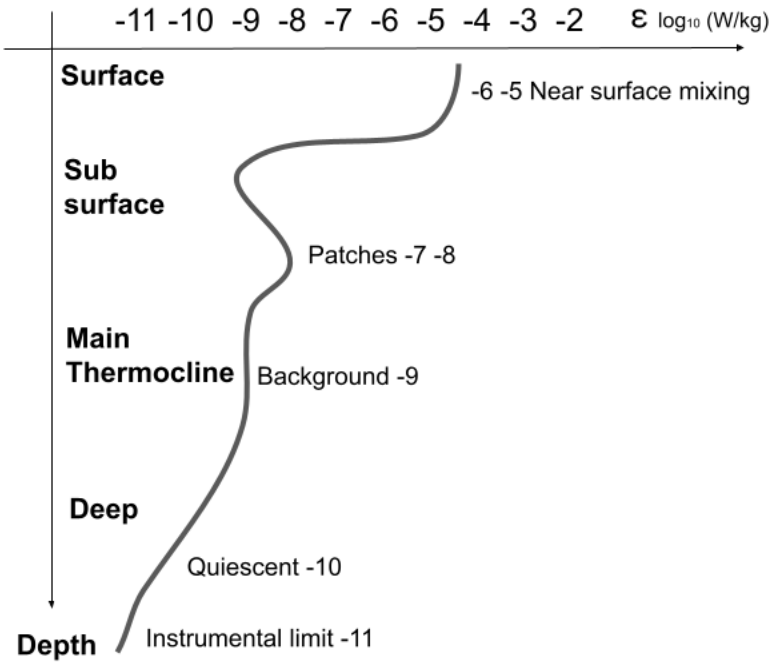
Reviewer 1

General comment: This data manuscript describes a unique and valuable microstructure dataset sampled with a glider over several years and seasons in the western Mediterranean. The manuscript provides and exhaustive overview of the data processing, which will be quite useful for anyone using this dataset in the future.
Few sections can be shortened without trying to produce a review or give arguments that such a dataset contains key parameters. All measured variables in the ocean are valuable and need to be publicly available, whether they are scarce (such as microstructure observations) or widely collected (such as CTDs).
You will find all my comments, questions and suggestions in the commented PDF attached to my review.
With kind regards.

We thank the reviewer for the careful and constructive review, and for the positive assessment of the value of the dataset. We have addressed all comments point by point below and revised the manuscript accordingly. All suggested clarifications, corrections, and simplifications have been implemented in the updated version of the manuscript (in green color).

PAGE 2 — Introduction
(1) from seasonal to interanual scales
(2) … to measure routinely small-scale…
(3) using
(4) sensorS to be consistent with the plural form of thermistor sensorS… or use singular form for both
(5) meaningless: all regions have "dynamics"
(6) avoid lengthy wording: simplify with: 'The present work provides one of the most….'
Thanks to the reviewer for these suggestions; we have taken them into account and corrected the manuscript accordingly.

PAGE 3 — Figure 1 caption / text
(1) 'schematize' is not explicit about what is shown: why don't you give a mean profile ?
(2) … of dissipation rates of turbulent kinetic energy (logarithmic scale) as a function of depth.
We chose to retain a schematic representation rather than a mean profile, as a mean vertical profile would provide an overly smooth view that does not adequately represent the patchy and intermittent nature of turbulent events encountered throughout the water column. The new schematic has been corrected and simplified as presented below, to better align with the synthetic vertical structure shown in Fig. 8 (renumbered, previously Fig. 7). The manuscript text, the Fig. 1 and its caption have been revised accordingly, including an updated description of the typical ε ranges with depth.



The new caption is :

Figure 1. Glider Teresa mission along its longitudinal transect, by years (in colors). The scheme indicates the deployment between Sardinia and Balearic Islands in the Western Mediterranean Sea. Sawing black lines indicate a schematic trajectory from surface up to 1000m-depth, encountering layers of the Atlantic Water (AW) in surface, the Eastern Intermediate Water (EIW) and the Western Mediterranean Deep Water (WMDW). Depth levels associated to these water masses (respectively 0-200m, 200-1000m, 1000m+) are indicative. The lower-right panel provides a schematic representation of typical ranges of turbulent kinetic energy dissipation rates (ε, logarithmic scale) as a function of depth.

PAGE 3 —
(3) already defined
The manuscript has been corrected accordingly.

PAGE 4 — Introduction
(1) quite redundant with the beginning of the introduction. No need to repeat.

The information has been integrated into the introductory paragraphs describing the water masses (new lines 39–43): "In the surface layers (0–150 m), Atlantic Water (AW) enters through the Strait of Gibraltar and undergoes progressive salinification as it circulates eastward and cyclonically through the western basin; below, the EIW core (200–600 m), characterized by salinities > 38.50 PSU and temperatures ~13.3°C, flows northwestward from the Eastern Basin and interacts with locally formed Western Intermediate Water (WIW) in winter". The redundant paragraph has been removed accordingly.

PAGE 4 — turbulence ranges

(2) ε and χ were already defined. Use these greek letters to shorten the text. Just cite Fig. 1.

(3) what is a natural mixed-layer wake ?

