

Content

- Point-by-point response to the reviewers
- List of relevant changes
- Marked manuscript

.....

Point-by-point response to the reviews

REVIEWER 1

We thank the reviewer for taking the time to review the manuscript thoroughly and for the helpful comments and suggestions. Here we provide answers to the specific comments and indications of how the manuscript has been improved to address the reviewer's concerns.

Specific comment:

S1- You have used the newspaper data as benchmark data for drought occurrence. Could you please elaborate on why you decided to choose newspaper data? How reliable do you think the newspaper data is? What is the false alarm rate?

The focus of our analysis was to identify indicators that could help managers in detecting drought conditions that may lead to impact and therefore we were looking for a benchmark based on drought impact occurrence.

The only available database of drought impact, the European Drought Impact report Inventory, doesn't provide enough records to allow an analysis at regional scale. And although there are a few reports and scientific articles that describe drought impacts in the area, especially in relation with the drought period 2005-2008, impact data in these studies is aggregated by year or even for the whole drought episode.

We selected newspaper records as a benchmark data source because it allowed us to systematically collect impact occurrence data of all affected sectors with a monthly time step for the whole period of analysis.

As mentioned in section 4.1. (second paragraph), a few reliability issues were noticed when collecting the data. The issues related to the wrong use of the word drought to report different situations such as summer shortages are the ones that may have the most impact on the reliability of the source. With this issue in mind, we classified the records of drought occurrence according to the source of the information about drought occurrence to make a distinction between official sources such as mandated authorities, managers and scientists (labelled in Fig 2 as "drought acknowledged by the authorities", "ongoing mitigation measures", and "periods retrospectively defined as anomalously dry", respectively) and non-official sources such as the journalist or the water users (labelled as "mention of drought occurrence"). This second type of source is the one that is susceptible to accuracy issues. In our case, only the mention of drought recorded in 2003 is not backed-up by mentions from official sources during the same period and may therefore be considered a false alarm. The rate of false records derived from a misuse of the word drought depends on the rigour of the newspaper.

Bias issues (over or understatement of drought caused by political or public interests) do not have an influence in our analysis since we are only considering binary data of occurrence or non-occurrence. However, we acknowledge that these can have a significant impact on reliability if the records are used to estimate the severity of the event. These clarifications will be added to the second paragraph of the discussion.

S2- For the daily remote sensing data, you have used “monthly aggregate”. Could you please explain more on how the aggregate was obtained?

Monthly aggregates were generated from precipitation, Land Surface Temperature (LST) and Soil Moisture (SM) daily datasets. Precipitation data in mm/day was aggregated by a sum of the daily values for each pixel to obtain monthly data in mm/month. LST and SM data were aggregated by averaging the daily values for each pixel. A clarification will be added to the remote sensing data section.

S3- You have used the cross-correlation function as expressed by (Chatfield, 2004), which assumes the bivariate data is stationary and ergodic (on page 122 of Chatfield, 2004). How did you test if your data was stationary and ergodic?

We acknowledge that stationarity as well as ergodicity of the data is a point of concern. Part of the input data (LST, SM, NDVI, ET, GPP and PsNet) present a seasonal trend. For these, monthly anomalies were obtained by subtracting the mean for the whole period from each monthly average value, using these anomaly time series as input for the cross-correlation function. We have added a comment to clarify this in the manuscript.

To detect possible issues related to the stationarity or ergodicity of the series, their time autocorrelation and partial autocorrelation plots were considered. Overall, time plots show no trends or discontinuities and the values in the autocorrelation plots show that the autocorrelation diminishes quickly with increasing lag. An exception are the series of the reservoir indices. In that case, for some of the series it is not clear from the plot if the series is stationary. For one of them (management unit 122) it clearly is not. This management unit corresponds to a reservoir (Rialb) that started to be filled in the year 2000 and therefore the levels cannot be considered stationary for the period of study. 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).

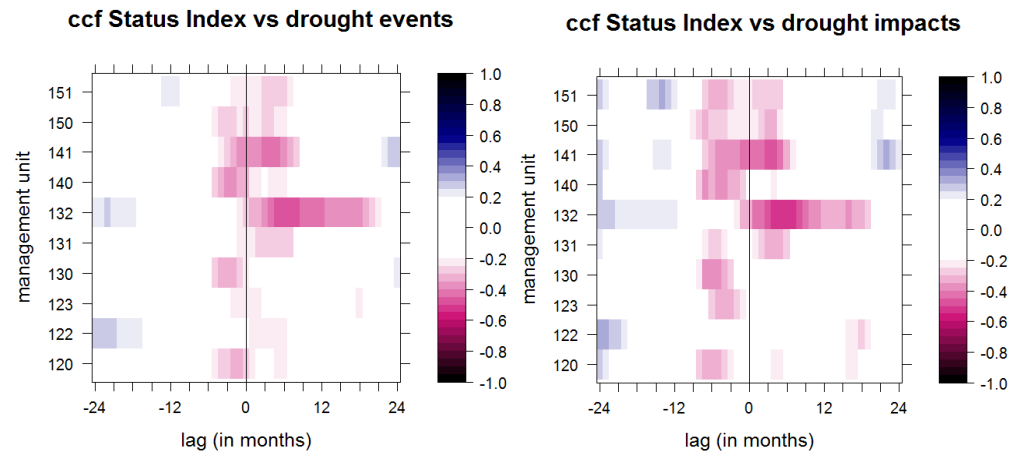
Additionally, we have run a stationarity test (Dickey-Fuller test) on the series. The results show that most of the datasets are stationary at the 95% confidence level. There are a few exceptions: for the LST series, two of the management units (120i and 130X) have a p-value of 0.07; while for the SPI-12 series, four of the management units have p-values higher than 0.05, these are management units 132 (p-value=0.07), 140 (0.10), 150 (0.19) and 151 (0.07); for the reservoir index series, again management units 122 and 132 present p-values over the significance level, 0.13 and 0.7 respectively.

We considered the remote sensing datasets to be stationary and ergodic enough to be used as input for the cross-correlation function. As for the reservoir data, a comment on the two reservoirs that don't satisfy the conditions could be included in the article.

We have verified the plots showing the anticipation to drought events. A small error was detected in these plots on careful review, which was corrected (Figure P3 shows the corrected plots). In these plots we can clearly see that the two management units that do not

satisfy the conditions for stationarity are those (at least two out of the three) that do not present anticipation.

This last sentence will be included in the results section.



P3. Corrected status index cross-correlation plots

S4- You have used the cross-correlation function as a method of drought anticipation. Can the same technique be applied for foreseeing the end of droughts?

This is a very interesting question, given the difficulty in identifying the cessation of drought events. By using records of direct drought impact as a benchmark to assess the different parameters we were aiming to identify the onset of the conditions that cause these impacts. A good knowledge of the conditions that lead to impacts would be useful to identify both the start and end of a drought event, since the end of those conditions (return to normal conditions) would be an indicator of the end of the drought event.

S5- As you explained in figure 3 and 4, the values on the left side of (negative side) of the central line, show how early the remote sensing data have anticipated the newspaper headlines on droughts, or anticipated the decline in crop yield. Could you please explain what the values on the right side (positive side) of the central line show?

The positive side of the plots reflects the correlation of the drought occurrence and impact series with the values of the different datasets in later months. For example at lag +1 the correlation between the impact and the value of the dataset the following month is shown. This type of correlation appears if the conditions that define the start of the event or impact occurrence last longer than one month.

For long term SPIs these correlation with positive lags are stronger since previous months are included in the SPI. For example for SPI-9 at lag +1 the drought event or impact occurrence is compared to an SPI built from the following month and the 8 previous months.

S6- Following from the previous comment, regarding figures 3 and 4, could you please explain what the values on both extremes of the graphs mean? They mostly seem to be happening around lag -24 and lag 24.

The positive correlations of the timelines of drought occurrence and impacts with the values of the indicator datasets with lags over one year (most notably at -15 and ± 24 lags) are

casual correlations. For example the figure P6 shows how the periods of reported drought impacts coincide with positive anomalies when the data is shifted by a lag of -15.

Since we are focusing our evaluation at anticipations within a period of one hydrological year, the correlations should not be affected by this issue. An explanation will be added to the results section.

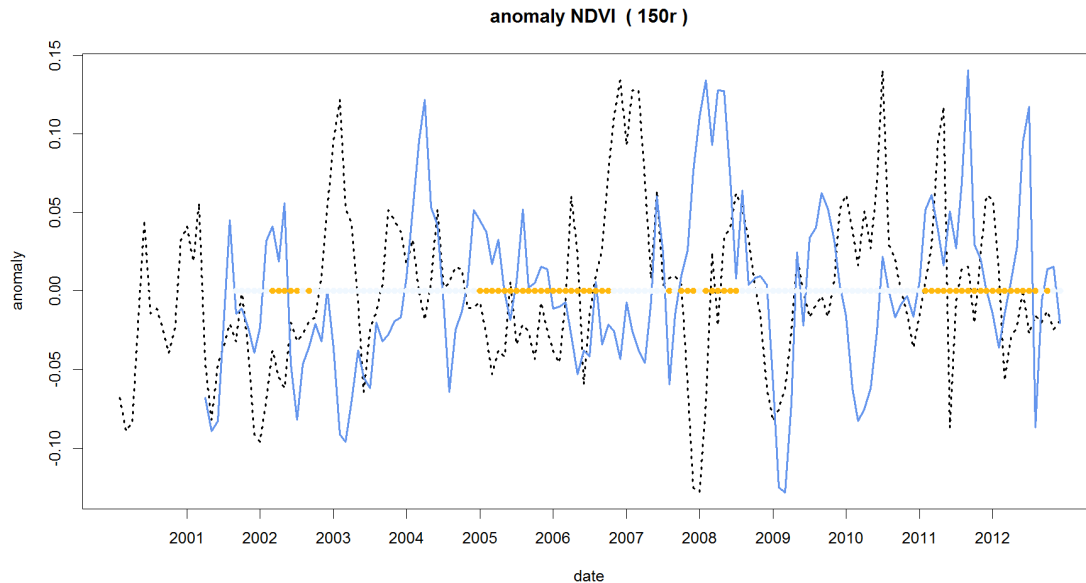


Figure P6. The dotted black line represents the anomaly of NDVI at zone 150r and the blue line is the same data with a lag of -15 months. The yellow and white points in the middle represent the months with and without reported drought impact respectively.

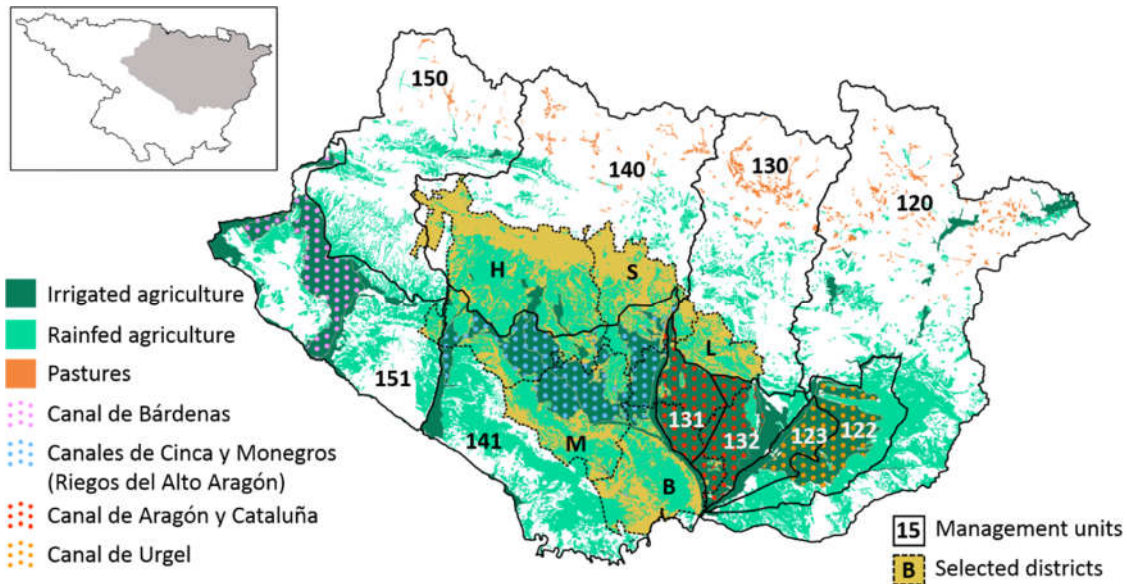
Technical suggestions:

T1- I suggest signposting the paper early in the introduction. The reason is that the introduction although being an interesting read, is rather long.

We agree with the suggestion and we will introduce the aim of the paper at the end of the second paragraph (page 2, line 10) of the introduction instead of in the last paragraph (page 3, lines 25-27).

T2- It would be easier for the reader to have the zone numbers on the map in figure 1.

We agree, and on inspection found that the figure seems to have lost the numbers in the conversion to pdf. We have amended the figure to ensure these numbers are preserved:



T3- On page 7 line 19, it would be interesting to state briefly why the 2nd largest news-paper in the Aragón region selected, and not the 1st largest? Is it because unlike all the other newspapers in the region, it had an online record?

Yes, the online record of the main newspaper only goes back to 2008. We also reviewed the available records for that newspaper, but didn't include them in the analysis to avoid having the later years better represented than others. This will be clarified in the manuscript.

T4- On page 8 line 9, why only “winter” cereal crops are selected?

Winter cereals are the cereal crops that are planted in the autumn and they are the crops that cover the largest surface area (especially barley and wheat). Their importance for the region results in better data availability than other crops and for this reason only this type of crops were selected for the analysis. A clarification will be added to the manuscript.

T5- On page 11, line 8, I think it would be an easier read if the lag example values appear with negative sign as they belong to the times before the central line.

T6- I would have changed “this side of the central line” on page 11 line 8 to the “left” side of the central line to make it clearer.

T7- Page 11, line 18, I think 0.6 should have a negative sign as it is referring to a negative correlation (as depicted in pink in figures 3 and 4).

T8- On page 13 in the header, 3.3 is repeated twice

Technical suggestions 5-8: We agree with the corrections suggested by the reviewer and will modify the text accordingly.

.....

REVIEWER 2

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).

1a.- 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 (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).

1b.- 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.

1c.- 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 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 predictability. 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). [finally added to section 2.2.3 Benchmark datasets].

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 autocorrelation 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>

List of relevant changes

1. Introduction

1. Sentence about the aim of the paper moved from the last paragraph of the introduction (page 3, lines 25-27) to the end of the second paragraph of the introduction (page 2, line 10) [R1 T1].
2. Reference to Stagge et al. 2015 added (p. 3, line 14) as an example of use of impact data to assess indicators for drought detection [R2 Suggestion]
3. Last paragraph of the introduction: clarified that “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” [R2 1b, R2 5]

2. Material and methods

2.1. The Ebro basin

1. Clarification about the variability of the period added: “The period 2000-2012, selected for the analysis, encompasses a wide range of different conditions: 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), while the hydrological year 2003-2004 is considered one of the wettest hydrological years of the country’s record (MMA, 2005)” [R2 1c].
2. Zone numbers added to figure 1 [R1 T2].

