Review of the manuscript nhess-2019-314, entitled "Meteorological drought in the Miño-Limia-Sil hydrographic demarcation: The role of atmospheric drivers", submitted by Rogert Sorí et al. for possible publication in Natural Hazards and Earth System Sciences. Recommendation Sorí et al. assessed drought characteristics in the Miño-Limia-Sil basin (NW Iberia) from 1980 to 2017 using the Standardized Precipitation Evapotranspiration Index (SPEI). The temporal variability of drought metrics was linked to changes in the dominant weather types and atmospheric circulation patterns in the region. Weather types in the study domain were classified using an automated version of the standard Lamb weather types' scheme. The study is interesting from the climatological point of view and falls within the scope of NHESS. The level of innovation of this work is fair/reasonable. Albeit with the availability of several studies, which assessed drought characteristics in the IP (most of them are referenced herein), this study applied well-established/existing methods in a standard fashion for a cross-boundary basin in the IP. The manuscript is generally well-written. However, the manuscript cannot be accepted for publication in its current version. Some methodological issues should be clarified. The structure of the manuscript should be refined. My major and specific comments are listed below:

I. Major comments

I want to point out one smaller aspect that deals with the general readability of the text. The readability is sometimes hampered by the use of various abbreviations (e.g. P, Y, M, MSRB, LRB, MLSHD, NWIP, VIMF, etc). If the reader is not very familiar with the terms and abbreviations, it is hard to follow. I would suggest using only those abbreviations that are reoccurring and central for the topic and trying to avoid others. This would strongly improve readability.

Following your advice some of the abbreviations were removed such as "y", "mo", NWIP, MSRB, LRB?

In the methods section, the authors need to clarify the final number of weather types retained in their work. Have they applied any regrouping/reclassification of hybrid types? What are the stopping criteria of their classification scheme? What statistical significance criteria are applied? How robust and reproducible are the results? What guided the reduction and grouping? Can you illustrate what the types represent, via typical flow or MSLP patterns? Also, a map illustrating the 16 points used for weather type classification should be included.

The total number of types was included in the study but, as suggested by the reviewer a regrouping of the hybrid types was realized in order to obtain the monthly percentage for every weather types. This and other issues addressed by the reviewer where added into the text to explain better the WTs computation process. The MSLP associated with every WTs is presented in Figure 4 and the points used in the computation are represented in Figure 1 (right panel). Transcription is presented below:

"The methodology here described is able to daily identify the weather pattern (from the 26 listed before) presented over the area of study. From this daily information, and in order to study the WTs influence on monthly SPEI series, the monthly frequency of occurrence for every pure WTs is computed for the period 1979-2017. In the frequency computation, the 26 WTs are regrouped in the 10 "pure" ones. This procedure was done following the same approach applied in Trigo and DaCamara (2000) in which the hybrid types count equally as a half occurrence to each of their pure types, being the 10 final number of WTs analysed in this study."

Regarding stopping criteria and statistically significance, it has been applied in other studies a 'stopping rule' to determine, for example, at what point to stop adding predictors (WTs) to obtain *the relative (%) contribution to total monthly Iberian rainfall by each WT. However, is not our case. However, as in other previous studies we distributed the few cases (<2%) with unclassified situations among the 26 classes. The unclassified situations occur when the module of the total shear vorticity (Z) is less than the standard deviation of Z/2 and simultaneously the total flow (F) is less than the square root of the sum of squared westerly and southerly flows.*

The types represent the MSLP. We considered the circulation associated to each WT to explain the role of directional flows. Several studies explain how to obtain the WT through the methodology here utilized, so, the results are perfectly reproducible.

I am wondering why the authors have not used some Mediterranean specific indices, such the WeMO and MO. Also, I am wondering why the authors did not consider SST as a driver of drought variability in their domain. This can be implemented using El Niño 3.4 index (SST anomalies in the central Pacific), El Niño 1 + 2 index (SST anomalies in the eastern Pacific) or SST anomalies in the tropical Atlantic. I am aware that the manuscript focuses mainly on specific atmospheric drivers, but at least in the discussion, the authors need to highlight the possible role of SST warming in drought reinforcement. See, for example: https://doi.org/10.1175/JCLI-D-11-00296.1

Thank you for your recommendation. These Index were also investigated and no significant correlation was obtained for any of El Niño Index (r values from 0.05 to 0.25 with the SPEI6 and TNA (r < 0.07). This information is included in the manuscript. The Index BEST already introduce information of SST of the region 3.4 and also takes into account the SOI (the pressure difference between Tahiti and Darwin), to give a better representation of the phenomenon. It has been argued in the manuscript.

