Linking European droughts to year-round weather regimes

O. Savary, C. Ardilouze, J. Cattiaux

We would like to thank the reviewers for their responses and constructive comments, which we have considered very carefully. Both reviewers raised the issue of insufficient justification for the methodological choices made, as well as the lack of meteorological interpretation of the results obtained. We hope that the revised version of the manuscript will satisfactorily address their concerns. We will reply point by point to each reviewer.

1 Reviewer #1

1.1 Major comments

Comment
Interpretation of the key result: You focus a lot on how important the weather regime frequency anomaly term, α_1 in Eq. 11, is for drought events, which makes sense as it's the key question of this paper. However, you leave (at least) me with the open but relevant question, what that other, more important term, α_2 in Eq. 11, actually is from an intuitive point of view and why it shapes the drought events to first order. Can you improve your discussion and conclusion with respect to this aspect? What are the "processes" hidden behind this term? Is it also related to smaller-scale/synoptic anomalies that do not appear in the low-frequency regime definition? Maybe it would be helpful for a better understanding to look into the regime evolution ahead of some individual events such as the one of 1960-07-04 in Fig. 8, for which regime frequency anomalies had no importance at all. In previous studies (I think some of them include one of your co-authors), this decomposition is applied to understand climate change signals. As far as I understand, in these studies your second term (α_2 in Eq. 11) can be interpreted as a more thermodynamically driven term. However, in the drought context here, this does not really make sense, which is why I struggle to understand it...

Reply: We further investigated the meaning of the second term α_2. On average for droughts, the term ΔC_k dominates. If ΔC_k is large for a regime k, this means that within pre-drought periods, the precipitations occurring during days assigned to regime k only loosely match the canonical precipitation pattern associated with k. We identify two main causes for this discrepancy: limits of the WR approach to describe the atmospheric circulation, and the local variability of precipitation associated to WR. This corresponds to either a difference in zg500 composite between pre-drought days and all days, or an alteration during pre-drought days in the relationship between zg500 and precipitation. We discuss these two points below.

Limit of the WR approach : we study North Atlantic circulation through weather regimes. To assess to what extent the Z500 patterns that occur on pre-drought periods match the canonical patterns of the regimes, we performed the same decomposition as for precipitation but for $zg500$ (see Eq. 1).

$$\overline{zg}^S = \sum_k \Delta f_k C_k + \sum_k f_k \Delta C_k + \sum_k \Delta f_k \Delta C_k \tag{1}$$

We therefore decomposed the zg500 anomaly during the 90 days preceding the droughts into a WR frequency anomaly term , a zg500 pattern anomaly term, and a second order term. This decomposition is shown in 1. The first column corresponds to the average zg500 over the 90 days prior to droughts, for all droughts in the region. There is consistently an anticyclonic anomaly centred on the region of interest. The second term (third column) still predominates in intensity. The centers of action are systematically more

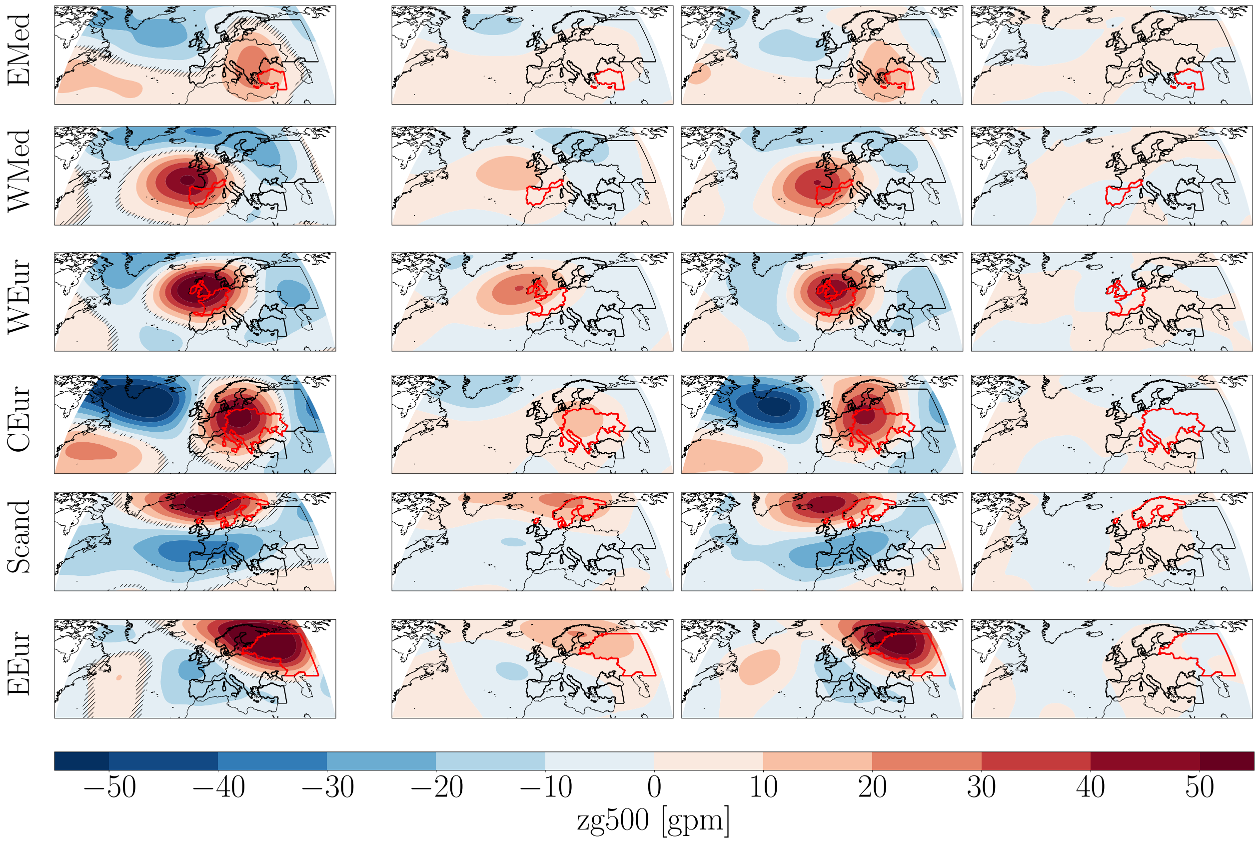

Figure 1: (I) Composite zg500 anomaly over the 3 months preceding the droughts for each region. Non-significant results (p-value > 0.05) are hatched. (II to IV) zg500 anomaly reconstructed.

intense for this term than for the frequency anomaly term (second column). However, the centers of action are located in the correct place.

Let us assume that each weather pattern corresponds perfectly to its associated precipitation pattern, regardless of the situation. We have shown that the zg500 pattern during the pre-drought period can be explained in small part by frequency anomalies and largely by a distortion of the canonical zg500 patterns.

Local variability of precipitation associated to WR : Some small-scale precipitation phenomena are not influenced by the large-scale circulation. These events cause precipitation anomalies to deviate from the canonical patterns associated to WRs. To highlight these phenomena, we calculate the precipitation standard deviation for each regime. The results are shown in Fig. 2. We normalized the intra-WR standard deviation by the precipitation standard deviation, in order to quantify the

We also notice a substantial variability over central Europe and near the Mediterranean coast, for which a more detailed analysis is provided at the end of the reply (see Fig. 8). We also present the seasonality of the standard deviation of precipitation.

Comment
Role of weather regime persistence and intensity: The frequency of daily regime occurrence 90d prior to a drought – as you define it here – does not tell us anything about regime persistence or regime intensity. For example, two drought events could be preceded by the same frequency of a regime, but in one case it could be one single, long-lasting regime event that is very strong, while in the other case, the regime could appear multiple times but in shorter and maybe weaker spells. I think this could make

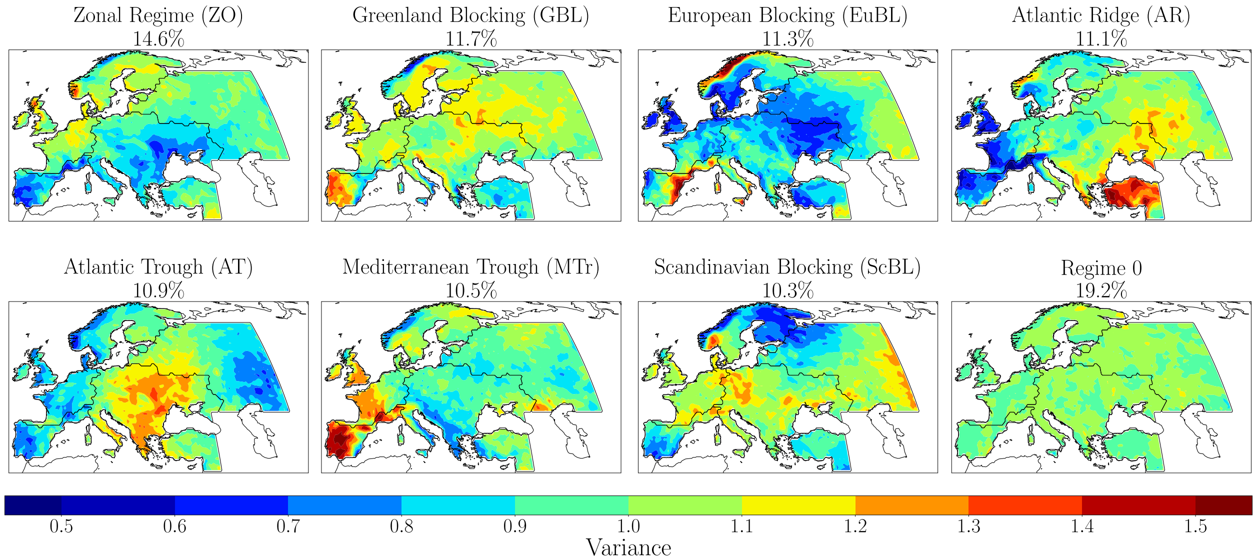

Figure 2: Intra-WR precipitation standard deviation [ ]

> a meteorological difference. Did or could you make some statistics over the regime duration before drought events, to understand if duration is an aspect that makes regime occurrence "more important" for droughts (for instance you could do your analysis in Fig. 8 but somehow split between longer-lasting and shorter-lasting regimes)? For regime intensity it might be harder to test this though, because you would have to define a regime index similar to other studies you cited. And a thought related to my first comment: Is it possible that your second term, α_2 in Eq. 11, is somehow also dependent on regime intensity, since the strength of a precipitation anomaly ΔC_k when being in regime k might be larger in a stronger regime period?

Reply: We distinguish between the size and the number of sequences for each regime. Fig. 3 presents these results. We find that in most cases, the frequency anomaly cannot be explained both by an increased persistence alone or an increased number of sequences alone, but rather by a combination of both. In some cases (such as the EuBL regime in the EMed region, for example), the positive frequency anomaly can be explained by an increased sequence size but an unchanged number of sequences.

