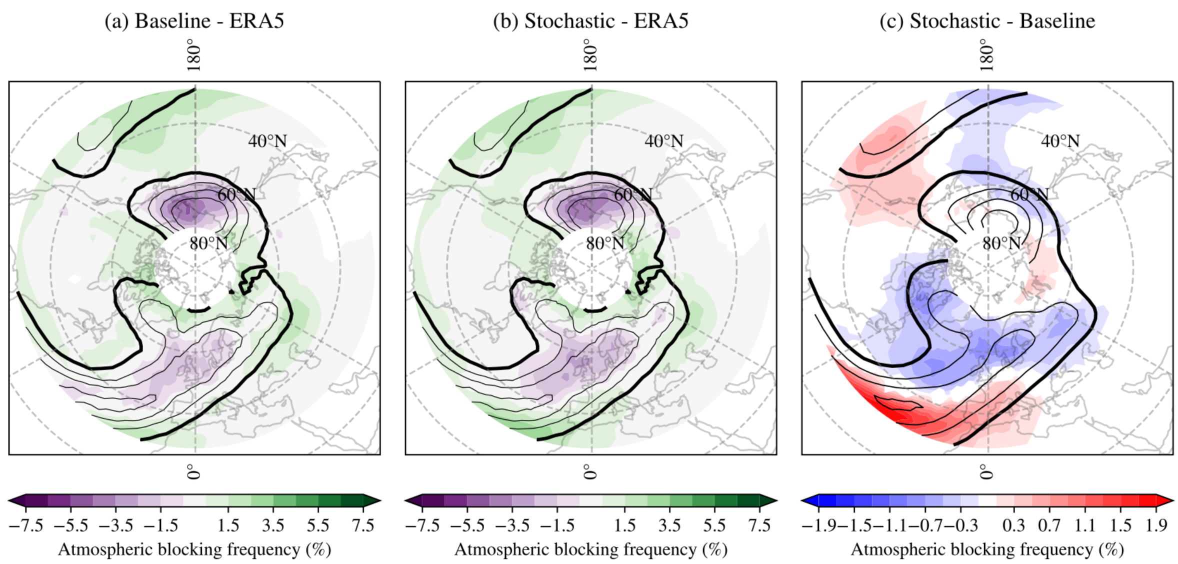
Response to reviewers' comments on 'Impact of stochastic physics on the representation of atmospheric blocking in EC-Earth3' - Michele Filippucci, Simona Bordoni and Paolo Davini

We thank the reviewers for their useful and thoughtful comments, which have significantly contributed to improving the quality of our study. We have revised the manuscript to address their concerns. Reviewers' comments are reproduced in blue, and our responses are in black.

List of relevant changes:

- During the final phase of our internal review process, we mistakenly renamed the plot of the blocking frequency (Figure 1) and reported on the manuscript the plot for the high-resolution runs (3 members, TL799), rather than the one referring to the average among all resolutions. This created a mismatch between what is written in the Section 3.1. and what is shown in Figure 1. We apologize for the confusion this mistake raised and we now show the correct plot in the revised manuscript.The corrected plot, which is also reproduced below for convenience, clearly shows how blocking frequency is under-estimated by the baseline model over Europe and the North Pacific.
  On the contrary, in the previously submitted manuscript the atmospheric blocking frequency over Europe seemed to be well represented by the baseline version of the model. This feature is a characteristic of the high-resolution runs, which are small in number and do not contribute much to the ensemble average. An earlier study (Davini, von Hardenberg, et al., 2017) attributed this apparent improvement to a compensation of errors due to the more resolved orography and a change in tropical precipitation (identified as the source of an upper-troposphere planetary wave).



- We slightly changed our conclusions by taking into account the work from Pickl et al. 2022, who showed how stochastic parameterizations have an asymmetric effect on the rapidly ascending motions in the tropical atmosphere (lines 429-456).
- We performed an additional analysis using a different blocking index, based on the geopotential height anomaly. This new analysis shows slightly different results than the main analysis but it does not change our conclusions. We rephrased some parts of the text

accordingly, we motivated the adoption of the gradient reversal index in the Methods Section (lines 180-186) and we included the new analysis in the Supplementary Material.

Following this important corrections, the point-to-point answers to the reviewers' comments are below:

REVIEWER 1

We thank the referee for taking the time for a detailed review and the constructive feedback on our manuscript.

