**Response to Reviewer #2**
The review is pasted below; our responses are interspersed in blue font.

General comments

In the manuscript by Garcia-Suarez and collaborators, the authors test the hypothesis of whether isotherm base criteria for the location of the Gulf Stream are appropriate in the context of a warming ocean. To do this they use two simulations, using two different model architectures, forced by two different scenarios. Their results suggest that widespread warming in the slope associated with the climate change signal can be erroneously interpreted as a northward shift in the current axis. The manuscript is well written and the language is clear. The overall framing of what current and future changes in the world's western boundary current systems tell us about the overturning circulation and continental shelf environment is a very relevant one and it would be of interest to a large part of the community. However, I'm afraid that the study lacks both the novelty and rigour that are expected for a peer reviewed scientific publication. The question that is posed by this study, presented as their hypothesis in lines 50-51, is already addressed in Todd and Ren 2023, which the authors already cite in the manuscript. Todd and Ren clearly articulate how to decompose temperature signals to isolate warming from axis shift based in the observations. A similar approach could have been taken here to isolate warming from shift. With regards to experimental design, while I value the authors' effort to demonstrate that their results hold across multiple models, the different model architecture and different forcings lead to very different GS behaviours, and the manuscript presents no evidence that the either model is able to reproduce variability. Unfortunately, I believe the work required to bring this manuscript into acceptable form is beyond the scope of the review process, and as such I have to recommend against publication.

**Response:** We appreciate the reviewer's comments that our manuscript is well written and clear and that the topic is a "very relevant one and it would be of interest to a large part of the community."

We are surprised by the comment that the study lacks the novelty and rigour of a peer-reviewed study. Regarding novelty, the reviewer seems to suggest that our hypothesis is already exhaustively addressed by Todd and Ren (2023). We respectfully disagree. As stated by the reviewer, our hypothesis is given on lines 50-51 as the "15 ºC isotherm [..] criterion overestimates GS northward shifts." Todd and Ren (2023) did not address this hypothesis at all, in fact, they don't use the 15 ºC isotherm criterion. Todd and Ren (2023) used in situ observations and provided an excellent and detailed analysis of Gulf Stream temperature, salinity, and density gradients, the warming and changes in stratification occurring in the Gulf Stream, and an estimate of the rate at which it moved laterally from 2001 (they find a northward movement of 6 +/- 3 km per decade). Todd and Ren (2023) do not address the question of whether the isotherm criterion is appropriate to track lateral shifts in Gulf Stream location in a warming ocean and they don't have at their disposal multi-decadal model projections, which we analyze here.

Regarding the reviewer's criticism that we do present no evidence that either model is able to reproduce variability, we would like to note that both model runs have been extensively discussed and used in the published literature, including in studies that make inferences about lateral shifts in the location of the Gulf Stream (Saba et al. 2016, Ceasar et al. 2021) or discuss mesoscale variability in the region (Martin et al. 2022, Martin and Biastoch 2023, Huo et al. 2024). More specifically, the same CM2.6 simulation we are using in our study is discussed and used by Saba et al. (2016 [cited 600 times]), Ceasar et al. (2021 [cited 460 times]), Claret et al. (2018 [cited 117]). The FOCI model we are using is more recent and not yet as frequently cited as CM2.6, but has been validated (Matthes et al. 2020 [cited 26 times]) and used in a number of follow-up studies including Martin and Biastoch (2022), Martin et al. (2023), and Huo et al. (2024). All the studies referring to CM2.6 are already cited in our manuscript. Regarding FOCI, we have only cited Matthes et al, (2020) thus far but can add the other studies in our revised manuscript.

Caesar, L., McCarthy, G. D., Thornalley, D. J. R., Cahill, N., and Rahmstorf, S.: Current Atlantic Meridional Overturning Circulation weakest in last millennium, Nat. Geosci., 14, 118–120, https://doi.org/10.1038/s41561-021-00699-z, 2021.