We also present the zg500 composites for each regime, distinguishing between situations that can be explained by anomalous frequencies of regimes' occurrences (top) and those that cannot be explained by the regimes (bottom), under drought conditions for the WMed region (see Fig. 4). The regimes are relatively well represented in terms of shape and intensity. The intensity of the regimes does not differ between cases that are well captured by the WRs and cases that are poorly captured. The difference between the two types of drought does not appear to stem from a marked distortion of atmospheric centres of action in one case compared to the other.

> **Comment**
> Sensitivity to drought definition: SPI3 droughts are often used as proxies for agricultural droughts due to the relatively long timescale that starts to affect soil moisture anomalies. As far as I know, SPI1 are rather the ones used as proxy for the classic meteorological droughts. Did you apply the same analysis to SPI1 and find out if regime frequency anomalies play the same (or a more important) role for this shorter, monthly timescale?

Reply: We agree that different SPI timescales reflect different types of droughts. But, in our study, we are interested in the seasonal-scale impact of atmospheric regimes on precipitation deficits, which is more relevant for SPI3. We have therefore not carried out the study for SPI1 here. This would require skilled work that is not immediately within our objectives given the different time scale.

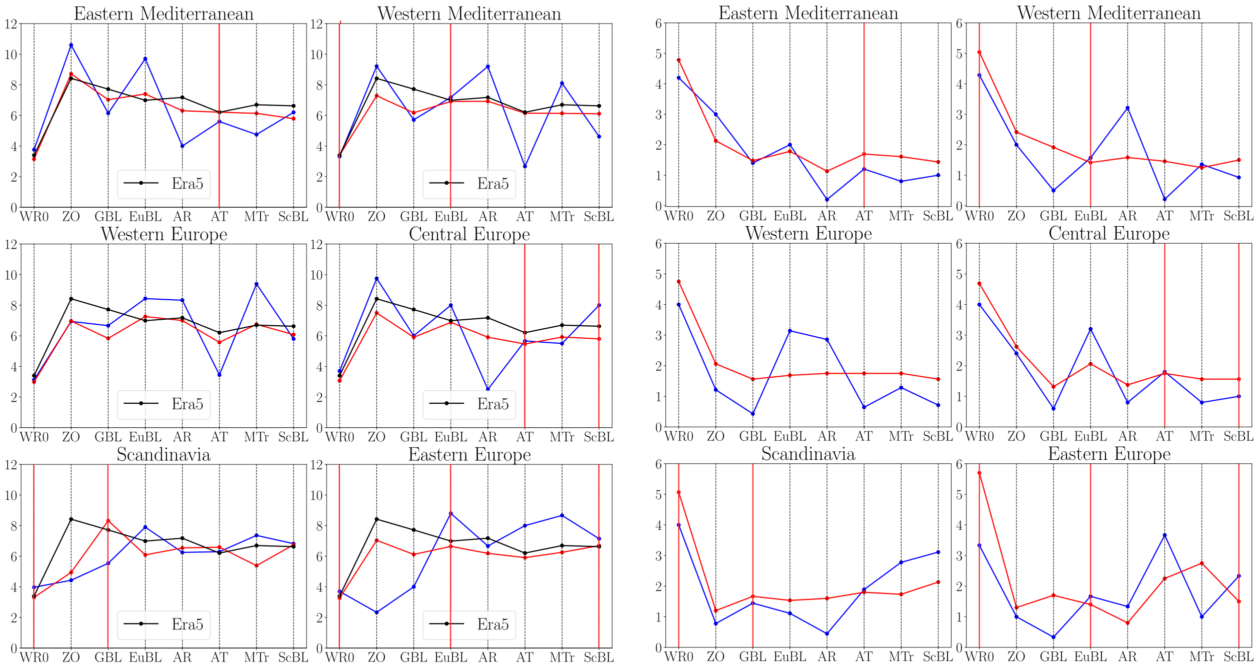

Figure 3: [LEFT] Mean duration of WR sequences (in days) for droughts explained by WR (blue), droughts not explained by WR (red), and for the entire ERA5 period (black). [RIGHT] Mean number of occurrences of WR sequences, for α_1 big (blue), α_1 small (red), and for the entire ERA5 period (black). WR with a non-significant frequency anomaly are shown with a vertical red line.

> **Comment**
> Role of "no regime": Where exactly did and where did you not include the "no regime" in your analysis? For instance, how do the "no regime" frequency anomalies before drought events look like in Fig. E1? Are there any interesting signals there? In principle, even in "no regime" periods you can have short-term extremes (for instance cut-off lows) that can yield significant amounts of precipitation and thus affect the occurrence or non-occurrence of droughts.

Reply: We added the zero regime to each of the figures. Its role was already taken into account in the reconstruction of precipitation anomalies. The frequency anomalies (see Fig. 5) associated with this regime are significantly different from zero but remain small.

However, "regime 0" pools together the transition states from one regime to another. If the regime signal is more pronounced in the period preceding droughts, i.e. if regimes are more frequent, we would expect to see shorter and rarer sequences of zero regimes. This translates into a zero or negative "regime 0" frequency anomaly, which is indeed what we observe.

> **Comment**
> Link to similar decomposition studies: A bit related to my first comment, have you considered to further decompose the weather regime frequency anomaly term as Fischer et al. 2025 have done (https://doi.org/10.5194/egusphere-2024-1253; see their Eqs. 3 and then 6), and from there, compute that ratio gamma that relates the contribution of the regime frequency anomaly/changes to the contribution from the intensity anomaly/changes? I haven't thought this through, so I'm not sure if it makes sense to apply this in your drought context (rather than in the climate change context), but it seems like it should be applicable too. Any thoughts on this?

Reply: Overall, we already make this comparison: in this article, the factor γ corresponds to a contribution ratio between two phenomena. In our article, we compare the factors α_1 and α_2, each expressing a contribution from two different phenomena. The factor γ in the article suggested here would broadly

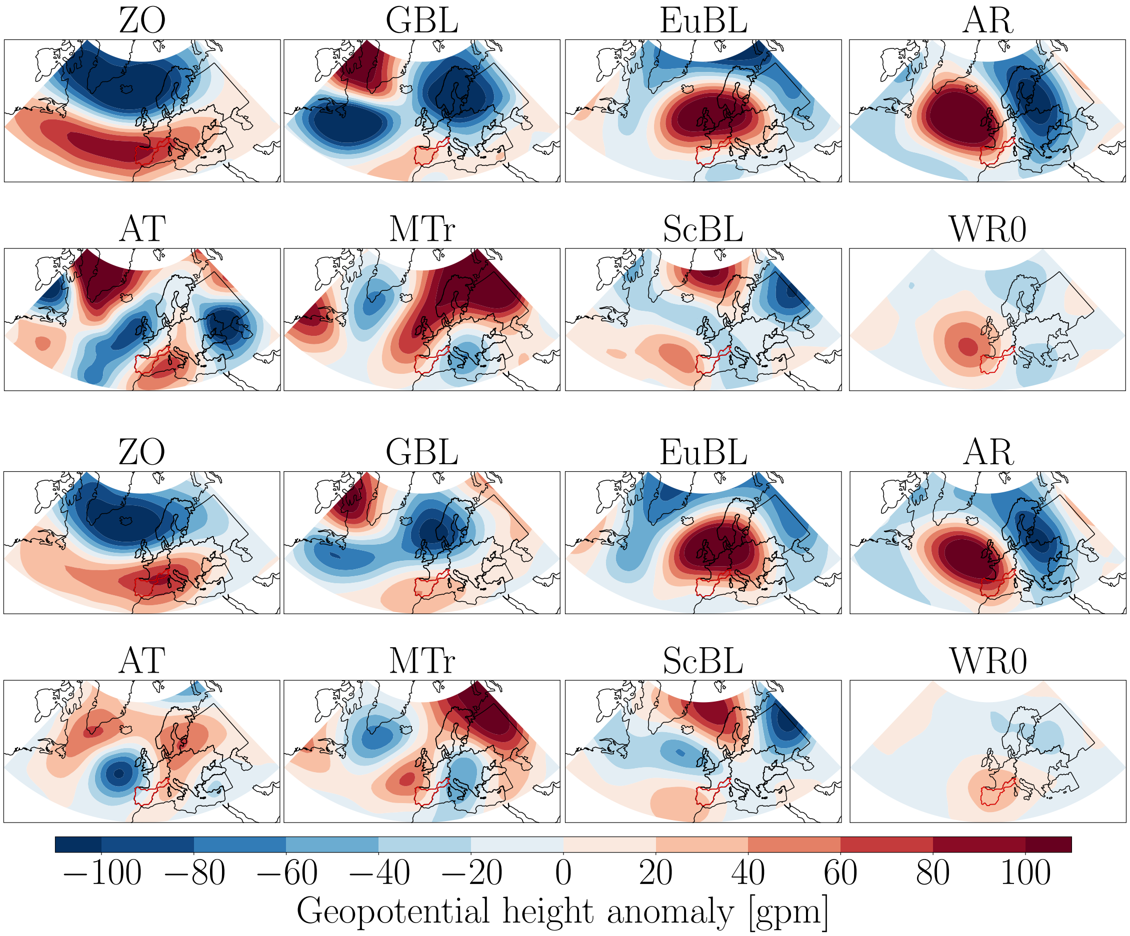

Figure 4: Z500 patterns for situations well explained by the WR approach (first two rows) and not well explained (rows 3 and 4).

correspond to calculating the ratio α_1/α_2. We could therefore do this, even though we are already doing it in a slightly different way by requiring this ratio to be greater than 0.33.
The main difference with this article is that the decomposition carried out in that study is performed on full-field variables and not on anomalies. We decompose precipitation anomalies. In particular, in their article the term $\phi^*_{\text{hist,i}}$ is the historical anomaly of the variable of interest ϕ_i. Equation 3b (see Eq.2 here) therefore presents the decomposition between the effect on precipitation climatology and the anomaly associated with the regime.

$$\Delta f_i \phi_{\text{hist,i}} = \Delta f_i \left(\phi_{\text{hist}} + \phi^*_{\text{hist,i}}\right) \tag{2}$$

Another major difference is that they compare a historical period with a future period, i.e. two separate periods. We, on the other hand, compare two periods, one of which is included in the other: the pre-drought period and the total period.

Comment
Quality of the text: Some specific parts of the manuscript need to be improved with respect to language / phrasing. Also, there are quite many typos (particularly in the second half) that can be removed when carefully reading the manuscript. As you will see in the minor comments, I corrected quite many of these little issues, but at some point, I stopped. One example is Section 4.1, which I find very hard to read and it took me quite some time to understand what the actual goal of this section is (especially the first paragraph there needs rephrasing).

