Comment from Referee #2

Below are my comments on the eddy-collocated temperature and salinity profile dataset by He et al.

This manuscript presents a global eddy-collocated T–S profile archive that links 2.35 million quality-controlled WOD profiles to satellite-detected mesoscale eddies over the altimetry era. The scale is impressive, with 2.35 million quality-controlled profiles spanning 1993–2021, broad spatial coverage, and a standardized radius-normalized eddy coordinate system. The manuscript is generally well structured, the motivation is clear. Generally, I think this manuscript is worthwhile to publish, but there are several major concerns must be addressed before the publication.

Response: Thank you very much for your positive evaluation of our manuscript. We greatly appreciate your constructive comments and suggestions. We have carefully studied all of your concerns and performed substantial revisions to improve the quality, clarity, and robustness of the manuscript and dataset. In particular, we updated the profile dataset using the CODC quality-controlled and bias-corrected observations, expanded the validation analyses, added comparisons with recently published eddy-collocated datasets, clarified methodological details, and strengthened the discussion of dataset limitations and potential applications. Below, we provide a detailed point-by-point response to all of your comments and suggestions.

Major Comments

1. Independent dataset cross-validation: One of the major limitations of the manuscript is the complete absence of any reference to, or comparison with, two datasets that are directly analogous in scope and purpose to the one proposed here.

First, Ioannou et al. (2024) presents a global TOEddies atlas explicitly integrated with nearly 3 million collocated Argo float profiles spanning 2000–2023, with the express aim of providing "a novel examination of eddy-induced subsurface variability and the role of mesoscale eddies in the transport of global ocean heat and biogeochemical properties", language that is nearly identical to the framing of the present manuscript. The accompanying open-access dataset (Laxenaire et al., 2024; SEANOE, https://www.seanoe.org/data/00917/102877/) is publicly available. Second, Simoes-Sousa et al. (2026, Earth System Science Data, https://doi.org/10.5194/essd-18-1089-2026) also presents another global ocean profile and altimetry-derived eddy colocation product published in the same journal. Neither work is cited anywhere in the manuscript.

These omissions are not peripheral. Any reader of an ESSD paper will immediately ask: how does this dataset differ from the TOEddies + Argo product, and from Simoes-Sousa et al. 2026? Which is preferable for which applications? What does the added value of the proposed dataset? The authors should directly and substantively engage with these prior works. At a minimum, this requires (a) citing and describing both datasets in the Introduction, (b) articulating clearly what the proposed product offers that the others do not, and (c) providing a quantitative cross-dataset comparison in at least one or two representative regions. Without this discussion, the novelty claim of the manuscript is substantially undermined.

Response: Thank you for this important suggestion. We carefully reviewed both of the two studies and datasets, which are indeed highly relevant to the present work. *Ioannou et al*. (2024) presented a global TOEddies atlas integrated with nearly 3 million collocated Argo float profiles spanning 2000–2023. Subsequently, *Simoes-Sousa et al*. (2026) expanded this framework by incorporating additional profile observations from other instruments, including CTD, XBT, autonomous pinniped bathythermograph (APB), and gliders, producing an eddy-collocated dataset based on 4.2 million profiles and the altimetry-derived META3.2 DT eddy product. However, in both of the two datasets, profiles were only classified according to whether they were located inside or outside eddies. The absence of information on the distance and azimuth between profiles and eddy centers limits the reconstruction of three-dimensional eddy thermohaline structures and characterization of the radial dependence of eddy impacts. In the present dataset, we fill this important gap by providing the relative distance and azimuthal angle of each profile relative to its collocated eddy, thereby facilitating reconstruction of regional mean three-dimensional eddy structures and analyses of eddy impacts on regional ocean environments.

In this revised version, we updated the profile database by adopting the quality-controlled and bias-corrected CODC temperature and salinity profiles recommended by you. This update increases the number of high-quality collocated profiles to 5.46 million, substantially improving the capability for investigating eddy-induced thermohaline anomalies at finer spatial and temporal scales.

Accordingly, we have added descriptions and citations of both of the two datasets in the Introduction (Lines 84–94 of the revised manuscript), clarified the differences and advantages of the present product (Lines 251–258 of the revised manuscript), and included cross-dataset comparisons between this new dataset, previous dataset, and the Simoes-Sousa et al. (2026) product in terms of the spatial distributions of eddy-induced temperature and salinity anomalies (Figs.7-10 and Lines 294–352 of the revised manuscript). We sincerely appreciate this valuable suggestion.

2. Validation is largely qualitative and insufficient: The current validation strategy relies primarily on reproducing mean eddy structures that are already established in the literature (Figs. 4, 7, and 8) and on spatial percentage maps (Figs. 9–12). While demonstrating consistency with prior results is a useful first step, it does not by itself persuasively establish the accuracy, uncertainty, or false-match rate of the proposed dataset. I would recommend two or three of the following additions:

(1) Can add quantitative metrics beyond visual pattern comparison. In selected representative regions, the authors should directly compare anomaly amplitudes, core depths, radial structure widths, and sign-consistency rates between their product and those of independently published composites. Particular attention should be paid to energetic regions, the Gulf Stream, Kuroshio Extension, ACC, and the Agulhas leakage region, where eddies cluster densely and the nearest-eddy assumption is most susceptible to failure (see Major Comment 6 below).