2.2. Input datasets

2.2.1. Remote sensing data

1. Description of the aggregations performed on the datasets moved from the first paragraph to the end of the section. Additional clarifications added [R1 S2 and R1 S3].
2. Added: reference to the monthly GPP and PsnNet product used in the analysis, downloaded from the Numerical Terradynamic Group (NTSG) (p.6, line 13) [R2 1b].

2.2.2. In situ data

2.2.3. Benchmarking data sets

1. Clarification on the reasons to use of a drought definition based on impacts occurrence rather than on a frequency analysis (p7, line 20): “The short length of the remote sensing 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.” [R2 1]

2. Specified that the impact data is also used to reconstruct the onset and evolution of drought conditions during the period of analysis (p8 line 5): “were used to reconstruct the onset and evolution of drought conditions during the period of analysis and as benchmark for the comparison of the remote sensing datasets.” [R2 4]

3. Added clarification on the reason to select newspaper records as a source of impact data (p8 line 6): “Newspaper records were selected as data source because it allowed a systematic collection of impact occurrence data of all affected sectors with a monthly time step for the whole period of analysis.” [R1 S1].

4. Added clarification of the purpose of the correlation analysis with crop yields and mention to the wide use of correlation between crop yields and drought indices (p8 line 9): “The correlation of remote sensing data, especially SPI and NDVI, to agriculture yield data has been widely researched and applied (see Bachmair et al., 2016, for a review). This second type of impact data was included 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 and discuss the advantages and limitations of one with respect to the other” [R2 3]

5. Added clarification of the reason to select the second largest newspaper (p8 line 14): “was selected for the review because it has an online record going back to September 2001”. [R1 T3]

6. Substituted “rainfed cereal” by “rainfed agriculture” (description of labels of figure 2, P8 lines 18) [R2 6].

7. Added clarification of the reason to classify the records and details on the classification moved to this section (P8 line 18): “The records of drought occurrence are classified according to the source of the information, making a distinction between non-official sources such as journalists and water users, labelled “mention of drought occurrence” in Figure 2, and official sources labelled as “drought acknowledged by the authorities”, “ongoing mitigation measures” and “periods retrospectively defined as anomalously dry”. This last type corresponds mainly to news about the publication or communication of analysis performed by the scientific community or the water managers describing an ongoing or past drought”. [R2 6 and R2 10]

8. Added clarification of the reason to focus the second test on winter cereal (P8 line 28): “Winter cereals are the cereal crops that are planted in the autumn and they are the crops that cover the largest surface area. Their importance for the region results in better data availability than other crops and for this reason this type of crops were selected for the analysis” [R1 T4]

2.3. Correlation between remote sensing data and the benchmarking data sets

1. First sentence of the paragraph rephrased to make clear what is correlated to what (p9 line 2): “The correlation 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) was analysed in terms of strength of the relationship and anticipation” [R2 7]

2. Substituted “is related to” with “is a function of” (P9 line 3) [R2 8]

3. Added clarification of the aim of the correlation analysis (P9 line 6): “The aim of this analysis 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.” [R2 2]

4. Sentence about the cross correlation function rephrased to have the reference text citation before the mention to the R function (P 9 line 16). Reference to R function moved to line 20. [R2 9]

5. Equations 3 and 4: substitution of Greek symbols by Latin symbols (r , c) [R2 14]

3. Results

3.1. Drought events and impacts

1. Figure 2: substituted “rainfed cereal” class label by “rainfed agriculture” [R2 6]

2. Description of the classes moved from this section to section 2.2.3. [R2 10]

3.2. Correlation of text-based records and remote sensing indicators

1. Sign changed to negative in the example of the lag values (p11 line 33) [R1 T5] and in the value correlation [R1 T7].

2. Sentence about the x-axis removed (p 11 first paragraph) as it is already mentioned in the caption of the figure (p14 line 1) [R2 12].

3. Page 11, line 33. Changed “this side of the central line” to “the left side of the central line”. [R1 T6].

4. Added clarification of the meaning of the correlations on the positive side of the x-axis (p11 line 31): “The positive side of the plots reflects the correlation of the drought occurrence and impact series with the values of the different datasets in later months. This type of correlation appears if the conditions that define the start of the event or impact occurrence last longer than one month.” [R1 S5]

5. Explanations regarding the possible reason for weaker correlations for soil moisture (p.14, second paragraph) and GPP-PsnNet (p.14, third paragraph) moved to discussion section [R2 13]

6. Figures 3 and 4. I_e plots fixed [R1 S3], header removed and symbol for the cross correlation added in the legend (r_{xy}) [R2 14]

7. Added clarification on the results of the time and autocorrelation plots (evaluated to detect possible issues with the input data) (p14 line 14) and a comment on the two reservoirs that don't satisfy the stationarity condition: “The time plots obtained for each of the parameters present no

trends or discontinuities and the values in the autocorrelation plots show that the autocorrelation diminishes quickly with increasing lag. An exception are the series of the reservoir indices. In that case, for some of the series it is not clear from the plot if the series is stationary. For one of them (management unit 122) it clearly is not. This management unit corresponds to a reservoir (Rialb) that started to be filled in the year 2000 and therefore the levels cannot be considered stationary for the period of study. 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). In the I_e plots in Figures 3 and 4 it can be clearly seen that the two management units that do not satisfy the conditions for stationarity (management units 122 and 132) are those (at least two out of the three) that do not present anticipation. 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.” [R1 S3]

8. Added a clarification of the meaning of the positive correlation values in the plots: “The positive correlations of the timelines of drought occurrence and impacts with the values of the indicator datasets with lags over one year (most notably at -15 and ± 24 lags) are casual correlations. Since the analysis presented here focuses on anticipations within a period of one hydrological year, the correlations should not be affected by this issue.” [R1 S6].

3.3. Correlation of crop yield and remote sensing indicators

1. Section header fixed [R1 T8]

1. “significance” substituted by “reliability” (p 14 line 35). The same sentence was repeated in the captions of figures 5 and 6 and has been removed. [R2 15]

2. Explanations regarding stronger correlations for ET and NDVI and the mention to local press reporting drought impacts of the hydrological year 2006-2007 that were not reflected in the regional press moved to the discussion section [R2 13]

3. Figures 5 and 6: Font size increased and p-values in the legend substituted by intervals [R2 15].

4. Discussion

4.1. Use of impact records as drought reference

1. List of caveats expanded with examples from the analysis presented in the paper (p19 line 8) [R2 17]

2. Added clarification of the implications of the length of the remote sensing series (p19 line 30): “The length of the period of analysis does not 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” [R2 1b]

3. Added reference to the work of Spinoni, corroborating our results in identifying three drought events during this short period (p19 line 34): “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 Catalonia, it was still low during that year (MMA, 2007).” [R2 1a]

4. Clarification regarding stronger correlations for ET and NDVI moved from results to this section (P20 line 19). Mention to local press reporting drought impacts of the hydrological year 2006-2007 that were not reflected in the regional press moved to this section as an example of scale issues (P19 line 22) [R2 13]

4.2. Early drought detection with remote sensing products

1. Added clarification about the implications of the autocorrelation of the different datasets (p21 line 6): “However, while for most products the autocorrelation dissipates at a lag of 2 months, for NDVI and SPI-6 it takes 3-4 months and this can have an influence in NDVI and SPI-6 showing stronger correlations. For SPIs with longer accumulation periods (SPI-9 and 12) the correlation dissipates even slower.” [R2 9]

2. Explanations regarding the possible reason for weaker correlations for soil moisture and GPP-PsnNet moved from results to this section (p21 line 9) [R2 13]

3. Added reference to Stagge et al 2015 as an example to illustrate 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” (p 22 line 22) [R2 Suggestion]

5. Conclusions

6. References

.....

Marked-up manuscript

The predictability of reported drought events and impacts in the Ebro Basin using six different remote sensing data sets

Clara Linés¹, Micha Werner^{1,2} and Wim Bastiaanssen^{1,3}

¹ UNESCO-IHE, Institute for Water Education, Delft, the Netherlands

² Deltares, Delft, the Netherlands

³ Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft, the Netherlands

Abstract

The implementation of drought management plans contributes to reduce the wide range of adverse impacts caused by water shortage. A crucial element of the development of drought management plans is the selection of appropriate indicators and their associated thresholds to detect drought events and monitor their evolution. Drought indicators should be able to detect emerging drought processes that will lead to impacts with sufficient anticipation to allow measures to be undertaken effectively. However, in the selection of appropriate drought indicators the connection to the final impacts is often disregarded. This paper explores the utility of remotely sensed data sets to detect early stages of drought at the river basin scale, and how much time can be gained to inform operational land and water management practices. Six different remote sensing data sets with different spectral origin and measurement frequency are considered, complemented by a group of classical in situ hydrologic indicators. Their predictive power to detect past drought events is tested in the Ebro basin. Qualitative (binary information based on media records) and quantitative (crop yields) data of drought events and impacts spanning a period of 12 years are used as a benchmark in the analysis. Results show that early signs of drought impacts can be detected up to some 6 months before impacts are reported in newspapers, with the best correlation-anticipation relationships for the Standard Precipitation Index (SPI), the Normalized Difference Vegetation Index (NDVI) and Evapotranspiration (ET). Soil Moisture (SM) and Land Surface Temperature (LST) offer also good anticipation, but with weaker correlations, while Gross Primary Production (GPP) presents moderate positive correlations only for some of the rainfed areas. Although classical hydrological information from water levels and water flows provided better anticipation than remote sensing indicators in most of the areas, correlations were found to be weaker. The indicators show a consistent behaviour with respect to the different levels of crop yield in rainfed areas among the analysed years, with SPI, NDVI and ET providing again the stronger correlations. Overall, the results confirm remote sensing products' ability to anticipate reported drought impacts and therefore appear as a useful source of information to support drought management decisions.

1. Introduction

Drought is defined as a temporary water shortage in part caused by anomalous climatic conditions but strongly influenced by socio-economic factors (Kallis, 2008). The effects of drought propagate

through all human and natural systems that depend on water directly or indirectly producing substantial losses (Wilhite et al., 2007). Various economic sectors are adversely affected, in particular agricultural production, energy generation, and water supply for domestic and industrial use. Habitat degradation, increased mortality of flora and fauna, and increased occurrence of wildfires are examples of the effects on the natural environment. Indirect impacts, such as increase on prices, unemployment or migration, arise as a consequence of the direct impacts, and may be felt in a much wider area, even reaching the global scale (Wilhite & Vanyarkho, 2000).

The occurrence and severity of drought impacts depends on the intensity and duration of the event, but also on the vulnerability of the society and the environment (Wilhite, 2000). As a consequence the conditions that produce negative socio-economic impacts are not necessarily the same for the different sectors that may be affected (Redmond, 2002). The timing of the event also influences the severity of impacts. Soil moisture deficit during the flowering stage of a crop or reduced domestic water supplies during the tourist season are examples of situations in which the socio-economic impact is aggravated due to the timing of the drought. The aim of the current paper is to identify earth observation data sets that can be used to detect early stages of drought at the basin scale, as well as to what extent these data sets can anticipate drought impacts, and be used to inform operational land and water management.

The implementation of drought management plans by governing agencies can contribute to reducing the negative effects of drought by guiding decision makers in taking appropriate mitigation actions. However, the effectiveness and cost-efficiency of these actions rely on the selection of suitable indicators to monitor drought conditions, and to detect drought events at an early stage, gaining valuable time for mitigation measures to be implemented effectively and impacts to be mitigated. Examples of actions that can be taken include retention of water; reallocation of available water resources; curtailment of current allocations; recommendations to plant less water-demanding or drought resistant crops; or prohibition of certain water uses (e.g. watering gardens or washing cars).

Indicator systems consist of drought indices with associated thresholds that allow classifying the event in categories of drought severity. A classical examples is the division of river flow into several categories. When the value of the indicator crosses one of the thresholds, managers should decide whether to activate the corresponding responses defined in the drought management plan for that situation. Indicators and associated thresholds should be problem, context and user-specific (Kallis, 2008), and therefore an integrated management of droughts in basins where there are multiple users requires advanced drought detection systems based on multiple indicators.

Measurements from in situ networks and from remote sensing are complementary sources that can be used to build the system of indicators for early detection and monitoring of drought conditions. In situ data is generally collected at specific points only. The advantage is the high temporal frequency of observations, and the availability of longer term records. Remote sensing techniques, on the other hand, offer cost effective and spatially continuous information over extended regions. Satellites allow drought events to be categorised over a certain area, rather than at point locations (Famiglietti et al., 2015; Kogan, 2001). Several satellite data sets are now available at a daily or at even at shorter time scales, offering excellent potential to develop sound drought monitoring

systems in real time, and allowing to overcome the shortcomings of classical indicators based on in situ data sets that lack the spatial scale (e.g. Sheffield et al., 2014; van Dijk & Renzullo, 2011).

Keyantash and Dracup (2002) analyse a set of criteria to assess the usefulness of drought indicators for the assessment of drought severity and point out that while the robustness of an indicator provides insight into its consistent behaviour across differing conditions, assessing the accuracy of the information provided by the indicator requires a standard or benchmark for comparison. This holds true for both remote sensing-based and ground-based indicators. A standard that offers an absolute metric of drought is not easily available and likely does not exist, and as a result a common approach to evaluating the performance of remote sensing-based drought indicators is to assess their robustness by comparing them with other indicators such as flow, reservoir levels or widely used drought indices (e.g. Morid et al., 2006; Tsakiris et al., 2006; Vasiliades et al., 2011). Expert knowledge may also be used in practical applications as a benchmark to assess drought indicators such as in Steinemann et al. (2015), who rely on regional water managers, drought decision-makers and other stakeholders' knowledge as a reference to develop, select and evaluate drought indicators. Expert judgement is also included, in combination with several indicators and model outputs, in the U.S. drought monitor (Svoboda et al., 2002) to develop a weekly map of drought conditions in the U.S. which itself is also frequently selected as a reference data set in the evaluation of the performance of drought indicators in the country (e.g. Anderson et al., 2011, 2013; Brown et al., 2008).

Since mitigating impacts is the purpose of drought indicators included in drought management strategies, impact data is especially suitable as a benchmark in this case. Drought impacts, however, are difficult to evaluate and are rarely monitored (Lackstrom et al., 2013; Wilhite, 2011). Several studies have analysed the connections of drought indices to quantifiable effects on agriculture, hydrology or forests (see Vicente-Serrano et al., 2012, for a review), but very few have applied the impact data (mostly crop yields) as a benchmark to assess indicators for drought detection (e.g. Potop, 2011; Sepulcre et al., 2012; Stagge et al., 2015)(e.g. Potop, 2011; Sepulcre et al., 2012). Recognising the potential of this kind of data for drought management, two large scale initiatives have recently been launched: the US Drought Impact Reporter (DIR) (Wilhite et al., 2007) and the European Drought Impact report Inventory (EDII) (Stahl et al., 2016). These have the objective to collect text-based impact records systematically with the aim to increase their availability and accessibility. Recent studies have explored the links of the EDII records to drought indicators (Bachmair et al., 2016; Bachmair et al., 2015; Blauhut et al., 2015), though these have focused on the national scale, and it is recognised that further development is required to allow analysis at the subnational scale (Stahl et al., 2016).

