Responses to Editor (Julian Mak)

We would like to thank the Editor for his constructive comments and for the numerous suggestions to help improve the manuscript. Our responses are given below in black type (Editor's comments in blue italics).

Public justification (visible to the public if the article is accepted and published):
It's a minor decision in that I will just go through it, but it's "major" in content (although I think it's all fixable). I think this work is useful and should eventually be published, but I am strongly of the opinion (which is consistent with a sample of the referees upon separate conversations) that the current presentation has issues (that's putting it mildly). Below are the things I can see or can be bothered to write down, and there will be more; **line numbers refer to the ones in the tracked file**.

* Common theme: The article writing noodles around and tangles itself up unnecessarily. There is extra information that is distracting, and the framing of the message (at the section and paragraph level) is pretty chaotic and/or non-existent at times. The main issue I have is that often the ordering of the messages are back-to-front (justification goes before the message), so then one has to go through a chunk of (distracting) things without necessarily seeing where the path goes or the reason for the content. It is particularly annoying because often the solution is a simple copy-paste or rephrasing. Some suggestions for fixes below, but the solution is likely non-unique.

* Motivation: I have a big issue with the "3d" motivation of this work (probably in line with two of the referees?). The authors make a big deal about how "3d" is necessary and there are compound effect relating to horizontal advection and transport etc. Sure the model evaluation here gives provides the "global coverage" (which is valuable by itself), but the authors never quantify the "advection/transport" aspect explicitly in any of their diagnostics. Yes there are there implicitly in the analysis in Section 3 and 4, but there is a distinct lack of explicit evidence in quantify how essential the "3d"ness is, which is particularly imbalanced since the authors try and make such a big deal about that it. Two ways I see this how one could get out of this hole, one being substantially easier than the other:

- Dig out of the hole but investigate changes in resulting circulation if any, diagonse and analyse the relevant horizontal/vertical divergence/diffusion trends relating to advection
- Don't dig into the hole in the first place but not making a big deal about "3d", but talk about "global coverage" instead

My personal opinion (and I think of the two referees also, although I hope I am not putting words in their mouths) is that this work seems to be trying very hard to be something that it isn't, and it really isn't working, but there is simply no need for that since the results are valuable as is.

* line 43: "...even interannual TIMESCALES" (because subject here could be "forcing", "time-scale", "events", and all of them would actually make sense)

Correction added line 43

* line 65: Could do "...these processes ACTING PRIMARILY IN THE VERTICAL DIRECTION occur at subgrid scales, and therefore..." (possible "...subgrid VERTICAL scales...")
Correction added line 65

* line 71: "independently" -> "independent"
Correction added line 71

* line 72: I know what it is that is trying to be said but it is probably bad investment in works, so I think just remove "that propagates counter-gradient of the stratification", because it's distracting without adding anything that essential that hasn't already been said.
Correction added line 72

* line 82 to end of sentence: Redundant text with what is at the end of paragraph, probably remove "motivating the need...convective turbulence."
Correction added line 82

* line 84: If "Fer (2014)" is cited later why is this not cited here when taking about "near-inertial wave breaking"? Suggesting fixes are not comprehensive and patchy.
Correction added line 83

* line 130: "quantities" is ambiguous because multiple things can be refereed to as "quantities" here, and "variables" is used above anyway, so do "...in representing the AFOREMENTIONED VARIABLES" (don't introduce new namings for no particular good reason)
Correction added line 128

* line 150: "5" -> "five"
Correction added line 148

* line 151: "z" should be in math font, "...WITH LOCAL thickness STARTING AT 1m at the surface..., and INCREASING up to 200 m near the bottom OF THE MODEL OCEAN."
Correction added line 149-150

* line 154: Scientific/modelling point here, and actually the authors raise a similar point in line 546 also (so shooting themselves in the foot as well...) Eddy-permitting requires something different done with GM (turning down the coefficient, resolution functions, what I proposed in Mak et al 2023), so what is done here? That itself leads to biases (as the authors say themselves), so at least give some details of what is done here, and/or some arguments on why the changes in modelling details and biases come from the schemes of interest here and not from some compound effect from the other things like GM choices.
There will be compound/nonlinear things probably, so instead of digging a hole that needs to be gotten out of, would suggest just state clearly what the choice of GM was made here, but this is applied consistently across all simulations, and acknowledge there can be compound effects that are beyond the scope of present investigation.

