We thank the reviewer for their careful review and for providing useful suggestions on how to improve this manuscript. Below, we first respond to the general comments, after which we provide detailed replies (in blue) to the specific comments.

The manuscript The effect of noise on the stability of convection in a conceptual model of the North Atlantic subpolar gyre explores the sensitivity of the subpolar gyre (SPG) convection to noise by adapting and building on the conceptual model of Born and Stocker (2014, hereafter BS14). The BS14 model is non-dimensionalized and made autonomous in order to perform a bifurcation analysis. In addition, stochastic noise is added to represent variability in surface current's salinity and in freshwater forcing. The bifurcation analysis shows two stable states in the system: convective and non-convective, where the non-convective state is defined by the total non-dimensional volumetric transport in the gyre M ≤ 22 Sv. The stability of the model is explored by analysing the sensitivity of convection to noise in salinity and freshwater forcing. It is found that the salinity noise impacts convection in the gyre significantly more than the noise in freshwater forcing. Additionally, it is found that the SPG recovers from the non-convective state across all tested parameters, a result which is not commonly found in Earth System Model (ESM) studies. The quality of the scientific analysis in the manuscript is good and the topic is both important and thematically suitable for the Earth System Dynamics Journal. However, the study could go further in terms of the impact of the results. In the present form, the novelty of the research presented in the manuscript is questionable. This review presents some suggestions as to how the authors could push their study further.

General comments

The current description of the conceptual model is lacking. At multiple instances, the authors provide citations to previous work without outlining how these choices fit into the current model (for example, ll. 125-127: how is the value for τX picked?; ll. 445-448 what is the effect of picking $k \gg 1$ on the model?).

Reply: We thank the reviewer for pointing this out. In an earlier version of the manuscript the reference to Dijkstra et al. (2023) (l. 126) was made to justify the use of an Ornstein-Uhlenbeck process. Later, we added the more relevant references to Penland and Ewald, 2008; Boers et al., 2022; Ditlevsen and Johnsen, 2010 in l. 122-123. These more relevant references make the reference to Dijkstra et al. (2023) redundant and we will therefore remove this reference from the manuscript. The choice of tau_X in the Ornstein-Uhlenbeck processes is motivated in the next paragraph in the manuscript (ll. 129 – 135, see also the discussion below).

Picking k >> 1 ensures that boxes 1 and 3 immediately mix if the stratification is unstable, in line with the original BS14 model. A lower value of k would result in slower mixing. Picking k >> 1 is a common choice to implement a step function analytically (Dijkstra, 2004, pp. 70).

Changes in text: we will remove the reference to Dijkstra (2023) and in l. 447 change "where k ≫1 (e.g. Dijkstra, 2004). In this study a value of k = 10^5 was used"

to

"where k >> 1 to ensure mixing between the boxes occurs instantaneously (e.g. Dijkstra, 2004, pp. 70). In this study a value of k = 10^5 was used."

We will also define all variables immediately after the Equations and expand our discussion of the parameters (see below).

BS14 provide an extensive discussion on the origin and physical meaning of the conceptual model parameters. Since this model is adapted in the current work, such in-depth discussion is not necessary – but sentence summary for different model parameters would greatly improve the transparency and clarity of the text. For example, mentioning that Ubtp corresponds to the volumetric transport of 20 Sv would be useful.

Reply: We agree that the current presentation of the model parameters in the text is lacking and will extend it.

Change in text: In l. 120 we change

“All other parameters are prescribed.”

to

“All other parameters are prescribed: U_btp is the barotropic component of the current and has a strength of 20 Sv, r represents the ratio of the surface and deep box height, eta the strength of the thermal wind, mu_H the horizontal mixing efficiency, mu_C the convection efficiency, and mu_A the atmosphere-ocean exchange efficiency.”