The paper at hand presents an analysis of the influence of stochastic perturbations on the representation of wintertime blocking over Europe in a global climate model. First, it is shown that stochastic parametrizations decrease blocking frequencies, especially over Europe, and thereby deteriorate the representation of atmospheric blocking. Thereafter, the authors investigate mechanisms how the blocking frequencies over Europe are changed by the stochastic perturbations. They show that the mid-latitude jet stream and the tropical overturning circulation (i.e. the Hadley cell) are strengthened through SPPT, and argue that these changes in the mean winter circulation are responsible for the deterioration of the blocking representation. The authors claim that the changes in the midlatitude blocking representation originates from modified diabatic heating in the tropics, which impacts stationary eddies in the tropics and hence modifies the extratropical circulation, and emphasize the importance of the correct representation of the tropical circulation for atmospheric blocking.

The paper provides an interesting perspective on atmospheric blocking in the context of stochastic parametrization and is a valuable contribution to the existing literature, as it aims at providing a process-level understanding of how stochastic parametrizations affect the mean state of the model – an aspect which has not received too much attention. The presentation of the results (figures, equations, formulations, etc.), however, requires substantial revision, especially in the first parts of the paper.

General comment:

- The presentation of the results partly does not meet the required scientific standards. For example, for equations 1 and 2, the used symbols are not explained. For all figures, the units are not displayed at the color bar. In Figure 6, wind vectors are displayed, but the caption does not explain what they show. Most abbreviations are used without explanation (MJO, IFS; ECMWF, …). Together, all these aspects disturb the readability of the manuscript and can be improved without a lot of effort.

Thanks for pointing out these mistakes. We added the description of all symbols in the first two equations, we added the definition of the wind vectors and introduced and explained acronyms at their first appearance in the text. Moreover, we added units in the colorbar of the plots.

Specific comments:

- Line 8 (Abstract): I would not use the word "Surprisingly" here, as there is no indication for the reader why this should be surprising at this stage of the paper. I suggest to use a more neutral formulation, such as: "We show that the activation of …".

We rephrased 'we show', as suggested.

- Lines 22-23: This sentence is not correct. Depending on the configuration of the blocking, winds can be particularly strong on the northern flank of the blocking (for example when the jet is deflected northward)

We substituted 'northern' with 'southern' (line 23).

- Lines 27-29: It might be worthwhile to add that not only climate models, but also numerical weather prediction models struggle with predicting blocking over Europe (see for example Quinting and Vitart, 2019)

Thank you for pointing this out. We added a sentence at the end of the paragraph (line 36): "*Numerical Weather Prediction (NWP) models also have similar struggles with representing atmospheric blocking onset, systematically overestimating the propagation distance of Rossby Wave packets (Quinting and Vitart, 2019).*"

- Line 41: It was not the initial purpose of stochastic parametrizations to be an alternative to resolution increases, but to represent uncertainty in the unresolved scales. Please rephrase this sentence, as it is misleading.

Here you are raising an important point, which we tried to address in the revised version by rephrasing the related paragraph in the Introduction (lines 45-52). As you rightly pointed out, originally SPPT was intended as a way to represent model uncertainty arising from the unresolved scales of the atmospheric flow. However, more recent studies highlighted how representing the influence of sub-grid scale variability on the synoptic-scale flow by introducing stochastic perturbations have a beneficial impact on the model mean state (Berner et al., 2012) (lines 57-58). These improvements are similar to those achieved by increasing the resolution and such findings serve as motivation for our study.

- Lines 64-65: Adding to the discussion of possible dynamical reasons leading to changes of the mean state of the model, you might be interested in the two following references that deal with a process-based understanding of the impact of stochastic perturbations: Pickl et al. 2022, and Deinhard and Grams, 2023. Even though they focus on different processes, their findings fit well into the scope of your discussion.

Thank you for this suggestion. We found these articles very useful and we cited them in the paper adding their considerations both to the introduction (line 61) and conclusions (line 441) sections.

- Lines 65-85: Consider reordering this paragraph. After line 65, it would be good to discuss the body of literature that has dealt with the impact of stochastic perturbations on different aspects of the model (i.e. MJO, ENSO, tropical cyclones, etc.) without a specific focus on the Climate Sphinx data set. After that, the research gap can be identified (i.e. connection to blocking), the question can be posed and the data set can be mentioned.

