Response to Reviewer 2 for "Eight years of continuous Rockall Trough transport observations using moorings and gliders"

We thank the reviewers for their constructive comments which help us improve the quality of our manuscript. Below, we provide detailed responses to each comment.

During the revision, we are implementing the following improvements to the transport calculations, which we are making independently to strengthen the analysis:

- **Extending the mooring dataset to 10 years**, now overlapping with the entire glider observation period. We are changing the title of the manuscript to reflect this update: "A decade of continuous Rockall Trough transport observations using moorings and gliders"
- **Correcting the EOF analysis** by using the original time steps of glider transects instead of 15-day averages, which previously included an irregular number of transects.
- **Correcting an error in the new methodology:** The EOF analysis and regression are applied to velocity anomalies. In the earlier version, we mistakenly subtracted the glider mean at EB1 and RTADCP positions from EB1 data and GLORYS12V1 output instead of subtracting the mean of each respective dataset. This introduced a systematic offset, which has now been corrected. The glider mean field is added at the final step to define the mean of the reconstructed section, eliminating the need for bias correction of the GLORYS12V1 output.

Recommendation:

Requires major revision before acceptance to "Ocean Science". Suggested improvements and key points are listed below.

Main points requiring revisions:

This study presents an updated methodology for analysis of observational data from Rockall Trough. The new methodology is supposedly better than existing methods, but there is no hard evidence that the new version is indeed an improvement. I recommend adding analyses that quantitatively show why this is the case. To do this, the following approaches come to my mind (but the authors may choose differently):

- One could test the different methods in a numerical model that has sufficient resolution and fidelity in the Rockall Trough (validate the model first - does not need to be GLORYS).

In the model world, the actual flow (of volume, heat, and freshwater) is known, and the observing system results can be simulated (in an OSSE type analysis).

Thank you for these suggestions. Houpert et al. (2020) performed a comprehensive error analysis. We are adding a subsection in the data and method section of the manuscript. Additional details are provided in the expanded comments below.

- If the biggest improvement is the better time resolution compared to earlier methods, one could also compare spectra of the resulting time series from old/new methods. If the new method is really superior, the shape and power levels in the spectra should demonstrate what kind of aliasing is now avoided, and what the presumably lowered noise floor at high frequencies looks like.

  Thank you for highlighting this point. We would like to clarify that both the old and the new methodologies for reconstructing the eastern section use the same temporal resolution, based on mooring data and GLORYS12v1 output at the RTADCP location.

  Since 2020, we have additional glider observations that provide detailed insight into the structure of the European Slope Current (ESC). However, these observations remain scattered in time and are therefore prone to temporal aliasing. For the first time, mooring and glider observations overlap, allowing direct comparison. We find that the old method agrees well with glider transports, but it fails to realistically reproduce extreme events, which are likely driven by enhanced mesoscale activity.

  During this review, we identified an error in the new approach. After revision, the updated method now successfully reproduces these extreme events (see Figure 1, page 3 in our response to Reviewer 1) by including the second EOF mode, which we interpret as representing mesoscale eddies. We agree with the reviewer that a spectral comparison of the time series would provide valuable additional information, and we are including this analysis in the revised manuscript.

The study does not quantify uncertainties in the resulting time series. Error bars, in the few places where they exist, denote the error of the mean when averaging e.g. several years into an annual cycle, but this says nothing about the uncertainty of the original time series itself. Can we believe the smaller wiggles in figure 5a, or are they instrumental noise?

Houpert et al. (2020) assessed the error arising from the horizontal extrapolation of the current meters and reliance on ocean reanalysis at the eastern wedge against lADCP data from cruises along the section. For the EW transports, they found a total mean bias error for the method of -0.21 Sv and a Root-Mean-Square Error (RMSE) of 0.59 Sv (see their SI text S1.3, Table S1, Figure S6 and S8). As mentioned above, we are adding information about the transport accuracy in the data and method section of the manuscript.