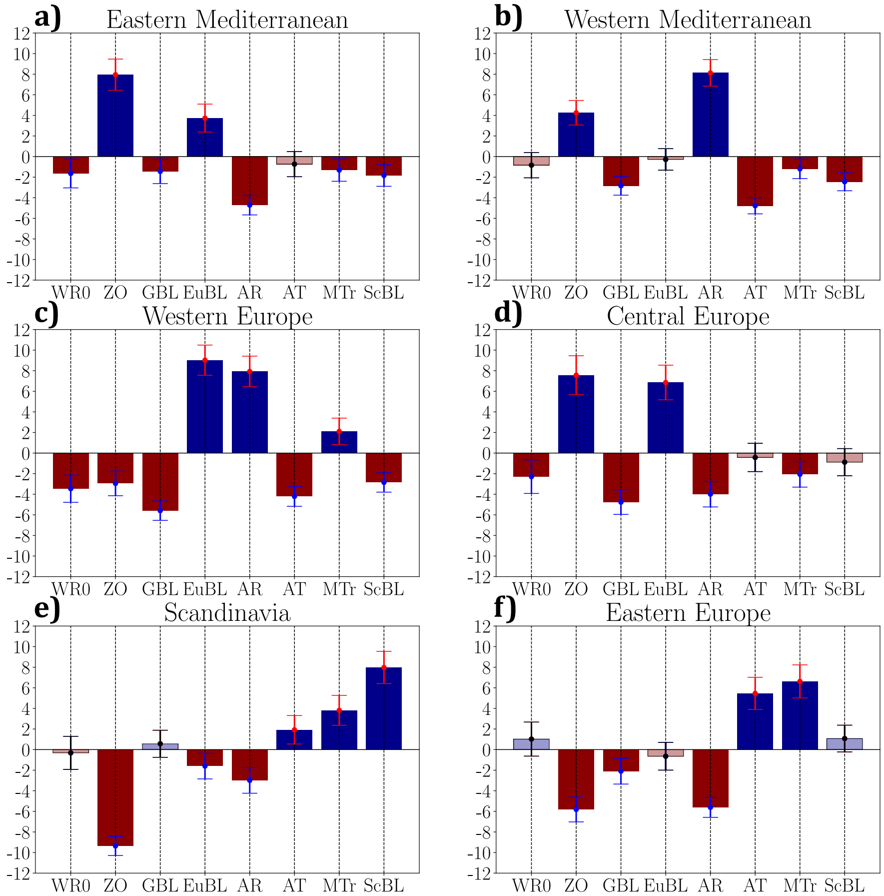

Figure 5: Frequency anomalies with WR0

Reply: Thank you for this comment. We carefully revised the parts of the paper that had wording issues, working to reformulate our reasoning as clearly as possible.

1.2 Minor comments

Comment
L15: ... consequences for society ...
L16: Maybe replace classically with typically
L17: Maybe better "depending on the physical variables they are associated with"?
L20: Maybe mention that this is mainly the case for the extratropics (the example with the winter drought).
L26: Rainfall where?
L29: ... which began between March and May and continued into June ...

Reply: Thank you, we have corrected or clarified this, lines X.

Comment
L34-36: I feel like there are two things mixed here and I would rephrase a bit: Yes, tropical teleconnections influence Europe less than, e.g., North America, but still, even in North America, rainfall anomalies are modulated by weather regimes (but these weather regimes are then modulated by teleconnections). So, I would not say it's either the teleconnections from the tropics or the local weather regimes, because it's rather both, but in one case, regimes are more driven by remote processes than in

the other case. Furthermore, also European regimes are driven by teleconnections, but for less from the tropics but for instance from the stratosphere etc

Reply: Thank you, we have corrected this accordingly with your comment and the same comment (see 2) from the R2.

Comment
L42: Maybe better "... are often defined separately for each season ..."?
L43: ... including the transitional ...
L62: I would add "daily mean precipitation and 500 hPa..."
L68: ... covers ...; also, isn't there a word missing after AR6?
L71: ... data at 500 hPa is determined ...; also, it sounds a bit odd, because it's not you "determining" this field – it's just a field in ERA5; do you refer to computing the anomalies?
L71-72: Is this the same domain as in Grams et al. 2017? Then maybe write this.
L83: Maybe add "... in terms of its climatological probability of occurrence ..." or something like this (i.e., the relation to the climatological distribution)?
L86: ... the SPI is calculated as the cumulative ... which is then subtracted from the long-term median cumulative precipitation ... : I'm not sure if my suggestion is proper, but I think you should somehow mention the fact that it's an anomaly

Reply: Thank you, we have corrected or clarified this in the next version of the paper..

Comment
L86: Does the cumulative 3-month window end on that day or is it centered over that day of reference? I think that's an important detail

Reply: The cumulative 3-month end on that day. Thank you for your comment, we corrected it.

Comment
Section 2.2.2: This is a quite nice way of proving the usability of the SPI in different regions. But given that your conclusion is that it's usable over most of Europe (i.e., your region of interest), I wonder if this section really needs to be in the main manuscript or could be moved to the supplement? I suggest this also because the manuscript is quite long already and includes many different methodological steps the reader has to get familiar with...

Reply: The main interest of this section is to raise an issue that is rarely addressed in studies using SPI : the fact that SPI is not a suitable tool in excessively arid regions. The second objective is to explain the reason why we excluded the North African coast from our study, even though it is initially part of the MED region as defined by AR6. We agree that the paper is already long, so we moved this discussion to the appendix, as we believe it is important but not central to our study.

Comment
L105: I wouldn't say you use "two distinct methods" – you rather compute two things but look at them together, so it's rather one approach you use.
L106: ... consists of ... Fig. 1: The hatching is too strong such that the underlying colors can hardly be seen – please change this.
L119: ... SPI3 time series ...
L120: Make sure you don't mix the tense in writing. In the previous sentence you write in present tense, and in the following in past tense.
L134: Section number is missing in brackets.
L136: What is AR6? Is that a region?
L141: Maybe write something like "spatio-temporal simultaneity" – which is the case, right?

Reply: Thank you, we have corrected or clarified this in the next version of the paper..

Comment
L155-156: I'm not sure I understand this last step with the seasonal variance. What exactly do you mean about judging the "relevance of this decision"? Which decision? Is this step the one that decides when to stop the creation of further regions? And aren't there various objective metrics to decide how many clusters to retain after a clustering (such as similarity indices etc.)? Couldn't you apply one of these?

Reply: The decision we are referring to concerns the number of study regions. It effectively refers to the point at which the creation of new regions is halted. The expression 'judging the relevance of this decision' refers to the fact that there is no single solution to our problem. This complicates the decision and therefore warrants a prior assessment of its relevance. The classic criterion for hierarchical clustering is the similarity criterion. We also wanted a reasonably small number of regions, in order to limit the dimensionality of the study. Too many regions would reduce the interest of the study. Figure 6 shows the dendrogram associated with this clustering. The similarity criterion suggests "cutting" the dendrogram at the number of clusters

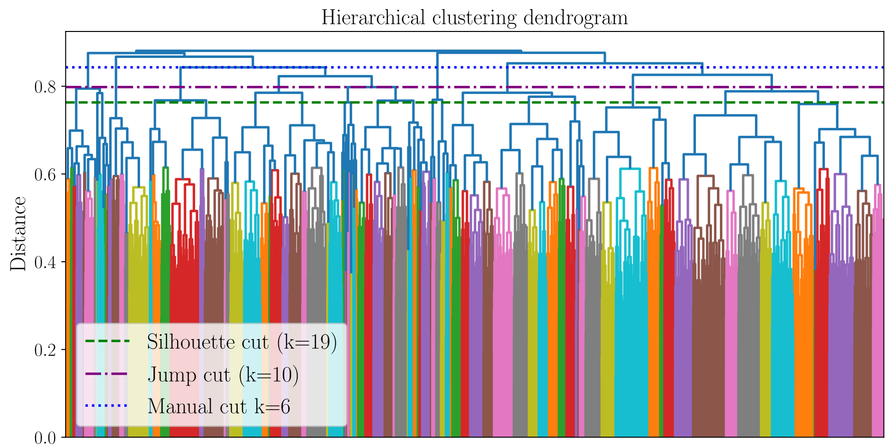

Figure 6: Dendrogram

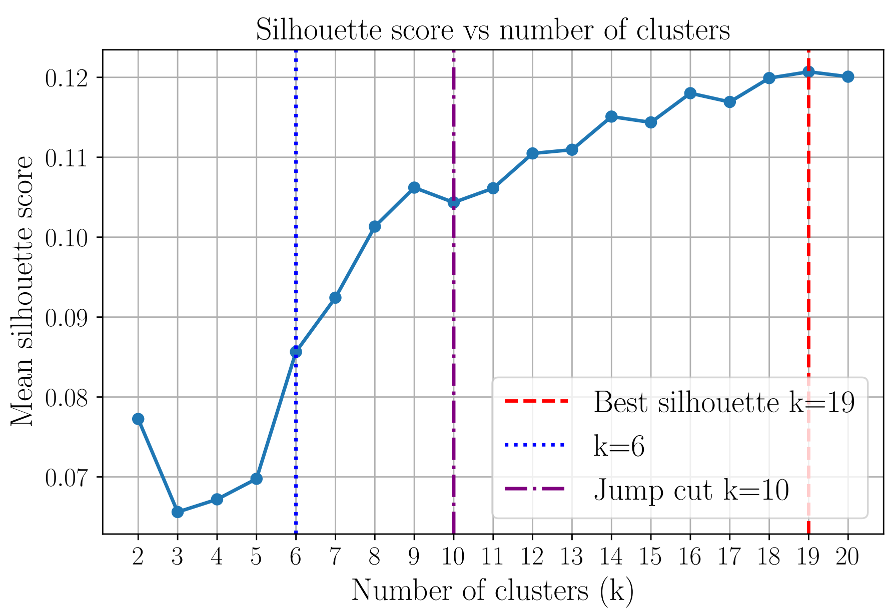

Figure 7: Silhouette scores

where the jump in this index is maximal. Applying this criterion, we obtain $k = 10$ clusters, which is too high. Another criterion is the silhouette index. We show in Fig. 7 the silhouette index values for an increasing number of clusters. The silhouette criterion suggests choosing $k = 19$ clusters, which is also too high.

In addition, we carry out this regionalization so that spatial averaging does not erase the SPI signal. We therefore want regions where the probability of drought occurrence remains the same for a given threshold, whether the analysis is done at the grid-point scale or after spatial averaging. This property is preserved if the grid-point SPIs within the same region are correlated, which the UPGMA algorithm ensures. We can check that this approximation is valid by verifying whether the properties of the SPI are preserved after spatial averaging. One of the property of SPI, linked to the probability of occurrence of an event for a fixed threshold, is the unit variance of SPI.

The number of clusters that optimises the criteria varies depending on the criteria chosen, so there is no single solution to the problem and no particular region seems to be more relevant than another. We therefore prioritise keeping the number of regions small, so we choose 6 regions.

Initially, we wanted a limited number of regions in Europe, such as the reference regions for Europe in the IPCC. The use of objective criteria such as those presented above does not allow us to refine this choice, as they propose too many regions. We cannot therefore rely on these criteria. However, given that the silhouette score shows a significant jump between a number of clusters of k=5 and k=6, choosing a number of 6 regions is consistent.

To go further, we also justify (see comments 1 and 2 from Reviewer II) that this regionalization is stable for different sub-periods, and including or suppressing one region does not change the critical regions (i.e. regions where the results obtained are good).

Comment
L162: "significantly equal to 1" – can you really say that? At least I never heard it like this.