The default NEMO parameters for the GM parameterization are used in all three simulations, following the formulation of Treguier et al. (1997), with a lateral diffusive velocity of 0.01 m s^{-1}. No specific adjustments to GM are made for the eddy-permitting configuration considered here.

Although the GM configuration is identical across all three simulations, we agree that changes in the density field induced by the choice of vertical mixing scheme can, in turn, affect the response of GM, potentially leading to further changes in the density field. A detailed investigation of these compound effects is indeed beyond the scope of the present study.

We thank the editor for this suggestion, this is now clarified line 151-156.

(Some of the words in line 817 to 823 should really be up here, contributing to my point that I feel a lot of the writing is back-to-front in terms of the ordering.)
Yes, this part was removed.

* line 160, 268, 270 and elsewhere: Either create a reference entry for the doi and use the reference entry, or the dois go into the "data availability" section with appropriate text (e.g. "The GREP at *1/4°* resolution can be accessed at [doi].")
Correction added lines 267-269, 841-845.

* line 172: Don't need the comma, "...ad hoc empirically-based..."
Correction added line 172

* line 181: "Effective" unnecessary, "...results in LARGE mixing..."
Correction added line 181

* line 191 and probably elsewhere: hyphenation inconsistent, "non-local" vs. "nonlocal", choose one and be consistent
Correction added line throughout the text. We did the same for “ad hoc” and “ad-hoc”.

* line 202: So this section really did irritate me and I think it boils down to an implicit assumption the authors are making that people who read this know the scheme as well as they do (which will generically not be the case). It is only on several reads the picture becomes clear(er) to me, and it probably boils to framing the parameterisation in a much more explicit way. Some description of the scheme and then something as simple as describing what the prognostic variables are EDMF will help substantially (at the moment this is like one of those cooking books that tell you about the cooking method but you the reader have to figure out the ingredients.) My reading is that the plume being parameterised leading to mass flux needs

w_p, a_p, T_p and S_p (and something else? not clear and not sure)

so why not say that...? Once that is said then the equations and model parameters would make sense, and the overall flow probably works better as a result.
We thank the reviewer for this suggestion which helps clarifying this section. Corrections added lines 199-202.

* line 209-212: Doesn't really answer the query raised by one of the referees, still seems entirely abrupt and not well justified, and for me the ordering also doesn't work. Consider stating what is known and supported first ("The coefficient \alpha blah blah, ... and it was found that \alpha=0.2 produced the best results in the EDMF 1d test cases (+ appropriate references).") Then say something like "However, because of [whatever reason from diagnoses related to some empirical metric], in the present 3d case we employ the value \alpha=1.", but provide at least some reason for that choice given

the discrepancy (it might as well be "it was revealed to me in a dream" at present). The ordering avoids the noodling, and some reason is needed to make the choice vaguely scientific.

Correction added lines 211-214.

line 214: "A" plume, not "the" plume (Is there plume variables for every grid cell grid cell? Not clear)
Correction added lines 217

line 224 and 240: If some parameters are re-tuned why are these not re-tuned or don't need re-tuning? Again this reads imbalanced and suggests arbitrarily-ness, when presumably there was a conscious decision made...? (Or maybe not, but either way it shouldn't read like that...)
Correction added line 214-215.

line 228: Either "plume properties" or "plumes' properties" (some grammar thing associated with the possessive)
Correction added line 234

line 228: Scientific question here. Is anything done to resolved (horizontal and/or vertical) momentum from this parameterisation? If there is a mass flux then by definition there will be a momentum flux, some represented in w_p (cf. EVD in NEMO has an option to do that although I think it's off by default). If nothing is done, say so and justify why (if that's the case I would defend/justify that choice when the prognostic variables are first stated back in the new line 209-212 bit, so the question is shut down before it has a chance to be queried). If something is done, then say what is done, and the next point maybe less relevant.
Correction added lines 202-203

line 229 and Eq 4: Don't like the \psi_p because to me it's entirely unnecessary, and it is never used again. Could just write the two equations equations for T_p and S_p, and maybe have "\epsilon_\psi" -> "\tilde{\epsilon} = \epsilon_T = \epsilon_S" or similar in Eq 5.
\psi_p is now defined line 201.