Is the ~0.8 Sv misfit (l. 150) the dominant contribution to uncertainty, or is it the sensor errors that are listed for glider and mooring CTD - how many Sv result from these sensor errors?

Thank you for raising this point. The ~0.8 Sv misfit was primarily caused by an error in the previous implementation of our methodology rather than by sensor inaccuracies. Specifically, the EOF analysis and regression were applied to velocity anomalies, and the glider mean field was added at the final step to define the mean of the reconstructed section. This approach eliminates the need for bias correction of the GLORYS12V1 output.

The misfit occurred because, in the earlier version, we mistakenly subtracted the glider mean at EB1 and RTADCP positions from EB1 data and GLORYS12V1 output instead of subtracting the mean of the respective datasets themselves. This introduced a systematic offset, which we have now corrected. After this revision, the updated method aligns well with glider observations and no longer exhibits the previous misfit.

I recommend including a systematic accounting of the major sources of uncertainty, and adding those to the relevant figures and results. Just copying the manufacturer's sensor specifications and not propagating them into the resulting volume, heat, and freshwater fluxes is, in my opinion, insufficient.

Thank you for these suggestions. Houpert et al. (2020) conducted a comprehensive error analysis of the transport estimates using a Monte Carlo approach to assess the impact of instrument errors, and evaluated methodological uncertainties using LADCP and CTD data from the Ellet Line hydrographic sections, alongside climatology data from MIMOC (Schmidtko et al., 2013).

In the mid basin, the Monte Carlo simulations showed the combined effect of pressure, temperature, and salinity resulted in a RMSE of 0.05 Sv. For the western wedges, current meter inaccuracy led to a maximum transport error of ±0.12 Sv at a 68% confidence level.

Methodological errors were assessed by comparing transport estimates derived from full LADCP/CTD sections with those derived from subsampled mooring data using the same processing methods. Errors were attributed to gridding and vertical extrapolation for the mid basin (bias error of 0.11 Sv, RMSE of 0.34 Sv for sull data return), gridding and horizontal extrapolation for the western wedge (bias error of -0.30 Sv, RMSE of 0.63 Sv) and horizontal extrapolation and reliance on ocean reanalysis for the eastern wedge (bias error of 0.21 Sv, RMSE of 0.59 Sv).

Mean bias and RMSE for the total transport were obtained by combining the errors of all three subsection. For optimal data return, the total RMSE was 0.93 Sv and the mean bias error was 0.03 Sv. However, data loss due to instrument failure in 2014-2015 and 2016-2017 increased uncertainty in the mid basin transports resulting in a bias error of -0.39 Sv and 0.38 Sv as well as a RMSE of 1.10 Sv and 0.93 Sv, respectively. For further details, we refer readers to Houpert et al. (2020). We are adding a subsection about the error estimates in the data and method section in the manuscript.

The methodology for heat and freshwater transports is flawed in that underlying data are not available at sufficient spatial resolution, let alone with appropriately co-located measurements. As a bare minimum, there needs to be a validation why the method should still give correct results. This could (again) be done with an OSSE-type numerical simulation. If these validations have been done in some of the referenced work, they need to be summarized here. I did not review the section with heat and freshwater transports at this time.

Thank you for your suggestion. To assess whether mooring-based hydrographic profiles are sufficient for estimating heat and freshwater transports, we compare transports derived from full ship and glider sections with those calculated using ship and glider profiles subsampled at mooring positions (Table 1). For the western and eastern wedges, mean bias errors were small (1–6% of the mean transport), and RMSE values were well below one standard deviation (4–8%), except for freshwater transport in the eastern wedge (19%). For the mid basin, errors were larger (bias: 22–31%; RMSE: 25%), but given the high natural variability, we consider these results as acceptable. Full details are provided in Table 1 and we are adding a subsection in the data and method section of the manuscript.

