Answer to Report #1

1 General comment

By addressing both reviewers' comments, the present study by Detring et al. has improved considerably. In particular, the analysis of the significance of observed trends supports the results importantly. Still, there are numerous minor issues where the text could be more specific or where the grammar and wording make the text difficult to understand. I have done my best to point out the numerous instances needing correction, and I have provided suggestions to help the authors. I therefore recommend minor revisions prior to acceptance in WCD.

We thank the reviewer for carefully reading our manuscript, and for the constructive comments. In the following we will respond to the comments and point out any changes we made. The line numbers and figure references in the reviewer's comments refer to the revised submission. The reviewer's comments are in black italic; our responses are in blue. We would like to mention that we have rewritten large parts of the paper in order to emphasize the statements more clearly and to give more guidance to the reader as was suggested by other reviewers. This mainly concerns the introduction, the results and the discussion. As the changes in these three parts are significant, we refrain from highlighting line-by-line changes and rather present the new text.

2 Minor Comments

Title: Please consider to specify the region for which occurrence and transition probabilities are investigated in this study (e.g., Europe or Atlantic-European region).
Thanks a lot for this advice! We changed the title accordingly.

l. 3: "single low" sounds a bit awkward. What about "isolated low" or very specific "cut-off low"?
Thanks for the suggestions: we changed "single" to "isolated".

l. 6: Please introduce acronym "NCEP-DOE". If not overly important you could say more general "Based on reanalysis data in the...".
Done.

l. 13: "state" instead of "states".
Done.

l. 15: The statement "...and the previous time step neglecting all previous time steps" could be confusing to the reader. "...an the previous time step neglecting all further time steps before" could be an alternative. In line 72 of the manuscript this is explained much

more clearly.
Thanks, we followed your suggestion.

l. 19: Please insert "to be" between "and" and "more".
Done.

l. 25: The way High-over-Low and Omega blocks are described is valid only for the Northern Hemisphere. To clarify this you could state in this line "Northern Hemisphere blocks are characterized by a steady high pressure area...".
Yes, you are right. We adapted the sentence.

l. 35: Please include a reference to a study showing that Omega blocking caused the widespread temperature records in 2019.
Bissoli et al. (2019, in German) states: "Darüber hinaus gab es im Westen und der Mitte Deutschlands verbreitet lokale Stationsrekorde." (Translated: In addition, there were widespread local station records in western and central Germany.). Deutscher Wetterdienst (2019) writes " Between 1881 and 2018, the 40-degree threshold was reached or exceeded a total of ten times in Germany: now, in July 2019, this threshold was exceeded 25 times in just three days!" Both references are given at the end of the sentence. Moreover, we added another reference who also states the record breaking temperature values at several stations (Vautard et al., 2020) and rewrote the sentence so that it is clearer that we talk about local (station) data.

l. 38: Please consider to delete this sentence. In its current form it is more or less a repetition of the sentence in lines 22 and 25.
Yes, you are right. We deleted the first part of the sentence and changed it to: "Moreover, blocking can lead to contrary weather situations. For example, ...".

l. 43: Many conceptual explanations of atmospheric blocking exist. Therefore, I would suggest to account for this by saying "One conceptual explanation of atmospheric blocking is...".
Done

l. 51: What is "both speeds" referring to? Please clarify.
We rewrote the sentence to make it clearer: "Stationarity of the the vortex system is explained if the westerly wind speed and the translation speed of the point vortex system are identical."

l. 54: Please consider to state that the point vortex theory allows "to classify and to distinguish the two blocking states High-over-Low and Omega". Also please, remove "blocking types" at the end of the sentence in l. 54.
Thanks, we adapted this and the previous sentence and wrote: "Application of the point vortex concept allows to identify and to locate each vortex that is associated with the

block. Moreover, it allows to classify and to distinguish the two blocking states *High-over-Low* and *Omega*."

l. 57: Please delete the sentence starting with "Using this method..." since it is redundant with the sentence in l. 52-53.
Done.

