Response to Referee Comment 1 *on* **"Seasonal weather regimes in the North Atlantic region: towards new seasonality?"** *by* **Florentin Breton et al.**

Comment

This study aims to investigate changes in the weather seasonality in the North Atlantic. Seasonal weather regimes (SWRs) are defined using cluster analysis based on the first principal component on raw Z500 data. Results for four and seven regimes are presented for ERA-Interim and 12 CMIP5 models. Models results are first compared to ERA-Interim for the period 1979-2017. The authors investigate changes in the patterns and frequency of the regimes by comparing CMIP5 simulations for present (1980-2008) and future (2071-2100) climate. The paper includes a lot of Figures (10 in the main document and 22 in the supplement), with little explanation and description. The authors compare SWRs to the classical weather regimes, but no apparent link exists between the two concepts. The seasonality analysis is based on daily data but it is not related to weather phenomena (i.e. during winter only 1 regime is found).

Response:

We thank the referee for the comments. However, some of these comments reflect misunderstandings of the paper which we will address in our response. We will update the manuscript with the suggestions and clarifications in relation to the comments. Our response to each comment can be found below.

Comment

1. Definition of seasonal weather regimes (SWRs) and comparison to classical "weather regime". I have trouble to see the link between the author's definition of SWRs and the classical weather regimes. The authors compute the first EOF of raw Z500,which should represent the seasonal cycle (no Figure of EOF1 is provided). SWRs are then defined by clustering PC1 (a single 1D variable). So these clusters should represent the strength of the seasonal cycle, which can be seen for example in Figure 2 by the Z500 contours (stronger Z500 gradient in winter R1, weaker in summer R4). It is also clear from Figure 1 that during winter only R1 occurs. Does it mean that there is no weather in winter and all days look like R1 (cf. for example with Figure 2 in Michel and Riviere (2011)? I find therefore misleading to match the patterns in Figure 2 to the classical weather regimes (see also comment on Figure 2), as those methods are considering two different things. To calculate classical weather regimes a cluster analysis is applied to more than 10 PCs (that explain at least 80% of the variance) and the seasonal cycle is removed from the raw data (e.g. Cassou 2008, Vrac and You 2010).Thus, the mix of different concepts makes it hard

to understand what is the main goal of the paper and what the authors attempt with their analysis.
a) If the goal is to explain changes in the seasonality by analyzing PC1 of raw Z500 (what it seems so far),I would recommend comparing EOF1 patterns and the distribution of PC1. Would this be a different way of investigating seasonality as compared to previous studies (stated in the Intro, l. 25, without many details)?
b) If the goal is to understand future changes in seasonality by looking at the changes in the frequency of weather regimes, those regimes should be defined accordingly (see for example Cattiaux et al. (2013) or Grams et al. (2017)). However, this task might be challenging, as most climate models have a large bias in weather phenomena (such as blocking, e.g. Masato et al 2013).

Response:
The link between our SWRs and the classical weather regimes is the analogy in the general way they are defined (i.e. clustering approach by classifying similar atmospheric situations). However, our preprocessing of the data and their constraints to define the regimes are different, and thus their temporal properties are also different, compared to classical weather regimes. Indeed, our goal is not to study the intra-seasonal variability with classical circulation regimes but rather to study the variability of the seasonal cycle via clustering-based regimes of circulations. Some of the seasonal structures that we find happen to be similar to some of the intra-seasonal structures. A detailed explanation is provided below.

Our SWRs are defined in the same way as in the original paper (Vrac et al. 2014) that investigated past changes in the seasonality of daily atmospheric circulation (here we compare historical simulations to a reanalysis, before investigating future simulations). So, they are defined similarly to “classical” weather regimes except that we take full-year data (instead of summer or winter only; lines 76-77) and raw values (instead of seasonal anomalies; lines 76-77). Indeed, we are studying how the seasonal structures might migrate in the calendar year (e.g. from summer towards winter or vice-versa). Our regimes are “seasonal” because they are defined (and studied) based on their seasonality (weather patterns and annual cycle). We compare them to classical weather regimes to highlight weather patterns that appear similar (lines 129-137), while remaining cautious about the differences in their definition and properties (lines 137-138). It is important to keep in mind that the SWRs are defined based on the daily PC1 values of Z500 and that the figures 1 and 2 represent averages of the atmospheric conditions (respectively the frequency of occurrence and the spatial patterns) for the days belonging to the SWRs over the period. Indeed, the days are attributed to the regimes based on their similarity in terms of atmospheric situations, and there are also many days of R2 in winter (not only R1), or R3 in summer (not only R4). Since our main goal is to study the response of the atmospheric circulation seasonality to climate change, this attribution is based rather on the long-term variability of seasonality rather than intra-seasonal variability.

Regarding PC1, it contains not only most of the seasonal cycle, but also a large part of the long-term variability, as shown on Figure A below.

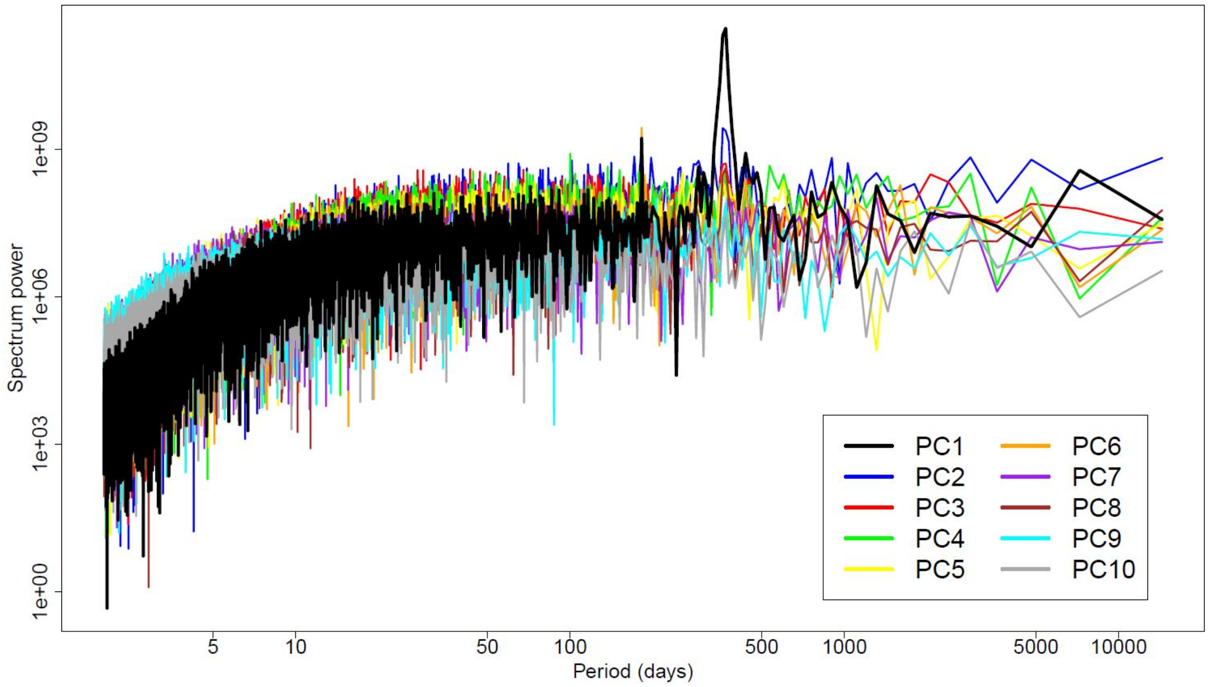

Figure A. Spectral power captured by each PC in function of the period. Dataset: Z500 from ERAI over 1979-2017.

Furthermore, the pattern given by the first EOF and the distribution of PC1 for ERAI and climate models appear generally similar, as shown respectively on Figures B and C below. The pdf of PC1 approximately corresponds to a bimodal Gaussian-like distribution (Fig. C).