Table 1: Comparison of heat (Q_h) and freshwater (Q_f) transports estimated from full temperature–salinity sections versus profiles at mooring positions. "Mean full" and "Std Dev full" represent the mean and one standard deviation calculated from complete ship sections (western wedge, mid basin) or full glider sections (upper 1000 m, eastern wedge). "Mean profile" and "Std Dev profile" represent calculations using data only at WB1/2 and EB1 positions from ship sections (western wedge, mid basin) or glider data (upper 1000 m, eastern wedge). Mean bias error and root-mean-square error (RMSE) between full-section and profile-based heat and freshwater transports are also shown. Only ship sections covering all Extended Ellet Array stations in the western wedge or the mid basin, respectively, were used.

	Western wedge 10 ship sections		**Mid basin** 8 ship sections		**Eastern wedge** 166 glider sections	
	Q_f (10^{-2} Sv)	**Q_h (10^{-2} PW)**	**Q_f (10^{-2} Sv)**	**Q_h (10^{-2} PW)**	**Q_f (10^{-2} Sv)**	**Q_h (10^{-2} PW)**
Mean full	0.23	-0.43	-2.79	6.20	-0.43	1.19
Mean profile	0.24	-0.42	-3.66	7.59	-0.44	1.12
Mean bias error	-0.01	0.00	0.86	-1.39	0.01	0.07
Std Dev full	1.57	3.33	5.62	11.04	0.66	1.67
Std Dev profile	1.62	3.42	5.12	10.13	0.67	1.60
RMSE	0.08	0.12	1.46	2.90	0.12	0.14

There is an inconsistency or flaw in the method, in that it uses EOF patterns but use of patterns higher than order one fails to improve the results in the validation step. I suspect this is due to how observational and model data are mixed in the methodology. There are comments about this below, as well as suggested steps to address this.

Thank you for this observation. We do not consider this a flaw in the methodology but rather a limitation related to the available data coverage. When the subsampled data (EB1 observations and GLORYS12V1 output at the RTADCP location) do not capture the main features of an EOF mode—such as being in an area of minimum amplitude—the corresponding regression coefficient will be small, and the mode cannot be reproduced accurately. In such cases, including higher-order modes does not improve the reconstruction and may introduce noise. Therefore, only modes whose dominant features are represented at the mooring locations contribute meaningfully to the transport estimates.

Our approach is designed to minimise reliance on model output, using it only where observational gaps cannot otherwise be filled. We are adding a paragraph in the discussion section to clarify this limitation and provide additional context in the comments below.

The overall presentation needs to be polished; I am including detailed comments below. For a publication-ready manuscript, I expect more consistency with labels, abbreviations, etc., across the figures and text.

Thank you for pointing this out. We are revising all figures and text to ensure clarity and consistency throughout the manuscript. This includes harmonising labels, abbreviations, and formatting across figures and captions, as well as improving readability in the main text.

Improvements to content:

Ll. 47-49:

In the transition between the paragraphs here, it is not clear whether this study solves the shortcomings of the glider observations, or reveals what they were in the first place. Maybe change the wording in the paragrph lines 49 ff. to make this clear.

Thank you for pointing this out. We are revising the entire introduction for clarity.

Ll. 47 and 85:

These state that the glider data are "scattered and sparse in time". It is not very clear what this means, nor what would be considered "good enough". Can this be clarified/quantified somehow? For the introduction, this is probably OK as is, but in the methodology section, I recommend being more quantitative.

Thank you for pointing this out. We are revising the text for clarity.

Ll. 61 ff.:

The glider CTD sensors are essentially the same as the mooring ones. You are listing what appears to be the factory specifications for sensor accuracy here. For the moorings described later, you list similar accuracies but presumably requiring the tedious calibration procedures referenced there (line 97; McCarthy et al. 2015). Are the glider data processed with similar methods? If not, they will not be as accurate as described.

