nature portfolio

Peer Review File

Submesoscales are a significant turbulence source in global ocean surface boundary layer

Corresponding Author: Professor Jihai Dong

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This study explores the contribution of upper-ocean, small-scale (i.e., submesoscale) flows or processes to turbulence budgets in the global ocean. By carefully calibrating outputs from a submesoscale-permitting, global ocean model, the authors reveal that geostrophic shear production at ocean fronts is a significant source of turbulent kinetic energy (TKE) within the ocean surface boundary layer. While the magnitudes appear somewhat high (35% during winter, 18% during summer), the message to the reader is clear: these processes rival present-day understanding of ocean surface boundary layer turbulence. A minor but important point is that these processes are intermittent in time and space--a point that appears to have been made elsewhere but that nonetheless agrees with the present study.

It is clear this study took a tremendous effort on the part of the authors to combine several different scalings in a consistent manner. While this has been done before (Buckingham et al. 2019; Zippel et al. 2022; also deLavergne et al. 2020 for scalings for internal wave breaking), here the authors argue that the lateral scale of fronts unresolved in most models and observations (even down to the 10- or 100-m scale) need to be accounted for in a proper assessment of the contribution of submesoscale processes to TKE budgets in the upper ocean. It therefore diverges from the study of Buckingham et al. (2019, JAMES), who found that winds and waves dominate over submesoscale processes in terms of their effect on turbulence budgets. While this message might be a difficult one to accept, if correct, it could revolutionize our understanding of turbulence budgets in the upper ocean.

My main criticism is that the results depend upon a particular numerical model. It is important because the claim is being made that up to 35% of the energy in the upper ocean is dissipated by geostrophic shear within the fronts, implying that this was previously attributed to some other process. However, these estimates depend heavily on how well the model is realizing upper ocean fronts. In the absence of a demonstrated fidelity of the model to reproduce the observed fronts in the ocean (i.e., the studies of Rocha et al. mentioned on L293 do not constitute a proper evaluation of the model), the study falls short of its objective: the results of the authors cannot be extrapolated to the observed or real ocean. While other parts of the manuscript concern me, this is the main hindrance to publication in my view.

Reviewer #2

(Remarks to the Author)
Review of Submesoscales are a significant turbulence source in global ocean surface boundary layer

This submission uses an existing approximate analysis for the scaling of turbulence sources in the ocean surface layer to make inferences as to the relative magnitude of frontal submesoscale-driven turbulence. Though the idea is not new, and has been published previously by Buckingham et al. (2019), the authors use a newly derived parameterization for a frontal arrest scale, and some newer highly resolved simulations to look at the problem again. They come to considerably different conclusions, and find that the submesoscales can be a significant source of turbulence in the ocean surface layer. This result would be important for the prediction and modeling of ocean mixed layer depths and for understanding the processes that are influential there. This is a very important topic that would have significant implications for global ocean modeling. I find the submission somewhat difficult to review, since the methods were too opaque for me, and raised many questions about exactly where the differences with the Buckingham et al. study come about. I outline my concerns below.

The biggest concern I have with the submission is that the methods are not very transparent, and disagree with a recent publication that takes a nearly identical approach, namely Buckingham et al. (2019) in JAMES. As far as I can tell the Geostrophic Shear Production (GSP) is calculated in a similar way to Buckingham et al., but I see no discussion of the orientations of the wind stress and buoyancy gradients. The scaling of GSP listed on lines 315 and 319 seem to assume that winds are always down-front. If this is not the case (which I hope to be the case), the authors must describe in more detail how this term is quantified. This seems especially important since the conclusions in this submission are so different from Buckingham et al. with such a similar analysis. I would very much appreciate a list of all the steps and assumptions that are needed to quantify the GSP term. It was helpful to have the discussion of how results would change as the frontal width is corrected (or not), showing that it does not seem to make a large difference, despite some very large amplifications. Such an analysis for each of the different assumptions and steps would be very nice to see where this analysis diverges from Buckingham et al.. Here are some questions I had in the calculated GSP:

- How is wind/buoyancy gradient alignment accounted for, and how does the calculation of GSP differ from that of Buckingham et al.?
- Is every front that is not resolved then assumed to be at the frontal arrest scale? If so, what is the justification for this?
- In calculating the frontal arrest scale, is it always assumed that surface buoyancy fluxes are destabilizing? What is m_star, ie exactly how is it specified? (Stating that the reader must combine two equations in another reference to get it is not sufficient.)
- The discussion of the frontal arrest scale and its dependence on turbulence closure was difficult to connect to equation (8). Perhaps because I am missing information on m_star? What information is fed into eqn (8) from the turbulence parameterization?

One reason for the discrepancy with the Buckingham et al. results was offered on line 398, but I would think that this could easily be tested with the OSMOSIS data, which seems like the best platform to test the predictions of the new parameterizations out on. What exactly is the discrepancy with the Buckingham analysis? If not accounting for the wind/buoyancy gradient as described above, it seems to be the estimation of the horizontal buoyancy gradients. By what factor are the horizontal buoyancy gradients amplified in the analysis of the OSMOSIS data and how does this compare to the histograms shown in Buckingham's figure 11? On line 376 the authors state that the frontal arrest scale at the OSMOSIS site is tens of meters. Are they suggesting that the horizontal scale of fronts at the OSMOSIS site should actually be tens of meters? I find this difficult to believe, especially since it comes from extrapolating using a parameterization (I assume), and I know of no observations of such behavior. (Although here it seems the authors have a supplementary text that was not included in my review materials. I have notified the journal about this.) The view of Buckingham et al. is particularly bleak when discussing the alignment of the wind and buoyancy gradients:
"First, few fronts of appreciable magnitude pass through the
moored domain. This is seen both in the inner and outer mooring gradients (cf. Figures 8a and 8b). Second,
though the wind stress is considerable, with magnitudes reaching as high as 0.6–0.7 N/m2 (higher when
examining the hourly winds), these winds coincide with strong lateral buoyancy gradients (i.e., fronts) and
produce considerable buoyancy fluxes only a few times."

One other point that occurred to me is the situation with up-front winds will tend to restratify the mixed layer, thus damping turbulence with a positive buoyancy flux. This opposite effect seems to be excluded by the definition of the "Ocean Surface Boundary Layer" (OSBL) being above the 10^-8 W/kg dissipation level. With this definition of the OSBL, I would actually call it the "active mixing layer" or something similar. This is not to be confused with the surface mixed layer, which would include the positive buoyancy flux of the restratifying submesoscales (as seems to be noted on line 268). I think this is at least worth some discussion and justification of the choice made.

More minor points:

- Lines 59, 60 list references for the statement that fronts contribute a significant amount of turbulence to the boundary layer. The first two references show this, but both are at the wall of the Gulf Stream, and the second two are not relevant to frontal turbulence. More and balanced references (ie not just in the extreme location of the Gulf Stream) are needed for supporting this statement.

- Fig. 1 would make much more sense to me visually if all axes have the same range. Then one could see from the position of the distribution what the relative contributions are. As it is now I can't even read the axis labels without a lot of difficulty.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
Dong et al. revision Review:

The revised version of Dong et al. has come some way to alleviating my initial concerns. They offer a reasonable explanation to why their results differ so significantly from those of Buckingham et al.. In addition, I do have a better idea of how the analysis was carried out (although I need even more clarification), and the authors have also stated the limitations of the analysis in the revised version. I still find the results interesting and surprising, and that the study can be a significant contribution to the field. That said however, I still have significant concerns as to what I see as the primary drawbacks of the study, that have now become more clear to me with the authors replies:

Drawbacks:

1) Limited conditions of applicability for the analysis. The scaling of GSP is only applicable in down-front wind conditions and that for VBP in destabilizing surface buoyancy fluxes. This means the results are for 31% of winter and 21% of summer conditions (if I have understood the analysis, which I am still not convinced I have). Perhaps the authors can justify at the outset why they want to focus on these conditions, I think it would help very much with the applicability of the analysis.

2) All unresolved fronts are assumed to reach a very small horizontal scale that was simulated in a recent LES study, but has never been observed (I could not find such small scales even in the reference listed by the authors). This does not drastically change the results, but does put GSP into second place rather than third place (Table S1 with and without correction). The authors note that this amplification factor exceeds 6 at mid latitudes. Maybe this is correct, but it is at least not very well supported as far as I can tell (a single LES study). Also, in their comparison to OSMOSIS observations the authors find a factor 4 times larger frontal width gives better results than the LES-suggested parameterization, and this value is then used in the analysis (different choice of C_L). On the one hand, it gives more conservative estimates, but on the other it is all quite arbitrary and not very justified, especially when these tuned values are then compared to observations to justify the correction.

3) There is a somewhat arbitrary depth dependence to the results. This was identified by the authors as an important factor in why they get different results to the Buckingham et al study, which makes a lot of sense. Buckingham et al chose the depth of 45 m for their analysis, whereas the authors take the mid-depth of the mixed layer. While I do agree that the mid-depth is a more general choice, I am left a little unsatisfied at the generality of the analysis as a whole. We will get different results depending on our relative position within the mixed layer. This is just how the physics work, but it leaves me a little uncomfortable with statements in the text like:
- "GSP is found to be a leading contributor to turbulence in the OSBL and the prevalent one in winter"
- "the global contribution of energy transfer by submesoscales"
- "contributing 34% to the total dissipation in winter and 17% in summer"
It feels like the authors are talking about a vertically integrated measure of "turbulence", but they have chosen a specific depth (relative to the mixed layer) and only certain conditions (see point 2). I am concerned that readers will get a skewed impression of the "total dissipation".

Further comments that are more specific:

The discussion of the observations with and without the GSP (submesoscale) component does not seem balanced to me (lines 419-429). Yes, it can be seen that without the GSP term the observed dissipation mean is slightly larger than predicted, but it seems just as over-estimated by the addition of GSP. Given the log-normal distribution, I also would expect to see some confidence intervals for these means. Do these differences between scaling estimates with and without GSP lie outside or inside the confidence intervals? The conclusion presented in the Methods of the manuscript is simply that the addition of GSP makes the estimates better. It also seems a little contrived given the factor 4 reduction in frontal length used to make the correction fit best, if I have understood what has been done.

How exactly are the production rates for GSP calculated? I feel I am getting closer, but still unsure. Is it that at each grid point at each time a horizontal buoyancy gradient is calculated (at mid-depth of mixed layer) and then if there are unstable buoyancy flux conditions and a component of down-front wind stress, the formulae listed in the Methods section are applied? Then, these conditions are found 31% and 21% of the time in winter, summer, and this is where our numbers come from?

Very minor stuff I noticed:

- Switch order of integration in eq'ns (5,6) on LHS.
- Typo in time periods on lines 434-5.

Reviewer #3

(Remarks to the Author)
Using the re-calibrated outputs from the submesoscale-permiting global circulation model simulation (LLC4320) and turbulence parameterizations, the authors demonstrate that submesoscale (SMS) geostrophic shear production (GSP) is a significant source of turbulent kinetic energy (TKE) in the ocean surface boundary layer (OSBL). Along with the wind, waves and convection, turbulent dynamics at SMS fronts can provide an important route for the dissipation of the energy. Their finding aligns with recent regional observational studies and SMS-resolving numerical simulations in the past two decades. The message that authors

try to send to the readers is clearly important and necessary at this time. The main concern is whether LLC4320 can be trusted to quantitatively deliver this important message. As pointed out by the two previous reviewers, there are two main points of skepticism about the analysis: (1) how well does LLC4320 represent the real ocean conditions? and (2) why does the authors' finding conflict with the analysis of Buckingham et al. (2019) who suggested that SMS dynamics is not a significant source of TKE relative to the wind, waves and convection? The authors have nicely addressed many aspects of the skepticism in their revision. However, a few issues remain as I elaborate below.

With regard to the first point of concern, the authors responded by pointing to the recent publication of Gallmeier et al. (2023) who used artificial intelligence techniques to compare the SST in LLC4320 with the observed SST from Level-2 Visible Infrared Imaging Radiometer Suite (VIIRS) dataset. Gallmeier et al. (2023) conclude that "the LLC4320 simulation reproduces, over a large fraction of the ocean, the observed distribution of SSTa patterns well, both globally and regionally." It should be noted that the conclusion is based on training a machine learning algorithm with the distribution of structures in SST anomaly of 10-80 km scales. The conclusion is not applicable for the SST structure smaller than 10-km which is the SMS range of interest in the present analysis. Furthermore, Gallmeier et al. (2023) also pointed out a "modest, latitude-dependent offset" between LLC4320 and VIIRS. It is unclear whether the authors have taken into account this offset.

Besides referencing Gallmeier et al. (2023), the authors also compare the PDF of firstorder structure of SST between LLC4320 and VIIRS at different scales and they found that the difference between LLC4320 and VIIRS is significant at the small scales, and thus, requires correction for the lateral buoyancy gradient. Although the authors indicate that GSP remains to be significant whether or not they apply the correction for the buoyancy gradient (on line 258), it is clear that LLC4320 does not have the appropriate resolution for the quantitative analysis of the SMS dynamics. Besides SST, the OSBL thickness h is a crucial parameter used in their analysis and it is difficult to assess how well LLC4320 can capture h in the real ocean.

With regard to the second point of concern, the authors pointed out numerous differences between the present analysis and that used in Buckingham et al. (2019) that potentially can lead to the conflicting conclusions. The two studies target two different depths: mid-depth in the former and fixed 45-m in the latter. Buckingham et al. (2019) used 24-hr time filtering and did not include the correction of buoyancy gradient, both of which potentially can lessen the significance of GSP. The simulation used in Buckingham et al. (2019) (namely, NATL60) was constraint to North Atlantic while LLC4320 is a global circulation model. While the justification for the different conclusions between the two studies sounds reasonable, the concern of whether the simulations (both LLC4320 and NATL60) can represent the real ocean conditions remains. If the authors carry out the analysis exactly as in Buckingham et al. (2019) (focusing only on the North Atlantic subdomain in LLC4320, targeting the same depth, applying the same time filtering, etc.), would they obtain the same conclusion that GSP is insignificant relative to the wind, wave and convection? This exercise would help validate the use of LLC4320 and strengthen the conclusions in the present study.

It should be emphasized that none of the subtle (but important) differences between the two studies is mentioned in the abstract. Would the conclusion about the GSP stated in the abstract only valid in the bottom half of the OSBL but not in the upper half at global scale? If that is indeed true (i.e., both the present conclusion and that of Buckingham et al. 2019 are realistic to real ocean conditions), then it should be clearly stated in the abstract so that there would be no misunderstanding about the different results between the two studies.

Minor comments:

1. Line 372: AG is assumed to be 0.5 following the symmetric instability (SI) simulations in Thomas and Taylor (2010). Would SI be the only or the main mixing mechanism at arrested SMS fronts? For meandering SMS fronts with active mixed-layer instability (MLI), SI doesn't occur homogeneously in space as in the setup of Thomas and Taylor (2010). The LES of forced MLI in Hamlington et al. (2014) suggested SI occurs over localized regions of the front. Hamlington et al. (2014) found only 33% of the frontal area are exposed to symmetric instability. Would assuming the dissipation of GSP is only due to SI as in Thomas and Taylor (2010) over the entire frontal region overestimate the significance of GSP contribution in the real ocean? Also, I wonder the value of 0.5 in Thomas and Taylor (2010) was obtained by integrating the TKE budget over the OSBL or at mid-depth.