Despite their important role in mitigation of drought impacts, the selection and use of indicators and thresholds for decision making often suffers from a lack of scientific justification: only a few studies have analysed the choice of drought indicators in relation to drought management in practice (Steinemann et al., 2015; Steinemann & Cavalcanti, 2006). Moreover, the thresholds that have been selected to declare droughts are only rarely connected to the specific impacts that need to be avoided (Wilhite, 2000). ~~The aim of the current paper is to identify earth observation data sets that can be used to detect early stages of drought at the basin scale, as well as to what extent these data sets can anticipate drought impacts, and be used to inform operational land and water management. The innovation is that~~ In this paper quantitative and qualitative drought impact

information are applied as a benchmark in evaluating the utility of indicators derived from six different remote sensing data sets at river basin scale. This implies that the analysis is 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. Two aspects are considered in the assessment of the indicators against the benchmark data: how well do these indicators reflect reported drought impacts and to what degree can these indicators be used to anticipate drought conditions and consequent impacts.

2. Material and methods

2.1. The Ebro basin

The Ebro basin, with an extent of 85,600 km², is the largest catchment in Spain. It is located in the north-east, bounded by the Pyrenees and Cantabrian mountain ranges to the north, and the Iberian System to the south. It is a highly regulated basin with 51 reservoirs (> 1 Mm³) and a total storage capacity of more than 7,500 Mm³, which supply water to more than 900,000 has of irrigated agriculture and more than 450 hydroelectrical plants (CHE, n.d.).

Analysis of the impacts of a recent drought event (2005–2008) revealed that agriculture and food production are the main sectors affected by drought in the area, but impacts to hydropower production, water supply to villages, food industry, recreational activities and ecosystem functions were also identified (Hernández-Mora et al., 2013; Perez y Perez & Barreiro-Hurlé, 2009).

The period 2000-2012, selected for the analysis, encompasses a wide range of different conditions: 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), while the hydrological year 2003–2004 is considered one of the wettest hydrological years of the country's record (MMA, 2005).

The Confederación Hidrográfica del Ebro (CHE) is the organization responsible for the management, regulation and conservation of water in the Ebro basin. The basin is divided into 18 management units, each of which has a board constituted of representatives of the different water users as well as of the basin authority to coordinate the use of the hydraulic infrastructures and water resources in their area.

A drought management plan for the basin was developed in 2007 to guide drought management actions (CHE, 2007). The plan defines a set of indicators to detect situations of hydrological drought in the Ebro basin and evaluate their severity. The indicators are built using observations from a rich network of in situ automatic stations. In the areas in which the flow is regulated by dams, water stored in reservoirs is considered the most robust indicator, but other variables such as water flow, snow depths or head levels in aquifers may also be taken into account. For areas with a natural or an almost natural flow regime without reservoirs, the three-month water flow measured at representative stations is selected as the main indicator. In one of the management units, where there is no regulation and no representative rivers, groundwater levels are used as indicator.

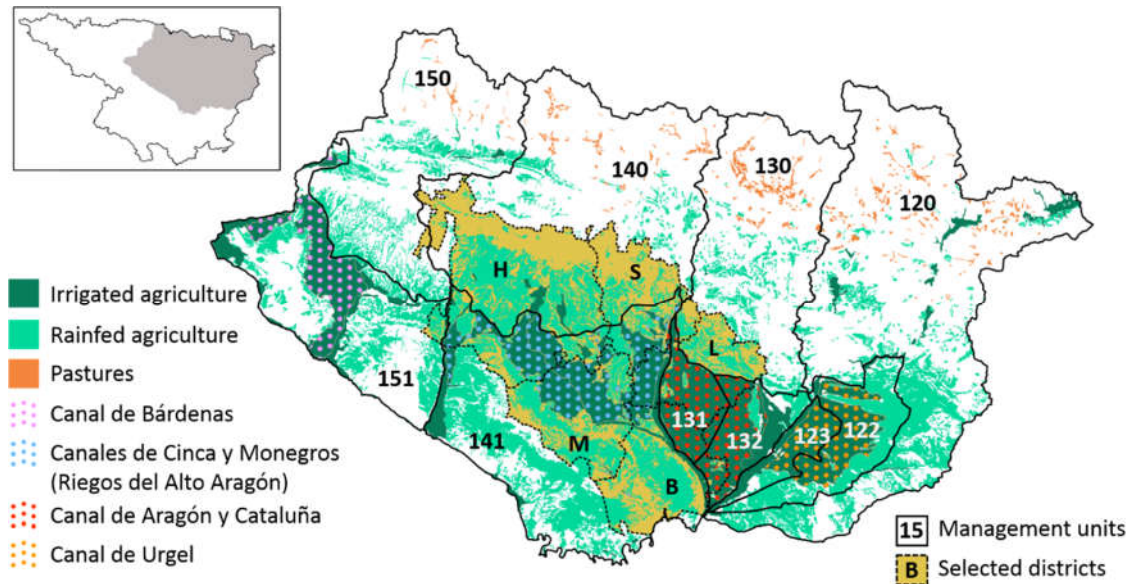


Figure 1. The north-eastern part of the Ebro basin with selected agricultural land cover information.

The north-east of the basin, where the larger irrigation districts of the Ebro basin are located, was selected to evaluate the set of drought indicators against the qualitative text reports (Figure 1). This area was also the most affected by the drought period 2005–2006 (Hernández-Mora et al., 2013). It is composed of four management units (management units 12 to 15). Figure 1 shows the management units further subdivided according to the main drought indicators currently selected in each: three-month water flow in the northern sectors (zones 120, 130, 140 and 150) and reservoir levels in the southern sectors. To differentiate non irrigated agricultural areas, Corine Land Cover map 2006 (CLC06) classes 211 (non-irrigated arable land), 242 (complex cultivation patterns) and 231 (pastures) have been used. The land cover class irrigated agriculture corresponds to the irrigation polygons provided by MAGRAMA (1997 version). The main irrigation districts are marked with dotted patterns and are identified by the name of the main canal that serves them. For the evaluation of quantitative crop yield data as a benchmark, only five of the districts within this area were selected: Hoya Huesca (H), Somontano (S), La Litera (L), Monegros (M), and Bajo Cinca (B), including irrigated and rainfed cropland.

2.2. Input data sets

2.2.1 Remote sensing data

The analysis focuses on medium resolution global remote sensing products that are related to land surface hydrological and vegetation growth processes. Six commonly used remote sensing parameters were investigated: Precipitation (P), Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), Gross Primary Production (GPP), top soil moisture (SM) and actual EvapoTranspiration (ET). In order to have one common time interval, monthly aggregated data sets were obtained for each parameter (Table 1). The remote sensing data has been aggregated per management unit. Pixels with at least 85% of their area within the management unit (30% in the case of soil moisture due to the coarse resolution) were considered in establishing the aggregate value for that unit. For the 1 km resolution LST, NDVI and ET remote sensing products, the results are also analysed by land cover type. This relates to pixels where at least 85% of the area consists of

irrigated and rainfed agriculture land cover classes (see Figure 1). The selected remote sensing data sets are the following:

Precipitation (P): The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a gridded precipitation data set based on satellite and station data, designed with the main objective to support agricultural drought monitoring. It is a daily, quasi global product, with a resolution of 0.05°. The data set is available from 1981 to near-present. It is based on top cloud temperature measured by geostationary satellites and the Tropical Rainfall Measuring Mission (TRMM) satellite with a rainfall radar aboard. A detailed description of the product can be found in Funk *et al.* (2015). In this study, monthly aggregated rainfall values have been converted into Standard Precipitation Index (SPI) data sets. The SPI (McKee *et al.*, 1993) is a normalised rainfall anomaly, computed by comparing the accumulated rainfall over a given period with the long-term record. The SPEI R-package developed by Beguería and Vicente-Serrano (2013) was used to calculate SPI for periods of 1, 3, 6, 9 and 12 months. The possibility to calculate the index for different periods is one of the strengths of SPI as it allows to explore the effects of rainfall anomalies of different duration. SPI values calculated for shorter periods are associated to meteorological drought, while those calculated for longer period are often associated to hydrological drought (World Meteorological Organization (WMO), 2012).

Land Surface Temperature (LST): The MODIS (Moderate Resolution Imaging Spectroradiometer) product MOD11A2 offers day and night LST data sets, available at 1 km resolution as daily and 8-day products (<http://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>). The daily daytime LST data have been used for the current study. LST is based on longwave emissions in the thermal infrared range (10 to 12 µm).

Vegetation health (Normalized Difference Vegetation Index, NDVI): The MODIS Vegetation Indices product (MOD13) provides information on the active leaf chlorophyll, and thus indirectly on the photosynthetically active process. NDVI describes the ratio of the difference and sum of reflected radiances in the red (0.65 µm) and near-infrared parts of the spectrum (0.9 µm). MOD13 is available at different resolutions: 16-day (250 m, 500 m and 1 km) and monthly (1 km) (<http://modis.gsfc.nasa.gov/data/dataproduct/mod13.php>). The monthly 1km data product has been used in the current analysis.

Gross Primary Production (GPP) and PsnNet: GPP describes the daily gross carbon flux as a result of the photosynthetic process and is thus suitable to detect the effects of drought on biomass production. The MODIS GPP product (MOD17) applies a light-use efficiency model based on MODIS FPAR (fraction of photosynthetically active radiation) data, meteorological data and biome-specific parameters. The product also includes Net Photosynthesis (PsnNet), which corresponds to the GPP minus the maintenance respiration for leaves and roots. It is available at a spatial resolution of 1 km as 8-day composites or as annual values (<http://modis.gsfc.nasa.gov/data/dataproduct/mod17.php>), and as monthly aggregates (Numerical Terradynamic Simulation Group, NTSG, <http://www.ntsug.umd.edu/>). ~~The product also includes Net Photosynthesis (PsnNet), which corresponds to the GPP minus the maintenance respiration for leaves and roots.~~ Additional background information can be found in Running and Zhao (2015).

Soil Moisture (SM): The soil moisture product considered is taken from the Soil Moisture CCI project, part of the ESA Programme on Global Monitoring of Essential Climate Variables (ECV) (Liu *et al.*,

2011, 2012; Wagner et al., 2012). Three daily products are available (data sets based on active, passive or merged microwave instruments) for the period 1978 to 2014 at a spatial resolution of 0.25 degrees (<http://www.esa-soilmoisture-cci.org/>). Higher spatial resolution products are only available for certain areas. The product used in the current analysis is the merged CCI SM data set. The data is based on C-band scatterometers and multifrequency radiometers.

EvapoTranspiration (ET): There are currently three global data sets of actual ET in the public domain. These are the MODIS ET product (MOD16, <https://modis.gsfc.nasa.gov/data/dataproduct/mod16.php>) developed by the University of Montana (Mu et al., 2007, 2011) and supported by NASA; the Surface Energy Balance System SEBS developed by Su (2002), and the Global Land Evaporation Amsterdam Model (GLEAM) developed by Miralles et al. (2011) which is available through www.gleam.eu. In addition, there are global ET products that are quasi open-access, including the Atmosphere–Land Exchange Inverse Model (ALEXI) being developed by Anderson et al. (1997) from the USDA in conjunction with Hain et al. (2009) from NOAA; the Operational SEBS (SEBSop) of the USGS (Chen et al., 2016; Senay et al., 2013) and the C.M.R.S. Evapotranspiration (CMRSET) published by Guerschman et al. (2009) from CSIRO in Australia. In this paper an ensemble product based on these individual products with accumulated monthly ET values at a pixel resolution of 250 m x 250 m is used. This ensemble ET product (ETens v1.0) is available from the Water Accounting Group of UNESCO-IHE (www.wateraccounting.org). The six individual ET models considered all use different parts of the spectrum, which reinforces the power of this tool.

Table 1. Selected remote sensing products

Parameter	Product	Pixel size	Original time interval
P	CHIRPS	0.05 deg	Daily
LST (day)	MOD11A2	1km	Daily
NDVI	MOD13A3	1 km	Monthly
GPP and PsnNet	MOD17 (NTSG)	1 km	Monthly
SM	Merged SM product (CCI project)	0.25 deg	Daily
ET	Ensemble	250 m	Monthly

In order to have one common time interval, precipitation data in mm/day was aggregated by a sum of the daily values for each pixel to obtain monthly data in mm/month and LST and SM data were aggregated by averaging the daily values for each pixel (Table 1). Part of the input data (LST, NDVI, SM, ET, GPP and PsnNet) present a seasonal trend. For these, monthly anomalies were obtained by subtracting the mean for the whole period from each monthly average value, using these anomaly time series as input for the correlation. The remote sensing data has been aggregated per management unit. Pixels with at least 85% of their area within the management unit (30% in the case of soil moisture due to the coarse resolution) were considered in establishing the aggregate value for that unit. For the 1 km resolution LST, NDVI and ET remote sensing products, the results are also analysed by land cover type. This relates to pixels where at least 85% of the area consists of irrigated and rainfed agriculture land cover classes (see Figure 1).

2.2.2 In situ data

In situ data of reservoir levels and inflow and river flow from the basin measurement network were used to calculate the Status Index (I_e), a normalised monthly index used by CHE to homogenise the different indicators (CHJ, 2007):

$$\text{If } V_i \geq V_{avg} \rightarrow I_e = \frac{1}{2} \left[1 + \frac{V_i - V_{avg}}{V_{max} - V_{avg}} \right]$$

$$\text{If } V_i < V_{avg} \rightarrow I_e = \frac{V_i - V_{min}}{2(V_{avg} - V_{min})}$$

where V_i is the value of the indicator for month i and V_{avg} , V_{max} and V_{min} are respectively the average, maximum and minimum values of the indicator derived from historical data. Based on this Status Index (which is a value between 0 and 1), the situation under analysis is classified by the authority as normal ($I_e > 0.5$), pre-alert ($0.5 > I_e > 0.3$), alert ($0.3 > I_e > 0.15$) or emergency ($I_e < 0.15$).

The indicators selected by CHE for each of the management areas were used for the analysis presented here. These are the values of reservoir volume for the regulated areas (122, 123, 131, 132, 141, 151), inflow into the corresponding reservoir(s) for the upstream areas (120, 140, 150), and runoff at a selected station for management area 130.

2.2.3 Benchmarking data sets

Two different tests were carried out using drought impact data sets as benchmark to assess the ability of remote sensing based indicators to provide early drought detection information during the period 2000-2012. The short length of the remote sensing 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. In the first test, text-based records of drought occurrence and impacts collected from a review of local news (i.e. qualitative information) were used to reconstruct the onset and evolution of drought conditions during the period of analysis and as benchmark for the comparison of the remote sensing datasets. Newspaper records were selected as data source because it allowed a systematic collection of impact occurrence data of all affected sectors with a monthly time step for the whole period of analysis. (i.e. qualitative information). In the second test, the use of crop yield statistics (i.e. quantitative information) as a benchmark of drought impact on agriculture is considered as a benchmark of drought impact on agriculture. explored (i.e. quantitative information). The correlation of remote sensing data, especially SPI and NDVI, to agriculture yield data has been widely researched and applied (see Bachmair et al., 2016, for a review). This second type of impact data was included 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 and discuss the advantages and limitations of one with respect to the other.

