



Corresponding author: a.czaja@imperial.ac.uk

The “scatter” of the stratification: a simple metric to study the finescale structure of the oceanic density field

Arnaud Czaja¹, Pascale Lherminier², and Herlé Mercier³

¹Imperial College, London, Department of Physics, Huxley Building, Room 726, SW7 2AZ London, UK.

^{2,3}Université de Brest, CNRS, Ifremer, IRD, Laboratoire de Physique des Océans, IUEM, 29280, Plouzané, France.

Abstract. A new metric, hereafter the “scatter” of the stratification, is introduced to study the finescale structure present in the observed oceanic density field. It is based on the linear regression of buoyancy against depth over a layer of specified thickness $h = O(10 - 100m)$. The scatter η is defined as the fraction of the total buoyancy variance contained in the residuals
($0 \leq \eta \leq 1$). Thus a small scatter indicates that the buoyancy profile is linear to a very good approximation on the vertical scale h , while it is a poor description of the buoyancy field on that scale when the scatter approaches unity. The new metric includes a formal method to estimate instrument errors on buoyancy, and to isolate the scatter arising from physical effects from the scatter due to instrument errors. The scatter can also be related to the potential energy (E_m) present in the buoyancy field on vertical scales within the $[h/2, 2h]$ interval. Specifically, it compares this energy with that (E_m^*) needed to remove the stratification of the layer (low scatter when $E_m \ll E_m^*$, large scatter when E_m is comparable or larger than E_m^*).

The method is applied to a set of repeated hydrographic sections in the North Atlantic between Greenland and Portugal (OVIDE 2002-2023). It is found that the background Garrett-Munk (GM) internal wavefield contributes to low values of scatter ($\eta < 0.1$ with $E_m \sim E_m^{GM} \ll E_m^*$), that thermohaline intrusions in the Irminger Sea raise the scatter slightly above this value ($0.1 \leq \eta \leq 0.2$ with $E_m^{GM} \leq E_m \leq E_m^*$), while the interaction of tidal (Reykjanes Ridge) and mesoscale (East Thulean Rise) flows with topography, the western boundary current regime of the East Greenland - Irminger Current and double diffusive convective events below the core of the Mediterranean Water all increase it even further ($\eta > 0.2$ and $E_m^{GM} \ll E_m \sim E_m^*$). The Iberian Abyssal Plain is found to be a region where the stratification is so low that only a small amount of wave activity is needed to create a large scatter. The probability density function of η does not show gaps but changes shape depending on the value of h .

It is hoped that the scatter of the stratification can offer a useful benchmark for high resolution numerical models of the oceans, and can help detect and understand changes in oceanic stratification, particularly in the abyss where finite instrument errors and low stratification pose significant challenges, among many other possible applications of the method.

1 Introduction

In his seminal review of internal waves, Munk (1981) emphasized how it has been known from the early days of oceanography that measured temperature and salinity profiles as a function of depth were “noisy”. Many types of motions can indeed con-



tribute to generating displacements of isotherms, isohalines and isopycnals: fast dynamics associated with internal waves, tides and high frequency submesoscale (Torres et al., 2022), balanced motions of either the submesoscale or mesoscale flavour (e.g., Müller, 1988), thermohaline intrusions (e.g., Shcherbina et al., 2009) and double diffusive convection (e.g., Radko, 2013). As the horizontal and vertical resolutions of numerical models of the ocean increase, they are becoming able to simulate these different class of motions, both in regional (e.g., Mazloff et al., 2020; Baker and Mashayek, 2022) and global (e.g., Qiu et al., 2018) set-ups. It is the aim of this study to provide a simple observational metric, building on the rich set of high quality, repeated, global hydrographic observations (Talley et al., 2016; Katsumata et al., 2022), which will help in validating the amount of “noise” generated by these models from the sea surface to the ocean sea floor.

With their typical vertical resolution of a few decibars hydrographic observations have allowed to study the variability of the density field on vertical scales of a few tens to a few hundred meters, the so-called “finescale” (e.g., Munk, 1981; Pinkel, 2014). Much interest has been put in particular on estimating turbulent kinetic energy dissipation rates and vertical diffusivity from the global hydrographic dataset, with the aim of informing parametrisations of mixing in Ocean General Circulation Models (OGCMs). In doing so the assumption that most of the finescale structure reflects the presence of a weakly interacting internal wave field is invoked (e.g., Gregg and Kunze, 1991; Polzin, 1995; Kunze, 2017). This assumption is however questioned in recent studies, where the use of both temperature, salinity and shipboard or lowered ADCP data or EM-APEX floats, allows a decomposition of the kinetic and potential energy spectra into internal waves and balanced, or “vortical” (Muller, 1984), modes of motion, revealing significant contributions from the latter on $O(10-100m)$ vertical scale (e.g., d’Asaro and Morehad, 1991; Kunze, 1993; Lien and Sandford, 2019; Vladioiu et al., 2024).

The use of ADCP or EM-APEX float observations, while a powerful tool to disentangle the dynamics, introduces further complications for the user. For example, the vertical resolution of these datasets and that of temperature, salinity data generally differ and, except for GOSHIP data where ADCP and CTD instruments are jointly deployed, there is a significantly larger amount of temperature and salinity data available globally and historically, especially since the development of ARGO floats. With this in mind, and with the goal of providing a straightforward diagnostic for high resolution ocean models, we seek here to build a metric of the (potential) energy contained in the finescale solely based on the temperature, salinity and pressure measurements.

The paper is structured as follows. We start in Section 2 by introducing the new metric, which we refer to as the “scatter of the stratification”. A formal definition and a physical interpretation of this quantity in terms of “finescale potential energy” is developed, followed by illustrative examples of its application to monochromatic waves in the linear and non linear regime, as well as to a field of random linear internal waves. The important issue of instrument errors on temperature, salinity and pressure is also addressed in this section, and we suggest how one can separate in the hydrographic data fluctuations of density associated with motions from those more likely resulting from instrument errors. In Section 3 the method is applied to a set of repeated hydrographic observations in the North Atlantic. A discussion and conclusion section are offered in Sections 4 and 5, respectively.



2 Methodology

2.1 The scatter of the stratification

2.1.1 Definition

We assume given a set of hydrographic vertical profiles, or casts, providing measurements of conservative temperature Θ and absolute salinity S_A (hereafter simply written as S) as a function of pressure p . We consider a pressure layer of thickness Δp centred on a value $p = p_o$ and partitioned into $2K + 1$ pressure levels with resolution $\delta p = \Delta p / 2K$:

$$p_k = p_o + (k - K - 1)\delta p, \quad \text{for } k = 1, \dots, 2K + 1. \quad (1)$$

At each of these levels, a buoyancy b_k is calculated from the potential density referenced to the pressure p_o (Millard et al., 1990; McDougall and Barker, 2014):

$$b_k = -g\rho\{S_k, \Theta_k, p_o\} / \rho_o, \quad \text{for } k = 1, \dots, 2K + 1 \quad (2)$$

where $g = 9.81 \text{ m s}^{-2}$, $\rho\{S, \Theta, p\}$ is the functional giving density ρ as a function of the state variables and ρ_o is a reference density for the layer. The geometric depth z_k at the pressure level p_k where b_k is calculated is obtained from the hydrostatic balance:

$$z_k = -\left(\frac{p_k - p_o}{g\rho_o}\right), \quad \text{for } k = 1, \dots, 2K + 1. \quad (3)$$

The geometric thickness of the layer is denoted by $h = (\Delta p / \rho_o g)$. Using the subscript m to denote vertical averaging over the layer, i.e., $(\cdot)_m = \sum (\cdot)_k / (2K + 1)$, we note that $z_m = 0$.

There are several possibilities to set ρ_o in (2). If repeated casts are available, one could choose ρ_o to be the long term mean density of the layer (referenced to $p = p_o$). This is in effect the approach taken in Section 2.2, where a background stratification is assumed, and buoyancy perturbations are prescribed. Alternatively, and this is the choice made here for the analysis of observations in Section 3, one can simply set ρ_o to be the averaged density of the layer, i.e., $\rho_o = \rho_m = \sum \rho\{S_k, \Theta_k, p_o\} / (2K + 1)$.

Performing a linear regression of the b_k s against the z_k s produces:

$$b_k = q + sz_k + e_k, \quad \text{for } k = 1, \dots, 2K + 1 \quad (4)$$

where q and s are, respectively, the offset and slope of the regression, while e_k is a set of residuals or “errors”. Note that s is just the squared Brunt-Väisälä frequency of the layer while q is also equal to the averaged buoyancy b_m of the layer. We define the scatter of the stratification $0 \leq \eta \leq 1$ as the ratio of variances:

$$\eta \triangleq \frac{\sum e_k^2}{\sum (b_k - b_m)^2}. \quad (5)$$



Accordingly, the success of the linear model (4) is quantified by low values of scatter, while its relevance becomes questionable when the scatter becomes large. For example, a value $\eta = 0.1$ indicates that 90 % of the variance of the buoyancy field within the layer is explained by this linear approximation.

It is useful to develop further (5) to link it explicitly to the measured stratification (i.e., the term sz in the linear regression).

90 Using $q = b_m$ and acknowledging that the residuals are decorrelated from the depth levels, i.e. $\sum z_k e_k = 0$, eq. (5) is rewritten as:

$$\eta = \frac{\sum e_k^2}{s^2 \sum z_k^2 + \sum e_k^2}. \quad (6)$$

Noticing further that the layer thickness is $h = 2K \delta p / (\rho_o g)$, we obtain, after using (1) and (3), that $\sum z_k^2 = h^2 (2K + 1)(K + 1) / (12K) \approx h^2 (2K + 1) / 12$, and so (6) can be rewritten as:

$$95 \quad \eta = \frac{\sigma_e^2}{\sigma_e^2 + s^2 h^2 / 12}, \quad (7)$$

where we have introduced the variance of the buoyancy residuals $\sigma_e^2 = \sum e_k^2 / (2K + 1)$. The form (7) will prove useful to understand the physical processes controlling the scatter and to discuss instrument errors in subsequent sections.

Finally, while the forms (5) or (7) make a clear connection to the variance of the residuals and that captured by the linear approximation $q + sz_k$, it can also be recognised that $\eta = 1 - r^2$ where r is the correlation coefficient between the b_k s and the

100 z_k s. This allows a straightforward estimate of the scatter η using standard linear regression packages.

2.1.2 Physical interpretation

The buoyancy variances appearing in (5)-(7) bring to mind an interpretation of the scatter in terms of (gravitational) potential energy. Following Holiday and McIntyre (1980), we consider the energy $E(z)$ required to create, adiabatically, at a given depth z , a “buoyancy excess” $b_e(z)$ from a reference state $b_o(z) = N_o^2 z$. A simplified form for E is introduced in Appendix A,

105 namely:

$$E \approx \frac{1}{2} \rho_o \frac{b_e^2}{N_o^2} \quad (8)$$

which allows to compute the average potential energy E_m of the layer as:

$$E_m \approx \frac{1}{2} \rho_o \left(\frac{1}{h} \int_{-h/2}^{h/2} b_e^2 dz \right) / N_o^2. \quad (9)$$

A natural benchmark for E_m is the energy E_m^* required to bring the reference state to a state of neutrality (i.e., $b = N_o^2 z + b_e =$

110 0). This can readily be computed from (9) as:

$$E_m^* \approx \frac{1}{2} \rho_o \left(\frac{1}{h} \int_{-h/2}^{h/2} (-N_o^2 z)^2 dz \right) / N_o^2 = \rho_o \frac{N_o^2 h^2}{24} \quad (10)$$



The scatter formula (7) is then identified with the ratio

$$\eta = \frac{E_m}{E_m + E_m^*} \quad (11)$$

once the substitutions $s \rightarrow N_o^2$ and $\sigma_e^2 \rightarrow \frac{1}{h} \int b_e^2 dz$ have been made in going from the discretized forms in Section 2.1.1 to the continuous forms in this section. It is crucial to emphasize that, in making these substitutions, the buoyancy excess $b_e(z)$ is identified with the residuals e_k and the reference state $b_o(z)$ with the linear approximation $q + sz_k$ in (4). As a result, the potential energy of motions with vertical scales much greater than h will not contribute to E_m as their impact on buoyancy will be included in the offset q rather than in the residuals e_k . For the range of values considered in this study ($h = 25m, 100m$) these motions are, in the terminology of Munk (1981), associated with the “gross” structure of the density field and typically include the quasi-geostrophic mesoscale and low modes internal waves.