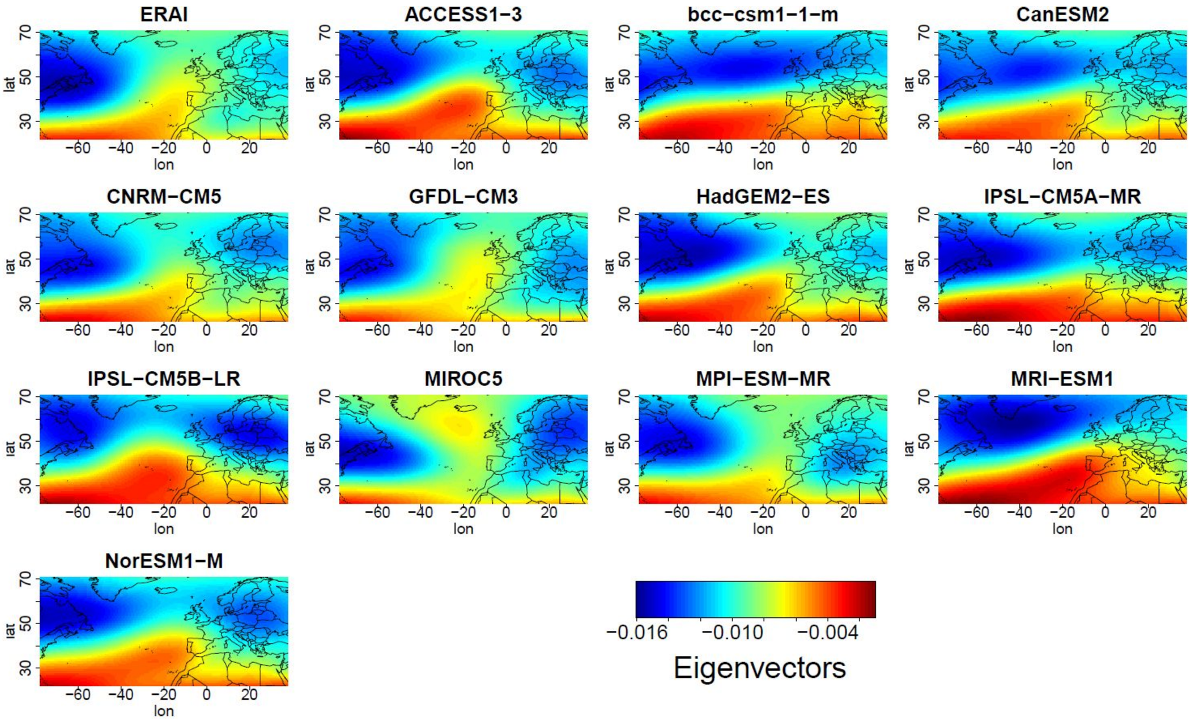

Figure B. Maps of the first EOF for ERAI and each climate model over 1979-2017 (exception of HadGEM2-ES: 1981-2017).

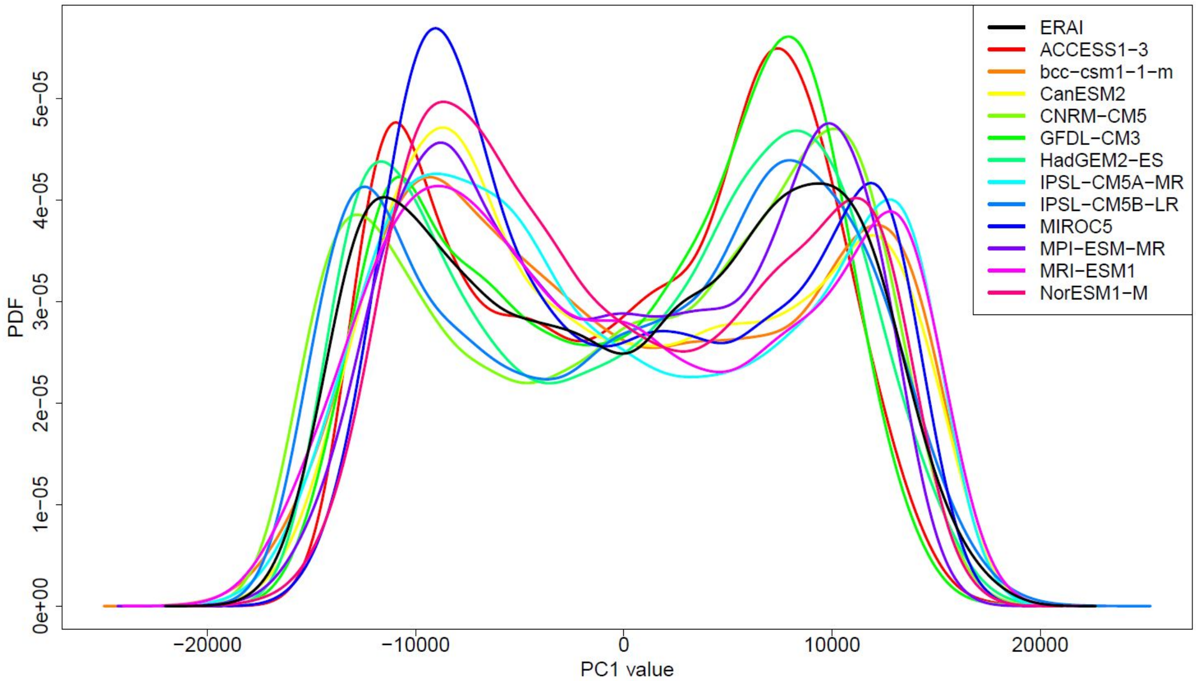

Figure C. Probability distribution function of PC1 for ERAI and each climate model over 1979-2017 (exception of HadGEM2-ES: 1981-2017).

Regarding the number of PCs to include in the analysis, adding more PCs increases the total variability but adds little seasonality (lines 84-87). The amount of variability and seasonality captured by the PCs are shown on Figure D below.

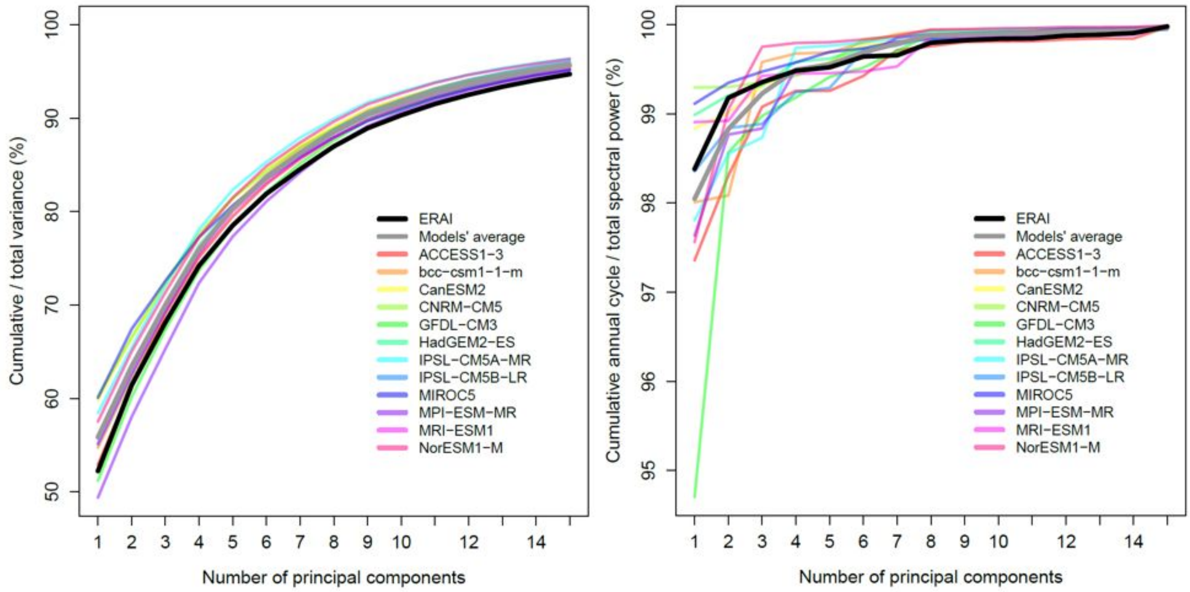

Figure D. Cumulative variance (left) and cumulative annual cycle (right) captured in function of the number of PCs for ERAI and each climate model over 1979-2017 (exception of HadGEM2-ES: 1981-2017).

Furthermore, we tested the sensitivity of the SWR results (lines 87-88) to using additional PCs in the clustering but in most cases including a few more PCs (e.g. PC2 to PC5) brings similar results (weather patterns, seasonal cycle), although the results become more dissimilar when a higher proportion of PCs is included. Using PC2 to PC5 has a small influence on the results of the clustering over 1979-2017 and this influence becomes very small over 1979-2100.

The comment from the reviewer states that "the mix of different concepts makes it hard to understand what is the main goal of the paper and what the authors attempt with their analysis" and suggests two alternatives (a and b) based on the desired goal. We argue that the different concepts are not mixed but are each introduced and explained, as they are the main goals of the paper (lines 48-56 and 58-65).

a) The reviewer's suggestion represents the first steps of our analysis, which we complemented with the weather regime approach to investigate more in depth the nature of the changes in terms of patterns, timing and drivers.
b) If we separate the data between seasons to define the weather regimes, then it defeats the purpose of studying if specific regimes are migrating within the year (e.g. from summer towards winter or vice-versa) which is one of the main goals of our paper. Moreover, we address the bias in climate models in the first part of the analysis by comparing the weather regimes from the simulations to those from the reanalyses, before looking at the future.

