Response to Referee Comment 2 *on* **"Seasonal weather regimes in the North Atlantic region: towards new seasonality?"** *by* **Florentin Breton et al.**

Comment:

The authors utilize the concept of seasonal weather regimes defined based on clustering daily fields of 500 hPa geopotential height from reanalysis and model data. By exploring their frequency and distribution throughout the year, they are able to trace changes in seasonality of European/North Atlantic climate, which are further explored in the course of the study. In addition to the material presented in the main text, the supplementary material contains an impressive body of additional figures illustrating the results of this study in more detail and may be helpful to better understand certain aspects of the authors' findings. The manuscript itself is concise, well-written and contains (almost) all necessary information needed to fully comprehend the presented analysis. I only have a few minor comments listed below that I would like to invite the authors to consider/address before this work may become published as a regular paper in ESD.

Response:

We thank the referee for the positive, constructive and helpful comments and detailed remarks. 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:

My main point concerns some minor clarification on the use of the EM algorithm for clustering the daily Z500 fields. Specifically, I have two questions:

1. Could you please elaborate a bit more explicitly on the meaning of the variable x that you use for clustering? Is it really just the (scalar) PC1 score (i.e. you collapse the full spatial pattern to the magnitude of the EOF-1 pattern), and you perform a one-dimensional cluster analysis?

Response:

Yes, we use only PC1 values in our clustering because our focus is on seasonality which is overwhelmingly captured by PC1 (see lines 82-88). In the preliminary steps of the analysis, we tested the sensitivity of the results to the number of PCs included in the clustering. In most cases, including a few more PCs (e.g. from 2 to 5 PCs) brings similar results (weather patterns, seasonal cycle). Therefore, we use only PC1 and the cluster analysis is indeed one-dimensional (see lines 92-93). We propose the following revisions (blue) in the manuscript (line 108):

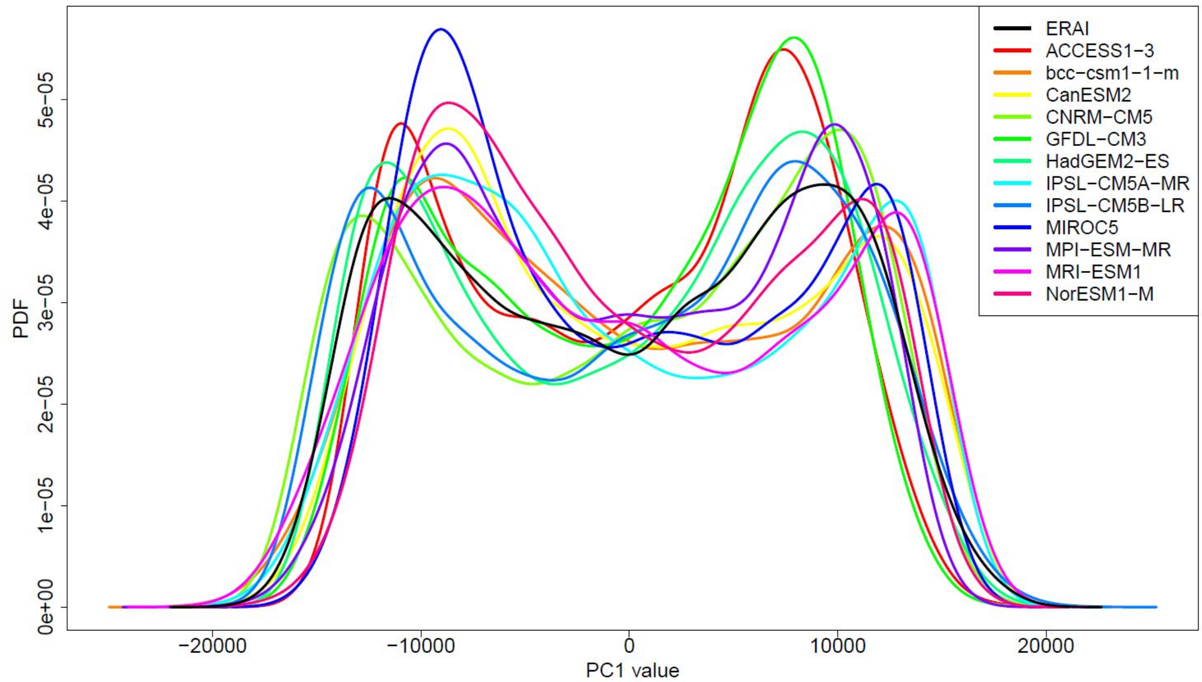
“*We also tested the sensitivity of the clustering results (weather patterns, annual cycle) to the number of PCs included (from 1 to 5). There was a small influence of additional PCs on the results (reanalyses, models) over 1979-2017 and very small influence over 1979-2100 (increasing with the number of PCs; not shown). This reinforced our choice of using only the first PC, considering that additional PCs represent little additional seasonality and difference in the long-term response of atmospheric circulation to climate change.*”