Equation (11) thus allows to interpret the scatter η as a measure of the “finescale potential energy” present in the layer considered. When the average value of this quantity (i.e., E_m) is small compared to E_m^* then (11) simplifies to $\eta \approx E_m/E_m^* \ll 1$. Only perturbations of the density field with E_m comparable to or larger than E_m^* will lead to large scatter $\eta \geq 1/2$. Given that E_m^* represents a rather extreme situation (completely removing the stratification of the layer), we expect the scatter to be generally small¹. These ideas are explored further in the next section by analysing the scatter found in idealised models of oceanic motions: first considering a single wave in both linear and non linear regimes, and then a field of linear internal waves as described by Garrett and Munk (1972).

2.2 Illustrative models

2.2.1 Monochromatic waves

We consider in this section the scatter resulting from a single sinusoidal perturbation and a buoyancy profile given, at each vertical level by:

$$b_k = N_o^2 z_k + \delta b \cos(\beta z_k + \phi) \quad \text{for } |z_k| \leq h/2 \text{ i.e., } k = 1, \dots, 2K + 1. \quad (12)$$

As in Section 2.1.2 N_o^2 measures the squared Brunt-Vaisala frequency of the undisturbed state of the layer, and δb is the amplitude of the buoyancy perturbation, β is its vertical wavenumber and ϕ its phase. As shown in Appendix B, the scatter obtained by applying (5) to the buoyancy field in (12) is controlled, for a given value of the phase ϕ , by two non dimensional numbers: the steepness of the wave $\sigma \triangleq \beta \delta b / N_o^2$, which measures its non linearity, and the vertical wavelength of the wave, as measured by the parameter $\kappa \triangleq \beta h$.

Figure 1a illustrates the buoyancy profiles and the scatter obtained in this model for the case of a steep wave ($\sigma = 0.9$) and $K = 12$, for a given value of the phase and κ . The blue curve depicts a $\kappa/2\pi \ll 1$ situation and we observe a nearly linear deformation of the buoyancy profile and, as a result, a weak scatter ($\eta = 0.006$). Indeed, in this case the buoyancy perturbation

¹Analysis of hydrographic data in Section 3 indeed suggests that E_m can exceed E_m^* but this is a very rare occurrence, most likely seen in an energetic region where unstable density overturnings are also present.

is nearly entirely captured by the offset q in the linear regression (4). The result is shown here for a value of the phase $\phi = 0$ but very similar plots are obtained for other choices. As the wavenumber increases however, shown here in orange for $\kappa/2\pi = O(1)$ and a value of the phase $\phi = \pi/6$, we observe an “S-shape” deformation of the buoyancy profile and accordingly an increased scatter ($\eta = 0.12$). Increasing the wavenumber further, $\kappa/2\pi \gg 1$ (green curve) means, at fixed σ , a large β and thus a small buoyancy perturbation δb , leading to a negligible scatter ($\eta = 0.001$). As for the $\kappa/2\pi \ll 1$ case, this result is independent of the phase of the perturbation (not shown).

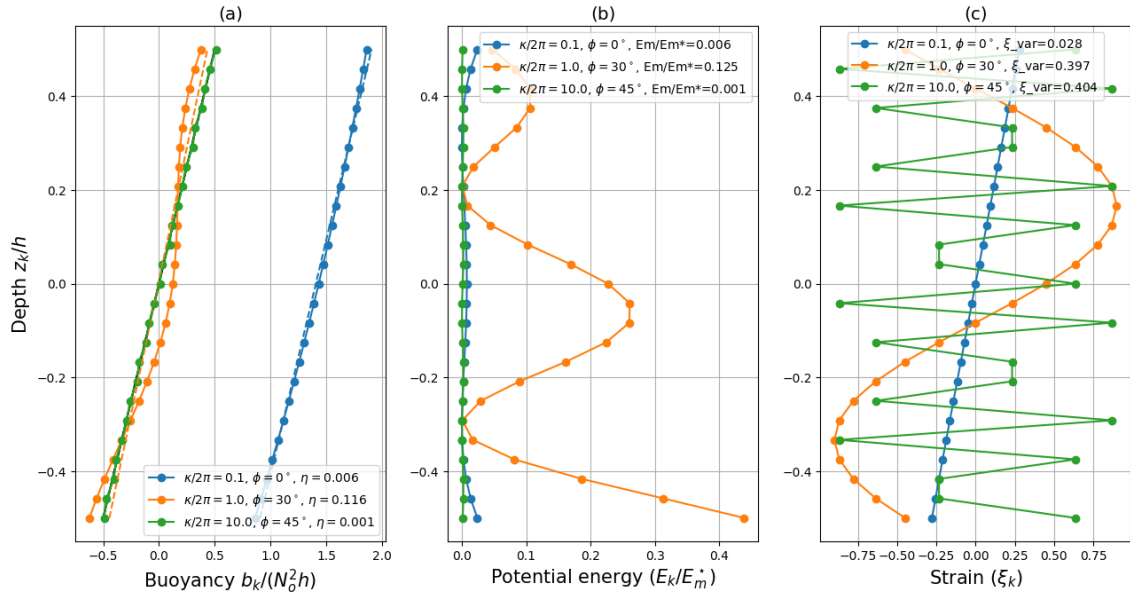


Figure 1. Profiles of (a) buoyancy b_k , (b) “finescale potential energy” E_k and (c) vertical strain ξ_k resulting from the model (12). Here $k = 0, \dots, 2K + 1$ with $K = 12$, the associated depth levels z_k being normalised by h , the buoyancy by $N_o^2 h$ and the potential energy by E_m^* . In (a) the dashed lines indicate the result of the linear regression while the scatter is given in the legend for each case with three decimals precision: (blue) $\kappa/2\pi = 0.1, \phi = 0^\circ$ (orange) $\kappa/2\pi = 1.0, \phi = 30^\circ$ (green) $\kappa/2\pi = 10.0, \phi = 45^\circ$. The mean normalised potential energy (E_m/E_m^*) and strain variance are indicated in the legend in panels (b) and (c), respectively.

The interpretation of the scatter in terms of “finescale potential energy” is confirmed by inspecting the profiles E_k , computed at each level from (A4) and displayed in Fig. 1b (see Appendix A for a discussion of the calculation). As emphasized in Section 2.1.2, the relevant “excess buoyancy” b_e is here the residual in the linear regression, and this can differ significantly from the sinusoidal perturbation imposed in (12). As a result, both the $\kappa/2\pi \ll 1$ and $\kappa/2\pi \gg 1$ profiles (blue and green curves, respectively) have little “finescale potential energy” $E_k \ll E_m^*$, the energy required to remove the layer’s stratification –see Section 2.1.2), although the $\kappa/2\pi \ll 1$ profile (blue curve) has a much larger δb than the $\kappa/2\pi \gg 1$ profile (green curve). The intermediate profile $\kappa/2\pi = O(1)$ however shows values $E_k = O(0.1 E_m^*)$, since there is now a clear distinction between the



linear approximation (dashed orange line) and the full buoyancy profile (continuous orange curve), in agreement with the larger
 155 scatter found in Fig. 1a. Despite the large value of steepness chosen here ($\sigma = 0.9$), the scatter remains nevertheless small, at
 most on the order of $\eta \approx 0.1$. As explored further in Appendix B, it would indeed require including strongly non linear waves
 ($\sigma \approx 2$), and thus mean potential energies E_m approaching E_m^* , to increase the scatter further.

The suggestion from Fig. 1a,b is that intermediate wavenumbers $\kappa/2\pi = O(1)$ lead to greater scatter and “finescale potential
 energy”. Indeed, and as explored further in Appendix B, the scatter can be thought of as a bandpass filter, being most sensitive
 160 to perturbations satisfying approximately $1/2 \leq \kappa/2\pi \leq 2$, or, in terms of the dimensional vertical wavenumber β :

$$\frac{\pi}{h} \leq \beta \leq \frac{4\pi}{h}. \quad (13)$$

This corresponds to vertical wavelengths between half and twice the layer thickness h .

2.2.2 Garrett-Munk internal wavefield

For a more realistic situation, we now study the scatter η produced by a field of random linear internal waves superimposed on
 165 an exponential background stratification:

$$N_o(z) = N_{surf} e^{z/D} \quad \text{where } -H \leq z \leq 0, \quad (14)$$

In this expression N_{surf} measures the stratification near the surface, H is the total ocean depth and D is a vertical scale for the
 thermocline. Within a given layer of thickness h , centred at a depth where the stratification follows (14), the buoyancy field is:

$$b_k = N_o^2 z_k + b'_{k,GM} \quad \text{for } |z_k| \leq h/2 \text{ i.e., } k = 1, \dots, 2K + 1. \quad (15)$$

170 Here $b'_{k,GM}$ is a buoyancy perturbation at level k whose vertical structure and amplitude satisfies the Garrett-Munk (Garrett
 and Munk, 1972 –hereafter GM) spectrum while N_o is understood to be the value of (14) at the centre of the layer. We aim to
 apply (4)-(5) to a large number of synthetic buoyancy profiles given by (15).

To do so, we follow Munk (1981) and the derivation in Gregg and Kunze (1991), who showed that the buoyancy perturbations
 b'_{GM} have a spectral density F_b^{GM} as a function of vertical wavenumber β (keeping with the notation used in Section 2.2.1)
 175 given by

$$F_b^{GM}(\beta) = \frac{N_o^4 E D j_\star}{\beta^2 + \beta_\star^2}. \quad (16)$$

Here the constant E is a non dimensional scale for the energy and $j_\star$ is a mode number setting the dominant (inverse) vertical
 scale $\beta_\star = \pi j_\star N_o / (D N_{surf})$. To generate the synthetic buoyancy profiles (15) satisfying (16) we used a Langevin equation.
 In essence the latter generates a random walk with depth and we can subsample the resulting buoyancy fluctuations within a
 180 set of layers, each of thickness h . The method is briefly discussed in Appendix C.

For illustration, five buoyancy profiles generated for the case $h = 100m$ are shown in Fig. 2a, each profile having been
 normalised by $N_o^2 h$ and centred on a different depth range (color coded). One notices the larger “horizontal swings” of the
 deeper profiles, reflecting that most of the potential energy in the GM wavefield is contained near $\beta = \beta_\star \propto N_o$ (i.e., as one

goes deeper in the water column the energy containing wavelength increases). These swings are picked up by the offset term q in (4) rather than by the residuals e_k because $\beta_*^{-1} \gg h$ for the deeper layers when $h = 100m^2$.