Reply: Thank you, we rewrote it with "not different from 1 with a 95% confidence level"

Comment
L163: Are there examples of other regions / clusters in the literature, either related to drought or something related? If yes, it would be nice if you compared yours to them in a few sentences. I am not aware of any, but are your regions, for instance, similar to the heatwave clusters Stefanon et al. 2012 or Pyrina and Domeisen 2023 found?

Reply: Examples of regionalisation using this method can be found in the literature. Notably Pappert et al. [2024], which directly inspired our method. Although they apply this method to extreme temperature data, they obtain regions that are quite similar to ours. In particular, their region 1 corresponds to our region WMed, their region 2 corresponds to our region WEur, their region 3 is slightly smaller than our region CEur, and their regions 5 and 6 correspond to our regions EEur, while their region 4 corresponds to our region Scand.

These two studies Stefanon et al. [2012], Pyrina and Domeisen [2023] indeed proposed regional analyses. However, they did not propose a physics-informed regionalization. They applied clustering directly on their impact data (equivalent to precipitation or SPI3 in our case). The different centers of action that appear from one centroid to another suggest the need for regionalization. Nevertheless, the centers of action are generally located over the regions we use.

We can directly link the regions we found to those affected by the clusters in the proposed studies. In Stefanon et al. [2012], the clusters RU, WE, EE, IB, and SC correspond respectively to our EEUr, WEur, CEur, WMed+EMed, and Scand regions. However, their NS cluster overlaps our Scand and WEur. Similarly, in Pyrina and Domeisen [2023], the clusters BSea, Ru, Sc, EEu, WEu, NSea correspond to our EMed, EEur, Scand, CEur, WEur+WMed regions, again with some overlap. The different approaches to regionalize the

European continent, even if applied to different types of extreme events, lead to a fair degree of similarity in the resulting regions

Comment
L166: How relevant is this 10-day threshold for your conclusions? And aren't single days with SPI3 below the threshold already extreme as well (given that they are already based on long 3-month windows), so isn't this threshold a bit too strict? How often do these shorter-than-10d-events occur at all?

Reply: We aim to study persistent droughts. In most studies Spinoni et al. [2019, 2014], SPI is computed at a monthly rather than daily frequency, and droughts considered persistent are those lasting more than 2 months. Our persistence threshold, by comparison, is actually quite low. Furthermore, reducing this threshold (i.e., including very short droughts of a few days) produces too many events, which act as noise in our data. As an illustration, with no threshold, we obtain :

- EMed : 64 droughts
- WMed : 56 droughts
- WEur : 46 droughts
- CEur : 43 droughts
- Scand : 50 droughts
- EEur : 35 droughts

Comment
L168: Maybe replace "characteristic date" with "time stamp" or "event date"?

Reply: Thank you, we have corrected or clarified this in the next version of the paper..

Comment
Section 2.2.5: What was the reason that you did not directly use the 7 regimes from Grams et al. (2017), but instead redefined them with a slightly adapted domain? Why did you use this different domain? Is it to include more of Eastern Europe (which would be good reason)? In any case, I think it would be good to briefly discuss these choices and also the differences of your regimes compared to the Grams regimes, because the methodological steps you perform seem to be very similar to Grams et al. (2017).

L173: Is there a reason you only used 1990-2020 for defining the regimes, rather than going further back to get a more robust picture? Also, wouldn't going back further make the regimes more similar to the ones in Grams et al. 2017 (which might be something desirable)?

L207-210: My gut feeling would tell me it is mainly the different domain you used, which made the difference to the Grams regimes, probably followed by the different time periods. I can hardly imagine though that ERA-Interim vs. ERA5 makes a difference, given that regimes are defined on this very coarsely smoothed field Z500.

Reply: We deliberately applied exactly the same procedure as Grams et al. [2017], except for the domain. We wanted a domain covering our full study area, including the EEur region. Following this comment, we will justify this choice in the next version of the paper.

We agree that using a longer period would probably yield regimes closer to those of Grams et al. [2017]. However, the assumption of a linear trend decline is more robust over 30 years, whereas this is less the case over a longer period.

Comment
What exactly is the criterion for deciding if a regime is cyclonic or anticyclonic? Please specify. Also, it's a bit unfortunate that the Mediterranean Trough regime is a blocking regime, although the trough over the Mediterranean seems to be the key feature (according to your interpretation). Regarding that MTr regime: It's interesting that this appears so distinctly in your regime definition (compared to Grams et al.). Do you think it comes from extending your domain much more to the east, which allows this blocking in northeastern Europe to appear, probably favoring the stationarity of a trough (probably associated with cut-off lows or just Mediterranean cyclone activity) in the Mediterranean? Maybe you can discuss this "additional regime" briefly?

Reply: We wanted to name our weather systems according to the centre of action with the highest intensity. Following this logic, MTr could be seen as a flavor of Scandinavian Blocking because of the north-eastern anticyclone, but this term was already taken for the last regime, which was more similar to Grams et al. [2017] regime. Out of respect for the original publication, we chose to name this regime after the second notable feature: the Mediterranean trough.

The major difference between Grams et al. [2017] and us lies in obtaining a ScTr regime, which we do not have (having the MTr regime instead). Grams' ScTr regime does not resemble our MTr regime. However, our MTr regime has an anticyclonic anomaly in the east of the domain, extending over Scandinavia and the United Kingdom. Its ScTr closely resembles its AR, with a high and low pressure tripole, with a shift in the centre of action. In this respect, our two AT and MTr regimes also present a tripolar pattern with a shift in the centre of action, while its AT is more similar to a zonal regime. The correspondence between Grams' regimes and ours is therefore not immediate and depends on the details of the location of the centre of action. All the differences between Grams' diets and ours are classification issues and may explain the emergence of a completely new diet. Furthermore, this classification is extremely dependent on the domain. A restricted domain tends to produce a dipolar field during PCA, whereas an extension of the domain is more likely to produce a tripolar field, which is what we observe with MTr. Hence, we believe that it is indeed the extension of the domain towards the east that allows this stationary structure to emerge.

Following this comment, we added the precision concerning the name of the WR in the next version of the paper. and we added the precision "We obtain a new WR "Mediterranean trough", characterised by an anticyclonic anomaly in northern Europe and a low over the Mediterranean. This pattern probably arises from the extended spatial domain to the east, which allows a zonal wave structure to appear over eastern Europe and thus a stationary anomaly over the Mediterranean." in the next version of the paper..

Comment
L207: I would mention Grams et al. first here, because this was the original definition.

Reply: Corrected.

Comment
211-212: I would personally find it more useful if you sorted the weather regimes according to some similarity principles or according to cyclonic/anticyclonic, rather than frequency of occurrence. But I guess that's a matter of taste and I'll leave it up to you

Reply: Thank you, we wanted to show first the predominant one since our study focuses on the frequency anomaly.

Comment
L223-225: You should be more specific here: the blocked regimes are not associate with negative P anomalies everywhere, but rather in the high-pressure regions, while there are important (and impactful) positive P anomalies as well! And vice versa for the cyclonic regimes

Reply: Thank you, we have clarified this in the next version of the paper..

Comment
Fig. C1: Interesting to see these patterns! Just to be sure, I assume you computed these seasonal anomalies with respect to seasonal climatologies? I guess that would be important as you observe and discuss stronger magnitudes in winter.

Reply: Thank you for your comment; we add this clarification in the paper. We indeed computed it with respect to seasonal climatologies.

Comment
L235: "particularly in the December-January-February (DJF) season" is not really needed because you already say before that it's the winter months in which the magnitudes are larger.
L240: I would stick to "weather regimes", because "weather patterns" is sometimes used for shorter-lasting / higher-frequency circulation anomalies than weather regimes.
L241: What does this specifically mean, 91 days prior to the drought? Since you look at SPI3, i.e., 3-months events, this should be clarified.
L251: ... shown in Fig. 5 ...
L262: Better to write "Similarly, on the drought-preceding periods S" or something similar. Eq. 9: Isn't there the sum operator missing before the last term, $f_k * C_k$ (or, alternatively, brackets around everything)?
L277: ... characterizes ...
L280: ... illustrated in Fig. 6 ...
L284: Which terms do you refer to? And over which regions are these spatial averages? Please specify in the text.
L291: I would not use the term "satisfactory" here, but something more objective...
Table 3: Can you also include the "no regime" here? Also, I personally like Fig. D1 better than this table here – could you consider just using the figure in the manuscript instead of this table?

Reply: Thank you, we have corrected this in the next version of the paper..

Comment
L297: How do you get to this number of 50 drought events? Isn't the number of events different from region to region, as shown in Table 2?

Reply: This was a typo; thank you for catching it.

Comment
L301-302: I don't understand this sentence about EMed. Do you mean it's the region in which the highest number of "NS" appears in Table 3? However, is it not a bit misleading to count the number of NS? Because it still seems that EMed is influenced by weather regimes, but by chance just by one (the Zonal Regime), but relatively strongly. So, I don't think the "number of NS" is a meaningful measure here for saying how strongly a region is influenced by weather regimes. You see my point?

Reply: Indeed, that sentence was unclear. We agree that it is misleading to count the number of NS. We corrected it.

Comment
L307: ... presents ... (instead of is presenting)
L309: What do you mean with "then the amplitude of the signal decreases with the distance"?
L311: Isn't it better to write "climatological precipitation" than "canonical precipitation"?
L316: Replace "coast" with "North Atlantic"
L317-318: "accurate" with respect to what / comparing to what? Please specify.
L321: Remove "easily"
L341: You repeat yourself here with the statement that you did this analysis for all regions Table 4: Please specify in the caption what exactly the middle and right columns show.
L350: ... are the ones in which droughts are the most influenced by ... (and same for the sentence afterwards)

L351: ... by the large-scale atmospheric circulation in the North Atlantic ...

Reply: Thank you, we have corrected this in the next version of the paper..

Comment
Fig. 8: Please explain this figure better in the caption / text, because it's very confusing. What do you mean in (a) with "for all cases where droughts cannot be explained with year-round weather-regimes"? Aren't just all drought events shown in (a), i.e., those that can and those that cannot be explained by weather regimes according to your definition? Also, the explanations for (b-d) are confusing. In the text you write "Panels a) and b) in Fig.8 show the mean of the precipitation anomalies and the mean of the reconstructed signal, in cases explained by anomalous WR frequencies. Figures c) and d) show the cases that are explained by anomalous WR frequencies." Isn't this twice the same?

Reply: Thank you for your remark. This was a caption inversion, which we corrected.

Comment
L360: "faithfully" is a strange term to be used in this context, in my opinion. Fig. E1: Since it's flipped, you should change "left" and "right" in the caption.
L370: What is the "WE contribution"? Figure 9: You don't really introduce this in the text, but you suddenly discuss Fig. 9b. Probably you missed it?
L380: Can you refer to the figure again where one can see the frequency anomalies of all droughts (to which you compare)?

Reply: Thank you, we have corrected this in the next version of the paper..

Comment
L386-387: What exactly do you mean here with "without taking into account the above-mentioned differences concerning the heterogeneity of the reconstruction quality"?