Comment:

2. I would appreciate it if you could motivate a bit more explicitly your use of a model-based clustering technique as opposed to other model-free approaches like hierarchical or k-means clustering. How is this related to the choice/type of the variable x that the clustering is based on? Why is a finite Gaussian mixture a reasonable model for the designated purpose? I think that the obtained results are correct and reasonable, I just want to better understand the rationale behind the approach followed here.

Response:

The choice of a finite Gaussian mixture is justified by the distribution of PC1 (x) that is Gaussian-like and bimodal (seasons), as shown on the figure below.



K-means is a distance-based clustering method that assumes spherical clusters and is more sensitive to noise (Estivill-Castron & Yang, 2000; Han et al., 2011; Rokach & Maimon, 2005). EM is a probabilistic model-based clustering method that can be seen as more statistically sophisticated than k-means, although it requires more computation and detailed prior knowledge (Estivill-Castron & Yang, 2000). The EM approach generalizes k-means in allowing more flexibility in the covariance structure of the clusters (Banfield and Raftery, 1993; Rust et al., 2010): k-means corresponds to a special case of EM with spherical Gaussian laws that is less able to capture non-spherical structures. Hierarchical clustering was considered less suitable than EM because of the inability to scale well with a large number of values (about 45000 for PC1 over 1979-2100; Rokach & Maimon, 2005) and with unsupervised decisions of merging or splitting without examining or evaluating many objects or clusters (Han et al., 2011). In the case of EM using only PC1 (univariate data), the initial partitioning of data is done by separating in quantiles, so it is less sensitive to noise, outliers and extremes. We tried both EM and k-means algorithms in the preliminary stage of our study to test the sensitivity of the results to the clustering method, they brought similar results so we focused on EM. We propose the following revisions (blue) in the manuscript (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:

3. I found the description of the detrending procedure a bit hard to follow in the maintext. I understand from Appendix C that you take the daily mean of the whole Z500 field and subtract it from all values (thereby removing not only long-term trends, but also seasonality in the regional-mean Z500 amplitudes)? Or is it something different?I have to confess that I am a bit lost with the term "calendar trend" (ll.111 and 215).

Response:

Yes, to clarify the description of the detrending procedure, we propose the following revisions (blue) in the manuscript (lines 110-112) and the removal of "Appendix C":

> *"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:

4. There are quite a few cases where additional technical results are "not shown". Since you provide a very detailed supplementary material with many additional figures, I was wondering if it would make sense to also include (some of) the results labelled as "not shown" in this supplement?

Response:

We considered at length which figures to show in the paper (and supplement) and chose to remove those that show very similar results or bring little more information to the analysis, trying to find the balance between making the research investigation understandable and keeping the paper readable (not too long).

Comment:

5. In Figure 3.1, I would find it more logical to start with showing and discussing the obtained SWRs (i.e., the associated spatial patterns) before focusing on their past changes.