l. 59: Isn't the question already answered in l. 54? There you state "Furthermore, this leads to the possibility to classify and analyse the two blocking states...". In its current form it reads a bit awkward to raise a question that seems to be rhethorical in nature as it has been addressed by Müller et al. 2015 and Hirth et al. 2018.
Thanks for your valid point. We decided to exclude the first question. But we want to point out that one interesting outcome of this study is also this method. We refined the trapezoid method published in Hirt et al. (2018) to (1) identify the blocking type at each time step and (2) to split IBL-identified blocking phases if necessary to preserve the Lagrangian aspect of the blocks. Now, this allows us to study - maybe for the first time - the transitions between blocking types. This can possibly be used to study the weather-related impacts associated with blocks in future work.

l. 92: Depending on your response to my previous comment (l. 59), please consider to exclude this question as it seems to be answered already.
Please, see our comment above.

l. 106: Please introduce acronym "NCEP-DOE" if not already done in the abstract.
Done.

l. 120: Here and elsewhere in the manuscript: Please note that the term "blockings" does not really exist in the English language (plural of gerund). Write "blocks" instead.
Thanks, we use "blocks" instead.

l. 130: Better write "obtain" instead of "get".
Done.

l. 134: "Longitude" instead of "longitudes".
Thanks!

l. 136: Better write "Note, that two blocking events can exist at the same time...". In its current form this is quite colloquial.
Done!

l. 148: Please replace "over" with "of the".
Thanks, done!

l. 153: Since you are dealing with data on a regular latitude-longitude grid: Have you used the corresponding equations in spherical coordinates? If yes, please provide the equations in spherical and not cartesian coordinates.
Thanks. Yes we use spherical coordinates, hence, we rewrote the equations with spherical coordinates as you suggested.

l. 166: What is the motivation for choosing a radius of 1500 km around $P_{max,neg}$? Is this related to the Rossby radius of deformation? As the reader might get the impression that the radius was chosen quite subjectively, please explain why a radius of 1500 km is reasonable. Could you also calculate the circulation as the sum of all grid points in an object which encloses only grid points with negative circulation?
Indeed we already did a sensitivity study to obtain the optimal radius in the previous work of Hirt et al. (2018). In Hirt et al. (2018) we started with a smaller radius of 500 km around the high centroid and increased the size gradually by 250 km steps up to 3000 km. We first saw a fast increase in the circulation magnitude of the high for increasing radii. The circulation magnitude stabilized around a radius of 1500 km, which was then chosen as threshold (see Hirt et al., 2018, for more details). On the other hand, the Rossby radius of deformation in a stably-stratified, dry high is indeed larger than that of a low pressure system because the Brunt-Väisälä-frequency is genereally higher than for a low. This further supports the choice of 1500 km.

l. 174: "Latitude" instead of "latitudes".
Done.

l. 170-184: Many parameters need to be chosen. Please include some discussion on the sensitivity of the results on the choice of the parameters (not necessarily here but somewhere in the manuscript).
Yes, you are right. One aim of section 4.1 is to study the sensitivity of the results on the chosen thresholds. In this section 4.1, we focus mainly on the distance criterion between the high centroids of successive time steps. The sensitivity on other parameters, such as the criterion to distinguish between High-over-Low and Omega block, have already been discussed in Hirt et al. (2018). We will add some lines to the end of chapter 3.4 and give a reference to Hirt et al. (2018).

l. 193: This sentence needs clarification/simplification. The usage of 4x "of" makes the sentence difficult to understand.
Thanks! We rewrote the text to make it clearer.

l. 203: Just to be sure: Do you mean "breached" or "reached"?
Thanks for your question. Since this seems to lead to confusion, we rewrote the passage:
"In order to obtain configurations associated with the same high, we split each blocking period into smaller periods, if the distance between the centroid locations of two highs in successive time steps (6 hours) is too large (*distance criterion*). We assume that these

two highs represent the same system, if the distance differs less than 10 degrees latitude ($\approx 1000\,\mathrm{km}$ in north-south direction) and less than 15 degrees longitude ($\approx 1000\,\mathrm{km}$ in west-east direction). Otherwise if the distance is larger, we assume that these two highs represent two different systems. Although this allows for slow motions of the blocks, "large jumps" in distances would rather indicate that a different high pressure system enters the configuration. If the lifetime of one or both events in the split period becomes less than 5 days, the event(s) is(are) removed from the analysis."

