

Review of WCD manuscript egusphere-2025-4969 “Atmospheric Blocking Representation in Storm-Resolving Climate Models under Historical and Future Forcing” by Dolores-Tesillos et al. (2026)

Assessment

I appreciate the authors' efforts in addressing my previous comments in the first round of the review process and revising the manuscript accordingly. The paper is now in much improved shape and clearer in several important aspects including a more concise title. However, some shortcomings remain. Some of them have only recently become apparent, while others persist from the earlier version.

The study demands further improvement in the discussion of the results as well as a more convincing justification for including the blocking trend estimates. The latter represents the only remaining major concern. Given the progress made and the issues that still require improvements, I recommend a **minor revision**. If the concerns outlined below are adequately resolved, the manuscript would be in a good position for acceptance in *Weather and Climate Dynamics*.

General comments

1.) The authors consider the evaluation of future blocking trends valuable, particularly when the projected signals are consistent with those from CMIP-based studies. This raises concerns about the reliability of the results in cases where km-scale models and CMIP models disagree. Such a selective agreement does not provide a sufficiently robust basis for confidence in the projected trends. Also, if the analysis emphasizes only trends that align with CMIP-based studies, it is unclear what the novel contribution is beyond showing that km-scale models seem to behave similarly to CMIP models.

In light of this, I encourage the authors to reconsider the relevance and robustness of these results. A meaningful evaluation of blocking trends requires a more careful and comprehensive treatment, accounting for both mean state biases and trend signals across the key contributing factors (background flow, SSTs, and storm tracks). If the section is to be retained, the magnitude of the trend signal needs to be critically assessed in a quantitative manner to ensure it can be distinguished from the influence of model biases. In particular, an estimation of SST biases would be highly beneficial, given that the AMIP simulation perform considerably better and the manuscript emphasizes the strong sensitivity of blocking to underlying SSTs in both seasons.

Additional minor comments are provided in the line-by-line comments below.

Line-by-line comments:

Line 40-43: 1) Horizontal resolution and 5) Parameterisations are technical model-related aspects, whereas 2), 3) and 4) are physical processes of the climate system. The authors should add this information and emphasize the interdependence.

Line 63-64: I would suggest to remove the second part of the sentence, as the AMV is not further addressed in the study.

Line 76: Add space between „realistically“ and reference.

Line 80-81: „...where biases of SST, storm tracks and moist processes...“

Line 82: What is meant by „expose new limitations“? Either expand or remove this phrase.

Line 100 & 105: The reference to nextGEMS is given twice here. I'd suggest to remove the second part of the sentence in line 105.

Line 118-124: This is the main motivation for assessing km-scale models. I'd recommend moving it to the introduction.

Line 212: Why do you use a southern border of the ATL domain at 50°N and for the PAC domain at 40°N? Please clarify.

Line 221: Remove „in multidecadal simulations“ in the 4.1 title. Same for 4.2 title in line 348.

Line 235: Consider adding the bold-written part: „suggesting that biases **of blocking frequency** in the coupled version“

Line 253, 258, 331, 401-402, 433, 452: These lines include overly too descriptive parts of what is shown in the respective figures. As I already mentioned in the first review, the authors should consider moving the figure reference to the end of the sentence and removing these overly too descriptive information, which distracts the reader from being fluently guided through the results.

Line 304-307: Same here: This paragraph is too long in terms of accompanying information, describing how the short following paragraph is structured. I'd suggest to remove this part.

Line 310: This conclusion about the baroclinicity can be *suggested*, but not drawn from any figure shown in the manuscript. I'd suggest rephrasing this sentence accordingly.

Line 436: The result about ICON hist and CMIP6 ensemble should be compared against ERA5 and not against IFS hist.

Line 440: Again insert the figure reference at the end of the sentence. The authors may consider adding labels to subplots (e.g. a), b), c) ...)

Line 478: Remove „nextGEMS“

Fig. 11: Does the trend patterns still hold significance, when the threshold is increased from 80%? The authors should rethink the value, upon which significance is given. As mentioned in the general comment 1, the authors could think about a way of indicating a trend significance by accounting for the uncertainties arising from the underlying mean state biases. Here, I would recommend to focus on the strength of the SST bias as the AMIP simulations indicates the importance of correct SSTs.

Line 531: You are writing „previous studies“, however, there are no studies cited. Please add the reference to the studies that you are addressing.

Line 576: „?“ - Reference list probably includes a mistake in the compilation of one reference.

Line 590 - 600: The subsection 5.5 provides a summary again rather than a discussion in the context of state-of-the-art research. Particularly, this section dedicated to future changes requires a thorough discussion due to the lack of reliability emerging from the mean state biases.

Line 659: Where is the reference to the „studies with similar results“?

Line 667: Why are these types of simulations considered particularly promising for a more process-based analysis? The model output ultimately determines which processes can be examined, largely independent of resolution. In general, the reader would also expect a combined assessment of the three processes considered. Specifically, which of the three factors (jet stream, SSTs, storm tracks) has the strongest impact on improving the representation of blocking, and which one requires the most improvement?