Response:

If we understand correctly, the comment from the reviewer suggests to show and discuss the spatial patterns of SWRs before looking at their annual cycle in Section 3.1. We agree with the comment and propose the following revisions (blue) in the manuscript (lines 125-138 and Figures 1-2):

We start by looking at the weather patterns of the regimes as shown by the composite maps in Figure 1. 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. The spatial patterns of the four average regimes found in the models are very similar to those from ERAI. They are also visually similar to the usual North-Atlantic weather regimes from the literature (e.g. Cassou (2008), Yiou and Nogaj (2004)).The first regime (R1) corresponds to the positive phase of the North Atlantic Oscillation (NAO+) and the second (R2) to its negative phase (NAO-; Hurrell et al. (2003)). The third and fourth regimes (R3 and R4) respectively represent the AtlanticRidge (AR) and Scandinavian Blocking (SB) atmospheric conditions, resembling the weather patterns from Yiou and Nogaj (2004), and Vrac et al. (2014). However, note that the temporal patterns of our SWRs are based on full years (like Vrac et al. (2014)), unlike the literature considering weather patterns in winter (Cassou (2008), Yiou and Nogaj (2004)) or in summer(e.g. Guemas et al. (2010)). Thus, if our SWRs resemble the usual regimes, they present differences in their definition and properties. The *annual cycle of our regimes' monthly frequencies over 1979-2017 is shown in Figure 2. Climate models reproduce a seasonal cycle of SWRs similar to ERAI, with regime 1 (hereafter R1) representing a winter-like season,R4 a summer-like season, and R2 and R3 transitional seasons (R2 around winter and R3 around summer).*

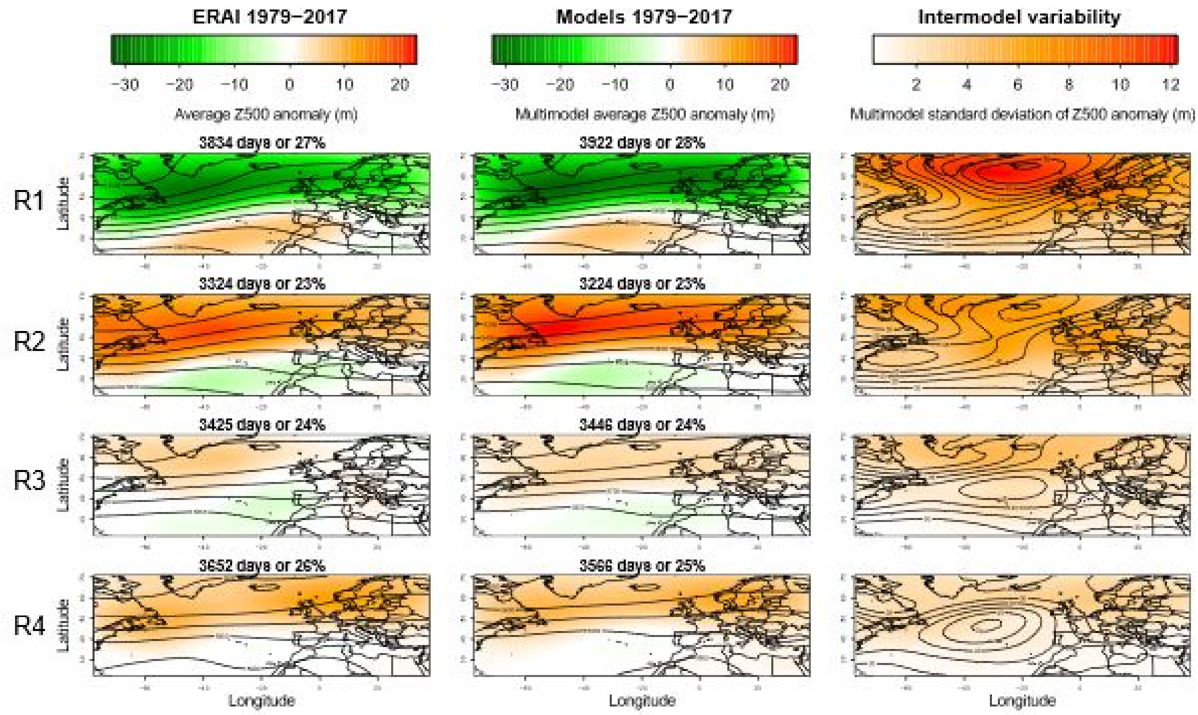


Figure 1. Composite maps of the four regimes (one per row) for ERAI (first column) and climate models (second column; each map shows the average of 12 composite maps; third column shows standard deviation of Z500 between the 12 composites). The maps are calculated by averaging the seasonal anomalies (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).

