

## Authors' replies to Referee #2

We really thank Referee #2 for their helpful suggestions and comments about our manuscript. Below, we provide our replies; the line numbers and sections refer to the old version of the manuscript.

### Major comments

#### Major suggestion #1

*You need to provide a strong rationale for why you use the weather typing approach for identifying blocks as opposed to a feature-based method that allows simultaneous tracking in space and time. What are its advantages? I don't think that simply stating that others have used it before is sufficient. There are places in the manuscript where you say things that really make me wonder, why did they choose such an approach, such as: Line 261-262: "This is due to the fact that the events have been identified via a partitioning algorithm (k-means) and not via blocking indexes designed for geopotential fields that are typical during atmospheric blocking." You don't need to remove these sentences, but earlier, in the methods section, you need to make a convincing argument as to why this method useful. Given that there are multiple blocking indices that are designed for geopotential field, and that you felt compelled to compare your method against one of these, I want to know: why use the weather type approach?*

*Related to this, at line 340-341, you write: "Our results are in agreement with previous findings where blocking events are defined with blocking indexes. This confirms that the application of the WTD is also a good strategy to analyse blocking event characteristics." For me, this is a bit confusing. Why do we need another strategy that gives us the same information that we already have?*

We thank the Referee to point out these observations. We explained the reason for choosing the WTD in the Introduction (please, see reply 1a to Referee #1) and in subsection 3.1 at line 94: "Overall, while identifying blocking via blocking indexes implies making several choices, identifying blocking via the WTD can be considered as a standard procedure. This motivated us to apply the WTD and to explore this methodology for identifying blocking events and then studying their main characteristics (frequency, duration, size)."

Moreover, it is maybe worth reminding that the WTD application allows one to analyse more (in this case four) weather regimes at the same time, otherwise, one index per weather regime should be used. This could lead to some inconsistency, for example, the attribution of multiple weather types to the same day.

Regarding the second part of the comment of the Referee: since there is not one only way nor one best way to identify blocking (as explained in the Introduction) we carried out this study to explore if the WTD can be used to identify blocking events. Like the blocking indexes, this strategy has pros and cons and will not substitute the indexes, rather, it is another possibility which can be favored or not according to the purpose of the study.

Moreover, we added Figure 1 (below) in the Supplement to show the agreement between the results obtained with the center method and the DG method: 70%-85% of the blocking durations identified via the WTD overlaps (or coincides) with the durations identified via the DG method, showing a quite high correlation ( $R > 0.74$ ). The underestimation of the DG-blocking duration is due to the fact they are identified within the WTD-blocking events (as

written at line 145).