(2) Can add time-series validation. Showing that the seasonal and interannual variability of eddy thermohaline anomalies recovered from the dataset is consistent with independent observations in selected regions would substantially strengthen confidence in the product.

(3) Carry out cross-product validation against the TOEddies + Argo dataset (Ioannou et al., 2024; Laxenaire et al., 2024) and/or the Simoes-Sousa et al. (2026) product. Differences in anomaly amplitude, radial structure, and regional data density would reveal where the methodological choices (eddy product, colocation rule, QC criteria) lead to materially different results, which is itself valuable scientific information.

Response: Thank you for your valuable suggestions. Following your recommendations (as well as Major Comments 4 and 5 below), we replaced the original profiles with the quality-controlled and bias-corrected CODC dataset, increasing the total number of high-quality profiles to 5.46 million. We then compared the revised dataset with both our previous product and the Simoes-Sousa et al. (2026) dataset. Specifically, we estimated mean eddy-induced temperature and salinity anomalies within each

2° × 2° grid box at representative depths of 20 m, 200 m, and 800 m, and compared the results with those derived from He et al. (2024b) and Simoes-Sousa et al. (2026). The horizontal distributions and vertical structures of eddy-induced temperature and salinity anomalies show high consistency among the three products, with spatial correlation coefficients exceeding 0.8 at most of the analyzed depths ($p<0.05$) ($p < 0.05$). (see Figs. 7–10 in the revised manuscript or the Figs. below)

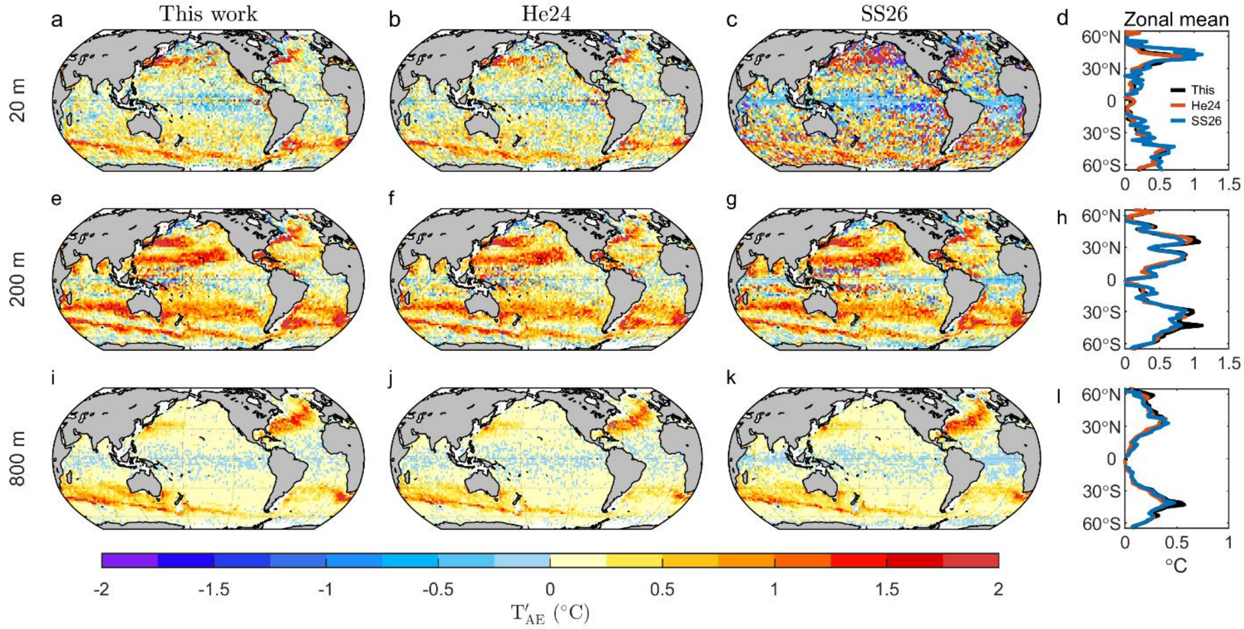

Fig.R1 Geographic distribution of mean temperature anomalies within anticyclonic eddies, estimated within each 2°×2° grid box, at the depths of (top to bottom) 20 m, 200 m, and 800 m. The left three panels are eddy-induced temperature anomalies estimated from the profile dataset in this study, *He et al.* (2024b) (He24), and *Simoes-Sousa et al.* (2026) (SS26), respectively. The right panel is the zonal mean of the left three panels.