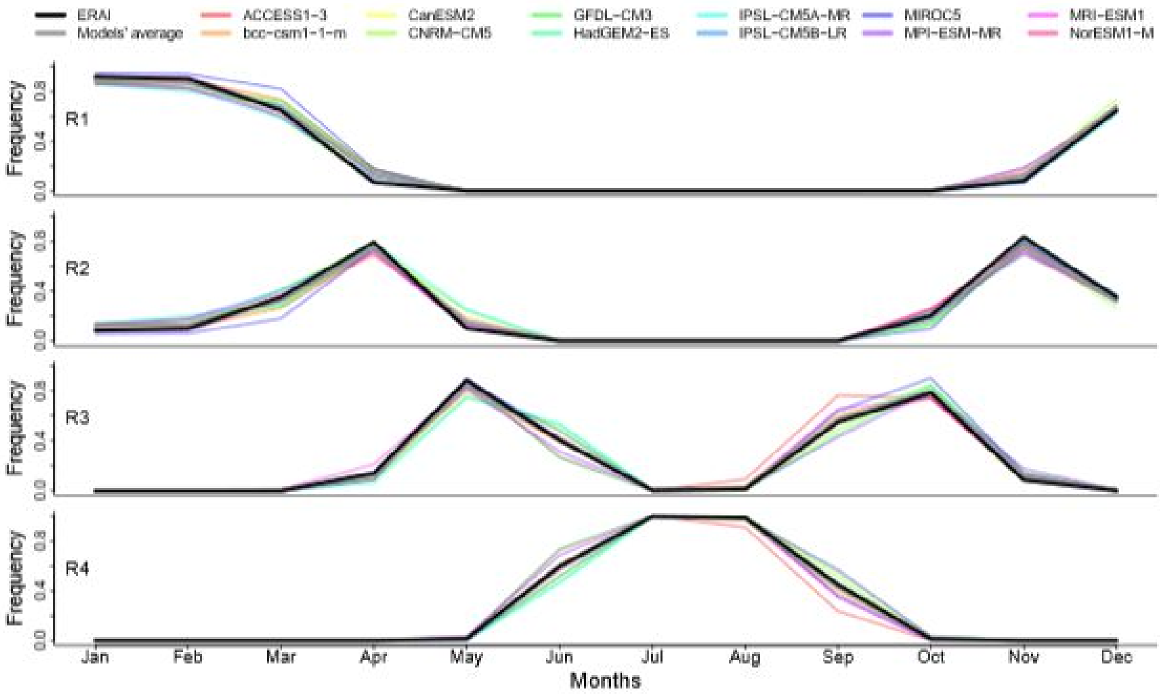


Figure 2. Average seasonal cycle of the frequencies of occurrences for the four regimes of ERAI and the 12 climate models, over 1979-2017. Monthly frequencies correspond to the number of days of regime occurrence divided by the number of days in the month.

Comment:

6. Relating to ll.139-147: In order to better understand the (dis)agreement between the Z500 fields and SWR distributions obtained from them, some quantitative metrics(like mean spatial correlation, bias,...) between the models and ERA-Interim could be added as a table to the supplementary material.

Response:

Thank you for this interesting idea. We calculated the coefficients of pattern correlation between the regimes from the different models and ERAI, and propose to include these results in Table 2 of the main materials of the paper (with the following changes highlighted in blue; lines 139-140):

"In general, the climate models reproduce atmospheric weather patterns that are very similar to ERAI, but individual models are less successful (see Table 2 and Supplementary Fig. 1-4)."

Table 2. Coefficients of pattern correlation between the regimes from ERAI and each climate model over 1979-2017.