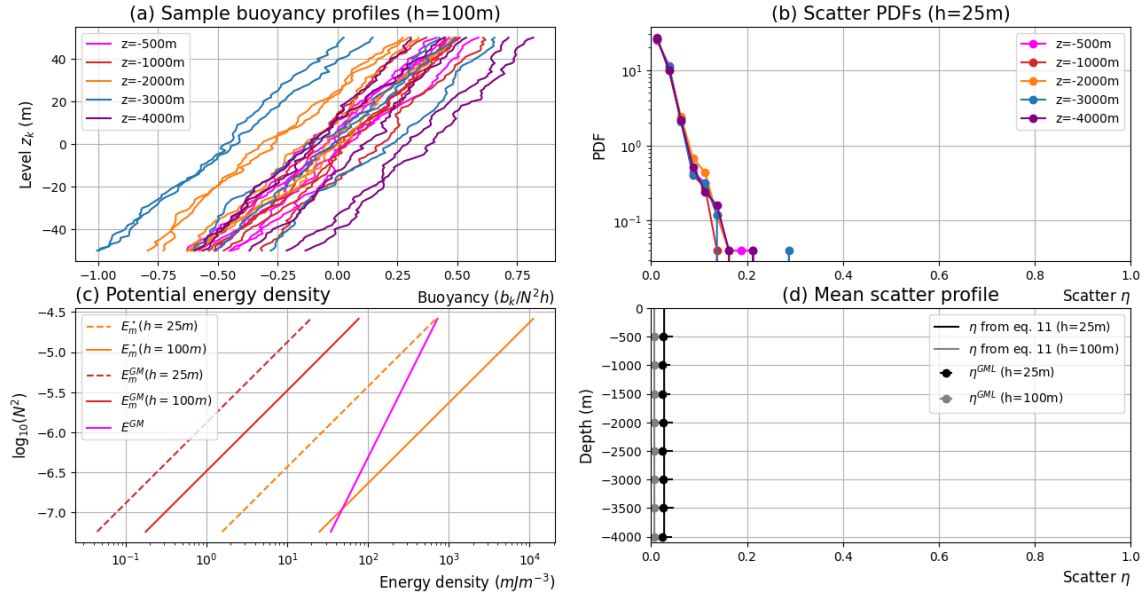


Figure 2. Scatter produced by the synthetic GML internal wavefield. (a) Examples of buoyancy profiles obtained for the case $h = 100m$, color coded by the central depth of the layer and normalised by $N_o^2 h$. (b) PDF of the scatter η obtained at different depths in the case $h = 25m$, using the same color coding as in (a). (c) Dependence of “finescale potential energy” E_m^{GM} (red) and E_m^* (orange), both on the x-axis and expressed in mJm^{-3} , on the stratification (y-axis, $\log_{10} N_o^2$). The case $h = 25m$ ($h = 100m$) is shown in dashed (continuous) line. (d) Vertical profiles of scatter from the numerical GML simulation (dots) and from the prediction (11) for the case $h = 25m$ (black) and $h = 100m$ (gray). For (b) and (d) 1000 realisations were produced at each depth (see Appendix C for details of the numerical calculation).

As expected from the small values of scatter obtained in the monochromatic wave model in Section 2.2.1, as well as the linearity assumed in deriving the GM spectrum, the scatter generated by the buoyancy profiles (15) is small. Its probability density function (PDF), shown in Fig. 2b for the case $h = 25m$, using the same color coding as in Fig. 2a, is very steep, with more than 90 % of the population contained in the interval $0 \leq \eta \leq 0.05$. The particular shape of the distribution is controlled by the Gaussian noise used in generating the Langevin equation, because this noise sets the buoyancy variance in each generated random walk segment of extent h (see Appendix C). It is possible that different stochastic processes, still obeying (16), might produce different shapes of distribution for η but this has not been explored further here. For this reason we will refer to this particular realisation of the GM internal wavefield as “GML” to emphasize that it is based on a Langevin equation. A particularly striking feature of the GML profiles seen in Fig. 2b is the lack of dependence on depth of the distributions.

²All layers satisfy this condition when $h = 25m$ and so the swings are seen at all depths in this case (not shown).



195 To explain this feature, we return to the interpretation of the scatter in terms of “finescale potential energy” (Section 2.1.2). We display in Fig. 2c the reference value E_m^* (orange curves) given by (10) using either $h = 25$ m (dashed line) or $h = 100$ m (continuous line). This potential energy is proportional to N_o^2 and thus varies with a +1 slope in the log-log plot used in this figure. The discussion in sections 2.1.2 and 2.2.1 suggests that the scatter measures how close to E_m^* is the average potential energy E_m of the perturbations with vertical wavelength between approximately $h/2$ and $2h$. We thus consider the potential
200 energy E_m^{GM} defined as:

$$E_m^{GM} \triangleq \frac{1}{2} \rho_o \left(\int_{\pi/h}^{4\pi/h} F_b^{GM} d\beta \right) / N_o^2. \quad (17)$$

This quantity is plotted in Fig. 2c for $h = 25$ m (red dashed line) and $h = 100$ m (red continuous line line). As expected from the redness of the GM buoyancy spectrum, E_m^{GM} is greater for the thicker layer, as is E_m^* because of its h^2 dependence. The “finescale potential energy” E_m^{GM} is seen to increase with stratification but at a slower rate than the “all wavelengths” potential
205 energy E^{GM} defined as:

$$E^{GM} \triangleq \frac{1}{2} \rho_o \left(\int_0^{+\infty} F_b^{GM} d\beta \right) / N_o^2 = \frac{1}{4} \rho_o E N_o N_{surf} D^2 \quad (18)$$

and shown in Fig. 2c by the magenta curve. A comparison of the scatter predicted from (11) with that simulated (mean $\pm$ one standard deviation) is shown in Fig. 2d for $h = 25$ m (black) and $h = 100$ m (gray). As can be seen, the agreement between predicted (continuous curves) and simulated profiles (dots at selected depths) is excellent. This explains the weak scatter
210 ($\eta \approx 0.025$ for $h = 25$ m and $\eta \approx 0.0065$ for $h = 100$ m) produced by the GML wavefield in terms of small ratio E_m^{GM} / E_m^* , and rationalizes its lack of vertical dependence by an approximate compensation between the decay of both E_m^{GM} and E_m^* with depth (the dependence of E_m^{GM} on N_o^2 is seen to be similar in Fig. 2c to that for E_m^* , i.e., a +1 slope in a log-log plot).

2.2.3 Scatter and vertical strain

It is useful to relate the scatter of the stratification η to the more familiar vertical strain field ξ , measuring the stretching
215 (positive strain) or squeezing (negative strain) of isopycnals. While the latter is mostly used in relation to linear motions (e.g., Lien and Muller, 1992), it can also be used in a more non linear regime, using the potential energy definition of Holiday and McIntyre (1980) –see Appendix A. Without going into technical details³ we display in Fig. 1c the vertical profile of strain $\xi_k = \sigma \sin(\beta z_k + \phi)$ corresponding to the buoyancy profiles in Fig. 1a. The “S-shape” of the buoyancy profile in the $\kappa/2\pi = O(1)$ case (orange curves) leads to a large fraction of the layer having large and positive strain $\xi_k \approx 0.8$ leading to a
220 weak stratification near $z_k/h = 0.15$, and a large and negative strain ($\xi_k \approx -0.8$) and enhanced stratification near $z_k/h = -0.3$. The resulting non linear buoyancy profile is picked up by the enhanced scatter in this case. However, as large a strain field is

³We consider a constant $N^2 = N_o^2$ over the layer and use Holiday and McIntyre eq. (2.20) to define the potential energy. It can be shown that, for the monochromatic wave model (12) this leads to a vertical displacement of isopycnals from the undisturbed state $\zeta_k = -\delta b \cos(\beta z_k + \phi) / N_o^2$ at each level k , and a strain field $\xi \equiv d\zeta/dz$ equal at each level to $\xi_k = \sigma \sin(\beta z_k + \phi)$, as in the linear case.



found for the $\kappa/2\pi \gg 1$ case (green curve) while the scatter is correspondingly very small. Indeed the strain variances are both equal to ≈ 0.4 for the green and orange profiles while the scatter are very different (0.12 and 0.001).

To further investigate the relationship between scatter and strain we display in Fig. 3 the strain PDF produced by the GML internal wavefield in the case $h = 25m$ (see section 2.2.2), conditioned on high (blue curve, $\eta > 0.05$) and low (orange curve, $\eta \leq 0.05$) values of the scatter⁴. Note that as was found for the scatter PDF associated with the GML internal wavefield, the strain PDF is depth independent (not shown) and so all depths were used in constructing the distributions (see Appendix C). In agreement with the behaviour in Fig. 1c, we see a significant overlap between the distribution of strain at high and low values of the scatter, indicating that on a case by case basis, strain is not a good predictor of scatter. Conversely, the variance of the distributions are similar (0.029 and 0.03 for high and low scatter, respectively). However, we see that positive strain (lower stratification) is more likely when the scatter is high, leading to a clear shift of the strain distribution (the means of the distribution are $\approx +0.2$ and ≈ 0 for high and low scatter, respectively). Thus the scatter captures a signal which is not contained in the variance of the strain field, but which, in a statistical sense, is captured by the mean value of the strain. We thus expect that the scatter will provide an information on the oceanic buoyancy field quite distinct from that based on finescale parameterisation of turbulent mixing which are based on strain spectra and thus strain variance. We will return to this aspect, in the context of repeated North Atlantic sections, in section 4.3.

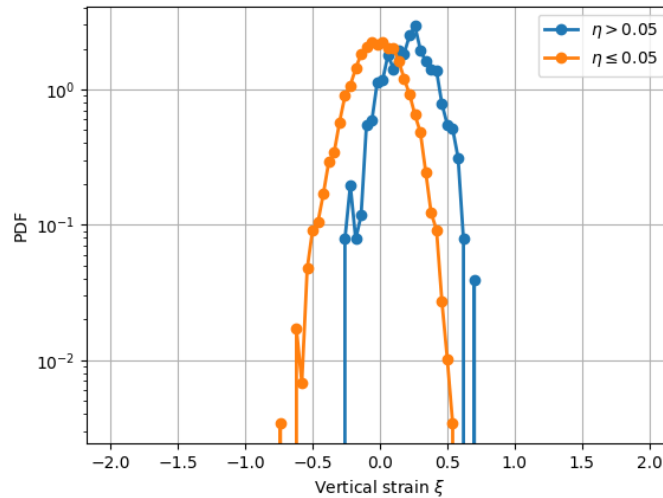


Figure 3. PDF of vertical strain computed from synthetic GML buoyancy profiles (see Appendix A), conditioned on high ($\eta > 0.05$, blue) and low ($\eta \leq 0.05$, orange) values of the scatter. The strain was computed following Kunze (2017), i.e. as $\xi = (N_o^2 - s)/N_o^2$, where N_o^2 refers to the background stratification, and s is the stratification computed from the linear regression (4) applied to the full (background+perturbation) buoyancy field. The sign convention is chosen such that positive (negative) strain corresponds to stretching (squeezing) of isopycnals.

⁴The strain in Fig. 3 is computed for each layer of thickness h , rather than at each level within the layer as was done in Fig. 1c, following Kunze (2017) –see Fig. 3’s caption for details of the calculation.



2.3 Scatter due to instrument errors

Residuals in eq. (4) reflect deviations from linearity caused by actual oceanic motions, such as those illustrated in Section 2.2, and hereafter denoted by d_k , but also caused by instrument errors. These can arise, for example, from random electronic noise affecting the CTD measurement of temperature and salinity, which will in turn affect the calculation of buoyancy field b_k in (2), but also possibly errors in pressure which will affect the calculation of the depths z_k in (3). We denote collectively these errors as ϵ_k and decompose the residual e_k s in (4) as:

$$e_k = d_k + \epsilon_k \quad \text{for } k = 1, \dots, 2K + 1. \quad (19)$$

To estimate the scatter η^{err} arising solely from instrument errors, we consider the opposite extreme to the calculation in section 2.2 which ignored instrument errors entirely. We thus focus on an ideal situation where there is a stratification undisturbed by finescale oceanic motions (denote this by $N = N_o$), and we write eq. (4) as:

$$b_k = N_o^2 z_k + \epsilon_k, \quad \text{for } k = 1, \dots, 2K + 1. \quad (20)$$

Application of (7) then produces:

$$\eta^{err} = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + N_o^4 h^2 / 12}. \quad (21)$$

where σ_ϵ^2 is the variance of the instrument errors ϵ_k s.

To estimate σ_ϵ^2 , we follow Gargett and Garner (2008) and assume that the variability of the buoyancy field in regions of low stratification reflects instrument errors and can thus be used to estimate them. Unlike Garrett and Garnett however, we do not specify a threshold to isolate regions of low stratification, but instead use the layers where the linear regression (4) is not statistically significant at the 95 % level in a two-sided t-test. We'll refer to these layers in the following as “no- N^2 ” layers and denote their number, in a given section, by n_s . For a given hydrographic section, the corresponding set of layers defines a population of buoyancy errors ϵ_k . The variance σ_ϵ^2 can then be computed as the median⁵ of all the sample variances $\tilde{\sigma}_{\epsilon,n}^2$ given by:

$$\tilde{\sigma}_{\epsilon,n}^2 = \frac{1}{2K + 1} \sum (b_k - b_m)^2 \quad (22)$$

where $n = 1, \dots, n_s$.