To clarify the manuscript, we propose the following revisions (in blue; lines 84-88 and 107-108):

"Only the first principal component (PC1) is kept and used for clustering because it captures between about 49% and 60% of the variance and between about 95% and 99% of the seasonal cycle (spectral power at 1/365 of frequency;1/360 for Hadley Center) over 1979-2017 for ERAI (similar to Vrac et al. (2014) on another reanalysis) and all climate models (not shown). A large part of the long-term variability is also contained in PC1, while the spatial pattern (eigenvectors) and statistical distribution (pdf) of PC1 are generally similar between ERAI and models over 1979-2017 (not shown). Including more PCs in the analysis provided similar results (not shown) but brought more noise (more variance but only little more seasonality)."

"*Different clustering methods can lead to different results (e.g., Philipp et al. (2010)) so we tested the sensitivity of the SWR results to using k-means instead of EM clustering, which brought very similar results (not shown). We also tested the sensitivity of the clustering results (weather patterns, annual cycle) to the number of PCs included (from 1 to 5). There was a small influence of additional PCs on the results (reanalyses, models) over 1979-2017 and very small influence over 1979-2100 (increasing with the number of PCs; not shown). This reinforced our choice of using only PC1, considering that additional PCs represent little additional seasonality and difference in the long-term response of atmospheric circulation to climate change.*"

Comment:

2. I miss several explanations in the data and methodology. For example, figure S5 shows seasonal anomalies of TAS, but I could not find how these are defined (neither in the methods, l. 75 nor in the text l. 153). Since one of the main points of the paper should be about changes in the season, the definition of seasonal anomalies needs to be clarified. The same is true for Z500 anomalies.

Response:

The reviewer is referring to Figure 5 in the supplement, which has a caption describing how the calculation is done (see text below). The same is true for Z500 anomalies (e.g. Figure 2).

"Composite TAS maps conditional to the four regimes (one per row) for ERAI (first column) and climate models (second column;each map shows the average of 12 composite maps; third column shows standard deviation of TAS between the 12 composites). The maps are calculated by averaging the seasonal anomalies (shading) and raw values (contour lines) over the days belonging to the regime. Seasonal anomalies correspond to the raw values minus the average seasonal cycle. The number of days per regime is shown above each map (average of 12 values for the climate models)."

To complement the caption and clarify in the main text how the composite maps of the regimes are calculated, we propose the following changes (highlighted in blue) in the manuscript (lines 127-129):

“The composite maps associated with each regime are shown in Figure 2. Each composite map is calculated by averaging the values of the Z500 fields corresponding to the days that belong to the regime, with color shading representing the seasonal anomalies and contour lines representing the raw values. Seasonal anomalies correspond to the raw values minus the average seasonal cycle over 1979-2017. For climate models, each regime composite map is determined individually (i.e. average map) and the multimodel composite is calculated as the mean of the distribution of the twelve composites.”

Comment:

3. The authors argue for 4 and 7 regimes based on the BIC (without showing it). However, it is not clear what is the main advantage of using 7 regimes. With 4 regimes,R1-4 patterns for 1979-2008 and 2071-2010 are still similar (Fig. S6), while this is not the case with 7. For example, R7 in 1979-2008 (Fig. 3) represents 54 days (0%) and it is argued that this regime becomes more frequent in the future (24%), but R7 for 2071-2010 is very different from R7 in 1979-2008. My recommendation is to present a complete analysis for either 4 or 7 regimes, but not jumping back and forth (i.e. some Figures for 4 are in the supplement, some in the main text).

Response:

The values of the BIC for the clustering over 1979-2017 are shown on Figure E below. The BIC values were normalized between 0 (best statistical model) and 1 (worst model) in order to show the results for all the datasets on the same figure. The best statistical model of EM is chosen between E (equal variance) and V (variable variance) for the univariate (i.e., one-dimensional) mixture. This figure illustrates that a plateau of the BIC is reached at 4 clusters with the V model of EM (lines 105-106). One later comment from the reviewer suggests to include the figure of the BIC in the paper, we agree and the details can be found in our response to that comment.

The second advantage of using 4 regimes (in addition to the plateau of BIC) is that it corresponds to the traditional (astronomical) number of seasons (lines 105-106), which makes sense (as a preliminary analysis) for investigating seasonality. However, using 4 clusters is too restrictive to properly allow the emergence of new structures in the clustering over the future (179-182). We started by using 4 regimes in the future to compare the results with the past in terms of changes, before using 7 regimes to investigate the emergence of new structures (we also tried 7 regimes in the past but no new structure emerged; lines 208-210).

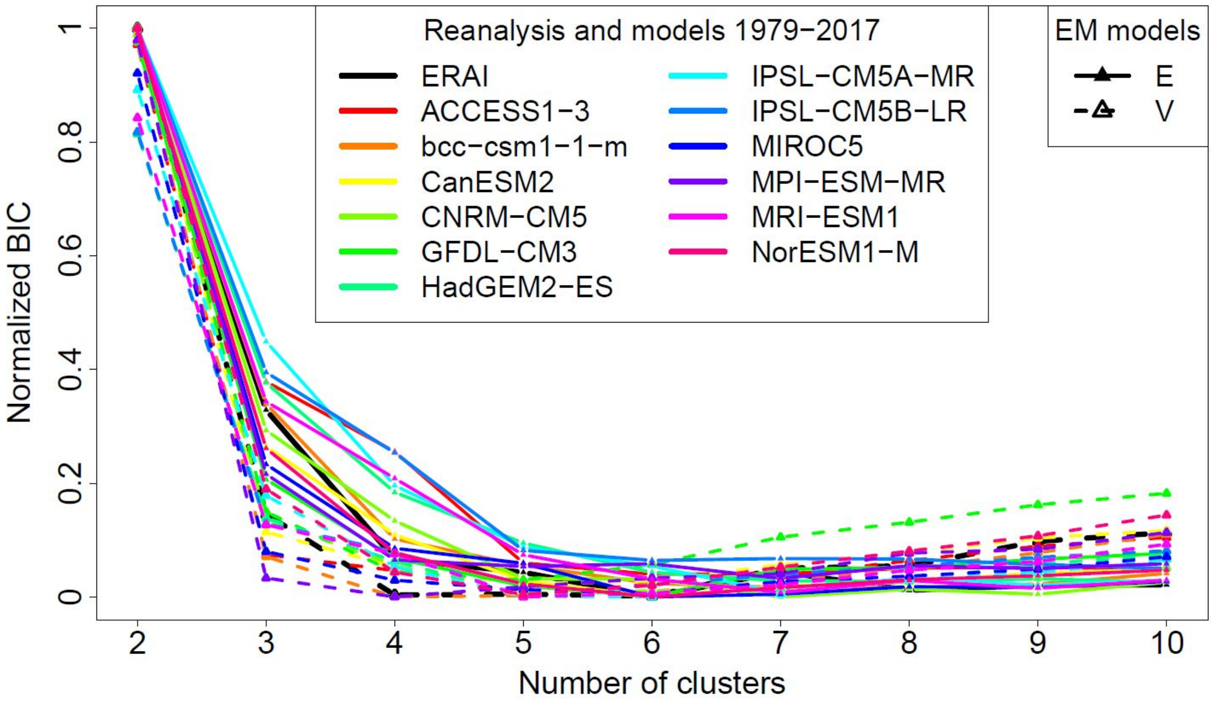

Figure E. Bayesian information criterion (BIC) of the clustering models in function of the number of clusters and model type for ERAI and each climate model over 1979-2017 (exception of HadGEM2-ES: 1981-2017). The BIC values are normalized between 0 (best) and 1 (worst model). E: equal variance, V: variable variance.

The regime spatial patterns appear different with 7 regimes in 2071-2100 compared to 1979-2008 because these patterns are calculated as anomalies relative to the average seasonal cycle of the considered period but this cycle changes a lot over 1979-2100 (see lines 161-162 and our response to another comment below). The regimes are based on clustering which puts similar patterns (of raw Z500 fields transformed into PC1 values) together by definition. Therefore, if future R7 doesn't appear similar to past R7 in terms of the pattern of Z500 **seasonal anomalies**, it is similar by definition (clustering consistency) in terms of the pattern of **raw** Z500 (or PC1).

The second advantage of using 7 regimes, in addition to the larger freedom in the clustering-based definition of regimes allowing for the emergence of new structures, is that it clearly illustrates the transition between the "past" (R6) and the "new" (R7) summer regimes. We heuristically chose 7 regimes after trying from 5 to 10 regimes. As an example, the results for 6 and 8 regimes are respectively shown on Fig. F-I below (weather patterns, annual cycle). They highlight seasonal changes similar to those identified from 7 regimes.