2. Inconsistent definition of OSBL thickness: "h is the OSBL thickness as determined by using an offline KPP scheme" (line 362) and "the OSBL thickness h is determined as

the depth where the observed dissipation rate decreases to a threshold value of 1×10^{-8} W kg^{-1}" (line 403-405). Please clarify.

3. Line 493: Is m_* set to be 0.5 as in Bodner et al. (2023)? If so, please include the value in the text.

4. Line 518: " ... while the GSP and horizontal shear production of the fronts themselves should contribute somewhat to the turbulence causing the arrest ..." Can the authors comment on the significance of the horizontal shear production with respect to the GSP contribution to the TKE especially in the limit of narrow (less than 100 m) arrested fronts? The horizontal shear production is not accounted for in the TKE budget (Equation 2).

References:
Bodner AS, Fox-Kemper B, Johnson L, Van Roekel LP, McWilliams JC, Sullivan PP, et al. Modifying the mixed layer eddy parameterization to include frontogenesis arrest by boundary layer turbulence. Journal of Physical Oceanography 2022.
Buckingham CE, Lucas NS, Belcher SE, Rippeth TP, Grant AL, Le Sommer J, et al. The contribution of surface and submesoscale processes to turbulence in the open ocean surface boundary layer. Journal of Advances in Modeling Earth Systems 2019, 11(12): 4066-4094.
Gallmeier KM, Prochaska JX, Cornillon P, Menemenlis D, Kelm M. An evaluation of the LLC4320 global ocean simulation based on the submesoscale structure of modeled sea surface temperature fields. Geoscientific Model Development Discussions 2023: 1-42.
Hamlington PE, Van Roekel LP, Fox-Kemper B, Julien K, Chini GP. Langmuir-submesoscale interactions: Descriptive analysis of multiscale frontal spindown simulations. J. Phys. Oceanogr. 2014, 44: 2249-2272.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
Review of version 3 of Dong et al.

In their responses and revised version of the manuscript, the authors have adequately addressed my concerns. They have provided qualifying statements in their revised version to narrow the scope of their findings, and in so doing, no longer overreach the results of the study. These results I still find significant and exciting, and informative for the ocean and climate community. The message that emerges is that the submesoscales contribute significantly to turbulent dissipation in the ocean surface mixed layer, and this source has been overlooked. The methodology that the authors used appears to be sound, and in line with our current understanding of these processes, and they have been careful in quantifying the new assumptions made. I congratulate the authors on a nice contribution to the field.

Minor comments:

- Lines 75-76: "... the prevalent one in winter." Please be more specific about which statistic is being referred to here. Is it the spatial prevalence (i.e. Fig. 4)?

- eq'ns (5,6): Again, I am pretty sure here that the order of integration should be d\theta dr (the limits of integration are fine, as the authors mentioned in their previous response).

Reviewer #3

(Remarks to the Author)
The authors have addressed all of my concerns. It's nice to know both LLC4320 and ENATL60 support the importance of GSP.

Reviewer #1

1) This study explores the contribution of upper-ocean, small-scale (i.e., submesoscale) flows or processes to turbulence budgets in the global ocean. By carefully calibrating outputs from a submesoscale-permitting, global ocean model, the authors reveal that geostrophic shear production at ocean fronts is a significant source of turbulent kinetic energy (TKE) within the ocean surface boundary layer. While the magnitudes appear somewhat high (35% during winter, 18% during summer), the message to the reader is clear: these processes rival present-day understanding of ocean surface boundary layer turbulence. A minor but important point is that these processes are intermittent in time and space--a point that appears to have been made elsewhere but that nonetheless agrees with the present study.

Response: Thanks for these comments. We believe these two points are the most important results of the work. Oceanographers have invested significant effort to understand turbulence in the surface boundary layer to improve the capability of ocean models. However, most works are focused on classical dynamic processes, such as convection and waves. In this work, submesoscales, which have become a focus in the recent decade, are quantitatively argued to be another important source of boundary layer turbulence. The significance of this work is that the result provides ideas on how to improve model simulation capability for oceanographers and ocean modelers. However, as suggested by Reviewer #2, the theory here is based on down-front wind conditions (i.e., winds that align down the frontal geostrophic currents), so in the revision we limit our results to down-front wind conditions. So, we recalculate these percentages which are slightly smaller (34% in winter and 17% in summer). Meanwhile, all numbers are changed and updated in the revised manuscript, and we have highlighted these changes in red in the main text.

2) It is clear this study took a tremendous effort on the part of the authors to combine several different scalings in a consistent manner. While this has been done before (Buckingham et al. 2019; Zippel et al. 2022; also deLavergne et al. 2020 for scalings for internal wave breaking), here the authors argue that the lateral scale of fronts unresolved in most models and observations (even down to the 10- or 100-m scale) need to be accounted for in a proper assessment of the contribution of submesoscale processes to TKE budgets in the upper ocean. It therefore diverges from the study of Buckingham et al. (2019, JAMES), who found that winds and waves dominate over submesoscale

processes in terms of their effect on turbulence budgets. While this message might be a difficult one to accept, if correct, it could revolutionize our understanding of turbulence budgets in the upper ocean.

Response: Thanks for the comment. We agree that the result challenges our existing knowledge of boundary layer turbulence. However, one point should be clarified. Our results here are different from Buckingham et al. (2019) but not contradictory to their results, because not only our method but our emphasis is slightly different.

The key difference between our and their works is the investigated depths are different (Figure R1), which is a likely reason for the difference in the results between ours and theirs. The dissipation rate from Buckingham et al. (2019) is at 45-m depth which is derived from a mounted 600-kHz ADCP (green line in Figure R1). By contrast, our dissipation rate is at the mid-depth of the mixed layer which is from dissipation profiles measured by a Seaglider (orange line in Figure R1). Our mid-depth choice is generally deeper than the fixed 45-m depth of Buckingham et al. in the winter period. Importantly, the relative magnitudes of the turbulence regimes are sensitive to the depth. This is because the vertical structures of these turbulence regimes are different. Taking the Langmuir turbulence (LSP) and submesoscale turbulence (LSP) as examples, LSP turbulence decreases more sharply with depth [if we look at the turbulence model with Langmuir used by Buckingham et al. (2019), the turbulence magnitude is roughly inversely proportional to depth]. But GSP turbulence decreases linearly with depth. It implies that the relative contribution of LSP turbulence to the total turbulence will decrease with depth while that of GSP turbulence tends to increase. This may explain why we get a larger submesoscale contribution compared to Buckingham et al. (2019). We note that other studies (e.g., Haney et al. 2015) find that the relative importance of the submesoscale varies with depth, even within the boundary layer.

We now emphasize that using a fixed depth near the surface to evaluate the contribution of submesoscale turbulence is a limited view of the boundary layer and mixed layer. One can only conclude that submesoscale turbulence is not significant compared to wave-driven turbulence near the surface from Buckingham et al. (2019) result, but it does not mean that submesoscale turbulence is not important away from the surface but still within the mixed layer. As another example, using the observations from the same project OSMOSIS, Yu et al. (2019) observed a strong mixing event in the OSBL during 8-10 April 2013. According to their analysis, the enhanced OSBL-integrated dissipation is mainly caused by submesoscales and the submesoscale-induced dissipation is about 3-4 times larger than the dissipation by buoyancy, wind, and waves [cf. Figure 4 of Yu et al. (2019)]. However, if we look at Figure 6 of Buckingham et al. (2019), their calculated submesoscale turbulence is always one order of magnitude smaller compared to the summation of other turbulence sources. Are their results contradictory? We do not think so. The difference between their results is because their research emphasis is different. Namely, it is an integral of the dissipation for Yu et al. (2019) but a fixed-depth dissipation near the surface for Buckingham et al. (2019). Nevertheless, their difference reminds us that comparing their relative contributions near the surface is only one aspect of boundary layer turbulence. Especially when we are interested in OSBL deepening, and the exchanges between the OSBL and ocean interior, a near-surface depth underestimates the importance of submesoscale turbulence.

As the vertical structures of these turbulence regimes are usually confined and scaled by the OSBL depth, we believe it is more dynamically reasonable to compare their contributions at a relative depth (i.e. mid-depth here) rather than a fixed depth. For the mid-depth here, if a significant contribution of submesoscale turbulence is obtained, we can conclude that submesoscale turbulence is important in the OSBL away from the sea surface and is crucial for the exchanges between the OSBL and the ocean interior. This choice is also consistent with related prior work such as Belcher et al. (2012).

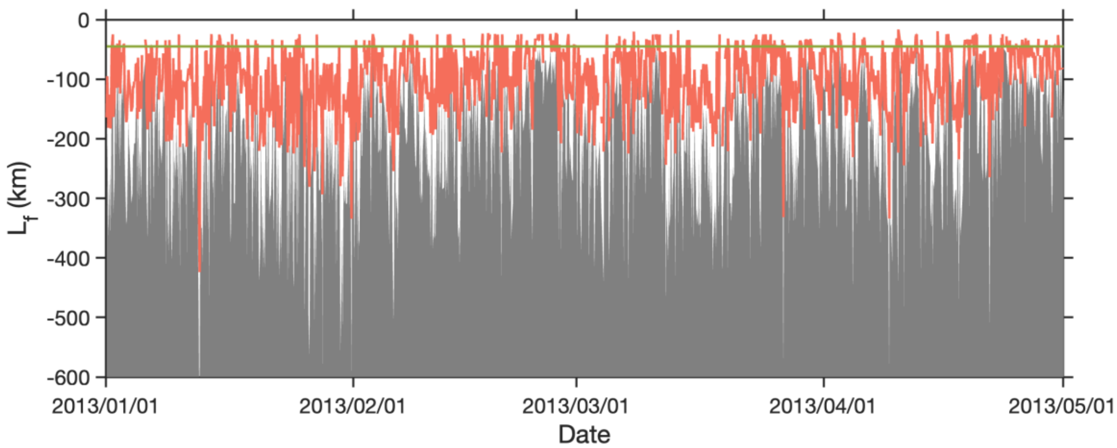

Figure R1 The mid depth of the mixed layer at the OSMOSIS site (orange line). The green line denotes the 45-m depth at which the dissipation rate is derived in Buckingham et al. (2019).

Moreover, there are also other potential reasons for the difference, some of which are important to the design of our study.

Firstly, submesoscale-induced turbulence has temporal scales of hours and shows strong intermittency, which is potentially removed by the 24-hour filter as used by Buckingham et al. (2019).

Second, as the reviewer states, the strength of submesoscale turbulence heavily relies on the intensity of fronts (or buoyancy gradients) in the boundary layer. As we emphasize in the manuscript, 10-km fronts would be represented in daily averages of the mooring-observed buoyancy, but many smaller more intense fronts are likely important. Faster signals from fronts as small as 2.6 km are directly detectable in the OSMOSIS data. So, Buckingham et al. (2019) potentially underestimate the frontal intensity due to their 24-hour filter applied to the buoyancy. Even for the original buoyancy gradients, they still fail to resolve local arrested fronts under the TTW balance which have spatial scales of less than 1km according to our estimation based on the theory of Bodner et al. (2023) (Figure R2). Fronts under the TTW balance is believed to be more dynamically rational in the OSBL and has been reported by a number of works (e.g., Gula et al., 2014; McWilliams et al., 2015; Wenegrat and McPhaden 2016). Directly using the original buoyancy gradients will underestimate the submesoscale turbulence intensity and the buoyancy gradient should be rescaled in analyzing its contribution to OSBL turbulence.

In the revised manuscript, we discuss this point further in Line 442-449.

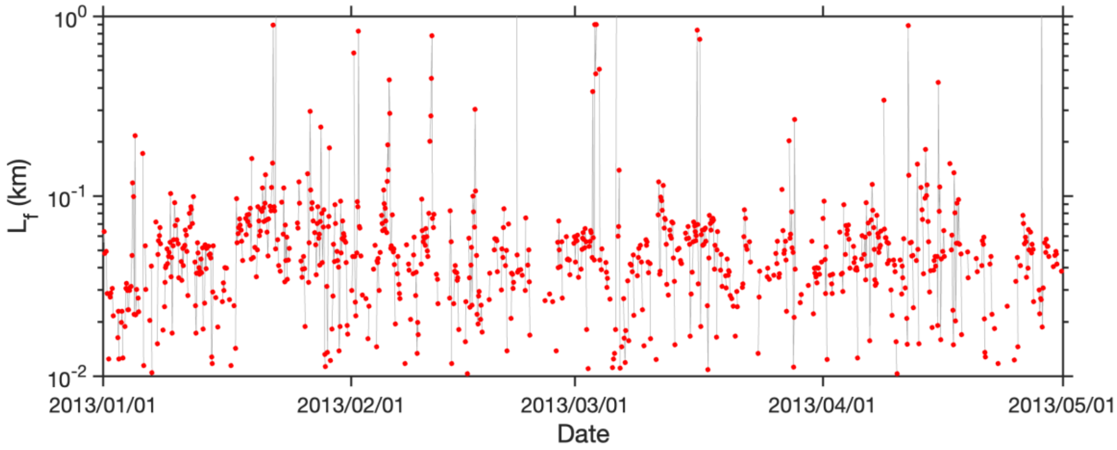

Figure R2 The estimated frontal arrested scale under the TTW balance at the OSMOSIS site based on the method of Bodner et al. (2023).

3) My main criticism is that the results depend upon a particular numerical model. It is important because the claim is being made that up to 35% of the energy in the upper ocean is dissipated by geostrophic shear within the fronts,

implying that this was previously attributed to some other process. However, these estimates depend heavily on how well the model is realizing upper ocean fronts. In the absence of a demonstrated fidelity of the model to reproduce the observed fronts in the ocean (i.e., the studies of Rocha et al. mentioned on L293 do not constitute a proper evaluation of the model), the study falls short of its objective: the results of the authors cannot be extrapolated to the observed or real ocean. While other parts of the manuscript concern me, this is the main hindrance to publication in my view.

Response: Thanks for pointing out the concern. The studies of Rocha et al. mainly evaluate LLC4320 in two regions, the Kuroshio Extension (Rocha et al., 2016, GRL) and the Drake Passage in the Southern Ocean (Rocha et al., 2016, JPO). In the GRL paper, they evaluate the upper ocean dynamics and summarize that the upper ocean density structure is well captured by the LLC4320, according to a comparison of the Kuroshio Extension stratification and its seasonal variability and eddy kinetic energy with Argo climatology and the Jason II across-track geostrophic eddy kinetic energy. In the JPO paper, the spectral slope features of the kinetic energy in the Drake Passage are similar between the LLC4320 and observations, despite they indeed find a higher intensity of the upper ocean kinetic energy in the LLC4320 simulation. However, a higher intensity may be not surprising since it is an averaged result of 14-year observations while it is only an 84-day result in 2012 from the simulation. For example, relatively larger mean wind velocities in the Southern Ocean during the simulated period have been reported by Flexas et al. (2019).

We agree that the results here heavily depend on the performance of the LLC4320 model, especially about the accuracy of simulated fronts in the upper ocean. In the revised version, we conduct more analysis to demonstrate the capability of the model output in reproducing frontal dynamics in the sea surface boundary layer and reference a recent paper examining fronts in this model versus satellite detection of fronts (Gallmeier et al., 2023).

Before the analysis results are shown, one point needs to be clarified first. Our result is mainly based on the high-resolution simulation of the frontal buoyancy (or density) gradients in the surface boundary layer. However, directly evaluating the buoyancy is impossible because there are no available high-resolution observations of buoyancy over the globe. So, we decided to evaluate the model output based on satellite-derived sea surface temperature (SST), whose L2 products usually have a spatial resolution of less than 1 km.