l. 211: Other types of blocking identification methods exist that are not solely based on indices. E.g., Schwierz et al. 2004. Hence, I suggest to explicitly reference the studies which use an index for blocking identification. Otherwise, the current statement is too general.
Thanks, you are right. We specified and rewrote the sentence: "Note, that this rather Lagrangian view on blocking differs from the rather Eulerian perspective we would get, if we solely use the instantaneous blocked longitudes to identify blocking."

l. 256: Please replace exclamation mark with full stop.
Done.

l. 312: In its current form it reads that in theory you can describe time-varying transition probabilities. Please rewrite the sentence so that it becomes clear that this is what you are actually doing.
Thanks, we specified the statement: "Therefore, occurrences of exceptional droughts, that were experienced 2018 and 2019 in central Europe (Hari et al., 2020) are likely to occur more frequently at the end of our study period than at the beginning."

l. 337: Insert blank between full stop and "For".
Done.

l. 344: Here and elsewhere in the manuscript: Please consider to write "propagation" instead of "travel".
Thanks, we changed the "travel speed(s)" to "propagation speed(s)".

l. 353: Replace "sensible" with "sensitive".
Done.

l. 360: Please merge this paragraph with the previous sentence since these seem to be related.
Thanks! Done.

l. 388: "This" instead of "his".
Done.

l. 404: What is the reason that the trend is highly significant in JJA when considering an entire season but not significant when considering individual months?
Trends estimated for individual months might be opposing and thus average out when the results are averaged over the season. This is now also explained in the manuscript.

l. 426: "about" instead of "About".
Done.

l. 442/443: What is the reasoning for distinguishing between decrease/slight decrease and increase/slight increase. Is there any objective criterion to make this distinction? If not I would suggest to only use increase and decrease to describe the trends. The quantitative information of the significance is more informative than the qualitative distinction between "slight" and "not slight".
You are right, there was no need to distinguish between "slight" and "not slight" decrease/increase. Due to our changes in the significance test (from Wald test do likelihood ratio test) now there is a need for this differentiation. Significance can no longer be reported for *High-over-Low* or *Omega* blocking individually, but only for the influence of a particular categorical term in the predictor (e.g. *Seas* or *month*). Therefore, we use terms like "slight" or "strong" to highlight which of the two blocking states shows larger or smaller changes.

l. 444: Similar to the previous comment, the distinction between "strongly" and "pronounced" seems to be a bit arbitrary.
Please, see our comment above.

l. 453: Are you sure that the trend in February is significant for HoL (solid) but not for Omega (dashed). This is striking given the marked increase of Omega blocks in February. Please double check.
We have checked the significance levels again and have rewritten the analysis of Fig.11. Due to the Hauk-Donner effect we now use the likelihood-ratio test instead of the Wald-test to determine the significance levels. A more detailed explanation can be found in our response to Report 2.

l. 456: "as" instead of "than".
Done.

l. 461: The final sentence is a bit colloquial. I guess what you want to say is that "All in all, the net trend is zero".
Thanks, we will use your version.

l. 463: Please specify at the beginning of this subsection that the analysis is only performed/shown for the Euro-Atlantic region.
Thank you for this helpful advice. We have adapted the text accordingly.

l. 467: Please remove "of" between "and" and "three".
Done.

l. 472 and after: Please correct quotation marks.
Done.

l. 508: "become" instead of "becomes".
Done.

l. 520: Depending on whether the first question is kept in the introduction or not, this line needs to be changed.
As explained in the comment on l.59, we have reformulated the question and adjusted the text here accordingly.

l. 537: "On the one hand" needs to be followed by "on the other hand" somewhere in the sentence. So, I guess you can delete "on the one hand" from the sentence.
Yes, you are right. We deleted "on the one hand".

l. 547: This sentence needs reordering. My suggestion: As shown in Fig. 7b the probabilities of blocks increase slightly, but significantly in summer.
Thanks, reordering the sentence makes it more clear.