Reply: This sentence was misleading; we removed it.

Comment
Figure 10: I think it's confusing that you now use red for those droughts that are not explained well by WRs, while red in the figures before was the term that stands for the importance of WR influence

Reply: Thank you, we have corrected this (green for well explained droughts, grey for not well explained).

Comment
Section 4: I assume the importance of the regimes for droughts also depends a bit on the set/definition of regimes you use. I would assume that they become more important if you went to a higher number of regimes, because they would explain more of the surface weather variability. Do you agree and can you maybe discuss this briefly in the discussion? However, I don't say it would make sense to use a higher number, because in terms of predictability, a lower number of regimes is generally more useful.

Reply: We agree that increasing the number of weather regimes would refine the description of the atmospheric circulation. However, even if increasing the number of regimes leads to a reduction in intra-cluster variance, it also leads to an increase of the inter-cluster variance. This reduces the robustness and stability of the regimes, as the attribution of a daily situation to a regime is less straightforward. We estimate that we would lose too much of the benefit of the WR approach which conveniently provides a description of the atmospheric circulation with a reduced set of attractors.

Comment
L389: "in contrast to" rather than "with respect to"

Reply: Thank you, we have corrected this in the next version of the paper..

> **Comment**
> This is a cool result! Could convective activity (or non-activity) and maybe soil-moisture-related feedbacks on precipitation during summer also contribute to the fact that WRs explain the droughts less well in this season (since convective activity is not so directly linked to the low-frequency large-scale circulation)?

Reply: We do indeed believe that this may be the cause. To be sure, we computed the standard deviation of the distribution of precipitation maps for each regime during droughts and for each season, normalized by the precipitation standard deviation (regardless of the regime or the season). The standard deviation is highest in summer, which confirms that our method (which relies only on mean precipitation) is less effective in summer. Winter variability is highest around the Mediterranean. The higher standard deviation around the Mediterranean shows that weather in these areas is mainly driven by the Mediterranean rather than by North Atlantic circulation.

Finally, to answer the question: we think that the higher standard deviation in summer across Europe

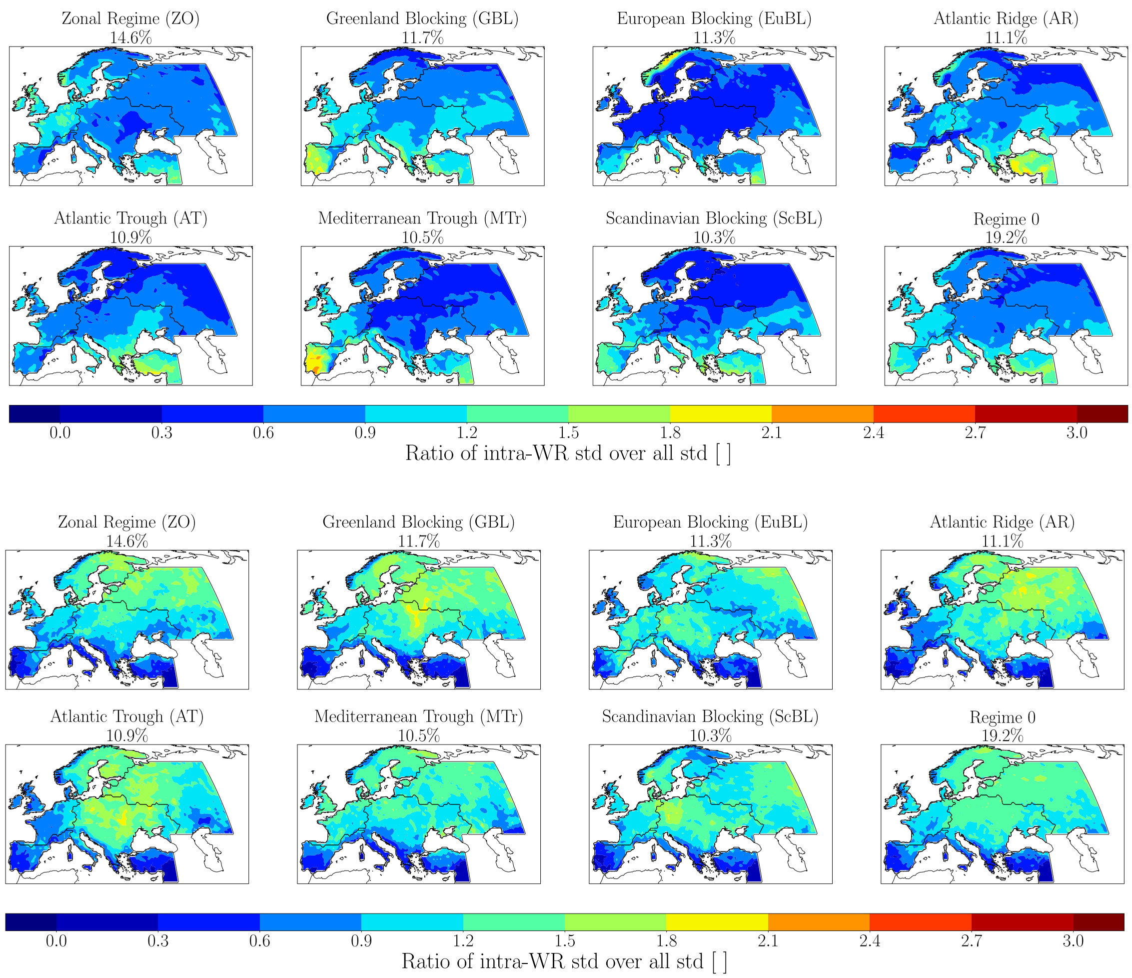

Figure 8: Standard deviation of precipitation (in mm), for each WR, during winter (DJF) and summer and (JJA)

reflects a predominance of convective precipitation, which accompanies small scale disturbances.

> **Comment**
> Figure H1: What exactly is the difference shown here? It's a bit hard to understand with the given

explanations. Can you maybe refer to the other, previous figures that "flow into" this figure?

Reply: For a given region, droughts are assumed to be evenly distributed throughout the year. In other words, we are not supposed to have more drought in winter than in summer. When reconstructing precipitation patterns, we use canonical patterns of regimes, i.e. average precipitation when regimes are active, regardless of the season. The precipitation patterns we use are considered to be invariant throughout the year. In reality, these precipitation patterns vary depending on the season (see Fig. C1). To be more accurate in our reconstruction, we should therefore carry out a reconstruction using seasonal patterns such as those shown in Figure C1.

Thus, as rainfall patterns change with the seasons, we make an error when we use the average pattern for the year. The error made depends on the seasonality of droughts for a given region. Thus, for each rainfall pattern and each season, we obtain a residual rainfall that depends on the seasonality of droughts in that region.

It is possible to calculate the average residual error made for each region by applying the decomposition method to the residual precipitation patterns. We present the results Fig. 9. This figure is the same as in the article, with the addition of a row showing the average residual error made for each region by applying the decomposition method to the residual precipitation patterns, and masking the regimes for which the frequency anomaly is not significant.

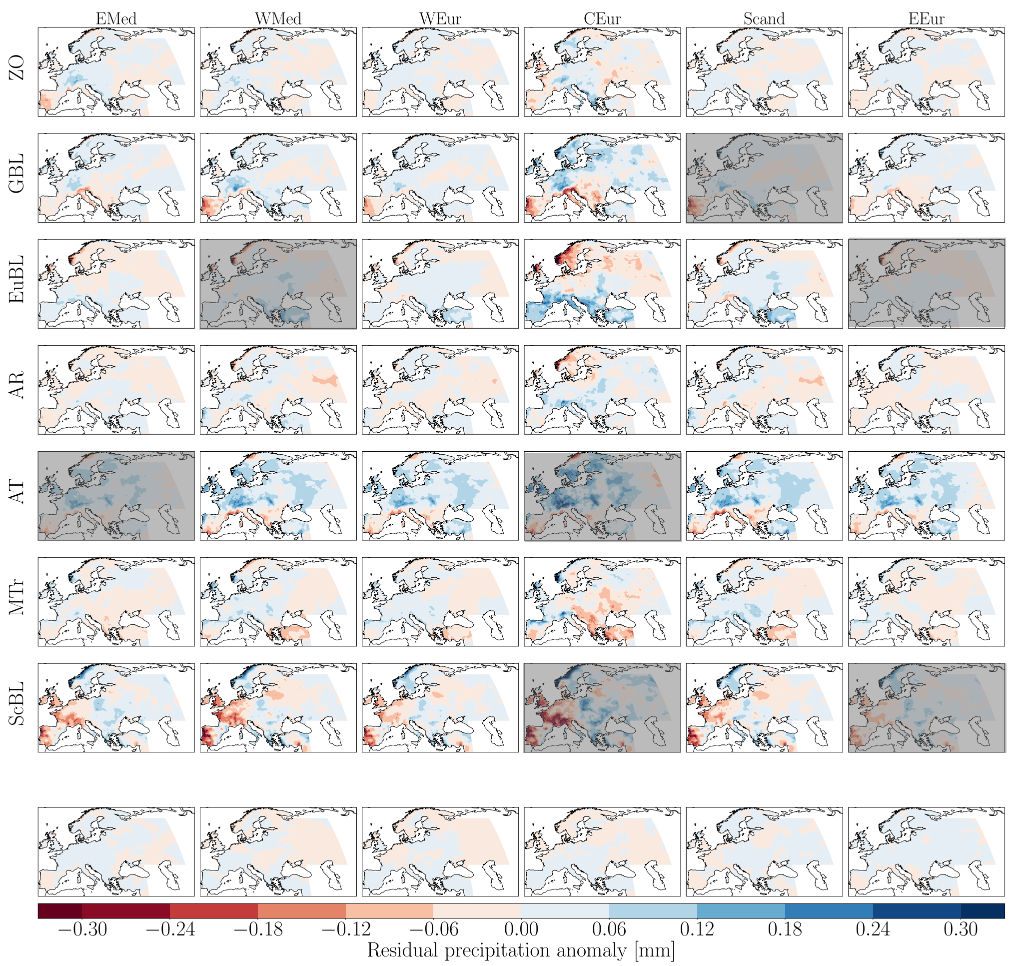

Figure 9: Residual precipitation anomaly after reconstruction due to the seasonnality in precipitation

We have replaced the figure in the article with this one, added the explanations we have just given in this comment, and added references to the figures used to create it.

Comment
L427: Why exactly do you explain the AT-WMed example of Fig. H1? I don't see a very strong difference/signal in the Western Mediterranean in this example, but much bigger magnitudes in other subpanels.

Reply: We wanted to describe the negative anomaly over the northern Mediterranean basin, common to all subpanels of the AT line. The use of the term "Western Mediterranean" was not appropriate here, since it refers to a specific region that we defined differently in the rest of the paper. We replaced it with "northern Mediterranean region."