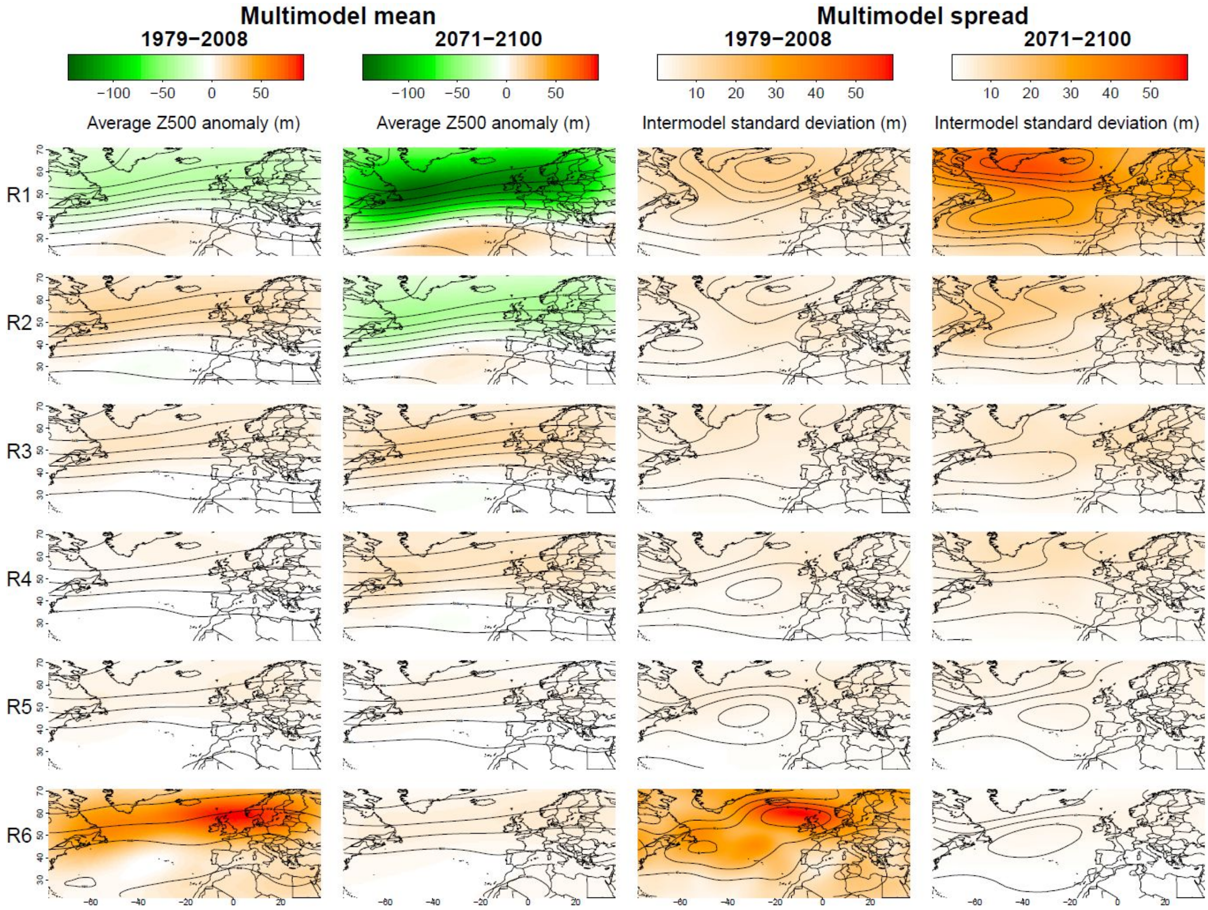

Figure F. Weather patterns for 6 regimes.

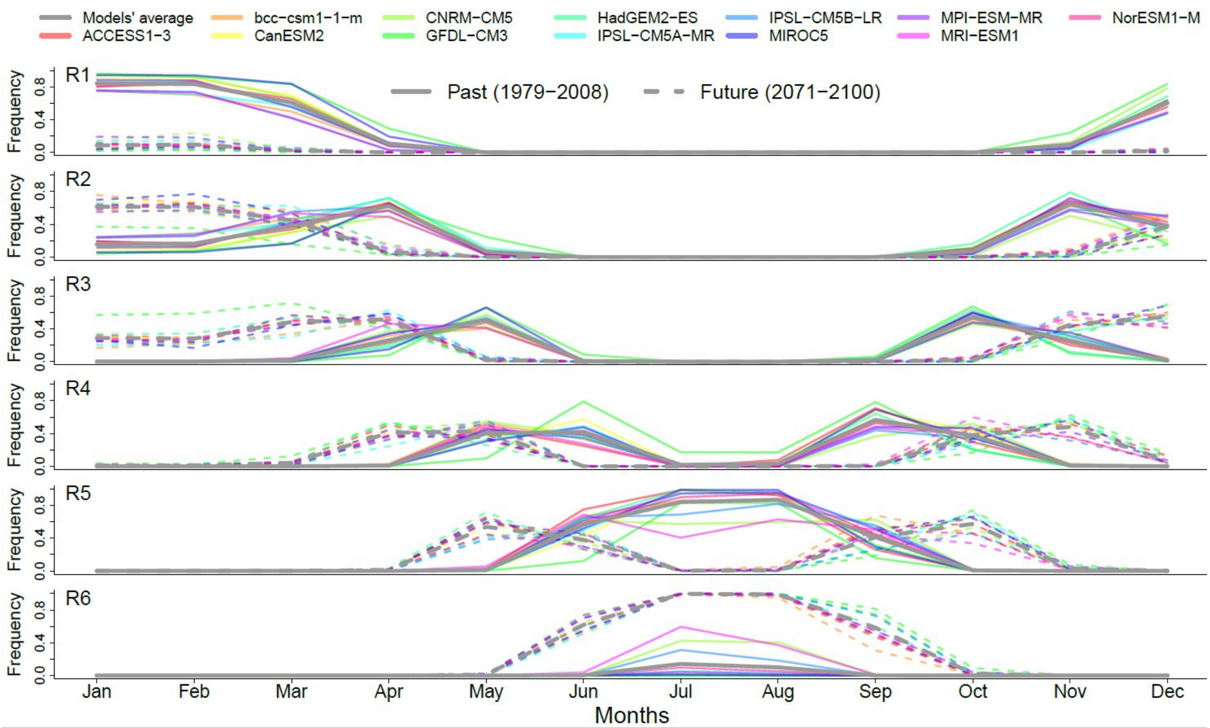

Figure G. Annual cycle for 6 regimes.

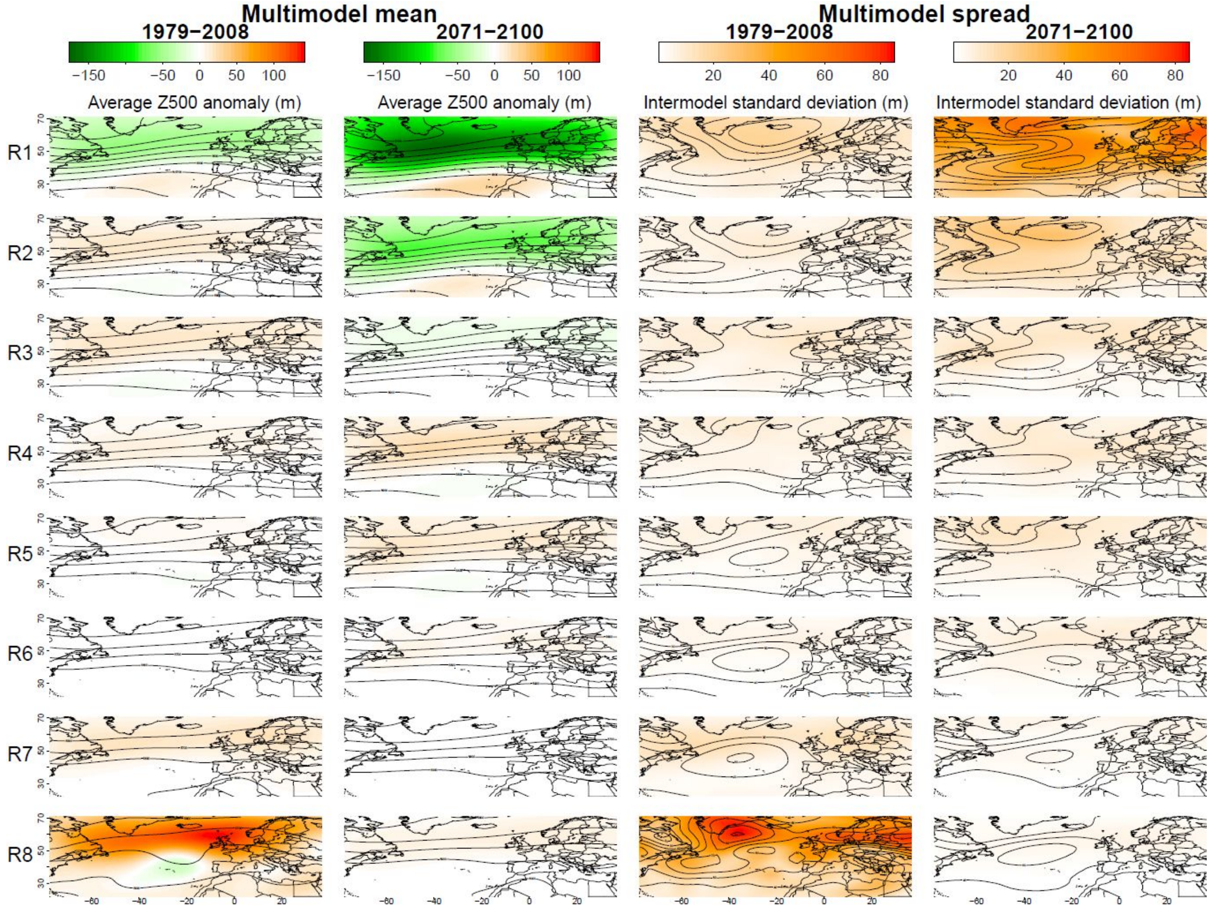

Figure H. Weather patterns for 8 regimes.