l. 548: This statement is not supported by results of this study. Of course, significant trends have been identified based on historical data, but no climate projections have been investigated. Therefore, the statement is speculative and I would suggest to delete it from the manuscript. As an alternative you could write "If this significant increase continues in a warmer climate, the occurrences of exceptional droughts could occur more often in the future".
We have adjusted the sentence and it is now worded as follows: "Therefore, occurrences of exceptional droughts, that were experienced 2018 and 2019 in central Europe (Hari et al., 2020) are likely to occur more frequently at the end of our study period than at the beginning."

l. 559: Replace "in" with "on".
Done.

l. 576: To ease the comparison with Drouard and Woolings (2018): Why are you not showing Fig. 15 for June to August only? Since you are analysing some of the results for each season separately, it should be fairly straightforward to show Fig. 15 for June to August. So, please revise the figure so that it is better comparable to the results of Drouard and Woolings (2018).
Thanks for your advice. We adopted the plot and even show both figures for the whole

data set as well as for the June-August subset.

l. 606: "relatively" instead of "relative".
Thanks, done.

l. 630: Consider to replace "decision" with "identification". Also, to my understanding the method is a continuation/refinement of the method introduced by Hirth et al. (2018). Therefore, I'd suggest to cite their work also in the concluding section.
Thanks, we changed "decision" to "identification" and added a sentence to cite the work of Hirt et al. (2018)

l. 648: "find" instead of "found" (present tense).
Done.

l. 651: Please specify that this is the subjective impression of the authors (or reference a paper which confirms the impression).
Thanks. We added some words to clarify that this is our subjective impression.

l. 656: In what sense could the analysis of temperature and wind shear help to better understand blocking? What about other processes such as diabatic processes, Rossby wave breaking etc?
Thanks for your comment. We specified our statement as follows: "In future studies, for a deeper understanding of atmospheric blocking, the relation of on- and offset of blocking to different parameters such as the North-Atlantic-Oscillation (NAO) index or the difference in midlatitude wind speeds from one time step to the next could be investigated. NAO, on the one hand, is related to the large-scale midlatitude pressure gradient, the speed and location of the jet that could influence the blocking process in Europe. Croci-Maspoli et al. (2007) confirm a strong correlation between blocking occurrence and the negative phase of the NAO. On the other hand, a sudden change of the large-scale midlatitude wind speeds might be used as a proxy to estimate a meandering of the jet stream, which might also affect the blocking behavior in Europe; see e.g. Luo et al. (2019) or Riboldi et al. (2020) who show that periods of reduced Rossby wave phase speed are systematically related to atmospheric blocking. Additionally, diabatic effects such as latent heat release play an important role in blocking dynamics (Pfahl et al., 2015; Steinfeld and Pfahl, 2019). Thus, studying this relation with respect to different blocking types would be very insightful, too. "

l. 658: What is the purpose of "however" at the beginning of the sentence?
Oh, you are right. It should emphasize yet another possible application. We changed "however" to "furthermore".

l. 663: Replace "development" with "trends" or "long-term changes".
Thanks, we replaced it with "trends".

Figure 1: Please specify in the figure caption that relative vorticity is shown.
Done.

Figure 3: Please specify in the figure caption that relative vorticity is shown. Also, in line 2 of the caption it should be "vorticity" instead of "vorticty". Here and elsewhere: Please use date formatting as described in the Submission guidelines of WCD: https://www.weather-climate-dynamics.net/submission.html. Further, the units for geopotential height are different between Figs. 1 and 3 (gpm vs dm). I would suggest to use units consistently throughout the study.
Thanks, we adapted the date formatting and the geopotential height units to "dm" in both figures. We also changed "vorticity" to "relative vorticity" in both figure captions.

Figure 10: Please consider to rewrite the Fig. caption in order to avoid 4x "for".
We rewrote the caption of Fig. 10: "Temporal development of Euro-Atlantic ($40° W -$ $30° E$) blocking probabilities for individual seasons: (a) *High-over-Low* and (b) *Omega* blocking."

