Comment from Referee #1

In this work by He and coauthors a dataset providing collocated temperature and salinity data with satellite-derived eddy information. The paper would require major changes to improve its clarity and refine the scope. I have three major concerns as described below, as well as a number of minor suggestions.

Response: Thank you very much for your valuable comments and suggestions. We have carefully studied all of your remarks and revised the manuscript accordingly to improve its clarity and better refine the scope of the study. We hope that the revised version now meets your expectations. Our point-by-point responses are provided below.

The novelty of this contribution is not clearly communicated, as the processing of the input data (WOD and AVISO) appears to be relatively basic (essentially interpolation). Beyond easier access to existing resources, the value added by the authors should be made explicit.

Response: Thank you for this insightful comment. To better highlight the novelty and scientific value of this dataset, we expanded the Introduction to include a more comprehensive review of the two previously published eddy-collocated profile datasets by Ioannou et al. (2024) and Simoes-Sousa et al. (2026). We note that, in these two datasets, profiles were only classified according to whether they were located inside or outside eddies. Whereas, our dataset additionally provides the relative distance and azimuthal angle between each profile and its collocated eddy. This information enables the reconstruction of three-dimensional eddy thermohaline structures and facilitates analyses of how eddy impacts vary with distance from the eddy center, as well as the reconstruction of regional mean three-dimensional eddy thermohaline structure (see revised manuscript, Lines 84–94).

In addition, we updated the hydrographic profile dataset in the revised manuscript by replacing the previously used profiles with a systematically re-quality-controlled and bias-corrected dataset (CODC-v1) (Zhang et al., 2024; Tan et al., 2025). This updated data retained as many valid observations as possible while ensuring data quality, increasing the total number of collocated profiles from approximately 2 million to 5.46 million (see revised manuscript, Lines 142–147). We further compared the updated dataset with previous products and demonstrated their strong consistency in representing regional mean eddy thermohaline impacts. The substantial increase in profile number may provide important support for investigating eddy-induced variability at finer spatiotemporal scales, including sub-basin variability and seasonal/interannual changes (see revised manuscript, Figs. 7–10 and Lines 294–352).

Furthermore, building upon the methodological descriptions provided by Ioannou et al. (2024) and Simoes-Sousa et al. (2026), we further discussed the potential applications of this dataset in studying eddy three-dimensional structures, eddy impacts on thermohaline distributions, stratification and mixing, heat and salt transport, extreme temperature and salinity events, and climate-change-related variability, thereby expanding the potential research applications of the dataset. (see revised manuscript, Lines 451–461)

References:

Simoes-Sousa, I. T., C. Rocha, A. Tandon, and A. Schmidt (2026), Integrating Global Ocean Profiles Data and Altimetry-Derived Eddies, *Earth System Science Data*, doi:10.5194/essd-2025-40.

Ioannou, A., L. Guez, R. Laxenaire, and S. Speich (2024), Global Assessment of Mesoscale Eddies with TOEddies: Comparison Between Multiple Datasets and Colocation with In Situ Measurements,

Remote Sensing, *16*(22), doi:10.3390/rs16224336.
Tan, Z., L. Cheng, V. Gouretski, B. Zhang, Y. Wang, F. Li, Z. Liu, and J. Zhu (2023), A new automatic quality control system for ocean profile observations and impact on ocean warming estimate, Deep Sea Research Part I: Oceanographic Research Papers, 194, doi:10.1016/j.dsr.2022.103961.
Zhang, B., et al. (2024), CODC-v1: a quality-controlled and bias-corrected ocean temperature profile database from 1940-2023, *Scientific data*, *11*(1), 666, doi:10.1038/s41597-024-03494-8.

Discussion on the effect of trends (e.g., salinity and temperature, eddy, other datasets used) is not provided adequately (global warming mentioned once in sect. 2.2), but may be very relevant. Please expand the discussion and provide context also for the possible influence of observational changes through time.