Besides these “no- N^2 ” layers, we also keep track of layers where the slope of the linear regression is statistically significant at the 95 % level in a two-sided t-test, but negative. We refer to these layers as “overturning layers” and denote their number, for each section, by n_o . With the physical interpretation of the scatter in Section 2.1.2 in mind, η is only computed over the set of layers excluding the “no- N^2 ” layers and the “overturning” layers, such that the stratification estimated for these layers make a physically plausible reference state.

⁵The median was found to be more reliable than the mean for this population of buoyancy variances. Indeed the probability distribution functions of instrument errors have long tails and are non normal, as illustrated in Section 3.



265 Finally, to be able to use (21) in observations, we also need an estimate of N_o^2 . The latter is assumed to be reflected in the coarser or, following the terminology in Munk (1981), the “gross” stratification found in the larger scale buoyancy field. Using this assumption, we set N_o^2 to be the slope of the linear regression of the b_k s v the z_k s over a thicker layer than that over which the scatter is computed, but centred on the same pressure level p_o . To do so we introduce a parameter n_l , counting the number of “extra” layers above and below the layer over which the scatter is computed (i.e., we use $(2K + 1)(1 + 2n_l)$ pressure levels
 270 for this calculation). For the observations discussed in Section 3 for example, we considered $n_l = 1$ for the $h = 100m$ case (N_o^2 being then estimated over a $\sim 300m$ layer) and, to use a comparable vertical scale, $n_l = 5$ for the case $h = 25m$ (N_o^2 being then estimated over a $275m$ layer).

2.4 “Motion” and “Error” dominated regimes

We now bring together the physical and statistical aspects of the scatter. We distinguish two regimes:

- 275 – a “motion dominated” regime where the measured scatter $\eta \gg \eta^{err}$. This is the situation where we can use the scatter as a tool to investigate the finescale structure. This will be achieved if the associated motion is vigorous enough to lead to large η or, if of smaller amplitude, it occurs in regions of high stratification where (21) predicts that η^{err} will be smaller.
- an “error dominated” regime where the measured scatter is dominated by that expected from instrument errors, $\eta \approx \eta^{err}$. In this regime, physical effects cannot be detected. This is typically expected to occur in abyssal parts of the ocean where
 280 the stratification is weak and where (21) predicts that η^{err} could be large if instrumental buoyancy errors are comparable to $N_o^2 h / \sqrt{12}$.

These two extremes can be distinguished by considering the ratio Γ :

$$\Gamma \triangleq \frac{\eta}{\eta^{err}} \quad (23)$$

With this definition, the motion-dominated limit corresponds to the situation $\Gamma \gg 1$ while the error-dominated regime to $\Gamma \approx 1$.
 285 The limit where $\Gamma \ll 1$ occurs when the measured scatter is much less than that expected from instrument errors. A large occurrence of such situations would indicate a failure of the method since it would imply negative scatter due to physical motion. As such it is interesting to keep track of all values of the parameter Γ to test the method (as we’ll see in Section 3 the occurrence of $\Gamma \ll 1$ is negligible).

Figure 4 summarises the steps involved in the analysis, highlighting the different types of layers encountered: “no- N^2 ” layers
 290 for which no statistically significant stratification can be detected and which form the basis of the instrument error calculation, “overturning” layers for which the linear regression slope s in (4) is statistically significant but negative, and “analysis” layers for which the parameter Γ can be estimated and from which the separation between “error dominated” and “motion dominated” regimes can be made. This figure also summarises the calculation of the finescale potential energies E_m and E_m^* (see Appendix A), as well as the calculation of the finescale potential energy E_m^{GM} defined in Section 2.2.2 and computed using the “gross”
 295 stratification N_o^2 based on a layer of thickness $(2n_l + 1)h$ –see Section 2.3.

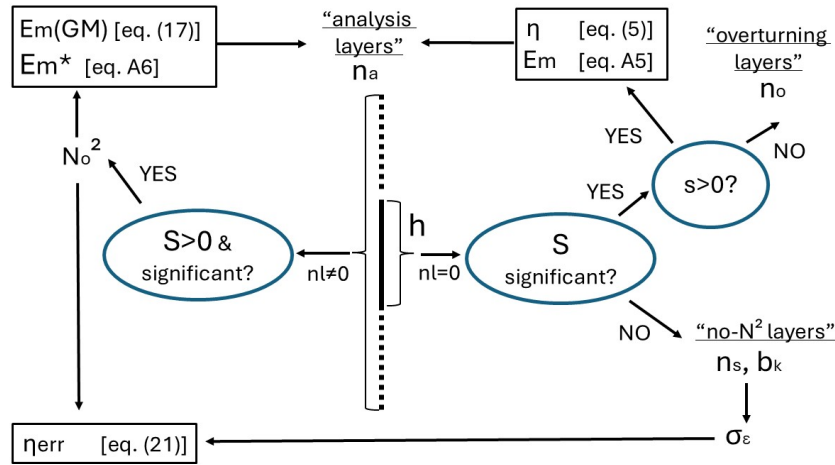


Figure 4. Flowchart depicting the calculation of the main variables in this study. Starting from the centre of the figure, a choice of layer thickness h (thick black line) is made. The calculation of the measured scatter η and “finescale” potential energy E_m , as well as the identification of the “no- N^2 ” and “overturning” layers proceed from the right through consideration of the slope s obtained in the linear regression (4) applied to that layer ($n_l = 0$). The calculation of the reference potential energies E_m^* and E_m^{GM} , as well the scatter η^{err} expected from instrument errors proceed from the left by applying (4) to a thicker layer of width $h(1 + 2n_l)$ (dotted black lines). The requirement that both linear regression slopes s be statistically significant and positive define the “analysis layers” for which the parameter Γ can be computed.

2.5 Other diagnostics

To help identifying processes responsible for low or high scatter we also calculate the Turner angle T_u (Ruddick, 1983). The latter requires knowledge of a temperature slope with depth ($d\Theta/dz$) and a salinity slope with depth (dS/dz), both of which can be estimated by linear regression of the Θ_k s and S_k s v the depths z_k s. We distinguish between a “local” Turner angle, estimated over the same layer as the one for which the scatter is calculated, and a “gross” Turner angle estimated over the same (thicker) layer used for the estimation of N_o^2 in Section 2.3.

3 Application to OVIDE sections (2002-2023)

3.1 Data and oceanographic context

We apply the scatter calculation to a set of hydrographic sections, hereafter OVIDE sections (also known as WOCE A25), between Greenland and Portugal, repeated nearly every two years between 2002 and 2023 –See Table 1. Seasonal dependence is minimised because all sections were taken during Summer months (May, June and July). The sections cross the Irminger Sea to the Northwest, the Reykjanes Ridge near $30^\circ W$, the Iceland basin, the West European basin and Iberian Abyssal Plain



further east, with an approximately equal spacing of 25 nautical miles ($\approx 46km$) between casts (see insets in Fig. 5). In terms of circulation, they sample parts of the subpolar gyre and the northeastward extension of the subtropical gyre, and contain both the East Greenland Current in the west (hereafter EGIC to include contributions from the Irminger current –see Pickart et al., 2005; Danialt et al., 2011), and multiple branches of the North Atlantic Current (NAC) centred around $20^\circ - 25^\circ W$. The NAC is the main contributor to the upper limb poleward transport of light waters associated with the Atlantic Meridional Overturning Circulation (AMOC) across the OVIDE sections. The mean AMOC estimated by inverse method in density class is $17.4 \pm 0.6 Sv$ with small yr-to-yr variations of $2.1 Sv$ over the 1997-2021 period (Mercier et al., 2024). Important watermasses are present in all these sections, including Labrador Sea Water (LSW), Mediterranean water (MW), Icelandic Slope Water (ISW), Denmark Strait Overflow Water (DSOW) and Iceland-Scotland Overflow Water (ISOW) –see for example Danialt et al., 2016.

CTD measurements were performed using a Neil Brown Mark III probe from 2002 to 2006 and with a Seabird 911 Plus probe for subsequent expeditions. All temperature (T) and practical salinity (S_P) profiles were obtained with a resolution of $1db$ after quality control (Branellec et al., 2019). In-situ temperature (T) and practical salinity (S_P) profiles were then converted to conservative temperature and absolute salinity profiles using the GSW package (McDougall and Barker, 2020).

Using GOSHIP precision standards of $\epsilon_T = 0.5mK$ for temperature and $\epsilon_S = 1mg/kg$ for absolute salinity (Katsumata, 2022), we define a scale for the instrument error on buoyancy via:

$$\sigma_\epsilon^{ref} = \frac{g}{\rho^{ref}} \sqrt{(\alpha^{ref} \epsilon_T)^2 + (\beta^{ref} \epsilon_S)^2} \quad (24)$$

where ρ^{ref} is the density estimated at $p^{ref} = 2000db$ and at the global mean absolute salinity and conservative temperature, $S^{ref} = 34.89g/kg$ and $\Theta^{ref} = 3.9^\circ C$, respectively (see Roquet et al., 2015). Likewise for the thermal and haline expansion coefficients $\alpha^{ref} = \alpha\{S^{ref}, \Theta^{ref}, p^{ref}\}$ and $\beta^{ref} = \beta\{S^{ref}, \Theta^{ref}, p^{ref}\}$, leading to a value $\sigma_\epsilon^{ref} = 7.36 \times 10^{-6} ms^{-2}$. Instrument errors will be expressed as a fraction of this value in the following.

3.2 Instrument errors

We illustrate the error calculation for a choice of layer thickness $h = 25m$ in Fig. 5a. This panel shows the number of these $25m$ layers which are identified as “no- N^2 layers” (Section 2.3), using all sections available and displaying the results in bins of $100m$ -depth by 1° -longitude. It is seen that most of these layers are found in the eastern portion of the section, over the seafloor of the Iberian Abyssal Plain. Other regions where such layers can be detected are over the Reykjanes Ridge near $30^\circ W$ and within the EGIC near $42^\circ W$. The number n_s of these layers for each section is given in Table 1 (third column).

Separate analysis of the early (2002-2006) and late (2008-2023) sections indicate that, despite having a very similar spatial distribution of “no- N^2 layers” (i.e., as in Fig. 5a, not shown), the two sets display significant differences in the statistical distribution of the errors. This is shown in the associated PDF in Fig. 6, with typically smaller values of the b_k (see Section 2.1.1) found in the “no- N^2 layers” in the later sections. Using the methodology described in Section 2.3 this leads to an estimate of the standard deviation of instrument errors on buoyancy σ_ϵ , significantly smaller in the later sections (Table 1, fifth column). Interestingly, as discussed in section 3.1, a change in the CTD instrument was made in 2007 so it is encouraging that

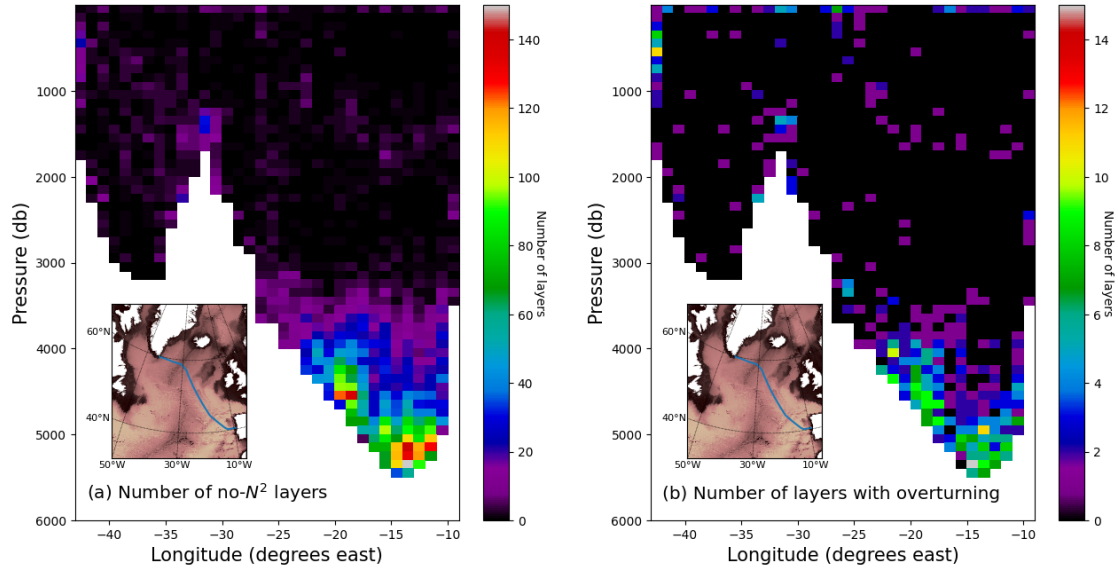


Figure 5. Number of 25m layers counted as (a) “no- N^2 layers” and (b) “overturning layers” using all available OVIDE sections (2002-2023). The count was performed over bins of size 1° longitude - 100m depth. Note the change of colorbar between the two panels. An inset shows the location of OVIDE tracks (blue) with bathymetry in pink shading.

our error calculation can detect this. An investigation of threshold errors on density by Kokoszka (2012) aimed at computing Thorpe scales in OVIDE sections led to a similar conclusion. Kokoszka estimated a density threshold of $17 \times 10^{-4} \text{kgm}^{-3}$ and $6 \times 10^{-4} \text{kgm}^{-3}$ for the early (2002-2006, Neil Brown CTD) and late (2008-2010, Seabird CTD) sections, respectively. We have indicated the corresponding range of values in Fig. 6 by thick horizontal lines, and one can see that the change in
 345 range between the two instruments match the change in the PDF. Actual calculation shows that the fraction of values with $|b_k|$ smaller than the threshold is indeed $\sim 98\%$ for both PDFs, corresponding to a threshold between 2σ (95%) and 3σ (99%) in terms of Gaussian statistics⁶.