Reference: C. Schwierz, M. Croci-Maspoli, and H. C. Davies, "Perspicacious indicators of atmospheric blocking," Geophys. Res. Lett., vol. 31, no. 6, p. L06125, Mar. 2004, doi: 10.1029/2003gl019341.
Thanks for the reference!

References

Croci-Maspoli, M., Schwierz, C., and Davies, H. C.: Atmospheric blocking: Space-time links to the NAO and PNA, Climate Dynamics, 29, 713–725, 2007.

Hari, V., Rakovec, O., Markonis, Y., Hanel, M., and Kumar, R.: Increased future occurrences of the exceptional 2018–2019 Central European drought under global warming, Sci. Rep-UK, 10, 1–10, 2020.

Hirt, M., Schielicke, L., Müller, A., and Névir, P.: Statistics and dynamics of blockings with a point vortex model, Tellus A: Dynamic Meteorology and Oceanography, 70, 1–20, 2018.

Luo, D., Zhang, W., Zhong, L., and Dai, A.: A nonlinear theory of atmospheric blocking: A potential vorticity gradient view, Journal of Atmospheric Sciences, 76, 2399–2427, 2019.

Pfahl, S., Schwierz, C., Croci-Maspoli, M., Grams, C. M., and Wernli, H.: Importance of latent heat release in ascending air streams for atmospheric blocking, Nature Geoscience, 8, 610–614, 2015.

Riboldi, J., Lott, F., d'Andrea, F., and Rivière, G.: On the Linkage Between Rossby Wave Phase Speed, Atmospheric Blocking, and Arctic Amplification, Geophysical Research Letters, 47, e2020GL087 796, 2020.

Steinfeld, D. and Pfahl, S.: The role of latent heating in atmospheric blocking dynamics: a global climatology, Climate Dynamics, 53, 6159–6180, 2019.

Vautard, R., van Aalst, M., Boucher, O., Drouin, A., Haustein, K., Kreienkamp, F., Van Oldenborgh, G. J., Otto, F. E., Ribes, A., Robin, Y., et al.: Human contribution to the record-breaking June and July 2019 heatwaves in Western Europe, Environmental Research Letters, 15, 094 077, 2020.

Answer to Report #2

1 General Comment

In response to the original reviews, the authors have done considerable revisions to the manuscript; in fact, large sections have been completely rewritten. On a superficial level, the authors have addressed a good fraction of my earlier issues. Nevertheless, I feel that the paper still lacks something to be desired in order to make it truly convincing. I encourage the authors to do further revisions and I try to give suggestions below.

We thank the reviewer for carefully reading our manuscript, and for the constructive comments. In the following we will respond to the comments and point out any changes we made. We would like to mention that we have rewritten large parts of the paper in order to emphasize the statements more clearly and to give more guidance to the reader as you suggested. This mainly concerns the introduction, the results and the discussion. As the changes in these three parts are significant, we refrain from highlighting line-by-line changes and rather present the new text. The line numbers and figure references in the reviewer's comments refer to the revised submission. The reviewer's comments are in black italic; our responses are in blue.

2 Provide better guidance to the reader

I cannot really judge the technique used to determine statistical significance. Having said this and assuming this technique to be sound, the paper is probably OK on a micro-level, i.e., by evaluating each sentence individually. However, I was sometimes missing guidance from the authors as to what the results really mean, providing a broader sense of interpretation and putting everything into the larger context. For instance, in the summary and concluding section I would expect less of a detailed repetition of what has been said before, but rather a discussion on a higher level of abstraction. This issue is particularly relevant regarding the monthly trends (see next issue).

We have rewritten parts of the paper in order to emphasize the statements more clearly and to give more guidance to the reader as was suggested. This mainly concerns the introduction, the results and the discussion.

2.1 What do the detected trends mean?

The authors made an effort to say more about the statistical significance of their results and they are now more careful when quoting results that are lacking statistical significance. I am not an expert in statistics, but I simply trust that their method to quantify

significance is sound. Yet, I still find some of the results puzzling, and it would be desirable if the authors could say a little more about how they interpret their own results.