Originally, we decided to separate the references dealing with the Climate SPHINX experiment from those dealing with other experiments. It seemed wrong to mention studies dealing with Climate SPHINX before even mentioning Climate SPHINX. However, we now tried to follow your suggestions and we found that the introduction works better. Thanks for this very useful suggestion.

- Line 86: “With these goals in mind”: When implementing the previous comment, the reference to the research goals are much clearer.
- Line 102: How is the ensemble for the baseline simulation generated, when no stochastic perturbations are used? Please clarify.

For each resolution, multiple ensemble members have been run by providing different initial conditions (ICs). A first experiment is run at each resolution for a few days, and it is used to create the ICs for the other experiments. For instance, for T255 experiments, the ICs for the 10 ensemble members are extracted using the midnight values (00:00) from each of the first 10 days respectively, and then reassigned to 1 January. This information can be found in the Climate SPHINX description paper (Davini, von Hardenberg, et al., 2017), as reported in the manuscript (line 104).

- Lines 104-106: This sentence is not very precise. Consider rephrasing it along the following lines: “For each of the two setups, simulations have been run with different resolutions ranging from 125 km to 16 km, with decreasing ensemble size for increasing resolution.”

We rephrased it following your suggestion (line 108).

- Lines 109-110: Isn’t the main reason that blocking is the most frequent in winter and spring time?

Here we don’t understand what you are referring to.

- Lines 121-122: If the TL1279 simulation is not considered, please remove it from Table 1.

It is not used indeed. We removed it from the table.

- Line 130: It might be helpful to add that by averaging all ensemble members, the low-resolution simulations have a larger weight than the high-resolution simulations.

We added this information in the methods section, at line 128.

- Line 138: I don’t really understand the sentence. What do you mean by “reference value for the 98% confidence interval for 29 ensemble members”?

We changed this sentence slightly:” *for 29 ensemble members”* now is *“considering the size of the ensemble.”(line 144).* When performing the student test it is necessary to consider the number of ensembles that have been used to perform the averages and standard deviations.

- Lines 147-148: This sentence is not very nice to read. Consider rephrasing this sentence along the lines: “Even though the SKEB scheme accounts for variability that is not represented in the deterministic version of the model, its impact on the model climate is negligible (Davini et al., 2017b).”

We rephrased the sentence as suggested.

- Lines 150-151: Also here, I suggest reformulating the sentence to something like: “SPPT introduces variability by perturbing the deterministic parametrization tendencies of the temperature, specific humidity and wind fields as follows:”

We rephrased the sentence as suggested.

- Line 158: In the context of your results, it is interesting to add that the magnitude of the SPPT-perturbations on average scale with the magnitude of the deterministic tendency, which are larger in the tropics than in the extra-tropics (see Leutbecher et al., 2017). Therefore, the most prominent effects of SPPT occur in the tropical regions.

Thank you for pointing this out. We added this consideration to the introduction (line 66).

- Lines 206-210: I cannot bring together your discussion of Figure 1a and the Figure itself. You write that the baseline model underestimates the blocking frequencies over Europe, but I mainly see white to light red colors over Europe in Figure 1a, indicating a slight overestimation of the frequencies compared to ERA5. You also describe that “the model significantly underestimates blocking at high latitudes in the Siberian region”, but overall, the reddish colors predominate. In contrast, the large biases over the East Pacific are not discussed at all. Please carefully review this paragraph.

This comment arises from a mistake we made while uploading the plots to the Latex document, as described in Introduction of this answer to the reviewers. The discussion should now be coherent with the updated plot. We changed ‘Siberian region’ with ‘Bering Strait region’ at line 235. That was inattention. Thanks for pointing this out.

- Lines 210-220: Even though this paragraph is important, it disturbs the discussion about the first results in Figure 1. I would suggest shifting the whole paragraph into the Methods section where also the blocking index is described, and once again refer to the issue here.

We followed your suggestion and moved the paragraph to the methods section (lines 198-203).

- Line 221: Again, I don’t like the term “surprisingly”, as it suggests that a different result is expected, but the reasons for the speculation are not clear to the reader. It would make sense to do this if you wrote something like “Unlike reported in REFERENCE, stochastic parametrizations do not improve the representation of … ”. However, if this is not the case, I would suggest to stick to a neutral formulation, such as “We show that stochastic parametrizations do not improve the representation of …”.