A test of the accuracy of the method is provided by comparing the calculation of σ_ϵ for different values of h . Indeed, if the errors truly reflect instrument errors, their value should be independent of the arbitrary choice of this parameter. In doing
 350 so it must be emphasized that the “no- N^2 ” layers might be different in these two calculations since they are based on the statistical significance of a linear regression applied to different sets of b_k s and z_k s in (4). Indeed, there is typically less “no- N^2 ” layers with $h = 100\text{m}$ than $h = 25\text{m}$ (Table 1, third column), consistent with the expectation that a stratification is more

⁶We only mention an equivalent Gaussian value for reference as the distributions in Fig. 6, with their extended tails, are clearly not Gaussian. We haven’t investigated this further but it seems consistent with the presence of large sporadic fluctuations, perhaps as a result of debris being trapped in the CTD instrument on occasion, or intermittent problems with the transmission cable.

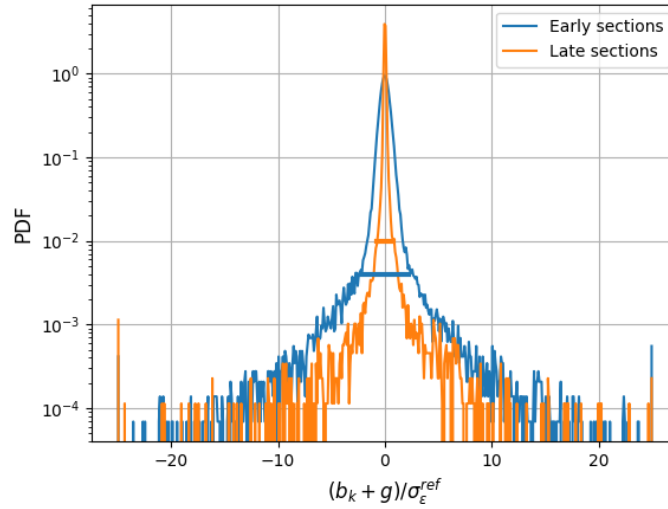


Figure 6. PDF of the buoyancy b_k found in the “no- N^2 layers” for the early (2002-2006, blue) and late OVIDE sections (2008-2023, orange). The values of buoyancy on the x-axis have been normalised by σ_ϵ^{ref} (see Section 2.3) after adding the offset g . The density thresholds estimated by Kokoszka (2012) for the Neil Brown CTD (early sections) and Seabird CTD (late sections) are shown as thick horizontal lines (the density was converted to buoyancy after multiplication by $\rho/g = 1030 \text{ kg m}^{-3}/9.81 \text{ m s}^{-2}$). The PDF was computed by splitting the $[-25, 25]$ interval into 500 equally spaced bins. Positive (negative) values outside this range have been added to the largest (lowest) bin, which explains the “spikes” seen at the extremities of the distributions.

readily detected in a thicker layer. The fifth column of this Table compares the values of σ_ϵ found in the two calculations for all sections (values found for $h = 100 \text{ m}$ given in parenthesis). As can be seen, a very good agreement is found, again giving
 355 confidence in the method.

Finally, we display in Fig. 5b the number of 25 m layers counted as “overturning layers” (Section 2.3). Their distribution largely mirrors that of the “no- N^2 layers” in Fig. 5a, with most of these layers found again over the Iberian Abyssal Plain, but with a much smaller count overall (their number for each section is given in Table 1, fourth column). One also notices the increased occurrence of overturning layers near the sea surface and within the EGIC.

360 3.3 Scatter maps

3.3.1 General features

To illustrate the calculation of the scatter, we show in Fig. 7 the resulting values obtained with $h = 25 \text{ m}$ for an earlier (2006) and a later (2023) section. Starting with the 2006 section, in panel (a) we show a map of η calculated with $n_l = 0$ as a function of the casts longitude and pressures while the scatter expected from instrument errors η^{err} is shown in panel (b), calculated



with $n_l = 5$ (see Fig. 4 for a summary of the methodology). Perhaps the most striking feature is the similarity of panels (a) and (b), suggesting that instrument errors dominate the residuals in (4) for that particular year.

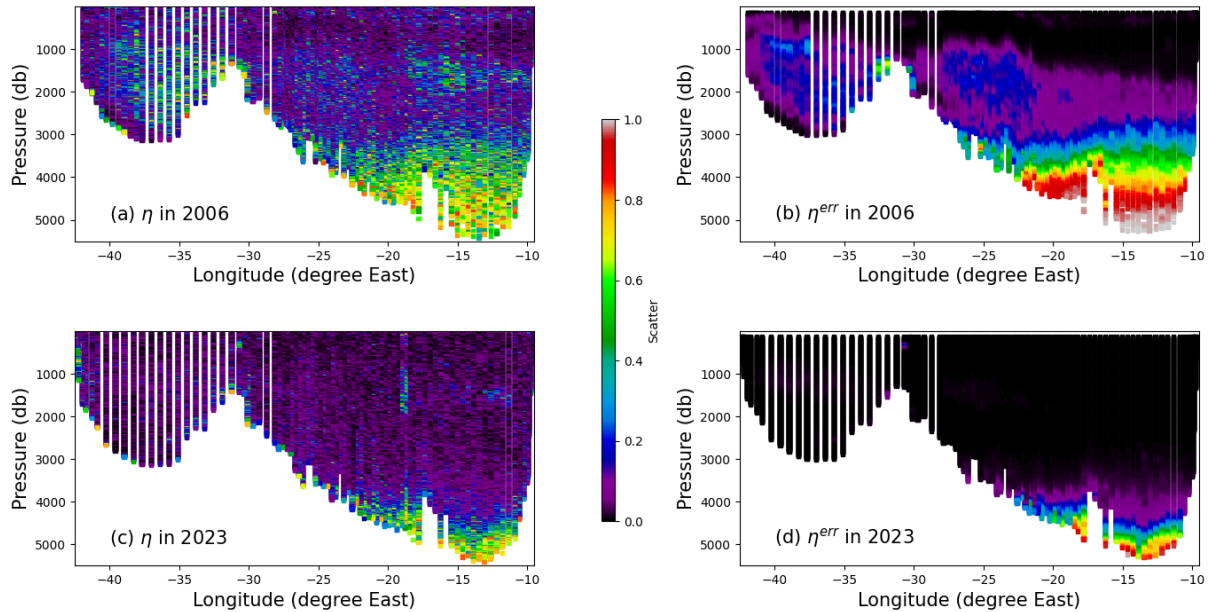


Figure 7. Scatter calculation for the 2006 (top) and 2023 (bottom) OVIDE sections with layer thickness $h = 25m$. The x-axis (y-axis) indicates longitude (pressure) along the sections. Panels (a) and (c) show the scatter η from eq. (5) while panels (b) and (d) show the scatter due to instrument errors η^{err} from eq. (21). Each cast is plotted independently from the sea surface at $p = 0db$ to the seafloor (white domain) but since the calculation of η^{err} used $n_l = 5$, there is a loss of data at the top and bottom of the section in panels (b) and (d) compared to (a) and (c). The vertical white stripes are proportional to the longitudinal distance between adjacent casts (shorter when the section is inclined with respect to a latitude circle).

The situation in 2006 is typical of the sections in 2002 and 2004 (not shown). It is however quite different from the subsequent sections where, as noted earlier, instrument errors are estimated to be smaller (Table 1, 4th column). An example is shown for the 2023 section in Fig. 7c,d which can be directly compared with Fig. 7a,b. We recover the general similarity between η and η^{err} near the sea floor, especially in the Iberian Abyssal Plain. One is struck however by the low values of η^{err} elsewhere (“dark regions” in panel d), and the presence of localised regions with values of scatter η much larger than this background (e.g., the “bright” band near $18^\circ W$ in panel c).

Another way to display the distribution of scatter is provided in Fig. 8, using ‘height above bottom’ as vertical coordinate rather than pressure. The Iberian Abyssal Plain (roughly east of $20^\circ W$) again stands out as a region of high scatter for both η and η^{err} but, interestingly, it is also seen that both variables are not systematically larger close to the seafloor –see for example the low values found in the Irminger Basin (roughly $40^\circ W - 30^\circ W$). Supplementary Figs. S1 and S2 display the corresponding

figures using a thicker layer ($h = 100m, n_l = 1$). Besides the features discussed for Figs. 7, 8 one notices the larger scatter found near the sea surface at this larger vertical scale, reflecting the non linear buoyancy profile created by the summertime stratification. The comparison of the scatter field obtained with $h = 25m$ and $h = 100m$ provides useful information and this will be highlighted in Sections 3.3.2 and 3.3.3.

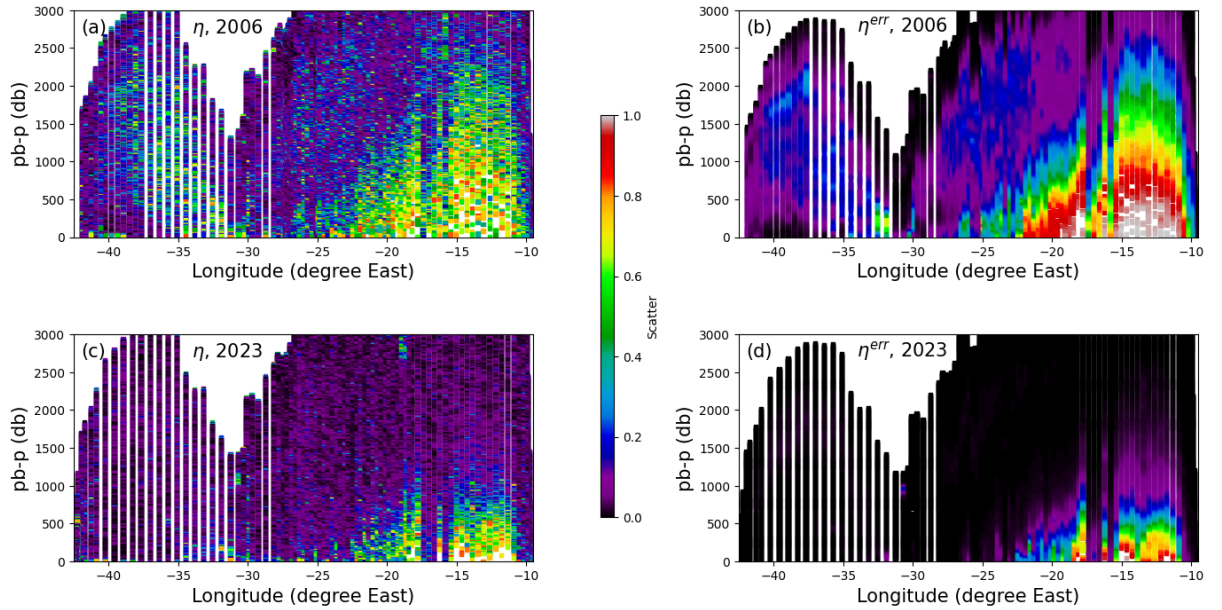


Figure 8. Same as Fig. 7 but plotted as a function of height above seafloor instead of pressure. The latter coordinate was obtained by plotting $p_b - p$ (where p_b is the deepest pressure level for that cast) as vertical coordinate.