(4) add units to all measured variables in the text of the manuscript

These sentences have been corrected accordingly and revised to clarify the typical ε ranges with depth (new lines 105–112): "Oceanic turbulent kinetic energy dissipation rates (ε) span over nine orders of magnitude with depth, as schematized in Figure 1. Typical dissipation levels range from enhanced near-surface mixing ($\varepsilon \approx 10^{-6}$–$10^{-5}$ W kg^{-1}), to intermittent subsurface patches ($\varepsilon \approx 10^{-7}$–$10^{-8}$ W kg^{-1}), background thermocline values ($\varepsilon \approx 10^{-9}$ W kg^{-1}), and increasingly quiescent deep waters approaching the instrumental detection limit ($\varepsilon \approx 10^{-10}$–$10^{-11}$ W kg^{-1}). These ranges are consistent with canonical dissipation levels reported in ocean mixing studies and widely used in vertical mixing parameterizations (e.g., Gregg, 1989; Waterhouse et al., 2014). A similar order-of-magnitude range applies to the thermal variance dissipation rate χ (not shown)."

PAGE 5 — Historical context

(1) l. 110 to 142: Keep focused. These paragraphs form a lengthy general discussion that does not bring significant useful information for the present dataset and its use. Either remove it or significantly shorten it, keeping only the essential and avoiding generalities and repetition. No need to do an 'historical' review of the evolution of methodologies or instrumental capabilities. No need to provide details on the methodology, which is explained later on.

Following the reviewer's recommendation, this section has been shortened to retain only the information directly relevant to the dataset, removing general discussion and historical context (new lines 112–124): "Autonomous underwater gliders equipped with airfoil shear probes and fast-response thermistors enable concurrent estimates of ε and χ across a wide range of magnitudes during multi-week missions, providing vertically resolved observations of mechanical and thermal mixing beyond the capabilities of traditional ship-based profilers (Sherman and Davis, 1995; Eriksen et al., 2001; Peterson and Fer, 2014). Because turbulent mixing controls vertical exchanges of heat, salt, nutrients, and carbon, ε and χ are increasingly recognized as key parameters for sustained ocean observing systems and are currently considered emerging (pilot) Essential Ocean Variables (Lindstrom et al., 2012; Le Boyer et al., 2021; https://goosocean.org/what-we-do/framework/essential-ocean-variables/). Recent advances in autonomous platforms and in processing, calibration, quality control, and uncertainty estimation methodologies have made the systematic acquisition and dissemination of turbulence microstructure datasets feasible (Lueck et al., 2002; Piccolroaz et al., 2021; Lueck et al., 2024), motivating the present data compilation and its alignment with FAIR practices and community standards such as ATOMIX (Fer et al., 2024)."

PAGE 5 — Typos
(2) data acquisition
(3) no ending date for that deployment

The manuscript has been corrected accordingly.

PAGE 6 — Methods, MicroRider description
(0) this is mentioned at line 173. No need to repeat
(1) … for tranferring…
(2) … dissipate kinetic energy.
(3) … the highest…
(4) of course it does… useless
(5) Why 'at the contrary' ?

The manuscript has been corrected accordingly.

PAGE 7,8 — Data and Methods
(1) already defined in the introduction, no need to repeat
(1) Already mentioned line 227

The manuscript has been corrected accordingly.

PAGE 10 — Glider vs pressure-derived vertical velocity
(1) How does the vertical component of the velocity inferred from the GFM compare with the vertical velocity derived from the pressure change recorded by the micro rider ? What are the typical differences (std, typical extrema) between those two ? Are there some moments for which the GFM vertical velocity differs significantly from the measured pressure change in time ?
Figure 1 - supplementary section: add labels to all ordinates of the angle of attack

Thanks to the reviewer for raising this interesting point. For the processing of microstructure data, the relevant quantity is the along-path glider speed, which is used to convert shear probe measurements into physical dissipation rates. While comparing vertical velocities inferred from different approaches can provide useful insight into glider flight dynamics or larger-scale vertical motions of the water column, this comparison is not required for the quality or validity of the turbulence data processing. The vertical component was therefore not retained in the present data products.
Regarding the Figure S1, the supplementary figure has been updated by adding labels to all ordinates of the angle of attack.

PAGE 11 – Typos
(1) … noise of the vertical velocity, which eases the detection of

The manuscript has been corrected accordingly.

PAGE 12 — Thermistor QC logic
(1) is the variance calculated over the whole up- or down-cast glider section ?
(2) unit missing
(3) Variance threshold undefined

Thanks to the reviewer for this question, which allowed us to clarify missing information. We corrected the description by explicitly referring to a low-variance (stuck-sensor) test, with its empirical threshold, and by specifying that the test is applied over the current gliding section, either upcast or downcast. The text has been revised accordingly (new lines 359–361):
"If none correlates significantly, the decision is made based on the variance of each FP07 channel, using a low-variance (stuck-sensor) test on the considered gliding section. If both variances are unusually below an empirically defined threshold of 1×10^{-5} °C², indicative of a stuck or flat-line thermistor signal, the function flags potential malfunction (flag 0);"

(4) Why not comparing to the T_gl_slow variance ?

The variance comparison is used as a relative criterion between the two FP07 sensors to select the less noisy channel. A comparison with T_gl_slow is not appropriate, as the glider temperature is low-pass filtered and does not represent fast thermistor variability.

