

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

Anonymous Referee #1

Received and published: 10 September 2020

1 General comments

This study investigates future changes in Atlantic-European and Pacific-North American weather regime occurrence according to CMIP projections. In a first part, the authors evaluate how well the models represent weather regimes (WRs) in their historical simulations compared to reanalysis, with a particular focus on how models from the 6th CMIP phase have improved compared to the 5th phase. In a second part, they investigate how WR frequency and persistence changes by the end of the century. In a last part, they trace these changes back to changes in the atmospheric mean state. The study provides an important basis for understanding most recent projections of future changes in surface weather from a large-scale dynamics point of view and opens

C1

up interesting questions for further research. Moreover, the detailed analysis of model biases can be a good guidance for the CMIP community to further improve their models. The paper has a clear and logical structure and is comprehensibly written. The methodological procedure is thorough and transparent. Aside from one major concern, I only have a (relatively large) set of minor comments that should be addressable relatively easily by providing some further explanation or making small adjustments in the text or figures. Therefore, I suggest the paper to be published after considering this major comment and the list of minor comments.

2 Major comment

- Section 4.2 (Drivers of future circulations): Although I appreciate the attempt to understand the origins of the projected changes in weather regimes (WRs) in more detail, I find this particular analysis in its current form not convincing enough from a causality perspective. In my opinion, this starts with calling the four indices (UTW, AA, PST, NAW) “remote drivers” and using them as statistical predictors for “predicting” WR changes based on linear regression. The reason is that there are strong dynamical links between mid-latitude storm track activity (and thus the WRs) and these indices. For instance, Ambaum and Hoskins, 2002 (<http://shorturl.at/ipDO9>) suggested a strong coupling between the NAO and the stratospheric polar vortex (which can be used as a proxy for your PST to first order) in the sense that a positive NAO can trigger a strong polar vortex, which in turn can strongly couple with the troposphere and induce persistent periods of positive NAO. Similarly, the effect of weak stratospheric polar vortex states on the troposphere (and thus on WRs) has shown to be strongly influenced by synoptic activity or WR occurrence beforehand (e.g., Kodera et al., 2016, <https://doi.org/10.1002/2015JD023359>; Domeisen et al., 2020, <https://doi.org/10.5194/wcd-2019-16>). Along the same lines, Garfinkel et

C2

al., 2015 (<https://doi.org/10.1002/2015JD023284>) showed that for instance SST anomalies (and thus the tropospheric state / storm track activity) can contribute to Arctic lower-stratospheric temperature changes (i.e., your PST). Likewise, I find it surprising to see such a strong link between the NAW and the PNA sector WRs. Although you mention that previous studies show a similar effect of North Atlantic temperatures on the troposphere in the Pacific sector – could it be that the link partly also acts the other way round, i.e. that the occurrence of certain PNA sector WRs (or a certain ENSO state) affects the SSTs in the North Atlantic and thus the NAW? At least this would be intuitive from a storm-track dynamical point of view, as the PNA sector WRs strongly influence the entrance of the North Atlantic storm track. Having these strong dynamical links between the four indices and the tropospheric dynamics in mind, I suggest that you either discuss / address this explicitly in your manuscript (by also “weakening” all the causality statements) or, optimally, that you gain some more insight in the causality by doing some kind of linear regression analysis considering time lags (similarly to Section 2.2. in Manzini et al., 2014, <https://doi.org/10.1002/2013JD021403>), if this a possible approach in your framework. The latter analysis may help to shed some more light on this chicken-and-egg-like problem. Summarizing my comment in other words: I think it is very helpful for further research to include these four indices into your study, but I just think one should treat them more as phenomena that may be strongly and mutually coupled to the WRs themselves.

3 Minor comments

- L28-30: I see that several studies find a poleward shift of the upper-level jet caused by the UTW. You additionally mention an intensification of the upper-level jet due to the UTW. However, it is not obvious to me why a stronger meridional temperature gradient in the upper troposphere strengthens the jet on the same

C3

level? According to the thermal wind balance, the upper-level jet should primarily be driven by the meridional temperature gradient in the lower troposphere (as, for instance, Hassanzadeh et al., 2014, considers by looking at changes of near-surface meridional temperature gradients on jet intensity). Can you explain this from a dynamical point of view?

- L32: I would add Pithan and Mauritsen, 2014 (<https://doi.org/10.1038/ngeo2071>) to the reference of Screen and Simmonds, 2010, who discussed the mentioned “several other positive feedbacks” in more detail.
- L75-76: I would add Michelangeli et al., 1995 (<http://shorturl.at/kmKW2>) here.
- L76: I would reword the sentence “each WR has a different impact on the climate of the downstream region” a bit, because it sounds like WRs are defined in a domain upstream, e.g., over the North Atlantic, to investigate surface weather downstream, e.g., over Europe.
- L76: I suggest to rename the abbreviation for the Pacific-North American sector in the whole manuscript to something like PAC or PNAM, because using PNA becomes confusing later on due to the two equally named regimes PNA+ and PNA- (for instance, at the end of line 214 it is not unambiguous whether you talk about all four PNA sector WRs or only PNA+/-).
- L70-85: Could you give a very brief summary (two to three sentences) of previous studies investigating WR changes in GCM simulations / CMIP projections? I think you partly do that later in the results section, but it may be helpful to get an overview of studies here already.
- L116-118: Can you elaborate a bit more on why it is necessary to detrend the historical data with the described approach before identifying the WRs? More specifically: How robust / meaningful is the described linear trend in the Northern Hemisphere area-averaged geopotential height, considering for instance the