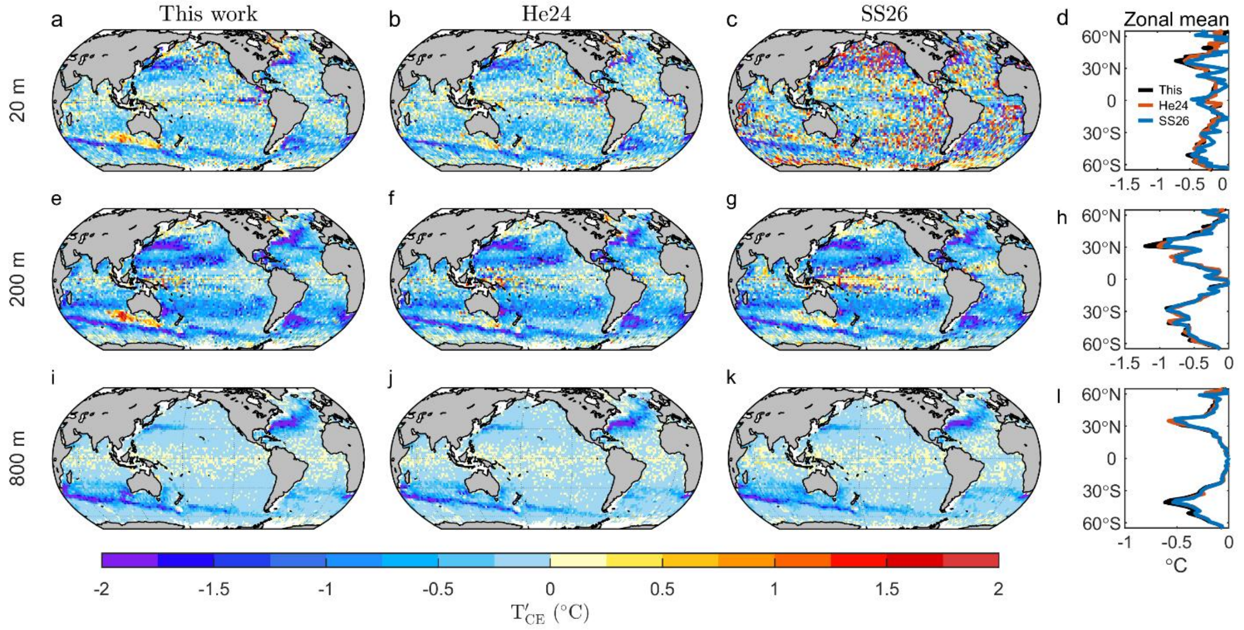

Fig.R2 The same as **Fig.R1**, but for mean temperature anomalies within cyclonic eddies.

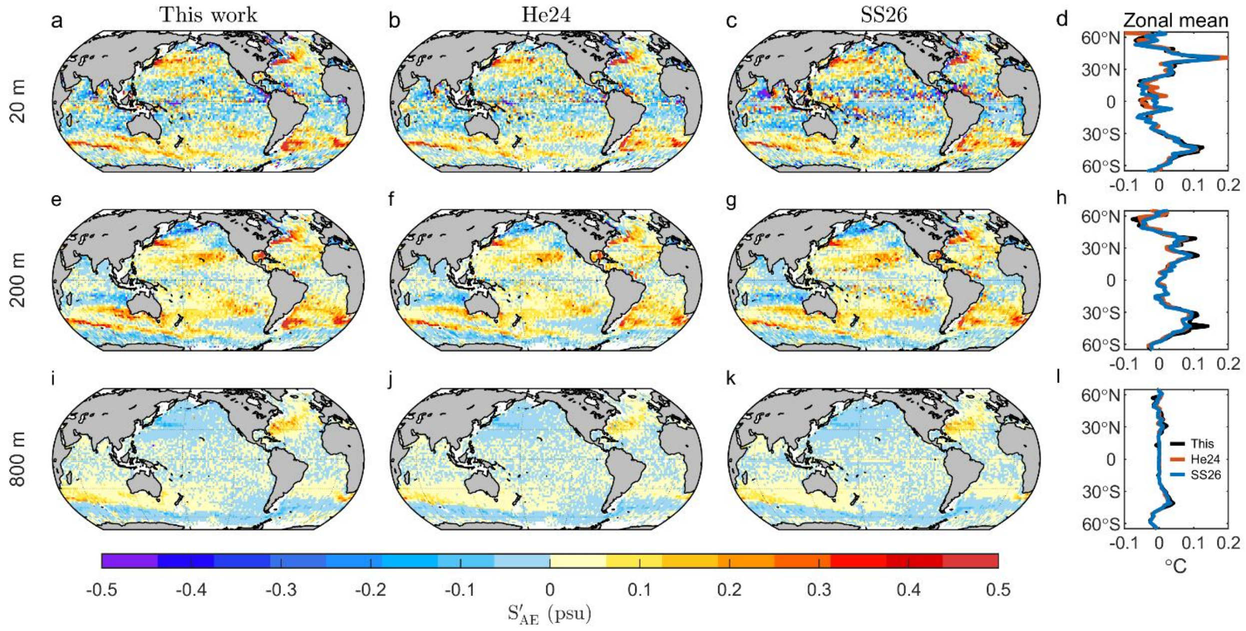

Fig.R3 The same as **Fig.R1**, but for mean salinity anomalies within anticyclonic eddies.