By coincidence, a quantitative comparison of SST between LLC4320 and

Visible Infrared Imaging Radiometer Suite (VIIRS) SST products is conducted recently by Gallmeier et al. (2023) using an AI method. As they argue in the abstract, “A principal finding is that the LLC4320 simulation reproduces well, over a large fraction of the ocean, the observed distribution of SST patterns, both globally and regionally.” We believe this statement is a great support to our results here.

Further, we believe only a comparison of SST broadly is not enough to directly evaluate the skill of the LLC4320 at the particular metrics used in our paper. As we focus on buoyancy gradients, we further compare the spatial variance of SST between LLC4320 and VIIRS using the first-order structure function (King et al., 2015; Yu et al., 2017). Here, the VIIRS L2 data have spatial resolutions around 1 km (from 0.75 km at nadir to 1.5 km at the swath edge) and are retrieved from the JPL Physical Oceanography Distributed Active Archive Center (https://doi.org/10.5067/GHVRS-2PO28). And the LLC4320 SST is defined as the uppermost 0.5 m level of the simulation. As the VIIRS L2 data have missing values due to clouds, the structure function can avoid the effect of these missing values and statistically demonstrates the capability of the LLC4320 model in reproducing SST variances.

The first-order structure function here is defined as the difference of SST between the pair of points, $\vec{x}$, and $\vec{x}+\vec{r}$, namely,

$$\delta = SST(\vec{x}+\vec{r}) - SST(\vec{x}) \tag{R1}$$

The structure function clearly relates to fronts, which are jumps in SST and other variables. We calculate the probability density functions (PDF) of SST structure functions δ at different scales (r = 100 km, 80 km, 60 km, 50km, 40 km, 30 km, 20 km ,10km and 5 km) based on VIIRS and LLC4320 data in the same period (February and August of 2012). To avoid the effect of the missing values in VIIRS, we interpolate the LLC4320 data onto the VIIRS grids at the corresponding dates, and then avoid the corresponding missing-value regions in the LLC4320 data. The structure function PDFs of larger separation distances are expected to show consistent distributions for these two datasets, as this is where model resolution is not an issue. But as r decreases, the PDFs from LLC4320 are speculated to underestimate the magnitude of fronts from VIIRS. The frontal rescaling to TTW based on Bodner et al. (2023) used in our paper is specifically intended to correct the small-separation-scale frontal strengths *vs.* their underestimated strength in the LLC4320.

The calculated PDFs of the first-order structure functions of SST in different regions from the VIIRS and LLC4320 data are shown (Figure R3 and Figure R4). Their differences shown in Figure R5 confirm the speculation that they agree in probably magnitude of fronts measured over large separations but disagree over small separations. The negligible differences between LLC4320 and VIIRS indicate that LLC4320 reproduces the probability of finding SST jumps well at large separation scales. However, as the scale decreases below the effective resolution, the biases begin to become more and more consequential. The positive biases at small SST magnitudes and negative biases at large SST magnitudes imply that LLC4320 increasingly underestimates the SST jumps (or gradients) in the real ocean. So, we need to correct the LLC4320 SST gradients due to the underestimation.

In conclusion, despite that our results are mainly based on a simulation, our comparison of the simulation to the observations indicates that the simulation can statistically reproduce OSBL fronts well in the real ocean at large scales, and we already correct the underestimation at small scales based on Bodner et al. (2023). We also note that the comparison of the theory and LLC4320 with the OSMOSIS observations provides another line of evidence supporting the analysis approach used here.

In the revised manuscript, we add some paragraphs to discuss the model capability in Line 312-344.

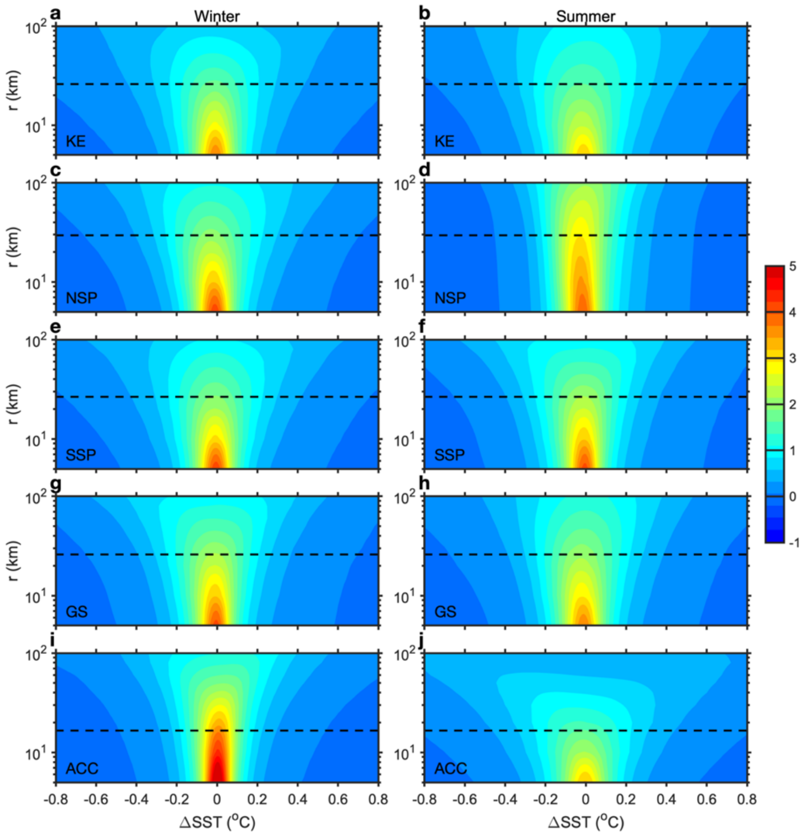

Figure R3 **Probability density functions of the first-order structure functions of SST in different regions derived from the VIIRS data. a, b,** the Kuroshio Extension (32~38 ^{o}N, 150~156 ^{o}E). **c, d,** the Northern Subtropical Pacific (15~21 ^{o}N, 180~186 ^{o}E). **e, f,** the Southern Subtropical Pacific (20~26 ^{o}S, 120~126 ^{o}W). **g, h,** the Gulf Stream (28~34 ^{o}N, 60~66 ^{o}W). **i, j,** the Antarctic Circumpolar Current (50~56 ^{o}S, 115~121 ^{o}E) (left: winter; right: summer). The dashed lines denote the minimum wavelengths that the effective resolution resolves (i.e., two times the effect resolution 7Δx).

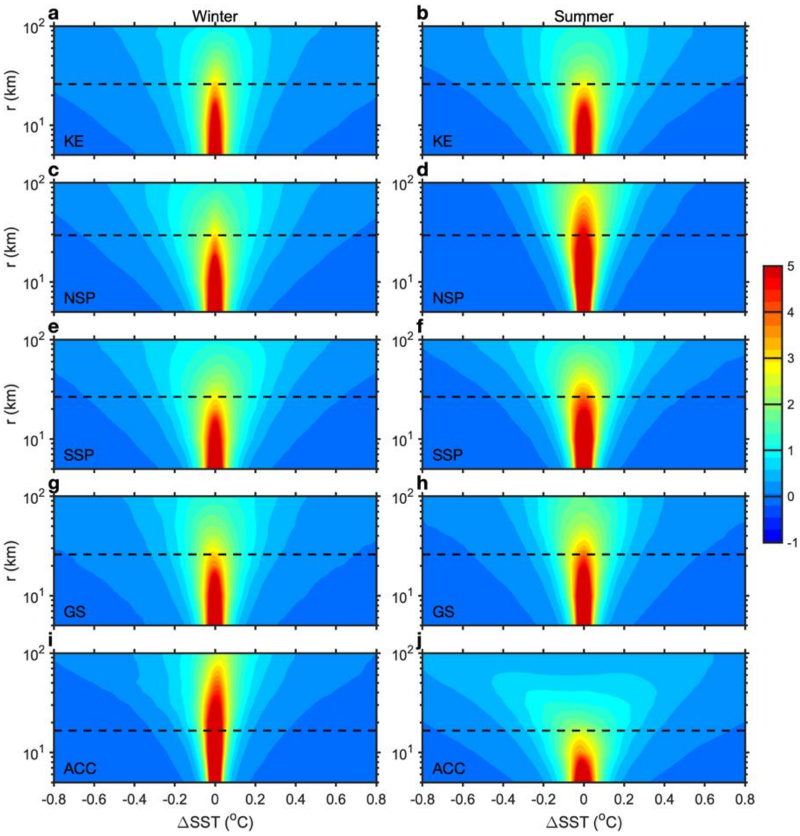

Figure R4 **Probability density functions of the first-order structure functions of SST in different regions derived from the LLC4320 data. a, b,** the Kuroshio Extension (32~38 °N, 150~156 °E). **c, d,** the Northern Subtropical Pacific (15~21 °N, 180~186 °E). **e, f,** the Southern Subtropical Pacific (20~26 °S, 120~126 °W). **g, h,** the Gulf Stream (28~34 °N, 60~66 °W). **i, j,** the Antarctic Circumpolar Current (50~56 °S, 115~121 °E) (left: winter; right: summer). The dashed lines denote the minimum wavelengths that the effective resolution resolves (i.e., two times the effect resolution 7Δx).

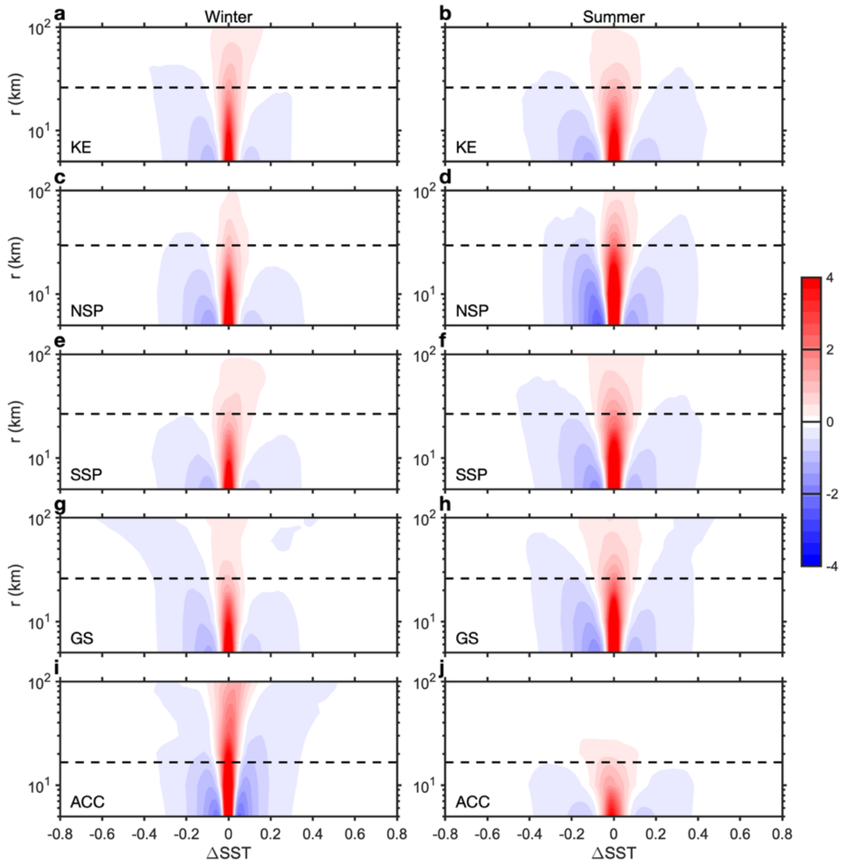

Figure R5 **Probability density functions differences of the first-order structure functions of SST in different regions. a, b,** the Kuroshio Extension (32~38 °N, 150~156 °E). **c, d,** the Northern Subtropical Pacific (15~21 °N, 180~186 °E). **e, f,** the Southern Subtropical Pacific (20~26 °S, 120~126 °W). **g, h,** the Gulf Stream (28~34 °N, 60~66 °W). **i, j,** the Antarctic Circumpolar Current (50~56 °S, 115~121 °E) (left: winter; right: summer). The dashed lines denote the minimum wavelengths that the effective resolution resolves (i.e., two times the effect resolution 7Δx).

Thanks again for your valuable comments!

Reviewer #2

1) This submission uses an existing approximate analysis for the scaling of turbulence sources in the ocean surface layer to make inferences as to the relative magnitude of frontal submesoscale-driven turbulence. Though the idea is not new, and has been published previously by Buckingham et al. (2019), the authors use a newly derived parameterization for a frontal arrest scale, and some newer highly resolved simulations to look at the problem again. They come to considerably different conclusions, and find that the submesoscales can be a significant source of turbulence in the ocean surface layer. This result would be important for the prediction and modeling of ocean mixed layer depths and for understanding the processes that are influential there. This is a very important topic that would have significant implications for global ocean modeling.

Response: Thanks for the comment. Buckingham et al. (2019) have evaluated the contributions from different processes to the boundary layer turbulence based on observations in the Northeastern Atlantic Ocean. They found that the relative contribution of submesoscale-turbulence is limited. However, several differences need to be pointed out.

First, the result by Buckingham et al. (2019) is based on single-point observations. The study region cannot be extrapolated to general conclusions about the role of submesoscale turbulence over the global ocean. For example, Dong et al. (2021) (Figure 8) also compared the wind-work reduction due to submesoscales with Buckingham et al. (2019). The result indicates that their magnitude derived based on the northeastern Atlantic region underestimated the globally averaged magnitude. We agree that investigating the importance of submesoscale turbulence is not a newly proposed idea, since a number of works have been conducted from a decade ago, such as D'Asaro et al. (2011), Thomas et al. (2016), however there are key innovations in our work that differentiate it.

Second, our work differs from Buckingham et al. (2019) not only in a newly derived parameterization for a frontal arrest scale but also in the turbulence budget theory. Buckingham et al. (2019) used three different parameterization schemes (wind&wave, submesoscale, and vertical convection) to evaluate their productions. But in our work, we extend the non-dimensional budget equation of Belcher et al. (2012) by including the submesoscale term.

Third, another key difference between our and their works is the investigated depths are different (Figure R1), which we hypothesize is a key reason for the difference in the results between ours and theirs. The dissipation rate from Buckingham et al. (2019) is at 45-m depth which is derived from a mounted 600-kHz ADCP (green line in Figure R1). By contrast, our dissipation rate is at the mid-depth of the mixed layer which is from dissipation profiles measured by a seaglider (orange line in Figure R1). It can be observed that the mid-depth is generally deeper than the fixed 45-m depth in the winter period. However, it should be noted that the relative magnitudes of the turbulence regimes are sensitive to the depth. This is because the vertical structures of these turbulence regimes are different. Taking the Langmuir turbulence (LSP) and submesoscale turbulence (LSP) as examples, LSP turbulence decreases more sharply with depth [if we look at the turbulence model with Langmuir used by Buckingham et al. (2019), the turbulence magnitude is roughly inversely proportional to depth]. But GSP turbulence decreases linearly with depth. It implies that the relative contribution of LSP turbulence to the total turbulence will decrease with depth while that of GSP turbulence tends to increase. This may explain why we get a larger submesoscale contribution compared to Buckingham et al. (2019). This distinction has now been highlighted in the text (see also the response to Reviewer 1).