Thank you for your comment. The accuracy values provided refer to manufacturer specifications and differ from those listed later for the moorings. Glider CTD sensors and compasses were calibrated in the laboratory before each mission, and an in-water compass calibration was performed at the start of each deployment. However, the processed glider data do not achieve the same accuracy as the fully calibrated mooring sensors. Fraser et al. (2022) found that glider temperature profiles at EB1 agree well with moored profiles, while salinity is consistently underestimated by approximately 0.02 g/kg. Importantly, Fraser et al. (2022) also showed that this salinity bias has minimal influence on geostrophic velocity shear and associated transport estimates. We are editing the text for clarity.

Ll. 93-95:

It is not clear how the WB1 and WB2 moorings are concatenated. The reference quoted in turn refers to another reference (McCarthy et al., 2015), which also leaves details somewhat open. There is a "correct" way to do this merger, assuming geostrophy and that the current and CTD data are actually available: One can start by integrating the CTD-derived specific volume anomaly upwards from the bottom of the deep mooring. Then, one can "jump" horizontally to the shallower mooring using the geostrophic equation and the currents at the depth of the jump (supposedly some average from nearby current meters on both moorings can be used). Then, continue integrating the CTD-derived data up along the shallower mooring. Is this what was done?

Thank you for pointing out we referencing to the wrong paper. We use the approach of Fraser et al. (2020), not Houpert et al. (2018).

For the mid basin section we generate WB1/2 by using potential temperature and absolute salinity values from WB1 above 1600m, and from WB2 below 1600m. We justify this given that the isopycnals are near horizontal between WB1 and WB2 meaning the transport here would not be captured by baroclinic shear. Below 1600 the meridional velocity values of WB2 current meters are used to fill in the region east of WB1/2. The flow over this small area is weak and contributes negligible to transport (Fraser et al., 2020). We are editing the text for clarity.

L. 99:

What is meant by correcting the velocity measurements for "sound"? The instruments are sonic current meters, but what corrections need to be made?

Thank you for pointing out the typo. The current meters are corrected for speed of sound using actual measurements from the nearby hydrographic mooring data. We are editing the text accordingly.

Ll. 107 ff.:

Clarify whether the RTADCP will be used, or why it is worth mentioning here. Else, remove.

Thank you for highlighting. We use the RTADCP to bias-correct the GLORYS12V1 output. The bias-correction of GLORYS12v1 output is only required in the old methodology. We are editing the data and method sections in the manuscript for clarity.

Ll. 138 ff.:

If the Hilbert EOF analysis is not used, I think this paragraph that refers to it can be removed.

Thank you for pointing this out. Our methodology is based on Brandt et al. (2014, 2016, 2021), who applied Hilbert EOF (HEOF) patterns to better capture vertical displacement of current cores. In our study, we evaluate both EOF and HEOF approaches and obtain similar results. For simplicity, we chose to use standard EOF patterns. We are clarifying this choice and are providing additional explanation in the data and methods section and we are removing the paragraph in the results section.

Ll. 140 ff.:

There is a fundamental inconsistency in the method used here, which needs to be reconciled:

For input parameters, the regression method uses a mix of observational data (from the EB1 mooring) and numerical model data (from the GLORYS analysis at a single nearby but separate location). The observational and numerical data reflect two different "realities" that may be inconsistent with each other. There is no quantitative reasoning supporting the choice made here over other available choices, other than the choice of using only observational data from EB1 (which is shown to be inferior). In order to justify the choice made, I recommend additional analyses:

- Validate that the EOF patterns and magnitudes in the model world reasonably match the observational ones, i.e. recreate figure 3 from GLORYS data alone.

- Validate the model against the existing EB1 and glider observations.

- Use the model transport alone (at full model resolution in space and time), and quantitatively test whether this is inferior to your choice.

- Try to find some combination of input from the model (other than the velocities at the single RTADCP location, but simpler than using everything) that might optimize agreement with your reference data.

Thank you for your suggestions. The use of GLORYS12V1 output at the RTADCP location for estimating transport in the eastern wedge was thoroughly tested by Houpert et al. (2020) and calibrated against eight months of ADCP observations at that site, which are not available elsewhere along the section. For the full eastern wedge, they reported a bias error of 0.21 Sv and an RMSE of 0.59 Sv. As noted earlier, our methodology is designed to minimise reliance on model data, using it only where observational gaps cannot otherwise be filled.