Claret, M., Galbraith, E. D., Palter, J. B., Bianchi, D., Fennel, K., Gilbert, D., and Dunne, J. P.: Rapid coastal deoxygenation due to ocean circulation shift in the northwest Atlantic, Nature Clim Change, 8, 868–872, https://doi.org/10.1038/s41558-018-0263-1, 2018.

Huo, W., Drews, A., Martin, T., & Wahl, S. (2024). Impacts of North Atlantic model biases on natural decadal climate variability. Journal of Geophysical Research: Atmospheres, 129, e2023JD039778. https://doi.org/10.1029/2023JD039778

Martin, T. and Biastoch, A.: On the ocean's response to enhanced Greenland runoff in model experiments: relevance of mesoscale dynamics and atmospheric coupling, Ocean Sci., 19, 141–167, https://doi.org/10.5194/os-19-141-2023, 2023.

Martin, T., Biastoch, A., Lohmann, G., Mikolajewicz, U., & Wang, X. (2022). On timescales and reversibility of the ocean's response to enhanced Greenland Ice Sheet melting in comprehensive climate models. Geophysical Research Letters, 49, e2021GL097114. https://doi.org/10.1029/2021GL097114

Matthes, K., Biastoch, A., Wahl, S., Harlaß, J., Martin, T., Brücher, T., Drews, A., Ehlert, D., Getzlaff, K., Krüger, F., Rath, W., Scheinert, M., Schwarzkopf, F. U., Bayr, T., Schmidt, H., and Park, W.: The Flexible Ocean and Climate Infrastructure version 1 (FOCI1): mean state and variability, Geosci. Model Dev., 13, 2533–2568, https://doi.org/10.5194/gmd-13-2533-2020, 2020.

Saba, V. S., Griffies, S. M., Anderson, W. G., Winton, M., Alexander, M. A., Delworth, T. L., Hare, J. A., Harrison, M. J., Rosati, A., Vecchi, G. A., and Zhang, R.: Enhanced warming of the Northwest Atlantic Ocean under climate change, J. Geophys. Res. Oceans, 121, 118–132, https://doi.org/10.1002/2015JC011346, 2016.

Specific comments

(1) The paper goes back and forth between the concept of the GS as a driver of the changes in the Slope (both introduction and discussion) and changes in the slope explaining the warming in the northern flank of the GS (the explanation for the mismatch between velocity and thermal proxies of the GS). I think there's a real opportunity for the authors to address this question, but a more systematic approach to decomposing changes in temperature and circulation in the western North Atlantic would be needed.

**Response:** We believe our approach is systematic and would be curious to know why the reviewers suggests it is not.

As a starting point considering the actual spatial patterns in warming may already reveal some interesting information. At the moment because the Slope temperature is calculated as an average over the region relative to the mean GS position in 2015, and figure 4 shows that at least in CM2.6 there is a non negligible trend in the actual GS path (both blue and black lines in panels c and d confirm, and the magnitude of the temperature trends over the entire period in FOCI are much less clear due to the nature of the forcing post 2040 I suspect), some of the warming attributed to the Slope is likely a real shift in the GS axis.

**Response:** The spatial patterns of warming are discussed in detail in Garcia-Suarez et al. (2024). This manuscript is focussed on the connection (or, more precisely, the lack thereof) between Gulf Stream location and warming in the northwest Nort Atlantic shelf and slope regions.

Garcia-Suarez, L. and Fennel, K.: Physical Drivers and Biogeochemical Effects of the Projected Decline of the Shelfbreak Jet in the Northwest North Atlantic Ocean, Journal of Advances in Modeling Earth Systems, 16, e2024MS004580, https://doi.org/10.1029/2024MS004580, 2024.