So, we argue that using a fixed depth near the surface to evaluate the contribution of submesoscale turbulence is an incomplete view. One can only conclude that submesoscale turbulence is not significant compared to waves near the surface from Buckingham et al. (2019) result, but it does not follow that submesoscale turbulence is not important away from the surface. Please also see our response to Comment 2 of Reviewer #1 on the comparison of our work to Buckingham et al. (2019) for examples where vertical structure matters for the interpretation of the submesoscale *vs.* other boundary layer processes.

In the revised manuscript, we highlight our difference from Buckingham et al. (2019) in Line 442-449.

2) I find the submission somewhat difficult to review, since the methods were too opaque for me, and raised many questions about exactly where the differences with the Buckingham et al. study come about. I outline my concerns below.

Response: Thanks for the comment. We have to admit that we use a number of methods in order to quantify the submesoscale turbulence from a global perspective. However, the methods are not totally newly proposed in our work.

The turbulent kinetic budget equation without the submesoscale production term actually has been applied by Belcher et al. (2012) and Li et al. (2019) to investigate the role of Langmuir turbulence over the globe. The submesoscale production (i.e. GSP) is derived from the work by Thomas and Taylor (2010). This work here linearly combines them to evaluate their relative contributions. The frontal arrest scale is a newly proposed method by Bodner et al. (2023) which is demonstrated to describe the frontal scale in the boundary layer more reasonably under thermal-wind balance. While we believe this is the most quantitatively accurate approach, the results comparing ‘raw’ model gradients show that the major conclusions of this study regarding the significant role of the GSP do not depend on this rescaling.

About the differences, we have discussed them in Comment #1. Hopefully that response will satisfy you. Also, in the revised manuscript, we highlight our differences from Buckingham et al. (2019) in Line 442-449.

3) How is wind/buoyancy gradient alignment accounted for, and how does the calculation of GSP differ from that of Buckingham et al.?

Response: Thanks for the comment, as addressing it has improved the paper. In the original version, we did not consider the angle between fronts and winds, since the small-scale arrested fronts do not necessarily share the same direction as the model-resolved fronts. Following your comments, we have decided to limit our result to the down-front wind conditions (i.e., mixing favorable). It implies that we must also assume that all unresolved small-scale arrested fronts (for which we use the Bodner et al. (2023) approach to approximate) share the same direction as the model-resolved fronts. In the revised manuscript, all results are under the condition of down-front winds, and the figures and main text are all updated.

According to our evaluation, it is about 31% and 21% of the time in winter and summer meet the down-front wind and destabilizing conditions. It means GSP contributes ~40% in a third of the winter. It should be claimed that this is the most conservative estimation since even in up-front wind conditions GSP is expected to have a vertical structure similar to AGSP and a comparable magnitude of the GSP contribution to the down-front case will be derived (see Line 264-271).

About the calculation of GSP, in this work, we directly use the method proposed by Thomas and Taylor (2010; TT10 hereinafter) which is obtained from a front with down-front winds. Buckingham et al. (2019) used the method proposed by Thomas et al. (2013; T13 hereinafter) which also considers the effect of sea surface forcing in addition to down-front winds. Under destabilizing sea surface buoyancy flux, the T13 method may get a more reasonable result compared to the TT10 method. However, given that the buoyancy flux has been parameterized in the unextended non-dimensional equation (Belcher et al., 2012), we decided to use the TT10 method. It should be noted that the TT10 method potentially underestimates the GSP magnitude compared to the method of T13. In the TT10 method, GSP is parameterized as

$$GSP = EBF\frac{z+H}{H} \tag{R2}$$

Here, $EBF = \frac{\boldsymbol{\tau}\times\boldsymbol{k}}{\rho_0 f}\cdot\nabla_h b$ is the Ekman buoyancy flux due to down-front winds, *H* is the OSBL thickness, *z* is the water depth ($-H \le z < 0$). But for the T13 method, the contribution of the destabilizing sea surface buoyancy flux (i.e., $B_0 > 0$) is considered, and the expression is modified as

$$GSP = \begin{cases} 0, z = 0 \\ (EBF + B_0)\frac{z+H}{H} - B_0\frac{z+h}{h}, -h \le z < 0 \\ (EBF + B_0)\frac{z+H}{H}, -H \le z < h \end{cases} \tag{R3}$$

Here, *h* is the convective layer thickness. The expression can be rewritten as

$$GSP = \begin{cases} 0, z = 0 \\ EBF\frac{z+H}{H} + B_0 z\frac{h-H}{Hh}, -h \le z < 0 \\ EBF\frac{z+H}{H} + B_0\frac{z+H}{H}, -H \le z < h \end{cases} \tag{R4}$$

It can be observed that the T13 method will always get a larger GSP compared to the TT10 method. So, because the interaction between GSP and vertical convection (i.e., what we classify as VBP) is not considered, the GSP turbulence is potentially underestimated in this work. However, as this avoids conflating the effects of VBP and GSP confounding our classification, we think that is the sensible approach. Caveats as to possible interaction between turbulence sources are given in the manuscript (Line 271-278).

In the revised manuscript, the angle between the wind and frontal axis is included in the definition of the geostrophic shear stability length (Line 365-367).

The underestimation due to the TT10 method is also discussed (Line 372-375).

4) Is every front that is not resolved then assumed to be at the frontal arrest scale? If so, what is the justification for this?

Response: Thanks for the comment. In this work, we indeed assume that the smallest unresolved fronts are under arrest. (Note that as explained in the methods, Equations (5) to (8), we use an integral over the spectrum to evaluate M^2, and only the cutoff length scale is set to the arrest scale.) In the real ocean, fronts evolve in time and can experience frontogenesis, arrest and frontolysis. It can be expected that the dissipation at frontogenesis and frontolysis will be different from the arrest stage. Nevertheless, two reasons make our results of practical significance. First, many have reported that the turbulence thermal-wind (TTW) balance is a very common stage of OSBL fronts (Gula et al., 2014; Bonder et al., 2023). We still have no idea about the proportion of the arrested fronts in the real ocean, a result that can only be found with super high-resolution global observations. However, in Bodner et al. (2023), the predicted frontal arrest scale is validated and demonstrated to statistically agree well with the result from Large Eddy Simulations and have a range in the Bay of Bengal similar to high-resolution observations collected there. In their comparison, a number of cases including surface cooling, wind with different directions, and surface waves are considered. One would expect that there are frontogenesis, frontolysis underway in addition to frontal arrest in these cases.

Second, OSBL fronts in the arrest state have the strongest intensity of the buoyancy gradient compared to the other two stages, leading to the strongest submesoscale turbulence. So, the estimation of the submesoscale M^2 that contributes to submesoscale turbulence generation here is a maximum magnitude that OSBL fronts can reach, not an average over their whole life.

Finally, we note that the major results of this work are qualitatively robust even when directly using the raw model output buoyancy gradients, suggesting the choice of frontal arrest scale does not itself determine the finding that GSP contributes significantly to the global OSBL TKE budget.

In the revised manuscript, we make this point clearer in Line 379-386.

5) In calculating the frontal arrest scale, is it always assumed that surface buoyancy fluxes are destabilizing? What is m_star, ie exactly how is it specified?

(Stating that the reader must combine two equations in another reference to get it is not sufficient.)

Response: Thanks for the comment. According to Reichl and Hallberg (2018), $\overline{u'w'} = (m_* u_*{}^3 + n_* w_*{}^3)^{2/3}$ scales the magnitude of the dissipation in the OSBL due to the mechanical wind forcing and the sea surface buoyancy forcing. Here, u_* and w_* are the frictional and convective velocities, respectively; m_* and n_* measure the efficiency of these two forcings on changing the TKE in the OSBL. The application of this parameterization does not require a destabilizing surface buoyance flux. However, according to Bodner et al. (2023), only the destabilizing case is considered in the validation of the theory of the frontal arrest scale. Meanwhile, our TKE model is also only applicable under destabilizing as well. As a result, we only consider destabilizing in our whole work. This approach is consistent with previous work estimating convective energy sources (Belcher et al. 2012, Li et al. 2019). Moreover, as suggested by Comment 3), we now further limit our results to down-front wind conditions.

In the revised work, we discuss this point in Line 264-271. Also, we explain m* and n* more explicitly in Line 492-496.

6) The discussion of the frontal arrest scale and its dependence on turbulence closure was difficult to connect to equation (8). Perhaps because I am missing information on m_star? What information is fed into eqn (8) from the turbulence parameterization?

Response: Sorry for the confusion. The derivation of the frontal arrest scale is complex and we will not show all details here. Please allow us to explain it briefly here. According to Bodner et al. (2023), the TTW balance can be expressed as

$$\nabla_h b = -f\boldsymbol{k} \times \frac{\partial \bar{\boldsymbol{u}}}{\partial z} + \frac{\partial^2}{\partial z^2}\left(\overline{\boldsymbol{u}'w'}\right) \quad \text{(R5)}$$

Here, the velocity field is decomposed into the temporal average (the basic velocity field) and the perturbation (small-scale turbulence) by using the Reynolds-averaging. By applying the planetary boundary layer scheme (ePBL; Reichl and Hallberg, 2018), the Reynolds stress term can be parameterized as

$$\overline{\boldsymbol{u}'w'} = (m_* u_*{}^3 + n_* w_*{}^3)^{2/3} \quad \text{(R6)}$$

Here, $u_* = \sqrt{\frac{|\tau_w|}{\rho}}$ is the friction velocity, $w_* = (B_0 h)^{1/3}$ is the convective velocity, and m_* and n_* measure the efficiency of these two forcings on changing the TKE in the OSBL. Here, the turbulence is assumed to be generated by sea surface wind and buoyancy forcings. By a scale analysis, the buoyancy gradient can be scaled as $\nabla_h b \sim \frac{\Delta b}{L_f}$. This is a rough explanation of how the turbulence closure connects to the frontal arrest scale. For detailed mathematical derivation please refer to Bodner et al. (2023). In the revised manuscript, we add more details about the connection in Line 485-489.

7) One reason for the discrepancy with the Buckingham et al. results was offered on line 398, but I would think that this could easily be tested with the OSMOSIS data, which seems like the best platform to test the predictions of the new parameterizations out on. What exactly is the discrepancy with the Buckingham analysis? If not accounting for the wind/buoyancy gradient as described above, it seems to be the estimation of the horizontal buoyancy gradients. By what factor are the horizontal buoyancy gradients amplified in the analysis of the OSMOSIS data and how does this compare to the histograms shown in Buckingham's figure 11? On line 376 the authors state that the frontal arrest scale at the OSMOSIS site is tens of meters. Are they suggesting that the horizontal scale of fronts at the OSMOSIS site should actually be tens of meters? I find this difficult to believe, especially since it comes from extrapolating using a parameterization (I assume), and I know of no observations of such behavior. (Although here it seems the authors have a supplementary text that was not included in my review materials. I have notified the journal about this.) The view of Buckingham et al. is particularly bleak when discussing the alignment of the wind and buoyancy gradients: "First, few fronts of appreciable magnitude pass through the moored domain. This is seen both in the inner and outer mooring gradients (cf. Figures 8a and 8b). Second, though the wind stress is considerable, with magnitudes reaching as high as 0.6–0.7 N/m2 (higher when examining the hourly winds), these winds coincide with strong lateral buoyancy gradients (i.e., fronts) and produce considerable buoyancy fluxes only a few times."

Response: Thanks for the valuable comment! Combining the comments here and above, we understand that your major concern actually is about the discrepancy between our results and Buckingham et al. (2019). Above and in the response to Reviewer #1 and in the revised text we have clarified the differences between our method and focus from those of Buckingham et al. (2019), please allow us to explain again according to your comments here.

First, our results are not directly comparable to Buckingham et al. (2019) and the difference between these two results is not a discrepancy. One important distinction is the depth. As we discuss in the comments above, the investigating depth of Buckingham et al. (2019) is a fixed depth of ~45 m (green line in Figure R1). By contrast, our investigating depth is the mid-depth of the OSBL. In the winter period, the mid-depth is much deeper than the fixed depth. In the method section "Validation of the TKE model" and the Supplementary information, we conduct validation of the method compared to the OSMOSIS observations (**Fig. S7**). The results show that the production without GSP tends to underestimate the observed dissipation. By contrast, high GSP events shift the probability density function (PDF) of the dissipation towards larger values, correcting the underestimation and the corrected PDF peak is more consistent with observations.

One can only conclude that submesoscale turbulence is not significant compared to waves near the surface from Buckingham et al. (2019) result, but it does not mean that submesoscale turbulence is not important away from the surface. In our response to Reviewer #1 above, we offer examples where depth of investigation has a controlling effect on the relative importance of submesoscales.

As the vertical structures of these turbulence regimes are usually confined and scaled by the OSBL depth, we believe it is more dynamically reasonable to compare their contributions at a relative depth (i.e. mid-depth here) rather than a fixed depth. For the mid-depth here, if a significant contribution of submesoscale turbulence is obtained, we can conclude that submesoscale turbulence is important in the OSBL away from the sea surface and is crucial for the exchanges between the OSBL and the ocean interior.

Second, about the correction of the horizontal buoyancy gradient, an important base is the frontal scale in the ocean. In the OSBL, the significant role of turbulence results in the failure of the geostrophic theory in describing OSBL frontal dynamics. Instead, the turbulent thermal-wind balance theory is demonstrated to be more suitable for describing OSBL front (McWilliams et al., 2015; Gula et al. 2014; Dauhajre and McWilliams 2018; Sullivan and McWilliams, 2018; 2019). Due to a high requirement of observations (you need to observe vertical profiles of dissipation, buoyancy and velocity in a high spatial resolution), direct observation of TTW is rarely reported. So, up to now, most studies on TTW are based on high-resolution simulations. However, according to these results, their frontal scales under TTW indeed can reach less than hundreds of meters. For example, Sullivan and McWilliams (2018) simulated a filament and the filament shrinks and finally is arrested at a frontal scale of O(100) m (see their Figure 17), despite that their simulations are conducted using idealized Large Eddy Simulations (LES) not in a realistic ocean region (Note that the Coriolis parameter of their cases corresponds to a latitude of 32~33^{o} and their horizontal resolution is ~ 1 m). In Bodner et al. (2023), they also validated the theoretically predicted frontal arrest scale with Large Eddy Simulations, and they considered more conditions (such as Stokes drift) and obtained roughly consistent frontal arrest scales as well (see their Figure 1). In their Figure 1, one can also observe frontal arrest scales less than 100 m.

It is a pity that there are still no direct global observations reporting arrested OSBL fronts. However, this could also be indirect evidence for the small arrest scale of OSBL fronts. Only with high-resolution observations, such as by towed high-frequency observations, can such surveys be carried out. The Bay of Bengal is one location where such surveys have occurred, and Bodner et al. show that the frontal widths from the theory agree with the range of observed fronts there.

To address the "bleak view" of Buckingham et al. regarding appreciably strong fronts, it is important to note that the rescaling theory (equations (5)-(8)) does not require that the smallest, arrested fronts be strong. Indeed, they are assumed to obey a spectral roll-off in buoyancy variance, so the smaller fronts are weaker than the larger fronts. The arrested frontal scale only enters the correction for M^2 as a lower bound on the integral across all of the eddying fronts (resolved and unresolved, per equations (5)-(8)).