As discussed in Section 2.4, one can use η and η^{err} to separate “error dominated” and “motion dominated” regimes by considering the ratio $\Gamma = \eta/\eta^{err}$. The latter was computed using ($h = 25m, n_l = 5$) for the recent OVIDE sections (2008-2023) and we display in Fig. 9a the resulting occurrence of layers satisfying $\Gamma > 3$, which we associate with the “motion dominated” regime, and that of layers with $1/3 < \Gamma \leq 3$, which we associate with the “error-dominated” regime in Fig. 9b. The number of such layers was computed in bins of 100db-pressure and 1°-degree of longitude, as in Fig. 5, and the result normalised by the total number of “analysis layers” in each bin. We recover that the Iberian Abyssal plain is a region “opaque” to the scatter analysis, with more than 80% of its layers in the “error dominated” regime (Fig. 9b), as is the core of the LSW water in the Irminger Sea (about 50% of the layers in this category). The eastern boundary, the wedge made up by waters of Atlantic origin, the westernmost portion of the sections and the abyssal Irminger Sea show up as the region with the largest occurrence of layers in the “motion dominated” regime (typically more than 90% of the layers), and thus as the most favourable to the scatter analysis. For completeness, we display in Fig. 9c the occurrence of layers with $\Gamma \leq 1/3$, which we associate with layers where the methodology fails (Section 2.4). As can be seen their occurrence is negligible and restricted to the deepest parts of the

Iberian Abyssal Plain. Supplementary Fig. S3 show the corresponding figure for the earlier OVIDE sections (2002-2006). As expected from the discussion of instrumental errors in section 3.2 the occurrence of layers in the “motion dominated” regime shrinks considerably and becomes restricted to the wedge made by incoming Atlantic waters. The bulk of the early sections is thus found in the “error dominated” regime, and the occurrence of layers where the method fails also increases significantly.

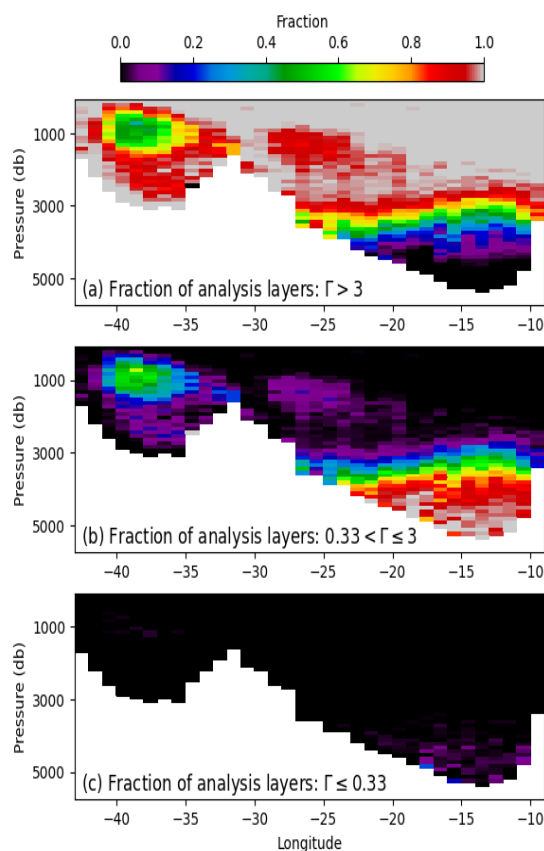


Figure 9. Partitioning of “analysis” layers into regimes based on the value of the parameter Γ . The colors indicate the fraction of such layers satisfying (a) $\Gamma > 3$ (“motion dominated” regime) (b) $1/3 < \Gamma \leq 3$ (“error dominated” regime) (c) $\Gamma \leq 1/3$ (“method failure” regime). Calculations based on late OVIDE sections (2008-2023) using ($h = 25m, n_l = 5$).

Figures 7a,c and 8a,c as well as those of other OVIDE sections not shown here, suggest that the scatter in the “motion dominated” regime is composed of a background of low to moderate values (black to dark blue color range, $\eta = 0 - 0.2$) and localised regions with much higher values ($\eta \geq 0.2$, light blue, green and warmer colors). We now discuss these in turn.



3.3.2 Low-to-moderate scatter regime

The distribution of “analysis” layers satisfying $\Gamma > 3$ and $\eta < 0.2$, i.e. low-to-moderate scatter events in the “motion-dominated regime”, is shown in Fig. 10a for a calculation using ($h = 25m, n_l = 5$) applied to the later OVIDE sections (2008-2023) –the analog for the early sections is given in Supplementary Fig. S4a. One is struck by the similarity of Fig. 9a and Fig. 10a, both in terms of spatial distribution and frequency, which indicates that the bulk of the layers with $\Gamma > 3$ in the OVIDE sections
 405 experience low-to-moderate scatter. This will be confirmed by looking at the PDF of scatter in Section 3.5.

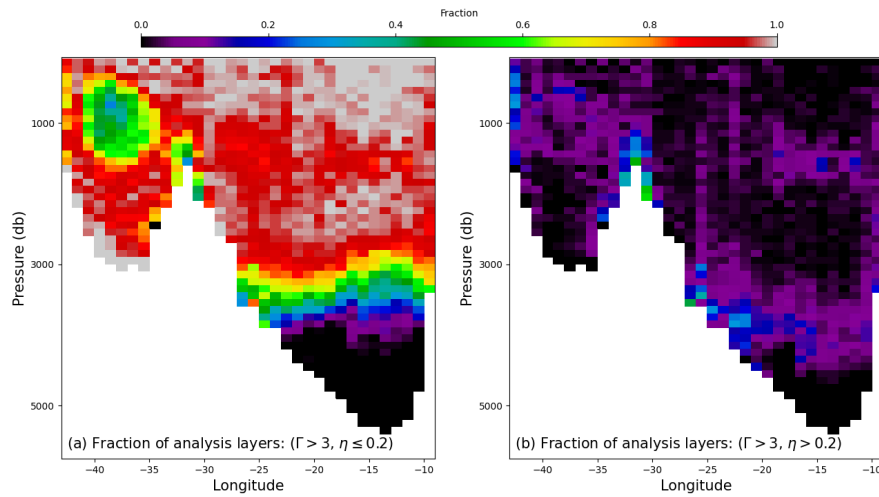


Figure 10. Same as Fig. 9 but for the occurrence of “analysis” layers satisfying $\Gamma > 3$ and (a) $\eta \leq 0.2$ (low-to-moderate scatter events) (b) $\eta > 0.2$ (high scatter events). Calculations based on late OVIDE sections (2008-2023) using ($h = 25m, n_l = 5$).

An example of cast in this regime is shown in Fig. 11 for the Irminger Sea at $36.4^\circ W - 59.36^\circ N$ (2023 section). The Θ and S profiles from the raw ($1db$ vertical resolution) data are plotted in blue in panels (a) and (b), respectively, while the scatter η calculated with ($h = 25m, n_l = 5$) is superimposed in orange, scaled appropriately to emphasize the links with features seen in Θ and S . Despite the large swings seen in the salinity profile, the vertical profile of scatter is in comparison fairly uniform
 410 with depth, with a mean value on the order of $\eta = 0.05$. This is consistent with the idea that this region hosts thermohaline intrusions efficient at creating vertical perturbations in the salinity profiles (e.g., Fernandez-Castro et al., 2024 –see their Fig. 7c,d) but that these are largely compensated in temperature and are thus less effective at creating vertical perturbations in the buoyancy profiles. The scatter η^{err} expected from instrument error for that section is shown in red in (a) and (b), and is small enough that most of the water column here is classified as in the “motion-dominated” regime. This is highlighted by the green
 415 dots in panel (c) which show where the profile of Γ satisfies $\Gamma > 3$.

Figure 11c displays the potential energies E_m (blue), E_m^* (orange) and E_m^{GM} (red). The low values of scatter seen in orange in panels (a) and (b) are consistent with the finescale potential energy E_m of the layers being generally much smaller than that required to remove the stratification of the layer (E_m^*). In addition, the depth-independent background of low values displayed

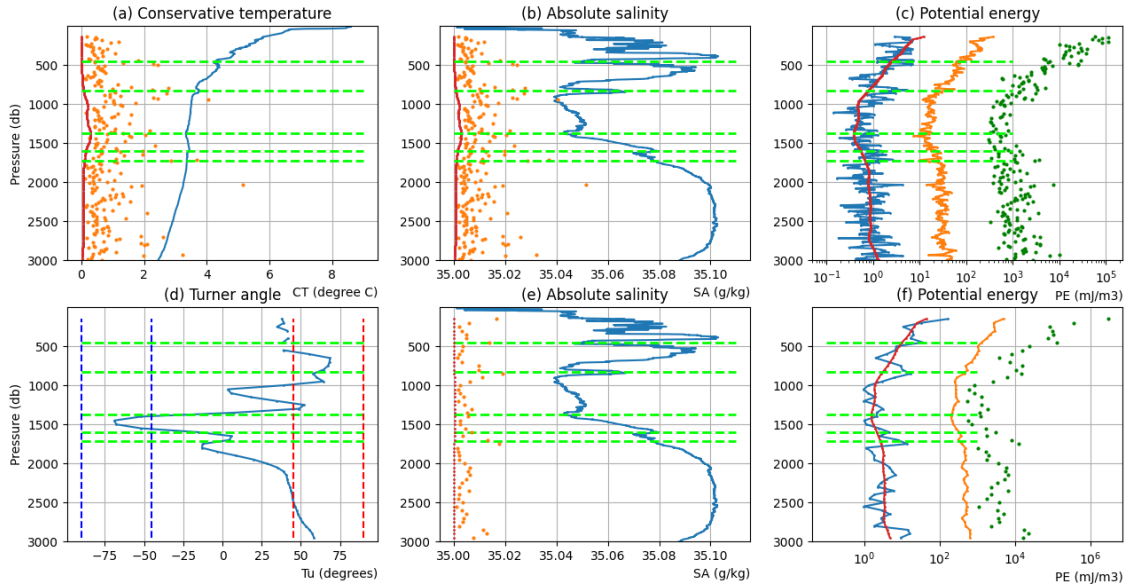


Figure 11. Vertical profiles at $36.4^\circ W - 59.36^\circ N$ in 2023 (cast 76), with the depth of selected thermohaline intrusions highlighted by lime dashed horizontal lines (same for all panels). (a) Conservative temperature Θ (blue, in $^\circ C$) and (b) Absolute salinity S (blue, in g/kg) from the raw data at $1db$ resolution, and (c) Potential energies E_m (blue), E_m^* (orange) and E_m^{GM} (red) estimated for $h = 25m, n_l = 5$ (all expressed in mJm^{-3}). The profile of scatter η and that expected from instrument error for that section, η^{err} , are plotted in orange and red, respectively in panels (a) and (b). Note that for visual display $20(\eta, \eta^{err})$ is plotted in (a) and $35 + 0.2(\eta, \eta^{err})$ is plotted in (b). Panel (c) also shows in green the values 100Γ but only displaying layers where $\Gamma > 3$. (d) “Gross” Turner angle T_u (in degrees). The “diffusive” and “salt finger” regimes of double diffusive convection have been highlighted by the vertical dashed lines at $T_u = -90^\circ, -45^\circ, +45^\circ$ and $+90^\circ$ (blue and red intervals, respectively). (e)-(f) As in (b)-(c) for a calculation using ($h = 100m, n_l = 1$). Note that in (e) (η, η^{err}) was rescaled by 0.5 instead of 0.2.

by E_m is suggestive of that expected from the GML internal wavefield (Section 2.2.2). Indeed, E_m is typically on the order of
 420 a few mJm^{-3} and follows the general expectation E_m^{GM} from the GM spectrum discussed in Section 2.2.2. Thus we identify
 the background of low scatter in Fig. 10a as that resulting from the GML internal wavefield.