(5) not clear. Suggestion if my understanding is correct: L2_temperature_source.m sets the reference temperature profile used to calculate the kinematic viscosity of seawater. In case Th_source_logic (please define it in the preceding paragraph) equals 11, 22, 1, 2, 0, 100 or 200), the glider temperature is reference temperature.
(6) simplify: if Th_source_logic equals 10 or 20, …
(7) At this point, you don't know which of T_gl_fast or (T1_fast or T2_fast) are good, do you ? If T_gl_fast is bad how are T1_fast and T2_fast calibrated ?

Thanks to the reviewer for highlighting this ambiguity. We simplified the description of L2_temperature_source.m and explicitly defined the role of Th_source_logic, clarifying the default use of the glider temperature and the specific case (flags 10 or 20) in which an FP07 thermistor is selected instead. FP07 thermistors are fast-response sensors that are not calibrated for absolute temperature, and are used here only as a fallback reference when the glider temperature is deemed unreliable. This has been clarified in the revised text (new lines 368–382): "L2_temperature_source.m defines the reference temperature profile used to compute the kinematic viscosity of seawater required for ε estimates. The selection is based on the Th_source_logic flag assigned during the thermistor comparison step. In the vast majority of cases (Th_source_logic = 11, 22, 1, 2, 0, 100, or 200), the glider temperature interpolated on the fast channel (T_gl_fast) is retained as the reference temperature. Cases where Th_source_logic equals 10 or 20, indicating that both FP07 thermistors exhibit sufficient variance but correlate poorly with the glider temperature, are rare. They represent about 2.3% of the processed sections and 1.2% of all dissipation estimates, and account for only 0.13% of the fully validated ("QC=0") ε estimates. These occurrences are restricted to the earliest part of the dataset (year 2015). In such cases, the FP07 thermistor with the lower variance (T1_fast or T2_fast) is used instead of T_gl_fast. For these sections, the temperature offset between FP07 and glider measurements is typically on the order of ~2 °C. Over the observed

temperature range (approximately 10–20 °C), this translates into a change in kinematic viscosity of less than 4%, which has a negligible impact on the resulting ε estimates compared to other sources of uncertainty. The selected signal is stored as temperature_for_dissipation together with a descriptive string and logic flag, and is used subsequently in get_diss_odas()".

PAGE 13–14 — FFT and spectra
(1) between what and what ?
(2) Not clear. Do you mean that an averaged spectrum is formed over a time window of 12 seconds from the mean of (four FFT independent spectra of 3 seconds + 3 spectra from the overlapping segments) ?

We clarified that the consistency check refers to agreement between the two independent FP07 thermistor–derived χ estimates (new lines 415–417): "As with ε, χ estimates are quality controlled using statistical thresholds and consistency between the two independent FP07 thermistor–derived estimates."
We clarified then about the FFT lengths used for estimations (new lines 428–433): " FFT spectra are first computed on short segments length (N_fft) of typically 3 s duration. Dissipation estimates are then obtained by averaging spectra over a longer window composed of Ntimes=4 FFT segments, using a 50% overlap between consecutive segments. This results in an effective averaging window of approximately 12 s and corresponds to averaging seven individual FFT spectra for each dissipation estimate

PAGE 14 — Temperature-gradient spectra
(1) Could you explicitly mention what fitting method is used ? Is it the MLE ?
(2) Please indicate if χ-estimates use ε-estimates from shear sensors (in which case this input is missing in your list) or if both χ and ε are estimated from MLE fits, as is usually done with temperature-gradient spectra and MLE method ? What are the limitations, given that the downcast velocity prevents resolving the temperature-gradient spectral peak and roll-off for enhanced ε ?

Thanks to the reviewer for this comment. We clarified that theoretical spectra are fitted using a Maximum Likelihood Estimation (MLE) method, and that χ estimates (and an independent ε estimate) are obtained solely from temperature-gradient spectra, without using shear-derived ε as input. We added a short statement acknowledging the limitation imposed by downcast speed on resolving temperature-gradient spectral roll-off at high turbulence levels, and clarified that estimate reliability is handled through the existing MLE-based quality metrics and QC flags. The text has been revised accordingly (new lines 443–455): "Spectral analysis of temperature gradient is performed in L3_spectra_th.m using the gradT_dis_spec() routine from Piccolroaz et al. (2021) to estimate temperature variance dissipation rates (χ). Inputs include pressure, vertical temperature gradients (dT/dz), and required parameters such as the kinematic viscosity of seawater and spectral models. Segments of the temperature gradient signal are analyzed, and theoretical Batchelor or Kraichnan spectra are fitted to the observed spectra using a Maximum Likelihood Estimation (MLE) method following Ruddick et al. (2000). The MLE fit is used both to correct χ for unresolved variance outside the integration range and to provide an independent estimate of ε derived from temperature-gradient spectra, without using shear-derived ε as input. Quality metrics returned by the routine include the Mean Absolute Deviation (MAD), the wavenumber range used for the fit, likelihood ratios, and QC flags for spectral fits.