Response: Thank you for the valuable suggestion. We have added a new subsection discussing the potential application of this dataset in the study of the impacts of climate change on eddy-induced thermohaline anomalies. Regarding the influence of temporal changes in observational coverage, we added a note in the revision that the number of profiles in this dataset has remained relatively stable at more than 200,000 profiles per year since 2004, which may support basin-scale trend analyses (see revised manuscript, Lines 545–561).

The level of detail and support of several pieces of text is insufficient; for example, the discussion of temperature extremes is missing a description of the methodology (and several key parameters, such as the baseline used to define such extremes), and the application section is very scarce and limited mostly to citation of works by the paper authors. The authors should clarify that the dataset is not gridded (even though a 2x2 grid is mentioned in various places).

Response: Thank you for the helpful suggestions. Following your recommendation, we added a clearer description of the methodology used to identify extreme temperature anomalies. Specifically, extreme high (low) temperature anomalies were defined as temperature anomalies exceeding the 95th percentile (falling below the 5th percentile) within each 2° × 2° grid box over the entire study period (1993–2021), relative to the local monthly climatology (see revised manuscript, Lines 500–503).
We also expanded the result section by adding comparisons between our dataset and previously published datasets in terms of the spatial patterns of eddy-induced thermohaline anomalies (Section 3.2 and Figs.7-10 in the revision), as well as comparisons of reconstructed vertical eddy structures with earlier studies (see revised manuscript, Lines 375-381, 427–434, and 481-485).
In addition, we revised the subsection titles from "3.1" to "Overview of the eddy-collocated profile dataset" and from "3.2" to "Validation of the eddy-collocated profile dataset" to clarify that the dataset itself is not gridded.

Minor suggestions:
36 why heat budgets?
Response: We have replaced "heat budgets" with "heat/salt transport". Thank you.

52 why freshwater?
Response: We have replaced "freshwater" with "salt". Thank you.

62 add reference

Response: Done. Thank you.

1) Chelton, D. B., M. G. Schlax, R. M. Samelson, and R. A. de Szoeke (2007), Global observations of large oceanic eddies, *Geophys Res Lett*, *34*(15), doi:10.1029/2007gl030812.
2) Dufau, C., M. Orsztynowicz, G. Dibarboure, R. Morrow, and P. Y. Le Traon (2016), Mesoscale resolution capability of altimetry: Present and future, J Geophys Res Oceans, 121(7), 4910-4927, doi:10.1002/2015jc010904.

86 this reads like a straw man argument

Response: We removed this sentence from the revised manuscript. Thank you.

89 collocation may be working at the surface, but how deep this would hold?

Response: Thank you for this insightful comment. Because the subsurface shape of eddies cannot be directly observed, we followed previous studies by assuming that eddies have vertically coherent columnar structures during the profile–eddy collocation process (*Ioannou et al.*, 2024; *Simoes-Sousa et al.*, 2026). The reconstructed three-dimensional thermohaline anomaly structures based on the collocated dataset support this assumption and show that eddy impacts can penetrate several hundred meters and even exceed 1000 m, with substantial regional variability (see Figs.11-12 in the revision or the Figs. below). We should note that, 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. We have added a discussion on this limitation to lines 571-577 in the revision.