The analysis for the WeMO index was incorporated and removed the AMO.

The authors should discuss their results in the context of some earlier studies whose results contradict the findings of the current study (particularly with respect to the significant role of NAO in drought variability in Europe in general and the IP in particular). See for example: https://doi.org/10.1007/s10584-007-9285-9. 10.1007/978-94-007-1372-7_3

We improve the discussion considering this advice. The link seems to be wrong.

The authors focused mainly on the temporal variability of drought and its connections with WTs and climate teleconnections. However, I have not seen any attempt to show the varying spatial response of drought to these drivers. The reasonable spatial resolution (0.1º) of E-OBS allows for a reliable assessment of the spatial variability of drought in response to the different atmospheric configurations. The authors indicate in the abstract "We concluded that regional patterns of land-use change and moisture recycling are important to consider in explaining runoff change, integrating land and water management, and informing water governance". I think decision-makers and water resources planners in any catchment seek for detailed information on the spatial variability of droughts so that they can adopt integrated policies and strategies for managing their catchment, taking into consideration the different conditions at both upstream and downstream.

To assess this, we decided to perform an EOF analysis for those months of the MLSHD when the average SPEI was ≤ - 0.84. In this analysis the PCA of the first EOF were correlated with each series of the WTs for the same chosen months. The first 6 EOF explain more than the 95% of the total *explained variance, and the two first around the 80%. Figures (EOF & PCA) has been added to the manuscript.*

In the Introduction, the authors lack the opportunity to comprehensively provide evidence on the hydrological, environmental and socioeconomic importance of MLSHD in the IP. A short description of the study domain, highlighting its main physiological, climatic and hydrological settings is needed. Section 3.1 is misplaced in the results section and should be forwarded to a new subsection called "study area".

Thank you for your advice. The Introduction has been improved by adding new information about its hydrological, environmental and socioeconomic characteristic of importance. The section 3.1 was also moved to the Introduction section in order to provide a more compressive information about the climate of the region.

A justification of the selection of the (-0.84) as a threshold for defining drought events is needed.

We improved this in section 2 as follows:

We avoid considering that any precipitation below the mean constitutes a drought. Therefore, the classification of drought categories according to SPI values proposed by Agnew (2000) (Table 1) was utilised in this study. Other authors have also employed this classification for investigating drought in the IP (e.g. Vicente-Serrano et al., 2006; Páscoa et al., 2017). This classification despite to be pre-established is built by probability classes rather than magnitudes of the SPI, and is therefore suggested as a more rational approach, with a most noticeable effect at the demarcation of mild and moderate droughts (Agnew, 2000). Then a drought episode was considered to occur (onset) when the SPEI at the temporal scale of 1 month fell below zero, reached a value of at least -0.84, and later returned to positive values. The threshold of -0.84 corresponds to 5% probabilities, whereby drought is only expected 2 years in 10, which reduce the incidence of mild meteorological droughts.

In the methodology, the authors should clarify how the different drought metrics were computed? How were the trend and its statistical significance assessed? Have they accounted for the possible presence of serial correlation in the series?

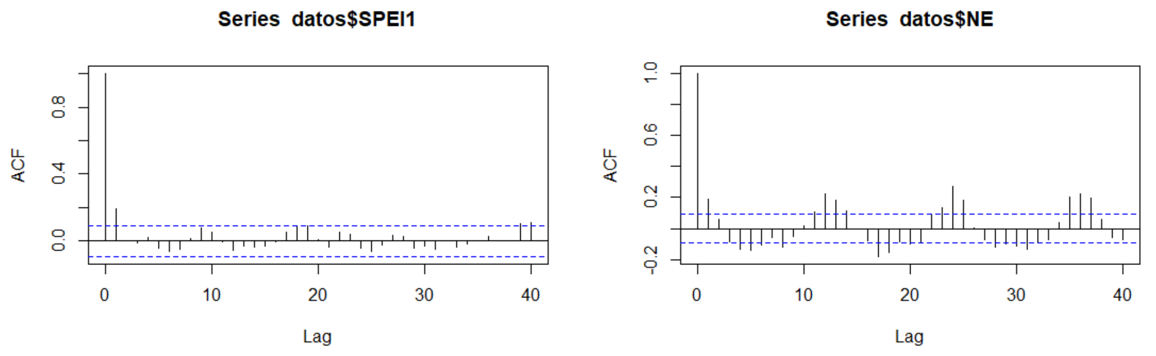
It is now clarified:

- *The duration is computed as the sum of all months from the onset with negative values, the peak is the month in which the episodes reach the highest value of SPEI1, and the severity is calculated as the sum of all SPEI values (in absolute values) during the episode.*

Besides, the whole section 2 was improved.
In the case of the trend we just made a regression and the significance is calculated using the Wald test with t-distribution of the test statistic, and the trend is considered significant when $p < 0.05$. In this new version the results were similar after utilize the Mann Kendall test and the Sen´s slope.

At first we didn't consider the presence of serial correlation in the series. I order to assess the possible autocorrelation of the series was utilised the function of autocorrelation (ACF) in the R program. For the SPEI1 the autocorrelogram show, as expected, non-significant values. However, for some WTs (e.g. NE) (see right panel below) it seems that exist a clear seasonality and several values are significant, confirming the autocorrelation. To solve this concern we applied the Mann-Kendall Test of Prewhitened Time Series Data in Presence of Serial Correlation Using the von Storch (1995)

Approach. A detailed description of this method was added to section 2. (next figures are just an example and won't be included into the manuscript).



The authors should clarify the rationale behind constraining their study from 1980, while EOBS dataset extends back to the 1950s (probably not for all climate variables).

Despite the EOBS data is available from 1950s other variables (such as Era-Interim data used for WTs computations) is from 1979 and at first moment the river discharge data were from 1980. Thus, we decided to reduce the study to the period 1980-2017.

I would recommend adding a new figure, in which the authors compare the accumulated SPEI values corresponding to each WT.

Here we follow your instruction: Figure 6: I would recommend plotting the events at the xaxis, while the stacked bars show the contribution of each WT to such events. This will deliver the message clearer.

So, the new figure is compound by 10 new figures (10 episodes) where are shown the SPEI1 and each WT anomaly. A similar figure could be accumulating the anomalies of each WT, but the first choice you recommend also illustrate better the real frequencies of each WTs highlighting the percentages at the SPEI peak, or at the end of the episode. Besides, on a single month several WTs occurs and we have the percentage of days for each WT every month, while just one SPEI value for each month, because it is calculated at monthly scale, so, it is not possible to accumulate the SPEI values corresponding to each WT.

Prior to calculating correlation, it is important to detrend the series of the frequency of WT.

The series of WTs and SPEI were de-trended and the correlation was computed again. In this case the spatial correlation was also made.

In the discussion, the authors should refer to the role of zonal and meridional circulation in drought characteristics (in the context of the directional WTs).

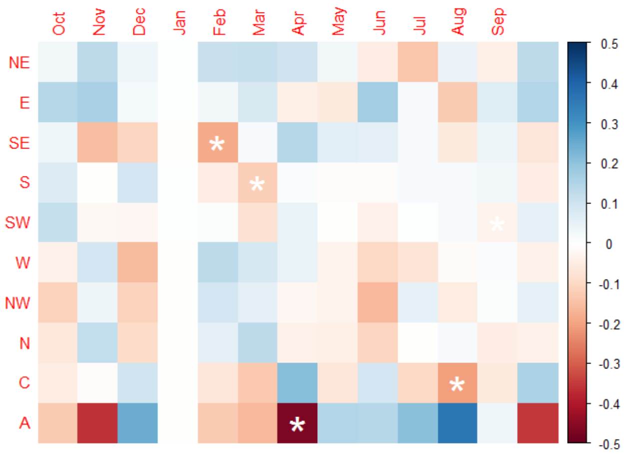
We improved the discussion considering this advice.

It is recommended to add a table, which summarizes trends of the 10 main WTs over the study period, their statistical significance and compares those with the trends observed for the different drought metrics (e.g. drought, severity, and occurrence).