Given the strong agreement between glider observations and the previous reconstruction method (Figures 1a; see also Figure 1, page 3 in our response to Reviewer 1), which are independent data products, together with the findings of Houpert et al. (2020) and Fraser et al. (2022), we consider the choice to use GLORYS12V1 output at the RTADCP location justified. As shown in Figure 2b and 2c, the new methodology still depends on this model input to achieve a realistic representation of ESC transport.

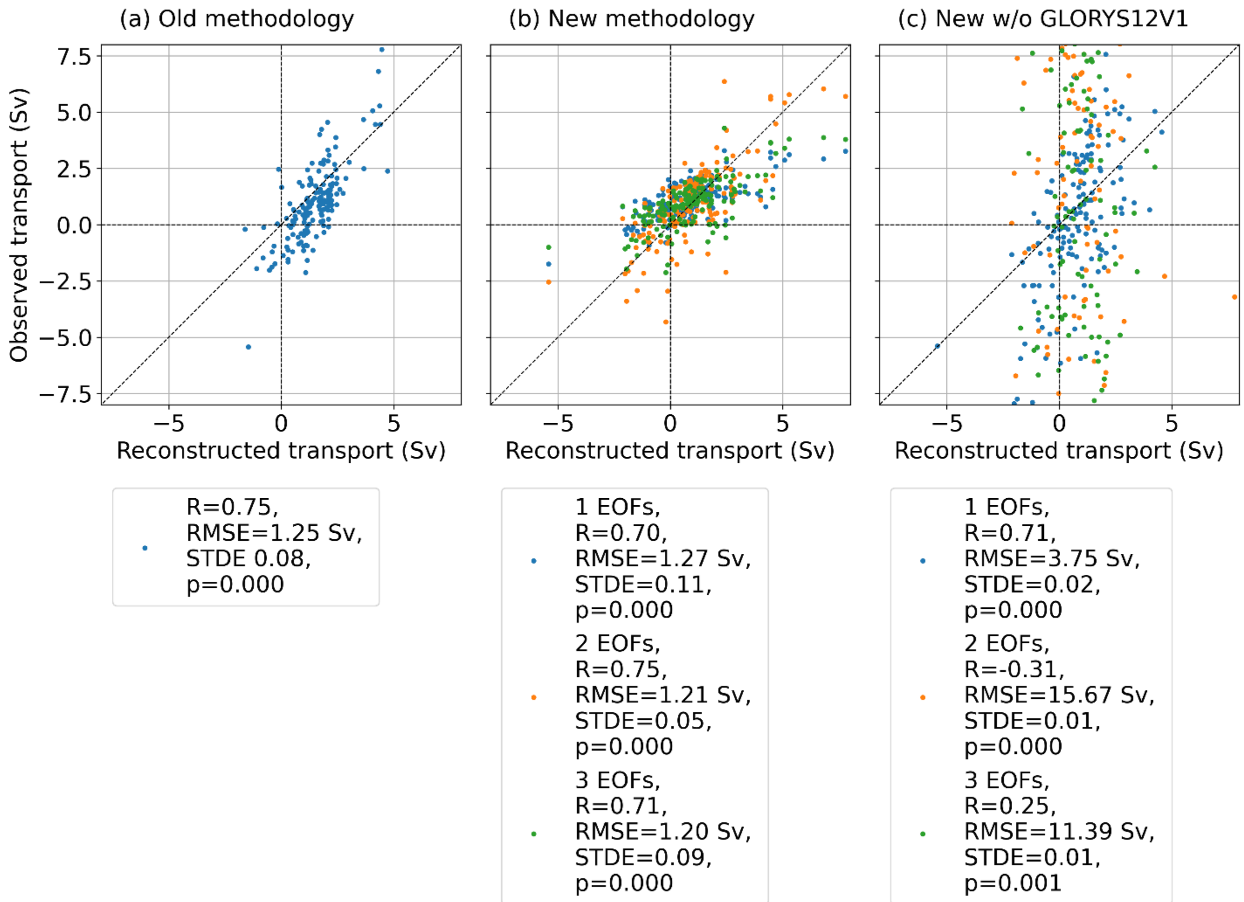

Figure 1 Linear regression of mooring-based reconstructions versus glider-derived transports. Panels show (a) the old methodology (Fraser et al 2022; Houpert et al. 2020), (b) the new methodology using EB1 observations and GLORYS12V1 output at RTADCP, and (c) the new methodology using only EB1 observations. Statistics shown: R = Pearson correlation coefficient; RMSE = Root Mean Square Error; STDE = Standard Error; p = two-sided p-value for the null hypothesis that the slope equals zero.

Ll. 155 ff.:

I am confused - you are only using the first EOF mode to reconstruct the velocity field, correct? Why does your figure 4b not look like figure 3 (mode 1) then? Shouldn't they be more similar?

Thank you for highlighting a lag of clarity here. It is common practice to remove the temporal mean before performing an EOF analysis as we are interested in the spatial pattern of its variability and not the mean field itself. The EOF pattern shows the dominant mode of variability. Before the regression we remove the temporal mean from the EB1 and GLORYIS12v1 data, reconstruct the section based on the EOF pattern and add the mean of the glider section to the reconstructed velocity anomalies. We are editing the method description for clarity.

Ll. 177-178:

I understand that the mid basin transport is not the primary focus of this study, but since you are mentioning it here and in figures 6-7: The way I read the reference, the mid basin transport is calculated using an assumption of no motion at ~1800 m. This is not consistent with how I read figure 1(b), in that there is more "red" at depth than "blue". I can only assume that there is a substantial amount of variability at that depth. In order to quantify the error from this reference level assumption, can you provide the time series of the velocity (averaged between WB2 and EB1) at 1800 m from the 17 ADCP sections from figure 1(b)? Multiplying this with the water depth and the section width will show you the error in terms of volume transport. I would not be surprised if that error were as large as your entire signal in figure 5. If I am doing the math right, 5 mm/s velocity variability will translate to 1 Sv error, but please double-check and provide the actual number.