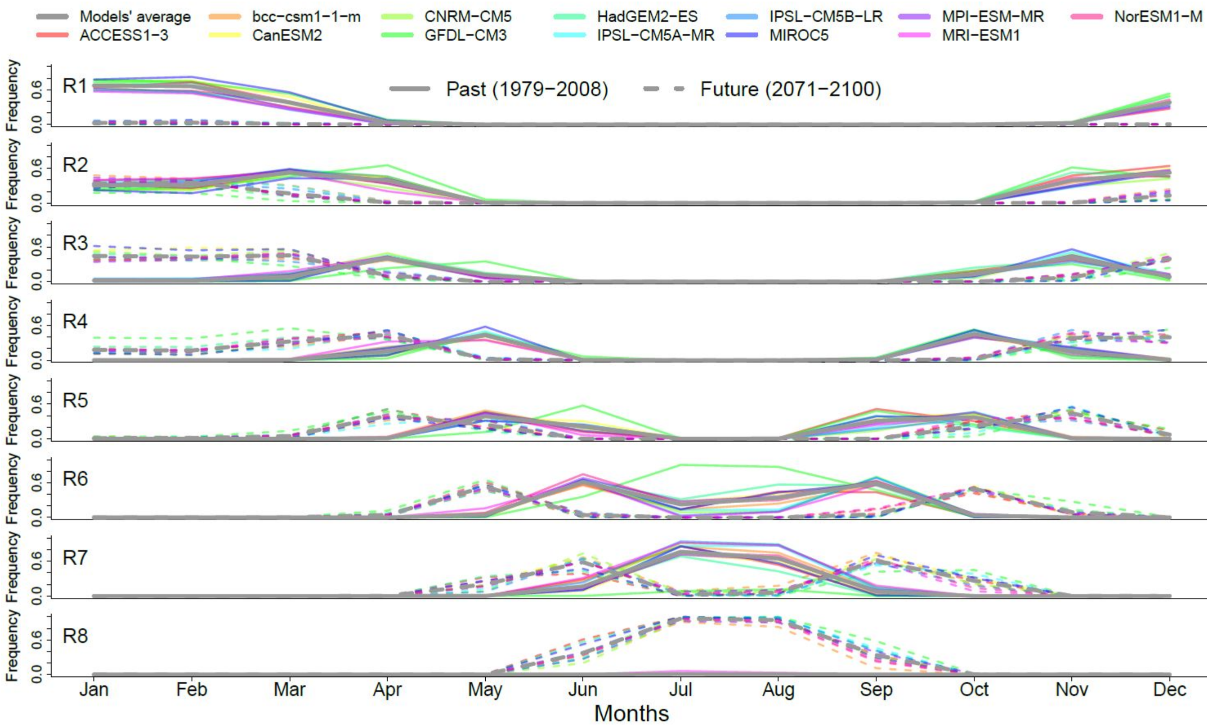

Figure I. Annual cycle for 8 regimes.

For the reasons stated above, we decided not to present a complete analysis with either only 4 or only 7 regimes, but rather to include both while explaining why (as detailed in the manuscript).

Comment:

Title: I find the word "weather" in the title non-appropriate and confusing. It is strange to call "weather" something that persists for one winter/season. Circulation regimes or seasonal regimes (without weather) might be a better option.

Response:

The term "seasonal weather regimes" corresponds to using the "weather regime" approach to study seasonality, as in Vrac et al. (2014). In our study as for classical weather regimes, the weather is represented by the daily atmospheric situations (Cassou (2008); Michel & Riviere (2011)). If our seasonal weather regimes do sometimes persist for weeks, this is not unusual from a synoptic perspective. Indeed, some atmospheric situations can persist for weeks to month based on the analysis of the jet-stream and weather regimes (e.g. Barnes & Hartmann (2010); Franzke et al. (2011); Hannachi et al. (2012); Woolings et al. (2010)).

Comment:

l. 32 "North Atlantic atmospheric patterns are the results of two systems operating at different scales". Please rephrase this paragraph, adding the relevant Literature. A few critical points: low-frequency structures are not "systems". What is the "atmospheric dynamic variability"?

l . 34 anticyclones are also important.

l. 30-40: There is missing relevant literature (for example, there is a bulk of literature on the eddy-driven jet more relevant than Cassou et al 2008, l. 35). See also some suggestions that could be useful at the end of the document.

Response:

The "atmospheric dynamic variability" represents the variability of atmospheric dynamics (i.e. circulation). We agree with the suggestions made by the reviewer and propose the following revised version (with changes in blue) of the manuscript (lines 32-35):

"*North Atlantic atmospheric patterns are the results of physical phenomena operating at different scales: "low-frequency" quasi-static structures such as the Icelandic Low and the Azores High (*Angell & Korshover (1974), Hurrell & Deser (2010), Marshall et al. (2001), Wang (2002)), and *"high-frequency" eddies or propagating synoptic waves such as cyclones and anticyclones (*Barnston & Livezey (1987), Price & Magaard (1986)) *associated to the eddy-driven jet stream (Blackmon et al. (1984), Franzke et al. (2011), Woolings et al. (2010), Woolings & Blackburn (2012)).*"

Comment:
l. 42 How can climate dynamics have a strong seasonal feature? The climate in the extratropics has a strong seasonality. Should atmospheric blocking (time scales of 1-2weeks) be an example of climate dynamics?
l. 49-51 what is the link between the representation of the atmospheric circulation and the seasonality in the model? I agree that it is important that both are correctly represented, but I can not see the link between the two concepts yet. Are the authors suggesting for example, that if the jet stream moves too slowly polewards through the season, this will have an impact on seasonality?

Response:
Examples of climate dynamics with a strong seasonal feature are the ITCZ (Hu & Liu (2007), Schneider et al. (2014), Xian & Miller (2008)) and the jet-stream (Iqbal et al. (2018); Kuang et al. (2014); Woolings et al. (2010)), which both typically migrate northward in boreal summer and southward in boreal winter. Atmospheric blocking is an example of peculiar climate dynamics in the form of quasi-stationary anticyclones (Sillmann & Croco-Maspoli (2009)) that can facilitate weather extremes (e.g. Sillmann et al. (2011); Schaller et al. (2018)).

The link between the representation of the atmospheric circulation and seasonality in the model corresponds to the effects of the seasons on the behaviour of the atmospheric circulation in terms of preferential flow regimes (e.g. jet-stream), and weather extremes (Cassou et al. (2005), Iqbal et al. (2018), Scaife et al. (2008), Wallace et al. (1993), Woolings et al. (2010)). To clarify, we propose the following revisions (in blue) in the manuscript (line 42).

"Extratropical climate variability is largely seasonally dependent (Wallace et al. (1993)), and both climate dynamics (e.g. Iqbal et al. (2018); Woolings et al. (2010)) and weather extremes (e.g. Cassou et al. (2005); Scaife et al. (2008)) have strong seasonal features."

Comment:
l. 83-84 I expect the first EOF of raw data to be the seasonal cycle. Is that correct?

Response:
Yes but not only, see our detailed response to the first comment.

Comment:
l. 84 What is the main advantage of a GMM if only a single PC is used for clustering? How different is this method from dividing PC1 into quantiles? Is the PDF of PC1 non-gaussian?