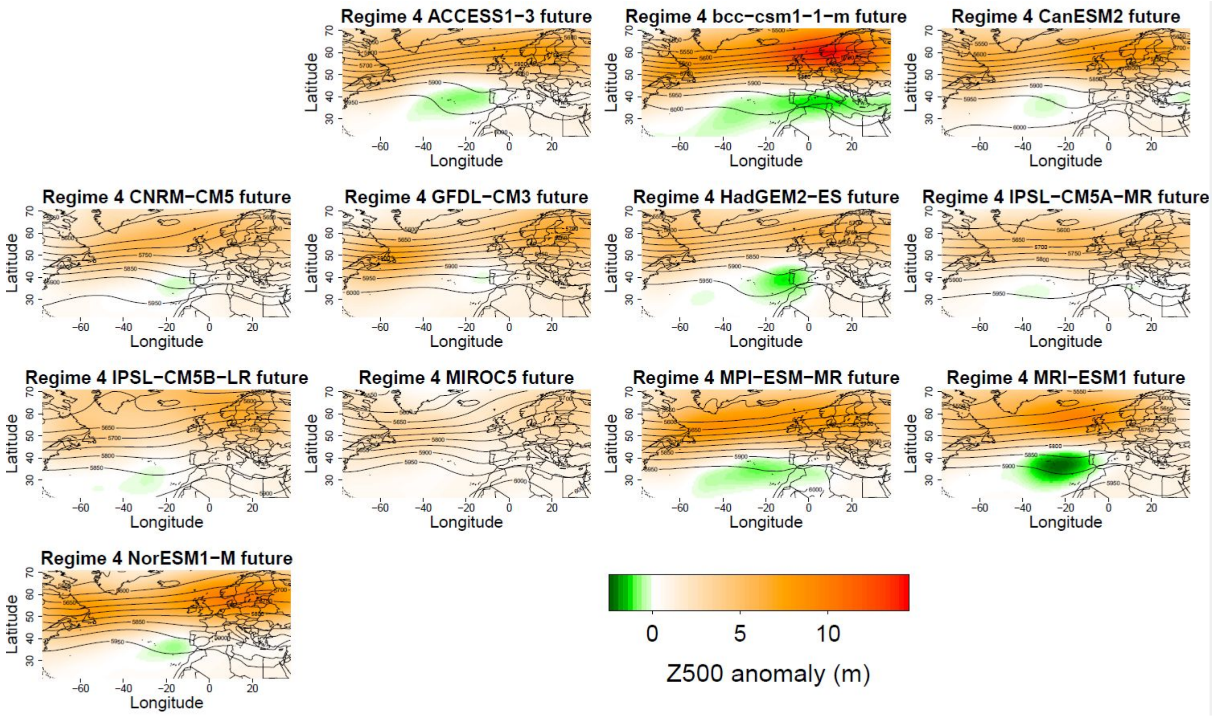
Model	R1	R2	R3	R4
Average	0.901	0.869	0.518	0.9
ACCESS1-3	0.944	0.888	0.275	0.805
bcc-csm1-1-m	0.928	0.842	0.757	0.959
CanESM2	0.935	0.888	0.781	0.931
CNRM-CM5	0.937	0.933	0.589	0.962
GFDL-CM3	0.934	0.868	0.486	0.932
HadGEM2-ES	0.926	0.879	0.887	0.916
IPSL-CM5A-MR	0.908	0.874	0.529	0.948
IPSL-CM5B-LR	0.912	0.859	0.65	0.933
MIROC5	0.553	0.674	-0.397	0.824
MPI-ESM-MR	0.968	0.934	0.59	0.837
MRI-ESM1	0.913	0.86	0.753	0.815
NorESM1-M	0.95	0.925	0.319	0.941

Comment:

7. L.165: Can you elaborate a bit more on the "new summer regime that did not exist in the past"?

Response:

The main finding expressed here (and developed further in the paper) is the emergence of a new regime of atmospheric circulation that was not present in the historical period. As the clustering has little freedom with 4 clusters, this emergence is more evident with 7 clusters (as shown in the later results). An interesting point is that GFDL is the only model showing the emergence of a new regime with only 4 clusters (i.e. despite large constraints on the definition of the regimes). It means that in the case of GFDL, the difference between historical and future Z500 conditions in summer is so large that a new regime was created by the clustering algorithm. However, future R4 appears relatively similar between models (see Figure 4 below from the Supplementary materials of the paper):



Since future R4 in GFDL is very similar to future R4 from other models (for which R4 was already well established in the past), the emergence of R4 in GFDL doesn't represent the emergence of a new regime from a general perspective. We propose the following changes (highlighted in blue) in the manuscript (lines 164-165):

> *"GFDL-CM3 stands out from the other GCMs by showing the emergence in the future of a new summer regime that almost did not exist in the past (one day, not shown). This emergence means that in the case of GFDL, the difference between historical and future Z500 conditions in summer is so large that a new regime was created in the clustering. As the clustering has little freedom with 4 clusters (i.e. large constraints on the definition of the regimes), this emergence is even more interesting, but it is consistent with higher increases of Z500 and TAS in this climate model by comparison to other models (not shown). However, since future R4 in GFDL is very similar to future R4 from other models (seasonal cycle and weather pattern, respectively shown in Supplementary Fig. 8 and not shown), and since R4 was already well established in the past for other models, this emergence of R4 in GFDL does not represent the emergence of a new regime from a general perspective."*

Comment:

8. The authors consider future changes in SWRs over the European/North Atlantic sector for the RCP8.5 scenario only. Did they check if the corresponding results for more moderate scenarios would be compatible with the reported findings (i.e. show consistent trends but smaller "magnitude" of changes)? I do not request to show any additional results for RCPs, but a brief discussion (e.g. along with what is stated in ll.308-309) could be interesting.

Response:
RCP8.5 was initially chosen because it contains the largest signal of anthropogenic climate change and therefore facilitates the emergence of long-term changes in the climate system (as in many climate studies). We decided not to try other scenarios because it was expected (as said by the reviewer) to find similar trends of smaller magnitude to RCP8.5 (increasing in proportion to the scenario forcing), and the later emergence if any of the new regime in summer (increasing earliness in proportion to the scenario forcing). As the current trajectory of the Earth system already approximately follows the RCP8.5 scenario so far (Allen et al., 2018), the choice of scenarios representing its future trajectory is debated (Burgess et al., 2020; Hausfather & Peters, 2020), but RCP8.5 appears to remain a solid figure for the decades to come (Schwalm et al., 2020).

We propose the following changes (highlighted in blue) in the manuscript (line 80):

> "*The choice of the RCP8.5 scenario is motivated by its approximative representation of the current climate trajectory (Allen et al., 2020), its plausibility for future climate trajectory (Schwalm et al., 2020), and its large magnitude of scenario forcing for facilitating the emergence of long-term changes in the climate system.*"

Comment:
** L.2: "insight into"*

Response:
Thank you, we revised the manuscript accordingly (line 2).
"As both have strong seasonal features, a better insight into their future seasonality is essential to anticipate changes in weather conditions for human and natural systems."

Comment:
** L.3: What do you mean by "for human and natural systems"? This reads a bit odd to me...*

Response:
We are differentiating between the impacts on nature (e.g. phenology of ecosystems) and society (e.g. health, transportation, energy). We propose the following changes (highlighted in blue) in the manuscript to make it clearer (lines 3 and 28-29).

“As both have strong seasonal features, a better insight into their future seasonality is essential to anticipate changes in weather conditions for human systems (e.g. health, transportation, energy) and natural systems (e.g. phenology).”

“The meteorological seasons are a prominent feature of climate variability, experienced by human systems (e.g. health, transportation, energy) and natural systems (e.g. phenology) through the seasonality of surface weather conditions.”

Comment:
* *L.125: “the regimes’ monthly frequencies”*

Response:
Thank you, we revised the manuscript accordingly (line 125).
“We start by looking at the annual cycle of the regimes' monthly frequencies over 1979-2017, shown in Figure 1.”