Thank you for highlighting the importance of the reference level in the mid basin transport calculation. We agree that the choice of reference level significantly affects the transport estimates. We adopt the method from Houpert et al. (2020). This method yields a basin-wide transport below 1,250 m of approximately −0.3 Sv, aligning with prior findings that deep northward flow is blocked by topography (Holliday et al., 2000), allowing only a small net southward transport of dense Wyville Thomson Overflow Water (-0.3 Sv, Johnson et al., 2017). For the area below the reference level of 1760m, a current of 5mm/s translates to an error of 0.2 Sv. As noted above, Houpert et al. (2020) estimated a total RMSE of 0.96 Sv for the mid basin transport, accounting for both instrumental and methodological uncertainties. We are adding additional information in the data and method section.

Ll. 184 ff.:

I think there is a flaw in how the heat and freshwater transports are derived, in that the underlying temperature and salinity observations do not provide data at locations where it is needed. The way I understand the explanations below equations 4 and 5, the temperature and salinity profiles are from the moorings at the western and eastern edges (and an average of these). However, at least for the mid basin transports, the section is much longer than typical mesoscale length scales (a case made obvious by figure 1). In order for equations 4 and 5 to hold, both the velocity field v as well as Theta and S profiles must be known at some sort of eddy-resolving resolution in situ. I would not have a lot of confidence in the outcome of these equations unless the methodology has been validated somehow (e.g. in a high-resolution numerical model that reasonably depicts the mesoscale eddy field inside Rockall Trough, where you can then compute the ground truth from equations 4 and 5 using the full model field and compare it with a version that resembles sparse observations). If this has already been done in one of the references, it should be explained here, together with some quantitative uncertainty estimate.