C4

substantial multi-decadal variability in the large-scale circulation over these 50 years? Could it be that the trend (and thus the WR identification) becomes significantly different when considering, for instance, only the last 40 instead of 50 years (which is often done when investigating the ERA-Interim period only)? How reasonable is it to detrend, for instance, the North Pacific with a linear trend that is obtained from an area average over the whole Northern Hemisphere (including the North Atlantic)? Does your WR pattern identification change if you do not detrend the historical data (which is often done in other studies to my knowledge)?

- L122-123: If I understand correctly, you ultimately apply the EOF analysis to unfiltered daily Z500 anomalies, right (apart from the running mean climatology you subtract)? What is your idea behind using the daily anomalies like this, without applying any low-pass filter beforehand? Would the latter change your result – did you test this?
- L123-127: Is it correct that you calculate the future Z500 anomalies (used for detecting the WRs) by subtracting the Z500 climatology (or the seasonal cycle, as you call it) based on the historical period and not based on the future period? Is this what you mean with the last sentence in this paragraph? I think this is crucial but may not be fully clear from the text.
- L129-131: Here it would be worthwhile citing some other studies using similar regional domains for the EOF analysis.
- L131-136: I am not sure if I fully understand the procedure: Do you apply an EOF analysis both on the observed anomalies (to get the 4 observed PCs), and also separately on the modeled anomalies (to get the 4 pseudo-PCs) to calculate the “computed regimes”? In the current form, it sounds like you apply the EOF analysis only on the observed anomalies, and you then project the modeled anomalies into this (observed) phase space, as a basis for both the “computed” and “projected” regimes.

C5

- L146-148: Out of curiosity, did you calculate the “computed regimes” also for the future simulations? If yes, are they different, and do they tell us something about changes in modes of variability in the large-scale circulation?
- L160-162: Could you briefly mention in the manuscript whether a higher or lower variance ratio is generally desirable for a WR definition (independent of the comparison between observations and model)? I guess a WR definition is “better” (i.e., the WRs are more distinct) the larger the variance ratio is, because a high variance ratio implies a relatively large distance between the cluster centroids compared to the distances within a centroid, right?
- L195-196: Out of curiosity, do you know whether there are preferred circuits / transitions between the PNA sector WRs, considering the fact that they resemble (different states of) Rossby wave trains originating from the tropics?
- Figure 2: I really like the way you compare the WR representation / biases in Figure 2! I think it could be of interest for the CMIP community to additionally see in the Taylor diagram how individual model centers improved (or worsened) their WR representation between phase 5 and 6. Could you use a specific symbol (instead of a dot for every simulation) for the same model (or for related model simulations / centers)? Or does this make the figure too overloaded?
- L212-214: I kind of see your argument of an overall improvement between CMIP5 and CMIP6 visually, but can you “proof” this with a certain measure of significance (considering the relatively small number of model simulations)? Is this the degree of overlap of the shaded blue and red ellipses (if yes, please specify in the manuscript)? I would also be careful with concluding that the two NAO WRs improve more than the others – this is not that convincing considering the large inter-model spread for instance in the NAO+. Also, in principle it can be that the same model center does not improve but rather worsen from phase 5 to phase 6, which would become visible if the symbols were changed as suggested in the

C6

previous comment (for instance, the very top-left red dot in the AR diagram probably indicates such a case). Can you elaborate a bit more on this in the manuscript and, in case you do not indicate the individual models with a symbol, say whether all (or most) models improved from phase 5 to phase 6?

- L212-214: The clearly smallest inter-model spread and generally smallest bias in the NAO- WR in the Taylor diagram may indicate a higher intrinsic predictability of the NAO- WR compared to the others (also compared to NAO+). Does this make sense, and did you think about this? And, if yes, do you have an explanation for this, or do you know whether this has been shown before? On the other hand, the PNA sector WRs generally seem to be harder to capture properly, considering the large inter-model spread. Do you have an explanation for this? Could it be related to the strong dependence on the tropics, implying that models with a bad representation of the tropical-extratropical interaction perform substantially worse in terms of PNA sector WRs? I know this is beyond the main focus of this study, but I think it would be helpful to briefly discuss these aspects and speculate about possible reasons in a few sentences.
- Figure 3: I understand the idea of Figure 3 from a perspective of condensing information, but I do not see the scientific reason for plotting the frequency bias against the variance ratio because there is no direct link between the two. Hence, it could confuse the reader because a linear relationship may be expected by this way of plotting. If there is a link, please clarify in the text. Otherwise, I suggest showing two vertical box-whisker plots, one for the frequency bias and one for the variance ratio (with a horizontal black line for the corresponding ERA-Interim value). Furthermore, please indicate the units for the frequency bias.
- Figures 2 and 3: How would you relate the biases shown in your figures to the well-known blocking biases in GCMs (e.g., Davini and D'Andrea, 2020, <https://doi.org/10.1175/JCLI-D-19-0862.1>)? Do we also see this somehow in your

C7

figures?