line 244-247: Out of order, the "half step" is used before the "leap frog" piece of information. Move the last two sentences up, probably to between "...same time step." and "In this study..."
Correction added lines 249-254.

line 251: Indicate end of list by "...parameterization); AND EDMF is the new..."
Correction added lines 260

line 254: The reply to the referee about "online calibration" to me doesn't make sense, and I am not convinced the "online" is relevant or the right adjective here (I honestly don't know). "Online" to me would imply the parameters are adjusted while the model is run (i.e. "online", "on the fly"), but this is not what happens here. I don't really know what the authors are trying to say here, but an easy fix is just removing the word "online"?
(Are you are talking about "history matching" like what people increasingly use for tuning climate models, i.e. tune several parameters at a time given a target, instead of tune one parameter after the other for the same given target? That would be "calibration" using "online runs", but there is nothing online about the "calibration" itself.)

We thank the reviewer for this suggestion and acknowledge the confusion, correction added line 255

table 1, line 256, experiment naming: EDMF also has TKE so why isn't that tagged on there also...? The cleaner way is to remove TKE is from everything and say it is always on.

EDMF already includes TKE in its naming through the “ED” part, this also mentioned line 191-192.

* text from Table 1 to Section 2.3: Block of text is arguably out of place, and probably needs more text to introduce why it is here. I assume this part exists because the authors want to talk about contributions to temperature fluxes later (Section 3.2 where it is first used), so why not introduce it there when you use it? Because if not the two criticisms I have are (1) why not salinity, and (2) there is a sizable gap between where things are defined and where things are used when there does not need to be. It also needs a bit more wording to motivate this bit, e.g. (if this was in Section 3.2) "To analyze the impact of the vertical schemes to the oceanic properties, we choose to compute the temperature trend..."
We thank the reviewer for this suggestion, this part was moved to Section 3.1.1, lines 364-370

The writing of the relevant text could be made more cleaner, less ambiguous and more efficient also (e.g. "In the case of EDMF simulations, THE TOTAL TREND FOR VERTICAL HEAT FLUXES IS DEFINED TO BE [blah], where as in the case of the EVG simulation it is [blah]."
Correction added line 369-370.

* rendering of d/dz here and many other places elsewhere: The one in Eq (7) looks fine but the in-line text ones are all over the place, and "\partial z" shows up more times than "\partial_z" (with the underscore and so subscript); I assume the former one is not the one intended (given line after Eq 7 and the entry in Fig 2 has some of it used correctly).
Correction added to the text lines 372 to 376, and to Fig. 2.

* definition of MF: Simply unnecessary, just write "a_p w_p" in Eq 7, expand "MF" as "mass flux" and adjust Fig 2 and other figures as necessary (I would then argue why not use "TF" instead of "heat flux"...)
Correction added to Eq. 7.

* "line 271 and elsewhere: "in situ" -> "in-situ" probably
Correction added throughout the text.

* line 282-283: Text suggesting for example that Argo data goes back to Jan 1970, which is ambiguous. One fix is the ordering (again), e.g. "...collected from the Argo program... profilers FOR THE APPROPRIATE TIME PERIODS JAN 1970 AND DEC 2021 WHERE DATA IS AVAILABLE." or similar.
Correction added lines 283-285

* line 284: "OMIP" itself does not do the "recommending", it is presumably "used" in "OMIP" for validation. Consider "...ocean models, SUCH AS THOSE PARTAKING IN OMIP..."
Correction added lines 286

* line 284: "accurate" is contradictory because by averaging accuracy has been lost (gaining "robustness" or similar maybe). "...for AN APPROPRIATE comparison" or similar, or just get rid of the adjective (don't need it).
Correction added lines 282

* paragraph starting line 290: Confusing because this suddenly jumps to tropical ocean and lacks a clear motivation starting the paragraph. Again a fix for this is the ordering, something like "For evaluating the skill of the scheme for shallow diurnal convection that is particularly prevalent in the tropics, we compare model output against observational data in the tropics, where in-situ observations from the standard operational TAO surface mooring etc. etc.... For a like-for-like comparison, the observational data were temporally averaged to have the same temporal frequency" or something similar.
Correction added lines 289-292.