We calculated the trend for each WT and also for individual months. We applied the Mann-Kendall Test of Prewhitened Time Series Data in Presence of Serial Correlation Using the von Storch (1995) Approach. As all p values are greater than 0.05, trends are not statistically significant.

	NE	E	SE	S	SW	W	NW	N	C	A
Z-value	1.04	1.43	- 0.29	4.11	0.64	-0.37	-0.22	-0.18	-0.16	-0.70
Slope	0.002	0.004	- 0.00	0.000	0.000	- 0.0002	- 0.0002	- 0.0001	- 0.0003	- 0.004
p-values	0.29	0.15	0.77	0.96	0.52	0.70	0.81	0.85	0.86	0.48

Also, no statistically significant trends appear in almost every WTs and months. Significant trends only appear for SE in February, S in March, SW in September, C in August and A in April. Moreover, in most of the cases the trend show low values. We did not added the table above into the manuscript, but commented the results.



HS method for calculating PET is a temperature-based method, which is more suitable for arid and semi-arid regions (not humid regions like the study domain). Have the authors tested the performance/accuracy of this method in their region? Several studies reported a less performance of temperature-based methods in assessing PET in humid climates.

We agree about your concern. The limitation to computed the Eto thought the Penman Monthie method lies on the lack of several variables at high resolution for a long study period. To support our choice, the next paragraph was added to the manuscript.

By this method we do not consider that relative humidity and wind speed are also important factors that determine the vapour density above the soil surface and the aerodynamic resistance for vapor transport, permitting more realistic Eto values and consequently drought assessment (Bittelli et al., 2008; Vicente-Serrano et al., 2010; WMO, 2012; Davarzani et al., 2014). However, even though the Penman-Monteith offers a more accurate estimation of reference Eto than the Hargreaves formula (Tomas-Burguera et al., 2017), results of Vicente-Serrano et al., 2014 showed that Eto in Spain *estimated by the Hagreaves-Samani method for the period 1961 – 2011, had the closest agreement with the Eto obtained by the Peaman Monthie method in terms of temporal evolution and magnitude respect other eleven methods. These authors also found high correlations between Eto obtained by both methods in the northwest IP.*

Vicente-Serrano et al., (2014). Reference evapotranspiration variability and trends in Spain, 1961–2011. Global and Planetary Change, Volume 121, October 2014, Pages 26-40. https://doi.org/10.1016/j.gloplacha.2014.06.005

Given the limited area of the study domain, the spatial resolution of SMroot data seems to be coarse (0.25◦ × 0.25◦) to provide a reliable assessment of the response of soil moisture to precipitation deficit. Also, in humid climates like those of the study area, the response of soil moisture to accumulated precipitation deficit is more pronounced at longer time scales (not 1-month time scale). The persistence of negative soil moisture anomalies is expected to be higher when there is a cumulative long-term decrease in the amount of precipitation. This aspect should be discussed thoroughly.

In order to perform a better analysis, the figure and explanation for this section was removed. We considered your concern regarding the low resolution of the SMroot, which is representative for around 10 grid points. A similar analysis was performed but utilizing the soil moisture from the high-resolution (~ 4km) global dataset of monthly climate and climatic water balance "Terraclimate". We also utilized the Standardized Soil Moisture Index (SSMI) (Haoand AghaKouchak, 2013)

Describe all symbols given in Eq. 1.

Symbols are now described:

we used the method proposed by Hargreaves and Samani (1985) based on temperature data to estimate the Eto according to Equation 2:

$$Eto=0.408*Ch*Ra*(\sqrt{(Tx-Tn)})+(Tm+17.8) \qquad (2)$$

Where Ch = 0.0023; Ra is the extraterrestrial radiation (derived from the latitude and the month of the year), and Tx, Tn and Tm the maximum, minimum and mean temperature respectively.

Which index exactly of the NAO, as well as ENSO, was used? Please, be more specific. There are different indices for quantifying each of them.

We added this and more information.

To obtain the Northern Hemisphere teleconnection indices utilized in the study the Climate Prediction Center (CPC) utilise the Rotated Principal Component Analysis (RPCA) method used by Barnston and Livezey (1987) but utilizing monthly mean standardized 500-mb height anomalies obtained from the CDAS in the analysis region 20°N-90°N. This procedure isolates the primary teleconnection patterns for all months and obtain the index time series.