- L222-225: The overall different variance ratios between the PNA and EAT WRs compared to observations are indeed interesting. In the case of the PNA sector WRs, however, a higher than observed variance ratio does not necessarily mean that the models perform better than for the EAT WRs, right? It just means that the models distinguish more (too?) strongly between the different WRs. So I do not really understand this statement here, because, regarding my previous comment (about the larger spread in the Taylor diagram for the PNA sector WRs), I would rather think the models overall have more problems for the PNA sector WRs. . .
- L230-233: I find it somehow surprising that the model biases in reproducing the observed regimes are smaller for projected regimes than for computed regimes. Can you explain? Is it simply because applying the k-means clustering to each model simulation (i.e., the computed regimes) yields slightly different and thus “new” regimes, hence, you kind of compare apples with oranges in the Taylor diagrams in Figure 2?
- Figures 4, 6, 7: Considering the relatively small number of models, I wonder how robust the distributions in the box-whisker plots are. Did you check whether certain distributions are skewed due to, for instance, a clustering of several model simulations from the same model center? I guess the Welch's t-test does consider that particular problem. Nevertheless, it may be helpful for the reader to replace the box-whisker plots with violin plots additionally indicating the density within the distribution.
- Figures 4, 6, 7: I would color the historic box in black (and all the future scenarios in color, as it is), just as a suggestion.
- L257-263: The apparently non-linear response of the NAO- to the CO₂ forcing is very interesting! Furthermore, the temporal development in the different simula-

C8

tions in Figure 5 shows an interesting multi-decadal variability. You mention that this will be analyzed in further studies. Can you nevertheless speculate about some potential reasons? Could it be related to some kind of tipping points in external forcings such as the Greenland ice sheet (which could affect the Greenland high) or sea surface temperature?

- L264: How does the temporal evolution for the PNA sector WRs in the CMIP simulations look like (analog to Figure 5)? Does it also exhibit any multi-decadal variability in specific WRs? I would suggest adding this figure to the supplement.
- L277-289: Thinking in terms of WR life cycles, the strong correlation between changes in WR frequency (Figures 4, 6) and WR persistence (Figure 7) implies that there does not seem to be changes in numbers of life cycles but rather changes in the duration of individual life cycles (which ultimately make the changes in WR frequency). Is that correct? If yes, can you discuss this with a few sentences in the manuscript? It could also be interesting / helpful to plot changes in WR frequency against changes in WR persistence. Depending on the robustness, this finding may to some degree also have implications for the (operational) predictability of WRs for instance on subseasonal-to-seasonal time scales.
- L310-311: What are the projections for future ENSO occurrence? Do we expect (significant) changes? Can you cite some of these studies here?
- Figure 8: Does NML stand for the hemispheric zonal mean?
- Figure 9: Just for clarification, does the shading in this figure show the mean Z500 (grid-point level) in the future simulation minus the zonal mean (at every latitude) shown in Figure 8? How does Figure 9 compare to a map that simply shows the future mean Z500 minus the historic mean Z500 (both on a grid-point level)?

C9

- L344-350: How sensitive is your analysis to the latitude / pressure boundaries used to define the four metrics?
- L362: Please add Figure S7 to the supplement, because it's missing.
- L366-376: I would add a reference to, e.g., Ambaum and Hoskins, 2002 (<http://shorturl.at/ipDO9>), who proposed a mechanism for the strong NAO-polar vortex coupling (see previous comment).
- L388-401: The changes in WRs in a future climate must be strongly linked to changes in extratropical cyclone activity and thus the storm track. Can you briefly discuss or at least speculate here whether and how some of your results (e.g., the increase in NAO+ frequency) might relate to the expected changes in extratropical cyclone frequency, location, and intensity? I assume this question is more complex than we think, but it would be nice to at least mention these questions in the conclusions and thus make a bridge toward the cyclone research community. Because in the end, changes in WRs are also a result of changes in cyclone activity. . .

4 Language comments

- Throughout the manuscript, you often write certain phenomena with capital letters, which I would not do. For instance, change "Weather Regimes" to "weather regimes", "Polar Stratospheric Temperature" to "polar stratospheric temperature", "Polar Vortex Strength" to "polar vortex strength" etc.
- You misspell "Pacific Through" (instead of "Pacific Trough") several times in the manuscript (including the Abstract)

C10

- L11: Change to “A major challenge for the climate community is to understand how a warmer climate . . .”
- L13: Change to “. . . inextricably related to regional impacts . . .”
- There are a few further grammatical inconsistencies throughout the manuscript, which should be detected when carefully revising the manuscript.

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