

Interactive comment on “A regime view of future atmospheric circulation changes in Northern mid-latitudes” by Federico Fabiano et al.

Anonymous Referee #2

Received and published: 23 October 2020

This paper deals with the changes in weather regimes in the Euro-Atlantic and Pacific-North American sectors, in both CMIP5 and CMIP6 models. A comparison of simulated and observed weather regimes is done and then the changes in frequency and persistence of the weather regimes are characterized. Potential drivers of those changes are finally discussed. This is a nice paper, well written, with interesting results. The paper is concise, despite the large amount of work necessary to do these analyses. I recommend publication of the paper after some moderate revisions.

Main comments

The weakest part of the paper (but also, potentially, the most interesting one) is probably section 4.2. The analyses themselves are interesting but I think the interpretation of the results should be more cautious. It is not because a significant (?) correlation be-

C1

tween changes in weather regimes frequency and what the authors call the "drivers" is found that a causal relationship exists. The weather regimes could themselves impact these drivers, and other factors, not studied by the authors may physically explain both the changes in weather regimes and in the drivers. More drivers could also have been considered: e.g. many studies have shown the potential impact of SST anomalies (especially in the Tropics, but not only) on large scale circulation, including on weather regime occurrence. I think it is difficult to discuss the potential drivers of weather regime changes without investigating the role of regional SST changes.

I also think that a discussion on the potential causes of the differences seen between CMIP5 and CMIP6, which are sometimes quite large, should be added to the paper. An interesting analysis would be to evaluate whether differences between the changes in weather regimes are linked to differences in the changes of the drivers between CMIP6 and CMIP5 (if the authors are right about the drivers, it should be the case).

Many methodological choices are necessary in a weather regime analysis: e.g. pre-processing through EOF analysis or not, if yes number of principal components retained, precise algorithm (many variants of k-means exist), choice of the number of clusters, whether all the days are classified or not, if not what is the criteria for not classifying some days, which variable is used (slp, zg), how to best take into account the mean increase in zg due to global warming etc. I somewhat understand why the authors don't want to discuss all these aspects, show sensitivity tests etc. The paper would be too long and less interesting. But nevertheless some aspects deserve to be better justified, at least with a sentence. Sensitivity tests could also be added in SI (as the authors decided to add SI to the paper). For example, the authors have chosen to first apply an EOF decomposition to zg and then to retain the 4 leading EOFs. They give no justification for this pre-processing step. The use of principal components for cluster analysis has been common in the past but I suspect that in many old studies, it was more a question of dimension reduction in order to run classification algorithms on computers with limited resources. The use of principal components is also some-

C2

times intended as a kind of filter. If it the case here, the authors should explain their objective(s). It is also unusual to retain such a small number of principal components for clustering, and therefore such a low level of explained variance. I would like to see a discussion on this point and potentially the result of a sensitivity test: e.g. changes in weather regimes without EOF analysis, or with enough principal components to keep at least 90% of the variance etc.

Finally, it would also be interesting to characterize the impact of internal variability on the changes in weather regime occurrence, as signal-to-noise may be small regarding atmospheric circulation changes. It would help to better evaluate the importance of future changes, of the differences between CMIP6 and CMIP5, and between scenarios etc. There are a few CMIP6 models with at least 10 members for which Zg is provided (for future scenarios). It would be very interesting to use the members of these models to characterize the impact of internal variability (for a single scenario, ssp585 for example, this is sufficient).

Specific comments

L15. Please define precisely what is meant by "low-frequency" in the paper.

L103. I think that Table S1 with the models and members used should be in the main document.

L105-108. I'm not sure it is good idea to combine the results of two reanalyses (ERA40 and ERA Interim). Moreover, both of these reanalyses have now been superseded by ERA5, with major improvements in ERA5. I understand that ERA5 is not (yet) available before 1979, but I'm not sure why the authors absolutely need to start in 1964. They could use the period 1979-2015. It would lead to a smaller sample, it is true, but I also think that the hypothesis of a linear trend (section 2.2.1) may be more reasonable on a shorter than on a longer period.

L118. Please explain (in the paper) why you need to "detrend" the data (by the way,

C3

I'm not sure that detrend is the good word considering what is done). And please justify why you use the average on the Northern Hemisphere (30-90N), rather than the averages on the domains of classification, as it is usually done. I think it may have some non-trivial implications, for example if the mean increase in geopotential height, largely controlled by warming, is different between the Euro-Atlantic and Pacific-North American sectors, as it may favor artificially some weather regimes in both domains.

L132-134. See my main comments. Why do the authors only use 4 EOFs for around 50% of explained variance?

L136. I'm not sure that it is a very good justification, but OK. . .

L142. How is the centroid exactly defined: average, median, real day closest to the average etc?

L145. Are all days classified and why? In some studies, the transition days are not classified, for example.

L146. Computed regimes: how are the computed regimes associated with the observed regimes? Is the associated observed regime the one with the stronger spatial correlation? Are there ambiguities? (e.g. a computed regime that looks like something intermediary between two observed regimes)

L176-177. Not clear to me.

L223-22. I'm not sure to understand the reasoning.

L263. "Further analysis. . ." I understand, but it is unfortunate. It is a very interesting (and strange) results. Discussing some hypotheses would have been nice. Note that it may be relevant regarding the discussion about the "drivers", as the "drivers" could be implicated in these variations (if they really are "drivers").

L315. I think the mean changes in geopotential height at 500 hPa (maps) should be shown somewhere, maybe as the first figure of the paper.

C4

L320 and legend of figure 8. It is not clear what the "trend" is here. In section 2.2.1 the trend is defined as filtered area-weighted average Northern Hemisphere geopotential height, and used to detrend the local geopotential height. Therefore, if my understanding is correct, here the authors look at the zonal mean "trend of detrended" data, right? It is not clear, and bit awkward from a vocabulary point of view. Please indicate the period used to compute the trends in the legend.

L323: zonal mean trend or zonal mean trend anomaly, as said in the legend?

L351. For the predictors I suppose that you also use the trends, as for the predictands (it is not said explicitly)? What is the period on which the trends are calculated for the predictors and the predictands? 2015-2100 I guess, but it is not explicitly said I think (also in the legend of Fig 10).

Section 4.2. Are there correlations between changes in EAT and PNA regime frequencies? I think knowing that might be useful for the discussion in this section.

L352. Why 2 or 3? Please justify. Are (i) all the models with the ssp585 scenarios used, or, at it seems to be the case based on Table S1 (ii) only the models with the 4 scenarios are used even for this analysis? I think (i) would be much better as it would lead to a larger sample of models, which is quite small with (ii). It is even truer since some of the models are nearly duplicates: models at different resolutions, ESM and AOGCM from the same group etc., which decreases the "effective" sample size.

L362. I don't see figure S7. Are these correlations significant (with the issue of effective sample size mentioned above it is difficult, or impossible, to do the test right, but it is still an interesting indication).

L387. Any idea of the reason(s) that might explain the improvements in CMIP6? Please discuss.

Interactive comment on Weather Clim. Dynam. Discuss., <https://doi.org/10.5194/wcd-2020-37>, 2020.