Closer inspection of the profiles shows, in addition, the presence of localised increase in scatter ($\eta \sim 0.1 - 0.2$) associated
 with thermohaline intrusions. These are highlighted by dashed lime lines in Fig. 11 for layers in the “physically dominated
 regime”, and they are seen to correspond approximately to regions where $E_m \geq E_m^{GM}$ (note that the salinity profiles are
 425 plotted on the raw $1db$ resolution data while the calculation of scatter and potential energies is carried out at the approximately
 $25db$ -spaced layer centers). The fact that the scatter is enhanced in these regions indicates that the intrusions are not perfectly
 compensated in density, and this imperfect compensation enhances the scatter above that produced by the GML wavefield.



This effect is large enough to double the vertical average of η compared to the GML averaged value of $\eta = 0.025$ found in Fig. 2d.

430 A repeat of this analysis for a choice of layer thickness $h = 100\text{m}$ leads to very similar conclusions (Fig. 11e,f). In this figure the salinity profile in panel (e) is the same as that in panel (b) since they are based on the raw 1db profiles, but the scatter (orange) is based on $h = 100\text{m}$ and has been rescaled for clarity. To help with the comparison between the $h = 25\text{m}$ and $h = 100\text{m}$ cases, we have kept the same horizontal dashed lime lines between the two rows. It can be seen that the $h = 100\text{m}$ calculation also displays a background of low scatter values, with the same enhanced scatter regions as those found in the $h = 25\text{m}$ case –compare the location of the orange dots in panels (b) and (e). The potential energies shown in panel (f) indicate again the relevance of the background E_m^{GM} in most layers, with the enhanced scatter regions found where $E_m > E_m^{GM}$ as was found in panel (c) for the ($h = 25\text{m}, n_l = 5$) calculation. In agreement with Section 2.2.2, the values of E_m^{GM} are greater in the $h = 100\text{m}$ case but since the values of E_m^* increase even further with a thicker layer, the background scatter is lower than in the $h = 25\text{m}$ case. It must be emphasized that the red curves (E_m^{GM}) in Fig. 11c,f are not low passed version of the blue curves (E_m) but rather the prediction from the analytical formula (17) whose sole input from the data is the “gross stratification” N_o^2 . The general agreement is truly remarkable and underscores the relevance of the background GML internal wavefield in this region.

The role of double diffusive convection in being possibly responsible for the thermohaline intrusions is investigated in Fig. 11d, by plotting the “gross” Turner angle for this cast. At the levels where the scatter is enhanced the Turner angle is not in a regime most favourable to double diffusive processes ($75^\circ \leq |T_u| \leq 90^\circ$, highlighted by the vertical dashed lines), with values of 40° , 60° , -24° , -16° and 0° at, respectively, depths of $450, 830, 1370, 1600$ and 1720db . This is consistent with the large vertical scale ($\approx 50 - 200\text{m}$) of the disturbances required to enhance the scatter for the $h = 100\text{m}$ calculation which contrasts with the shorter vertical scale (a few tens of meters) expected from double diffusive convection (e.g., Radko, 2013).

3.3.3 High scatter regime

450 The distribution of “analysis” layers satisfying $\Gamma > 3$ and $\eta > 0.2$, i.e. high scatter events in the “motion-dominated regime”, is shown in Fig. 10b for a calculation using ($h = 25\text{m}, n_l = 5$) applied to the later OVIDE sections (2008-2023). The occurrence of such events has a spatial distribution quite distinct from that seen in Fig. 9a and 10a, with local maxima above 20% at the Greenland continental shelf near 42°W (top-to-bottom), over the Reykjanes Ridge near 30°W , over the Eriador Sea Mount near 25°W and over the East Thulean Rise between $20^\circ - 25^\circ\text{W}$. These features also dominate the distribution of high scatter events in the “motion dominated” regime in a calculation with ($h = 100\text{m}, n_l = 1$) –not shown–, and we illustrate some of the corresponding profiles below and in the Supplementary Material. Before doing so, we note that the occurrences in Fig. 10b are also enhanced in the Irminger Sea and the region below the Mediterranean Water layer ($\approx 10^\circ\text{W} - 20^\circ\text{W}$, $1000 - 2000\text{db}$), the latter being a prominent feature in the earlier sections (see Supplementary Fig. S4b). Inspection of the Θ , S_A and buoyancy profiles in this region suggest that this signal is not of climatic origin but reflects salinity spikes present in the early sections



below the Mediterranean Water core (not shown)⁷. We also note in Fig. 10b the extension to more eastward longitudes and greater depths of the large occurrences mentioned above over the East Thulean Rise. This feature will be discussed more fully in section 3.4.

An example of cast within the Greenland continental shelf at $42.23^\circ W - 59.8^\circ N$ is shown for the 2016 section in Fig. 12 (top row), in a format similar to Fig. 11a,b,c. The salinity profile in (b) indicates the presence of fresh surface waters in the top 200db layer which control the stratification there, an intrusion of salty and warm water near 350db, and a more uniform profile below 400db. Unlike the situation in Fig. 11, potential energies E_m (blue) in (c) systematically exceed E_m^{GM} (red curve) with typical values $O(10mJm^{-3})$. This is not entirely unexpected since the GM model is not meant to apply near boundaries, nor near regions of large velocity shear, which can be expected here because of the presence of the EGIC. Consequently, the scatter is large throughout the column (average of 0.24 here and 0.11 for the ($h = 100m, n_l = 1$) calculation), suggesting sources of potential energy over a broad range of vertical wavelengths (12.5 – 200m) in this energetic western boundary current regime. We also note the larger occurrence of “overturning layers” along the continental shelf in Fig. 5b, consistent with the idea that this location is conducive to mixing events (Pickart et al., 2005).

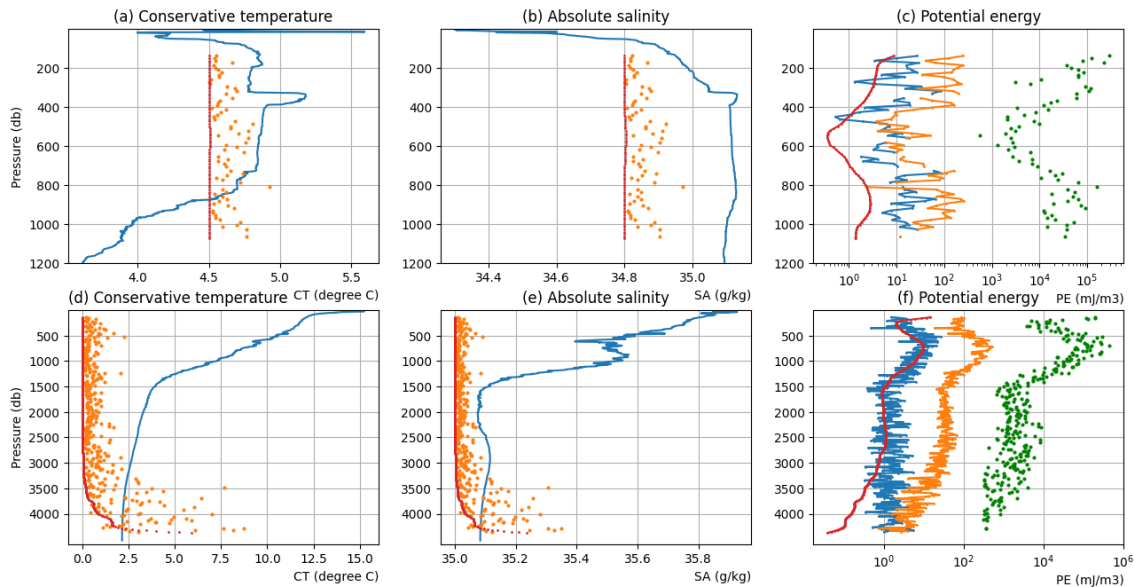


Figure 12. (a)-(b)-(c): as Fig. 11a,b,c but at $42.23^\circ W - 59.8^\circ N$ in 2016 (cast 88). The scatter (η, η^{err}) have been plotted as $4.5 + (\eta, \eta^{err})/2$ in (a) and $34.8 + 0.2(\eta, \eta^{err})$ in (b). (d)-(e)-(f): as Fig. 11a,b,c but at $20.85^\circ W - 48.03^\circ N$ in 2008 (cast 37). The scatter (η, η^{err}) have been plotted as $10(\eta, \eta^{err})$ in (a) and $35 + 0.4(\eta, \eta^{err})$ in (b).

⁷These arise as a result of a different response time of the temperature and conductivity sensors in the CTD instrument. These response times are better known for the Seabird instrument than the Neil Brown, and so more accurately corrected during calibration (Kokoszka, 2012).



A very different behaviour is seen over the East Thulean Rise in a cast near $21^\circ W$ obtained in 2008 (Fig. 12d,e,f). There, despite isolated thermohaline intrusions enhancing the scatter near 500db and 1250db, there is generally an excellent agreement
475 between E_m and E_m^{GM} from the sea surface to about 2500db and thus a low scatter over this layer. However, as one moves to deeper pressures the two curves increasingly depart from one another, E_m approaching E_m^* to within 1000db from the seafloor, with a resulting increase in the scatter towards the bottom of the profile. The seafloor near $21^\circ W$ is a hotspot of turbulent kinetic energy (TKE) dissipation in the 2008 OVIDE section (Ferron et al., 2014) –in fact we chose the profile in Fig. 12d,e,f to match that in their Figs. 5a, 11a. Further analysis by Ferron et al. (2016) with sections collected over 2002-2012
480 suggested that this enhancement of bottom dissipation is due to the generation of topographic internal waves by mesoscale driven flows, quite variable in this region, and associated with the NAC regime. Likewise, the Reykjanes Ridge was found to be a region of enhanced TKE dissipation associated with tidal forcing in Ferron et al. (2016) and the tendency to show increased scatter near the seafloor is also found there (Fig. S5a,b,c). We discuss in more depth the relationship between scatter and diagnostics of turbulent mixing in section 4.3.

485 Besides western boundary current regimes and flow topography interactions we also found that double diffusive convection was able to enhance the scatter well beyond the low background value provided by the GML wavefield. An example is provided in Fig. S5d,e,f for a cast obtained near $15^\circ W$ in 2012 which displays thermohaline staircases below the core of the Mediterranean Water and a scatter enhanced to values ≥ 0.5 , with potential energies $E_m = O(10mJm^{-3})$, as large as those seen in the EGIC in Fig. 12c.

490 3.4 Horizontal averaging of the scatter

Having looked in details at the regional distribution and vertical profiles of the scatter η in the late OVIDE sections (2008-2023), we now average this variable horizontally, considering all longitudes and focusing on the ($h = 25m, n_l = 5$) calculation. Figure 13a (red curve) shows the resulting profiles as a function of depth for the scatter η_{err} due to instrument errors. As expected from the general weakening of the stratification with depth and the definition of η^{err} –see equation (21), the scatter
495 expected from instrument error shows a pronounced increase with depth, reaching values approaching 0.6 at the deepest levels. In sharp contrast, the calculation carried out by averaging at fixed height above the seafloor (Fig. 13b, red curve), shows a much reduced depth dependence, with values below 0.2 even closest to the seafloor. Inspection of Fig. 7b,d indicates that this contrasting behaviour reflects the presence of very large instrumental scatter in the Iberian Abyssal Plain caused by the very weak stratification at pressure exceeding 4000db –the only region where such depths are reached in the OVIDE sections. The
500 scatter η^{err} is not however systematically the largest near the seafloor, a fact clearly illustrated in Fig. 8b,d in the Irminger Sea where the stratification at the bottom is actually greater than that at 1000db. This is why when averaging as in Fig. 13b, the profile of η^{err} shows less vertical structure compared to Fig. 13a.