Thank you for your very constructive and adequate scepticism here! Although we use advanced but standard statistical models here, it is very important to double check and challenge statistical significance. As reported in our manuscript, we use the framework of vector generalized linear models comprehensively described in the text book of Yee [2015]. With the ambitions to address significance of trends in occurrence probability over time (years) *individually* for each state of the three-state model, we aimed to test the coefficients in the logistic linear model for compatibility with zero using the z-statistic (Wald test). This is a standard procedure. However, for logistic regression, it is also known that this strategy can fail in some particular cases. This problem is illustrated in Yee [2015, Chap. 2.3.6.2]. This effect is named after Hauk and Donner and has been discovered already in 1977. Although this is a particular case, it is prominent enough to be mentioned also in Wikipedia's entry on the Wald test. This effect leads to the discrepancies you observed in Fig. 11a for Omega blocks in February: the increase in occurrence probability is strongest for Omega blocks in February in this figure but this is the only one which is *not* labelled as significant according to the test of the associated coefficient for compatibility with zero (z-statistic). This is a consequence of the Hauk-Donner effect and we did not report the results of the z-statistic for a significance test as this is not reliable here. The remedy here is to use a likelihood-ratio test, see also Yee [2015]. However, the setting of a likelihood-ratio test is slightly different than testing coefficients for compatibility with zero. The likelihood-ratio test does not test individual model coefficients for compatibility with zero but test for a significant improvement of the model due to inclusion of *years* as an additional covariate in the model. The result of this test is then "including years as covariate significantly improves the model describing the occurrence probabilities for the three states". This test does not allow to break this result down into a significant increase in the individual occurrence probabilities for HoL or Omega but is now robust against the Hauk-Donner effect. We thus change the strategy for significance tests associated with Figs. 10, 11, 13 and 14 to likelihood-ratio and explain that in the manuscript as follows: To avoid problems with significance testing due to the Hauk-Donner effect [Yee, 2015, Chap. 2.3.6.2], we avoid the Wald test setting in favor of a likelihood-ratio test. This implies that we test for an improvement of the model due to inclusion of the covariate Year instead of testing model coefficients for compatibility with zero. This implies that we cannot infer significance for trends in occurrence probability individually for *High-over-Low* or *Omega*.

As it turns out, there are no significant trends on a year-by-year basis; instead, some trends emerge only if one does the analysis on a seasonal or even month-by-month basis. At first sight this may appear surprising, because the monthly analysis is based on a much smaller data basis suggesting a lower degree of statistical significance than the year-to-year analysis. Especially those instances when a monthly trend reverses as one proceeds from one month to the next appear somewhat suspicious to me. Could you

possibly comment in order to make me feel less uneasy? Is there anything that I have overlooked?

The variance of statistical estimators (and thus the uncertainty) is indeed generally dependent on the number of data points. For a test of a signal against a null hypothesis, not only the variance of the estimate but also the value of the estimate itself is important. While signals on a monthly time scale can be prominent (although with large uncertainty) they might average out on seasonal time scales (although uncertainty is smaller), leading to a non-significant result for the larger time scale. Good examples for this are the winter month shown in Fig. 8 showing a prominent (compared to the uncertainty) decrease in December and a prominent increase in February; increase in January is moderate. These effects average out on a seasonal scale and are not significant there (Fig. 7). Another (unconventional) choice of seasonal grouping, e.g. January, February and March would have let to a significant increase in that season. We add on this in the manuscript.

My broader question is whether those monthly trends (which, allegedly, are statistically significant) have to be considered as a signal that goes beyond the noise of natural variability. In other words: if we had 40 realizations of the three decades 1990 to 2019 that are subject to the same external "forcing" and that only differ in their natural variability, would you expect those monthly trends to be more or less identical in all 40 realizations (see Deser et al., 2020)? If this were the case, these monthly trends could be attributed to the "forcing" from, e.g., increased greenhouse gases — otherwise they would be just noise. If the latter were true, I wonder what we have learned. I think the reader would appreciate more guidance from the authors.