We avoided using the term “surprisingly” by rephrasing the paragraph (line 231).

- Line 230: According to Figure 1, the blocking frequencies in the North Pacific are as large as in the North Atlantic, so please find a different motivation to focus on the North Atlantic region.

We removed that sentence. In the updated plot the reason why we focus on the Atlantic is more evident. Changes introduced by the stochastic parameterizations are much larger in the North Atlantic than in the North Pacific (line 243).

- Lines 240-244: It would be easier to understand the message of Figure 2a if the paragraph starts with a simple description of the differences between the baseline and stochastic experiments. For example: “The stochastic parametrizations result in a strengthening of the upper-tropospheric winds. This is evident for both the northern and southern hemispheres, even though they are in different solsticial conditions”.

We reordered the sentences as suggested. We also modified the paragraph by moving the discussion of the observed differences at the end of the Section (lines 283-287).

- Lines 258-259: Can you give more details of why the strengthened Hadley circulation can explain the strengthened jet?

We reorganized this section by moving the discussion of the changes in the mean state at the end. The connection between Hadley circulation and the midlatitude jet should be clearer now (lines 283-287).

- Lines 260-261: As the zonal mean wind or the mean overturning circulation, also TKE and EGR are quantities to characterize the mean state of the model. Please rephrase this sentence.

We rephrased the sentence to: *“In addition to changes in the climatological winds, we investigate the impact of the stochastic parameterization on the high-frequency atmospheric variability by analyzing two mean-state diagnostics.” (line 269).*

- Line 268: “This is compatible with what has already been found by …”

Corrected

- Lines 304-308: I suggest to shift this paragraph to the Methods section, in which the blocking index is introduced, and specify where you are using grid-point based frequencies and where you use other definitions.

We followed your suggestion and moved the paragraph to the methods section (line 212).

- Lines 415-419: Here again, the findings of Pickl et al. 2022 are of interest, who discuss a very similar threshold-behavior of rapidly ascending air streams that is observed with SPPT.

As above.

- Lines 449-451: I don't understand the last part of the sentence, after the comma. Please rephrase.

We rephrased the sentence *'More specifically, a slightly modified zonal wind climatological pattern can result in large differences in atmospheric blocking behavior, result also confirmed in a later study'* to *'More specifically, even a slight alteration in the zonal wind climatological pattern can lead to significant changes in atmospheric blocking behavior, a finding that was corroborated by a later study' (line 466).*

- Line 452-454: I don't fully agree with this sentence, as you do not verify if the stochastic parametrizations improve the Hadley circulation, the jet stream, etc. However, I agree that your analysis shows that the representation of blocking is affected by the tropical circulation, and hence it is important to accurately capture the latter.

We partially rewrote the conclusions of our study taking into account the results from Pickl et al., 2022 (lines 429-456). We also avoided using the words 'mean-state biases', 'deterioration', as we do not compare the model with observations. We rather argue that a retuning of the model when the SPPT parameterization is activated seems necessary, as such a parameterization leads to significant mean-state changes.

Technical comments:

We addressed all your technical comments in the revised manuscript with few changes:

We corrected 'Global Climate Model' by using the term 'General Circulation Model' (line 27), as suggested in one of the comments.

We used moderate progress rather than moderate advances (line 29). Probably you were suggesting to use the word 'small progress' but we did not feel it would have been the right word. For example, the representation of blocking in the Pacific has been substantially improved in recent GCMs

- Line 27: GCM stands for "Global Circulation Models" and not "Global Climate Models"
- Lines 29-30: Consider choosing a different formulation here, as it is not clear what you mean with "moderate". Do you mean that only little progress has been made in the field of blocking?
- Line 39: "trade-off" instead of "trade off"
- Line 45-46: "more accurately": The propagation of errors cannot be represented more accurately, but it can be accounted for.
- Line 54: "MJOs": Introduce the term Madden-Julian-Oscillation after the first usage of the abbreviation. Further, there is no plural of the MJO (leave out the "s").
- Line 56 and other instances: "Northern Atlantic": Please use "North Atlantic", as this is much more common on the literature.