Barnston and Livezey (1987, Mon. Wea. Rev., 115, 1083-1126)

Section 2.5 should be placed earlier in the materials and methods section (before the description of drought calculation).

The section was moved

The role of aerodynamic components in drought evolution should be discussed, given that these influences are not considered in HS method.

The next paragraph was added to section 2 in order to expose this topic, and along the manuscript are discussed other issues:

By this method (HS) we do not considered that relative humidity and wind speed are also important factors that determine the vapour density above the soil surface and the aerodynamic resistance for vapor transport, permitting to obtain more realistic Eto values and consequently drought assessment (Bittelli et al., 2008; Vicente-Serrano et al., 2010; WMO, 2012; Davarzani et al., 2014). However, some studies have shown similar Eto estimations by means of the Penman-Monteith and Hargreaves methods in Spain (López-Urrea et al. 2006; Gavilán et al. 2008; Vanderlinden et al. 2008; López-Moreno et al. 2009), although the Penman-Monteith offers a more accurate estimation of reference Eto than the Hargreaves formula (Tomas-Burguera et al., 2017).

II. Minor comments

Title: It is recommended to indicate the location of the study domain (i.e. NW Iberia), as the majority of the NHESS readers are not familiar with the study basin. Also, it is important to include "hydrological droughts" in the title.

The title was changed as follow: ''Hydrometeorological droughts in the Miño-Limia-Sil hydrographic demarcation (NW Iberian Peninsula): The role of atmospheric drivers"

In order to emphasize more on the hydrological issue, two figures showing the temporal evolution of the Standardized Soil Moisture Index (SSMI) (Haoand AghaKouchak, 2013) and the Standardized Runoff Index (SRI)(Shukla and Wood, 2008) were added. A detailed explanation of both index was added in section 2.

P1 - L14 and other parts of the ms: "period of" <> "period".

Thank you. It has been changed along the text

L16 and other parts of the ms: "mo" <> "month".

It was changed along the text

L17: For a study that covers 38 years, the use of the term "historically" is misleading; please, define the confidence interval at which the significance was assessed.

The historically term was removed from the text and the confidence level was provided (95%).

L18: "different" <> "the different".

Changed

L19: Based on which scheme this classification was made? The abstract should stand alone based on this basic information.

It was included

- *… a daily weather type classification based on standard Lamb scheme was utilised for the entire Iberian Peninsula.*

L20: “were directly related to dry and wet conditions” This statement is vague, with no clear phrasing. It does not make a clear conclusion on whether these weather types are favoring for above-normal or below-normal precipitation.

This sentence was removed and rewritten as follow:

Frequency and correlation analysis show that weather types conditioning an atmospheric circulation from the south-east/west, east/ west, north-east/west and the pure anticyclonic/cyclonic are associated to dry/wet conditions in the Miño-Limia-Sil Hydrographic Demarcation.

L25: Please, define the rainy season.

It was now defined: (October – May)

L27 and other parts of the ms: “1 y” <> “1 year”.

It was changed along the text

It is unclear how meteorological droughts assessed at 1-month time scale can be linked with land use changes (which almost occur at a coarse temporal scale”.

You are right, it was our mistake because at the same time were writing of a similar manuscript with a different scope. The sentence was removed.

P2 - L5: Please, give some examples of these thermodynamic factors (e.g. wind speed, air pressure).

In this case we modified the paragraph. Besides, in section 2 are explained the role of humidity and wind on the modulation of the Evapotranspiration and consequently the occurrence of dry conditions.

- *This phenomenon is usually considered a prolonged dry period in the natural climate cycle that can occur anywhere in the world. It is initially caused by a lack of rainfall as well as for thermodynamics processes that affect content of water on the soil, which are induced by the wind speed, temperature, relative humidity and solar and long-wave radiation (WMO & GWP, 2016; Vicente-Serrrano et al., 2010; Sereviratne, 2012; Miralles et al., 2019).*

- L15: Delete “e.g.” –

It was deleted

L19: “the precipitation” <> “precipitation”.

Changed

L25: “land” Do you mean air temperature? LST has a different conception and is mostly assessed using remote sensing products (e.g. MODIS, AVHRR), which are only available for the most recent decades.