Profiles of χ are obtained independently for each FP07 sensor. For enhanced turbulence levels, the downcast glider speed may limit the resolution of the high-wavenumber roll-off of the temperature-gradient spectra; in such cases, the reliability of χ and ε estimates is evaluated using the MLE-based quality metrics and associated QC flags.", and "Although shear-derived ε estimates are available elsewhere in the dataset, they are not used to constrain χ estimation or the associated spectral QC, which relies exclusively on temperature-gradient MLE fits following Piccolroaz et al. (2021), and we employ the QC flag in output from their routines,…"

(3) Is this criterion from Piccolroaz et al. appropriate here if you provide ε estimates from shear sensors ? (In Piccolroaz et al., the uncertainty of MLE fits strongly depends on resolving temperature-gradient spectral peaks and rolls-off when ε-estimates are not provided from shear sensors)

The Piccolroaz et al. (2021) QC criteria remain appropriate here because χ estimates, together with an independent ε estimate, are derived solely from temperature-gradient spectra using the MLE framework. Shear-derived ε is not used as an input to the χ estimation, so the applicability and limitations discussed in Piccolroaz et al. directly apply. This point has been clarified in the manuscript

PAGE 15 — QC table comments
(1) not fully clear: failure condition is 'none' but, as I understand, Flags 4 and 64 have not been passed yet. So a bit misleading. It could be better to indicate that 'none' refers to an 'individual' estimate that passed 'individual-QC' #1,2,8,16,32 but has not passed 'coupled-QC' #4, 64 yet.
(2) change e1 vs e2 into ε1 vs ε2
(3) the largest among ε1 and ε2…

Thanks to the reviewer for pointing out this ambiguity. We clarified that QC flags are assigned on a probe-wise basis and that inter-probe ratio and method-consistency checks (flags 4 and 64) are already included in the final bitwise QC flag, potentially rejecting only one of the two estimates. The table and text have been revised accordingly, in the same time than the typos (new lines 477–481): "The final QC flag is a bitwise sum of failed tests. In the QC table, the failure condition "none" refers to an individual probe estimate for which no QC test has failed. All individual and inter-probe consistency checks (including ratio and method-consistency tests, flags 4 and 64) are applied at the probe level. As a result, inter-probe tests may reject only one of the two estimates (ε_1 or ε_2), while the other can remain valid."

PAGE 16
(1) Why chosing different segment lengths ?

Thanks to the reviewer for this question. The same segment definitions are prescribed, while the internal segment handling differs between the two different processing codes from RSI and Piccolroaz et al. (2021), leading to slight differences in the number and alignment of estimates within a given gliding section. This may result in one additional or missing estimate and a non-exact alignment in depth or time, which can be reconciled a posteriori after QC filtering if needed.

PAGE 19 — QC counting
(1) Fig. 3 a,b,c,d: to improve immediate readability, indicate the shear probe # in abcissa labels

The figure has been corrected accordingly.

(2) The counting is interesting. If I understand, only 10% of shear 1 and 2 passed all individual QC and provide dissipation rate estimates that agree within a factor of 2.72 σ(ln ε). How does this compare to other deployments or platforms ? Why such a low percentage ? Is it due to the occurence of broken shear sensor during glides, loss of shear sensor sensitivity during long deployments, ... ? I was also quite surprised by the huge scatters of Figure 4ab. Why such a scatter between sensors (either shear or temperature) ?

Thanks to the reviewer for this comment. The relatively low fraction of cross-validated estimates reflects a combination of sensor degradation and variable flight conditions challenging to constrain during long deployments, rather than an intrinsic limitation of the processing. Based on data inspection and available deployment information, shear and FP07 probes were often degraded or damaged as deployments progressed, and replaced between deployments, which strongly impacts inter-sensor consistency. In addition, glider flight conditions were not optimized for microstructure sampling and include periods of unsteady pitch and speed, which contribute to the observed scatter between sensors. As a result, stricter cross-checks naturally lead to lower retention rates than typically reported for dedicated microstructure platforms or short, well-controlled deployments. These limitations are now explicitly acknowledged in the manuscript (lines 638-639): "The low proportion of cross-validated estimates primarily reflects sensor degradation during deployments and variable glider flight conditions, rather than a lack of usable individual measurements."

PAGE 24 — PDFs (ε, χ)
(1) I don't see any bimodal distribution in summer (red line).
(2) spring is just like summer and autumn. Winter differs.
(3) Both diapycnal mixing and isopycnal mixing have an impact on χ. Isopycnal stirring has a stronger influence on χ in some oceanic regions depending on the mesoscale activity, the existence of frontal region with possible intrusion processes for instance.

Thanks to the reviewer, we corrected the seasonal interpretation by removing the bimodal description for summer, assigning bimodality to autumn, and revising the ε ranges with explicit physical units (W kg^{-1}) (lines 701-715): "Seasonal decomposition shows distinct patterns. Spring and summer exhibit relatively weak dissipation levels, with distributions centered between $\varepsilon \approx 10^{-11}$ and 10^{-10} W kg^{-1}. Autumn shows a bimodal distribution, with a primary mode similar to spring and summer ($\varepsilon \approx 10^{-11}$–$10^{-10}$ W kg^{-1}) and a secondary mode at higher dissipation levels ($\varepsilon \approx 10^{-9}$–$10^{-8}$ W kg^{-1}). Winter is characterized by substantially higher dissipation rates, typically spanning $\varepsilon \approx 10^{-8}$–$10^{-6}$ W kg^{-1}, although based on a smaller number of observations. These patterns suggest that autumn captures intermittent energetic events superimposed on a generally weakly dissipative background, while winter conditions favor more frequent and intense turbulence associated with enhanced vertical mixing.