- Line 62: Introduce "Integrated Forecasting System"
- Lines 56-58: Please rephrase the second part of the sentence, along the lines: "… weather regimes, which are recurring and quasi-stable patterns of regional weather conditions …"
- Line 75: "atmospheric circulation" instead of "atmospheric circulations"
- Line 89: "dynamical" or "process-oriented" instead of "mechanistic"
- Line 94: Introduce the meaning of Climate SPHINX at the earliest appearance.
- Line 99: "considered here" instead of "here considered"
- Line 120: "scale-aware" instead of "scale aware"
- Line 127: Delete "in fact"
- Line 138: Add information of the equation, i.e. what is $<x1>$, $\sigma 1$
- Line 143 and 147: For the claim "The SKEB scheme", you use different references in two occasions. Please either unify these citations, or leave them out the second time. Further, I placing the reference at the end of the sentence.
- Line 146: As above, please introduce the symbols and terms that you are using in the equation
- Line 156: Use "independent patterns" instead of "independent perturbations"
- Lines 166-170: You use the term "particular" very often in this paragraph. Please avoid such repetitions.
- Line 172: Add GHGN and GHGS in brackets after "northward" and "southward", respectively.
- Lines 178 and 180: Add units after $\phi N = \phi 0 + 15$ (i.e. $\phi N = \phi 0 + 15°$ lat) and in equation 6 (-10m °lat-1)
- Line 182: "occurrence" instead of "timing and positioning"
- Line 230: "motivated" instead of "supported"
- Line 324: "it becomes clear that" instead of "it becomes clear how".
- Line 332: "larger" instead of "greater"
- Line 337: "use" instead of "implementation"
- Line 432: "To determine underlying mechanisms …" instead of "Underlying mechanisms …"

References:

Deinhard and Grams, 2023: https://egusphere.copernicus.org/preprints/2023/egusphere-2023-1938/

Leutbecher et al, 2017: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3094

Pickl et al, 2022: https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4257

Quinting and Vitart, 2019: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2018GL081381

REVIEWER 2

General comment

We thank the referee for taking the time to review our paper and for providing useful comments. We tried to improve our manuscript by following your suggestions.

The paper assesses the impact of stochastic parameterization on the blocking representation in a climate model. Although the study is in general well written, I have several concerns about the results and their interpretation, which are stated below.

Specific comments

Several times the authors state that the blocking frequency is underestimated in climate models, in particular over Europe (e.g. L28). This serves as a main motivation for the study. However, I don't see this underestimation in the baseline simulations presented in Fig. 1. If anything I would say the blocking frequency is slightly overestimated over Europe. Furthermore, the baseline model seems to have a much larger problem over the north Pacific with respect to blocking. Hence, the motivation and the focus on the European/Atlantic region is not evident to me.

In the updated plot (Figure 1), it is possible to observe how the European atmospheric blocking frequency is generally underestimated by EC-Earth, while in the Pacific the frequency is overestimated at low latitudes and underestimated at high latitudes (line 233).

Furthermore, the statement that the blocking frequency is underestimated in climate models, especially over Europe, is sustained by a large body of literature (Davini & d'Andrea, 2020; Masato et al., 2013; Woollings et al., 2018). The same papers highlight how biases in the Northern Atlantic persist even in recent generations of the CMIP project (e.g CMIP6), while biases in the Pacific have improved, and are more diverse among different Earth System Models (lines 27-36).
At the same time, as pointed out in the first lines of this answer to the reviewers, biases in the Northern Atlantic decrease when the resolution is increased. This fact was already noted in an earlier study (Davini, Corti, et al., 2017) and has been attributed to a compensation of errors due to the more resolved orography and a change in tropical precipitation (identified as the source of an upper-troposphere planetary wave). Even though the high-resolution version of EC-Earth may represent well blocking climatology over Europe, the motivation of our study holds, as stochastic parameterizations may represent a computationally cheaper alternative to improve blocking representation in the Atlantic (lines 50-53). Moreover, to further motivate our focus in the European region, by looking at Figure 1 and S1-S4, it is evident how stochastic parameterizations mostly impact the Northern Atlantic region (lines 242-243).

While the blocking frequency is compared to observations (ERA5), features of the mean state circulation (explored in Figures 2-5) is not. There, only the differences between baseline and stochastic runs are compared, which seem systematic and physically understandable, but rather small. I would like to see how these differences compare to the mean state model biases with respect to the ERA5 analysis.