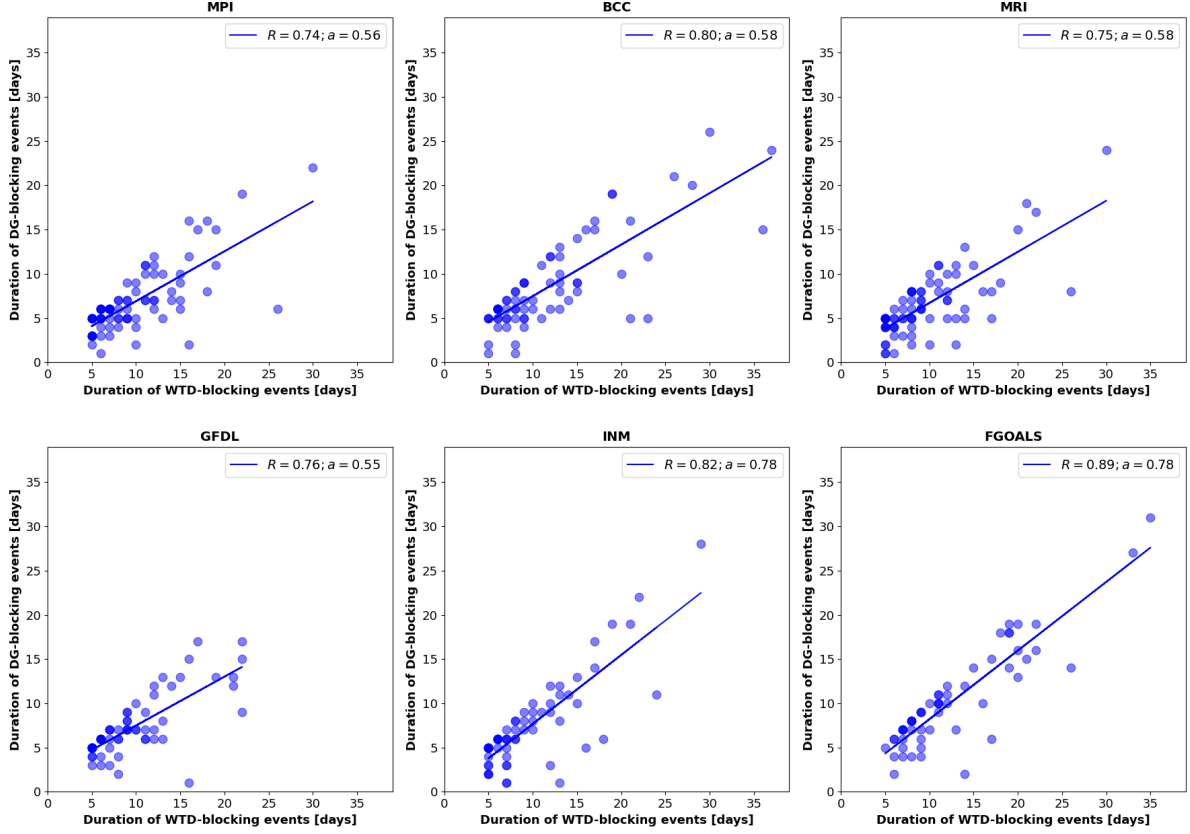


Figure 1: Correlation between duration of WTD-blocking events and DG-blocking events during the HIST period. 70%-85% of the total blocking events identified via the WTD are shown; those events which are not identified with the DG method are not displayed.

## Major suggestion #2

*Your choice to use shorthand names for the anomalies based on the climatology used is a bit confusing. For instance, when I first examined Figure 3, I thought there was a typo. Is there any chance that you might consider creating shorthand names that refer to both the removed and the dataset used, then, for instance: Z500\_HIST composites during the winter HIST period (1980-2009) could be named: Z500\_HIST\_HIST. Whereas, Z500\_HIST composites during the winter SSP2 period (2070-2099) could have the shorthand name: Z500\_HIST\_SSP2. This change would also help in clarifying what the x- and y-axes are referring to in Figure 4, are all of these showing results using the historical climatology, or is the x-axis the future climate with the historical climate removed? I eventually sorted out the answer to these questions, but with your existing format, it was more difficult than it needs to be.*

We thank the Referee for suggesting this new notation. We agree that referring to both the removed and the used datasets is more intuitive, so we followed this notation in the

new version of the manuscript. Actually, since we emphasized the results obtained with Z500\_SSP2 and Z500\_SSP5, we used the notation  $\Delta Z500_{\text{SSP2}}$  and  $\Delta Z500_{\text{SSP5}}$  for the anomalies in the future period (2070-2099) obtained by subtracting the climatology computed in the future; we thought that it is not necessary (and not nice) to write SSP2-SSP2 in subscript (i.e.  $\Delta Z500_{\text{SSP2-SSP2}}$ ). Consistently, we kept the notation  $\Delta Z500_{\text{HIST}}$  for the past. Instead, for the results obtained with Z500\_HIST, we used the notation  $\Delta Z500_{\text{SSP2-HIST}}$  and  $\Delta Z500_{\text{SSP5-HIST}}$ .

In the plots of Figure 4, the x-axis is for the results obtained in the HIST period (i.e. using  $\Delta Z500_{\text{HIST}}$ ), while the y-axis is for the results obtained in the future using  $\Delta Z500_{\text{SSP}}$  (i.e.  $\Delta Z500_{\text{SSP2-HIST}}$  and  $\Delta Z500_{\text{SSP5-HIST}}$ ). This type of plot wants to show the comparison between future and past in an easy way, as an alternative to reading the table. We improved the caption of Fig. 4 using the new notation.

## **Minor comments**

*Line 99: possible typo, I think the word “so” should be “to”.*

Done.

*Line 130 and multiple places elsewhere: when you use the word extension, do you mean the same thing as extent? For me, extension suggests an action, such as expansion or shifting in the location of the block, whereas, extent suggests the instantaneous location of the block. I am curious to see if the other reviewers or the editor agree with me on this. If they do not, you can leave it as is.*

We replaced the word “extension” with the term “size” in order to avoid any misunderstanding, as also suggested by Referee #1.