The reviewer is right, it was clarified

L28: The study of Vicente-Serrano et al. (2011) does not provide any assessment of future projections of precipitation.

You are right. The reference was removed and added the correct.

Spinoni, J., Vogt, J.V., Naumann, G., Barbosa, P. and Dosio, A. (2018), Will drought events become more frequent and severe in Europe?. Int. J. Climatol, 38: 1718-1736. doi:10.1002/joc.5291

P3 - L8: "a homogeneous region in terms of the total P variance over the IP". This statement should be elaborated thoroughly.

This argument has been better discussed in the analysis of EOF of the SPEI.

P4 - L25: "A drought episode was considered to occur when the SPEI at the temporal scale of 1 mo fell below zero, reached a value of at least -0.84, and later returned to positive values". This definition should be made simpler.

The episode characterization was rewritten.

A drought episode occurs everytime the SPEI1 is continuously negative and reaches the value of −0.84 or less.

P5 - L1: "Results" <> "Results aepisond discussion".

Changed

L9: Language and style should be revised.

Thank you. We improved it.

L12: "modulate" <> "impact".

Changed

It is unclear why the classification of weather types is only restricted to the period 1989-2017.

This period is not correct, we apologize. The correct is 1980-2017. I was corrected in the text. This period was selected based on the available ERA Interim reanalysis data together with the rest of datasets in the moment we began to do this research (in particular the river discharge availability).

L15: Please, define this spatial window.

It was defined

For classifying weather types, the authors should clarify how SF, WF, ZS, ZW, F, and Z were computed?

The complete section was improved and the formulas were added:

$$SF = 1.305[0.25(p_5 + 2 \times p_9 + p_{13}) - 0.25(p_4 + 2 \times p_8 + p_{12})]$$

$$WF = [0.5(p_{12} + p_{13}) - 0.5(p_4 + p_5)]$$

$$ZS = 0.85 \times [0.25(p_6 + 2 \times p_{10} + p_{14}) - 0.25(p_5 + 2 \times p_9 + p_{13}) - 0.25 \times (p_4 + 2 \times p_8 + p_{12})$$

$+0.25(p_3 + 2 \times p_7 \; + p_{11} \;)]$

$ZW \; = 1.12 \times [0.5 \times (p_{15} + p_{16}) - 0.5 \times (p_8 + p_9)] - 0.91 \times [0.5 \times (p_8 \; + p_9) - 0.5 \times (p_1 \; + p_2 \;)]$

$F \; = (SF^2 + WF^2)^{1/2}$

$Z = ZS + ZW$

P6 - L20: "from daily values" <> "aggregated from daily values".

Modified

L21: The name of the station "Albufeira Do Alto" does not fit with that labeled in Figure 1.

This Figure has been modified.

P7 - L5: "the annual cycle" <> "the year"; "western" <> "the western".

Changed

L20: What is the difference between "extensive" and "intense"? Do you refer to drought duration and severity?

Yes, it was clarified into the text.

P8 - L30: "for in" <> "for".

This expression does not appear in P8, it was changed in P18.

The acronyms "WTs" and "CWTs" are used interchangeably er in the text. - P23 (L10): "the soil" <> "soil".

It was now corrected and only WTs is now used in the text

P23 (L10): "the moisture" <> "moisture".

Changed

This work emphasized that drought did not respond linearly to most of the dominant circulation patterns in this region (apart from SCAN, AO) at 1-month timescale (Figure 10a). This finding should be discussed thoroughly in the text and linked with available literature.

We agree and the explanation was improved.

Tables

Table 1: There is a refinement of the drought categories of Agnew (mild drought is masked with another category).

Modified!

Table 2: Trends in SPEI values should be expressed in z-units/year.

Done

Figures

Figure 1: In the legend and caption, “rivers” <> “streams”. The negative symbol corresponding to the longitudes should be deleted, given that the direction “W” is already included. It is important to include a distribution of the meteorological stations whose data were used for SPEI calculation.

This figure was totally modified in order to follow your suggestions but also for representing the grid points utilized to compute the WTs.

Figure 3: how were drought episodes defined? Have you applied n consecutive months with SPEI <-0.84?

