Review of Perez-Santos et al. – BG

Presented is a study of turbulent mixing and dissolved oxygen fluxes in the Patagonian fjords region. The topic is highly relevant, and the large number of microstructure profiles coupled with dissolved oxygen measurements covering the entire Patagonian fjords region across three distinct seasons is impressive. However, I don't quite follow the chosen formulation of the dissolved oxygen flux, which lacks adequate flux units, and I find the presentation of the turbulence results lacking a robust assessment. The discussion and conclusion are limited to the region analyzed, and there is no attempt to broaden the findings to other systems. There's evidence of haste in the writing, and while reading, I was constantly distracted by inconsistencies in the presentation and/or missing details. Altogether, I believe the paper's analysis is a bit too light given the data's high relevance.

Major comments

1. Flux calculations
   Despite the given reference, it is hard to follow why you aim to model vertical oxygen fluxes (J) using the 2nd Fick's law. The quantity you report, d(O2)/dt, has units of mass/volume/time, whereas flux is the mass passing through a surface per unit time (i.e., with units of mass/area/time) and for vertical flux, it is modeled by the 1st Fick's law: J = – Kz·d(O2)/dz, where Kz is the vertical diffusivity. From your VMP measurements, you obtain coupled microstructure and fast oxygen measurements, which yield Kz and d(O2)/dz at high vertical resolution. I simply don't understand the choice to compute the derivative of J (although without the key negative sign in the formulation) in the vertical domain to obtain d(O2)/dt. Also, I'm not sure how to interpret this quantity—it would make sense in the control volume of a reactor, though. If the authors want to keep the current analysis, they should provide justification and have the editor find it sound. But in any case, I request that the authors also report fluxes using Fick's 1st law, as it would greatly benefit the broader community – you have all the information needed.

   Your definition of transport direction in terms of d(O2)/dt is also misleading because diffusive fluxes will occur against the gradient (i.e., from high to low concentrations) independently of how the concentration varies as a function of time. Regardless, I believe upward should be strictly toward the surface, downward toward the bottom, and the terms positive and negative should be dropped throughout the manuscript.

Given the type of measurements performed, I would have expected to find further calculations using the oxygen fluxes, particularly time scales. Using averaged flux estimates as boundary conditions for some of the hypoxic/anoxic deep basins and a simplified 1D solution of the diffusion equation (using an averaged diffusivity from your microstructure turbulence analysis), you could calculate time scales for substantial changes in oxygen concentrations due to vertical transport. I believe the same could also be achieved with a clever scaling. Comparing those time scales with the lateral intrusion time scales of modified water from the Jacaf channel into the Puyuhuapi fjord would be particularly interesting. This approach would allow you to discuss the relevance of vertical versus lateral mixing across the different semi-enclosed systems.

2. Lack of robust assessment of turbulence estimates and further analyses

   Spectral fitting is difficult and prone to analytical errors. My impression is that no effort was made to convince the reader that the data were treated to the highest possible standards. I understand that, since the paper was submitted to Biogeoscience, details of the microstructure procedures may have been omitted because they are not the journal's primary focus. However, I'd like to see some evidence in the Appendix showing you really know what you're doing. Answering some specific questions detailed below may help.

   - What binning was used for the turbulence microstructure analysis, e.g., 2-3 m, or was it a constant time window? Are the general characteristics of the turbulence profiles preserved if you increase or decrease the binning? Also, did you use any windowing? ATOMIX may suggest otherwise, but I would use an FFT window of 1/3 the bin length with a 50% overlap to obtain 3 spectra that are then averaged. The resulting averaged spectrum is then used to estimate epsilon using the Nasmyth procedure. This is done to ensure the reliability of the provided estimates.

   - Including some spectral fits in the supplementary material (for bins that pass and fail the spectral fit) would help the reader understand how finicky microstructure measurements are and therefore appreciate the effort put into assembling this dataset for the analysis of vertical transport of dissolved oxygen.

   - It might be helpful to share with the reader how much data was discarded during QA/QC procedures. I did not see any NaN values in the reported epsilon profiles of your open-access data, which made me suspicious about QA/QC. Also, highlighting how closely the estimates from the quasi-independent shear probes matched would be a nice addition. In general, if the agreement between the two independent estimates is within a factor of 10, one reports the average value; otherwise, one discards the higher value.