Response:

The pdf of PC1 corresponds to a bimodal Gaussian-like distribution (see our detailed response to the first comment). The first advantage of the GMM is that it is more flexible than k-means (Estivill-Castro and Yang (2000), Han et al. (2011), Lior and Maimon (2005); see also our response to a similar comment from RC2). Its second advantage is being better than k-means at accounting for a mixture of several PCs in the clustering. Indeed, we initially tested the sensitivity of the results to the number of PCs used (from 1 to 5), which brought similar results to using only PC1 (lines 84-88 in the paper, and more details in our response to the first comment). Since we finally only use PC1, the initialization of EM divides the distribution of PC1 into quantiles. The clusters corresponding to these quantiles are then optimized through the steps E and M towards more Gaussian-like distributions, with the possibility of overlap between the distributions of the clusters in the attribution of PC1 values. This is an advantage of the GMM (even with univariate data) by allowing a probabilistic clustering, attributing for each PC1 value a probabilistic of belonging to each cluster, unlike k-means which is binary (exclusive clusters without overlap), ultimately giving a better model approximation of the distribution of PC1 (in terms of maximum likelihood estimation of the statistical parameters). The possibility to use the BIC to estimate the optimal number of clusters is another advantage of the GMM. To clarify the manuscript, we propose the following revisions (changes in blues; lines 107-108):

"*Different clustering methods can lead to different results (e.g., Philipp et al. (2010)) so we tested the sensitivity of the SWR results to using the k-means clustering algorithm (more popular but less flexible; Estivill-Castro and Yang (2000), Han et al. (2011), Lior and Maimon (2005)) instead of EM, which brought very similar results (not shown). EM can be seen as a generalization of k-means with less constraint on the shape of clusters and better ability to account for structures of arbitrary shape (Han et al. (2011), Lior and Maimon (2005).*"

Comment:

l.104: Instead of only adding the formula in the appendix, it would be very useful to have a figure showing the BIC for the different k in the appendix/supplement

Response:

We agree with the comment from the reviewer and propose to add Figure E (shown below) in the appendix, as well as making the following revisions (changes shown in blue; lines 105-106 and 461-463):

"Four SWRs (hereafter SWR4) correspond both to a plateau of BIC (Figure B1 in Appendix) and to the traditional (astronomical) number of seasons."

“An additional constraint on the definition of clusters is on the covariance matrix. Our GMM is univariate (since we only use PC1) so the variance can be equal (E model) or different (V model) between clusters (i.e. constraint on volume but not on shape or orientation of clusters.”

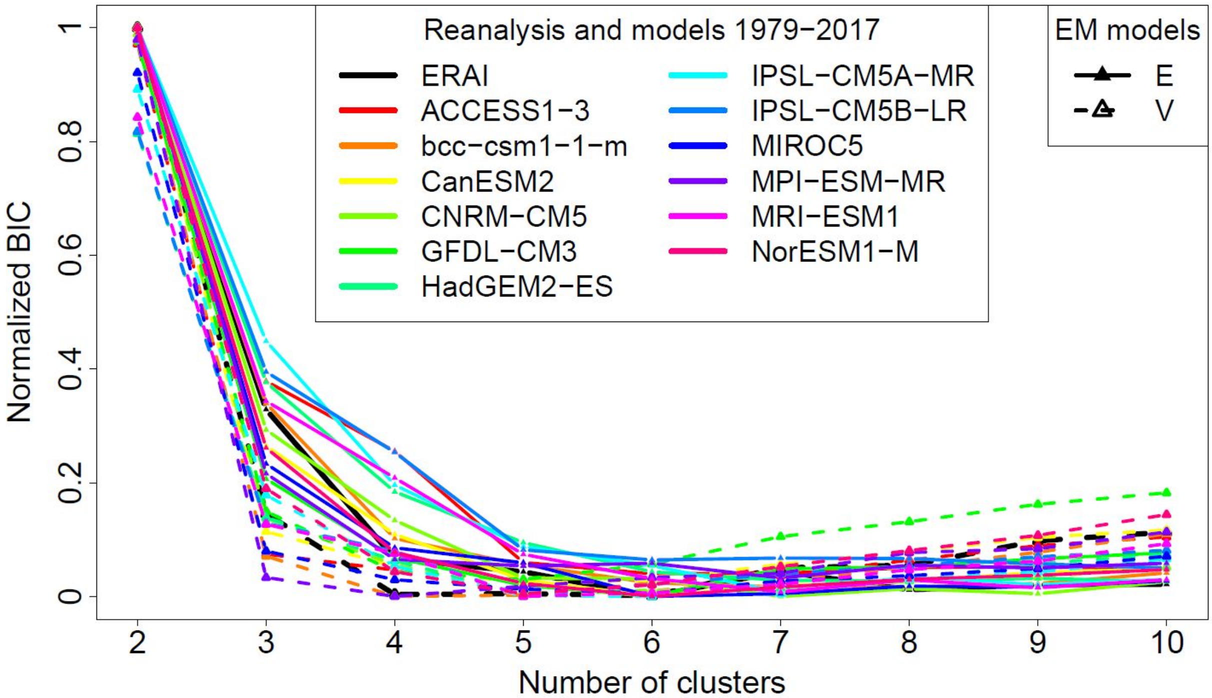

Figure E. Bayesian information criterion (BIC) of the clustering models in function of the number of clusters and model type for ERAI and each climate model over 1979-2017 (exception of HadGEM2-ES: 1981-2017). The BIC values are normalized between 0 (best) and 1 (worst model). E: equal variance, V: variable variance.

Comment:

l.112: why adding the seasonal cycle of 2017 and not the seasonal cycle of 1979-2017? What is the reason beyond this choice? This is particularly relevant if the same is done for temperatures

Response:

The suggestion from the reviewer comment is relevant. Nevertheless, since we are looking at change with regard to a reference period, we expect that changing that reference period would not change the result significantly. The choice of 2017 is heuristic because it is the last year contained both in the reanalysis and climate models (the same was done for temperatures). Furthermore, the added seasonal cycle of 2017 is the estimation for this year based on the calendar trend over the 122 years, which is more robust (and smooth) than the raw values of the year 2017.

Comment:

l. 130 “They are also visually similar to the usual North-Atlantic weather regimes from the literature (e.g. Cassou (2008), Yiou and Nogaj (2004)).” I do not think that this is true. Weather regimes are defined by removing the seasonal cycle and mostly using only winter months, why here the “regimes” represent the seasonal cycle. I can not see any Atlantic ridge, or blocking regime here! R3 does not have a ridge over the Atlantic(see comments below Figure 2)

Figure 2: I do not see any resemblance of R1-4 with the weather regimes (e.g. from Cassou 2008, https://www.nature.com/articles/nature07286, Figure 1), and I would not expect to see any.

<u>Response:</u>
We are comparing the spatial patterns between our regimes and the “classical ones” from the literature (lines 129-135) and acknowledge that their definition and temporal properties are different (lines 135-138). Although our regimes are based on raw values, their spatial patterns are calculated similarly to classical regimes (i.e. average maps of daily fields (seasonal anomalies) belonging to the regimes). Our response regarding Figure 2 follows. For the regime pattern of Atlantic ridge (R3), although our models are slightly biased over the pattern, ERAI restitutes an Atlantic Ridge pattern that resembles that of Cassou (2008) with anticyclonic conditions over the Northeastern Atlantic and cyclonic conditions from the Southwest towards the Northeast of the North Atlantic. For the regime pattern of Scandinavian blocking (R4), our blocking pattern (R4) represents strong anticyclonic conditions over Scandinavia (as in Cassou (2008) and Yiou & Nogaj (2004)) and also over Eastern North-America (as in Vrac et al. (2014)). Therefore, we consider that it qualifies as “Scandinavian blocking” as found in the literature. Regarding the NAO+ (R1), it is very similar to the NAO+ pattern from Cassou (2008) with strong cyclonic conditions over the Northernmost part and moderate anticyclonic conditions over the Southern part of the North Atlantic. Similarly for our NAO- pattern (R2) with strong anticyclonic conditions over the Northernmost part and cyclonic conditions over the Southern part of the North Atlantic.

<u>Comment:</u>
l.162 What is increasing 90hPa? Z500 should be in m or gpm.

<u>Response:</u>
Thank you, we revised the manuscript accordingly (lines 161-162).
“Note that these patterns are relative to the seasonal mean, which increases substantially over the North Atlantic between the first and last three decades (averaging about +90 m for Z500 and +4°C for TAS; not shown.”

<u>Comment:</u>

l. 165 “GFDL-CM3 stands out from the other GCMs by showing the emergence in the future of a new summer regime that did not exist in the past.” I see actually a discrepancy between the GFDL-CM3 model for 1979-2008 (figure S8, solid line) and 1979-2017 (Figure 1), so I have trouble to understand this statement.

Response:
Yes, this discrepancy comes from the fact that these results are not directly comparable because Figure S8 shows results based on clustering over 1979-2100 (lines 156-157), whereas Figure 1 shows results based on clustering over 1979-2017 (line 125). On Figure S8, we show the average frequency calculated over 1979-2008 and 2071-2100 of the regimes that were determined by clustering over 1979-2100.