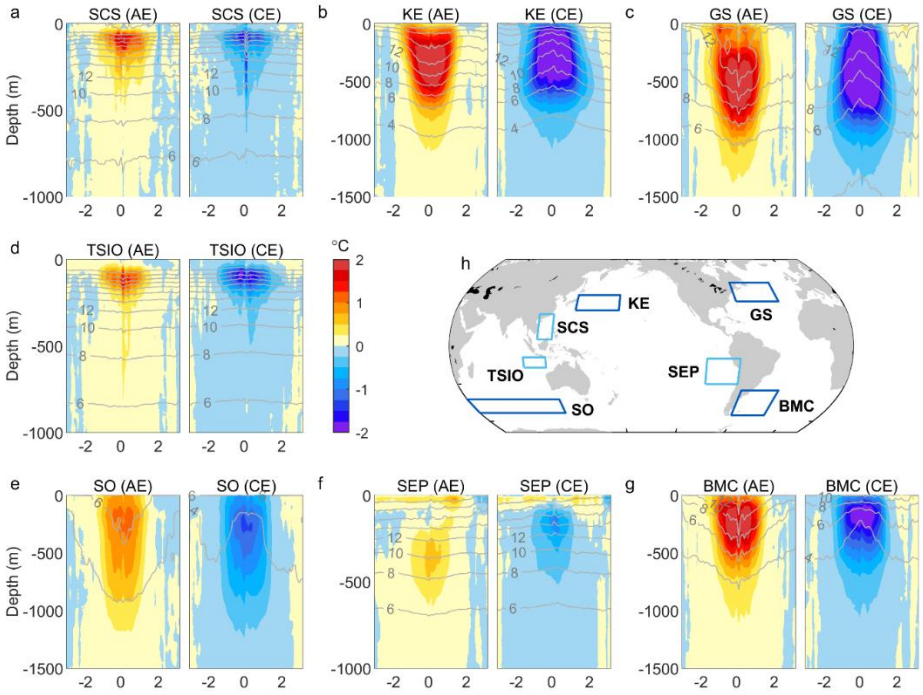

Fig.R1 Composite mean eddy temperature anomaly structures in representative regions of the global

ocean. **a-g**, West-east sections of mean temperature anomalies across the composite centers of anticyclonic (AE) and cyclonic (CE) eddies in the South China Sea (SCS), the Kuroshio Extension (KE), the Gulf Stream (GS), the tropical southeastern Indian Ocean (TSIO), the southeastern Pacific Ocean (SEP), the Southern Ocean (SO), and the Brazil-Malvinas Confluence (BMC). **h**, The statistical regions (boxes) of (**a-g**).

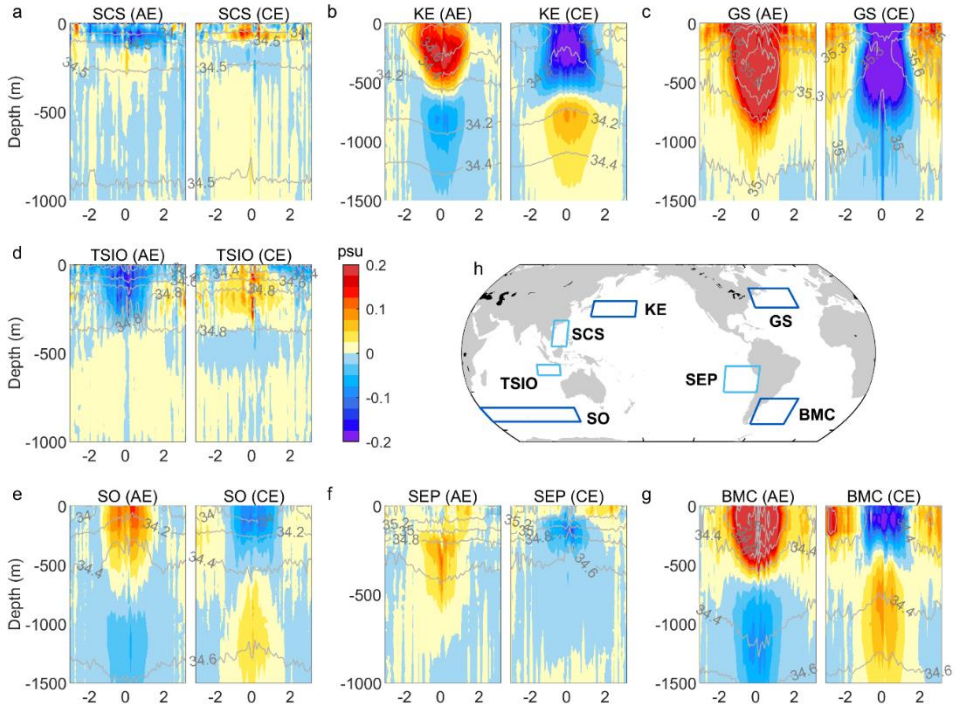

Fig.R2 The same as **Fig.R1**, but for eddy salinity anomaly structures.

