, R., Paredes, D., Trigo, R. M., Franco Trigo, I., Hernández, E., Barriopedro, D., & Mendes, M. A. (2007). The Outstanding 2004/05 Drought in the Iberian Peninsula: Associated Atmospheric Circulation. American Meteorological Society, 8, 483-498.

MMA. (2005). Informe balance del año hidrológico 2004-2005 (Assessment of the hydrological year 2004-2005) (Programa A.G.U.A.) (p. 141). Madrid (Spain): Ministerio de Medio Ambiente.

Spinoni, J., Naumann, G., & Vogt, J. V. (2015). The biggest drought events in Europe from 1950 to 2012. Journal of Hydrology: Regional Studies, 3, 509-524. http://doi.org/10.1016/j.ejrh.2015.01.001