Comment
L433: The last sentence doesn't make sense – you just did explain the ScBL!
L439-440: "Hence, we can expect from this region to present the more heterogeneous pattern of droughts, thereby exhibiting a limitation of the choices to keep only 6 regions during the regionalization method." – This is a heavy sentence to read and I'm not sure I understand. Can you simplify?
L445: I would either directly write "weather regimes" here or then specify "low-frequency large-scale North Atlantic circulation", because you don't just look at the instantaneous large-scale circulation.
L448: What do you mean with "original regionalization method"?
L451-452: At various places in the manuscript – and here – you use "large-scale circulation" (or simply "circulation") as a synonym for weather regimes, but this is not true. The large-scale circulation likely influences every drought, but the regimes apparently don't (for instance, a single convective event can be strongly driven by the instantaneous large-scale circulation but not necessarily by a regime). This is because the regimes are a categorization of the lower-frequency circulation and are persistent states. Please rephrase this a bit to avoid mis-interpretation by the reader.
L460: Not sure you introduce ESM before...

Reply: Thank you, we have corrected this in the next version of the paper..

2 Reviewer #2

Comment
Droughts are defined as the negative exceedance of a chosen SPI index averaged over defined regions of Europe, which are derived from drought concurrence. However, the weather regimes project onto opposite anomalies in some of these regions, even when analysed over the entire year (see Figure 5: EuBL over Scandinavia, AT over West Europe, ScBL over Scandinavia and the Iberian Peninsula). This raises the question of how much signal is lost due to this specific regionalization of Europe. To investigate the sensitivity of the regionalization, the reviewer suggests to investigate the sensitivity to subsampling the time period and analysing the consistency of the identified regions, as well as showing the regions identified when selecting 5 or 7 regions in the Appendix.

Reply: The proposed regionalization indeed leads to regions where the impact of weather regimes is not spatially homogeneous. A single regime—such as those given as examples in this comment—can have opposite impacts within the same region. These results should appear in the regionalization: a regime with a heterogeneous pattern indicates an anti-correlation within a region, with precipitation anomalies of opposite signs when that regime is active. As suggested, we performed clustering for 5, 6, and 7 clusters. The results are shown in Fig. 10. The regions presented in this study result from a split of the dark red region, forming the WEur and CEur regions. Increasing the number of clusters further leads to an North-South division of the EEur region. This shows strong stability in regions where contrasts within weather regimes exist.

This stability is likely explained by the difference in temporal scale: precipitation anomalies linked to weather regimes are computed daily, while SPI3 can be viewed as a 3-month moving average. Two time series within the same region may be anti-correlated at high frequency, but the SPI3 calculation produces

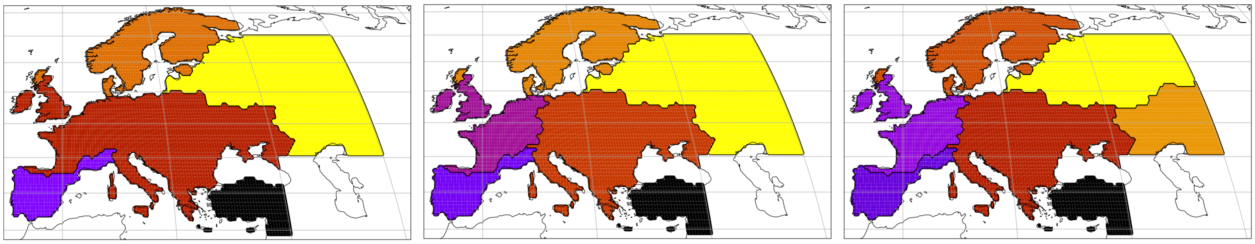

Figure 10: Regionalization for [LEFT] 5 clusters [CENTER] 6 clusters [RIGHT] 7 clusters

correlated series at low frequency.

The article originally lacked a clarification on the period used for SPI computation; this has now been added. Following this review, we also performed clustering on two disjoint sub-periods to test the temporal stability of our results. Figure 11 shows the regionalization for 1960–1990, 1991–2020, and for the full period. The EEur, EMed, and Scand regions are very stable, but there is greater uncertainty over Western Europe. In particular, the CEur region splits into two in both sub-periods. Regionalization is therefore sensitive to the sub-period considered in Western Europe, but much more robust in the East. We continue this discussion in our Reply to the next comment.

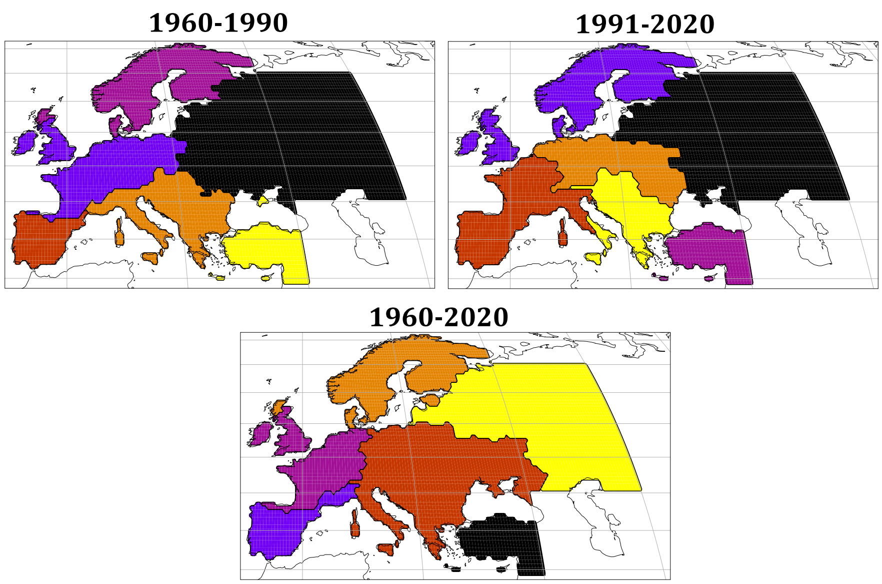

Figure 11: Regionalization for two different subperiods and for the entire period

Comment
Furthermore, other approaches to regionalization could be discussed and optionally investigated—such as clustering drought occurrences, and emerging approaches to identify circulation patterns targeted to explain specific impacts could be mentioned or discussed (Bloomfield et al. 2020, Rouges et al. 2023, Spuler et al. 2024, 2025).

Furthermore, as discussed in section 4.2 and shown in the Appendix (Fig C1), weather regimes can have different precipitation signatures depending on the season which is quite significant for some regions, regimes and seasons. However, the year-round analysis conducted here averages out this signal (see further comments on this in the next bullet point)

Reply: Several other regionalization approaches can indeed be considered. Regionalization is not central to our paper, and the analysis can be carried out regardless of the method used. Slightly different regions would not have drastically changed the results of the article. The method proposed here aims to define regions that can be considered "reference regions" for studying droughts. The primary goal is to reduce the dimensionality of the study and allow regional conclusions. They are not designed to perfectly match the impacts of weather regimes. However, as suggested in this comment, it is possible to design regions specifically for our study. We discuss some options below.

One option would be to avoid regionalization at the start and continue working at the grid-point level. For each grid point, we would identify drought periods and decompose Eq. 11 (see 3) averaged over each drought. This yields a spatial scalar field on which clustering would produce regions with homogeneous impacts of regime frequency anomalies.

$$1 = \underbrace{\frac{\left\langle \sum_k \Delta f_k C_k \right\rangle}{\left\langle \overline{p}^S \right\rangle}}_{\alpha_1} + \underbrace{\frac{\left\langle \sum_k f_k \Delta C_k \right\rangle}{\left\langle \overline{p}^S \right\rangle}}_{\alpha_2} + \underbrace{\frac{\left\langle \sum_k \Delta f_k \Delta C_k \right\rangle}{\left\langle \overline{p}^S \right\rangle}}_{\alpha_3} \tag{3}$$

Another option would be to compute, at each grid point and for each drought, the regime frequency anomalies during the 91 days preceding the drought. We could then perform clustering by placing all grid points into the seven-dimensional space of regime frequency anomalies. This would produce regions homogeneous in terms of regime frequency anomalies.

Finally, this comment suggests discussing new approaches for identifying links between circulation anomalies and specific impacts. The three articles mentioned present emerging methods. In Bloomfield et al. [2020], classification is performed on zg500 but also on demand variables. In our case, this would correspond to computing weather regimes and associated precipitation composites, then computing regimes from precipitation and looking at associated circulation, or computing regimes directly on SPI. Figure 12 shows clustering performed after PCA and k-means directly on SPI data. This shows the most frequent characteristic states of the SPI. Different types of drought can be observed, but the precipitation patterns characteristic of certain weather regimes can be found in the SPI. This breakdown of the SPI can be used to characterise our droughts. However, the SPI has a 'memory' effect due to its calculation over a cumulative 90-day period. The corresponding zg500 composites are probably more variable: from one day to the next, the SPI3 varies very little due to its 90-day rolling cumulative average, so several consecutive days in the same regime may present very different zg500 anomalies.

We also observe that 'SPI regimes' 2, 3, 5, 6 and 7 present droughts centred on a single region as defined: regime 2 dries out the WMed and EMed regions (even if this is actually more a trace of the NAO+ regime), while regime 3 dries out the WMed region only. Regime 5 dries out Scand, regime 6 dries out EEur and regime 7 dries out WEur and CEur. This approach demonstrates the relevance of our regionalisation. To construct

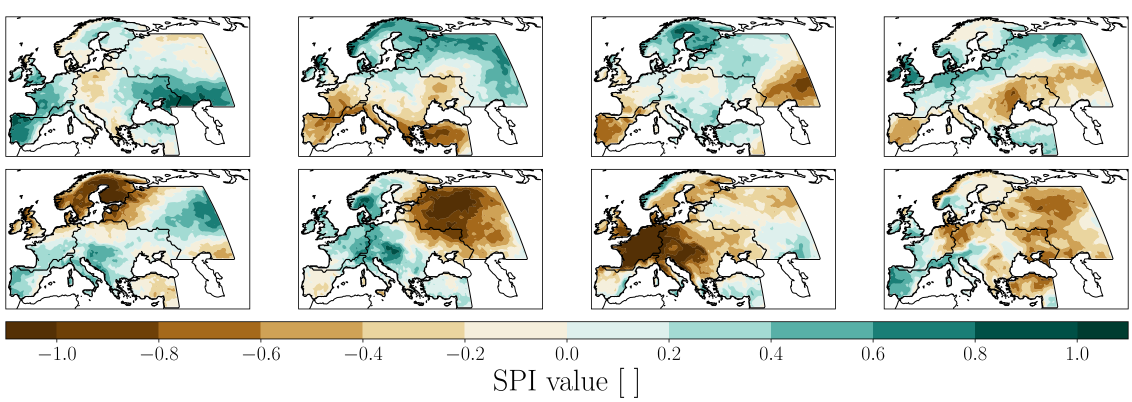

Figure 12: Weather regimes computation methodology applied to SPI data

circulation patterns targeted to explain specific impact, it is possible to perform a Maximum Covariance Analysis (MCA) to directly extract covariant circulation and precipitation patterns. We can then go further and apply the same regime construction method, clustering the zg500-PCs of the MCA (instead of PCA). By subsequently applying the same criteria as for WR construction (persistence, distance, explained variability, etc.), we classify each day into a regime. This allows us to plot new zg500 and precipitation composites and propose a new set of weather regimes designed specifically to maximize the circulation–precipitation link. Some of these new regimes show similarities with the year-round regimes studied in our work. We present these new regimes in Fig. 13. A complementary study would be needed to validate the relevance of these new regimes for drought analysis.