Yes. The explanation was improved in section 2. An episode was defined as the consecutive negative SPEI1 that reach at least the values of -0.84 or less.

Figure 4: I would recommend using the anomalies (not the actual values) of SLP corresponding to the different WTs. This will facilitate defining the positive and negative centers of action that control air advection at the surface.

Ok. Anomalies were calculated and will be added to the manuscript.

Figure 5: The use of the symbol “x” should be described. The use of the legend in a vertical form is confusing, given that all WTs at the top of the panel show a negative correlation (shown in blue). I would recommend reversing the legend so that negative values of correlation are shown at the top, while negative correlations are illustrated below. Why the authors did not use a portrait diagram showing the interpolated surface of the correlation coefficient, with some contour lines to show the significance of the correlation? This will facilitate the readability of the figure.

The X values indicate those correlations no statistically significant at $p < 0.05$. We reversed the legend. We considered to contour the correlations following your instructions, but, the function contour automatically interpolates and in this analysis the correlations are made for different WTs, so, in order to show the more precisely, we prefer to keep this format of circles.

A new figure showing the spatial correlations of the SPEI with each WT was made. This is NOT the last version of this figure, we will show just the demarcation, change the color and limits of the color bar and add a line for indicate the statistically significant values.

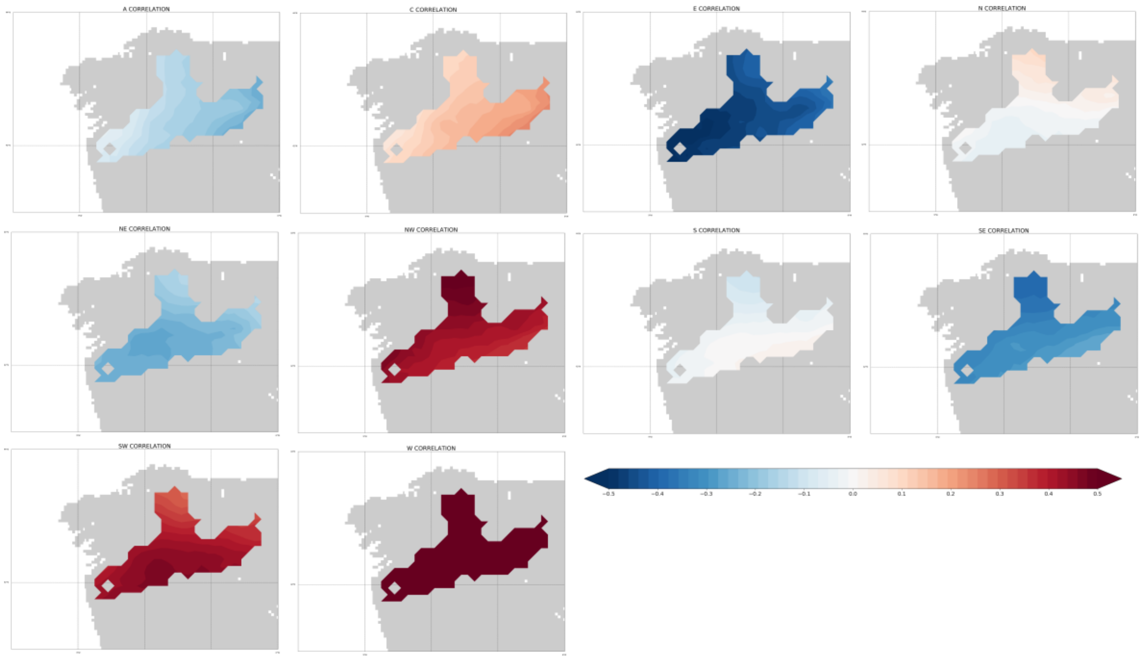


Figure 6: I would recommend plotting the events at the xaxis, while the stacked bars show the contribution of each WT to such events. This will deliver the message clearer.

We change this figure following your advice.

Figure 7: why the percentages are given in negative? To which WT refers the "red" color? I would recommend adding a column to the three drought categories, which refers to wet conditions (i.e. SPEI values >0). This contrast can show interesting results about the role of each WT during dry vs. wet conditions

This figure was improved following your suggestions. The red color don´t refer to any WT; it indicates the number of months under each drought category. The same figure was added for the wet conditions.