Through your comment you are raising an important issue: are the stochastic parameterizations generally improving the representation of the mean-state circulation in the Climate SPHINX experiment? Several previous studies focused on this point and are mentioned in the Introduction (lines 53-79) (Davini, von Hardenberg, et al., 2017; Strommen, Watson, et al., 2019; Vidale et al., 2021; Yang et al., 2019). For example, the fact that the Northern Hemisphere jet stream is slightly strengthened was already observed in Davini, von Hardenberg, et al., 2017, while changes in the Hadley circulation have been investigated by Strommen, Christensen, et al., 2019. In this paper we focus on the impact of these parameterizations on blocking representation. Figures 2-6 refer to the differences between stochastic and baseline simulations because our focus is on blocking representation biases, which in turn are influenced by these circulation differences. Bringing ERA5 reanalysis in the discussion of these circulation modifications could cause confusion and it is not the scope of this paper.

For completeness, we include in the supplementary material the comparison between the stochastic and baseline versions of the model and ERA5 reanalysis with regard to the transient kinetic energy and the Eady growth rate, which are not documented in other papers on the same dataset. We found that changes are rather small compared to the original bias, in line with the broader result highlighted by previous literature which show that stochastic parameterizations improve - but usually marginally - the mean state (Leutbecher et al., 2017).

The statement that an analysis grouped by resolution did not reveal significant differences (L130) deserves more explanation. This would be surprising and contradicting previous studies mentioned on L35, even on the same dataset. Figures to supposedly show this are presented in the supplement. However, I wonder if the figures S2-S4 are really correct. They seem bit-wise identical. After staring at them for several minutes I could not spot a single pixel which is different. This seems unrealistic and should be checked. If it is really true, I wonder what the explanation is and why the increasing resolution is here not beneficial for blocking simulation.

We corrected the plots in the supplementary materials. Thank you for pointing this out, we apologize for the mistake. It is now possible to observe how increasing resolution is beneficial for blocking representation, even though this apparent improvement is rather the effect of a compensation of errors (Davini et al. 2017).

Regarding the impact of stochastic representation at different model resolutions, figures S2-S4 now show how the differences induced by the stochastic parameterizations do not depend on the model resolution. This fact is the most relevant for our analysis (lines 135-139). Moreover, it was surprising as the relative importance of the stochastic parameterization should decrease as resolution is increased (see next comments and Buizza et al., 1999; Palmer, 2001). However, as discussed in the paper, this systematic effect can be explained by an asymmetric behavior of condensation and convective processes (lines 429-456).

The authors state that there are several methods of blocking detection (L161), which can lead to significantly different results. The study however only uses one method. I wonder how robust the findings are and if they can be reproduced if another method was used.

In our work we adopt a Lagrangian Tracking Algorithm that is based on a geopotential height gradient reversal index. Implementing the same algorithm for an anomaly based index would imply several considerations that are outside the scope of our study. However, we included in the supplementary material a plot that makes use of a simple implementation of an anomaly based index (see Supplementary Material Figure S8). We refer to this additional analysis in the Methods (line 175) and Conclusions (line 467) Sections. Through this simple algorithm we find that the overall impact of the stochastic parameterizations, regardless of the resolution, is smaller than the one obtained through the gradient reversal index and does not lead to a deterioration or improvement of blocking climatology representation. Yet, we find that the stochastic parameterizations yield to an eastward displacement of the blocking climatological maximum both at the exit of the Pacific and Northern Atlantic storm track. For a brief explanation of the method see the Supplementary Material. This additional analysis highlights how the gradient reversal index is more sensitive to changes in the model mean state than the anomaly index. This fact was expected, as the former takes into account the full field, while the latter examines the anomaly relative to the mean state. Earlier literature on the topic highlighted the same feature (Scaife et al., 2010, 2011).

We think this additional analysis does not change our main results, as anomaly indices and absolute indices can detect different kinds of blocking. Moreover, we rephrased our conclusions to account for the fact that different indices can give slightly different results. We discuss this topic briefly in the Supplementary Material.

It is not plausible that scales of 500-2000km, which are used in the SPPT correlation patterns represent the sub-grid variability (L158). I would rather argue that they represent flow-dependent biases of the parameterizations which are correlated over the size of weather systems. I think the authors should better explain why such large correlation patterns are used in the model setup and why the authors hypothesize that they may beneficially effect the blocking representation as suggested in L30-L47 (see also last comment).