Also, for OSBL fronts under TTW, the front arrest scale should depend on latitude and should become smaller and smaller as latitude increases. We can easily get this dependence from the TTW theory. The TTW balance is expressed as

$$f\boldsymbol{k}\times\frac{\partial \boldsymbol{u}}{\partial z}=-\nabla_h b+\frac{\partial^2}{\partial z^2}\left(\kappa\frac{\partial \boldsymbol{u}}{\partial z}\right) \tag{R5}$$

Following Bodner et al. (2023), the buoyancy gradient variance equation can be scaled by (see their Section 2),

$$|\nabla_h b|^2=\left|f\boldsymbol{k}\times\frac{\partial \boldsymbol{u}}{\partial z}\right|^2+\left|\frac{\partial^2}{\partial z^2}\left(\kappa\frac{\partial \boldsymbol{u}}{\partial z}\right)\right|^2 \tag{R6}$$

One can roughly estimate the frontal scale magnitude as follows:

$$L_f{}^2 = \frac{\Delta b^2}{\left|f\boldsymbol{k}\times\frac{\partial \boldsymbol{u}}{\partial z}\right|^2 + \left|\frac{\partial^2}{\partial z^2}\left(\kappa\frac{\partial \boldsymbol{u}}{\partial z}\right)\right|^2} \tag{R7}$$

As result, for fronts with the same velocity shear and eddy viscosity, the front at higher latitude tends to have smaller arrest scales. The OSMOSIS mooring array is located at about 48°N. According to the LES simulation result from Sullivan and McWilliams (2018), one could expect that the frontal arrest scale should be even smaller due to the higher latitude.

Thus our estimation of the arrest scale at tens of meters is reasonable—especially noting that appreciably strong fronts may not be arrested, it is only a smallest possible frontal scale. The following figure is the frontal arrest scale that we estimate over the globe using the method from Bodner et al. (2023) in the supplementary file (Figure R6). It can be observed that the front arrest scale can reach several kilometers at low latitudes, but decrease sharply as latitude increases. It is noteworthy that the calculated scale is not very sensitive to which data are used. We use two independent datasets with different vertical turbulent closures to calculate the frontal arrest scale, and the calculated frontal arrest scales are almost consistent and show similarity in seasonality.

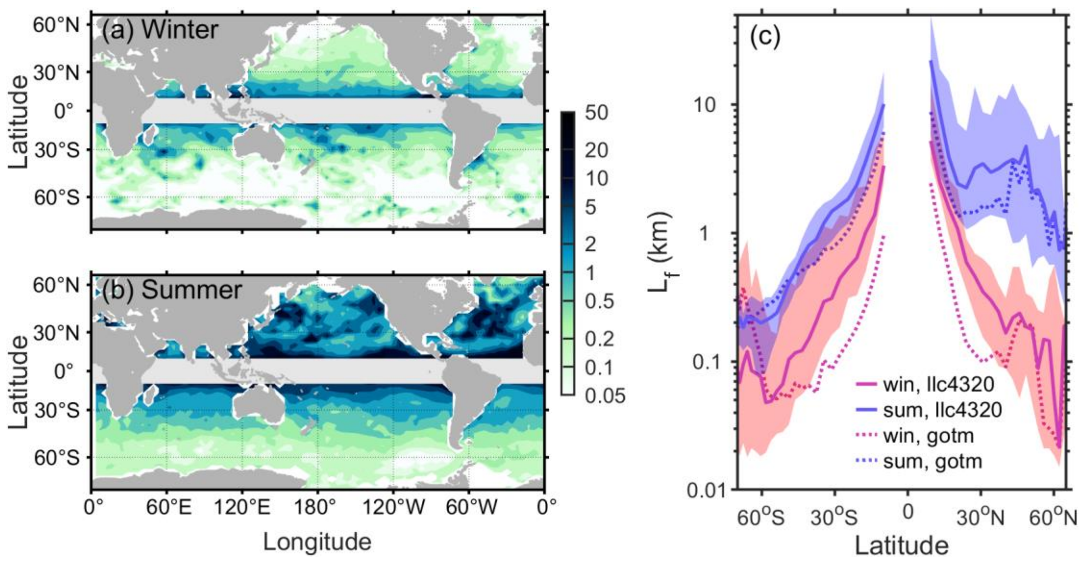

Figure R6 **Global distributions of the frontal scale L_f (km). a,** L_f in winter. **b,** L_f in summer. **c,** the zonal median L_f (winter in pink and summer in blue). The solid and dashed lines in **c** denote the values derived from the LCC4320 and GOTM results and the shaded intervals denote the corresponding bounds of the 10th and 90th percentile L_f values of the LLC4320 zonally.

Third, about the Buckingham et al. (2019) view, we believe their arguments are based on what is directly observed from the moorings. However, one important point that we should keep in mind is that their buoyancy gradients rely on the mooring spatial intervals. Are these buoyancy gradients really the frontal intensity in the real ocean? This remains an open question. Here in our work, we argue that the buoyancy gradients derived from the moorings are limited by the mooring spatial intervals based on the TTW theory and the frontal arrest scale proposed by Bodner et al. (2023). This is another key difference between our work and Buckingham et al. (2019).

Lastly, as discussed in Comment 3), we only consider down-front wind conditions in the revised manuscript. And we find that for the down-front wind conditions, the probability density function of our estimated production is much closer to the observed one (**Fig. S7**), which further supports the performance of our theory.

8) One other point that occurred to me is the situation with up-front winds will tend to restratify the mixed layer, thus damping turbulence with a positive buoyancy flux. This opposite effect seems to be excluded by the definition of the "Ocean Surface Boundary Layer" (OSBL) being above the 10^-8 W/kg dissipation level. With this definition of the OSBL, I would actually call it the "active mixing layer" or something similar. This is not to be confused with the surface mixed layer, which would include the positive buoyancy flux of the restratifying submesoscales (as seems to be noted on line 268). I think this is at least worth some discussion and justification of the choice made.

Response: Thanks for the comment. Thanks for clarifying the use of “mixing layer” and “mixed layer”. Here, we should claim that the “Ocean Surface Boundary Layer” that we use here is actually equal to “mixing layer”, rather than “mixed layer”. This is why we use a threshold value of the observed dissipation rates to get the OSBL depth. Furthermore, in response to your comments, we have decided not to include restratifying or stabilizing conditions, as evidence from observations and simulations of the stable boundary layer (e.g., Beare et al., 2006) indicate that it is not just the same scaling as for mixing run in reverse. Monin-Obukhov stability functions, etc., are well-understood to not be very different under mixing and stabilizing conditions and we know that the same is true for turbulence near fronts under upfront and down-front winds (e.g., Hamlington et al. 2014).

Minor comments:

1) Lines 59, 60 list references for the statement that fronts contribute a significant amount of turbulence to the boundary layer. The first two references show this, but both are at the wall of the Gulf Stream, and the second two are not relevant to frontal turbulence. More and balanced references (ie not just in the extreme location of the Gulf Stream) are needed for supporting this statement.

Response: Thanks for the comment. Yes, the first two are about frontal turbulence, but not all at the Gulf Stream. The first paper is at the Kuroshio Extension region, and the second one is at the Gulf Stream. The last two papers are focused on scaling the OSBL turbulence. However, deficiency or biases compared to observations are found from these classical scalings which fail to consider the contribution from the frontal turbulence. So, the last two papers are used to emphasize the first statement. To avoid potential confusion, we rewrite the statement. Meanwhile, we add another paper by Yu et al. (2019) who also found a significant contribution from the frontal turbulence at the OSMOSIS site. The statement is rewritten in the revised manuscript. (Line 59-60):

2) Fig. 1 would make much more sense to me visually if all axes have the same range. Then one could see from the position of the distribution what the relative contributions are. As it is now I can't even read the axis labels without a lot of difficulty.

Response: Thanks for the comment! We admit that the figure is a little difficult to understand. However, these three parameters have different magnitudes, and using the same axis will make the plot not at the center of the axes and so the figure looks weird. Anyway, we relabel the axis and try to make it clearer. The figure is revised in the manuscript and also see the copy as follows, and hope it will satisfy you.

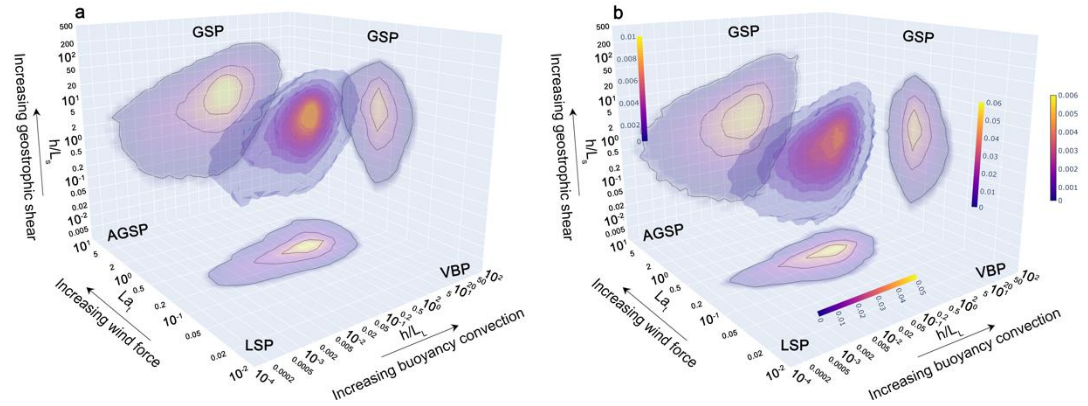

Figure R7 The copy of Fig. 1 in the main text.

Thanks again for your valuable comments!

References:

Beare, R. J., Macvean, M. K., Holtslag, A. A., Cuxart, J., Esau, I., Golaz, J. C., ... & Sullivan, P. (2006). An intercomparison of large-eddy simulations of the stable boundary layer. *Boundary-Layer Meteorology*, *118*, 247-272.

Belcher, S. E., Grant, A. L., Hanley, K. E., Fox-Kemper, B., Van Roekel, L., Sullivan, P. P., ... & Polton, J. A. (2012). A global perspective on Langmuir turbulence in the ocean surface boundary layer. *Geophysical Research Letters*, *39*(18).

Bodner, A. S., Fox-Kemper, B., Johnson, L., Van Roekel, L. P., McWilliams, J. C., Sullivan, P. P., ... & Dong, J. (2023). Modifying the mixed layer eddy parameterization to include frontogenesis arrest by boundary layer turbulence. *Journal of Physical Oceanography*, *53*(1), 323-339.

Buckingham, C. E., Lucas, N. S., Belcher, S. E., Rippeth, T. P., Grant, A. L., Le Sommer, J., ... & Naveira Garabato, A. C. (2019). The contribution of surface and submesoscale processes to turbulence in the open ocean surface boundary layer. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4066-4094.

D'asaro, E., Lee, C., Rainville, L., Harcourt, R., & Thomas, L. (2011). Enhanced turbulence and energy dissipation at ocean fronts. *science*, *332*(6027), 318-322.

Dauhajre, D. P., & McWilliams, J. C. (2018). Diurnal evolution of submesoscale front and filament circulations. *Journal of Physical Oceanography*, *48*(10), 2343-2361.

Dong, J., Fox-Kemper, B., Zhang, H., & Dong, C. (2021). The scale and activity of symmetric instability estimated from a global submesoscale-permitting ocean model. *Journal of Physical Oceanography*, *51*(5), 1655-1670.

Flexas, M. M., Thompson, A. F., Torres, H. S., Klein, P., Farrar, J. T., Zhang, H., & Menemenlis, D. (2019). Global estimates of the energy transfer from the wind to the ocean, with emphasis on near-inertial oscillations. *Journal of Geophysical Research: Oceans*, *124*(8), 5723-5746.

Gallmeier, K. M., Prochaska, J. X., Cornillon, P., Menemenlis, D., & Kelm, M. (2023). An evaluation of the LLC4320 global ocean simulation based on the submesoscale structure of modeled sea surface temperature fields. *Geoscientific Model Development Discussions*, 1-42.

Gula, J., Molemaker, M. J., & McWilliams, J. C. (2014). Submesoscale cold filaments in the Gulf Stream. *Journal of Physical Oceanography*, *44*(10), 2617-2643.

Hamlington, P. E., Van Roekel, L. P., Fox-Kemper, B., Julien, K., & Chini, G. P. (2014). Langmuir–submesoscale interactions: Descriptive analysis of multiscale frontal spindown simulations. *Journal of Physical Oceanography*, *44*(9), 2249-2272.

Haney, S., Fox-Kemper, B., Julien, K., & Webb, A. (2015). Symmetric and geostrophic instabilities in the wave-forced ocean mixed layer. *Journal of Physical Oceanography*, *45*(12), 3033-3056.

Li, Q., Reichl, B. G., Fox-Kemper, B., Adcroft, A. J., Belcher, S. E., Danabasoglu, G., ... & Zheng, Z. (2019). Comparing ocean surface boundary vertical mixing schemes including Langmuir turbulence. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3545-3592.

McWilliams, J. C., Gula, J., Molemaker, M. J., Renault, L., & Shchepetkin, A. F. (2015). Filament frontogenesis by boundary layer turbulence. *Journal of Physical Oceanography*, *45*(8), 1988-2005.

Reichl, B. G., & Hallberg, R. (2018). A simplified energetics based planetary boundary layer (ePBL) approach for ocean climate simulations. *Ocean Modelling*, *132*, 112-129.

Rocha, C. B., Chereskin, T. K., Gille, S. T., & Menemenlis, D. (2016). Mesoscale to submesoscale wavenumber spectra in Drake Passage. *Journal of Physical Oceanography*, *46*(2), 601-620.

Rocha, C. B., Gille, S. T., Chereskin, T. K., & Menemenlis, D. (2016). Seasonality of submesoscale dynamics in the Kuroshio Extension. *Geophysical Research Letters*, *43*(21), 11-304.

Sullivan, P. P., & McWilliams, J. C. (2018). Frontogenesis and frontal arrest of a dense filament in the oceanic surface boundary layer. *Journal of Fluid Mechanics*, *837*, 341-380.

Sullivan, P. P., & McWilliams, J. C. (2019). Langmuir turbulence and filament frontogenesis in the oceanic surface boundary layer. *Journal of Fluid Mechanics*, *879*, 512-553.

Thomas, L. N., & Taylor, J. R. (2010). Reduction of the usable wind-work on the general circulation by forced symmetric instability. *Geophysical Research Letters*, *37*(18).

Thomas, L. N., Taylor, J. R., D'Asaro, E. A., Lee, C. M., Klymak, J. M., & Shcherbina, A. (2016). Symmetric instability, inertial oscillations, and turbulence at the Gulf Stream front. *Journal of Physical Oceanography*, *46*(1), 197-217.

Thomas, L. N., Taylor, J. R., Ferrari, R., & Joyce, T. M. (2013). Symmetric instability in the Gulf Stream. *Deep Sea Research Part II: Topical Studies in Oceanography*, *91*, 96-110.

Wenegrat, J. O., & McPhaden, M. J. (2016). Wind, waves, and fronts: Frictional effects in a generalized Ekman model. *Journal of Physical Oceanography*, *46*(2), 371-394.

Yu, K., Dong, C., & King, G. P. (2017). Turbulent kinetic energy of the ocean winds over the K uroshio E xtension from Q uik SCAT winds (1999–2009). *Journal of Geophysical Research: Oceans*, *122*(6), 4482-4499.

Yu, X., Naveira Garabato, A. C., Martin, A. P., Gwyn Evans, D., & Su, Z. (2019). Wind-forced symmetric instability at a transient mid-ocean front. *Geophysical Research Letters*, *46*(20), 11281-11291.