* line 304, 305 and elsewhere: I am going to be more militant than the referee and I think the choice to use "mooring" is unnecessarily confusing, because it is really the data from a numerical grid cell. Why not just use "grid cell", particularly since there is no comparison in the Labrador sea with an actual mooring, and it's not like there is some up-sampling interpolation going on to talk about (numerical) data at a certain geographic location in longitude and latitude? Then this also helps the text everywhere else: when "grid cell" is used it lets the reader know the quantity is from the numerical model, while when "mooring" is used then it really means observation data. At the moment it's a mix and distracting.
We would like to clarify the rationale behind our original use of the term “mooring” for the two sites selected to represent the different convective regimes, as this terminology led to confusion. NEMO provides a feature commonly referred to as a “numerical mooring” or “virtual mooring”, which allows high-frequency time series to be extracted at specific geographical locations for direct comparison with observations from instrumented ocean moorings. For example, we extract hourly model output at 2°S, 140°W to compare it directly with observations from the TAO mooring located at the same geographical position.
We acknowledge that this terminology was potentially misleading in our manuscript, particularly for the Labrador Sea site, where no direct comparison with observations from an actual mooring is performed. We have therefore revised the terminology throughout Sections 3.1.1 and 3.1.2, replacing “mooring” with “grid cell” or similar when referring to model output.

* line 310: The use of "accuracy" to me is implying EDMF is already accurate, when no evidence has in fact been presented. Could also argue something being "accurate" but "not robust" is of limited use, so "accuracy" is more restrictive than it needs to be me. "Skill" or "performance" are alternatives I suppose.
Correction added line 309.

* line 316: Expand "MF" and elsewhere.
Correction added throughout the text.

* line 317: "...WHERE positive..."
Correction added line 317.

* line 318 (probably): No introduction of Fig. 1, add "The results are shown in Fig. 1" at the end of that paragraph.
Correction added line 318.

* line 319: "ThE RESULTS IN FIG. 1 INDICATE that the MASS FLUX component is..."
Correction added line 319.

* line 320: No "s" on "characteristic"
Correction added line 320.

* line 326: The "known as the few sites of the ocean" is not adding anything, probably remove.
Correction added line 326

* Fig 1: We strongly push for not using the "jet/rainbow" colour scheme (I think it's ugly, but the official reason is for colour perception and not accessibility friendly). And also this is a "yes" or "no" kind of graph so shouldn't be using a "jet/rainbow" scheme anyway, maybe something like "heat", "warm", or a one-sided "red blue", just not "jet". Also expand "MF".
Figure 1 was modified accordingly.

* line 334 to 340: Again, ordering. First don't use "numerical mooring" for arguments above. Second the ordering is back to front, consider e.g. "We are interested in a similar analysis for the deep convection, and we focus on the Labrador Sea region. No continuous observational time series are directly available... We choose the grid cell containing the location (whatever) as a representative point, analysing the model outputs also at hourly temporal resolution." or similar.
Correction added lines 334-336.

* line 349, within this paragraph, and elsewhere: line 349 and 350 is a long dash while connecting the years, while 354 is a short dash or hyphen, and this is inconsistent, not only here, but many other places. Choose one and stick with it; probably do the short dash or hyphen (the long dash might want to be reserved for minus signs and e.g. "Clausius--Clapeyron").
Correction added throughout the text.

* line 359-360: Reads like a bad sentence to me and I am not sure why, probably because a positive and negative clause are forced together with no smoothing between it. Try "...w_p, ALTHOUGH such OBSERVATIONAL measurements... difficult IN samplING small..."
Correction added lines 357-358.

* line 366: Don't use "mooring site".
Correction added.

* paragraph of lie 366: Wrong/inconsistent "\partial z" ("\partial_z" is wanted surely)
Correction added lines 364-374.

* line 381: The subclause is tagged to the wrong part, "...2015 (selected as it falls... Fig 2a), are shown in Fig 2b-f".
Correction added line 376

* line 383: The subpanels are referring to the wrong quantities, e.g. T_p < T_env is shown in Fig 2f not Fig 2e, and everywhere else in this paragraph. (There may be others in the article.)
Correction added throughout the paragraph.