Thank you for highlighting. Please see answer to main comment above.

Ll. 188 ff., eqn. (5):

I think the sign of the salinities is wrong. As it stands, high salinity would give you higher Qf.

Thank you for highlighting. We are correcting the typo.

Ll. 205-207:

If you discuss the undercurrent here, you should refer to it in figure 4 and also include a panel in figure 4 that shows the glider data for comparison. Reg. the overestimate: The error bars overlap, so you could also say that the measurements agree within the given uncertainty, couldn't you?

Thank you for your suggestions. We are updating Figure 3 to include the mean glider section alongside the EOF patterns for direct comparison. For the revised and extended transport calculations (April 2020 – Feb 2023), the mean transport for the glider and the new eastern wedge reconstruction is 1.0±0.3 Sv, while the old methodology gives 1.5±0.2 Sv. The uncertainties overlap, so although the agreement is marginal, the estimates agree within their respective uncertainties. We are editing the text accordingly.

Ll. 208 ff.:

There needs to be an explanation of the bias correction, why it is needed, and what it improves. Perhaps a reference to a publication plus a one-sentence summary is sufficient, but the way it stands, it sounds as if the bias "correction" actually makes agreement with observations here worse.

Thank you for this suggestion. We are moving the explanation of the bias correction in section 3.2 to the data and method section and are editing the paragraph for clarity.

Ll. 257-258:

The sentence here uses the words "better" and "robust". I don't think the manuscript in its present state actually demonstrates that the new data are "better", although there are good reasons to believe this is true. It should, however, be demonstrated. As for "robust", I am not sure what that is supposed to mean - it refers to the seasonal cycle, but figure 5 shows that one can basically draw a straight line through the plot at about 1 Sv and be within all the error bars. If anything, this shows that the data cannot determine the presence of an annual cycle with certainty, doesn't it?

Thank you for your comment. As noted in our response to Reviewer 1, we see the following main advantages of the revised methodology:

- Improved accuracy of the mean strength and structure of the ESC, based on multiyear glider observations rather than a bias correction of GLORYS12v1 data using only eight months of ADCP measurements.
- Enhanced ability to reproduce extreme events, likely associated with mesoscale variability, through inclusion of the first two EOF modes.
- High-resolution ESC product in both space and time, reducing aliasing effects caused by temporally scattered glider data.

We are editing the text throughout the manuscript to clarify these points and ensure that the discussion section clearly communicates the value of this methodology.

Ll. 269-271:

This sentence claims that the ESC is "disproportionately important" for something. When I look at figure 6, I find this to be completely untrue on two accounts: One, the heat and freshwater contributions are fairly proportional to the volume transport, and two, the EW contribution is not very important (instead, the total is dominated by the mid basin transport).

Thank you for pointing this out. As noted in our response to Reviewer 1, this statement refers to findings from previous studies (Clark et al., 2022; Daly et al., 2024), which emphasize the importance of the ESC for on-shelf heat and freshwater transports. The term “disproportionately” reflects its relatively small transport compared to the NAC in the Rockall Trough; however, NAC transport in this region does not significantly affect European shelf exchanges, whereas the ESC does. We are revising the sentence for clarity.

Improvements to text (readability/appearance/typos):

Abbreviations and acronyms:

The amount of abbreviations is overwhelming. Some are not used consistently throughout the text (e.g. RTWB1 vs. WB1). Some are never spelled out (geogr. names from fig. 1). To make things a bit easier, be sure to:

- spell out each abbreviation at its first occurrence and additionally in each figure caption if it occurs inside a figure,

- avoid inconsistencies,

- consider adding a list of all abbreviations in the supplemental materials.

Thank you for highlighting this point. We are revising the text and figures to ensure that abbreviations and acronyms are used consistently throughout the manuscript. Each abbreviation is now spelled out at its first occurrence and in figure captions where applicable. In addition, we are considering adding a comprehensive list of abbreviations in the supplementary materials for ease of reference.