For χ, the seasonal contrast is weaker. Spring, summer, and autumn display broadly similar distributions, whereas winter stands out with systematically lower χ values. χ is influenced by both diapycnal mixing and isopycnal stirring, and its seasonal variability likely reflects changes in stratification as well as mesoscale and frontal activity. The reduced χ levels observed in winter are consistent with a weaker stratification and a reduced role of isopycnal stirring, while enhanced χ during the stratified seasons may reflect a combination of small-scale turbulence and lateral stirring processes."

PAGE 26 — Conclusions

(1) I am a bit confused here. Are you talking about the final QC (l.462) or QC flags as in previous sections.

Thanks to the reviewer for this comment. We clarified that QC=0 refers to probe-wise validity of individual estimates in the final L4 product, while cross-checked subsets require both probes to simultaneously reach QC=0. The text has been revised to remove any ambiguity between probe-wise and paired validation (new lines 723–735): "Quality control flags discussed here refer to the final L4 QC assigned on a probe-wise basis. The final QC flag is computed as a bitwise combination of all applied quality-control tests. Estimates flagged as QC=0 correspond to individual shear-probe estimates for which no QC test has failed. This includes spectral fit quality (figure of merit $\leq$ 1.4), spike contamination, variance resolution, angle-of-attack constraints, and inter-probe ratio and method-consistency checks, noting that these latter tests may reject only one of the two paired estimates. QC=0 therefore indicates a valid estimate for a given probe, while stricter cross-checked subsets additionally require both probes to simultaneously reach QC=0. QC=0 estimates form the core of the L4 product and can be used with high confidence in scientific analyses.

The broader category QC $\leq$ 1 includes both QC=0 and QC=1 estimates. QC=1 flags estimates for which the figure of merit exceeds 1.4, indicating a poorer fit to the theoretical spectral model (e.g., Nasmyth or Lueck), while other QC criteria remain satisfied. As discussed by Lueck et al. (2024), no universal threshold exists for the figure of merit, as its interpretation depends on platform and environmental conditions."

PAGE 31 — Other questions

– Did you check if the sensitivity of shear/temperature sensor changed during long-term deployments ? (sensor calibration before and after deployments)

Thanks to the reviewer for this question. Sensors were typically installed new at the start of each deployment and replaced between missions when degraded or broken, which prevented a systematic assessment of long-term drift within deployments. Given the limited persistence of high-quality microstructure measurements over long missions, sensor drift was not a primary limiting factor for this dataset and is not further analyzed.

Reviewer 2

Summary of the paper

The paper "Advancing Turbulence Essential Ocean Variable: A Reference Glider-Based Microstructure Dataset from the Western Mediterranean" presents a decade-long (2015–2024) dataset of turbulence measurements collected by a Slocum Deep Glider equipped with a Rockland MicroRider. The missions repeatedly crossed the Sardinia–Balearic transect, producing over 3,000 high-quality profiles of turbulent kinetic energy dissipation (ε) and thermal variance dissipation (χ) down to 1000 m depth. The dataset, processed through standardized L0–L4 levels and validated against strict QC metrics, represents one of the largest autonomous microstructure archives in the Mediterranean.

Review summary

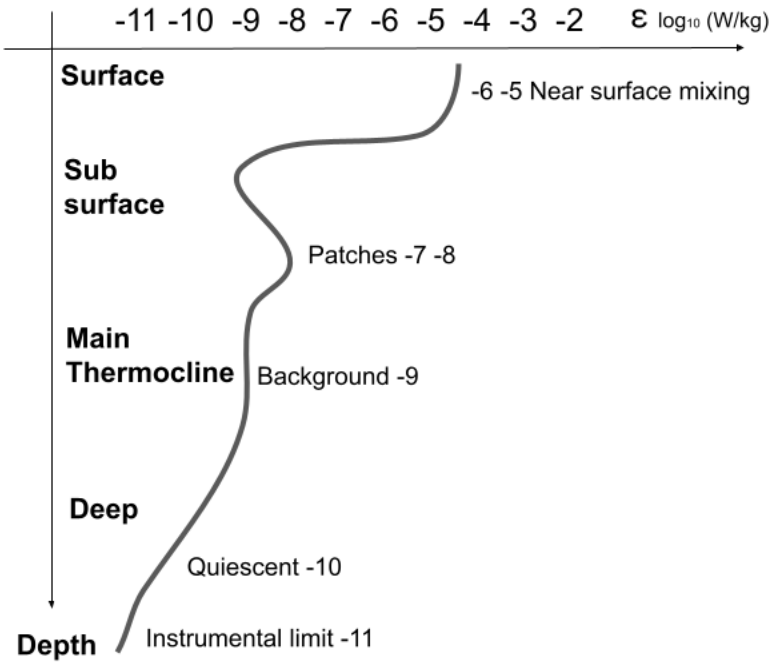
I am OK to publish the paper with minor revisions. The dataset is of high value and well-documented. The main improvements should focus on clarifying and tightening the introduction, replacing subjective or informal wording, and adding representative turbulence spectra to better demonstrate data quality and satisfy the expectations of turbulence specialists.

We thank Reviewer 2 for the careful reading of the manuscript and for the positive assessment of the dataset. We appreciate the constructive comments and suggestions, which helped us to clarify and tighten the introduction, improve the wording, and better document data quality. We respond below to each point raised by the reviewer. All modifications made in the manuscript are indicated in blue text.