To clarify the manuscript, we propose the following revisions (changes in blue; lines 156-159):
“We now use the same method as before to define SWRs but based on the full simulation datasets over 1979-2100 to detect potential future changes. The first approach is to use four regimes (SWR4). Between the first three decades (1979-2008) and the last three decades (2071-2100) of the period, R1 (NAO+) occurs less often but is more intense for both Z500 and TAS (Supplementary Fig. 6-7). We emphasize here that the regimes are defined over 1979-2100 and that we investigate their main properties (weather patterns, annual cycle) over the subperiods (1979-2008 and 2071-2100) by selecting the results of the full-period clustering over these subperiods.”

Comment:
l. 173-177 It is not clear which trends are meant here and how they are calculated. I do not understand how these trends are calculated since the regimes are not continuous.More explanation is needed here.

Response:
The regimes are defined by clustering over 1979-2100 but they are indeed not continuous since there are periods of alternance with different days belonging to different regimes. The regime-conditional trends are therefore calculated by multiple linear regression according to the different regimes (y ~ x1 + x2 + x3 + x4) and then averaged between models (see lines 174-176 and the caption of Figure 11 in the Supplement). To clarify, we propose the following changes (in blue) in the manuscript (lines 174-175):
“These maps of linear trends are obtained by calculating the linear regression of the evolution of the variable (raw values) by gridcell, grey areas correspond to trends that are not significant (p-value > 0.05). The unconditional trend corresponds to the linear fit over the whole period (all days), whereas the regime-conditional trends are

calculated by multiple linear regressions to account for the distribution of days between regimes."

Comment:

l.183-184 Are the regimes (EOF and GMM) calculated for 1979-2100 and then sep-arated in 1979-2008 and 2071-2100? If so I would expect to see a change in the frequency, but not a change in the patterns. For the 7 regimes, I do not see a good correspondence between the patterns (e.g. R2 and R7 in Fig 3 and 4, both shading and contours). Also, I expect the period 2071-2010 to be warmer than 1979-2008, but there are no regimes with warm TAS. This might be linked to how TAS anomalies are defined (please see main comment 2).

Response:

Yes, both the PC1 and SWRs are calculated over 1979-2100. The additional calculations (regime frequency, weather patterns) are then done according to the subperiods. The reason we observe a change in the patterns is that these patterns are relative to the evolution of the seasonal cycle which increases a lot over the period (lines 161-162). Indeed, the Z500 pattern (color shading) represents the average of seasonal anomalies over the days belonging to the regime, with seasonal anomalies calculated by removing the average Z500 seasonal cycle to the raw Z500 values over a reference period. We disagree regarding the contour lines of raw Z500 according to R2 and R7 in Fig. 3 and 4, they are similar. The period 2071-2100 is indeed warmer and this is reflected in the results by having many more days of high Z500 or TAS fields (e.g. R7; see contour lines) than low (e.g. R1; ditto). We addressed how TAS anomalies are defined in the response to comment 2 above.

Comment:

l. 185 I see that R7 occurs in summer, but I do not see from any Figures that this is linked to blocking over Scandinavia (Z500 anomalies are over most of the North Atlantic). Also, the percentage (54 days or 0%) is very low compared to blocking frequency (see for example Figures 2 and 3 in Pfahl et al 2012). Moreover, why is R7 much more frequent in the future, but not showing any temperature anomalies? Heat-waves are expected to be linked to blocking also in the future (see e.g. Schaller et al., 2018).

Response:

Regarding the geographic conditions associated to R7, we propose the following changes (highlighted in blue) in the manuscript (lines 185-186):

"Past (1979-2008) R7 corresponds to rare and very intense anticyclonic conditions over the Northern half of the North Atlantic in association with summer heatwaves over the continents of the North Atlantic region (except North Africa and northernmost Canada."

Our methods are very different from those of Pfahl et al. (2012), so the results are not directly comparable. Regarding the blocking frequency over 1979-2008 (Fig. 3 of our paper), you can see that a few regimes (R4, R6) have a similar pattern to R7 (although less intense), meaning that blocking days were distributed between different regimes. R7 is more frequent in the future but doesn't show relatively strong temperature anomalies because the future seasonal cycle of Z500 in summer is very high. The methods from Schaller et al. (2018) are also very different from ours, so again the results are not directly comparable. We emphasize that our regimes represent differences of conditions relatively to the average seasonal cycle (that evolves), and that the study is not designed specifically to investigate extremes. Although the pattern of future R7 is less intense than in the past, it shows anticyclonic conditions analogous to blocking (+10 to +20 m), relatively to a seasonal cycle (summer) with much higher Z500 than in the past. The situation is similar for TAS.

Comment:

l. 218-220: I am not sure to understand the logic behind detrending the data and then calculate the trend of the detrended data. I would understand to i) detrend the data to compute TAS and Z500 anomalies and then compare these for 1979-2008 and 2071-2100 or ii) detrend the data to compute the regimes and look at the trends in the regime occurrence (e.g. trends of Figure 6 and 7), but I do not understand why computing the trends of the detrended anomalies.

Response:

The goal of computing the trends of the detrended anomalies is to investigate local trends that are residual to the regional trend (lines 114-118 and 212-215), in order to further investigate changes in Z500 (i.e. atmospheric circulation) patterns after removing the effect (i.e. thermodynamic) of the large-scale increase of Z500 due to human influence (Christidis & Stott (2015)). To clarify this in the manuscript, we propose the following changes (in blue; line 110):

"The goal now is to remove the large-scale increase of Z500 to further investigate changes in Z500 patterns. This requires to preserve both the spatial structures and the seasonality while removing the large-scale effect."

Comment:

l. 220: I can not see the disagreement between the models in Fig. S17

Response:

The large greyed areas in Fig. S17 correspond to the average of the p-values from model trends being superior to 0.05 (lines 218-220). This can be caused by a few individual models with largely insignificant trends (i.e. high p-values) or many models with insignificant trends (i.e. p-values superior to 0.05).

Comment:
l. 225-235: As I do not understand what has been done to "examine the regime spatial trends with LGI but without the seasonal shift" I can not comment on this part. Also,what are the "large-scale increases in Z500"?

Response:
The calculation of the spatial trends with LGI but without the seasonal shift is explained lines 228-231. It simply corresponds to taking the clusters defined based on detrended data, and to calculate their spatial patterns by using the original (non-detrended) data. The "large-scale increases in Z500" correspond to the thermodynamic response of Z500 to human-caused large-scale warming (lines 58-59 and 212-213).

Comment:
l. 240-245: Which Figure leads to this conclusion? Why are ERAI and CMIP5 models similar in Figure 1, but very different in the supplement (Figure S8, 1979-2008, solid lines). Also, what is the "increasing frequencies of historical summer conditions of atmospheric dynamics"?

Response:
This section discusses the results from the corresponding section (3.1 Evaluation of past seasonal weather regimes in climate models (1979-2017); lines 125-154) and is therefore supported by the corresponding findings (analysis, figures). For the same reason as explained previously in response to another similar comment, Figures 1 and S8 are not comparable since Fig. 1 is based on clustering over 1979-2017 while Fig. S8 is based on clustering over 1979-2100. The "increasing frequencies of historical summer conditions of atmospheric dynamics" refers to the increase in frequency of R4 over 1979-2017 (lines 148-154).

Comment:
l. 248-270 Which Figures are showing that? Please, add a reference to help to follow the train of thoughts. (I.e. where are cold spells and heatwaves in the analysis? I see only temperature anomalies of a few degrees)? Also, how can a regime be replaced?

Response:
As the previous section of the discussion (4.1) is based on the findings of the corresponding section (3.1), this section (4.2) is based on the corresponding findings (3.2) and discusses them with regard to the literature. We use the terms "cold spells"

and “heatwaves” to designate robust anomalies (average of more than 50 days i.e. 50 values) over large areas (continents) of about 3°C relatively to the average seasonal cycle. We write line 193 that “R7 … replaces R6”, meaning that future R7 happens at the same time as past R6 did (Fig. 5) and with similar spatial patterns (Fig. 3 and 4). To clarify the terms “cold spell” and “heatwave”, we propose the following changes (in blue) in the manuscript (lines 185-187):

“Past (1979-2008) R7 corresponds to rare and very intense conditions of Scandinavian Blocking associated with summer heatwaves over Northern continents.Future (2071-2100) R1 corresponds to rare and very intense NAO+ conditions associated with cold spells over Northeastern America, Greenland and Scandinavia. Here, we use the terms "cold spell" and "heatwave" to designate robust anomalies (average of more than 50 days i.e. 50 values) over large areas (continents) of about 3°C relatively to the average seasonal cycle.“

Comment:

Appendix C: Why not having everything in the section “Seasonal weather regimes based on detrended data”. Is the same trend removed at each grid cell? Calculated over which region?