* line 385: Confusing to me, just do "...they accelerate AND reach vertical..."
Correction added line 383

* line 389: Tenses, "...the plumes lose their gainED buoyancy..."
Correction added line 387

* Fig. 2: where to start...

- what are the vertical gray lines and what is their purpose?
- middle of Fig 2b is basically a waste of space, could remove it and just say it in caption, do a zoom in, have split panels to show near surface and deep ocean activity at the same time, or analogous
- don't need any of the RHS of the equations really, since this is already in an equation (or remove the equations if this is explicitly used here in the panel titles anyway)
- this is me being pedantic, but to me it's weird to have "K" and "degC" at the same time. At least here "K" is used consistently when the heat flux quantities are being talked about, but not a big fan of this choice even if I know NEMO outputs it with the units involving "K"
- don't use "mooring location" in caption
- "\partial z" -> "\partial_z"
- caption entry for (e) should be a minus sign not a hyphen/short-dash here
Corrections added to Fig. 2 and caption.

* line 411: This one I don't know, is this an actual mooring or a numerical grid cell?
Here the model outputs are from the grid cell located at 2°S, 140°W and are directly compared to the observational dataset from the TAO mooring at the same location. This is clarified in the text lines 408-410, 430-431.

* line 418: Tenses, because there is "peaking". "As the night progressES, the plume deepenS and intensifieS..., the plume weakenS and eventually stopS at noon."
Correction added lines 416- 418.

* line 423: Expand MF
Correction added throughout the text.

* line 425: See "grid cell" is used here, so what is the objection of using "numerical grid cell" above!?
Correction added throughout this section.

* fig 3: Don't actually know if this is mooring or not. Also minus sign formatting in entry (d).
Figure 3 presents the model outputs at 2°S, 140°W. This is clarified in the caption, and correction was added to Fig. 3.

* line 436: Tenses, "...measures ARE available"
Correction added line 432

* line 446: "best agreement" but in what measure? Some of it seems to be said at the end of the paragraph, so then there is a re-ordering that makes the flow more logical.
Correction added line 442.

* line 457: "...tides that the simulations DO NOT REPRESENT.", to make explicit the subject is "simulations" and not "observations".
Correction added line 453

* Fig 4 caption: There are no green squares here. The "--" in "b--d" should be a short dash to be consistent with elsewhere.
Correction added to Fig. 4 caption.

* line 478: Choice of number of decimal places extremely inconsistent. Looks like it is 3 decimal places wanted, so use that everywhere (I mean everywhere, e.g. line 488, 599). Or change it to 2 decimal places, just be consistent (e.g. line 599 is inconsistent with both of those choices).

The numerical precision has been harmonized in Figs. 5 and 8, using three decimal places. In the main text, however, values introduced by approximate expressions such as “below”, “above” or “~” are generally rounded to one significant figure, as reporting several decimal places would imply an unjustified level of precision (e.g. “~1 K” rather than “~1.00 K”).

* Fig 5 caption: "0-700" or "0 to 700", choose one
Correction added to Fig. 5 caption.

* line 515: "Southern Ocean" with the capitalisation probably
Correction added line 511

* Fig 6 caption: as two points ago, probably short dash
Correction added to Fig. 6 caption.

* line 530 to 536: formatting of dashes and number of decimal places for temperature
Correction added lines 523 to 530

* line 544: This part is confusing for me, since this is an ocean-only model forced with a product (right?) and but "climate model" and "low-level clouds" are both mentioned. What is meant here is surely the "the warm biases are consistent with known biases in the forcing products" or something similar?
Correction added lines 536-537

* Fig 7: Similar comment to Fig 1 about jet. Also the choice of RMSD is confusing because 1 degC bias in the tropics is not the same as 1 degC bias in the polar regions, so if anything some normalisation needs to be done here.