Text-based data sets were collected from a review of regional news. “El periódico de Aragón”, the second largest newspaper in average daily circulation in the Aragón region was selected for the review because it has an online record going back to September 2001, was selected for the review. All news items containing the word “drought” were reviewed and relevant records of

drought events and impacts referring to the area of study were tabulated. For each entry the location, period, description, and, in case of reported impacts, the affected sector were noted. The affected sectors were labelled as *rainfed agriculturecereals*, *irrigated agriculture*, *livestock*, *water quality*, *fire*, *water supply*, *energy* and *others*. The records of drought occurrence are classified according to the source of the information, making a distinction between non-official sources such as journalists and water users, labelled "mention of drought occurrence" in Figure 2, and official sources labelled as "drought acknowledged by the authorities", "ongoing mitigation measures" and "periods retrospectively defined as anomalously dry". A distinction in the analysis is made between the mention of drought, the acknowledgement of drought by the authorities, the reported mitigation measures, and news items retrospectively reporting on droughts. This last type corresponds mainly to news about the publication or communication of analysis performed by the scientific community or the water managers describing an ongoing or past drought.

The limit between indicators and impacts is not always clear. For example, low flow or reservoir levels are considered an impact of meteorological drought in some analyses, while these serve as indicators of hydrological drought in others. Here we limit the definition of drought impacts as the effects of drought on people, economy and/or the environment.

Crop yield data of winter cereals both for irrigated and rainfed cropping systems was obtained for the five selected districts in Huesca (H, S, L, M and B). Winter cereals are the cereal crops that are planted in the autumn and they are the crops that cover the largest surface area. Their importance for the region results in better data availability than other crops and for this reason this type of crops were selected for the analysis. Only winter cereal crops with larger cultivated areas were ~~considered~~selected: two and six-row barley (irrigated and rainfed), wheat (irrigated and rainfed), and rice (irrigated). The two and six-row barley types refer to the number of fertile spikelets in the spike.

2.3. Correlation between remote sensing data and the benchmark data sets

The correlation between each of the remote sensing ~~parameters~~ time-series-withand both the timelines that aggregates all types of drought events records and the timeline that aggregates all types of drought impacts (Figure 2)-of drought occurrence and impacts was analysed in terms of strength of the relationship and anticipation. The strength of the relationship is related to a function of the predictability of the occurrence of drought and drought impacts provided by the remote sensing time series.

Anticipation reflects the ability of the remote sensing data sets in providing early information and gaining time to undertake actions. The aim of this analysis 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. The benchmark datasets~~These two characteristics of the relationship were examined using a lagged correlation analysis, i.e. establishing the correlation where the remote sensing data set is~~

shifted in time with respect to the benchmarking data set. In this way the impact occurrence was were compared to the variables represented by the remote sensing time series in the 24 preceding and following months. 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.

The sample cross correlation function (CCF), $I_{x,y}$, was used for the analysis. The CCF cross-correlation function ($\rho_{x,y}$) can be expressed as (Chatfield, 2004):

$$c\gamma_{x,y}(\tau) = cov(X_t, Y_{t+\tau}) \quad (\text{Eq. 1})$$

$$r\rho_{xy}(\tau) = \frac{c\gamma_{x,y}(\tau)}{\sigma_x \sigma_y} \quad (\text{Eq. 2})$$

For $\tau = \pm 1, \pm 2, \dots$, where τ is the lag and σ_x and σ_y are the standard deviations of the time series x_t and y_t . The set of $c\gamma_{x,y}$ coefficients corresponds to the cross-covariance function. The CCF function as implemented in R (R Core Team, 2016) was used for the calculations.

To detect possible issues related to the stationarity or ergodicity of the series, their time autocorrelation and partial autocorrelation plots were considered.

The reference drought periods used for the correlation provide a binary record; indicating the occurrence or non-occurrence of drought events in each month, without quantifying their intensity. To obtain an insight into the severity of the events, the use of annual crop yield data was explored. The correlation of each annual crop yield value to the monthly values of the remote sensing time series from the start of the hydrological year in September to the end of the following calendar year was analysed. This was done to detect the key months in which the occurrence of drought conditions lead to impacts on the (annual) crop yield. The comparison was performed for three rainfed areas and three irrigated areas. These were selected to correspond with the management units so that a relation could be established with the areas of influence of the reservoirs (Figure 1). The areas selected included the rainfed agriculture areas of Hoya Huesca (HH0, corresponding to management unit 140), Monegros-Bajo Cinca (MBO, corresponding to management unit 141), and in the five districts together (AA0) and the irrigated agriculture in Hoya Huesca-Monegros (HM1, corresponding to management unit 141), La Litera-Bajo Cinca (LB1, corresponding to management unit 131), and in the five districts together (AA1).

3. Results

3.1. Drought events and impacts

The timelines of drought events and impacts derived from the review of local news are illustrated in Figure 2. Three drought events can be distinguished in the Ebro basin: a short drought event at the beginning of 2002, a multi-year drought from the end of 2004 to the spring of 2008 and a shorter duration drought during 2011 and 2012.

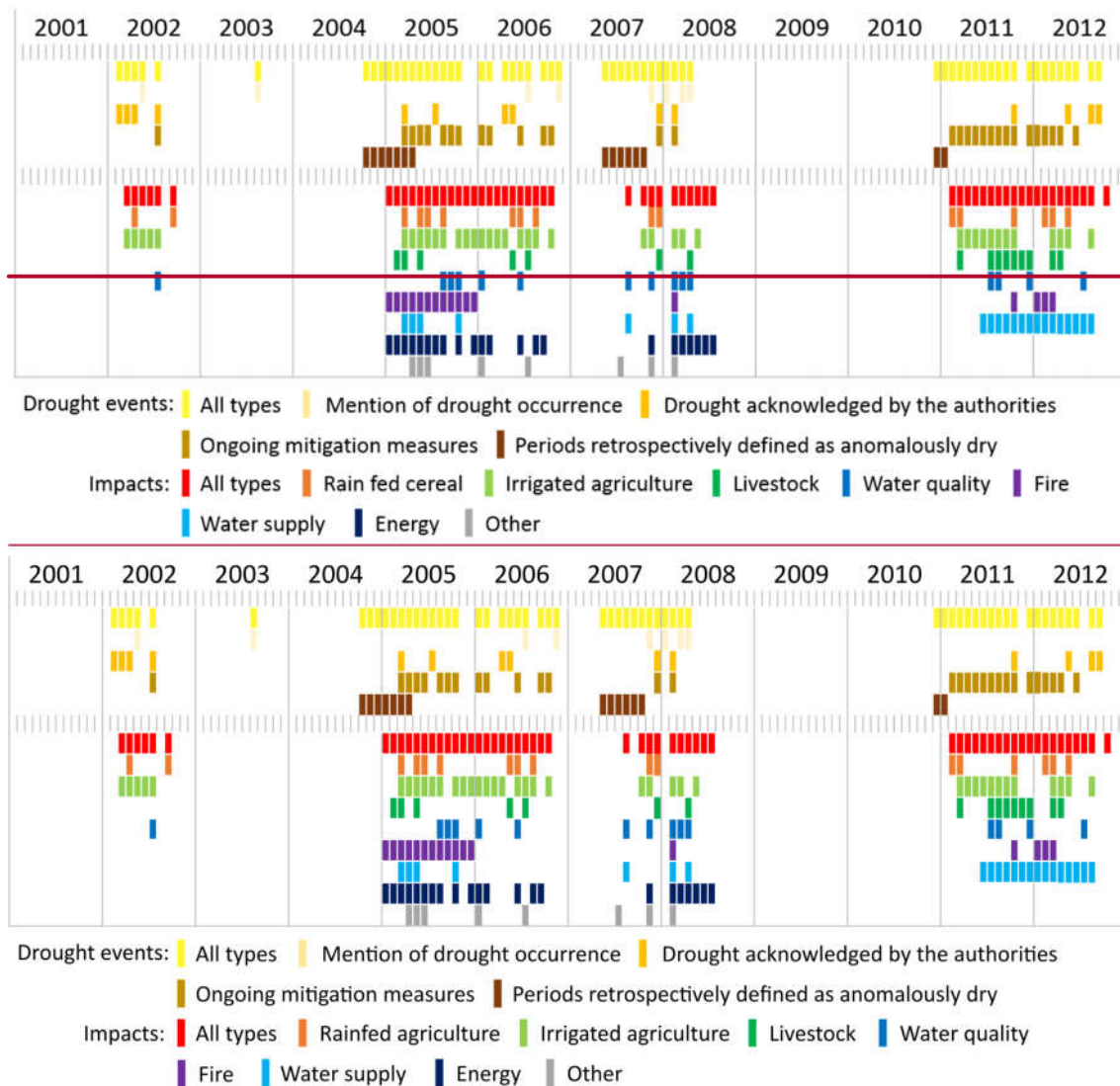


Figure 2. Timeline of drought events (upper part) and impacts (lower part) for the north-eastern Ebro basin during the period 2001 to 2012 based on a comprehensive local newspaper review.

The first colored row (yellow) in the figure represents the months in which drought was taking place according to the records found in the newspapers. ~~Four types of record can be distinguished for this first row, represented by the following four rows: (1) general mention of drought conditions, (2) mention of the acknowledgment of drought by authorities (national, local government or water authority), (3) mention of ongoing drought mitigation measures, and (4) mention of earlier periods of rainfall deficit in the newspaper but described in retrospect. This last type corresponds mainly to news about the publication or communication of analysis performed by the scientific community or the water managers describing an ongoing or past drought.~~ The first line of the second block (red) reflects the occurrence of drought impacts described in the newspaper, while in the following rows these impacts are disaggregated by the affected sector.

Based on the records gathered from the newspaper records, the following descriptions of the hydrological years affected by drought episodes were constructed:

In 2002, after a dry winter the availability of water in the reservoirs was low. A first reference to drought in the press appeared in February 2002. At the start of the spring, which is the beginning of the irrigation season, water curtailments were reported for Bardenas irrigation system. In the beginning of April, Agricultural Associations reported losses of 20% of rainfed cereal crops in Aragon and at the end of the month the impact of drought in the area was acknowledged by the Ministry as well as by the local government. In July, the flow of the Ebro in Zaragoza was half of the minimum 30 m³/s that has been set to warrant water quality. In September a reduction of 40–70% in olive oil production in the areas of Bajo Cinca, Medio Cinca and La Litera was reported in the news. General mention of impacts on pastures, hydroelectricity production and employment in the primary sector during that drought period appeared in retrospect, but these reports did not go into detail.

The hydrological year 2004–2005 was depicted as the driest on record. The combination of cold and dry conditions during the first part of 2005 produced significant losses in the agriculture and livestock sectors. First impacts were reported in February 2005 (lack of pastures production after five months without rain). From then, and until September 2006, the newspaper reflected a succession of impacts in different sectors, including all crop types, pastures, forests, livestock production, water supply to the population, wildlife, economy, recreational activities, hydroelectricity, water quality, employment and politics. The drought was already acknowledged by the authorities since March 2005, and the first mitigation measures were announced shortly after. This was that the regional government increased to 50% the area of land, rainfed or irrigated, that could be set aside to remain fallow. In June, aid measures were approved by royal decree.

Reservoir levels increased during the first half of the hydrological year 2005–2006, but the system failed to recover completely from drought before levels started decreasing again in April 2006, and at the beginning of the summer levels were lower than the previous year. After a hot summer, storage started to recover again and in December 2006 the government considered the drought as having ended. Intense rains starting in February 2007 were followed by a period of precipitation deficit from May to February 2008. A few problems of water supply to certain villages were reported in August 2007 and flows were below the minimum required to warrant water quality in October. Impacts on agriculture and hydroelectricity started to be reported again in October. Abundant rains during spring 2008 constituted a first step towards the end of the drought episode.

The hydrological year 2010–2011 was characterised by lower than average precipitation and high temperatures. In February 2011 the newspaper showed the first reference to an emerging drought and its impact on the sprouting of winter cereal. This drought especially affected the Bardenas irrigation district. The Riegos del Alto Aragón and Canal de Aragón y Cataluña districts were also affected. All the systems managed to reach the end of the irrigation season, but with restrictions of more than 60% on water quotas. Grapes and olives were the most damaged crops, but in general the food production in the area was defined as satisfactory at the end of the season. The following hydrological year (2011–2012) started with low reserves and a dry winter and spring, with the exception of November, which was a particularly wet month. In particular, the middle sector of Huesca revealed drought affected areas. Extensive livestock farming, fodder and cereal production were the most impacted sectors. The risk of fire was reported to be high, even during the winter, which translated in a higher number of fires.

3.2 Correlation of text-based records and remote sensing indicators

The information on drought occurrence and impacts obtained in the previous step was used as a benchmark data set to assess the ability of the remote sensing based data sets to provide early detection. Figures 3 and 4 present the results of the cross correlation of the remote sensing data sets to the timelines of drought events (i.e. upper records in Figure 2) and impacts (i.e. all other records in Figure 2) respectively. The x-axis expresses the time lag in months between the remote sensing data sets and the timeline of drought events and impacts (τ in equations 3 and 4). The central line ($x=0$) corresponds to the correlation of the two data sets in the same month. Negative values of x refer to correlations between impact time series at time t and remote sensing values at each of the 24 months before t ($\tau=-1, \tau=-2, \dots, \tau=-24$). Strong correlations on the left side of the central line reflect the ability of the data set to anticipate the occurrence of drought events and impacts. The positive side of the plots reflects the correlation of the drought occurrence and impact series with the values of the different datasets in later months. This type of correlation appears if the conditions that define the start of the event or impact occurrence last longer than one month.

Figures 3 and 4 have similar correlation patterns, with the second showing higher anticipation. This result was expected because Figure 3 is based on records reflecting climatic and hydrologic anomalies and deficits, and these processes precede the impacts.