Here you are raising an important matter which we try to further clarify: unrepresented sub-grid scale processes may influence synoptic scale systems by introducing a source variability that is otherwise missing in the model. The correlation scales of the stochastic noise are therefore chosen in order to account for this missing variability which, originating from sub-grid scale processes, influences the large-scale flow through scale-scale interactions (Leutbecher et al., 2017; Palmer, 2001) (line 166). Note that 500km, the smallest autocorrelation scale, is chosen in order to be significantly larger than the grid cell size. If the autocorrelation scale were smaller than the grid cell size, it would have been indistinguishable from uncorrelated random noise.

Along those lines, I don't see why the impact of the stochastic parameterizations, especially SPPT, would decrease with increased resolution (L126). SPPT uses large spatial and temporal correlation patterns that modify the tendencies from the parameterizations in a similar way, regardless the resolution.

The parameterizations implemented in the atmospheric model (IFS) of EC-Earth are scale aware. This means that the relative importance of the parameterizations changes together with the resolution. In particular, the parameterization tendencies are expected to be smaller in magnitude when the resolution is increased, as more and more scales of variability are explicitly resolved by the model (line 132). Since the stochastic noise in the SPPT scheme is multiplied by the parameterization tendencies, the impact of the parameterisation is expected to decrease when the resolution is increased (Buizza et al. 1999, Palmer 2001) (line 161).

L206: I can't follow the description here. Overestimation means red, right? So I do see an overestimation over the Pacific south at around 40°N. I don't see an underestimation over Europe and also not over Siberia. I do see a pronounced underestimation over the north Pacific and Alaska. Very confusing.

The underestimation over Europe is evident in the updated plot. We also substituted 'Siberia' with 'Bering Strait' (line 235).

L223: I don't agree. The anomaly over Siberia has also gotten worse with the stochastic parameterization.

We corrected 'high latitudes' with 'over Greenland and in the northernmost regions of the American continent' (line 239).

The authors consistently attribute the detrimental changes in blocking frequency seen in the stochastic runs to changes in the mean circulation, caused by SPPT. A retuning of the model is suggested (L441). Suppose one did that, do the authors expect any impact of the stochastic parameterization on blocking if the mean state was not changed? Are there any reasons for investigating this further, given that higher resolution simulations don't seem to affect blocking (L130, or does it?) and stochastic physics is used as “a cheaper alternative to increasing resolution” (L41). Some discussion on those issues would be helpful.

We changed the Conclusions Section of our paper by implementing the following changes:

- When referring to the model mean circulation we avoid terms such as ‘deteriorate’ or ‘misrepresentation’, rather arguing that the activation of SPPT leads to ‘significant changes’ in the model mean state (line 455).
- We argue that a retuning of the model when the SPPT parameterization is activated seems necessary because of unexpected significant changes to the mean state (line 447). This is likely due to an asymmetric response of condensation and rapidly ascending motions to stochastic parameterizations (lines 429-456).
- After a retuning, it would be possible to assess potential improvements of blocking representation that are now offset by these mean state changes. These improvements however are likely to be small compared to EC-Earth bias (lines 457-461).

Please note that a resolution increase does have an impact on blocking representation (Berckmans et al., 2013; Schiemann et al., 2017). On the contrary, in the Climate SPHINX model setup, a resolution increase does not change the effect of stochastic parameterizations on the mean state and on blocking climatology (see updated figures S1-S4). We again apologize for the confusion we caused by updating the wrong Figure and we hope our motivation is now clear.

Minor comments

L59: what is seen

The blocking improvement described above (lines 39-40).

The section “In this paper…” of the introduction should be the last part. I suggest to move the paragraph L76-L85 up.

We reordered the introduction following your advice and the advice from the 1st reviewer.

L143: SKEB is not part of Buizza et al., 1999

Thank you for pointing out the mistake.

L154: large-scale water processes -> microphysics

‘Large-scale water processes’ stands for the water cycle, which is a different thing from microphysics. (line 163)

L181: I don’t quite understand this paragraph. In particular the choices of rejecting some events seem arbitrary.

The algorithm is further explained in the Appendix (line 493). We added a sentence at line 209: *‘These thresholds have been chosen following earlier Literature on the topic. By applying these filters we select blocking events that interest the synoptic scale and are quasi-persistent events, coherently with common blocking definitions.’*

L406: “anomalous”. Do we really know that? Or is it just altered compared to the baseline runs?