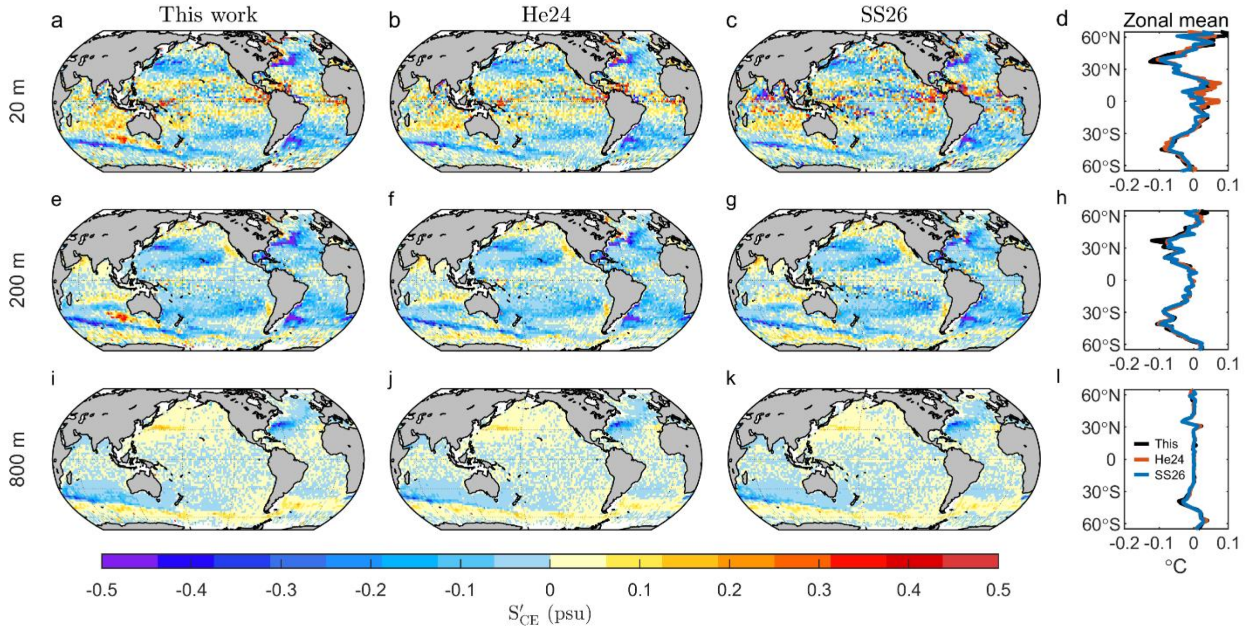

Fig.R4 The same as **Fig.R2**, but for mean salinity anomalies within cyclonic eddies.

Notably, temperature anomalies derived from the Simoes-Sousa et al. (2026) dataset exhibit stronger noisy features near the sea surface than those from the other two datasets (Figs. 7a–7c and 8a–8c). A likely explanation is that their dataset excluded eddies with lifetimes shorter than 30 days or mean amplitudes smaller than 2.5 cm. As a result, some profiles that would otherwise be associated with these weak eddies may instead be assigned to neighboring eddies, introducing additional uncertainty into the estimated temperature anomalies. Because these excluded eddies are typically weak and have relatively limited vertical influence, the impact of their exclusion is expected to decrease with depth, consistent with the reduced differences among the datasets in the subsurface ocean.

Because approximately 80% of the profiles used in *He et al.* (2024b) were derived from Argo observations, the corresponding results can be regarded as largely Argo-based, similar to the TOEddies + Argo dataset (*Ioannou et al.*, 2024; *Laxenaire et al.*, 2024). Thus, we did not specifically provide the results of TOEddies + Argo dataset in the manuscript. In the updated dataset in present study, a substantial number of observations from additional platforms were incorporated. Yet, the resulting temperature anomaly patterns remain highly consistent with the earlier results (Figs.7 and 8). This agreement suggests that the inclusion of observations from multiple observing systems does not materially alter the statistical characteristics of the estimated eddy imprints. Rather, the increased sampling density may enhance the reconstruction of mean three-dimensional eddy structures, particularly at regional scales (Fig.4).

In addition, we expanded the discussion and comparison of eddy-induced temperature anomaly intensities and core depths in representative regions, including the Bay of Bengal (*Sarma et al.*, 2018; *Sarma et al.*, 2020), the South China Sea (*He et al.*, 2018; *Sun et al.*, 2018; *Yang et al.*, 2015), tropical southeastern Indian Ocean (*Yang et al.*, 2015), southeastern Pacific Ocean (*Chaigneau et al.*, 2011), Kuroshio Extension (*Dong et al.*, 2017; *Sun et al.*, 2017), Southern Ocean (*Frenger et al.*, 2015), and Brazil–Malvinas Confluence (*Mason et al.*, 2017). (see revised manuscript, Lines 365-367 and 375-381)

Due to limitations in observational sampling, previous studies have rarely examined radial structure widths or the seasonal/interannual variability of eddy thermohaline anomalies. Therefore, we did not include additional validation for these aspects here. Nevertheless, the substantially enlarged profiles in this updated produced provides strong support for future analyses in these directions. We have added a note of this to Lines 455-458 in the revised manuscript. We sincerely appreciate your valuable suggestions.