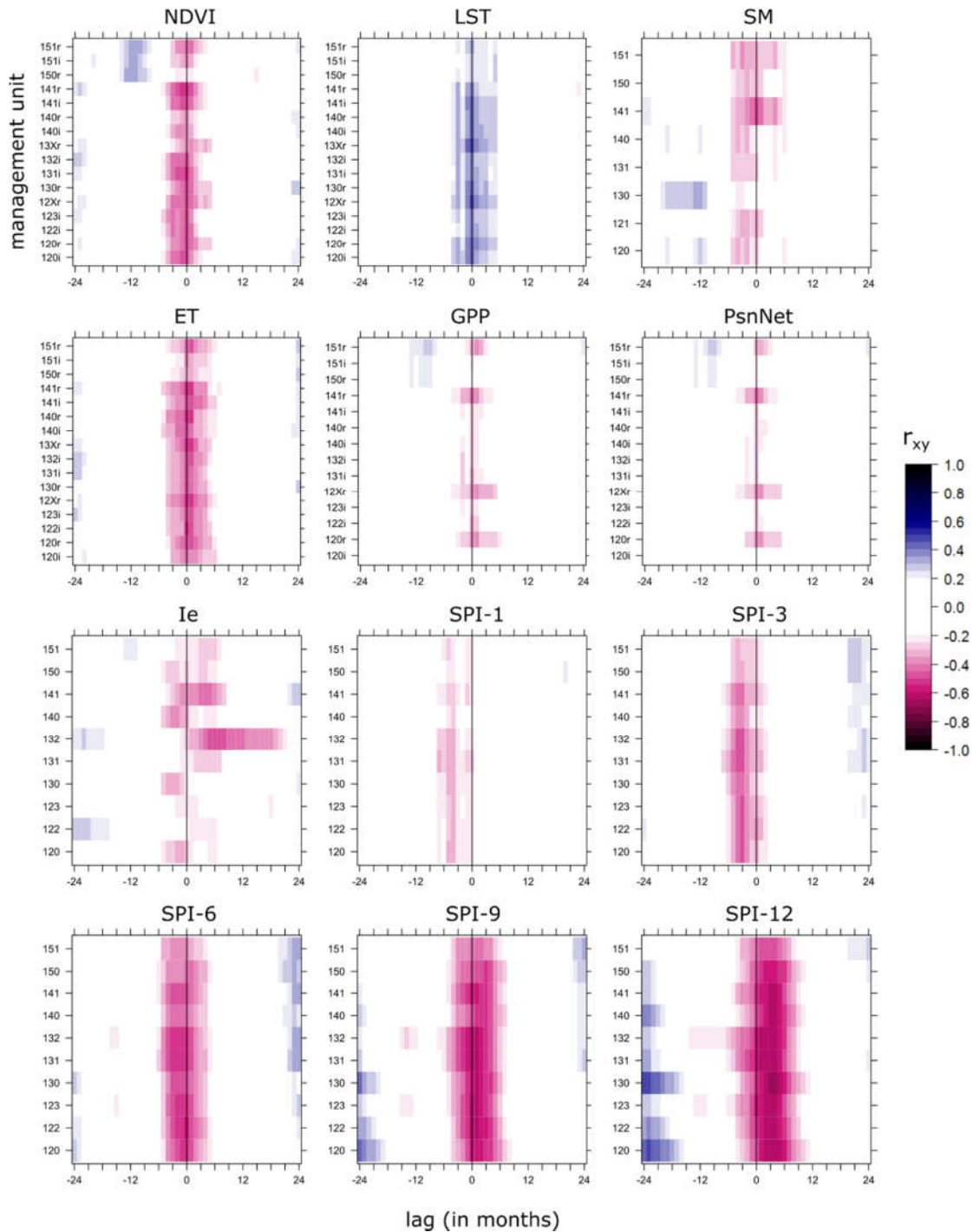
SPI shows the strongest correlations for both events and impacts. For SPI values calculated for longer aggregation periods the correlation grows stronger, while the anticipation is slightly reduced. The best correlation-anticipation relationship is obtained for SPI-6 and for SPI-9. For these indicators, the correlation is also stronger in the southern areas. This is probably because most of the socio-economic activities are concentrated in these managed areas, and therefore the impacts and media attention are likely to be higher. The results show that SPI-6 and SPI-9 are most suitable for predicting impacts, together with NDVI and ET; achieving an anticipation of 6 months with a sufficient correlation ($r^2 > 0.6$). This provides useful information for activating drought mitigation measures.

Soil moisture also shows good anticipation, albeit with weaker correlations. This may be due to the coarser spatial resolution of the data set. Higher resolution soil moisture products (e.g. Alexandridis et al., 2016; Scott et al., 2003) could be considered for future studies. NDVI and ET data sets show a strong negative correlation with drought occurrence and impacts, which would be expected from a biophysical perspective. NDVI shows better anticipation, preceding the impacts in most of the units by more than 6 months. ET shows a slightly stronger correlation in the rainfed areas, while no distinction is seen between rainfed and irrigated areas for LST and NDVI. LST has a positive correlation because evaporative cooling is diminished during drought events which prompts the land surface temperature to rise. LST correlation to events is stronger than to impacts, but the degree of anticipation is lower for the former.

Indices derived from both GPP and PsnNet present weak or no correlation for most of the areas, with only some of the rainfed areas showing moderate positive correlations.

In situ indicators based on water levels and flows show varied level of anticipation for the different areas. Most of them provide early information on drought occurrence and impacts from 6 to 9 months in advance, but there are two areas where the indicator offers no anticipation (management

area 140) or even no correlation with the benchmark data sets (management area 123). ~~The reason for these exceptions is unclear. Indices derived from both GPP and PsnNet present weak or no correlation for most of the areas, with only some of the rainfed areas showing moderate positive correlations. The reason for this may lie in the formulation of the MOD17 product. Indeed, limitations of the product in capturing spatial and temporal variability in croplands have been reported (Verma et al., 2014; Zhang et al., 2012).~~



Indicators vs. drought events

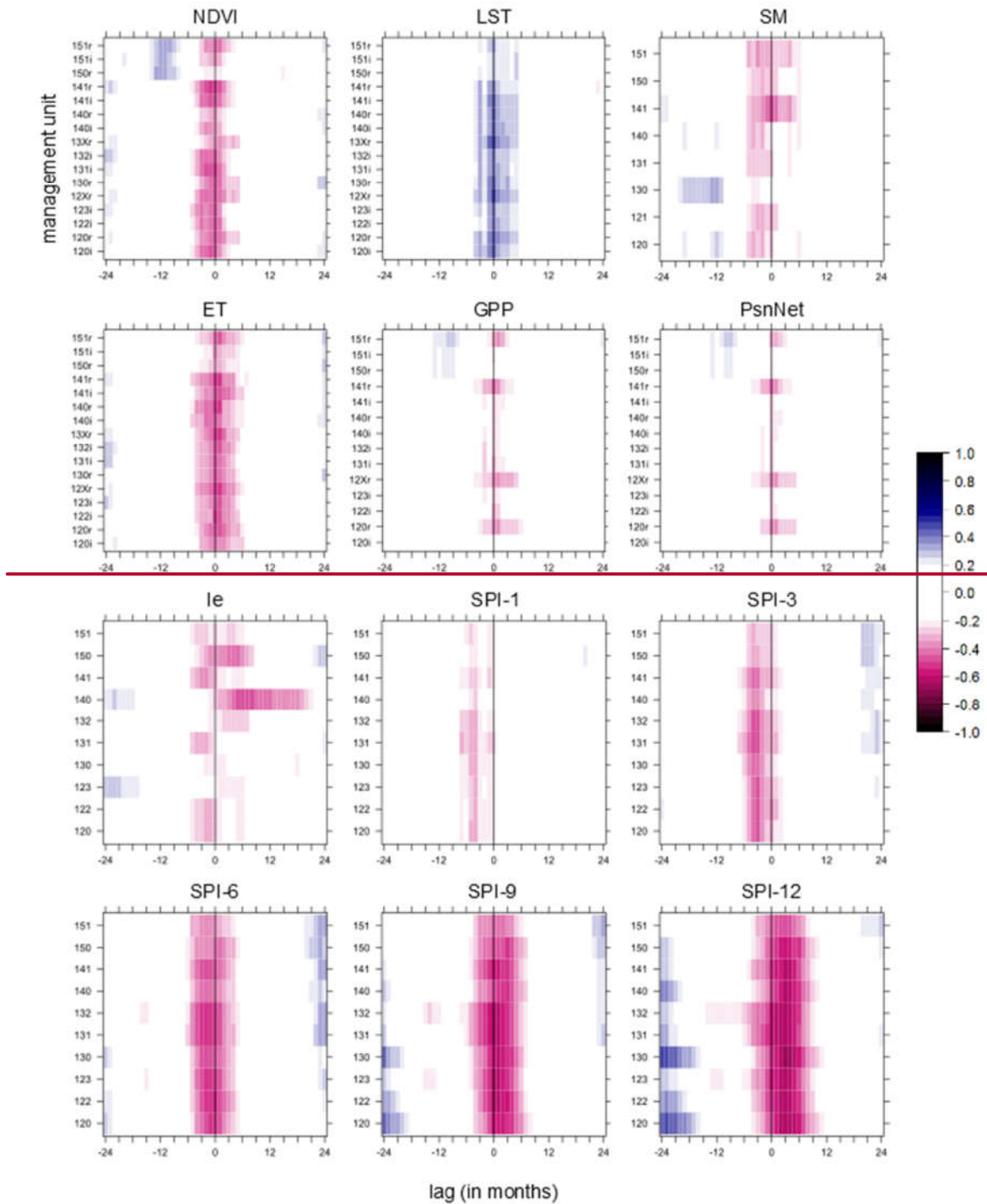


Figure 3 Cross correlation of drought indicators and drought events at multiple time lags. The numbers in the y-axis represent the management areas depicted in Fig. 1. For NDVI and LST, irrigated (i) and rainfed (r) crops within the areas are distinguished. The x-axis represents the shift in months between the two data sets. The indicator built from in situ data (I_e) is also included.

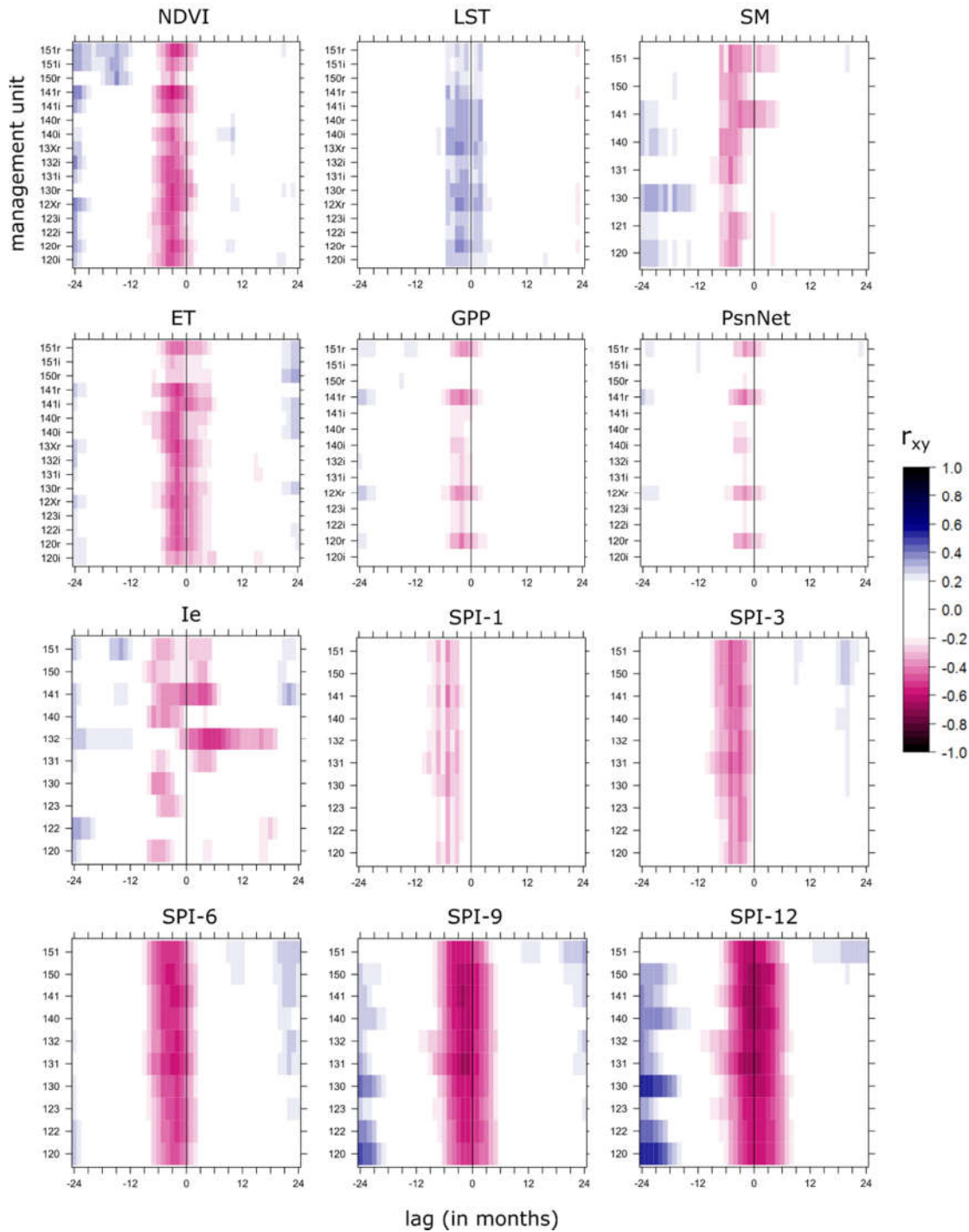


Figure 4. Cross correlation of remote sensing data sets against the timeline of reported drought impacts (all types).

The time plots obtained for each of the parameters present no trends or discontinuities and the values in the autocorrelation plots show that the autocorrelation diminishes quickly with increasing lag. An exception are the series of the reservoir indices. In that case, for some of the series it is not clear from the plot if the series is stationary. For one of them (management unit 122) it clearly is not. This management unit corresponds to a reservoir (Rialb) that started to be filled in the year 2000 and therefore the levels cannot be considered stationary for the period of study. 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). In the I_e plots in Figures 3 and 4 it can be clearly seen that the two management units that do not satisfy the conditions for stationarity (management units 122 and 132) are those (at least two out of the three) that do not present anticipation. 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.