In the description of Table 1 we change

“The values were calculated from the default model parameters as outlined in Table 1 of Born and Stocker (2014) and Appendix A. No values are given for the parameters eta, mu_H, mu_C, and mu_A, as these values do not have an intuitive interpretation.”

to

Suggestion: “We use the parameter values in Born and Stocker (2014) (their Table 1, here Table A1), to compute the parameter values for our non-dimensionalised model. Their parameter estimates are based on observations and expert assessment. The non-dimensionalisation introduces some additional dimensionless parameters (eta, mu_H, mu_C, and mu_A), for which no dimensional values are given for lack of interpretability.”

The choices made in connection with extending the BS14 model should also be clarified. On which basis were the values for τS and τF picked?

Reply: We selected the values of tau_S and tau_F to represent the different time scales of variability of ocean and atmosphere (ll. 132 – 135). Changes in the gyre salinity propagate on a time scale of several years (e.g. the Great Salinity Anomalies), whereas changes in the freshwater flux happen on a shorter (seasonal) timescale. For clarity we will rephrase the corresponding lines in the manuscript.

Change in text: In l. 132-135 we change

"To simulate the different intrinsic time scales of variability in ocean and atmosphere, correlation time scales of tau_S =1 yr and tau_F = 90 days were used unless specified otherwise. With these time scales, the stochastic variations in gyre salinity S2 (described by zeta_S) can be interpreted as being driven by external variations in for example sea ice cover, and the stochastic variations in freshwater forcing F (described by zeta_F) as quasi-seasonal variations in precipitation."

to

"We select the correlation timescales tau_S and tau_F such that the noise processes represent the time scales of variability in ocean and atmosphere; tau_S =1 yr and tau_F = 90 days (unless specified otherwise). This means that changes in the gyre salinity have a timescale of years, corresponding to that of the Great Salinity anomalies and driven by e.g. sea ice cover variations. Changes in the freshwater forcing correspond to quasi-seasonal precipitation variability."

What is the relation between parameters $c*$ in BS14, and $c1$ and $c2$ in the adapted version of the model? An alternative mechanism for convective mixing is introduced without sufficient justification or description. How is the value for $c2$ chosen?

Reply: Parameter c in BS14 corresponds to c1 in our model, with c1 = c* A / V (BS14 eq. 15) and c* = 0.03 denotes the mixing efficiency.

We parameterize convection analytically, rather than handling it computationally at every time step (i.e. mix when the box 1 density exceeds that of box 3). Conceptually, the analytic convection term contains the terms.

[strength of convection] * [step function that ensures convection only occurs under unstable stratification] * [difference in T/S].

Because we take a high value of k (see discussion above), this approach is equivalent to that of BS14. The difference is that our model does not contain a conditional step in the integration and therefore can be studied with continuation software.

Term c2 in our equations represents the strength of convection. We determine the value of c2 in relation to that of c1, since it is the relative importance of the horizontal and convective mixing terms that matters. We choose c2 such that the ratio c2/c1 = 10^3 to ensure box 1 and 3 are mixed very fast, staying close to the instantaneous mixing of BS14. This choice does not affect the bifurcation structure of the model, but only the time scale at which the steady state solution is reached.

Summarizing, instead of a discrete (conditional) mechanism for convective mixing as in BS14, we consider a continuous one by using a step function. We take the values of k and c2 such that this continuous mechanism is as close to the original BS14 model as possible.

Changes in text: In l. 441 we add

"Note that the (dimensional) parameter c_1 in our formulation is equal to c in BS14, such that c_1 = c* A / V."

In l. 452 we add

"Taking high values for both c_2 and k ensures that convective mixing between box 1 and 3 happens nearly instantaneously, in line with the implementation of convection in BS14."

We also add the values of c*, A, V to Table A1.

The discussion about the realism of the model is somewhat contradictory throughout the text. In the model description, the amplitude of the noise is described as unrealistic and the choice is motivated by exploring the mechanistic aspects of the system (ll. 136-140). In the discussion, the noise values are instead described as "on the high end of realistic values" (ll. 355-356). I agree with the authors that the realistic frequency of the non-convective state under the current oceanographic conditions can be seen as an argument for the robustness of the model and the magnitude of the noise parameters used. The discussion on this aspect of the model could be streamlined throughout the text.