The colormaps in Fig. 7, and Figs. S4, S5 and S6 have been modified accordingly.
We acknowledge the editor’s point that the same temperature error may have different implications in tropical and polar regions. However, RMSD fields are widely used in ocean model validation studies (e.g. Lellouche et al., 2013 https://doi.org/10.5194/os-9-57-2013; Liu et al., 2023 10.1088/1742-6596/2486/1/01202; Qin et al., 2023 https://doi.org/10.3390/jmse11122255; Fu et al., 2023

https://doi.org/10.3389/fmars.2023.1084186) to quantify the magnitude of model errors in physical units. We have therefore retained the unnormalised RMSD, which allows the absolute magnitude of temperature errors to be assessed directly. Normalisation would provide a relative measure of error, but its interpretation can become problematic where the reference field is close to zero.

* line 591, Fig 8 caption: Figures have "Northern" and "Southern" with the capitalisation, so inconsistent formatting (e.g. with Fig 8).
Correction added to Fig. 8.

* Fig 9 caption: Capitalisaiton on "Same" since it is a beginning of a sentence of sorts.
Correction added line 643

* line 656: Why not compare to TKE+MLP+EVD? Imbalance, choice not defended, in-line with referee comment(?)
Correction added line 646-650.

* Section 4.1: again, where to start...this whole section to me is a mess notation-wise and logic-wise, particularly since the key point that is truing to be make is really just that unlabelled equation above Eq. (13). Would suggest reading all the sub bullet-points to the end first.

- the writing implies here that "heat" is in units of "K", which is not true, and "E(0, t)" is not an energy/heat (the unit of "K" is correct), but that is not an extensive quantity then so seems a bit weird trying to compare it...it is also inconsistent with later when it is the "ocean mean temperature" that is being talked about
- why not just use A_T(x,y,z=0) = e1t(x,y) * e2t(x,y) * tmask(x,y,z=0) which is what you really want to write?
- or even more drastically (but cleaner and more general in some sense), why not just use the continuous version, which is (probably)

E(z=0,t) = \int Q_net A / (V \rho_0 C_p) dt

in Eq 8 and 12? Then you don't even need to define the metric and mask factors.

- I am seeing the in-line parts render as "Qnet", "SWnet" etc. when it should be "Q_net", "SW_net" etc.
- probably be explicit and say "downward" heat is "energy gain in the ocean" or similar
- writing, "THE downward longwave... simulations, WHILE the ..."
- I fail to see the point of introducing E(0, t) at all given Eq 11
- I think I see what is going on maybe. I am guessing "E" was introduced as a placeholder for the entries in Eq 13, but then why not define a "Q_net(0, t)" but use "E" there? Then all those "SW_net(0,t)" etc. are undefined, contributing to notation mess
- the capital Deltas (the triangles) are formatted inconsistently in Eq (13)
- why not start the whole section with

"At the global scale, assuming negligible contributions... [unlabelled equation about Eq (13)] and the components of SURF are expressed as [Eq (13) formatted properly], where the contributions are [longwave, shortwave blah blah]. This leads to changes in the simulated global mean temperature defined as [not using E because it's confusing]

f(z=0,t) = \int f A / (V \rho_0 C_p) dt

[then define how downward things are given by ERA5, upward is diagnosed, others from bulk formulae etc. etc.]"

(Honestly I can't visualise the end result of so many changes in one go, and I don't feel this is really my responsibility given this is not my paper nor am I the orginator of this, so give it a try in tidying this part up and I'll comment on the end result maybe...)

We thank the editor for those very valuable suggestions, we have reorganized Section 4.1 accordingly.

* Fig 10:

- it's weird that "K" is used here when "degC" is used elsewhere given it is a temperature that is being talked about
- the italic "W m-2" notation I guess is ok although I don't like it (because it is not defined in the caption and is relying on the reader to put that implicit link in).
- the minus signs in the bottom left box is formatted inconsistently (first line between the two brackets.
- "Tair" should really be "T_air"
- - math font for caption *y* axis
- why only "\Delta LW up" being talked here but not the other one? Imbalance and sticks out relative to the others. (I am guessing this is actually a mistake and it is "\Delta LW" without the arrow that is meant here and in the rest of the section [e.g. in the bottom left hand box of figure], but honestly given the abundance of inconsistent formatting I don't even know anymore.)

We thank the editor for these suggestions. All temperature units were homogeneized into K, throughout the text and figures. Fig. 10 and caption were updated according to these recommendations. Regarding the last point: \Delta LW_downward is provided by ERA5, therefore values are common between the two simulations, that's why we only provide \Delta LW_upward which are diagnosed for each simulation, lines 688-690.