The hypothesis tests employed here (Wald-type or likelihood-ratio) are to be interpreted as follows: given 1000 realizations from a system with constant occurrence probabilities, i.e. **without** the trends shown here (null hypothesis), we would observe significance, i.e. classify 1 (if given ***), 10 (if given **) or 50 (if given *) of them falsely as significantly different from our null hypothesis and thus postulate falsely a trend. However, the limit of a statistical modelling approach is that we cannot easily specify specific patterns of the global climate's natural variability as they arises from a complex and slowly (on the scale of decades) reacting climate system, e.g. decades of increasing probability followed by decades of decreasing probability. Our null hypothesis are constant occurrence probabilities with occurrence itself being subject to random noise. The trends we detect are significantly different from these. Based on our results, one could now design a fingerprinting study with specific patterns of natural variability as they arise from pre-industrial control simulations and potentially also alternative hypothesis derived form climate projections. Such a study is beyond the scope of this work.

2.2 Improvements to quality of the English

I think that the level of the English could and should be improved in numerous places. I have returned an annotated version of the revised manuscript where I give a few examples. I trust that the senior authors can help towards this goal. Alternatively, it might be beneficial to consult, e.g., the excellent book of David M. Schultz ("Eloquent Science", published by the American Meteorological Society).

Thanks for you advice. We have checked the language and the senior authors have read and corrected the entire document. The book is indeed excellent!

Reference:
C. Deser, F. Lehner, K. B. Rodgers, and collaborators, 2020: Insights from earth system model initial-condition large ensembles and future prospects. Nature Climate Change, 10, 277–286
Thanks for the reference!

3 Minor comments

Thanks, we added the comments from the pdf-supplement below. We have taken all linguistic comments into account. We answer comments on content in more detail here.

l.75 ff. It is still not clear to me why the 6h time resolution should make such a big difference in the general blocking behavior.
Thanks for this question: You are right that the general blocking behavior, i.e. the location and life time, is not very different from another temporal resolution. However, we want to study the transition between High-over-Low and Omega blocking patterns that occur on a much smaller time scale and can be captured by a time interval of 6 hours. More precisely, we first identify blocking periods with a minimum life time of 5 days and then inspect the blocking state of each time step and search for transitions between the two blocking states within this blocking period. We rewrote the sentence to emphasize the focus on transitions: "However to the best our knowledge, the transition between different blocking types, i.e.Omega and High-over-Low blocks, and the unblocked state has not been studied so far on a sub-daily, 6-hourly time scale".

l. 85 what do you mean by "model" here? An observed trend should not (or at least hardly) depend on a numerical model.

Barnes et al. [2014] states: "While differences across indices are to be expected, the differences among the reanalyses for a given index may be surprising. For example, the D2D MERRA frequencies (Figures 3c and 3d) exhibit very different values than the other three reanalyses. Similarly, the M2D ERA-Interim frequencies also differ from the

4

other three reanalyses over Asia (Figure 3e). This suggests that studies using different reanalysis products could disagree on "observed" trends in blocking, even if the same blocking identification method is employed." Note, that D2D (Dunn-Sigouin et al., 2013) and M2D (Masato et al., 2013) are different blocking indices Barnes et al. [2014] use in their analysis. Moreover, Berckmans et al. [2013] states: "An underestimate of atmospheric blocking occurrence is a well-known limitation of many climate models." So yes, the identification of blocking can depend on the data used and hence on the model that produced the data. This implies that trends might depend on the model, too, or that there is at least some amount of uncertainty. We replaced "model" with "model data" in the sentence.

l.109 suggest to write: .... that requires two-dimensional non-divergent flow.
Thanks, done.

l.124 "on the southern side..." is not clear/ incomplete: within a certain range of latitudes?
Thanks, we added more meat to the description of the method: "A blocking is identified if the geopotential height gradients on the northern (GHGN) and on the southern (GHGS) side of the CRBL satisfy the following criteria [following Richling et al., 2015]:

$$GHGS = \frac{Z(\phi_M) - Z(\phi_S)}{\phi_M - \phi_S} > \quad 0\ \frac{gpm}{^\circ N} \qquad : \text{corresponding to an easterly directed flow} \tag{1}$$

$$GHGN = \frac{Z(\phi_N) - Z(\phi_M)}{\phi_N - \phi_M} < -10\ \frac{gpm}{^\circ N} \qquad : \text{similar to a westerly flow} > 8 \text{ m/s} \tag{2}$$