Strommen et al., 2019 highlight how the cloud liquid water content of tropical clouds increases in an unrealistic manner when SPPT is activated. However, we relax the term ‘anomalous’ to ‘different’ as in our analysis we do not compare the tropical circulation to reanalysis.

Some new results are also presented in the conclusion section (L428). I find that confusing.

We moved the discussion on the top of the atmosphere radiative budget to Section 4.2. (lines 402-409)

Fig 1: I don’t understand what you mean by shaded contours.

We changed the term to ‘shading’.

Section 2.4: What is Fourier transformed? I assume only the time? And then it is transformed back with only time scales between 2-6 days maintained?

We added ‘temporal’ (line 219).

Answer to reviewers bibliography:

Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, *14*(1), 34–40.

Berner, J., Jung, T., & Palmer, T. N. (2012). Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *Journal of Climate*, *25*(14), 4946–4962.

Buizza, R., Milleer, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *125*(560), 2887–2908.

Davini, P., Corti, S., D’Andrea, F., Rivière, G., & von Hardenberg, J. (2017). Improved winter European atmospheric blocking frequencies in high-resolution global climate simulations. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2615–2634.

Davini, P., & d’Andrea, F. (2020). From CMIP3 to CMIP6: Northern Hemisphere atmospheric blocking simulation in present and future climate. *Journal of Climate*, *33*(23), 10021–10038.

Davini, P., von Hardenberg, J., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A. G., Weisheimer, A., & Palmer, T. N. (2017). Climate SPHINX: evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model. *Geoscientific Model Development*, *10*(3), 1383–1402.

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., Bonavita, M., Christensen, H. M., Diamantakis, M., Dutra, E., & Others. (2017). Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, *143*(707), 2315–2339.

Masato, G., Hoskins, B. J., & Woollings, T. (2013). Winter and summer Northern Hemisphere blocking in CMIP5 models. *Journal of Climate*, *26*(18), 7044–7059.

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, *127*(572), 279–304.

Pickl, M., Lang, S. T. K., Leutbecher, M., & Grams, C. M. (2022). The effect of stochastically perturbed parametrisation tendencies (SPPT) on rapidly ascending air streams. *Quarterly Journal of the Royal Meteorological Society*, *148*(744), 1242–1261.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., & Williams, K. (2011). Improved Atlantic winter blocking in a climate model. *Geophysical Research Letters*, *38*(23).

Scaife, A. A., Woollings, T., Knight, J., Martin, G., & Hinton, T. (2010). Atmospheric blocking and mean biases in climate models. *Journal of Climate*, *23*(23), 6143–6152.

Schiemann, R., Demory, M.-E., Shaffrey, L. C., Strachan, J., Vidale, P. L., Mizielinski, M. S., Roberts, M. J., Matsueda, M., Wehner, M. F., & Jung, T. (2017). The resolution sensitivity of Northern Hemisphere blocking in four 25-km atmospheric global circulation models. *Journal of Climate*, *30*(1), 337–358.

Strommen, K., Christensen, H. M., MacLeod, D., Juricke, S., & Palmer, T. N. (2019). Progress towards a probabilistic Earth system model: examining the impact of stochasticity in the atmosphere and land component of EC-Earth v3. 2. *Geoscientific Model Development*, *12*(7), 3099–3118.

Strommen, K., Watson, P. A. G., & Palmer, T. N. (2019). The impact of a stochastic parameterization scheme on climate sensitivity in EC-Earth. *Journal of Geophysical Research, D: Atmospheres*, *124*(23), 12726–12740.

Vidale, P. L., Hodges, K., Vannière, B., Davini, P., Roberts, M. J., Strommen, K., Weisheimer, A., Plesca, E., & Corti, S. (2021). Impact of stochastic physics and model resolution on the simulation of tropical cyclones in climate GCMs. *Journal of Climate*, *34*(11), 4315–4341.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A.

R., & Seneviratne, S. (2018). Blocking and its response to climate change. *Current Climate Change Reports*, *4*, 287–300.

Yang, C., Christensen, H. M., Corti, S., von Hardenberg, J., & Davini, P. (2019). The impact of stochastic physics on the El Niño Southern Oscillation in the EC-Earth coupled model. *Climate Dynamics*, *53*, 2843–2859.