3. The Agulhas retroflection and the Cape Basin are one of the hotspots of global eddy activity (Schubert et al., 2021) and have been studied in detail using eddy-profile colocation methods by Laxenaire et al. (2019, 2020). Given that it is precisely this region where the limitations of the chosen eddy product and colocation methodology are most acute (see Major Comment 6 below), It would therefore be valuable for the authors to include more discussion of this region (e.g., Figs 7-8). This region deserves a dedicated attention. For example, can reference or compare with the individual-eddy reconstructions of Laxenaire et al. (2019, 2020).

Response: Thank you for your valuable suggestion. As you noted, the Agulhas retroflection and Cape Basin are among the global hotspots of mesoscale eddy activity. Unfortunately, we were unable to locate the specific reference listed as “Schubert et al. (2021)” in your comments. In the revised manuscript, we present the vertical structures of eddy-induced temperature and salinity anomalies in the Brazil–Malvinas Confluence region and compare them with the results reported by Mason et al. (2017) (see Figs.11g and 12g). We also carefully studied the work by Laxenaire et al. (2019, 2020), which elegantly demonstrated the evolution of Agulhas ring thermohaline structures during their westward propagation. Following your suggestion, we added citations and discussion of these studies in Lines 455–458 of the revised manuscript to further highlight the potential application of the present dataset for investigating temporal evolution of eddy structures.

4. The quality control processes of the original WOD profile: In Sections 2.2 and 2.3, the dataset is derived using WOD quality flags.

First, the ambiguity in the flag specification must be resolved. In Line 134, the author said that 'we extracted temperature and salinity profiles with quality control flags marked as '0' (accepted) during the period of satellite'. Actually, WOD provides several types of quality control flags (Salinity_WODflag, Salinity_WODprofileflag, Depth_WODflag etc.), and different types of flags have different performance to identify outliers (an example can be found in the Fig. 9a of Tan et al., 2025). The manuscript does not specify which flags were applied or how they were combined. This must be clarified.

Second, the WOD quality-control flags may be too weak to identify outliers. This issue has been discussed in detail by Tan et al. (2023, Figs. 14–15), Tan et al. (2025, Fig. 14), and Good et al. (2023, Fig. 2). These serial studies suggest that, even after quality control based only on WOD flags, the temperature and salinity data may still contain a substantial number of non-negligible outliers, especially in observations in the pre-Argo era. Visual evidence of residual outliers appears to be present in the author's manuscript: the vertical sections in Figs. 8a, 8e, and 8f display suspicious spikes or 'dirty points' at approximately 700 m, 1000 m, and 500 m depth, respectively, and Fig. 10c shows an anomalously large spike in the orange line near 35°N that is absent from the corresponding blue line at the same latitude. Such outliers can strongly bias mean and standard deviation estimates (as shown in Fig. 8 of Tan et al. (2025) and Supplementary Fig. S1 of Zhang et al. (2024)) and propagate errors through vertical interpolation into the final eddy anomaly composites. Another example is that when I tried to validate the author's netCDF dataset, if I calculate the salinity standard deviation in each grid box at some selected standard depth levels, I can find that there are many 'suspicious spikes' or 'discontinuous blood-red spots' in the open seas (see my attachment), and this is very likely due to the QC performance. Moreover, if the authors try to calculate the standard deviation map of the Fig. 7 and Fig. 8, it is likely that the these 'discontinuous blood-red spots' occurs.

Anyway, quality control is non-negligible and is one of the largest uncertainty sources in the T(OHC)/S estimate (Boyer et al., 2016), especially to resolve the mesoscale process (Tan et al., 2022 states 'Although the large-scale pattern is similar on a global-basin scale, its meso-micro scale features are visibly different'; Yuan et al., 2026 states that 'An investigation indicates that the WOD local climatological range in its QC check, which is constructed by all historical data, mainly represents the historical ocean conditions and thus removes more positive but realistic positive temperature anomalies in the eddy-rich regions (Boundary Currents and Antarctic Circumpolar Currents regions) than CODC'). Therefore, the impact of QC on the final dataset should be taken carefully.

One possible solution is that the author could use the CODC (CAS-Oceanographic Data Center, Global Ocean Science Database) quality-controlled and bias-corrected in situ ocean temperature and salinity observation dataset (http://www.ocean.iap.ac.cn/ftp/cheng/CODCv2_Insitu_T_S_database/) either as a primary data source or as a cross-check. The main data source of the CODC is also the WOD, but it provides the quality control flag that can remove the outliers as much as possible (more than the WOD-QC) with minimizing the possibility of mistakenly flagging good data (more details could be found in Zhang et al., 2024, and Tan et al., 2025).

In addition, the author should remove profiles on the Argo grey list (WOD doesn't remove them in their quality control scheme), which may contain significant salinity drift. I didn't find any information about whether the author removed the grey list or not in the manuscript. If the author hasn't removed it yet, please remove it.