Thus, regardless of the method used, we obtain broadly the same regimes in zg500. The proportion of

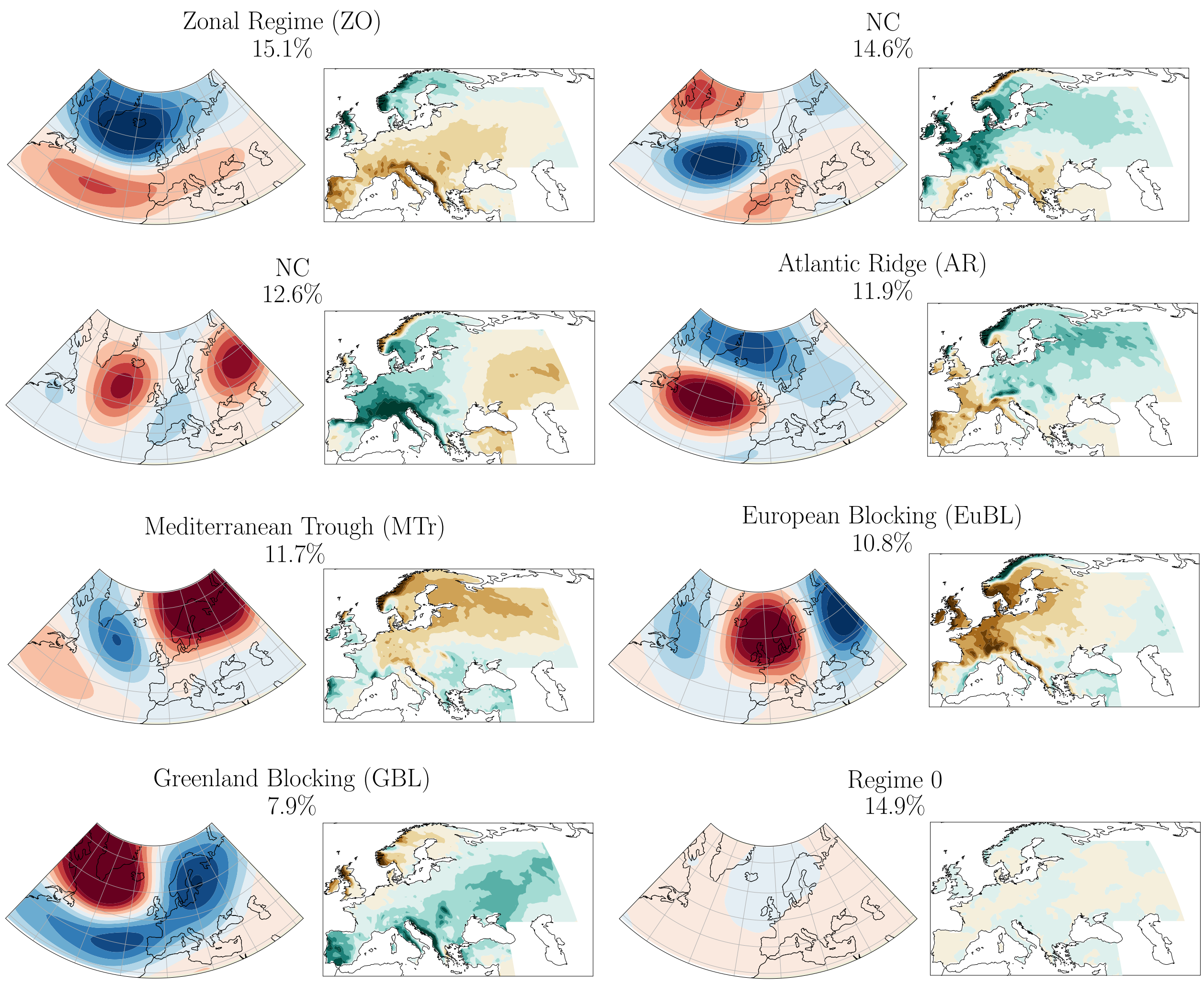

Figure 13: Seven new weather regimes, designed to maximize the link between atmospheric circulation and precipitation

droughts driven by zg500 is therefore fairly well represented in our article, and the sensitivity of the results to the classification method is therefore fairly low. We will add a section on the MCA approach to the discussion section of the article, as we believe that this is the approach most likely to yield satisfactory results.

Comment

The paper, including the abstract, emphasizes the added value of using year-round weather regimes ”as a unifying framework for understanding drought drivers, overcoming the constraints of purely seasonal classification" (Line 11). However, from the existing text, it is not quite clear what the benefits are of a year-round classification. There are significant differences in the precipitation impact of regimes in different seasons, as discussed later in the paper which reduce the potential explanatory power of regimes for drought conditions, but aside from the obvious avoidance of splitting the data into categories, no clear benefit of the year-round definition is discussed in the paper. If this is presented as a key advantage of the paper, the benefits of the year-round definition should be more clearly discussed.

Related to the point above, in line 42 the authors write that WR defined by season hinders a systematic analysis of the drought–circulation relationship throughout the year. This is not quite true; it is just that the analysis would have to be conducted season by season. This statement should be explained in more detail to be included.

Reply: The main advantage is that they reduce the dimensionality of the problem, as the study does not need to be repeated separately for each season. Similarly, classic weather patterns are well established for the winter (DJF) and summer (JJA) seasons, but not for the intermediate seasons (MAM and SON - spring and fall). Finally, the phenomenon under study – droughts – can occur at any time of the year. Any division of the year into sub-periods will most certainly lead to discontinuity in the study of the link between zg500 and drought. Following this comment, we will clarify in the paper the actual advantages of year-round regimes compared with seasonal regimes.

Comment
There is no weighting of the data with the cosine of the latitude mentioned to account for the different area that grid cells at different latitudes cover. This is common practice when dimensionality reducing circulation, in particular when computing weather regimes. Is there a particular reason that this was omitted here? For the weather regimes shown in Figure 4, in particular MTr and ScBL, the center of action is quite far North - if the data was indeed not weighted, this might be a possible reason. This reviewer would suggest trying out the computation of regimes including the weighting, or include convincing reasoning for not doing so.

Reply: We thank you for this comment. The data are indeed weighted by the cosine of latitude. The function *EOF* from the *eofs.xarray* library takes the cosine of latitude as an argument and therefore provides modes computed on weighted maps. The k-means algorithm then clusters these modes. The article will be modified subsequently.

Comment
The authors write that the z500 data was detrended between 1991 and 2020 prior to computing the regimes; however, for the full period, 1960– 2022, no detrending is mentioned. This would be inconsistent and likely lead to artificial trends in the regimes. Is there a particular reason that this was omitted?

Reply: This clarification has been added to the paper. The trend is removed over the 1991-2020 to compute the WR. Once this is done, we remove the same trend (1991-2020 trend) over the 1960-2022 period.

Comment
Section 2.2.2 – Is there a particular reason that a Gamma distribution is fit to the data, of which the skewness coefficient is then calculated? Why is the skewness not directly estimated from the data, circumventing the uncertainty around the distribution fit?

Reply: The proposed solution—to compute skewness directly without fitting a Gamma distribution—is valid. However, with a sample of this size, the difference between the fitted skewness coefficient and the raw skewness is negligible. Yet, the advantage of fitting a Gamma distribution is that it allows a direct analogy with SPI computation: the mathematical relation between a fitted Gamma distribution and the variance of its normalized distribution is demonstrable. This is not the case when computing skewness without assuming a distributional model.

Comment
Section 2.2.2 – The threshold of 7 for skewness coefficients is chosen to define regions where the SPI is a useful metric. This seems slightly arbitrary – why not 3 or 4?

Reply: The goal here is to define a threshold such that the variance of SPI is approximately equal to 1. We considered this to be the case when the variance was not significantly different from 1, as shown in Fig. 1c and 1d. Areas where the variance differs significantly from 1 correspond to those where the skewness coefficient is greater than or equal to 7. We agree that the comparison is somewhat arbitrary, and following this comment we will state this more clearly in the paper. A more rigorous demonstration of the correspondence between both methods would be needed, which justifies the previous point. We added "Of course, this threshold remains arbitrary and would require more rigorous demonstration in order to determine the exact value to be applied." at the end of the paragraph "There is a significant similarity between the skewness coefficient and the variance. The skewness coefficient, which is calculated before the SPI is obtained, is therefore an acceptable proxy for the validity of the SPI. However, there is no quantifiable criterion for the value of this skewness coefficient. In the context of our study, a threshold of 7 for the skewness coefficient seems appropriate."

Comment
Appendix A: A robustness criterion, i.e. the robustness of the cluster centres to subsampling the data on which they are computed, is a reasonable metric to choose clusters, as implemented for example. The authors could consider investigating this metric.

Reply: We could also have tried a robustness criterion, which is an interesting criterion in clustering. However, we tested two methods that evaluate the ratio between intra-cluster variability and inter-cluster variability. The robustness criterion is another way of evaluating this. Since we ultimately decided to stick with the initial choice of seven clusters, we do not believe it is necessary to conduct this additional study. Furthermore, the other evaluation methods do not converge on a particular choice. Even if this new method converged on a choice of seven clusters (or another number), the evaluation of the fifth method would not be any more robust.

Comment
Section 4.1: This section could be improved by mentioning possible physical mechanisms that motivate the investigation of variations in weather regimes and drought frequencies. It should further at least be discussed that calculating a linear trend over the entire period cannot capture non-stationarity due to decadal variability. Furthermore, Non-stationarity or Stationarity would seem a more suitable section title compared to Representativity.

Reply: Thank you for your comment. We will expand this section by mentioning the physical mechanisms (see below) that may drive variations in regime frequency. We also modified the title of this section for "Non stationarity". We added the following paragraph at the end of the renamed "Non stationarity" discussion :

"On scales ranging from decades to multiple decades, low-frequency ocean variability becomes a key factor in modulating patterns. In particular, Atlantic Multidecadal Variability (AMV) influences the probability of certain regimes occurring, especially in spring and summer, through its sea surface temperature (SST) anomalies, which alter the position of the jet stream and large-scale circulation over Europe and the Mediterranean [Hertig, 2014]. Cassou et al. [2004] show that the spatial and temporal stationarity of winter regimes is sensitive to slow oceanic forcings. Thus, calculating a single linear trend over a long period does not account for this non-stationarity, as the relationship between the state of the ocean floor and atmospheric dynamics varies according to the phases of multi-decadal variability. Although the link between AMV and the frequency of regimes is supported by several studies, the causal attribution between ocean forcing and intrinsic atmospheric variability remains uncertain, and projections from climate models show a high degree of dispersion in the representation of these long-term variations "

Comment
Sentence starting with 'The Mediterranean area ...' (Lines 24-26). Given the rich literature on future projections of precipitation over the Mediterranean, there should be other citations here to complement

the IPCC.