Reply: We thank the reviewer for pointing this out. We will streamline the discussion and use "on the high end of realistic values" to describe the noise consistently throughout.

Section 4 does not convey that the results contribute significantly to the understanding of the dynamics of the SPG. It is not obvious to me that the study goes far enough beyond the analysis of the conceptual model dynamics in the BS14 paper. One of the main results of the study is that the SPG convection is more sensitive to noise in the gyre salinity compared to freshwater forcing. However, as the authors themselves point out, this may be due to the structure of the conceptual model (ll. 339-340). Could the robustness of this result be tested in additional experiments?

Reply: The structure of the conceptual model is such that it best represents the physical mechanisms relevant for the gyre variability. Hence, changing the structure of the model would mean it no longer accurately represents the mechanisms of freshwater forcing and gyre salinity changes. It is encouraging to see that with such a simple representation of the mechanisms results that are in line with observations can be found, i.e. that convection is more sensitive to variability in the gyre salinity than precipitation.

There is ample evidence that supports the hypothesis that convection in the Labrador sea is more sensitive to noise in the gyre salinity (brought in via the boundary currents) than noise in the freshwater forcing. For example, Yashayaev (2024) identified the freshening of the Labrador sea, caused by a release of low-salinity water from the Beaufort Gyre, as the main cause of the 2023 convective shutdown. Similarly, Gelderloos et al. (2012) argued that the Great Salinity Anomaly (a blob of anomalously low salinity moving through the region) initiated the shutdown of convection in 1969. In addition, they found the advection of saltier waters to be one of the reasons convection restarted in 1972. Conversely, we are not aware of analyses that show noise in the atmospheric freshwater forcing (i.e., precipitation in the center of the Labrador sea) contributing meaningfully to the variability of convection here.

It is worth noting here that the different response to noise in the gyre salinity and the freshwater forcing is not unique to the noise, but is a result of the difference in physical mechanism between the freshwater flux and gyre salinity terms in the salinity budget of S_1. Varying F directly increases or decreases the amount of freshwater that's added to box 1. By comparison, varying S_2 only indirectly changes the salinity of box 1 by first changing the strength of the baroclinic current U_s, and then the magnitude of horizontal eddy transport mu_H*U_s*(S_2 – S_1). This is a highly nonlinear process which reflects the physical mechanism by which salinity anomalies are transported from the boundary current to the convective core of the gyre.

Changes in text: We will add a discussion of observed shutdowns of deep convection in the Labrador sea, how these shutdowns often depend on changes in the salinity of the boundary current (and not on changes in freshwater forcing) and how our conceptual model adds to the mechanistic understanding of these observed shutdowns.

Another main result of the study is the resilience of the SPG convective state. The collapse and recovery of the SPG has been observed in at least one ESM study (Jochum et al. 2012). The physical mechanism which allows SPG to recover in the ESM is the freshwater flux through the Bafflin Bay. This and other ESM studies of the SPG dynamics could serve as a basis for a more exhaustive discussion on the physical meaning of the results in the present manuscript, and perhaps aid to design additional experiments that push the exploration of the idealized BS14 model with the inclusion of noise tipping further.

Reply: We thank the reviewer for this suggestion and will incorporate this reference. Deep convection in the Labrador Sea is also intermittent in a 12-model historical ensemble of EC-Earth (Brodeau and Koenigk, 2016), showing that at least this ESM is in principle able to simulate a recovery of convection in the SPG region. In addition, convection in the region has been observed to stop and restart in at least one CMIP6 model (NorESM-LM) (Swingedouw et al, 2021).