Ll. 17/18:

Remove "Fu et al." reference if unpublished, else update here and add proper citation to the reference list.

Fu et al. (2025) is now published and we are adding it to the reference list accordingly.

Ll. 61-62:

Remove "PSU".

This is being done.

L. 64: Change "Avaraged" to "Averaged"

This is being done.

L. 105: Change "gab" to "gap"

This is being done.

GLORYS references: When I search for "GLORYS" in the manuscript, I see inconsistent occurrences of GLORYS12, GLORYS21, GLORYS2, followed by equally inconsistent version numbers 1, 2, or 12. Please determine the exact version number used, and correct this throughout the manuscript. Why not just call it "GLORYS" everywhere and reference the full name once and once only with the dataset citation?

This is being done.

Ll. 158-159: I found the reference to the future equations confusing. This sentence can perhaps be removed.

This is being done.

L. 250:

I am wondering if "Conclusions" would be a more appropriate title than "Discussion" here.

We agree with the reviewer and are renaming the section to “Conclusions.” For clarity, we are revising the overall paper structure and are adding a dedicated discussion section.

L. 272:

Change "targetted" to "targeted"

This is being done.

Suggestions for figures and captions:

Use consistent labels and spelling across figures:

Figures 1(b), 3, 4, 6, S1 all have depth as vertical axis, but sometimes it counts positive up, sometimes down, and the axis label is spelled differently almost every time.

Ditto for 1(a), 1(b), 3, 4, S1 for longitude.

Ditto for time in 2, 7, 8, S2.

Thank you for highlighting this point. We are revising all figures and captions to ensure clarity and consistency throughout the manuscript. This includes harmonizing labels, abbreviations, and formatting across figures and captions.

Figure 1:

Add (a) to top panel.

Increase size of panel (b) such that it is roughly as wide as panel (a).

Increase font size of panel (b) such that it roughly matches that of panel (a).

In caption, start the panel (b) part with verbiage that says what is shown (e.g., "cross-section view of Rockall Trough section").

In caption, be consistent with abbreviations - use either RTWB1 or WB1, ditto for ...2.

Panel (b) shows things that are not explained in the caption - explain or remove these: two dashed lines with red dots on top, one solid line with blue dot on top. I assume these are the same as in figure 4, but this needs clarification.

Use the same green symbols on panels (a) and (b), i.e. either circles on both panels, or triangles on both.

Either explain all abbreviations in the caption (include the geographic names of panel (a)), or point to a list (potentially in the supplemental materials) where they are explained.

In caption, change "focusses" to "focuses".

Thank you for your suggestions. We are editing Figure 1 and caption accordingly.

Figure 2:

Add cruise IDs to potential list of acronyms, and refer to this list in the caption.

In caption, change "CTD sensors (d-f)." to "CTD sensors (c-f)."

Thank you for your suggestions. We are removing the cruise labels and are editing the caption accordingly.

Figure 3:

The figure caption here refers to "meridional velocity", whereas figure 4 mentions "across section velocity". These two things are almost identical, but not 100%. Please confirm that each caption correctly describes the quantity shown, or correct as appropriate.

Make the axes limits identical to those from figure 4, and add the same vertical lines/symbols for orientation.

Thank you for highlighting. We are revising all relevant figures to show meridional velocities and to share the same limits on the vertical axis.

Figure 4:

Use the same symbols and colors as figure 1 (a) and (b) for the mooring etc. locations.

This is being done.

Figure 5:

The figure seems to show volume transport, but the caption calls the data "velocities depth-averaged…". Correct the caption such that it calls out the correct physical quantity.

Add the time periods for the blue and green curves in the legend (why just the orange and black?).

Thank you for spotting this. We are updating the relevant figure and are clarifying the caption to ensure accuracy (see updated version in Figure 2).

Figure 8:

The font size is inconsistently large compared to the other figures. Reduce font size to make it look similar to the others.

This is being done.