Fig. 1 & L70–110 – Replace "expected" ε profiles with observed average data; clarify whether values are measured or estimated.

L105–110 – Rephrase "abyssal peace" and "natural mixed layer"; use neutral scientific language.

We chose to retain a schematic representation rather than a mean profile, as a mean vertical profile would provide an overly smooth view that does not adequately represent the patchy and intermittent nature of turbulent events encountered throughout the water column. The new schematic has been corrected and simplified as presented below, to better align with the synthetic vertical structure shown in Fig. 8 (renumbered, previously Fig. 7). The manuscript text, the Fig. 1 and its caption have been revised accordingly, including an updated description of the typical ε ranges with depth.



The new caption is :

Figure 1. Glider Teresa mission along its longitudinal transect, by years (in colors). The scheme indicates the deployment between Sardinia and Balearic Islands in the Western Mediterranean Sea. Sawing black lines indicate a schematic trajectory from surface up to 1000m-depth, encountering layers of the Atlantic Water (AW) in surface, the Eastern Intermediate Water (EIW) and the Western Mediterranean Deep Water (WMDW). Depth levels associated to these water masses (respectively 0-200m, 200-1000m, 1000m+) are indicative. The lower-right panel provides a schematic representation of typical ranges of turbulent kinetic energy dissipation rates (ε, logarithmic scale) as a function of depth.

L83–100 – Add citations for regional circulation and mesoscale dynamics; clarify that the author is not an expert in the region.

We thank the reviewer for this comment. We have revised this paragraph to include additional references documenting regional circulation and mesoscale dynamics, and we clarified that the description is not intended as a comprehensive regional synthesis but rather as contextual background for the dataset (revised lines 86–89): "The circulation framework of this region has been described in several foundational studies of the western Mediterranean, which highlight the role of interconnected currents, boundary exchanges, and water-mass transformations linking the Balearic and Sardinian sub-basins (Millot, 1999; Send et al., 1999; Pinardi et al., 2015)"

Millot, C.: Circulation in the Western Mediterranean Sea, Journal of Marine Systems, 20, 423–442, https://doi.org/10.1016/S0924-7963(98)00078-5, 1999.

Pinardi, N., Zavatarelli, M., Adani, M., Coppini, G., Fratianni, C., Oddo, P., Simoncelli, S., Tonani, M., Lyubartsev, V., Dobricic, S., and Bonaduce, A.: Mediterranean Sea large-scale low-frequency ocean variability and water mass formation rates from 1987 to 2007: A

retrospective analysis, Progress in Oceanography, 132, 318–332, https://doi.org/10.1016/j.pocean.2013.11.003, 2015.
Send, U., Font, J., Krahmann, G., Millot, C., Rhein, M., and Tintoré, J.: Recent advances in observing the physical oceanography of the western Mediterranean Sea, Progress in Oceanography, 44, 37–64, https://doi.org/10.1016/S0079-6611(99)00020-8, 1999.

L120–135 – Include citations for turbulence parameterization and EOV framework; clarify what "expected dissipation levels" mean.

This paragraph has been revised and clarified; appropriate references have been added accordingly (new lines 105–124). "Oceanic turbulent kinetic energy dissipation rates (ε) span over nine orders of magnitude with depth, as schematized in Figure 1. Typical dissipation levels range from enhanced near-surface mixing ($\varepsilon \approx 10^{-6}$–$10^{-5}$ W kg^{-1}), to intermittent subsurface patches ($\varepsilon \approx 10^{-7}$–$10^{-8}$ W kg^{-1}), background thermocline values ($\varepsilon \approx 10^{-9}$ W kg^{-1}), and increasingly quiescent deep waters approaching the instrumental detection limit ($\varepsilon \approx 10^{-10}$–$10^{-11}$ W kg^{-1}). These ranges are consistent with canonical dissipation levels reported in ocean mixing studies and widely used in vertical mixing parameterizations (e.g., Gregg, 1989; Waterhouse et al., 2014). A similar order-of-magnitude range applies to the thermal variance dissipation rate χ (not shown). Autonomous underwater gliders equipped with airfoil shear probes and fast-response thermistors enable concurrent estimates of ε and χ across a wide range of magnitudes during multi-week missions, providing vertically resolved observations of mechanical and thermal mixing beyond the capabilities of traditional ship-based profilers (Sherman and Davis, 1995; Eriksen et al., 2001; Peterson and Fer, 2014). Because turbulent mixing controls vertical exchanges of heat, salt, nutrients, and carbon, ε and χ are increasingly recognized as key parameters for sustained ocean observing systems and are currently considered emerging (pilot) Essential Ocean Variables (Lindstrom et al., 2012; Le Boyer et al., 2021; https://goosocean.org/what-we-do/framework/essential-ocean-variables/). Recent advances in autonomous platforms and in processing, calibration, quality control, and uncertainty estimation methodologies have made the systematic acquisition and dissemination of turbulence microstructure datasets feasible (Lueck et al., 2002; Piccolroaz et al., 2021; Lueck et al., 2024), motivating the present data compilation and its alignment with FAIR practices and community standards such as ATOMIX (Fer et al., 2024)."