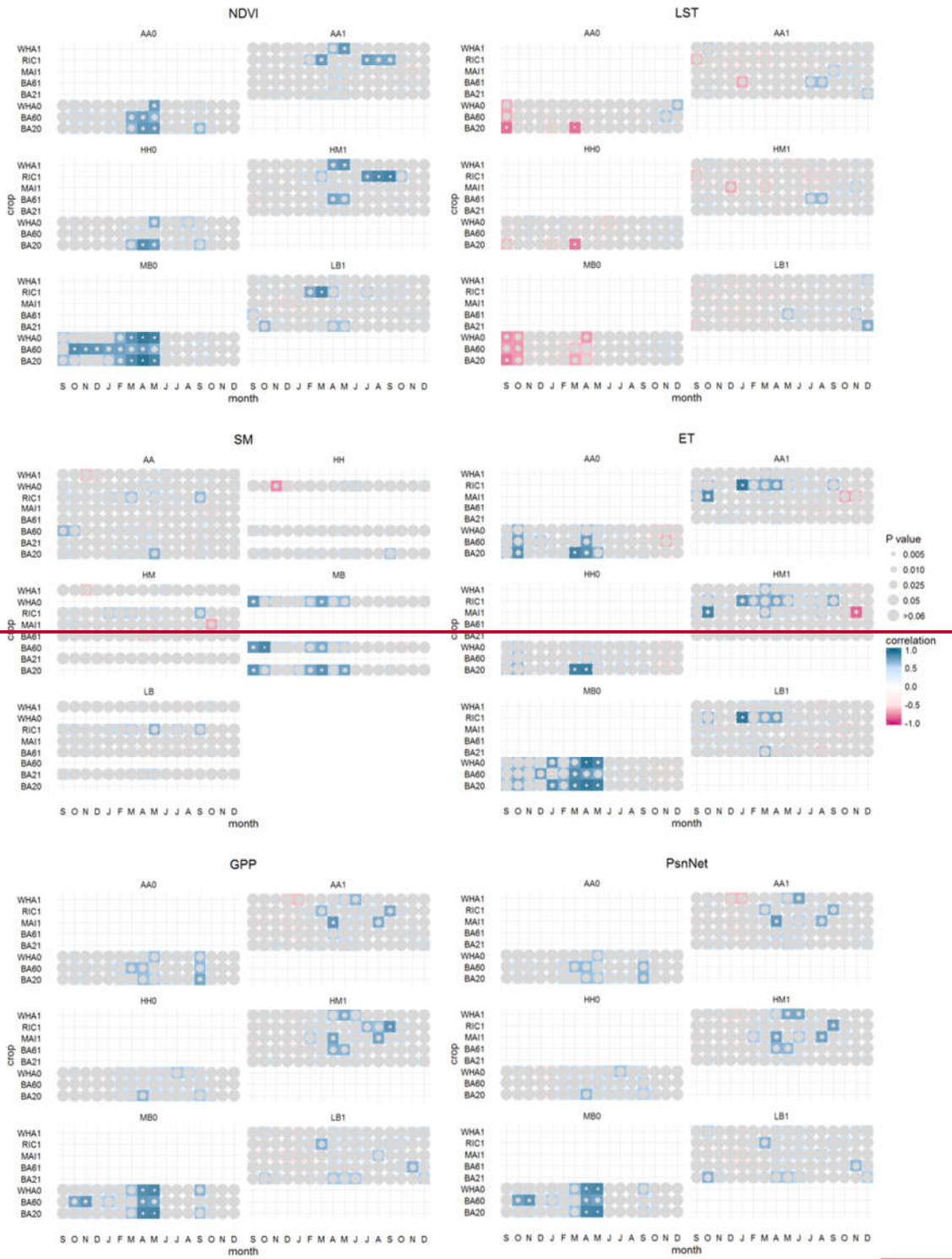
The positive correlations of the timelines of drought occurrence and impacts with the values of the indicator datasets with lags over one year (most notably at -15 and ± 24 lags) are casual correlations. Since the analysis presented here focuses on anticipations within a period of one hydrological year, the correlations should not be affected by this issue.

3.3 Correlation of crop yield and remote sensing indicators

The results of the correlation analysis between the remote sensing data time series and the annual crop yield for the main irrigated and rainfed cereal crop types in the selected districts in Huesca are represented in Figure 5 and Figure 6. Every parameter is tested for the six areas. The crop types are represented on the y-axis and the months on the x-axis. The latter spans from the start of the agricultural year in September to the end of the following calendar year. The colour gradient reflects the sign and strength of the correlation, while the size of the inner grey circle corresponds to the reliability of the correlation.

NDVI and ET present some of the strongest positive correlations, especially between the remote sensing measurement during the spring (MAM) and the yield of rainfed crops. ~~This is in agreement with the results obtained by Vicente Serrano et al. (2006) who observed a higher correlation between barley crop yield and NDVI for the month of March at a location in the Ebro River valley.~~ LST shows also strong correlations with rainfed crops in March and at the beginning of the season in September (S). The pattern is less clear for irrigated crops, probably because their water supply is less dependent of the rainfall. The strongest correlations in this case appear for rice crops with ET and NDVI, mainly at the start of the year.

Despite irrigated crops directly depending on reservoir supply, only rice shows significant positive correlations with the index based on reservoir levels for the two irrigated areas tested (HM, corresponding to management area 141 and LB corresponding to management area 131). The reason can be that rice is especially drought sensitive, since it has shallow roots and therefore a low depth of readily available soil water, which is the fraction of total available soil water that crops can obtain from the root zone without experiencing water stress. This fraction is 0.2 for rice (Allen et al., 1998). The rest of the tested crops have a higher value of this fraction and they experience stress when more moisture is depleted. The rainfed crops in the HH area (corresponding to management area 140) show correlation with the Status Index based on reservoir inflow, in April.



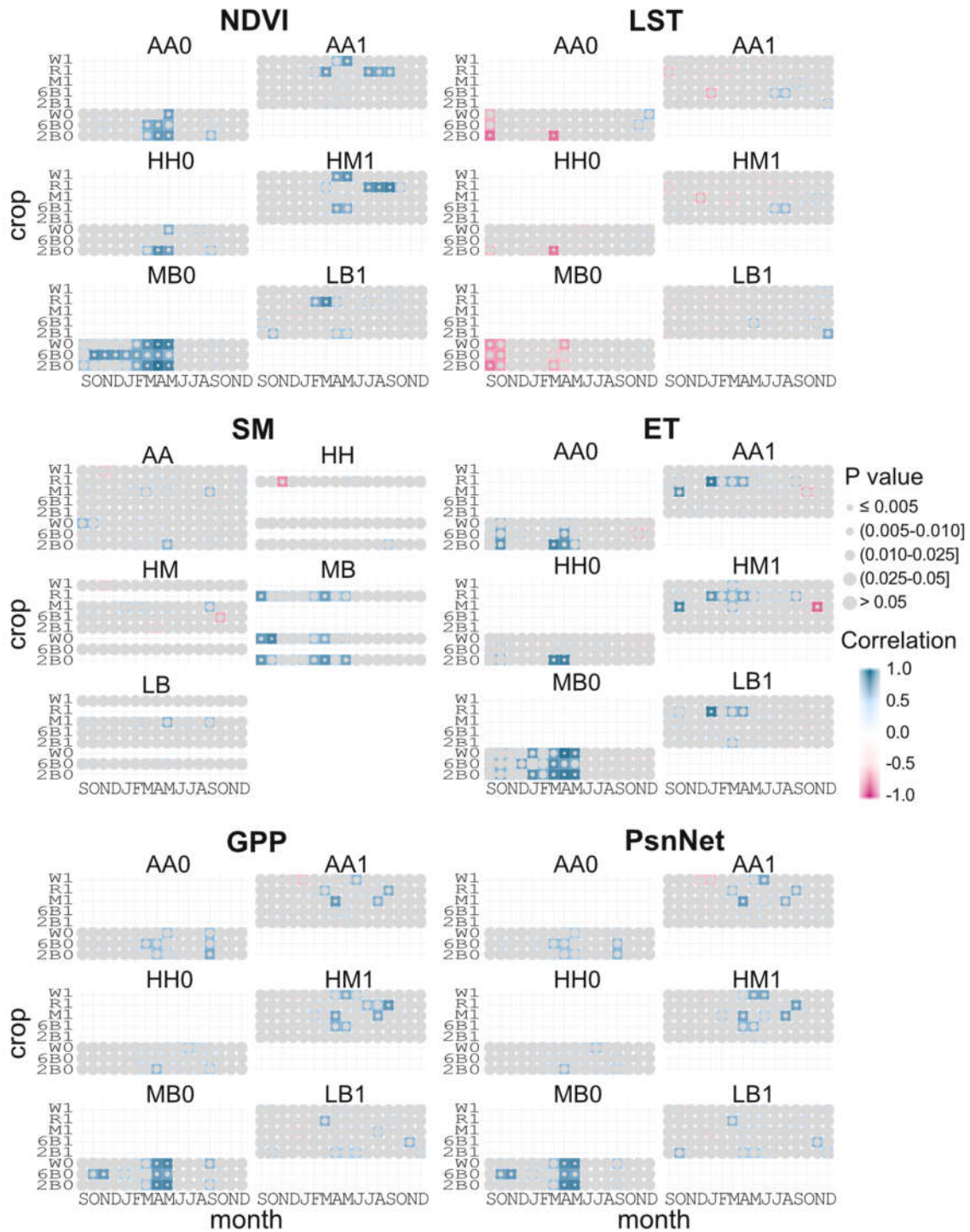
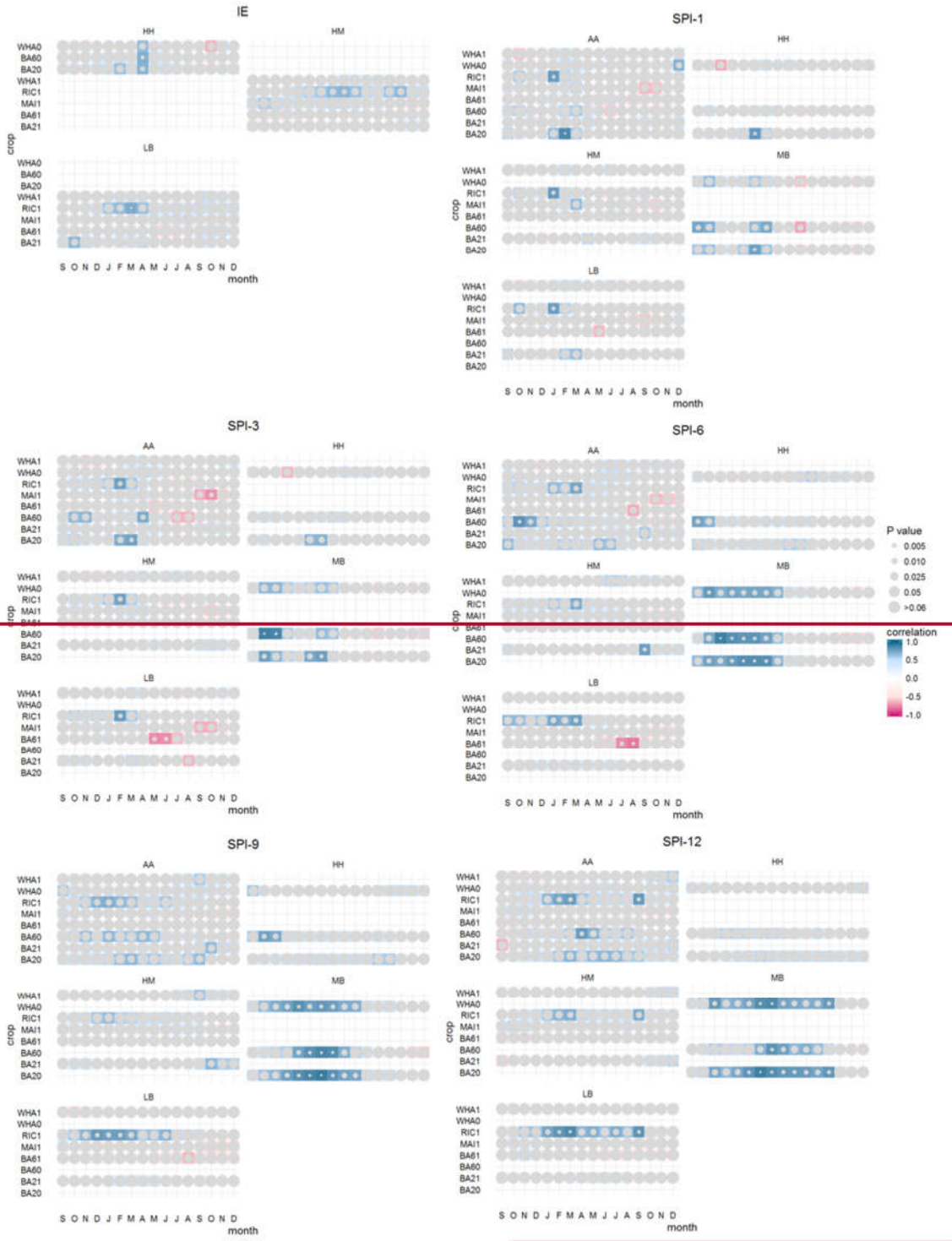


Figure 5. Correlation between remote sensing drought indicators and crop yield data. Rainfed areas and crops are marked with 0 and irrigated areas and crops with 1. The crops are irrigated and rain-fed wheat (W1 and W0), irrigated rice (R1), irrigated maize (M1), irrigated and rain-fed six-row barley (6B1 and 6B0), and irrigated and rain-fed two-row barley (2B1 and 2B0). The colour gradient reflects the sign and strength of the correlation, while the size of the inner grey circle corresponds to the significance of the correlation.



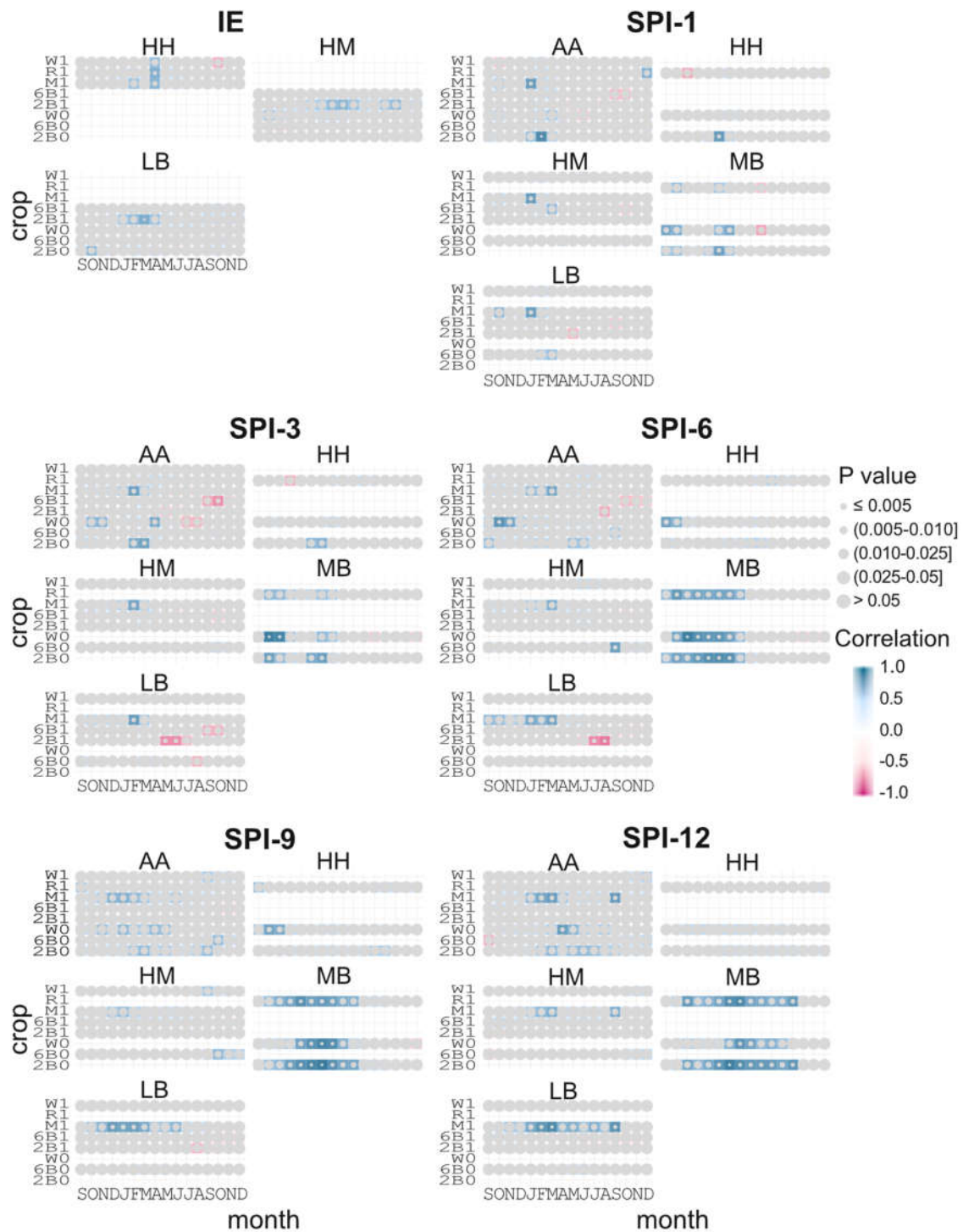


Figure 6. Correlation between drought indices (SPI and Reservoir index) and crop yield. Rainfed areas and crops are marked with 0 and irrigated areas and crops with 1. The crops types are irrigated and rain-fed wheat (W1 and W0), irrigated rice (R1), irrigated maize (M1), irrigated and rain-fed six-row barley (6B1 and 6B0), and irrigated and rain-fed two-row barley (2B1 and 2B0). The colour gradient reflects the sign and strength of the correlation, while the size of the inner grey circle corresponds to the significance of the correlation.