References:
Simoes-Sousa, I. T., C. Rocha, A. Tandon, and A. Schmidt (2026), Integrating Global Ocean Profiles Data and Altimetry-Derived Eddies, *Earth System Science Data*, doi:10.5194/essd-2025-40.
Ioannou, A., L. Guez, R. Laxenaire, and S. Speich (2024), Global Assessment of Mesoscale Eddies with TOEddies: Comparison Between Multiple Datasets and Colocation with In Situ Measurements, *Remote Sensing*, *16*(22), doi:10.3390/rs16224336.

90 text in brackets is puzzling

Response: We removed the text in brackets in the revised manuscript. Thank you.

103 fix typo in flowchart

Response: Corrected. Thank you.

128 I am confused by the reference on Captain Cook; please clearly list which data sources (Argo, moorings,...) are used in this dataset. Profiles are only possible with Argo, right?

Response: Our original intention was to briefly indicate that the World Ocean Database contains hydrographic observations spanning a long historical period. Actually, the present study only used

profiles collected during 1993–2021, when satellite-derived eddy observations are available. To avoid redundancy and confusion, we removed the sentence referring to Captain Cook.
To maximize the number of usable profiles, we included all systematically quality-controlled and bias-corrected temperature and salinity profiles, including observations from Argo, Conductivity-Temperature-Depth (CTD) instruments, expendable bathythermographs (XBT), autonomous pinniped bathythermograph (APB), gliders, and other platforms (Fig.5c in the revision). These details are now clarified in Lines 137–147 of the revised manuscript. Thank you.

134 while eddies are daily, are the WOD data provided with the same resolution? This means that profiles for a period shorted than a day are aggregated?

Response: The World Ocean Database profiles used in this study are discrete profile observations. For each profile, we searched for and collocated the nearest eddy identified on the same sampling day (see revised manuscript, Line 158). Therefore, no temporal averaging was applied, even when multiple profiles were collected within the same day. This situation may occur for ship-based CTD, and high frequency Argo, gliders, and/or APB observations, while the sampling location varies for each profile. Therefore, we consider daily averaging unnecessary.

137 the QC applied by WOD should be outlined

Response: Thank you for the suggestion. In the revised manuscript, we replaced the previously used World Ocean Database profiles with systematically re-quality-controlled and bias-corrected dataset (CODC-v1) from *Zhang et al.* (2024) and *Tan et al*. (2025), making the original World Ocean Database quality-control description unnecessary. We therefore revised this section accordingly (see revised manuscript, Lines 142-147) and removed the original sentence.

143 well then the grid is not uniform in general, only by layer. How do you define this layers, anyway?

Response: Thank you for pointing this out. We reduced the vertical grid resolution from 5 m to 10 m below 1,000 m and to 50 m below 1,500 m mainly to decrease storage requirements and computational memory usage. Considering that deep-ocean temperature variations are much weaker and smoother than those near the surface, this adjustment may not substantially affect the statistical characteristics of eddy-induced thermohaline anomalies (see Figs.11-12 in the revised manuscript). We have added a explanation of this to Lines 150-153 in the revision.

149 what is the maximum distance allowed?

Response: Thank you for bringing this to our attention. Considering that the influence of mesoscale eddies typically extends to approximately 1.5 times the eddy radius, profiles located beyond twice the eddy radius are all treated as background profiles unaffected by the eddy. Therefore, no strict maximum search distance is required.
However, to improve computational efficiency during the collocation process, we added a restriction in the revised manuscript such that eddy searching is only performed within a 10° × 10° grid box centered on each profile location (see revised manuscript, Line 157). Thank you for the helpful suggestion.