Unfortunately, the results from most ESM studies on the SPG region and collapse of cannot be compared directly to our results. We assume a constant background climatic state (i.e. T0 and Ubtp are constant), whereas most relevant ESM modelling studies study the response of the SPG region to a changing background state. Assuming this constant background state we conclude that convection in the SPG is stable to salinity perturbations in our current climate. It is possible that the stability of convection can change in this model as other climate parameters (e.g. T0, Ubtp) vary. This is an interesting avenue for future research.

Changes to text: We will add a discussion of ESM studies that show the intermittency of convection in the SPG to show that this result is not unique to the simple model we use, while stressing that direct comparison between this work and most ESM modelling work is not possible.

3 Minor comments

Larger figure labels would improve readability.

Reply: We thank the reviewer for pointing this out and will increase the label size for all figures.

Punctuation should be edited throughout the text.

Reply: We will critically revise the punctuation, grammar and phrasing throughout the manuscript.

ll. 69-71: Why is it worrisome? Clarifying the magnitude of the SPG effect on the AMOC here would strengthen this statement.

Reply: We thank the reviewer for this suggestion and will clarify the link between the SPG and the AMOC.

Changes to text: In l.71 we add

"In addition, there are indications that a collapse of convection in the SPG can precede an AMOC collapse (Danabosoglu et al. (2016), Drijfhout et al. (2025)."

l. 311: influence → influences

Reply: We will correct this.

ll. 325-326: Is this not just due to the form of the equation of state (β > α)?

Reply: This is an interesting point, although it is difficult to compare alpha and beta directly since their units are different and they are multiplied with variables (S and T) of different magnitude. The equation of state we used here is often used in idealized models. If the equation of state indeed causes strong dependence on noise in the salinity in our model, other studies using box models of the AMOC with noise (which often use the same equation of state) should show a similar dependence. We are not aware of similar results in such studies and have found no reason to believe the equation of state is causing the stronger sensitivity to haline perturbations.

Table A1: r as a symbol for radius of the inner box and ratio of the surface and deep box heights should be distinguished; S4 is the salinity of the deep gyre box.

Reply: We thank the reviewer for pointing this out. We will define the radius of the inner box as r_sd instead of r and correct the definition of S4.

References

Brodeau, L., & Koenigk, T. (2016). Extinction of the northern oceanic deep convection in an ensemble of climate model simulations of the 20th and 21st centuries. *Climate Dynamics*, *46*(9), 2863–2882. https://doi.org/10.1007/s00382-015-2736-5

Danabasoglu, G., Yeager, S. G., Kim, W. M., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Bleck, R., Böning, C., Bozec, A., Canuto, V. M., Cassou, C., Chassignet, E., Coward, A. C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., … Yashayaev, I. (2016). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. Ocean Modelling, 97, 65–90. https://doi.org/10.1016/j.ocemod.2015.11.007

Drijfhout, S., Angevaare, J. R., Mecking, J., van Westen, R. M., & Rahmstorf, S. (2025). Shutdown of northern Atlantic overturning after 2100 following deep mixing collapse in CMIP6 projections. *Environmental Research Letters*, *20*(9), 094062. https://doi.org/10.1088/1748-9326/adfa3b

Gelderloos, R., Straneo, F., & Katsman, C. A. (2012). Mechanisms behind the Temporary Shutdown of Deep Convection in the Labrador Sea: Lessons from the Great Salinity Anomaly Years 1968–71. *Journal of Climate*, *25*(19), 6743–6755. https://doi.org/10.1175/JCLI-D-11-00549.1

Swingedouw, D., Bily, A., Esquerdo, C., Borchert, L. F., Sgubin, G., Mignot, J., & Menary, M. (2021). On the risk of abrupt changes in the North Atlantic subpolar gyre in CMIP6 models. *Annals of the New York Academy of Sciences*, *1504*(1), 187–201. https://doi.org/10.1111/nyas.14659

Yashayaev, I. (2024). Intensification and shutdown of deep convection in the Labrador Sea were caused by changes in atmospheric and freshwater dynamics. *Communications Earth & Environment*, *5*(1), 156. https://doi.org/10.1038/s43247-024-01296-9