Reply: Thank you for your comment. We have added recent references concerning Mediterranean climate projections: Tramblay et al. [2020], Essa et al. [2023], Vicente-Serrano et al. [2022], Spinoni et al. [2018]

Comment
The conclusions section would benefit from more detail, in particular the last paragraph on potential implications of this paper. In particular the wording 'Our findings could have substantial implications in the field of sub-seasonal to seasonal drought predictions [...] is a bit too generic

Reply: We will modify the conclusion in the next version of the article, in particular to take into account the new results presented here, which explain why the term in ΔC_k is predominant during reconstruction.

Regarding this somewhat generic sentence, we will modify the conclusion in the article. The general idea is that seasonal precipitation predictability is very low in Europe. If weather patterns are somewhat predictable, we can hope to anticipate the risk of drought.

Comment
Abstract (Line 2), please specify the following part of the sentence: "due to the complex interplay between regional climate variability and large-scale atmospheric circulation." This distinction is ambiguous, especially made without prior context in the first sentence of the abstract. Large-scale atmospheric circulation is mostly understood to be a part of regional climate variability. Do you mean the interplay of atmospheric, ocean and land drivers? Or the interplay of thermodynamic and dynamic processes?

Reply: We agree with this remark and will specify in the abstract that we mean "due to the complex interplay between large-scale dynamic circulation patterns and local thermodynamic processes"

Comment
Line 34 and following: 'In this region, seasonal and subseasonal climatic variations are less dependent on tropical teleconnections than in other regions of the world (). Surface climate in this region is driven by more regional phenomena particularly the North Atlantic atmospheric circulation'. These statements are quite ambiguous and not well evidenced by the cited literature. There is a wide range of papers evidencing the influence of tropical teleconnections in Europe - including the Shaman and Tziperman 2011 paper cited, the North Atlantic circulation itself is influenced by tropical teleconnections, as investigated for example in Cassou 2008 which is cited a few sentences later. The statement 'compared to other regions of the world' is ambiguous

Reply: Thank you for your comment. We did not mean to suggest that climatic variations are not influenced at all by tropical teleconnections. We wanted to point out that they are less influenced than other regions of the world. The papers cited assess the strength of tropical teleconnections, showing mainly that despite their existence, they remain weak: Van Oldenborgh et al. [2000] shows a maximum correlation of 0.3 over a narrow band of Europe (corresponding to our CEur and WEur regions) between DJF NINO3 and MAM precipitation. Similarly, Cassou [2008] shows that there is an interaction between the MJO and conventional weather patterns. However, this only concerns winter weather patterns, with an interaction involving only the first two modes (i.e. NAO+ and NAO-). Finally, Fraedrich [1994] attempts to summarise the knowledge on the link between ENSO and the European climate. He concludes that it is difficult to prove the existence of such a link, that a possible link exists when certain conditions are met, and that the degree of confidence associated with such a statement is fairly low.

However, we are rephrasing our statement to make it clearer that we are only asserting the weakness of such a link, not its non-existence. We replaced "In this region, seasonal and sub-seasonal climatic variations are less dependent on tropical connections than in other regions of the world" by "Although teleconnections between this region and the tropics exist on seasonal and sub-seasonal scales, the signal of their influence remains weak and variable".

Comment
Line 160: Why does the result that the variance of the SPI3 averaged over each region is not different from 1 at the 95% confidence level confirm the suitability of the regionalization for the purpose of this study?

Reply: The SPI in grid points is valid for drought analysis if its variance is 1. The disadvantage of creating regions is that the spatial average of the SPI can create a flat average SPI (and therefore reduced variance) if the droughts within that region are not synchronous. We regionalise to create homogeneous regions by clustering neighbouring grid points whose droughts are as concurrent as possible. To study droughts in these regions, we then calculate a 'regional' SPI equal to the spatial average of the SPI of the grid points. If the variance of the average SPI within these regions is still 1, then our division is well suited for drought analysis at the scale of the regions obtained.

Comment
In a number of locations, for example Line 241, the authors write that the average frequency of each weather pattern during the 91 days preceding the droughts is calculated. This wording could be clearer, as the weather regime frequency is calculated precisely during those days over which the SPI is calculated. A clearer way of phrasing this would be 'during the 91 days over which the SPI3 is calculated.

Reply: Thank you for your comment. We added in the methodology the precision that the specific period of 91 days is chosen to correspond with the 3 months over which the SPI is calculated.

Comment
Line 436: References should be added to this statement.
Figure G1: Add axis labels.
Page 33: More details on Author Contributions

Reply: Thank you for your comment, we added references, axis labels and we specified the author contributions.

References

Hannah C. Bloomfield, David J. Brayshaw, and Andrew J. Charlton-Perez. Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types. *Meteorological Applications*, 27(1):e1858, January 2020. ISSN 1350-4827, 1469-8080. doi: 10.1002/met.1858. URL https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.1858.

Christophe Cassou. Intraseasonal interaction between the madden–julian oscillation and the north atlantic oscillation. 455(7212):523–527, 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature07286. URL https://www.nature.com/articles/nature07286.

Christophe Cassou, Laurent Terray, James W. Hurrell, and Clara Deser. North Atlantic Winter Climate Regimes: Spatial Asymmetry, Stationarity with Time, and Oceanic Forcing. *Journal of Climate*, 17(5):1055–1068, March 2004. ISSN 0894-8755, 1520-0442. doi: 10.1175/1520-0442(2004)017⟨1055:NAWCRS⟩2.0.CO;2. URL https://journals.ametsoc.org/view/journals/clim/17/5/1520-0442_2004_017_1055_nawcrs_2.0.co_2.xml. Publisher: American Meteorological Society Section: Journal of Climate.

Yassmin H. Essa, Martin Hirschi, Wim Thiery, Ahmed M. El-Kenawy, and Chunxue Yang. Drought characteristics in Mediterranean under future climate change. *npj Climate and Atmospheric Science*, 6(1):133, September 2023. ISSN 2397-3722. doi: 10.1038/s41612-023-00458-4. URL https://www.nature.com/articles/s41612-023-00458-4.

Klaus Fraedrich. An ENSO impact on europe? 46(4):541–552, 1994. ISSN 0280-6495, 1600-0870. doi: 10.1034/j.1600-0870.1994.00015.x. URL http://tellusa.net/index.php/tellusa/article/view/15643.

Christian M. Grams, Remo Beerli, Stefan Pfenninger, Iain Staffell, and Heini Wernli. Balancing europe's wind-power output through spatial deployment informed by weather regimes. 7(8):557–562, 2017. ISSN 1758-678X, 1758-6798. doi: 10.1038/nclimate3338. URL https://www.nature.com/articles/nclimate3338.

Jacobeit J. Hertig, E. Variability of weather regimes in the North Atlantic-European area: past and future. 2014. doi: 10.1002/asl2.505. URL https://rmets.onlinelibrary.wiley.com/doi/10.1002/asl2.505.

Duncan Pappert, Alexandre Tuel, Dim Coumou, Mathieu Vrac, and Olivia Martius. Long vs. Short: Understanding the dynamics of persistent summer hot spells in Europe, October 2024. URL https://egusphere.copernicus.org/preprints/2024/egusphere-2024-2980/.

Maria Pyrina and Daniela I. V. Domeisen. Subseasonal predictability of onset, duration, and intensity of European heat extremes. *Quarterly Journal of the Royal Meteorological Society*, 149(750):84–101, 2023. ISSN 1477-870X. doi: 10.1002/qj.4394. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.4394. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4394.

Jonathan Spinoni, Gustavo Naumann, Hugo Carrao, Paulo Barbosa, and Jürgen Vogt. World drought frequency, duration, and severity for 1951–2010. *International Journal of Climatology*, 34(8):2792–2804, 2014. ISSN 1097-0088. doi: 10.1002/joc.3875. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.3875. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.3875.

Jonathan Spinoni, Jürgen V. Vogt, Gustavo Naumann, Paulo Barbosa, and Alessandro Dosio. Will drought events become more frequent and severe in europe? 38(4):1718–1736, 2018. ISSN 0899-8418, 1097-0088. doi: 10.1002/joc.5291. URL https://rmets.onlinelibrary.wiley.com/doi/10.1002/joc.5291.

Jonathan Spinoni, Paulo Barbosa, Alfred De Jager, Niall McCormick, Gustavo Naumann, Jürgen V. Vogt, Diego Magni, Dario Masante, and Marco Mazzeschi. A new global database of meteorological drought events from 1951 to 2016. *Journal of Hydrology: Regional Studies*, 22:100593, April 2019. ISSN 2214-5818. doi: 10.1016/j.ejrh.2019.100593. URL https://www.sciencedirect.com/science/article/pii/S2214581818303136.

Marc Stefanon, Fabio D'Andrea, and Philippe Drobinski. Heatwave classification over Europe and the Mediterranean region. *Environmental Research Letters*, 7(1):014023, February 2012. ISSN 1748-9326. doi: 10.1088/1748-9326/7/1/014023. URL https://dx.doi.org/10.1088/1748-9326/7/1/014023. Publisher: IOP Publishing.

Yves Tramblay, Aristeidis Koutroulis, Luis Samaniego, Sergio M. Vicente-Serrano, Florence Volaire, Aaron Boone, Michel Le Page, Maria Carmen Llasat, Clement Albergel, Selmin Burak, Maxime Cailleret, Ksenija Cindrić Kalin, Hendrik Davi, Jean-Luc Dupuy, Peter Greve, Manolis Grillakis, Lahoucine Hanich, Lionel Jarlan, Nicolas Martin-StPaul, Jordi Martínez-Vilalta, Florent Mouillot, David Pulido-Velazquez, Pere Quintana-Seguí, Delphine Renard, Marco Turco, Murat Türkeş, Ricardo Trigo, Jean-Philippe Vidal, Alberto Vilagrosa, Mehrez Zribi, and Jan Polcher. Challenges for drought assessment in the Mediterranean region under future climate scenarios. *Earth-Science Reviews*, 210:103348, November 2020. ISSN 0012-8252. doi: 10.1016/j.earscirev.2020.103348. URL https://www.sciencedirect.com/science/article/pii/S0012825220303949.

Geert Jan Van Oldenborgh, Gerrit Burgers, and Albert Klein Tank. On the el nino teleconnection to spring precipitation in europe. 20(5):565–574, 2000. ISSN 0899-8418, 1097-0088. doi: 10.1002/(SICI)1097-0088(200004)20:5⟨565::AID-JOC488⟩3.0.CO;2-5. URL https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0088(200004)20:5<565::AID-JOC488>3.0.CO;2-5.

Sergio M. Vicente-Serrano, Dhais Peña-Angulo, Santiago Beguería, Fernando Domínguez-Castro, Miquel Tomás-Burguera, Iván Noguera, Luis Gimeno-Sotelo, and Ahmed El Kenawy. Global drought trends and future projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380(2238):20210285, December 2022. ISSN 1364-503X, 1471-2962. doi: 10.1098/rsta.2021.0285. URL https://royalsocietypublishing.org/doi/10.1098/rsta.2021.0285.