*Lines 131-135: This explanation did not make sense to me, mainly because you sometime mention composites and other times you do not, i.e., on line 132, you write: “The center method starts from the detection of the center of each blocking event.” Okay, for me, because you say “for each blocking event”, I think to myself: this does not involve composites. But then the next sentence talks about defining the centers based on the anomaly for the composite. Please try to re-write this description to remove any confusion.*

We reviewed the usage of the expressions “composites” and “blocking events” and we modified the text according to the Referee’s comment.

*Line 186: I think Figure S3 should be included in manuscript as a result figure not a supplemental figure. In the current form of the manuscript, a comparison of Figures 2 and 3 gives a strong suggestion that blocking changes significantly with climate change. But that is not the result that you are presenting. Instead, the differences in Figures 2 and 3 is mainly a due to a difference in the mean state for the 20 and 21 st centuries. Right? If you see it differently, then please explain.a*

Following also the suggestions of Referee #1, we showed the DeltaZ500\_SSP results in the revised version of the manuscript and we moved most of the DeltaZ500\_HIST results to the Supplement (or we removed them). Thus, also Figures 2, 3, and S3 changed as asked by the

Referee.

*Lines 252-257: This discussion of block intensity when considering delta-Z500HIST vs delta-Z500SSP is a bit awkward for me. Given that in the present climate we define blocks as anomalies with respect to the climatology, your discussion of future block “intensity” relative to the historical climatology, seems a bit arbitrary. For context, if we were discussing heat waves, we know that humans feel uncomfortable when the temperature is above a certain threshold, e.g. 32°C. So, when someone studies future warm events, there is a reason to look at anomalies with respect to our current climatology. But for block intensity, does that same thing hold? i.e., when surface pressure is above some threshold are there specific impacts on humans? If so, please explain. If not, then I wonder if you might consider streamlining this section and only discussing the intensity changes when comparing delta-Z500HIST for the historical runs with delta-Z500SSP for the future runs. Especially since your discussion of Figure 3 already makes the point about the impact of the change in the climatology on the Z500 anomalies.*

Please, see our previous reply.

*Figure 7, one result you have found, that I don’t think is mentioned (pardon me if I missed it), is that the model-to-model differences in intensity are larger than the intensity differences for SSP2 vs SSP5. This suggests to me that even though the models are getting some aspects of the physics of blocking correct, there is room for improvement.*

The fact that the model-to-model differences are larger than SSP2 vs SSP5 was mentioned at line 209, but we added this observation also at line 245. This could suggest that the signal of climate change on this particular quantity (blocking frequency/duration) is not significant and that taking a stronger warming scenario does not change much the results; we commented this in the Conclusions.

*Line 331, you write: “Climate change will significantly increase the extension of blocking events in the future especially in the worst-case scenario.” I don’t think your results agree with this statement. As you discuss, the difference is related to a change in the climatology. For me, this implies that the blocks themselves, in terms of their size and impact on the circulation, will be similar to what we observe in the current climate. Perhaps a method for testing this would be to address the question: are the pressure gradients associated with the blocks strengthening?*

As we showed the DeltaZ500\_SSP results in the revised version of the manuscript, as written before, this sentence is not present in the Conclusions.

*Philosophical Suggestion: Whenever a “trends” manuscript is written, I think the authors should take time to ponder the question: Are we (or somebody else perhaps) going to have to write this same paper in N years when the next generation of climate models are released? If the answer is yes, then why do we need this paper now? If the answer is no, or not exactly, then why not – what have we established here that is robust to potential changes in models? And maybe some elements will be updated, but are there at least some elements that are novel in this study that are not going to be superseded by results from an improved set of climate models, and/or are the hypotheses or theories developed in this work that provide a simpler guide to how we can interpret the trends? Obviously, I have not reviewed the majority of trends papers written, so most of the time this question is not addressed. So, I won’t hold you to a higher standard than what exists in the literature. However, I encourage you to consider*

*engaging in this activity and perhaps including a paragraph at the end of the paper addressing this issue. No problem if not.*

Investigating future blocking with the last generation of models is motivated by the fact that controversial results have been found so far about future blocking and CMIP6 models were shown to better represent the blocking weather type. It would be natural to propose again such an investigation if the previous conditions persist. At the same time, we would like to stress that in this study we introduce a different methodology to study spatio-temporal characteristics, showing that this can be an alternative to the blocking indexes; this methodology can be reused in future studies with newer generations of models.