Reviewer #2:

The revised version of Dong et al. has come some way to alleviating my initial concerns. They offer a reasonable explanation to why their results differ so significantly from those of Buckingham et al.. In addition, I do have a better idea of how the analysis was carried out (although I need even more clarification), and the authors have also stated the limitations of the analysis in the revised version. I still find the results interesting and surprising, and that the study can be a significant contribution to the field. That said however, I still have significant concerns as to what I see as the primary drawbacks of the study, that have now become more clear to me with the authors replies.

Response: Thanks for the general comment. Your concerns will be addressed in the following text, and we believe the revised manuscript clarifies the points you have raised.

Major comments:

1) Limited conditions of applicability for the analysis. The scaling of GSP is only applicable in down-front wind conditions and that for VBP in destabilizing surface buoyancy fluxes. This means the results are for 31% of winter and 21% of summer conditions (if I have understood the analysis, which I am still not convinced I have). Perhaps the authors can justify at the outset why they want to focus on these conditions, I think it would help very much with the applicability of the analysis.

Response: Thanks for the valuable comment. First, the work here focuses on the dynamic processes that lead to dissipation and OSBL deepening. These processes can be characterized using the turbulent kinetic energy (TKE) budget. However, restratifying due to processes such as the stabilizing surface buoyancy flux is a significantly different regime (e.g., Pearson et al., 2015) and cannot be described with an analogous approach in the TKE budget so is not considered here. This approach is consistent with prior publications (e.g., Belcher et al. 2012, Li et al. 2019) that did not consider fronts, and is why only the destabilizing surface buoyancy flux is considered.

Second, as discussed in the main text, a steady state is necessary but reasonable for the TKE budget equation. For an OSBL front, GSP is expected to occur when its PV is negative. However, a steady state of the balance between the sustained GSP and dissipation requires negative PV injection by sea surface forcing (Thomas, 2005). For an OSBL front without a sustained negative PV-injection force, but just negative PV in initial conditions, frontal instabilities will occur but rapidly restratify the front and arrive at marginal stability. In this case the GSP magnitude steadily decreases as the front restratifies, and a steady state cannot be reached. This is the basis of the SI parameterization scheme of Bachman et al. (2017), summarizing and synthesizing the results of many of the preceding LES and observational studies. Here, the Ekman buoyancy flux due to the down-front wind leads to negative PV injection and ensures sustained GSP. This amounts to a conservative estimate of the GSP contribution, as transient negative PV that might occur through a variety of mechanisms is neglected.

The two points above are why we focus on the conditions of down-front winds and destabilizing buoyancy fluxes. In the revised manuscript, we have clarified this on Lines 395-398.

2) All unresolved fronts are assumed to reach a very small horizontal scale that was simulated in a recent LES study, but has never been observed (I could not find such small scales even in the reference listed by the authors). This does not drastically change the results, but does put GSP into second place rather than third place (Table S1 with and without correction). The authors note that this amplification factor exceeds 6 at mid latitudes. Maybe this is correct, but it is at least not very well supported as far as I can tell (a single LES study).

Also, in their comparison to OSMOSIS observations the authors find a factor 4 times larger frontal width gives better results than the LES-suggested parameterization, and this value is then used in the analysis (different choice of C_L). On the one hand, it gives more conservative estimates, but on the other it is all quite arbitrary and not very justified, especially when these tuned values are then compared to observations to justify the correction.

Response: Thanks for pointing out your concern. Please allow us to explain point by point.

a) *About the robustness of the theory of the frontal arrest scale*

First, we must admit that it is a pity that no direct observations of the arrested OSBL fronts have been reported globally till now. An analysis of frontal statistics based on global observations is presently under review (M. Alford, personal communication). However, this could also be taken as indirect evidence for the small arrest scale of OSBL fronts. Only with high-resolution observations, such as by towed high-frequency observations, can such surveys be carried out. The bow-mounted thermistor chain technology recently developed at Scripps is changing this lack of observations of fronts from the meter-scale to km-scale. In the Bay of Bengal, Ramachandran et al. (2018) used such towed high-frequency observations to capture submesoscale fronts in the OSBL. According to their observations, these fronts have spatial scales around 2-5 km (see their Figure 3) and a weak adherence to geostrophy (see their Section 3b) yet the spatial scale and frontal arrest mechanisms are not discussed explicitly in their main text. As shown in Bodner et al. (2023; orange bar in Figure 2), the frontal widths from the theory agree well with the scale of observed fronts there. So, the comparison between Ramachandran et al. (2018) and Bodner et al. (2023) could be indirect evidence for the robustness of the frontal arrest theory, and other related observations are under analysis.

Second, using Large Eddy Simulations (LES), the work by Sullivan and

McWilliams (2018) explicitly claims in their abstract that the simulated filament shrinks and finally is arrested at a frontal scale of O(100) m (also see their Figure 17). It should be noted that the Coriolis parameter of their cases corresponds to a latitude of 32~33° and their horizontal resolution is ~ 1 m. Despite the result from Sullivan and McWilliams (2018) being derived from idealized LES, the magnitude of the arrested scale is close to the one from the frontal arrest theory at the corresponding latitudes [see Figure 3 of Bodner et al. (2023)]. In Bodner et al. (2023), they also validated the theoretically predicted frontal arrest scale with a variety of other Large Eddy Simulations under varying conditions, and they considered more forcing (such as Stokes drift) and obtained roughly consistent frontal arrest scales as well (see their Figure 1). Right now, no direct observation evidence for the frontal arrest scale can be provided, but the method is established based on the turbulent thermal-wind balance theory which is dynamically reasonable, and both indirect observations and LES results demonstrate the robustness of the arrest scale. It is for this reason that our paper is framed in the abstract and elsewhere as "based on theory predicting the frontal arrest scale" rather than based on observations.

Third, according to Equation 8, the rescaling of the horizontal buoyancy gradient requires a reference scale, but does not assume every front reach its arrested scale. The arrest scale provides the dynamically lower bound of the frontal width as the reference scale over which is the buoyancy gradient is rescaled.

Last, according to Table S1, the correction does not change the rank of the GSP place. GSP is in second place with a prevalence of 25% for the uncorrected case. GSP is still in second place for the corrected case despite an increase in the prevalence to 37%.

In the revised manuscript, we have added more discussion about the robustness of the frontal arrested scale (Lines 561-565).

b) *The choice of the factor C_L*

As discussed in Bonder et al. (2023), the parameter C_L is on the same order of magnitude as the Richardson number Ri, i.e., C_L ~ Ri. We choose a different C_L from the suggestion of Bodner et al. (2023), however the choice of C_L is not arbitrary here. In Bodner et al. (2023), shear turbulence is believed to be important according to LES simulations, shifting Ri to Ri ~ 0.25. But in the real ocean, the OSBL tends to stay near a neutral state with Ri ~ 1 and PV ~ 0 due to restratification processes (Bachman et al. 2017) and geostrophic adjustment (Tandon and Garrett 1994, 1995), especially at frontal regions focused here. So, the choice of Ri ~ 1 may be better supported by theoretical considerations. We believe this is why using Ri ~ 1 tends to reproduce dissipation rates closer to the observations at OSMOSIS. Conservatively speaking, the value of Ri in the OSBL is regionally dependent over the globe, and cannot be described by a constant value. Anyway, many observations suggests that Ri in the OSBL during convective and mixing events tends to stay at the range of [0.25 1] near the marginal stability of gravitational and symmetric instabilities. Given that it is still impossible to get the distribution of Ri over the globe based on observations, using Ri ~ 1 is an expedient choice. So, just as you comment, our estimate of the GSP magnitude in this work is a conservative lower bound. Nevertheless, even as a lower bound, the result still highlights the importance of the GSP turbulence.

In the revised manuscript, we have added more discussion about the choice of the parameter C_L (Lines 452-461).

3) There is a somewhat arbitrary depth dependence to the results. This was identified by the authors as an important factor in why they get different results to the Buckingham et al study, which makes a lot of sense. Buckingham

et al chose the depth of 45 m for their analysis, whereas the authors take the mid-depth of the mixed layer. While I do agree that the mid-depth is a more general choice, I am left a little unsatisfied at the generality of the analysis as a whole. We will get different results depending on our relative position within the mixed layer. This is just how the physics work, but it leaves me a little uncomfortable with statements in the text like:

- "GSP is found to be a leading contributor to turbulence in the OSBL and the prevalent one in winter"

- "the global contribution of energy transfer by submesoscales"

- "contributing 34% to the total dissipation in winter and 17% in summer"

It feels like the authors are talking about a vertically integrated measure of "turbulence", but they have chosen a specific depth (relative to the mixed layer) and only certain conditions (see point 2). I am concerned that readers will get a skewed impression of the "total dissipation".

Response: Thanks for the comment. Usually, using the OSBL depth to nondimensionalize the vertical direction is convenient for the study of the vertical structure of the OSBL dynamic processes. So dynamically, the mid-depth is a more general and rational choice for studying the vertical distribution of OSLB processes, and is consistent with how prior estimates of the sources of frontal turbulence (without fronts) have been calculated (e.g., Belcher et al. 2012; Li et al. 2019). We agree that some of our statements related to GSP are not rigorous enough. To avoid being potentially misleading, we have reworded these statements in the revised manuscript.

In the abstract, we have explicitly shown that our work is about the GSP contribution at the mid-depth of OSBL away from the sea surface (Lines 32-35). In the last paragraph of the introduction section, we have explicitly claimed the results are focused at the OSBL mid-depth away from the sea surface (Line 69; Line 75). Meanwhile, in the last paragraph of the discussion section, we have also highlighted that the result suggests a significant role of GSP turbulence in the exchanges between the OSBL and the ocean interior (Lines 284-286).

More specific comments:

1) The discussion of the observations with and without the GSP (submesoscale) component does not seem balanced to me (lines 419-429). Yes, it can be seen that without the GSP term the observed dissipation mean is slightly larger than predicted, but it seems just as over-estimated by the addition of GSP. Given the log-normal distribution, I also would expect to see some confidence intervals for these means. Do these differences between scaling estimates with and without GSP lie outside or inside the confidence intervals? The conclusion presented in the Methods of the manuscript is simply that the addition of GSP makes the estimates better. It also seems a little contrived given the factor 4 reduction in frontal length used to make the correction fit best, if I have understood what has been done.

Response: Thanks for the comment! For the first point about the comparison with observations, we want to use the mean values to argue that the means are largely determined by the high values, rather than to demonstrate the improvement by GSP turbulence. By using the bootstrap method, the mean production without GSP is 1.43 [1.25 1.75] $\times 10^{-7}$ W kg^{-1}, but is corrected as 1.87 [1.67 2.11] $\times 10^{-7}$ W kg^{-1}, compared to the observed value 1.62 [1.45 1.81] $\times 10^{-7}$ W kg^{-1}. The confidence intervals of these mean values do overlap. However, more importantly, due to the introduction of GSP, the PDF shape have been significantly improved. For example, the calculated skewnesses and kurtoses are 0.89 and 2.19 for the GSP-included PDF, which is much closer to the observed ones (0.9 and 2.25) compared to the PDF without GSP (1.07 and 2.75).

By considering both the changes in the mean value and the shape of the probability distribution curve, we can explain why the introduction of GSP has a significant impact on the curve's shape but only a minor effect on the mean value. At the OSMOSIS site, the intensity of GSP turbulence is moderate (the probability of high-value turbulent dissipation is consistent with observations even without GSP). Since the magnitude of the mean is often determined by the distribution of high turbulent dissipation values, the inclusion of GSP at this location does not significantly affect the mean value. However, excluding moderate-intensity GSP turbulence has a noticeable impact on the PDF shape. So, to avoid potential misleading about this point, we have removed the dots of the mean values in Fig. S8, and also provided a more quantitative comparison in Lines 474-477.

For the second point about the choice of C_L value, we have explained in the major comment 2 and also added more explanation in the main text (Lines 452-461).

2) How exactly are the production rates for GSP calculated? I feel I am getting closer, but still unsure. Is it that at each grid point at each time a horizontal buoyancy gradient is calculated (at mid-depth of mixed layer) and then if there are unstable buoyancy flux conditions and a component of down-front wind stress, the formulae listed in the Methods section are applied? Then, these conditions are found 31% and 21% of the time in winter, summer, and this is where our numbers come from?

Response: Thanks for the comment. We believe that your interpretation is exactly what we have done to calculate GSP turbulence. For each LLC4320 grid point, the sea surface buoyancy flux and angle between the wind and frontal current are calculated. Only at those grid points with a wind vector component of down-front winds and destabilizing sea surface buoyancy flux, the horizontal buoyancy gradients are corrected and then the TKE model is applied. In the revised manuscript, we have added statements to make it clear in Line 405-407.

According to the Large Eddy Simulations of an OSBL front by Thomas and Taylor (2010), the dissipation rate at the front is mainly balanced by the calculated GSP production with down-front winds ($-\overline{\boldsymbol{u}'w'}\frac{\partial \overline{\boldsymbol{u}_g}}{\partial z}$, cf. Equation 3 in the main text) which peaks near the surface at a value approaching the Ekman buoyancy flux due to the down-front wind component and follows a near-linear profile with depth in the surface boundary layer (see their discussion section). As a result, GSP is scaled as

$$GSP = EBF\frac{z+h}{h} \tag{R1}$$

Here, $EBF = \frac{\boldsymbol{\tau}\times\boldsymbol{k}}{\rho_0 f}\cdot\nabla_h b$ is the Ekman buoyancy flux due to down-front winds, *h* is the OSBL thickness, *z* is the water depth ($-h \le z < 0$). The expression can be nondimensionalized by using the wind forced production $u_*{}^3/h$, and the GSP production rate the OSBL mid-depth is obtained as shown in Equation 4 of Methods. In the revised manuscript, we have added more explanation about this in Lines 386-391.

About these numbers, your interpretation is also correct. They are the globally-averaged percentages of times with down-front wind component and destabilizing conditions over the whole months. For example, 31% means there are 31% of the time that meets the conditions of the down-front wind component and destabilizing in winter, and 31% of the time that the TKE model is applied. In the revised manuscript, we have added more words to explain how these numbers are derived (Lines 268-269).

Minor comments:

1) Switch order of integration in eq'ns (5,6) on LHS.

Response: Thanks for the comment. However, we believe the order of the integration is correct in Equations 5 and 6 after our careful examination. On LHS, it is an integration of the spatial distance, so L_{eff} and L_f should be the lower limit of the integration, which is different from the one on RHS. Note that this is an application of Parseval's theorem, as the left integrals are over distance and the right integrals are over wavenumber.

2) Typo in time periods on lines 434-5.

Response: Sorry for the typo. It should be 2013 for the observations and we have corrected it in the revised manuscript (Line 482).

Thanks again for these valuable comments!

Reviewer #3:

Using the re-calibrated outputs from the submesoscale-permiting global circulation model simulation (LLC4320) and turbulence parameterizations, the authors demonstrate that submesoscale (SMS) geostrophic shear production (GSP) is a significant source of turbulent kinetic energy (TKE) in the ocean surface boundary layer (OSBL). Along with the wind, waves and convection, turbulent dynamics at SMS fronts can provide an important route for the dissipation of the energy. Their finding aligns with recent regional observational studies and SMS-resolving numerical simulations in the past two decades. The message that authors try to send to the readers is clearly important and necessary at this time. The main concern is whether LLC4320 can be trusted to quantitatively deliver this important message. As pointed out by the two previous reviewers, there are two main points of skepticism about the analysis: (1) how well does LLC4320 represent the real ocean conditions? and (2) why does the authors' finding conflict with the analysis of Buckingham et al. (2019) who suggested that SMS dynamics is not a significant source of TKE relative to the wind, waves and convection? The authors have nicely addressed many aspects of the skepticism in their revision. However, a few issues remain as I elaborate below.