Response: Thank you for your careful analysis and valuable suggestions regarding quality control. In the original manuscript, our quality-control strategy consisted of selecting profiles with "WOD_observation_flag = 0" (accepted), and further excluding profiles with course vertical resolution, including profiles that (1) lacked measurements shallower than 20 m or deeper than 200 m, (2) contained fewer than 10 unique samples within the upper 200 m, or (3) exhibited vertical sampling intervals larger than 15 m between 0–100 m or larger than 25 m between 100–200 m. As you pointed out, these procedures may be insufficient to fully identify outliers. Following your recommendation, we replaced the original WOD profiles with the CODC quality-controlled and bias-corrected in situ temperature and salinity dataset in this revised version. We then compared the resulting product with both our previous dataset and the *Simoes-Sousa et al*. (2026) product, and found high consistency among the three datasets in terms of the horizontal and vertical distributions of eddy-induced thermohaline anomalies. (see Figs. 7–10 and Lines 294-352 in the revised manuscript)

After adopting the CODC dataset, the spikes and "dirty points" previously visible in Figs. 8a, 8e, 8f, and Fig. 10c disappeared (see Figs. 9,10, and 12 in the revised manuscript). At the same time, we also noticed that even this systematically quality-controlled and bias-corrected dataset cannot completely eliminate all problematic profiles during large-sample automated processing. Some isolated spikes remain in tropical and subtropical regions, although they are substantially fewer than in the previous datasets. We note that these residual outliers may have little influence on basin-scale or global-scale analyses of eddy impacts. However, for studies focusing on specific small regions, additional region-specific quality control may still be necessary. We have added discussion of this issue in Lines 338-348 of the revised manuscript.

5. Systematic instrument biases: Figure 5c shows that XBT, bottle (OSD), and APB data are included in the dataset, although as a small fraction of the total. Each of these instrument types is known to have systematic instrumental biases in the WOD archive: XBT depth and temperature errors have been documented and corrected in many studies (e.g., Cheng et al., 2014); bottle–CTD temperature inconsistencies have been quantified by Gouretski et al. (2022); and APB temperature biases, which are especially relevant for Southern Ocean coverage, have been characterised by Gouretski et al. 2024. These biases are not negligible when the goal is to quantify mesoscale eddy thermohaline anomalies at the level of tenths of a degree Celsius or hundredths of a practical salinity unit.

The authors should either apply the published bias corrections for each instrument type, or demonstrate quantitatively that including these instruments does not materially affect the eddy anomaly estimates in the relevant regions and time periods. For example, the CODC quality-controlled and bias-corrected in situ ocean temperature and salinity dataset mentioned in my previous point also includes temperature profiles that had been bias-corrected. It would be valuable if the authors could consider using bias-corrected data in the proposed dataset, or at least assess whether excluding XBT, OSD, and APB data would have a disproportionate effect on certain regions or time periods.

Response: Thank you for your valuable suggestion. Following your recommendation, we replaced the original profiles with the CODC quality-controlled and bias-corrected temperature and salinity dataset in this revised version. We sincerely appreciate your valuable suggestion, which substantially improved the quality and reliability of the dataset.

6. The choice of eddy detection product and the nearest-eddy colocation assumption: The manuscript uses the META3.2 product for eddy detection and assigns each profile to the nearest eddy on its

sampling day. Both choices carry limitations that deserve explicit discussion:

(1) META3.2 versus more sophisticated eddy atlases. Laxenaire et al. (2018) developed the TOEddies algorithm and validated it systematically against an independent dataset of upper-ocean eddies identified from surface drifters. Their results show that TOEddies correctly identifies approximately 10-15% more validated eddies than META, with lower polarity-mismatch rates, particularly for structures in the 25–60 km radius range. Critically, META3.2 does not detect eddy merging and splitting events. Laxenaire et al. (2018) and Ioannou et al. (2024) demonstrate that these events are abundant, concentrated in the most energetic regions, such as the Cape Basin, western boundary currents, and the ACC, and affect approximately 3% of all detected eddies. When a splitting event occurs while a profile is sampled, META will assign the profile to one fragment while the hydrographic anomaly may be centred in another. The manuscript neither acknowledges this source of contamination nor discusses the choice of META over alternatives.

(2) Subsurface-intensified eddies. Laxenaire et al. (2019) demonstrate through a Lagrangian reconstruction of a single Agulhas ring that eddies can transition from surface-intensified to subsurface-intensified structures as they propagate, retaining large density and temperature anomalies at depth (200–1200 m) while their sea-surface height signature diminishes. Laxenaire et al. (2020) confirm statistically that the majority of Agulhas rings in the South Atlantic are subsurface-intensified. For these eddies, the eddy centre and radius inferred from altimetry will not accurately reflect the location of the thermohaline anomaly core, meaning that radius-normalised distances assigned to the colocated profiles are systematically in error. This limitation may apply not only in the Agulhas region but also to any subsurface-intensified eddy family.