163 explain in the text how d, D, and R are defined

Response: The variable *d* denotes the distance between a profile and its collocated eddy center, *D*

represents the distance from the eddy center to the eddy boundary along the same azimuthal direction as the profile, and *R* is eddy radius. We have added these definitions to Lines 177–176 in the revised manuscript. Thank you.

167 I don't understand this. If profiles are extracted only in the vicinity of eddies, wouldn't this induce a bias in the selection, which would not be random?

Response: We apologize for this incorrect wording. This sentence is revised as: "As most of the profiles were discretely sampled around or within different locations of different eddies, we cannot obtain the three-dimensional structure of a specific eddy". (see revised manuscript, Lines 178-179)

173 this point is confusing. Being a 30-year period, trends may be relevant. How are these accounted for in your method? And aren't somehow eddy signatures visible in areas where those are more persistent? Also, how good is this product closer to the coasts?

Response: We apologize for the unclear wording. Our original intention was to indicate that monthly climatological means were removed from each profile to suppress seasonal-cycle signals and reduce aliasing caused by spatially heterogeneous historical sampling. We agree with you that ocean warming trend may be relevant. We intentionally retained this signal to facilitate applications of investigating recent trends in eddy-induced thermohaline anomalies and their potential driving mechanisms. Accordingly, we revised the sentence as follows: "To remove seasonal cycle signals and diminish aliasing due to sparse historical sampling, temperature/salinity anomalies were estimated for each profile by subtracting the corresponding climatological monthly mean value (Swart et al., 2018)." (see revised manuscript, Lines 183-184)

Satellite-observed eddy trajectories indicate that most oceanic eddies are continuously propagating features rather than stationary structures. Therefore, removing the local climatological monthly mean helps isolate eddy-induced anomalies from the background ocean state. We also note that eddy-induced anomalies could alternatively be estimated by directly computing mean differences between profile observations inside and outside eddies. We therefore provided the original profile data in the dataset to allow users to apply alternative approaches if desired. (see revised manuscript, Table 1)

As shown in the spatial distributions of profiles and collocated eddies (Figs.5a–5b), the vast majority of observations are located in the open ocean, where mesoscale eddies are more likely to persist. Therefore, we did not apply additional filtering for shallow coastal regions. The successful reconstruction of eddy three-dimensional structures in the northern Bay of Bengal example (Fig. 4) also suggests that this issue does not substantially affect the analysis. Nevertheless, users interested in coastal applications may choose to exclude shallow-water profiles depending on their research objectives.

186 "ambientes"?

Response: We have replaced "ambientes" with "eddy peripheries". Thank you.

Fig 4a typo

Respo**nse:** Corrected. Thank you.

Table 1 is missing units

Response: Units for all variables have been added in Table 1. Thank you.

Fig. 5 I don't understand if/how (a) and (b) should differ. Can also "glider" and "others" provide profiles? It would be useful to report data density by depth

Response: Thank you for bringing this to our attention. You are right that the spatial distributions of the profiles and their collocated eddies are identical because each profile is associated with one eddy. Therefore, we replaced panel (b) with the spatial distribution of all altimeter-detected eddy numbers. In addition, we added new panels showing the vertical distribution of data density and the seasonal variation of profile numbers (see Fig.5 in the revised manuscript or below).

The WOD dataset does include temperature and salinity profiles collected by gliders (*Mishonov et al.*, 2024; *Zhang et al.*, 2024). The "others" category mainly includes observations from ocean stations, moored buoys, and Mechanical Bathythermographs (MBTs). Since the numbers of these observations are relatively small, they were grouped together.

References:

Mishonov A.V., T. P. Boyer, O. K. Baranova, C. N. Bouchard, S. Cross, H. E. Garcia, R. A. Locarnini, C. R. Paver, J. R. Reagan, Z. Wang, D. Seidov, A. I. Grodsky, J. G. Beauchamp, (2024): World Ocean Database 2023. C. Bouchard, Technical Ed., NOAA Atlas NESDIS 97, 206 pp., doi.org/10.25923/z885-h264,

Zhang, B., et al. (2024), CODC-v1: a quality-controlled and bias-corrected ocean temperature profile database from 1940-2023, *Scientific data*, *11*(1), 666, doi:10.1038/s41597-024-03494-8.