Response:

Thank you for this interesting idea, we agree with the comment about moving the information of Appendix C to section 2.3. Yes, the same trend is removed at each grid cell, calculated over the whole region, but depending on calendar day. To clarify further, we propose the following changes in the manuscript (line 110):

“The goal now is to remove the large-scale increase of Z500 to further investigate changes in Z500 patterns. This requires to preserve both the spatial structures and the seasonality while removing the large-scale effect. Calculating and removing the trend by gridpoint would result in losing the spatial structures while doing so without a year of reference would result in losing the seasons. Therefore, the trend is calculated on the spatial mean of the whole area for each calendar day, with reference to 2017. This means that for each specific day of the calendar year (January 1st only, ..., December 31st only), the trend is calculated with the 122 values (from 1979 to 2100) of the spatial mean for this specific day. We took 2017 as the reference year because it is the last year contained in both reanalyses and models. The trend was estimated best by using a cubic smoothing spline.”

Comment:

I think it would be very useful to show EOF1 and the explained variance for ERA-Interim and CMIP5 and compare them, before starting calculating and comparing regimes.

Response:
Yes, see our response to the first comment.

Comment:
Figure 10: Why are TAS patterns opposite over land and ocean in 1979-2008 and 2071-2010? Is this because the same trend (if I understood it correctly) is removed at each grid point? I expect trends over land and over ocean to be very different (see e.g. Hegerl et al. 2018, Figure 1). Why we do not see this behaviour in Z500? Are the trends in TAS much larger than Z500?

Response:
For TAS patterns over land and ocean, yes, exactly! Continents are indeed warming much faster than oceans. This behaviour is absent in Z500 because it is mainly driven by the thermodynamic response to warming: almost uniform Z500 increase over the whole region (see Unconditional Z500 change in Figure S15, and Figure S19). The trends in TAS do not need to be much larger than Z500 to have this particular behaviour, it is mostly created by the land-sea differential warming (see Figures S18 and S20).

Comment:
Supplementary Figure 18: Why is the sum of the detrended trends not zero? I would expect some regimes to have positive trends and other regions to have negative trends at the same grid point. Since it is not clear how these trends are calculated, it is difficult to interpret these Figures.

Response:
The sum of the detrended trends is very close to zero, you can see it if you examine closely the scale of the colorbar (red corresponds to higher magnitude of values, compared to blue) and the distribution of trends (geographic, between regimes). This is because these local trends are calculated relatively to the average spatial trend over the whole region (lines 110-118). This figure (S18) is very related to Fig. 10, it shows the differential warming between land and ocean by regime (see our response to another comment above).

References:

Angell, J. K., & Korshover, T. (1974). Quasi-biennial and long-term fluctuations in the centers of action. Monthly Weather Review, 102(10), 669-678.

Barnes, E. A., & Hartmann, D. L. (2010). Dynamical feedbacks and the persistence of the NAO. Journal of the atmospheric sciences, 67(3), 851-865.

Barnston, A. G., & Livezey, R. E. (1987). Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly weather review*, *115*(6), 1083-1126.

Blackmon, M. L., Lee, Y. H., & Wallace, J. M. (1984). Horizontal structure of 500 mb height fluctuations with long, intermediate and short time scales. *Journal of the atmospheric sciences*, *41*(6), 961-980.

Cassou, C., Terray, L., & Phillips, A. S. (2005). Tropical Atlantic influence on European heat waves. Journal of climate, 18(15), 2805-2811.

Cassou, C.: Intraseasonal interaction between the Madden–Julian oscillation and the North Atlantic Oscillation, Nature, 455, 523–527, 2008

Cattiaux, J., Douville, H., and Peings, Y. (2013). European temperatures in CMIP5:origins of present-day biases and future uncertainties. Climate dynamics, 41(11-12),2889-2907.

Christidis, N., & Stott, P. A. (2015). Changes in the geopotential height at 500 hPa under the influence of external climatic forcings. *Geophysical Research Letters*, *42*(24), 10-798.

Estivill-Castro, V., & Yang, J. (2000, August). Fast and robust general purpose clustering algorithms. In *Pacific Rim International Conference on Artificial Intelligence* (pp. 208-218). Springer, Berlin, Heidelberg.

Franzke, C., Woollings, T., & Martius, O. (2011). Persistent circulation regimes and preferred regime transitions in the North Atlantic. *Journal of the Atmospheric Sciences*, *68*(12), 2809-2825.

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H. (2017). Balancing Europe's wind-power output through spatial deployment informed by weather regimes.Nature Climate Change, 7(8), 557-562.

Han, J., Kamber, M., & Pei, J. (2011). Data mining concepts and techniques third edition. *The Morgan Kaufmann Series in Data Management Systems*, 83-124.

Hannachi, A., Woollings, T., & Fraedrich, K. (2012). The North Atlantic jet stream: A look at preferred positions, paths and transitions. *Quarterly Journal of the Royal Meteorological Society*, *138*(665), 862-877.

Hegerl, G. C., Brönnimann, S., Schurer, A., and Cowan, T. (2018). The early 20th cen-tury warming: anomalies, causes, and consequences. Wiley Interdisciplinary Reviews:Climate Change, 9(4), e522.

Hu, Y., Li, D., & Liu, J. (2007). Abrupt seasonal variation of the ITCZ and the Hadley circulation. *Geophysical research letters*, *34*(18).

Hurrell, J. W., & Deser, C. (2010). North Atlantic climate variability: the role of the North Atlantic Oscillation. *Journal of marine systems*, *79*(3-4), 231-244.

Iqbal, W., Leung, W. N., & Hannachi, A. (2018). Analysis of the variability of the North Atlantic eddy-driven jet stream in CMIP5. *Climate Dynamics*, *51*(1-2), 235-247.

Kuang, X., Zhang, Y., Huang, Y., & Huang, D. (2014). Spatial differences in seasonal variation of the upper-tropospheric jet stream in the Northern Hemisphere and its thermal dynamic mechanism. *Theoretical and applied climatology*, *117*(1-2), 103-112.

Lior, R., & Maimon, O. (2005). Clustering methods. *Data mining and knowledge discovery handbook*, 321-352.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., ... & Visbeck, M. (2001). North Atlantic climate variability: phenomena, impacts and mechanisms. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *21*(15), 1863-1898.

Masato, G., Hoskins, B. J., and Woollings, T. (2013). Winter and summer Northern Hemisphere blocking in CMIP5 models. Journal of Climate, 26(18), 7044-7059.

Michel, C., and Rivière, G. (2011). "The link between Rossby wave breakings and weather regime transitions." Journal of the Atmospheric Sciences 68.8: 1730-1748.

Pfahl, S., and Wernli, H. (2012). Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-) daily timescales. Geophysical Research Letters, 39(12).

Price, J. M., & Magaard, L. (1986). Interannual baroclinic Rossby waves in the midlatitude North Atlantic. *Journal of physical oceanography*, *16*(12), 2061-2070.

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., & Knight, J. R. (2008). European climate extremes and the North Atlantic Oscillation. *Journal of Climate*, *21*(1), 72-83.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., & Russo, S. (2018). Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles. *Environmental Research Letters*, *13*(5), 054015.

Schneider, T., Bischoff, T., & Haug, G. H. (2014). Migrations and dynamics of the intertropical convergence zone. *Nature*, *513*(7516), 45-53.

Sillmann, J., & Croci-Maspoli, M. (2009). Present and future atmospheric blocking and its impact on European mean and extreme climate. *Geophysical Research Letters*, *36*(10).

Sillmann, J., Croci-Maspoli, M., Kallache, M., & Katz, R. W. (2011). Extreme cold winter temperatures in Europe under the influence of North Atlantic atmospheric blocking. *Journal of Climate*, *24*(22), 5899-5913.

Vrac, M., & Yiou, P. (2010). Weather regimes designed for local precipitation modeling: Application to the Mediterranean basin. *Journal of Geophysical Research: Atmospheres*, *115*(D12).

Vrac, M., Vaittinada Ayar, P., & Yiou, P. (2014). Trends and variability of seasonal weather regimes. International journal of climatology, 34(2), 472-480.

Wallace, J. M., Zhang, Y., & Lau, K. H. (1993). Structure and seasonality of interannual and interdecadal variability of the geopotential height and temperature fields in the Northern Hemisphere troposphere. *Journal of Climate*, *6*(11), 2063-2082.

Wang, C. (2002). Atlantic climate variability and its associated atmospheric circulation cells. *Journal of climate*, *15*(13), 1516-1536.

Woollings, T., Hannachi, A., & Hoskins, B. (2010). Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 856-868.

Woollings, T., & Blackburn, M. (2012). The North Atlantic jet stream under climate change and its relation to the NAO and EA patterns. *Journal of Climate*, *25*(3), 886-902.

Xian, P., & Miller, R. L. (2008). Abrupt seasonal migration of the ITCZ into the summer hemisphere. *Journal of the atmospheric sciences*, *65*(6), 1878-1895.