Response: Thanks for the general comment! Your concerns have been carefully addressed as below, and hope the response will satisfy you.

1) With regard to the first point of concern, the authors responded by pointing to the recent publication of Gallmeier et al. (2023) who used artificial intelligence techniques to compare the SST in LLC4320 with the observed SST from Level-2 Visible Infrared Imaging Radiometer Suite (VIIRS) dataset. Gallmeier et al. (2023) conclude that "the LLC4320 simulation reproduces, over a large fraction of the ocean, the observed distribution of SSTa patterns well,

both globally and regionally.” It should be noted that the conclusion is based on training a machine learning algorithm with the distribution of structures in SST anomaly of 10-80 km scales. The conclusion is not applicable for the SST structure smaller than 10-km which is the SMS range of interest in the present analysis. Furthermore, Gallmeier et al. (2023) also pointed out a “modest, latitude-dependent offset” between LLC4320 and VIIRS. It is unclear whether the authors have taken into account this offset.

Response: Thanks for the comment! The work by Gallmeier et al. (2023) is a piece of auxiliary evidence here to demonstrate the performance of the LLC4320 simulation. We totally agree that the argument from Gallmeier et al. (2023) is only applicable to SST structures larger than 10 km, and SST structures smaller than 10 km are near the effective resolution of the LLC4320, hence are not well resolved. This is precisely the point underlying the rescaling of the model buoyancy gradients we do here—to use the LLC4320 spectra only on scales larger than its effective resolution to infer the magnitude of fronts below its effective resolution based on the new theory of Bodner et al. (2023). In the manuscript, we note that our analysis based on the first-order structure function has also confirmed this result (Lines 340-347 and Fig. S6). The PDFs of the structure function from the LLC4320 simulation show consistency with the VIIRS observations at large scales but large biases as small scales below the effective resolution.

As a result, the frontal buoyancy gradient from the LLC4320 simulation needs to be corrected in most regions before being used to scale GSP. In this work, we rescale the buoyancy gradient by assuming a constant spectral slope from the effective model resolution to the estimated frontal arrest scale (see Methods: Buoyancy gradient rescaling). So, it is not the original buoyancy gradient that is used for our most accurate analysis but a rescaled one. However, we also provide the key results without this rescaling and just using the raw buoyancy gradients from the LLC4320 in the supplemental material.

Although quantitatively different than our best estimate, these raw gradient estimates also indicate that GSP is a non-negligible source of the mid-OSBL TKE.

About the offset of the LLC4320 simulation pointed out by Gallmeier et al. (2023), we do not consider this offset in our work. First, as you mention, Gallmeier et al. (2023) used artificial intelligence techniques to compare the SST in LLC4320 with the VIIRS dataset. They find that LLC4320 reproduces the observed distribution of SST patterns well both globally and regionally, despite a modest, latitude-dependent offset. Two reasons explain why we do not consider this offset. On one hand, a log-likelihood metric by the machine learning algorithm is proposed for the evaluation. Despite the fact that this parameter can qualitatively describe the LLC4320 performance, the explanation about its relation to physical dynamics to us seems inadequate which makes it impossible to correct this offset based on their results. And they directly compare SSTA fields, which is an indirect rather than direct reflection of the front intensity.

On the other hand, just as we mention above, the work by Gallmeier et al. (2023) is a piece of auxiliary evidence, the primary evidence here is our first-order structure function analysis. The first-order structure function analysis is explicitly focused on the spatial variance, which we believe is more suitable and convincing for the implications of LLC4320 accuracy for our work. According to the analysis, the performance of the LLC4320 simulation at the spatial scales above the effective resolution is demonstrated. Meanwhile, the underestimation of the frontal buoyancy gradient is also explicitly shown, strengthening the motivation for our rescaled best estimation procedure.

In conclusion, there are no perfect ocean models and each model has its own shortcomings, but we believe that our analysis result based on the first-order structure function combined with supporting evidence of the related skill from Gallmeier et al. (2023) strongly demonstrates the performance of the

LLC4320 in terms of OBSL fronts on scales larger than its effective resolution.

2) Besides referencing Gallmeier et al. (2023), the authors also compare the PDF of first order structure of SST between LLC4320 and VIIRS at different scales and they found that the difference between LLC4320 and VIIRS is significant at the small scales, and thus, requires correction for the lateral buoyancy gradient. Although the authors indicate that GSP remains to be significant whether or not they apply the correction for the buoyancy gradient (on line 258), it is clear that LLC4320 does not have the appropriate resolution for the quantitative analysis of the SMS dynamics. Besides SST, the OSBL thickness h is a crucial parameter used in their analysis and it is difficult to assess how well LLC4320 can capture h in the real ocean.

Response: Thanks for the comment.

a) About the capability of LLC4320 in simulating submesoscales

We agree that LLC4320 does not have the appropriate resolution for the quantitative analysis of submesoscale dynamics, whose resolution cannot completely resolve submesoscale processes (Dong et al., 2020; Dong et al., 2021), and is far away from resolving OSBL arrested fronts (Bodner et al., 2023). That is why the buoyancy gradient needs to be corrected. We believe there is a misunderstanding about the correction. The reason why the buoyancy gradient needs to be corrected is that the LLC4320 simulation fails to resolve submesoscale processes completely. Dynamically, the correction is an offline estimate to reproduce the buoyancy gradient associated with small-scale submesoscale fronts that are not resolved by the LLC4320 simulation, rather than amplifying the buoyancy gradient associated with the resolved submesoscale fronts.

The LLC4320 is a groundbreaking simulation using some of the world's fastest supercomputers to their limit. Were we to wait the century or so until the processes leading to resolved frontal arrest could be directly resolved globally (Hewitt et al. 2023), the study here would be of limited interest as GSP could be directly calculated in a computation of similar effort in that era to the LLC4320 without further theory or refinements. However, a century is a long time to wait for a conservative estimate based mostly on previously published methods like the one provided here!

According to the rescaling theory (see Methods: Buoyancy gradient rescaling), the buoyancy gradient is rescaled based on its spectral characteristics (i.e., the spectrum of the horizontal buoyancy gradient tends to be flat). As a result, in the wavenumber spectral space, the non-resolved part of the power spectrum can be extended based on this flat spectral slope from the effective resolution down to the frontal arrested scale. Then this spectral characteristic can be projected into the physical space (see Equations 5-8). Based on the method, theoretically, one can rescale the buoyancy gradient to any scale from any model resolution. For example, in Fox-Kemper (2011), they corrected the buoyancy gradient in 1° simulation to a 1/6 ° grid which is consistent in magnitude with the one directly simulated by a 1/6 ° model (see their Figure 2). This is a common approach used to implement submesoscale parameterizations into more coarsely resolved ocean models, implemented in roughly a dozen of the climate models used for the IPCC AR6/WCRP CMIP6.

In the revised manuscript, we have made the physical meaning of the correction more explicit (Lines 536-538).

b) About the OSBL thickness in LLC4320

We agree that it is an important parameter in our analysis. Here, the OSBL thickness *h* is a mixing layer determined by OSBL dynamic processes. So, how well an ocean model reproduces the real ocean *h* depends on our understanding of OSBL processes, and one can expect biases of *h* from model simulations due to the lack of our understanding. So far, it is still impossible to evaluate the LLC4320-simulated *h* based on direct observations, as Argo-based estimates (e.g., http://mixedlayer.ucsd.edu/) and climatological estimates (e.g., https://www.seanoe.org/data/00806/91774) fail to capture many of the small-scale variations that are part of our analysis. Nevertheless, comparing the mixed layer thickness is an alternative to demonstrate the capability of the LLC4320 simulation, since *h* should be close to it after temporally averaging.

To evaluate the performance of LLC4320 on the OSBL thickness, we compare it to the surface mixed layer from LLC4320 and Argo observations. Here, one must admit that all current models have biases in the simulation of the mixed layer, and no ocean model can fully quantitatively reproduce the real ocean mixed layer and the LLC4320 simulation is no exception. It can be found that the temporally averaged *h* is close to the mixed layer thickness in LLC4320 (the root mean squares of the bias are less than 5 m over the globe, similar to that of typical climate models that include parameterized submesoscale restratification; Fox-Kemper et al. 2011) which tends to simulate relatively deeper mixed layer depths compared to the observations, especially in winter months (when resolving submesoscale restratification processes is most important). The root mean square of the global mixed layer thickness bias is 13.5 m in February but 24.4 m in August. This may be partially attributed to the unresolved restratifying processes such as small-scale mixed layer instability and symmetric instability (Dong et al., 2020; Dong et al., 2021). Nevertheless, compared to the Argo observations, despite the quantitative bias, the global pattern of the mixed layer from LLC4320 resembles the observed one in different seasons (Figure R1; also Fig. S7).

In the revised manuscript, we have discussed the LLC4320 capability in simulating the mixed layer in Lines 349-361.

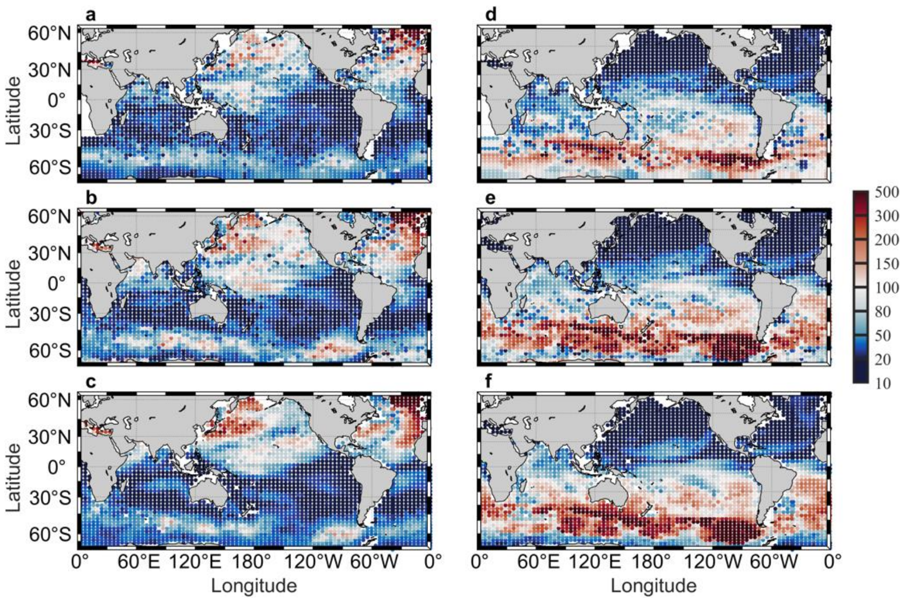

Figure R1 **Comparison of the OSBL and sea surface mixed layer thicknesses in different seasons (m)**. **a, b, c,** the thicknesses in February 2012. **d, e, f,** the thicknesses in August 2012. The upper, middle, and lower panels show the mixed layer thickness from Argo, LLC4320, and the OSBL thickness from LLC4320, respectively. The mixed layer depth is defined as the depth where a temperature variance of 0.2°C occurs compared to the 10-m depth temperature. The similarity of the global pattern demonstrates the capability of LLC4320 to reproduce the ocean surface layer.

3) With regard to the second point of concern, the authors pointed out numerous differences between the present analysis and that used in Buckingham et al. (2019) that potentially can lead to the conflicting conclusions. The two studies target two different depths: mid-depth in the former and fixed 45-m in the latter. Buckingham et al. (2019) used 24-hr time filtering and did not include the correction of buoyancy gradient, both of which potentially can lessen the significance of GSP. The simulation used in Buckingham et al. (2019) (namely, NATL60) was constraint to North Atlantic while LLC4320 is a global circulation model. While the justification for the different conclusions between the two studies sounds reasonable, the concern of whether the simulations (both LLC4320 and NATL60) can represent the real ocean conditions remains. If the authors carry out the analysis exactly as in Buckingham et al. (2019) (focusing only on the North Atlantic subdomain in LLC4320, targeting the same

depth, applying the same time filtering, etc.), would they obtain the same conclusion that GSP is insignificant relative to the wind, wave and convection? This exercise would help validate the use of LLC4320 and strengthen the conclusions in the present study.

Response: Thanks for the suggestion! We agree that a comparison between LLC4320 and eNATL60 will strengthen our conclusions in this study. There are two ways to compare them. One is that we conduct a comparison using LLC4320 following the idea from Buckingham et al. (2019) to see if GSP is insignificant as the reviewer suggest. However, this method is totally different from ours in this work, and is impossible to put the result in the manuscript. The other one is that we conduct analysis using eNATL60 to see if the GSP role in this work is robust to ocean models, which we believe is more consistent with the whole manuscript.

eNATL60 covers the North Atlantic with a spatial resolution of 1/60° and hourly outputs. It should be clarified that we only conduct the comparison in February since GSP is much stronger in winter. Meanwhile, the two datasets have different simulating periods and eNATL60 data in July 2009 ~ June 2010. Due to the different simulating periods, the PDFs of the non-dimensional dissipation rates from the four sources are compared.

As the density gradient is important for evaluating GSP magnitude, we first compare the power spectra between the two simulations. As shown in Figure R2, the two spectra are similar as their spatial resolutions are close, despite that the two models have different simulated periods (it is February 2012 for LLC4320 but 2010 for eNATL60). Their spectral slopes are also very close, and the slope from eNATL60 is slightly steeper. The similarity suggests that NATL60 is expected to produce GSP magnitudes similar to LLC4320, since the correction is based on the spectral characteristics.

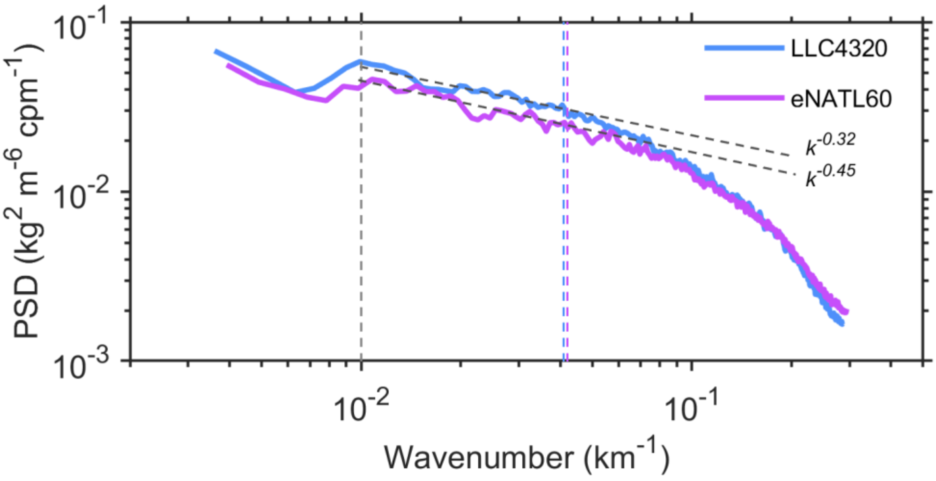

Figure R2 Power spectra of the horizontal density gradient. The blue and pink lines denote the results in the Gulf Stream region from LLC4320 and NATL60. The vertical dash lines show the 100-km wavelength (the scale where kinetic energy usually peaks and below which the kinetic energy begins to decrease) and the effective resolutions of LLC4320 and NATL60. The two simulations produce similar power spectra.