In the discussion of the warming in the slope as an explanation for the apparent shift in the GS inferred by T15, the authors show evidence that the slope is warming at an accelerated rate compared to the subtropical ocean (panels c/d versus e/f). If this is the case, the thermal gradient across the GS must also be increasing. For the historical period, because a big portion of the observed temperature changes are salinity compensated, as discussed in Todd and Ren (2023), they have no impact in the current structure. Using the heave/spice framework as presented could help distinguish which part of the warming is a driver and which is an impact. This framework along with the vertical structure of the temperature anomalies shown in Figure 5 offers some likely candidates for what are the actual drivers of the warming in the continental slope and shelf (lines 156-161) in these models.

**Response:** The reviewer seems to suggest our paper should address the drivers of the warming in the northwest North Atlantic. While we agree that it would be a worthwhile and interesting study, this is not the intention with the present manuscript. As stated in our response to reviewer 1, the motivation for our study is that the rapid warming of the northwest

North Atlantic shelf and adjacent slope regions, which are among the fastest warming ocean regions in the world, is commonly attributed to a northward shift in the Gulf Stream (see references on line 25). This northward shift is commonly diagnosed using a temperature-based criterion which, as we show in the manuscript, is problematic. We believe we provide strong evidence that the warming of the shelf and slope region can occur without a northward shift of the Gulf Stream (see FOCI model projections) and that isotherm-based criteria are misleading in that they suggest a northward shift even when no such shift is occurring.

(2) Amongst the key differences between models that make it hard to assess whether these models capture the real behaviour of the Western North Atlantic are the following.

Firstly, in the assessment of the mean GS path (Figure 1) in CM2-6 the SSH and maximum velocity criteria show important differences (with the sign of the difference varying across longitudes), this is compared to both observations and FOCI for which the two criteria tend to agree to the west of the Grand Banks. Secondly in the discussion about the Slope supported by figure 5, in CM2.6 the slope continuously cools down for most of the control run, while it oscillates around 0 in FOCI. The magnitude of the warming for different layers is also quite different between models, with both the 200 and 450m levels warming faster than the surface in FOCI, but only the 450m layer warming faster than the surface in CM2.6 according to Figure 5.

**Response:** We agree with the reviewer that the models have shortcomings, but would like to point out again that they are widely used in the literature, including in the widely cited study by Ceasar et al. (2021) that suggests a causal link between the warming in shelf and slope region and a northward movement of the Gulf Stream (which we believe Ceasar et al misdiagnosed by using the isotherm criterion).

(3) Lastly, as I final note, the North Wall is not expected to match the maximum velocity core (as stated in lines 145-146), but instead represents the northern flank of the current. As such it is expected to be located to the north of the GS core even in equilibrium. The question is whether the two metrics vary in a coherent way. Figure 4 suggests that within the eastern part of the domain, where the Gulf Stream is more coherent (A1 domain according to the labelling), a deeper isotherm could be used to track the evolution of the current core quite accurately in both models. A shallower isotherm, such as T15 does overestimate the evolution based on a simple linear fit. However, in close inspection, for FOCI (A1), most of the difference between T15 and T12 occurs at the beginning of the record, prior to 2024, while they approximately track each other from that point onwards. As I said, I think this work presents some very interesting preliminary results that the authors could build on in a future study.

**Response:** We agree with the reviewer that the North Wall is expected to be north of the Gulf Stream core. However, the point is that the North Wall index suggests a strong northward movement, which has been proposed as a driver of the rapid warming in the northwest North Atlantic, when the core may not actually move northward at all (as is the case in the FOCI simulation) or only minimally (as in the CM2.6 simulation.) In other words, the two metrics **do not** move in a coherent way which we clearly show.

The reviewer states that our study “presents some very interesting pre-liminary results that the authors [we] could build on in a future study” without providing any clear arguments as to why they perceive our reasoning as flawed.

We believe we have clarified the motivation of our study, have explained that spatial patterns of warming are analyzed elsewhere and are not within the intended scope of this study, and that the models have been widely used already in the published literature.