* line 737, 745: Capitalisation of "southern" and "northern" (cf. line 747).
Correction added lines 740-757

* line 740, 741: Should be dash not hyphen in "Clausius--Clapeyron" (the second one is using a hyphen).
Correction added lines 728

* line 758 to 760: The usage of those brackets is jumpy, particularly since "vice-versa" (used with the dash) is used in line 762. Mirror the formatting accordingly.
Correction added lines 745-749

* Paragraph of line 796: Again, ordering back-to-front surely. "The global 3d configuration..." part should go first since that is the raison d'etre for this study according to the authors, followed by the "Based on a 16-year simulation..." etc. which is the doing and findings part.

We thank the reviewer for this suggestion, correction added lines 783-787

* Same paragraph: The "which uses a unique formulation" is unjustified and redundant probably, because the scheme was never elaborated on that much, and the sentence makes sense and is fine without that clause.
Correction added lines 787

* line 817 to 823: Some/all of this text should have been in the methods section, but is not for reasons I don't understand.
This part was removed.

* line 825 to 829: This part is about the "3d"ness, so why is it here and not with two paragraphs before when that the "3d"ness is being talked about?

We would prefer to keep this part (now corresponding to lines 808-829) in its current position, as we believe it follows the logical progression of the Conclusion. In section 5, we first briefly recall the context, objectives and methodology of the study. We then summarize the main results, starting with the two site-specific case studies presented in Section 3.1, followed by the global temperature biases discussed in Section 3.2, and finally the results concerning surface heat fluxes presented in Section 4, which is the paragraph mentioned in this remark.

* line 840 to 843: Yeah no, I don't think I have ever had so many issues with one single sentence in a scientific article.

- GEOMETRIC at least as is shipped with NEMO deals with GM, which is not designed for "sub-mesoscales" (e.g. Mak et al, 2018, 2022)
- "Torres et al (2025)" is not GEOMETRIC as such because the energy budget used is different to what is proposed and referred to as GEOMETRIC (e.g. Mak et al, 2022)
- "Torres et al (2025)" to my mind makes a big focus on "isoneutral mixing", which is not to do with "restratification" (source: I am the last author on that paper)
- the GEOMETRIC reference for NEMO as intended should really be Mak et al. (2022) (source: me and Gurvan wrote GEOMETRIC into NEMO)
- in eddy-permitting models like the one considered here there are fundamental issues with using GM-class of parameterisations (e.g. Mak et al 2023), so further text should be added in accordingly if this sentence is to be kept
- honestly, I would just remove the sentence because it just feels like a terrible investment in words for basically no gain
We apologize for the confusion caused by this sentence and thank the Editor for the clarifications and helpful explanations. As recommended, we have removed the sentence from the manuscript.

* line 848: This part seems to be further evidence to suggest the "online" in "online calibration" in fact has nothing to do with "online"...
We thank the reviewer for pointing out this potential ambiguity. The sentence referring to High-Tune does not describe an online calibration procedure. High-Tune relies on an ensemble of simulations and on the analysis of their outputs to guide parameter adjustment. It is therefore an offline, semi-automatic calibration method. We have clarified this point in the revised manuscript, lines 829-831.

* data availability: Googling shows that the ftp link does exist, although I have no way of accessing it (but I also didn't try very hard). For best practice sub-samples of data and the relevant modification code should be archived and be immutable with appropriate doi anyway (e.g. Zenodo) as per the FAIR policy that Copernicus subscribes to. Probably do this at some point because I probably won't let this issue drop, but I will want to see the revised version again anyway so it's not the most urgent thing on the agenda.
We have create a Zenodo repository (https://doi.org/10.5281/zenodo.23536533), where the simulations' namelists, as well as the code for EDMF are available. We also provide information on how to access the model outputs via ftp (from linux or mac). Note that we could not directly upload the model outputs on Zenodo since our subset exceeds 200GB, which exceeds the 50GB authorized on Zenodo (using the free access).

* acknowledgments: Spelling, "...under granT number..." Also maybe thank the referees (and I suppose myself as handling editor and unofficial referee number 4) if you think they actually helped with the article in terms of science and/or presentation.
Correction added lines 856-858, 861.