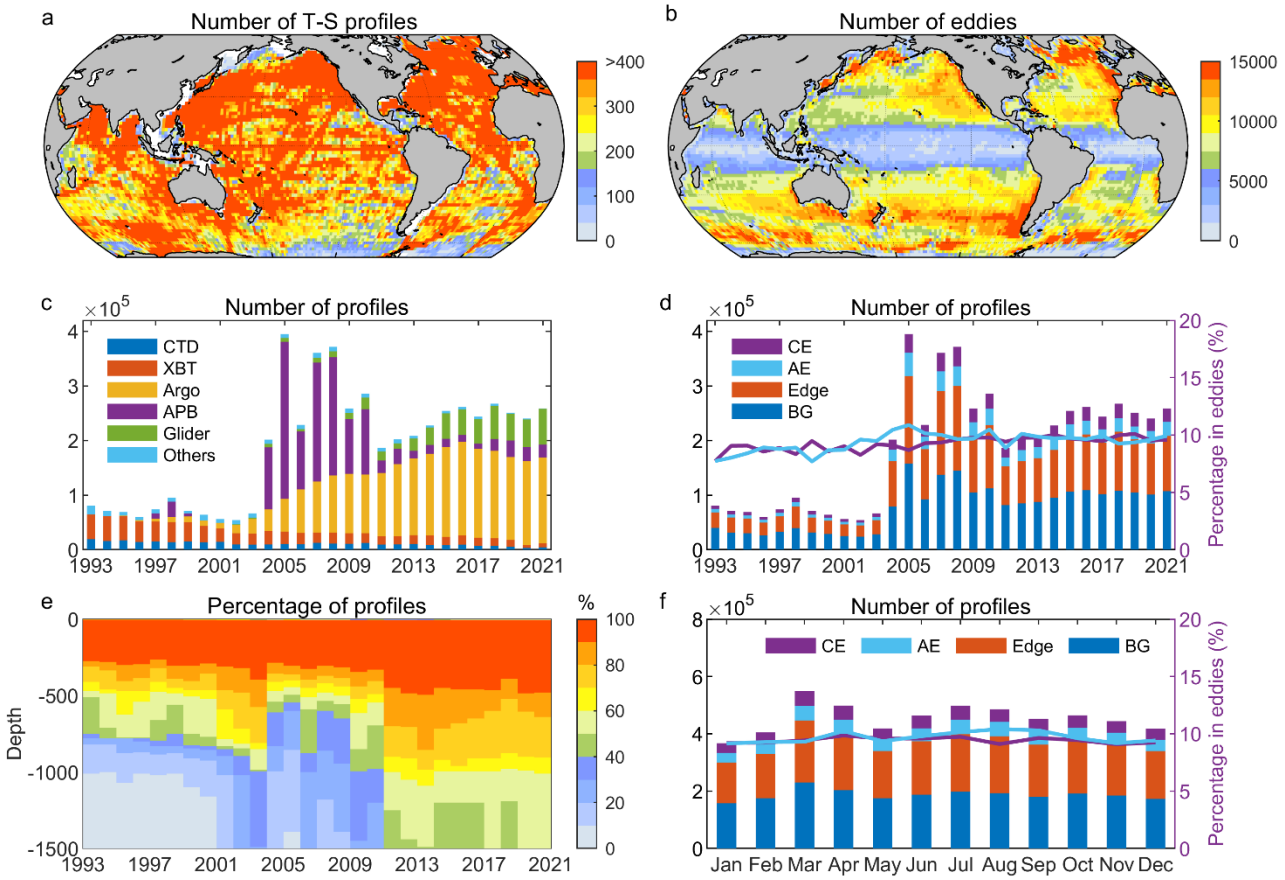

Fig.R3 Spatial and temporal distributions of eddy-collocated temperature and salinity (T-S) profile data in the global ocean between 1993 and 2021. **a**, Geographic distribution of the number of T-S profiles within 2°×2° grid boxes. **b**, The same as a but for the number of satellite-detected eddies. **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.