- I understand the paper is about diapycnal mixing, but from a superficial literature review, I gather that horizontal transport of different water masses sustains the hypoxic/anoxic layers in the Patagonian fjord system. It would be very nice if you could also quantify lateral dissolved oxygen fluxes due to isopycnal mixing (see the possible methodology in Fernandez Castro et al. 2025) and compare them with vertical transport through diapycnal mixing. I understand if this may feel like part of a completely different paper, but I am commenting because it could offer new insight. I believe it at least warrants discussion.

  Fernández Castro, B., Fernández Román, D., Ferron, B., Fontela, M., Lherminier, P., Naveira Garabato, A., et al. (2024). Isopycnal eddy stirring dominates thermohaline mixing in the upper subpolar North Atlantic. Journal of Geophysical Research: Oceans, 129, e2023JC020817. https://doi.org/10.1029/2023JC020817

3. Numerical model
   The model is disconnected from the paper. From my reading, it is mostly used to show that the Chiloe inner sea is energetic, and that explains the elevated turbulence estimates in that region compared to the rest of the dataset presented. However, since the model has been validated and published elsewhere, all this could be a sentence in the discussion or, at best, the material moved to the supplementary. Moreover, in Fig. 10c, you are comparing field estimates and model results for quantities with different units, and, as expected, since you are mingling scales, the quantities do not show agreement. Altogether, in my opinion, you would need to further develop your ideas and analyses to justify including the model results. One idea would be to use it to calculate residence times for layers with distinct oxygen levels across the different semi-enclosed basins to assess the relative importance of lateral transport versus vertical mixing. I have the impression that vertical transport may dominate in single-mouth fjords with a sill, but in the Jacaf-Puyuhuapi system, lateral oxygen transport from the adjacent deep, wide channel may dominate instead.

Minor comments

1. Broader context. While the introduction clearly and nicely states the relevance of mapping oxygen concentrations and understanding oxygen transport in the Patagonian fjords, the opening feels quite region-specific. I think the broader community would appreciate a brief review and contextualization of why the presented measurements and analyses are relevant to similar fjord systems (e.g., British Columbia, Scandinavia, and New Zealand). Moreover, I understand that Elin Darelius has been studying hypoxia in Norwegian fjords in recent years. I do not know whether that work has been published, but it would be interesting to compare.

2. For the transects performed within fjords, was any care taken to conduct entire transects during either ebb or flood tides? Measurements during different phases of the tidal cycle may encounter different currents that affect shearing layers and ultimately turbulence levels.

3. Fig. 1. I'm colorblind and can't discern the information displayed there clearly. Aside from the color choice, it looks cluttered and seems to include information that may not be necessary (e.g., lakes and snow). The text and symbols are hard to read in both panels. I'd recommend changing the color palette, reducing the text and information to what is absolutely relevant, and using buffered text to improve readability. Also, you can check whether your figures are suitable for individuals with different types of colorblindness at: https://www.color-blindness.com/coblis-color-blindness-simulator/

4. Table 1. Unclear. Alignment fails to provide information. Is it possible that the last row in the Measurements column, labeled 'Turbulence', is a typo? An in-text description is missing. Finally, were full transects performed at each campaign?

5. Lines 106-107: "…as well as two vertical shear probes (which measure mechanical turbulence in the water)". This is inaccurate; shear probes do not directly measure turbulence. They are used to estimate the turbulent kinetic energy dissipation rate through fitting of the empirical Nasmyth spectra. Please correct.

6. Lines 106 and 121. What do you mean by high-response? FP07 is a fast thermistor with a short response time (Fast Probe 07 [ms] with a nominally 10 ms response time). For the DO probe, is there a way to convert that % definition into an equivalent response time? I don't understand what is meant there.

7. Section 2.2. This is an incomplete and misleading explanation of how TKE dissipation rates are obtained from microstructure shear measurements. Some important details are provided around lines 112-118. Merging all the turbulence methods into one section would be appreciated.

8. Lines 148-152. It is stated here that eq. 3 is "the most frequently used formulation for Kshear" and Osborn (1980) is cited. This is incorrect. The most used formulation is by far that of Osborn (1980) using a constant diapycnal flux coefficient, Γ:

$$Kz = \Gamma\ \varepsilon/N^2$$

with Γ=0.2. Check the references below for further explanation of why this holds.