(3) The nearest-eddy assumption in energetic regions. In regions where eddies cluster densely, including the Gulf Stream, Kuroshio Extension, ACC, and Cape Basin, many profiles will be located near multiple eddies simultaneously, and the nearest eddy in terms of centre-to-centre distance may not be the one whose dynamics dominate the observed hydrography. The manuscript acknowledges this issue briefly (lines 155–161) but does not quantify its magnitude. A useful diagnostic would be to compute, as a function of region, the fraction of profiles for which the nearest eddy is also the eddy within whose boundary (outer contour) the profile falls, information that is directly available from the META3.2 eddy contour data.

Response: Thank you for your insightful comments. We fully agree that mesoscale eddies exhibit substantial individual variability, and that the present dataset still has several important limitations. Following your suggestions and our own further considerations, we added a dedicated discussion section (Section 5) in the revised manuscript addressing these limitations and providing references for future users.

Specifically, we now clarify that:

1) First, mesoscale eddies were identified from satellite altimetry and therefore only eddies with sufficiently strong sea surface height anomaly signatures can be detected. Consequently, eddies during their formation and dissipation stages may be underrepresented. In addition, although the META eddy product used in this study is among the most widely used global eddy datasets, it has known limitations in representing eddy splitting and merging processes, which may introduce uncertainties in profile–eddy associations during such events (*Ioannou et al.*, 2024; *Laxenaire et al.*, 2018).
2) Second, satellite observations characterize only the surface expression of mesoscale eddies. Beneath the surface, eddies may exhibit vertically tilted structures, subsurface-intensified cores,

or other forms of structural variability (*Laxenaire et al.*, 2019; *Zhang et al.*, 2016; *Zhang et al.*, 2017). Through composite averaging, the present dataset can reconstruct the mean three-dimensional eddy structure within a target region and reveal systematic regional differences among eddies. However, individual eddies may deviate substantially from the regional mean structure, and such composite analyses may not capture the full range of individual eddy variability.

3) Third, during the collocation procedure, each profile was assigned to its nearest eddy to avoid multiple associations. Although this approach provides a consistent framework for large-scale statistical analyses, hydrographic conditions may occasionally be influenced by multiple neighboring eddies, particularly in regions of dense eddy activity or strong eddy–eddy interactions. Such effects are expected to occur primarily near eddy boundaries and are therefore likely to have a limited influence on estimates of eddy-core thermohaline anomalies.
4) Finally, subsurface and deep-ocean eddies that do not produce detectable sea surface height anomaly signatures cannot be represented in the present dataset. Consequently, the thermohaline structures, transports, and environmental impacts associated with these eddies remain unresolved and require dedicated observational approaches beyond satellite-altimetry-based eddy identification.

We sincerely appreciate your thoughtful comments, which helped us substantially improve the rigor and clarity of the manuscript.

Minor Comments

1. Dataset temporal coverage: The abstract and methods emphasize '1993–2021', whereas the conclusions state '1993–2022'. This must be reconciled throughout.

Response: Thank you for your careful reading. The temporal coverage of the dataset is 1993–2021. We have carefully checked and corrected this inconsistency throughout the manuscript.

2. Table 1 should include units for each variable.

Response: Done, thank you.

3. Section 2.3: How did the author handle the large vertical gap below 200m? Are there any criteria for these intervals below 200m? Maybe the author can also refer to the methods used in Gouretski 2018 ($h = 20 + 0.24 \cdot z$, where z is the mean distance between the two adjacent levels in meters).

Response: Thank you for your suggestion. Since we replaced the original profiles with the quality-controlled and bias-corrected CODC dataset, this quality-control criterion is no longer applied and thus deleted in the revised manuscript.

4. Figure 5a and Figure 5b: add the upper triangle to the colorbar to indicate 'more than' 400 profiles.

Response: Corrected. Thank you.

5. Figure 7a: The grey central band visible in the South China Sea cyclonic eddy panel of Fig. 7a is unexplained. Please add the explanation.

Response: Thank you for your careful observation. This feature is likely caused by the very limited number of profiles located near eddy centers, which may introduce artifacts during spatial interpolation. After adopting the new CODC data with more than doubled profiles, this band vanished. (see Fig.11a in the revised manuscript)

6. Figure 11 and Section 4.2: how did the author define the MLD? Please add the details about the definition. For example, the density threshold, temperature threshold, or hybrid algorithm used (and the reference depth).

Response: In this study, mixed layer depth (MLD) was estimated for each profile using a density-threshold method, defined as the depth at which the increase in potential density relative to the surface equals the increase in surface potential density associated with a 0.5°C decrease in sea-surface temperature (de Boyer Montégut et al., 2004). We added this description in Lines 466-469 of the revised manuscript. Thank you.

7. For Figures 9–12, it would be great to overlay or provide sample count maps/contour lines, because visual interpretation of low-data regions is otherwise difficult.