By applying the same methods, the magnitudes of the four turbulence sources are derived based on the eNATL60 data. Considering the different simulating periods, the non-dimensional probability density functions (PDF) of the four turbulence sources at the North Atlantic are compared. As shown in Figure R3 (also Fig. S9), the corresponding PDF structures from the two simulations resemble. It is noteworthy that both simulations show a distribution of GSP PDF between LSP and VBP, demonstrating the robust role of GSP turbulence which is insensitive to ocean models.

As additional information, a comparison of these two models at one site (Uchida et al., 2022) suggests that the numerics and subgrid schemes used in eNATL60 remove significantly more vortical submesoscale variability (leaving more fronts and fewer eddies) than LLC4320, so even though their nominal resolutions are similar, the small-scale phenomena are fairly different. Nonetheless, our analysis based only on gradients estimated above the effective resolution of each model provides robust estimates of GSP (Fig. S9).

In the revised manuscript, we have added more discussion about the comparison in Lines 490-498.

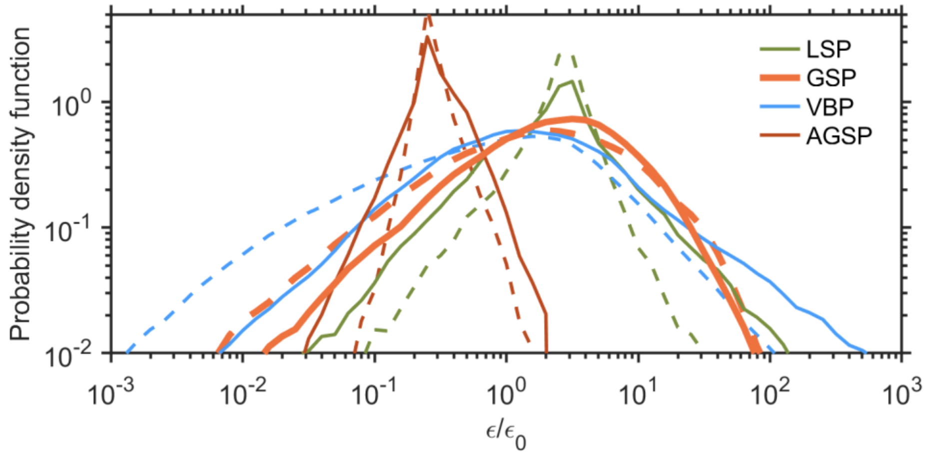

Figure R3 PDFs of the turbulence sources from different simulations. The dash and solid lines show the results from LLC4320 and eNATL60, respectively. The similarity of the PDF distributions demonstrates the robust role of GSP turbulence.

4) It should be emphasized that none of the subtle (but important) differences between the two studies is mentioned in the abstract. Would the conclusion about the GSP stated in the abstract only valid in the bottom half of the OSBL but not in the upper half at global scale? If that is indeed true (i.e., both the present conclusion and that of Buckingham et al. 2019 are realistic to real ocean conditions), then it should be clearly stated in the abstract so that there would be no misunderstanding about the different results between the two studies.

Response: Thanks for the suggestion. In the revised manuscript, we have explicitly shown that the conclusion in this work is valid within the OSBL away from the sea surface, and the quantitative contributions are at the OSBL mid-depth (Lines 32-35; Line 69; Line 75; Lines 284-286). However, we decided not to mention the difference between our work and Buckingham et al. (2019) in the abstract, as we have discussed the difference in the main text and the methods are totally different. While the comparison with Buckingham et al. (2019) is informative as to the specific measures and methods of our study, we do not see refuting that study's conclusions or inferences as the primary contribution of this work, rather it is the global estimates of the GSP contribution to the TKE budget.

Minor comments:

1) Line 372: AG is assumed to be 0.5 following the symmetric instability (SI) simulations in Thomas and Taylor (2010). Would SI be the only or the main mixing mechanism at arrested SMS fronts? For meandering SMS fronts with active mixed-layer instability (MLI), SI doesn't occur homogeneously in space as in the setup of Thomas and Taylor (2010). The LES of forced MLI in Hamlington et al. (2014) suggested SI occurs over localized regions of the front. Hamlington et al. (2014) found only 33% of the frontal area are exposed to symmetric instability. Would assuming the dissipation of GSP is only due to SI as in Thomas and Taylor (2010) over the entire frontal region overestimate the significance of GSP contribution in the real ocean? Also, I wonder the value of 0.5 in Thomas and Taylor (2010) was obtained by integrating the TKE budget over the OSBL or at mid-depth.

Response: Thanks for the comment. For meandering submesoscale fronts due to mixed layer instability (MLI), it is unsurprising that SI doesn't occur homogeneously in space. The work here investigates the forced SI which requires down-front winds. For front setup such as Hamlington et al. (2014), the front is completely destroyed by MLI. Under this condition, SI is not expected to occur everywhere over the front since winds are not always along the down-front direction. In Hamlington et al. (2014), the criterion for SI is negative Potential Vorticity (PV; see their Section 5). Their analysis is based on the LES result after 12-days evolution of the front. A negative PV source is required since the sea surface cooling is turned off after 10 days. According to Thomas (2005), the Ekman buoyancy flux due to down-front winds becomes the only source for PV injection. As a result, the frontal regions with SI based on negative

PV in Hamlington et al. (2014) should be regions with down-front winds. In conclusion, both results from ours and Hamlington et al. (2014) are consistent and actually derived under down-front conditions.

About the value of A_G=0.5, according to the Large Eddy Simulations of an OSBL front by Thomas and Taylor (2010), the dissipation rate at the front is mainly balanced by the calculated GSP production with down-front winds ($-\overline{\boldsymbol{u}'w'}\frac{\partial \overline{\boldsymbol{u}_g}}{\partial z}$, cf. Equation 3 in the main text) which peaks near the surface at a value approaching the Ekman buoyancy flux due to the down-front winds and follows a near-linear profile with depth in the surface boundary layer (see their discussion section). As a result, GSP is scaled as

$$GSP = EBF\frac{z+h}{h} \tag{R2}$$

Here, $EBF = \frac{\boldsymbol{\tau}\times\boldsymbol{k}}{\rho_0 f}\cdot\nabla_h b$ is the Ekman buoyancy flux due to down-front winds, *h* is the OSBL thickness, *z* is the water depth ($-h \leq z < 0$). The expression can be nondimensionalized by using the wind forced production $u_*{}^3/h$, and the GSP production rate the OSBL mid-depth is obtained as shown in Equation 4 of Methods. In the revised manuscript, we have added more explanation about this in Lines 386-391.

2) Inconsistent definition of OSBL thickness: “h is the OSBL thickness as determined by using an offline KPP scheme” (line 362) and “the OSBL thickness h is determined as the depth where the observed dissipation rate decreases to a threshold value of 1×10−8 W kg−1” (line 403-405). Please clarify.

Response: Thanks for the comment. Dynamically speaking, these two methods are consistent. Here, the OSBL we refer to is actually the mixing layer, characterized by strong turbulent mixing. Theoretically, both methods determine the thickness of the mixing layer based on the intensity of turbulent mixing. The difference between the two methods lies in the data used. If turbulent dissipation data is available, determining the mixing layer thickness based on the turbulent dissipation threshold is undoubtedly the better approach, as "mixing" is directly related to turbulent dissipation. This is why we use the turbulent dissipation threshold method to determine the mixing layer thickness at OSIMOSIS. However, for ocean models like LLC4320, which uses the KPP (K-Profile Parameterization) turbulence closure scheme, the thickness of the mixing layer is theoretically determined by the Richardson number (a parameter typically related to the generation of turbulent kinetic energy due to shear instability of the flow field) and does not output turbulent dissipation data. Therefore, to maintain dynamical consistency with the model results, we used an offline method to calculate the mixing layer thickness based on KPP. In the revised manuscript, we have made this point clearer in Lines 433-439.

3) Line 493: Is m_* set to be 0.5 as in Bodner et al. (2023)? If so, please include the value in the text.

Response: Thanks for the suggestion. Based on Reichl and Hallberg (2018), m_* is a variable that depends on surface buoyancy flux, friction velocity, and OSBL thickness, rather than being a constant. Therefore, in our calculations, we do not take m_* as a constant as in Bodner et al. (2023), but instead calculate it using Equations 29 and 36 from Reichl and Hallberg (2018). In the revised manuscript, we have made this explicit in Lines 553-555.

4) Line 518: “ ... while the GSP and horizontal shear production of the fronts themselves should contribute somewhat to the turbulence causing the arrest ...” Can the authors comment on the significance of the horizontal shear production with respect to the GSP contribution to the TKE especially in the limit of narrow (less than 100 m) arrested fronts? The horizontal shear production is not accounted for in the TKE budget (Equation 2).

Response: Thanks for the comment. To date, most studies on the TKE budget have assumed a horizontally homogeneous velocity field, thereby neglecting the horizontal shear production. As you mentioned, in the actual ocean, when the spatial scale is on the order of hundreds of meters, the horizontal shear production at a front can become significant and cannot be ignored. Based on the magnitude of the horizontal buoyancy gradient, M^2, we can discuss it briefly.

The magnitude of the geostrophic shear production can be expressed as

$$GSP = -\overline{\boldsymbol{u}'w'}\frac{\partial \overline{\boldsymbol{u}_g}}{\partial z} \sim \left|\overline{\boldsymbol{u}'w'}\frac{M^2}{f}\right|, \quad \text{(R3)}$$

while the magnitude of the horizontal shear production can be estimated as

$$HSP = -\overline{\boldsymbol{u}'v'}\frac{\partial \overline{\boldsymbol{u}_g}}{\partial y} - \overline{\boldsymbol{u}'u'}\frac{\partial \overline{\boldsymbol{u}_g}}{\partial x} \sim \left|\overline{\boldsymbol{u}'v'}\frac{1}{f}\frac{M^2}{L_f}h\right|. \quad \text{(R4)}$$

Assuming $u' \sim v' \sim w'$, the relative magnitudes of the two productions are roughly scaled as

$$\frac{HSP}{GSP} \sim \frac{h}{L_f} \quad \text{(R5)}$$

As the frontal scale is O(100) m, which is comparable to the OSBL thickness, h, HSP is expected to be non-negligible at OSBL frontal regions. In the revised manuscript, we have discussed the potential role of HSP in Lines 391-393.

Thanks again for these valuable comments!

References:

[1] Bachman, S. D., Fox-Kemper, B., Taylor, J. R., & Thomas, L. N. (2017). Parameterization of frontal symmetric instabilities. I: Theory for resolved fronts. *Ocean Modelling*, *109*, 72-95.

[2] Belcher, S. E., Grant, A. L., Hanley, K. E., Fox-Kemper, B., Van Roekel, L., Sullivan, P. P., ... & Polton, J. A. (2012). A global perspective on Langmuir turbulence in the ocean surface boundary layer. *Geophysical Research Letters*, *39*(18).

[3] Bodner, A. S., Fox-Kemper, B., Johnson, L., Van Roekel, L. P., McWilliams, J. C., Sullivan, P. P., ... & Dong, J. (2023). Modifying the mixed layer eddy parameterization to include frontogenesis arrest by boundary layer turbulence. *Journal of Physical Oceanography*, *53*(1), 323-339.

[4] Buckingham, C. E., Lucas, N. S., Belcher, S. E., Rippeth, T. P., Grant, A. L., Le Sommer, J., ... & Naveira Garabato, A. C. (2019). The contribution of surface and submesoscale processes to turbulence in the open ocean surface boundary layer. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4066-4094.

[5] Dong, J., Fox-Kemper, B., Zhang, H., & Dong, C. (2020). The scale of submesoscale baroclinic instability globally. *Journal of Physical Oceanography*, *50*(9), 2649-2667.

[6] Dong, J., Fox-Kemper, B., Zhang, H., & Dong, C. (2021). The scale and activity of symmetric instability estimated from a global submesoscale-permitting ocean model. *Journal of Physical Oceanography*, *51*(5), 1655-1670.

[7] Hamlington, P. E., Van Roekel, L. P., Fox-Kemper, B., Julien, K., & Chini, G. P. (2014). Langmuir–submesoscale interactions: Descriptive analysis of multiscale frontal spindown simulations. *Journal of Physical Oceanography*, *44*(9), 2249-2272.

[8] Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Griffies, S. M., Hallberg, R. W., Holland, M. M., ... & Samuels, B. L. (2011). Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Modelling*, *39*(1-2), 61-78.

[9] Gallmeier, K., Prochaska, J. X., Cornillon, P., Menemenlis, D., & Kelm, M. (2023). An evaluation of the LLC4320 global-ocean simulation based on the submesoscale structure of modeled sea surface temperature fields. *Geoscientific Model Development*, *16*(23), 7143-7170.

[10] Hewitt, H., Fox-Kemper, B., Pearson, B., Roberts, M., & Klocke, D. (2022). The small scales of the ocean may hold the key to surprises. *Nature Climate Change*, *12*(6), 496-499.

[11] Li, Q., Reichl, B. G., Fox-Kemper, B., Adcroft, A. J., Belcher, S. E., Danabasoglu, G., ... & Zheng, Z. (2019). Comparing ocean surface boundary vertical mixing schemes including Langmuir turbulence. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3545-3592.

[12] Pearson, B. C., Grant, A. L., Polton, J. A., & Belcher, S. E. (2015). Langmuir turbulence and surface heating in the ocean surface boundary layer. *Journal of Physical Oceanography*, *45*(12), 2897-2911.

[13] Ramachandran, S., Tandon, A., Mackinnon, J., Lucas, A. J., Pinkel, R., Waterhouse, A. F., ... & Farrar, J. T. (2018). Submesoscale processes at shallow salinity fronts in the Bay of Bengal: Observations during the winter monsoon. *Journal of Physical Oceanography*, *48*(3), 479-509.

[14] Reichl, B. G., & Hallberg, R. (2018). A simplified energetics based planetary boundary layer (ePBL) approach for ocean climate simulations. *Ocean Modelling*, *132*, 112-129.

[15] Sullivan, P. P., & McWilliams, J. C. (2018). Frontogenesis and frontal arrest of a dense filament in the oceanic surface boundary layer. *Journal of Fluid Mechanics*, *837*, 341-380.

[16] Thomas, L. N. (2005). Destruction of potential vorticity by winds. *Journal of physical oceanography*, *35*(12), 2457-2466.

[17] Uchida, T., Le Sommer, J., Stern, C., Abernathey, R., Holdgraf, C., Albert, A., ... & Dewar, W. (2022). Cloud-based framework for inter-comparing submesoscale permitting realistic ocean models. *Geoscientific Model Development Discussions*, *2022*, 1-32.

Reviewer #2:

(1) Lines 75-76: "... the prevalent one in winter." Please be more specific about which statistic is being referred to here. Is it the spatial prevalence (i.e. Fig. 4)?

Response: Thanks for the comment. It is the spatial prevalence. We agree that the statement is not specific enough. Meanwhile, we believe that the statement is a little repetitive and misleading since it has been claimed that GSP is found to be a leading contributor to turbulence. In the revised manuscript, we have shortened the sentence as (Lines 74-74): GSP is found to be a leading contributor to turbulence at the mid-depth of the OSBL in winter.

(2) eq'ns (5,6): Again, I am pretty sure here that the order of integration should be d\theta dr (the limits of integration are fine, as the authors mentioned in their previous response).

Response: Thanks for the correction! Yes! We agree that the expression is not rigorous enough. In the revised manuscript, we have reversed the order in Equations 5-7 as $d\theta dr$ (Lines 533, 535, 538).