Monismith et al. (2018): https://doi.org/10.1029/2018GL077229

Smyth (2020): https://doi.org/10.1175/JPO-D-20-0083.1

9. Fig 2: Make sure you mention that these images belong to Rockland (RSI) if that's the case.

10. All transect figures would benefit from the inclusion of bathymetry. For example, in Fig. 3, I infer that the stations are the same (not mentioned in the methods), but the profiles' depths differ.

11. How do you calculate density? [If you use TEOS-10, has it been validated for Patagonian fjord waters?] Also, in the results section, you present density anomaly, but in Fig. 3, density is presented as neutral density, which is a completely different quantity. I guess this is just a labeling mistake.

12. Buoyancy frequency and Brunt-Väisälä frequency are used interchangeably, e.g., in eq. 3 and Fig. 4, but are not defined as equivalent. I'd recommend using buoyancy frequency throughout the manuscript. It would also be appreciated if you included a definition of N^2 and the method used to calculate it. Note that, in several profiles I checked in your open data, N^2 appears to be a linear function, and for the reported range of values, the units must be in cph rather than $1/s^2$.

    I also noted in your open data that some of the Kz profiles show spikes. I suspect some of those spikes are generated by locally unstable N^2 values that are being masked as very small values close to 0. A clever way to handle data when calculating Kz from dissipation and buoyancy frequency is to use an N^2 computed from a monotonically reordered (i.e., stable) density profile.

13. Figs. 4 and 5. I don't see a compelling reason to use a diverging colormap in these figures. There's no clear pivot, such as a sign change or a marked regime shift. I'd suggest using a perceptually uniform, sequential colormap for these plots (e.g., cmocean matter).

14. Figure 7. I really struggle to interpret this figure. For example, in the large bar near 200 m depth in panel (a), does this mean there are 0 values at 1.6 uM/s and 2 values at 1.9 uM/s? There must be a better way to show the seasonal differences noted in the in-text description.

15. Figure 8. After struggling to follow the lower panel of Fig. 8 and reviewing your open data, I realized that Figs. 8q, r and s are identical to Figs. 8l, m and n. That should be corrected. Also, using the provided open data, e.g., the profile from Fig. 8p, I could not reproduce the reported Kz and d(O2)/dt profiles because the N^2 data are spurious (see minor comment 12 on N^2). Finally, what are those intermittent black lines in all the Kz profiles? Whatever the reason, this can definitely be handled better for presentation.

16. T-S diagrams color coded for O2 concentrations and J = – Kz·d(O2)/dz for single-mouth fjords (grouped) and the Jacaf-Puyuhuapi system to compare the different water masses that may intrude into the basins vs diapycnal mixing would be interesting additions.

17. Lines 458-466. I can't quite follow the proposed idea that double diffusion convection may play a role in diapycnal mixing processes of oxygen throughout the water column. You are working in a high-turbulence region and the resulting mixing would smooth out any signature of DDC at short timescales. The state of the art is quite categorical in that double diffusion may take place when background shear is negligible (Gregg et al. 2018), so the difference between the diffusivities of temperature and salinity remains 2 orders of magnitude. (i.e., at the molecular level). DDC would be expected in deep basins with low turbulence levels, but not in the upper regions as suggested in Fig A1,b and c. Layering in natural environments is ubiquitous. (e.g., Middleton et al. 2025). Turner angle alone is misleading because it only provides an after-the-fact indication that offers no information about the formation mechanism. Unless you can provide a theoretical framework that can help reconcile what you are stating about DDC, I think this bit does not belong here. You may, however, be in the presence of differential diffusion.

M.C. Gregg, E.A. D'Asaro, J.J. Riley, E. Kunze. 2018. Mixing Efficiency in the Ocean. Annual Review of Marine Science 10:443-473.
https://doi.org/10.1146/annurev-marine-121916-063643

Middleton L, Brown JM, Taylor JR. 2025. Reconciling layering mechanisms in double-diffusive and single-diffusive fluids. Journal of Fluid Mechanics 1019:R1.
https://doi.org/10.1017/jfm.2025.10539