Soil moisture, GPP and PsnNet do not show a clear pattern against reported crop yield, except for a strong positive correlation in one of the areas (MB). These correlations appear in the spring (especially for GPP and PsnNet) and at the beginning of the hydrological year.

SPI has positive correlations at the start of the season, which are particularly strong in the MB area. In some cases also a negative correlation during the summer emerges, especially for shorter term SPIs. It can also be observed that the stronger correlations appear later with longer term SPIs.

Rainfed two-row barley (2B0) in March stands out as the crop with the stronger overall correlation with the different indicators. 2B0 is one of the major crops in the area, with a maximum cultivated surface for the period 2000–2012 of 170,914 ha of the total of 204,614 dedicated to herbaceous crops (in 2008), and a minimum cultivated surface of 130,764 ha (2012). Maize is the second commonest crop with an average crop surface of 41,292 ha during the period. [Figure 7](#) illustrates the correlation of two-row barley to each of the indicators for the rainfed crops in the Monegros-Bajo Cinca area (MB0), which has been selected as an example. This shows the crop yield for the different years against the value of each respective indicator.

Three years stand out in [Figure 7](#) for having extreme low indicator values (high in the case of LST) for all variables: 2005, 2008 and 2012. Values are also low for year 2000 for the data sets for which it is available. These three years correspond with hydrological years of reported impact in the area identified in the previous section.

The lowest crop yields were obtained in 2012 (1342.5 kg/ha). This is in agreement with the information in newspapers reporting that drought that year especially affected cereal production in the middle sector of Huesca, the area of focus of this test. Accordingly the remote sensing parameters present some of the lowest (highest in the case of LST) values for the period. Only SPI-6 presents a value that is well above the minimum. This is caused by November 2011 being a particularly wet month in the middle of the drought period, thus moderating the value of SPI-6. For longer term SPI values this positive anomaly is compensated by the negative anomalies of the rest of the months.

Crop yields were very similar in 2005 and 2008 (1662.3 and 1800.1 kg/ha respectively) and so is the behaviour of most of the variables. The main differences appear in LST, with the 2005 LST for March being more than 2 degrees higher than in 2008 (23.8 and 21.3 degrees respectively), and SPI-3, which is less extreme in 2008 (-0.68 compared to -1.51 in 2005). The reason for this difference is the earlier start of spring rains in 2008. Both hydrological years start with an exceptionally dry period that extends to April in 2005 and to March in 2008 after which the spring rains improve the situation.

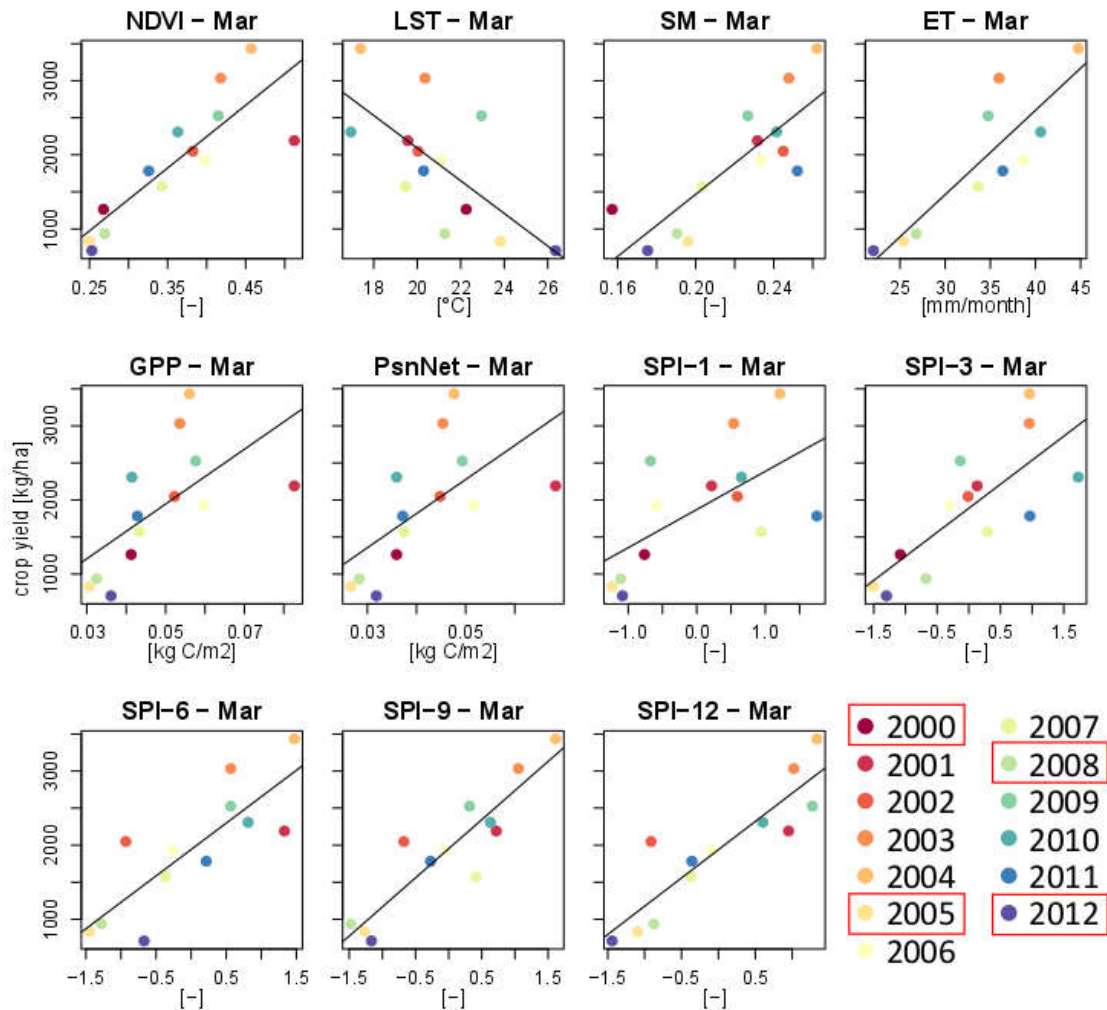


Figure 7. Correlation of the remote sensing drought indicators for the month of March to annual rainfed two-row barley yield in Monegros-Bajo Cinca districts (MB). For LST, NDVI, SM, ET, GPP and PsnNet monthly means were used.

There is a second group in the middle sector of the plots that includes the rest of the years for which drought impacts on rainfed agriculture were reported in the analysed media, 2011 (3551 kg/ha), 2006 (3857 kg/ha) and 2002 (4249 kg/ha), together with the hydrological year 2006–2007 (3115 kg/ha), for which no impact was reported in the regional press. ~~Local press for the area of 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.~~

March values of SPI-3 for these years are close to the mean and only 2002 presents strong negative anomalies for SPI-6. In 2006 and 2007 the precipitation deficits start in April and May respectively and for 2006 the impacts are reported only after that month. Hydrological year 2001–2002 shows dryer autumn-winter conditions according to SPI-6.

The results of this second tests present a consistent behaviour of the indicators with respect to the different levels of crop yield among the analysed years in rainfed areas. As in the previous test, NDVI, ET and SPI stand out for having stronger correlations. Most indicators present ~~very~~ similar March values for the years of severe drought, clearly differentiated from the behaviour of years of

moderate drought and years of no drought. The only exception is LST, in which a year where drought was not reported and yields were normal, such as 2009, has similar LST values in March than the years of severe drought. This indicates that LST may not be a good indicator of drought on its own, but can still be useful in combination with other indicators.

4. Discussion

4.1. Use of impact records as drought reference

The review of text-based records allowed a detailed reconstruction of the drought events during the period studied. The cross-correlation of the timelines of drought events derived from this review to the set of indices derived from remote sensing data sets revealed the potential of the latter to provide early detection of drought events. However, this binary information has the limitation that it does not allow to objectively quantify the severity of the events. For example, in the case of rainfed agriculture, the information on impacts collected from the newspaper does not allow for differentiation between those years in which production was extremely low as a consequence of drought conditions and those years in which production was only partially affected by drought. Other studies have also suggested a link between impact severity and the number of records reporting them (Bachmair et al., 2015; Hernández Varela et al., 2003), but this needs to be taken carefully since the media coverage of a drought event is also highly influenced by the socio-political context in the affected area (Llasat et al., 2009; Sonnett et al., 2006).

A few additional aspects concerning reliability were noticed while processing the records from the press:

- Accuracy: The information on drought occurrence reported in the newspaper may not be accurate. For example impacts due to other causes may be attributed to drought, or other phenomena such as normal summer shortages may be described as drought. 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: Reporting of drought occurrence in the newspaper is not systematic and therefore some impacts may be missing. In Figure 2 some unlikely situations can be identified. For example impacts to livestock in May and July 2006, but not in June. Records referring to specific types of impacts are more likely to have gaps. However, when all types are aggregated, part of the gaps in each of the disaggregated data sets will likely be filled with records from the other data sets.

- Scale: drought events affecting only a small area within the region covered by the newspaper may not be reported. 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: Public or political interest or concern about drought (or even scarcity of other relevant news) can motivate overstatement of drought impacts. 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.

The length of the period of analysis does not 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 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 Catalonia, it was still low during that year (MMA, 2007).

Crop yield data, on the other hand, allowed for a more objective identification of the drought events that had higher impact on agriculture, though the yield data does have the disadvantage that may only be reported on an annual basis. March was the month that presented higher correlations.-This is in agreement with the results obtained by Vicente-Serrano et al. (2006) who observed a higher correlation between barley crop yield and NDVI for the month of March at a location in the Ebro River valley.- The examination of the behaviour of the remote sensing parameters in the years with similar yield data provided an insight on the reliability of the parameter as an index. Similar values of the parameter for years of similar final crop yield indicate the robustness of the indicator. For the period of analysis, the occurrence of drought in the years in which low yields were obtained is confirmed by the media records, but despite that the water availability is a determinant factor for rainfed winter cereal yield, other factors such as frost, floods, plagues and diseases could have further reduced the annual yield. However, for March values, there are no anomalies that suggest these factors had a strong influence in the low annual yield values.

Crop yield data can also be a useful reference to identify thresholds of drought severity classes. These thresholds could be derived based on the differences observed between the groups of years

with severe, moderate and no drought conditions, although a longer data series than was used in this study is recommended to provide a more robust estimate of threshold values.

There are several factors that play a role in the severity of the impacts due to drought conditions, including coping capacities and water management (e.g. drought may not lead to impacts in irrigated areas). Variations in these factors can alter the relationship between the indicators and the impact. It should also be noted when using drought impacts as a benchmark of drought occurrence, that the absence of certain types of impacts as a result of sound drought management does not imply that there is no drought (Smakhtin & Schipper, 2008), though even with perfect management there will always be some kind of impact. For example, a reduction of income as a consequence of substituting the usual crops by less productive alternatives with lower water requirements constitutes a clear impact, even if the yield in kg/ha is not affected. The influence of management is probably also the reason for irrigated land showing less clear correlation patterns than drought in rainfed areas in both analyses. A wider view that considers as many different types of impacts and affected sectors as possible can help overcome the effect of management when using this type of data as a benchmark of drought occurrence. Initiatives such as the US Drought Impact Reporter and The European Drought Impact report Inventory can play a useful role in providing that broader view.

4.2. Early drought detection with remote sensing products

Early information on emerging droughts benefits mitigation strategies by increasing the time available for managers and affected communities to take action. The requirements for drought early warning range from a few weeks to several months (UNISDR, 2009). The results show the potential of the tested products to anticipate up to some 6 months reported drought impacts at basin scale. SPI, NDVI and ET products stood out in both analyses as particularly suitable data sets to detect early stages of drought at the basin scale and anticipate drought impacts. However, while for most products the autocorrelation dissipates at a lag of 2 months, for NDVI and SPI-6 it takes 3-4 months and this can have an influence in NDVI and SPI-6 showing stronger correlations. For SPIs with longer accumulation periods (SPI-9 and 12) the correlation dissipates even slower.

The weaker correlations obtained for SM data in the first test This may be due to the coarser spatial resolution of the data set. Higher resolution soil moisture products (e.g. Alexandridis et al., 2016; Scott et al., 2003) could be considered for future studies. The reason for the weak or no correlations between both GPP and PsnNet and the text-based records may lie The reason for this may lie in the formulation of the MOD17 product. Indeed, limitations of the product in capturing spatial and temporal variability in croplands have been reported (Verma et al., 2014; Zhang et al., 2012).

The trade-off between the anticipation of the information and its reliability is also illustrated by the results. The lower reliability associated with earlier information detection of conditions that may lead to drought implies that often the situation may not evolve into a drought event. However, that information is still highly valuable as it allows the stakeholders to get ready to undertake mitigation actions if necessary.

The remote sensing products tested can enhance early warning capacity and therefore contribute to the shift from reactive to proactive management recommended by the European Commission (Commission of the European Communities, 2007) and the United Nations (UNISDR, 2009) and being undertaken by many institutions (Iglesias et al., 2009). As remote sensing data products generally

have a global coverage, this contribution would therefore be especially useful in areas with less in situ data available. Yet 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 (Stagge et al., 2015).

Remote sensing products also have the potential to provide information at a finer spatial detail than the management units and land cover classes considered in this study, allowing the detection of local drought events that may remain unnoticed when the pixels are aggregated to the scale of the land cover classes considered.

5. Conclusions

The aim of this research was to test the ability of remotely sensed data sets to detect early stages of drought at the river basin scale, with particular attention to their capacity to anticipate drought impacts and gain time to inform operational land and water management. Media records from regional newspaper proved to be a helpful source of information that allowed a detailed reconstruction of drought events and impacts. The analysis using this data as a benchmark revealed the potential of the tested medium resolution remote sensing products to anticipate reported drought impacts on irrigated and rainfed areas at basin scale up to some 6 months. The best correlation-anticipation relationships were obtained for SPI, NDVI and ET. SM and LST also showed potential to anticipate drought, but with weaker correlations. GPP and PsnNet from MOD17 presented weak or no correlation for most of the areas, with only some of the rainfed areas having moderate positive correlations. The index based on in situ data currently used in the basin also provides early detection, with the exception of two of the management units, and the anticipation of drought impacts is better than that provided by the remote sensing indicators. However, the correlation of the indices based on SPI, NDVI and ET to anticipate drought impacts was found to be stronger.