Comment:
* *L.176: “regression coefficients”*

Response:
Thank you, we revised the manuscript accordingly (line 176).

“Both regression coefficients and p-values are calculated individually by climate model, and then averaged over the twelve values.”

Comment:
* *L.186: “the northern continents” – please be a bit more specific, as you do not seem to consider the whole hemisphere*

Response:
Yes, we revised the manuscript accordingly (line 185-186).

“Past (1979-2008) R7 corresponds to rare and very intense conditions of Scandinavian Blocking associated with summer heatwaves over the continents of the North Atlantic region (except North Africa and northernmost Canada).”

Comment:
* *Ll.189-190: “the R1 pattern...and the R7 pattern”*

Response:
Thank you, we revised the manuscript accordingly (lines 188-191).

"Overall, we observe a shift in the spatial patterns (Z500 and TAS) of the regimes (Figures 3-4) with past R1 patterns becoming future R2 patterns, past R2 patterns becoming future R3 patterns, and so on until R6, while the R1 pattern becomes seasonally more extreme (rarer and more intense pattern) and the R7 pattern becomes seasonally more normal (more frequent and less intense pattern)."

Comment:
** L.195: "the regimes' spatial patterns"*

Response:
Thank you, we revised the manuscript accordingly (lines 188-191).

"This shift in the seasonal cycle of the regimes between past and future appears very consistent with the shift in the regimes' spatial patterns."

Comment:
** L.225: I don't quite get what "LGI" stands for – please explain*

Response:
We just noticed that "LSI" should be used to designate "large-scale increase" instead of "LGI" and propose to replace the term in the manuscript. LSI stands for the "large-scale increase" of TAS and Z500 driven by human influence (lines 58-59; Christidis & Stott, 2015), corresponding for Z500 to the atmospheric thermodynamic response to human-caused global warming.

Comment:
** L.227: "the regimes' spatial trends" (what is a "spatial trend"?)*

Response:
Thank you, we revised the manuscript accordingly. We used the term "spatial trends" to denote the geographic variation of the trends calculated for each gridpoint. To clarify the manuscript, we propose to replace the term by "spatial pattern of trends" or "trends" where suitable.

Comment:
** L.449, Eq. (A4): Please replace $\sum_{k}^{t+1}$ by Σ_k^{t+1} in the LATEXsource, the presently shown mathematical symbol is not appropriate here*

Response:
Thank you, we revised the manuscript accordingly (Equation 4).

"\begin{equation}

\Sigma_k^{t+1} \, = \, \frac{1}{n \, \pi_k^{t+1}} \, p_{ik}^t \, (\, x_i \, - \, u_k^{t+1} \,)' \, (\, x_i \, - \, \mu_k^{t+1} \,)
\end{equation}”

References:

Allen, M.R., O.P. Dube, W. Solecki, F. Aragón-Durand, W. Cramer, S. Humphreys, M. Kainuma, J. Kala, N. Mahowald, Y. Mulugetta, R. Perez, M. Wairiu, and K. Zickfeld, 2018: Framing and Context. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press

Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 803-821.

Burgess, M. G., Ritchie, J., Shapland, J., & Pielke Jr, R. (2020). IPCC baseline scenarios over-project CO2 emissions and economic growth.

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.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.

Hausfather, Z., & Peters, G. P. (2020). Emissions–the ‘business as usual’story is misleading.

Rokach L., Maimon O. (2005) Clustering Methods. In: Maimon O., Rokach L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA. https://doi.org/10.1007/0-387-25465-X_15

Rust, H. W., Vrac, M., Lengaigne, M., & Sultan, B. (2010). Quantifying differences in circulation patterns based on probabilistic models: IPCC AR4 multimodel comparison for the North Atlantic. *Journal of Climate*, *23*(24), 6573-6589.

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8. 5 tracks cumulative CO2 emissions. *Proceedings of the National Academy of Sciences*, *117*(33), 19656-19657.

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8. 5 tracks cumulative CO2 emissions. *Proceedings of the National Academy of Sciences*, *117*(33), 19656-19657.

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