Gregg, M. C.: Scaling turbulent dissipation in the thermocline, J. Geophys. Res., 94, 9686–9698, https://doi.org/10.1029/JC094iC07p09686, 1989.

Waterhouse, A. F., MacKinnon, J. A., Nash, J. D., Alford, M. H., Kunze, E., Simmons, H. L., Polzin, K. L., St. Laurent, L. C., Sun, O. M., Pinkel, R., Talley, L. D., Whalen, C. B., Huussen, T. N., Carter, G. S., Fer, I., Waterman, S., Naveira Garabato, A. C., Sanford, T. B., and Lee, C. M.: Global Patterns of Diapycnal Mixing from Measurements of the Turbulent Dissipation Rate, Journal of Physical Oceanography, 44, 1854–1872, https://doi.org/10.1175/JPO-D-13-0104.1, 2014.

L125–130 – Add explicit mention of spectral methods and noise floor corrections.

Under the recommendation of Reviewer 1 to simplify and focus the introduction, this part of the text was removed rather than expanded. Spectral methods and noise-floor corrections are now described later in the Methods section.

L165–195 (Data & Methods) – Define Level 0 clearly here; improve flow—some content fits better under "Processing."

Following the reviewer's suggestion, we clarified the definition of Level 0 data at the beginning of the Data & Methods section and improved the flow by focusing this paragraph on raw measurements and instrumentation. Processing-related aspects have been moved to the dedicated processing section. Revised text (new lines 145–160):"Level 0 (L0) data consist of raw, high-frequency time series recorded by the MicroRider prior to any processing or spectral transformation. Shear and thermistor sensors provide turbulence measurements sampled at 512 Hz, which are internally logged by the Rockland MicroRider (model 1000-LP), a microstructure module designed for deployment on moving and fixed platforms such as gliders, moorings, or wire walkers. The instrument hosts two orthogonally oriented shear probes (sh1, sh2) mounted on the front bulkhead, measuring the vertical gradients of the horizontal velocity fluctuations, $\frac{\partial u'}{\partial z}, \frac{\partial v'}{\partial z}$ where the prime denotes deviations from the mean flow. Two fast-response FP07 thermistors similarly resolve small-scale vertical temperature gradients $\frac{\partial T'}{\partial z}$. These sensors capture variance at spatial scales ranging from centimeters to decimeters, where turbulent motions are energetic, down to millimeter scales where where viscosity will act to finally dissipate kinetic energy. In addition, the MicroRider includes a pair of piezo-accelerometers to monitor platform vibrations and a two-axis inclinometer providing pitch and roll angles with a nominal accuracy of 0.1°, allowing characterization of instrument dynamics during flight. For reliable shear-based turbulence measurements, the glider incident speed is expected to remain within approximately 0.2–0.6 m s^{-1}, fast enough to satisfy Taylor's frozen-turbulence hypothesis while slow enough to resolve the highest wavenumbers. Subsequent processing steps, including spectral analysis and dissipation estimation, are described in the following processing section."

L175–180 – Replace "averaging involved" by "segmentation or windowing"; clarify that raw data are not averaged prior to spectra.
L195–200 – "On the contrary" instead of "At the contrary."

We replaced "averaging involved" with "segmentation and windowing" and clarified that raw time series are not averaged prior to spectral computation. The text has been revised accordingly to better describe the processing workflow. Revised text (new lines 190–193): "The integration of turbulence signals relies on spectral segmentation and windowing using overlapping windows—commonly four 3-second segments—rather than on averaging of the raw time series. Spectra are computed for each segment and subsequently combined, yielding an effective vertical resolution on the order of 1.5 m. This processing step substantially reduces data volume, with final profile products typically ranging between 1 and 10 MB per file."

L200–225 – Explicitly state which spectral model and integration limits are used for ε and χ (Nasmyth vs Batchelor); mention the electronic noise floor.
L230–240, Fig. 2 – Cite Lueck et al. 2024 explicitly when stating the workflow follows it "exactly."

We clarified which spectral models and integration limits are used for ε and χ, and explicitly mentioned the role of the electronic noise floor in defining the usable wavenumber range. We also added an explicit citation to Lueck et al. (2024) where the workflow is stated to follow that reference**.** Revised text (new lines 208–214): "Due to non-turbulent variance and electronic noise present in the signal, spectra are only integrated over a restricted wavenumber range where turbulent variance dominates above the instrument noise floor. For shear-derived dissipation, the well-resolved portion of the shear spectrum is fitted to the reference empirical spectrum (Nasmyth, 1970; Osborn and Crawford, 1980; and in our case Lueck et al., 2002), and the fit is used to correct for unresolved variance outside the integration limits. Dissipation rates ε_1 and ε_2 are obtained independently from shear probes sh1 and sh2." Revised text (new lines 226–228): "The overall processing workflow follows the methodology described in Lueck et al. (2024), including spectral segmentation, noise-floor treatment, model fitting, and quality-control procedures, and is summarized in Fig. 2."

L250–265 – In Table 1, note that this is a list of routines, not processing "steps."

The manuscript has been corrected accordingly.

L480–495 – Simplify data structure description; too detailed for ESSD, could be moved to Supplement.