The use of quantitative impact data of crop yields as benchmark showed a consistent behaviour of the remote sensing indicators with respect to the different levels of crop yield in rainfed areas among the analysed years. SPI, NDVI and ET stand out for having stronger correlations, reinforcing the findings of the first analysis. In both analyses drought on irrigated land showed less clear correlation patterns than drought in rainfed areas.

Altogether, the results confirm remote sensing products' ability to anticipate reported drought impacts and therefore provide a useful source of information to support drought management decisions at the basin scale. However, further analysis of managers' information requirements and response options is required to better assess the usefulness of these type of products in informing specific operational drought management decisions.

Acknowledgments

This research received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 603608, "Global Earth Observation for integrated

water resource assessment": earth2Observe. This work is a contribution to the Hymex Drought and Water Resources Science Team.

6. References

- [Alexandridis, T., Cherif, I., Bilas, G., Almeida, W., Hartanto, I., van Andel, S. & Araujo, A. \(2016\). Spatial and Temporal Distribution of Soil Moisture at the Catchment Scale Using Remotely-Sensed Energy Fluxes. *Water* 8\(2\): 32.](#)
- [Allen, R.G., Pereira, L.S., Raes, D. & Smith, M. \(1998\). *Crop evaporation - Guidelines for computing crop water requirements*. Rome \(Italy\): FAO. Retrieved from <http://www.fao.org/docrep/x0490e/x0490e00.htm#Contents>](#)
- [Anderson, M.C., Hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., ... Pimstein, A. \(2013\). An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 Simulations with U.S. Drought Monitoring Classifications. *Journal of Hydrometeorology* 14: 1035–1056.](#)
- [Anderson, M.C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J.R. & Kustas, W.P. \(2011\). Evaluation of Drought Indices Based on Thermal Remote Sensing of Evapotranspiration over the Continental United States. *Journal of Climate* 24.](#)
- [Anderson, M.C., Norman, J.M., Diak, G.R., Kustas, W.P. & Mecikalski, J.R. \(1997\). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment* 60\(2\): 195–216.](#)
- [Bachmair, S., Kohn, I. & Stahl, K. \(2015\). Exploring the link between drought indicators and impacts. *Natural Hazards and Earth System Sciences* 15: 1381–1397.](#)
- [Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., ... Overton, I.C. \(2016\). Drought indicators revisited: the need for a wider consideration of environment and society: Drought indicators revisited. *Wiley Interdisciplinary Reviews: Water* 3\(4\): 516–536.](#)
- [Bachmair, S., Svensson, C., Hannaford, J., Barker, L.J. & Stahl, K. \(2016\). A quantitative analysis to objectively appraise drought indicators and model drought impacts. *Hydrology and Earth System Sciences* 20\(7\): 2589–2609.](#)
- [Beguería, S. & Vicente-Serrano, S.M. \(2013\). *SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index. R package version 1.6*. Retrieved from <https://CRAN.R-project.org/package=SPEI>](#)
- [Blauhut, V., Gudmundsson, L. & Stahl, K. \(2015\). Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts. *Environmental Research Letters* 10: 10.](#)
- [Brown, J.F., Wardlow, B.D., Tadesse, T., Hayes, M.J. & Reed, B.C. \(2008\). The Vegetation Drought Response Index \(VegDRI\): A new integrated approach for monitoring drought stress in vegetation. *GIScience & Remote Sensing* 45\(1\): 16–46.](#)
- [Chatfield, C. \(2004\). *Time-series forecasting*. New York: Chapman & Hall/CRC.](#)

- CHE. (2007). *Plan Especial de Actuación en Situaciones de Alerta y Eventual Sequía en la Cuenca Hidrográfica del Ebro*. Confederación Hidrográfica del Ebro. Retrieved from <http://www.chebro.es/contenido.streamFichero.do?idBinario=5889>
- CHE. (n.d.). Portal de CHEbro. Retrieved from <http://www.chebro.es/>
- Chen, M., Senay, G.B., Singh, R.K. & Verdin, J.P. (2016). Uncertainty analysis of the Operational Simplified Surface Energy Balance (SSEBop) model at multiple flux tower sites. *Journal of Hydrology*.
- CHJ. (2007). *Plan especial de alerta y eventual sequía en la Confederación Hidrográfica del Júcar*. Confederación Hidrográfica del Júcar. Retrieved from <http://www.chj.es/es-es/medioambiente/gestionsequia/Paginas/PlanEspecialdeAlertayEventualSequia.aspx>
- Commission of the European Communities. Communication from the Commission to the European Parliament and the Council - Addressing the challenge of water scarcity and droughts in the European Union. , Pub. L. No. COM(1007) 414 final (2007). Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52007DC0414&from=EN>
- Famiglietti, J.S., Cazenave, A., Eicker, A., Reager, J.T., Rodell, M., Velicogna, I., ... Zhulidov, A.V. (2015). Watching water: From sky or stream? *Science* 349(6249): 684.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2.
- 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.
- Guerschman, J.P., Van Dijk, A.I.J.M., Mattersdorf, G., Beringer, J., Hutley, L.B., Leuning, R., ... Sherman, B.S. (2009). Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. *Journal of Hydrology* 369(1–2): 107–119.
- Hain, C.R., Mecikalski, J.R. & Anderson, M.C. (2009). Retrieval of an Available Water-Based Soil Moisture Proxy from Thermal Infrared Remote Sensing. Part I: Methodology and Validation. *Journal of Hydrometeorology* 10(3): 665–683.
- Hernández Varela, L., Lozano Valencia, M.Á. & Soleto García, C. (2003). Estudio de los acontecimientos meteorológicos extraordinarios en la comunidad autónoma del País Vasco (1870-1954) a través de la prensa. *Investigaciones Geográficas* 30: 165–180.
- Hernández-Mora, N., Gil, M., Garrido, A. & Rodríguez-Casado, R. (2013). *La sequía 2005-2008 en la cuenca del Ebro. Vulnerabilidad, Impactos y Medidas de Gestión*. Universidad Politécnica de Madrid - CEIGRAM.
- Iglesias, A., Garrote, L. & Martín-Carrasco, F. (2009). Drought risk management in mediterranean river basins. *Integrated Environmental Assessment and Management* 5(1): 11–16.
- Kallis, G. (2008). Droughts. *Annual Review of Environment and Resources* 33: 85–118.

- Keyantash, J. & Dracup, J.A. (2002). The Quantification of Drought: An Evaluation of Drought Indices. *Bulletin of the American Meteorological Society* 83(8): 1167–1180.
- Kogan, F.N. (2001). Operational Space Technology for Global Vegetation Assessment. *Bulletin of the American Meteorological Society* 82(9): 1949–1964.
- Lackstrom, K., Brennan, A., Ferguson, D., Crimmins, M., Darby, L., Dow, K., ... Smith, K. (2013). *The Missing Piece: Drought Impacts Monitoring* (Report from a Workshop in Tucson, AZ). Carolinas Integrated Sciences & Assessments program and the Climate Assessment for the Southwest.
- Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M., Wagner, W., McCabe, M.F., ... van Dijk, A.I.J.M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment* 123: 280–297.
- Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., van Dijk, A.I.J.M., ... Evans, J.P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.* 15(2): 425–436.
- Llasat, M.C., Llasat-Botija, M., Barnolas, M., López, L. & Altava-Ortiz, V. (2009). An analysis of the evolution of hydrometeorological extremes in newspapers: the case of Catalonia, 1982-2006. *Natural Hazards and Earth System Sciences* 9: 1201–1212.
- McKee, T.B., Doesken, N.J. & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17). American Meteorological Society Boston, MA.
- Miralles, D.G., Holmes, T.R.H., de Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A. & Dolman, A.J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences* 15: 453–469.
- MMA. (2005). *Informe balance del año hidrológico 2004-2005 (Assessment of the hydrological year 2004-2005)* (Programa A.G.U.A.). Madrid (Spain): Ministerio de Medio Ambiente. Retrieved from http://www.mapama.gob.es/ministerio/pags/Biblioteca/Revistas/pdf_IBH%2FBAH_2004_05.pdf
- Morid, S., Smakhtin, V. & Moghaddasi, M. (2006). Comparison of seven meteorological indices for drought monitoring in Iran. *International Journal of Climatology* 26(7): 971–985.
- Mu, Q., Heinsch, F.A., Zhao, M. & Running, S.W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment* 111(4): 519–536.
- Mu, Q., Zhao, M. & Running, S.W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment* 115: 1781–1800.
- Perez y Perez, L. & Barreiro-Hurlé, J. (2009). Assessing the socio-economic impacts of drought in the Ebro River Basin. *Spanish Journal of Agricultural Research* 7(2): 269–280.
- Potop, V. (2011). Evolution of drought severity and its impact on corn in the Republic of Moldova. *Theoretical and Applied Climatology* 105(3): 469–483.

- Redmond, K.T. (2002). The depiction of drought - A commentary. *Bulletin of the American Meteorological Society* 83: 1143–1147.
- Running, S.W. & Zhao, M. (2015, October 7). User's Guide - Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Observing System MODIS Land Algorithm (version 3.0). Retrieved from http://www.ntsug.umd.edu/sites/ntsug.umd.edu/files/modis/MOD17UsersGuide2015_v3.pdf
- Scott, C.A., Bastiaanssen, W.G.M. & Ahmad, M.-D. (2003). Mapping Root Zone Soil Moisture Using Remotely Sensed Optical Imagery. *Journal of Irrigation and Drainage Engineering* 129(5): 326–335.
- Senay, G.B., Bohms, S., Singh, R.K., Gowda, P.H., Velpuri, N.M., Alemu, H. & Verdin, J.P. (2013). Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *Journal of American Water Resources Association* 49(3): 577–591.
- Sepulcre, G., Horion, S.M.A.F., Singleton, A., Carrao, H. & Vogt, J. (2012). Development of a Combined Drought Indicator to detect agricultural drought in Europe. *Natural Hazards and Earth System Sciences* 12(11): 3519–3531.
- Sheffield, J., Wood, E.F., Chaney, N., Guan, K., Sadri, S., Yuan, X., ... Ogallo, L. (2014). A drought monitoring and forecasting system for sub-Saharan African water resources and food security. *American Meteorological Society* 861–882.
- Smakhtin, V.U. & Schipper, L.F. (2008). Droughts: The impact of semantics and perceptions. *Water Policy* 10: 131–143.
- Sonnet, J., Morehouse, B.J., Finger, T.D., Garfin, G. & Rattray, N. (2006). Drought and declining reservoirs: Comparing media discourse in Arizona and New Mexico, 2002-2004. *Global Environmental Change* 16: 95–113.
- 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.
- 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.
- Stahl, K., Kohn, I., Blauhut, V., Urquijo Reguera, J., De Stefano, L., Acácio, V., ... van Lanen, H.A.J. (2016). Impact of European drought events: international database of text-based reports. *Natural Hazards and Earth System Sciences* 16: 801–819.
- Steinemann, A.C. & Cavalcanti, L.F. (2006). Developing multiple indicators and triggers for drought plans. *Journal of Water Resources Planning and Management* 132(3): 164–174.
- Steinemann, A.C., Iacobellis, S.F. & Cayan, D.R. (2015). Developing and Evaluating Drought Indicators for Decision-Making. *Journal of Hydrometeorology* 16: 1793–1803.
- Su, Z. (2002). The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences* 6(1): 85–99.
- Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... others. (2002). The drought monitor. *Bulletin of the American Meteorological Society* 83(8): 1181–1190.

- Tsakiris, G., Pangalou, D. & Vangelis, H. (2006). Regional Drought Assessment Based on the Reconnaissance Drought Index (RDI). *Water Resources Management* 21: 821–833.
- UNISDR. (2009). *Drought Risk Reduction Framework and Practices: Contributing to the Implementation of the Hyogo Framework for Action*. Geneva, Switzerland: United Nations secretariat of the International Strategy for Disaster Reduction (UNISDR). Retrieved from http://www.unisdr.org/files/11541_DroughtRiskReduction2009library.pdf
- van Dijk, A.I.J.. & Renzullo, L.J. (2011). Water resource monitoring system and the role of satellite observations. *Hydrology and Earth System Sciences* 15: 39–55.
- Vasiliades, L., Loukas, A. & Liberis, N. (2011). A water balance derived drought index for Pinios River Basin, Greece. *Water resources management* 25(4): 1087–1101.
- Verma, M., Friedl, M.A., Richardson, A.D., Kiely, G., Cescatti, A., Law, B.E., ... Propastin, P. (2014). Remote sensing of annual terrestrial gross primary productivity from MODIS: an assessment using the FLUXNET La Thuile data set. *Biogeosciences* 11: 2185–2200.
- Vicente-Serrano, S.M., Beguería, S., Lorenzo Lacruz, J., Camarero, J.J., López-Moreno, J.I., Azorin-Molina, C., ... Sanchez-Lorenzo, A. (2012). Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications. *Earth Interactions* 16: 27.
- Vicente-Serrano, S.M., Cuadrat-Prats, J.M. & Romo, A. (2006). Early prediction of crop production using drought indices at different time-scales and remote sensing data: application in the Ebro Valley (north-east Spain). *International Journal of Remote Sensing* 27(3): 511–518.
- Wagner, W., Dorigo, W., De Jeu, R., Fernandez, D., Benveniste, J., Haas, E. & Ertl, M. (2012). Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. I-7). Presented at the XXII ISPRS Congress, Melbourne.
- Wilhite, D.A. (2000). Drought as a Natural Hazard: Concepts and Definitions. In D. A. Wilhite (ed.), *Drought: A Global Assessment* (Vol. Vol. I). London: Routledge.
- Wilhite, D.A. (2011). National Drought Policies: Addressing Impacts and Societal Vulnerability. In M. V. K. Sivakumar, R. P. Motha, D. A. Wilhite & J. J. Qu (eds.), *Towards a Compendium on National Drought Policies: Proceedings of an Expert Meeting, July 14-15, 2011, Washington, D.C. USA*. Geneva, Switzerland: World Meteorological Organization. Retrieved from <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1079&context=droughtfacpub>
- Wilhite, D.A., Svoboda, M.D. & Hayes, M.J. (2007). Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness. *Water Resources Management* 21(5): 763–774.
- Wilhite, D.A. & Vanyarkho, O. (2000). Drought: Pervasive Impacts of a Creeping Phenomenon. In D. A. Wilhite (ed.), *Drought: A Global Assessment* (Vol. I). London: Routledge.
- World Meteorological Organization (WMO). (2012). *Standardized Precipitation Index - User guide*. Geneva, Switzerland: World Meteorological Organization. Retrieved from http://www.wamis.org/agm/pubs/SPI/WMO_1090_EN.pdf

Zhang, F., Chen, J.M., Chen, J., Gough, C.M., Martin, T.A. & Dragoni, D. (2012). Evaluating spatial and temporal patterns of MODIS GPP over the conterminous U.S. against flux measurements and a process model. *Remote Sensing of Environment* 2012: 717–729.