245 repetition, rephrase

Response: This sentence is revised as "Although the number of profiles exhibits pronounced interannual variability, the fraction of profiles located within mesoscale eddies remains relatively stable, with nearly 10 % occurring within cyclonic eddies and another 10 % within anticyclonic eddies, throughout the study period" (see revised manuscript, Lines 268-270). Thank you.

250 I am not following your reasoning. Wouldn't eddy trap lagrangian sensors?

Response: Thank you for pointing this out. We fully agree that eddies can trap Lagrangian sensors. However, XBT and CTD profiles are mainly collected from moving research vessels, APB observations follow marine animal trajectories, and gliders sample along pre-designed routes. Among the observing systems used here, only Argo floats drift with ocean currents. However, the deployment locations of Argo floats are largely random, with some floats entering eddies and others remaining outside. Additionally, most Argo floats park near 1000 m, where they are generally less influenced by mesoscale eddies, except in a few cases involving exceptionally strong eddies or shallow parking depths.

As shown in Figs.5-6, although the yearly number of profiles varies substantially, the fraction of profiles located within eddies stays at a relatively level of 10%. Additionally, the spatial distribution of the fraction of profiles within eddies is highly consistent with the occurrence probability of mesoscale eddies. Therefore, we conclude that, statistically, these profiles can be regarded as approximately randomly distributed relative to mesoscale eddy fields. We have added a detailed discussion on this to Lines 268-274 in the revised manuscript.

Fig. 6 why is the resolution different between left and right maps? What are Gamma and P?

Response: Thank you for pointing this out. In the left panels of Fig.6, percentages of profiles located within eddies were estimated within 2° × 2° grid boxes. This grid size represents a compromise between maintaining sufficiently high spatial resolution and ensuring an adequate number of profile samples within each grid box for the subsequent estimation of eddy thermohaline impacts.

In contrast, the right panels show the fraction of days during which each location was occupied by mesoscale eddies (within eddy interiors) over the study period. Since this calculation is not limited by sample numbers, we chose to provide a high spatial detail by performing the statistics directly on the original altimetry grid points (0.25° × 0.25°).

Gamma (Γ) denotes the percentage of profile data within eddies and P represents the occurrence probability of mesoscale eddies. These definitions have been added to the revised Figure caption.

204 please explain how interpolation is made

Response: We added a description of the interpolation method in the revised manuscript. Specifically,

at each depth level, profile observations were interpolated onto standard 0.1R×0.1R grid boxes using ordinary kriging interpolation. (see revised manuscript, Lines 216-217)

271 this is not clear

Response: This sentence has been revised as: "The resulting profiles are relatively evenly distributed both within eddies and in the surrounding regions" (see revised manuscript, Line 359-360). Thank you.

296 this should be shown with your dataset, as it should be possible

Response: Thank you for your suggestion. To maintain the overall organization and clarity of the figures, we did not include a separate comparison of the vertical stratification structures (e.g., vertical temperature gradients) between the two regions. Instead, we added contours of the vertical temperature structure within eddies (gray contours) to the Figs.11 and 12. These contours show that the Gulf Stream region exhibits denser temperature contours than the Kuroshio Extension region, indicating stronger vertical temperature gradients and thus weaker vertical stratification in the Gulf Stream region. Since this section mainly aims to demonstrate potential applications of the dataset in eddy studies, we removed the speculative discussion regarding the underlying mechanisms in the revised manuscript to avoid overinterpretation.

344 no references in these and close sections, only self-citation?

Response: Thank you for pointing this out. We have added citations to additional relevant previous studies to lines 425, 429, 432, 443, 457, and 459 in the revised manuscript.