where Z is the geopotential height at 500 hPa and the latitudes are given as $\phi_S = (\phi_C - 0.5\,\delta_\phi) + \Delta$, $\phi_M = (\phi_C + 0.5\,\delta_\phi) + \Delta$ and $\phi_N = (\phi_C + 1.5\,\delta_\phi) + \Delta$. Here, δ_ϕ is set to 15° latitude. In the calculation performed for this study, the spatio-temporally varying CRBL ϕ_C is determined based on the 30-year climatology (1990-2019) of the 500 hPa geopotential height field. In order to capture blocking that are not directly located at the CRBL, a possible shift Δ to the north and south is set to 10° latitude. For each time step, we obtain a (1d) series of longitudes (either 1: blocked or 0: unblocked) that is saved for further analysis. For more details on the method and the specific configurations used in the analysis see the Supplementary Material and Richling et al. [2015]."

l. 164 vertical component of (relative?!) vorticity
Yes, we mean the vertical component of relative vorticity-

l.254 "commonly used combination" is unclear.
The combination of directs and combined effects is commonly used. We remove "commonly used" as this seems to be confusing.

l.267 is "asymptotic normality" a good assumption in the present case?
Asymptotic normality refers to the limiting distribution of the estimator and makes the strategy transparent we base our uncertainty estimates on. We added "of the estimator" here to make that clear. This is fairly standard and frequently not even mentioned. An alternative choice for uncertainty estimates would be, e.g., a bootstrap approach. We think that it is important to reveal the assumptions here. This assumption does *not* refer to a distribution we assume for the data used for modelling! We model occurrence probabilities with a binomial (two-state) or multinomial (three-state) distribution.

l. 402 "can pinpoint the drivers", what do you mean by that? in what sense is a specific month a "driver" of blocking?
Sorry for expressing this vaguely. We changed this now to "we have more temporal detail for interpreting the observed trends".

l. 478 how can I see this from Fig. 12?
The arrows indicate the direction of the transition. The probability of transition from "no blocking" to Omega is higher than for transition to a High-over-Low blocking. We added a reference to Fig. 12 and the arrows.

l. 549 this conclusion is only valid if one is allowed to assume that the observed trend continues.
Indeed, this was not precise. We changed that to: "are likely to occur more frequently at the end of our study period than at the beginning.

l. 562 and, therefore, lower statistical significance!
Thanks for pointing us towards this potential misunderstanding. We changed the last two sentences of the paragraph to : Adding more detail to the analysis by considering blocking types and individual months leads on the one hand to larger uncertainties as more parameters are needed to be estimated but on the other hand to stronger signals making statistical significant results possible.

Figure 1: I know what you want to say but still the sentence is not clear.
We have rewritten the sentence to: "The point vortex systems become stationary if the typical westerly flow of the mid-latitudes and the propagation speed of the dipole/tripole point vortex systems are of equal magnitude."

Caption Fig.3 this construction with multple options in parantheses is extremely hard to read, please replace by a more straightforward way to express your thought.
We changed the description to: circles are the circulation centroids of the identified high (red) and low(s) (blue) for the *Omega* pattern. In the same way, crosses are used for *High-over-Low*) pattern.

References

Elizabeth A. Barnes, Etienne Dunn-Sigouin, Giacomo Masato, and Tim Woollings. Exploring recent trends in Northern Hemisphere blocking. *Geophysical Research Letters*, 41(2):638–644, 2014. ISSN 19448007.

Julie Berckmans, Tim Woollings, Marie-Estelle Demory, Pier-Luigi Vidale, and Malcolm Roberts. Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14(1):34–40, 2013.

Andy Richling, Chistopher Kadow, Sebastian Illing, and Oliver Kunst. Freie Universität Berlin evaluation system (Freva) - blocking. `https://freva.met.fu-berlin.de/about/blocking/`, 2015. (Documentation of the blocking plugin), Accessed: 2020-09-10.

Thomas W. Yee. *Vector generalized linear and additive models: With an implementation in R.* Springer, New York, USA, 2015.