We simplified the description of the data structure to focus on the essential elements relevant to ESSD readers. Detailed information on internal MATLAB structures and file-level implementation has been shortened and moved to the Supplementary Material. Revised text (L498–506):"The script make_structure.m compiles metadata, processing parameters, quality-control diagnostics, and turbulence estimates into structured MATLAB files for each glider profile section. Two levels of data products are generated: a comprehensive internal structure containing all intermediate signals, spectra, and diagnostics, and a lighter structure retaining only essential metadata, processing parameters, final dissipation estimates (ε and χ), and QC flags. To balance data traceability with storage and computational constraints, only the lighter product is published, while the full structure is retained internally for reproducibility and potential reprocessing. Each section is exported as an individual .mat file using the -v7.3 format and named following a standardized convention encoding key metadata (mission, time, location, depth range, and glider direction, see Supplementary Text 1)."

L545–570 (QC section) – Explain rationale for focusing on QC=2 rather than others; the purpose and statistical relevance are unclear.

Thanks to the reviewer for this clarification. In the approach described by Lueck et al. (2024), despiking is applied at the section level, while spike indices can subsequently be used to compute spike fractions at the dissipation-length segment level. In our processing, despiking is likewise performed per section, but the spike fraction is evaluated and applied at the section level rather than redistributed across individual dissipation segments. Given the observed distribution of spikes, which are typically sparse and localized, this choice is not expected to materially change which dissipation estimates are retained or rejected. The rationale for this implementation and the role of QC=2 and QC=32 have been clarified in the revised text (L553–558):"We implemented quality control metrics related to despiking following the section-based approach described by Lueck et al. (2024). Despiking is applied to each gliding section prior to spectral analysis, and spike occurrences are identified during this step. In principle, the resulting spike indices can be used to compute spike fractions at the dissipation-length segment level. In the present processing, spike statistics are instead summarized at the section level and the corresponding QC flags are applied to the dissipation estimates derived from that section.

A spike fraction test assigns QC=2 when more than 15% of the clean section data points are affected by despiking. This flag indicates that core validation criteria—acceptable spectral fit, agreement between probes, and sufficient resolved variance—are satisfied, but that elevated spiking is present. Such cases are not necessarily invalid, as spikes may arise from dense biological layers, brief mechanical disturbances, or localized contamination superimposed on otherwise valid turbulent signals. A second despiking diagnostic assigns QC=8 when more than nine despiking iterations are required. As noted by Lueck et al. (2024), there is no strong consensus on acceptable iteration limits, and ODAS processing caps the number of iterations by default, which limits the interpretive value of this metric alone.

Our statistical analysis (Fig. 3) reveals a subset of estimates with spike fractions between 5% and 15% and iteration counts below the rejection threshold. To avoid overly conservative rejection of such cases, we introduce a relaxed quality flag, QC=32, marking these estimates as conditionally valid. For long glider profiles extending to 1000 m, localized spikes may occur without substantially affecting the overall spectral estimate. Applying a strict section-level rejection in these cases would unnecessarily reduce data availability. The QC=32 flag therefore provides a practical compromise, retaining these estimates for secondary analyses while clearly identifying them as non-core data and leaving the final selection to the user."

L560–565 – Use FOM = 1.15 as the nominal threshold (per Lueck et al. 2024) and note that users can apply their own.

We followed the FOM threshold of 1.4, as recommended in Lueck et al. (2024) (pp15-16), who explicitly note that no robust universal threshold can be defined and suggest 1.4 as a reasonable practical limit. We clarified this point in the text and explicitly state that users are free to apply alternative thresholds depending on their application. Revised text (L578–583): "Despite its frequent use, there is no universal consensus on an optimal figure-of-merit (FOM) threshold, as its interpretation depends on probe characteristics, platform type, and environmental conditions (Lueck et al., 2024). Following the guidance of Lueck et al. (2024), we adopt a nominal threshold of FOM = 1.4 as a practical limit, while acknowledging its subjective nature. Users of the dataset are encouraged to apply alternative FOM thresholds if more appropriate for their specific scientific objectives or observational context".

L565–570 – Remove statements implying that removal is "unnecessary"; just report the data and FOM values.

We followed the reviewer's recommendation by removing subjective statements regarding "unnecessary" data removal and by clearly accepting the applied QC criteria as part of the published dataset. The text now emphasizes transparency and user-driven data selection through explicit QC flags, rather than justifying alternative retention strategies. Revised text (L585-588): "All dissipation estimates are published together with their associated QC flags, which explicitly document their quality status. This approach ensures transparency and allows users to apply their own selection criteria depending on their scientific objectives, while fully respecting the quality-control thresholds adopted in this dataset."

Fig. 3–4 – Add a few examples (panels or zoomed insets) illustrating good vs problematic QC spectra. Additional suggestion – Include representative shear and thermistor spectra in the Results or Supplement (always appreciated by turbulence experts) + Additional suggestion – Include representative shear and thermistor spectra in the Results or Supplement (always appreciated by turbulence experts).

We have added a new figure (Figure 4) presenting representative examples of accepted and rejected shear and thermistor-based spectra, illustrating the application of the spectral fitting, integration ranges, and QC criteria discussed in the text. This new figure has been included and described in an additional subsection to better demonstrate data quality and the rationale behind the QC flags. Given this insertion, the following figures and sub sections have been renumbered accordingly.

Minor editorial – Fix "than" → "then" (L580), typos like "enventually," "affixable," and consistent use of italics for ε and χ.

The manuscript has been corrected accordingly.