352 as in 296, this remains speculative and should be verified

Response: Thank you for your suggestion. We agree that this is an interesting phenomenon whose underlying mechanisms deserve dedicated investigation in future studies. Since the primary focus of this manuscript is to introduce the dataset and demonstrate its potential applications, we did not attempt a detailed mechanistic analysis here. In the revised manuscript, we replaced the mechanism discussion with a state that "This result highlights that eddies with similar surface signatures can exhibit markedly different subsurface structures. The dataset therefore provides important observational constraints on the subsurface characteristics of mesoscale eddies that are not discernible from satellite observations alone". (see revised manuscript, Lines 434-437)

386 was this defined already?

Response: Thank you for your careful reading. MLD here was calculated for each profile 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 have added the definition of MLD to Lines 466–469 in the revised manuscript.

408 include references and key details, e.g. the climatology used

Response: Thank you for your suggestion. We added the references regarding marine heatwaves and cold spells (Oliver et al., 2021; Schlegel et al., 2021) to Lines 495–496 of the revised manuscript.

References:

Oliver, E. C. J., J. A. Benthuysen, S. Darmaraki, M. G. Donat, A. J. Hobday, N. J. Holbrook, R. W.

Schlegel, and A. Sen Gupta (2021), Marine Heatwaves, Annual review of marine science, 13, 313-342, doi:10.1146/annurev-marine-032720-095144.
Schlegel, R. W., S. Darmaraki, J. A. Benthuysen, K. Filbee-Dexter, and E. C. J. Oliver (2021), Marine cold-spells, Prog Oceanogr, 198, 102684, doi:10.1016/j.pocean.2021.102684.

416 daily or coarser climatology?

Response: Because the profile observations are not temporally continuous, the threshold for identifying extreme temperature anomalies was defined as the 95th percentile of all temperature anomalies within the study period (1993–2021). We added this clarification in Lines 500–503 of the revised manuscript. Thank you.

Fig. 12 is the spatial resolution still 2 deg in these plots?

Response: Yes. Owing to the substantial increase in the number of collocated profiles in the updated dataset, we used a 2° × 2° statistical grid in this analysis. The resulting spatial patterns are highly consistent with our previous results based on 5° × 5° grid boxes, suggesting that the updated dataset enables investigation of eddy impacts on extreme temperature events at finer spatial scales. We added discussion of this point in Lines 518-522 of the revised manuscript.

436 this section seems too sketched and unsupported

Response: Thank you for this helpful comment. We agree that eddy-induced heat and salt transport is a complex topic that cannot be adequately addressed through a brief analysis. Reliable estimates of eddy-induced heat, salt, and water-mass transports require not only detailed thermohaline information but also constraints on eddy velocity structures and associated dynamical fields, which remain challenging to obtain from observations alone. Consequently, substantial uncertainties still exist in observational estimates of eddy transports, and a comprehensive assessment is beyond the scope of the present study.

Our intention in this section was not to provide quantitative estimates of eddy-induced heat or salt transport, but rather to highlight a potential application of the dataset. As this manuscript focuses on the description and evaluation of the dataset, we believe it is useful to discuss how the expanded profile archive may support future observational studies of eddy transport processes. Following your comment, we have revised this section in a more cautious discussion-oriented tone. The revised text now emphasizes the potential of the dataset to provide observational constraints on future transport estimates, rather than implying that such estimates are performed or validated in the present study. (see revised manuscript, Lines 529-544)

472 a single 10GB file is not ideal; why not splitting by basin, for example? The file is also lacking metadata (reference, data sources) and units. Why is "contour" and "char" set to 20?

Response: Thank you for your valuable suggestion. We have split the original large dataset into multiple smaller files by year and added metadata information, including data sources and references, and units. The variables “contour” and “char” remain set to 20 because these values were directly inherited from the original eddy dataset, and we did not modify them.

475 the specific link should be provided

Response: Thank you for pointing this out. We updated the manuscript with the specific link to the eddy dataset: https://www.aviso.altimetry.fr/en/data/products/value-added-products/global-mesoscale-eddy-trajectory-product/meta3-2-dt.html. (see revised manuscript, Lines 616-617)