Response: Thank you for your suggestion. We tested adding sample-count contours to these figures, but found that the contours substantially reduced figure readability. Instead, we provided the spatial distribution of profile numbers in Fig. 5a, the vertical distribution of profile density in Fig. 5e, and the seasonal variations of both total profiles and the fraction located within eddies in Fig. 5f, to help users access data coverage. (see Fig. 5 in the revised manuscript or the Fig. below)

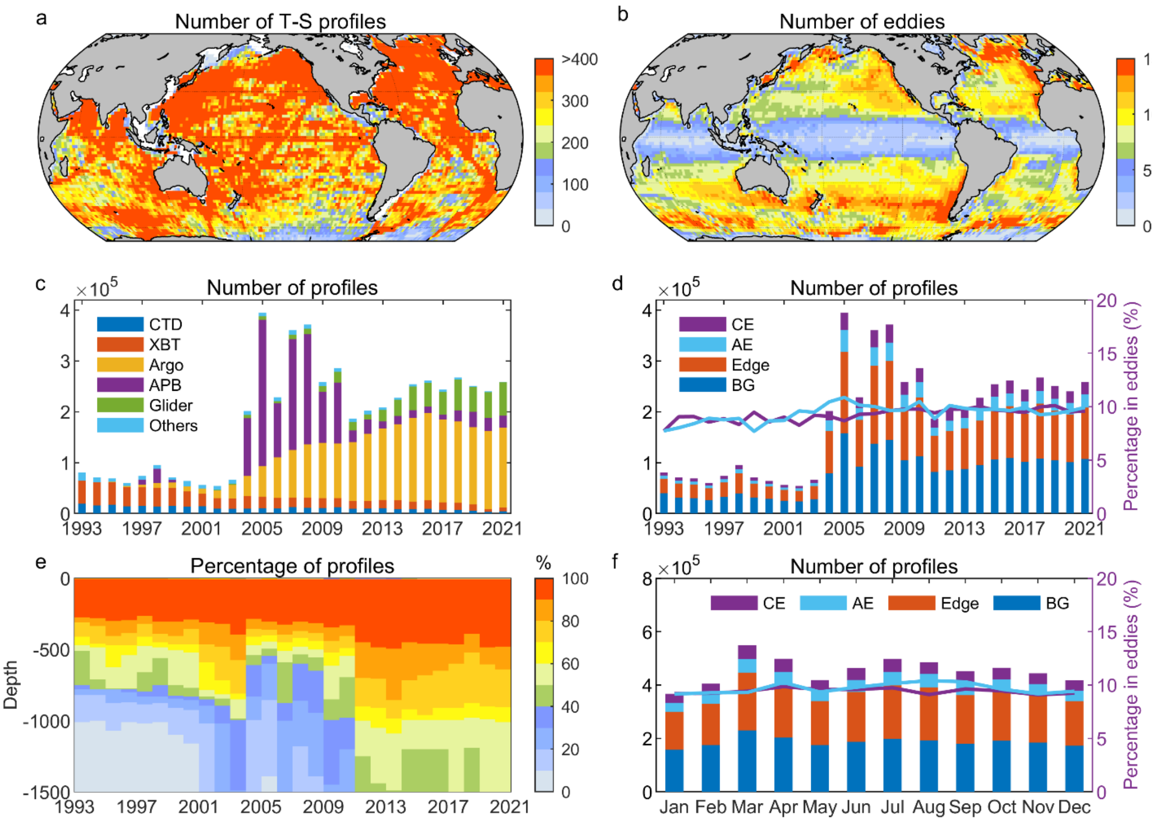

Fig.R5 Spatial and temporal distributions of eddy-collocated temperature and salinity (T-S) profile data in the global ocean between 1993 and 2021. **a**, Geographic distribution of the number of T-S profiles within 2°×2° grid boxes. **b**, The same as a but for the number of satellite-detected eddies. **c**, Yearly statistics of T-S profiles from different instruments. **d**, Yearly statistics of T-S profiles within cyclonic eddies (CE, d<R), anticyclonic eddies (AE, d<R), at eddy edges (R<d<2R), and at background fields (BG, d>2R). The purple and cyan lines are the percentages of profiles within CEs and AEs, respectively. **e**, The same as (**c**), but for the density of profile observations as a function of depth. **f**, The same as (**d**), but for monthly statistics of the profile data.

8. Laxenaire et al., 2020 shows that composite methods systematically underestimate peak heat content anomalies at the eddy centre relative to individual reconstructions, because eddies of different sizes and intensities are pooled in a common normalised coordinate system. The manuscript should acknowledge this limitation and note that the mean anomaly fields it provides do not capture the full range of eddy variability.

Response: Thank you for this important suggestion. We added a discussion of this limitation in the revised manuscript, noting that composite averaging may underestimate peak eddy-core anomalies and cannot capture the full range of eddy variability. (see revised manuscript, Lines 571-577)

9. Some typos should be taken care of. For example:

Line 90: “form the WOD” should be “from the WOD.”

Response: Corrected. Thank you. (see revised manuscript, Line 97)

Line 193: “the choose of profiles” should be “the choice of profiles”

Response: Corrected. Thank you. (see revised manuscript, Line 204)

Line 251: “random sampled” should be “randomly sampled.”

Response: Corrected. Thank you. (see revised manuscript, Line 274)

Line 453: “to what extend” should be “to what extent.”

Response: The paragraph is rewritten and these words are deleted in the revision. Thank you.

Line 471: “Zendo” should be “Zenodo” in the data availability section

Response: Corrected. Thank you. (see revised manuscript, Line 613)