The profiles for the actual scatter η (orange curves for analysis layers satisfying $\Gamma > 3$) show a similar behaviour. One notices an even more pronounced tendency to approach a uniform profile when considering averages at fixed height above the seafloor,
505 with values departing only weakly from $\eta = 0.07$ in Fig. 13b. The averaging at fixed depth (Fig. 13a) shows an enhancement of the scatter near 1500db which is associated with the high scatter regime found below the Mediterranean Water (Figs. 10b and

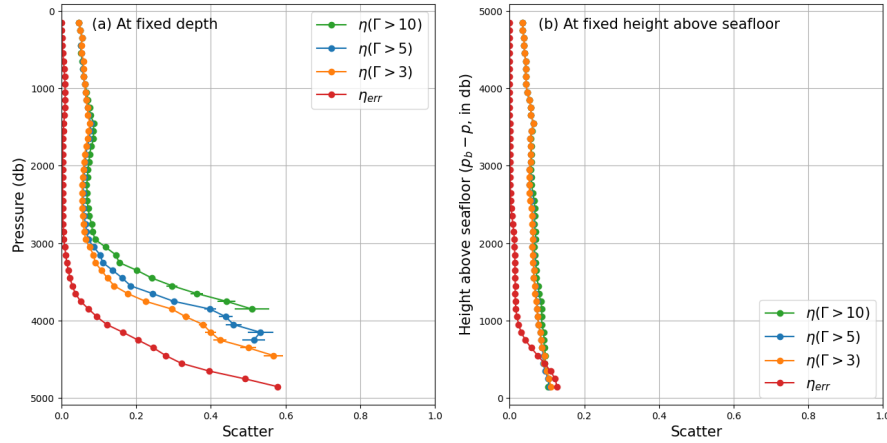


Figure 13. Vertical profiles of horizontally (isobarically) averaged scatter over the late OVIDE sections (2008-2023) based on analysis layers obtained using ($h = 25m, n_l = 5$). Averaging (a) at fixed pressure (b) at fixed height above the seafloor (computed as in Fig. 8). The scatter due to instrument errors η_{err} is plotted in red while the actual scatter is shown in orange (if $\Gamma > 3$), in blue (if $\Gamma > 5$) and in green (if $\Gamma > 10$). All scatters were averaged over $45^\circ W - 0^\circ$ in bins of $100db$ width. The standard error of the mean, taking into account the varying number of layers used for each depth bin (at least 10 were required to estimate the mean), is indicated as thin horizontal bars on all profiles.

Fig. S5d,e,f). It also shows a very striking increase with pressure for $p > 3000db$. Inspection of Fig. 10b suggests that this is not only reflecting the high scatter location over the East Thulean Rise discussed in section 3.3.3, but rather the tongue of high scatter occurrence extending further east and deeper down from it. Indeed, repeating the analysis but restricting the averaging to longitudes between $15^\circ W$ and $0^\circ E$ produces a very similar profile for η (not shown). The study of the internal wavefield by Ferron et al. (2014, 2016) show that the Iberian Abyssal Plain is quiescent in terms of TKE dissipation but contains both upward and downward propagating waves with a frequency close to inertial. We infer from the monochromatic wave model in section 2.2.1 that the stratification is so low in this region that even weak vertical velocity perturbations are associated with large wave steepness $\sigma > 1$. For a mean stratification $N_o^2 \simeq 0.6 \times 10^{-7} s^{-2}$ (see section 4.2) and $h = 25m$, inspection of Fig. 2c (dashed orange curve) provides $E_m^* \sim 1mJm^{-3}$. Thus little wave energy is needed to make $E_m \sim E_m^*$, and in this regime both large scatter and the presence of “no- N^2 layers” is expected (see an illustration in Appendix B, Figs. B1b, B2b), in agreement with the observations reported here for this region⁸. Figure 13 also displays the result of calculations with a different threshold for Γ which are discussed further in section 4.1.

⁸To make this argument more quantitative, we can relate the steepness $\sigma = \beta \delta b / N_o^2$ in section 2.2.1 to the potential energies $E_m \approx \rho_o (\delta b)^2 / 4 N_o^2$ and $E_m^* = \rho_o N_o^2 h^2 / 24$ via $\sigma \approx \beta h (E_m / 6 E_m^*)^{1/2}$. In the Iberian Abyssal Plain, $E_m \approx E_m^*$ leading to $\sigma \approx \beta h / \sqrt{6} > 1$ for β on the low wavenumber side of the interval (13).



3.5 Statistical distribution of the scatter

520 A compact representation of the scatter is provided by its PDF, shown here in Fig. 14 based on analysis layers with $\Gamma > 3$ (“motion dominated regime”) in the late (2008-2023) OVIDE sections. The calculation was carried out for the ($h = 25m, n_l = 5$) case in panel (a), based on more than 146,000 layers, and for the ($h = 100m, n_l = 1$) case in panel (b), using more than 39,000 layers. In both cases, we also superimpose the scatter expected at the corresponding resolution from the GML internal wavefield discussed in section 2.2.2 (black continuous curve, $j_* = 3$). The 25m case in panel (a) shows that scatter events satisfying
 525 $0.025 \leq \eta < 0.05$ (2nd bin in the PDF) have a probability of occurrence comparable to that seen in the GML wavefield. However, the latter is not able to generate the long tail of the distribution which, from the discussion in Sections 3.3 and 3.4, we associate with thermohaline intrusions, the EGIC regime, mesoscale driven flows (East Thulean Rise) and tidal flows (Reykjanes Ridge) within 1000db of the seafloor, double diffusive convective events below the core of the Mediterranean Water, and weak inertial waves in the very low stratification environment of the Iberian Abyssal Plain. Consequently there is a much
 530 smaller occurrence of very low scatter events $\eta < 0.025$ (1st bin in the PDF) in the observations compared to the expectation from GML. Overall, about 83 % of the layers in the $h = 25m$ calculation have a scatter $\eta \leq 0.1$. In other words, even at the small vertical scales 12.5 – 50m, buoyancy profiles captured by the OVIDE sections are linear to an excellent degree and the sampled ocean is best described as “quiescent”. The calculation with $h = 100m$ (Fig. 14b) shows even more linearity on the 50 – 200m scale, with more than 98 % of the layers having $\eta \leq 0.1$. Unlike the situation with $h = 25m$, this PDF is monotonic
 535 and thus more qualitatively consistent with the GML behaviour at this larger vertical scale. The inability of the GML wavefield to generate long tails in the PDF is however even more pronounced at $h = 100m$ than it was at $h = 25m$. For reference we have indicated in red two slopes which characterize the distribution at low and high values of the scatter and suggest a power law behaviour in Fig. 14a,b.

Out of curiosity we have produced synthetic GML buoyancy profiles with different values for the dominant wavenumber j_* .
 540 Figure 14 illustrates the resulting scatter PDFs as gray lines for $h = 25m$ ($j_* = 8$) and for $h = 100m$ ($j_* = 13$). It can be seen that these values produce a better agreement, although not yet perfect, with the observed scatter compared to the canonical model $j_* = 3$. This is interesting as it shows that while it is probably possible to fit a GML-type model to the observed scatter, application of this “big data” approach (e.g., Pollmann, 2020; Dematteis et al., 2024) would miss here the correct physics: as we suggested in Sections 3.3 and 3.4, the enhanced scatter is not due to weak interactions of internal waves, as has been
 545 suggested to govern Garrett-Munk type models (e.g., McComas and Müller, 1981), but to several other processes discussed in these sections. Indeed the canonical model with $j_* = 3$ is clearly relevant in terms of scatter, as was demonstrated by the comparison of E_m and E_m^{GM} in Figs. 11, 12d,e,f, S5 and should thus not be dismissed. Its relevance to the overall internal wave field across the OVIDE section was also emphasized by Ferron et al. (2016).

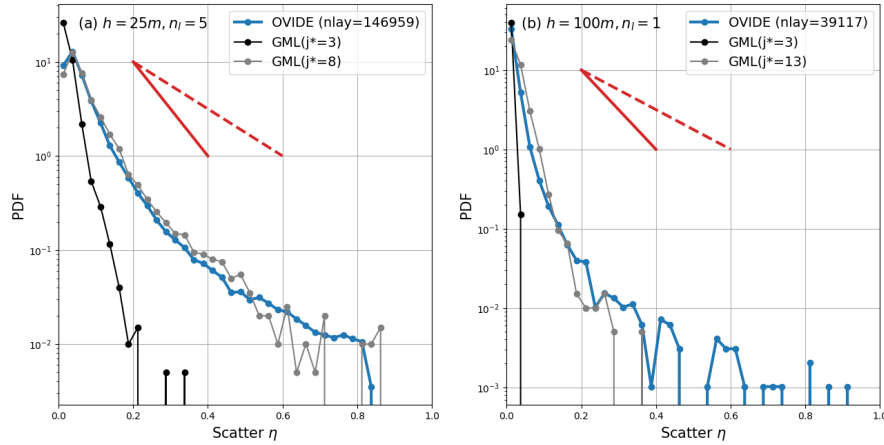


Figure 14. PDFs of the scatter η (blue curves) built over the late OVIDE sections (2008-2023) with all analysis layers satisfying $\Gamma > 3$ (total number indicated in the inset in parenthesis) (a) $h = 25m, n_l = 5$ and (b) $h = 100m, n_l = 1$. The corresponding expectation from the GML internal wavefield simulated in Appendix C is shown in continuous black. Additional GML calculations were performed with $j_* = 8$ and $j_* = 13$, and these are shown as gray lines in panels (a) and (b), respectively. Red lines indicate reference slopes (one decade per 0.2 increase in η in continuous line, and one decade per 0.4 increase in η in dashed line). All PDFs were constructed by normalising the histograms obtained using 40 bins over the $[0, 1]$ interval.

4 Discussion

550 4.1 Sensitivity of the scatter metric

The method presented here contains very few parameters (Fig. 4): the thickness of the layer h over which the scatter is computed, the number n_l which controls the thickness $(2n_l + 1)h$ of the layer used to estimate the “gross stability” N_o^2 , and the threshold on the parameter Γ which separates the “motion dominated regime” from the “error dominated” regime. We discuss these in turn.

555 The value of h sets the vertical scale of the density perturbations of interest because the scatter acts as a band pass filter highlighting scales between $h/2$ and $2h$ —see equation (13). Because we focused on the oceanic finescale, i.e., vertical scales on the order of a few tens to a few hundred meters, we restricted ourselves to two values $h = 25m$ and $h = 100m$. This choice makes low mode internal waves or mesoscale motions quite invisible in this study as their impact on the density field occurs on a larger vertical scale and is captured by the offset q rather than the slope s in the linear decomposition (4). While it could
 560 be interesting to increase h further to values of a few hundred meters to study the impact of these motions on the deep and abyssal density field, assuming these regions have a simple density structure, it would be difficult to do so for intermediate and upper layers of the ocean where the permanent thermocline creates non linear buoyancy profiles on these scales (Feucher et al.,



2019). But even near the seafloor the stratification can be enhanced over a small $\sim 100db$ vertical scale as a result of transport of deep water masses of different origins, as we found in the Irminger Sea in this study for example.

565 The value of n_l affects the calculation of N_o^2 in (21), and thus the calculation of η_{err} . We found that our results were not particularly sensitive to this choice (for $h = 25m$ we considered $n_l = 0, 1, 2, 3, 5$ while we limited ourselves to $n_l = 1$ when $h = 100m$ for the reasons mentioned in the previous paragraph). Rather than affecting much the local value of η_{err} , increasing n_l has the effect of removing analysis layers at the top and bottom of the cast. For example, with the choice $n_l = 5, h = 25m$, the top and bottom $125m$ of the ocean are not sampled. When many repeats of a section are available, an alternative to using
 570 a thicker layer of size $2n_l + 1$ to estimate N_o^2 could be to define an ensemble mean N^2 profile by averaging all sections. This method would remove the need for the parameter n_l but it has not been attempted here.

Finally we tested the sensitivity of our analysis to the threshold used for the parameter Γ by repeating the calculations with $\Gamma > 5$ and $\Gamma > 10$. We found that the main impact of increasing this threshold was to reduce the number of layers available for analysis since the criterion used to isolate the motion dominated regime then became more stringent. All qualitative features
 575 discussed here were not found to be sensitive to changes in the threshold, an example of which is illustrated in Fig. 13 (orange line $\Gamma > 3$, blue line $\Gamma > 5$, green line $\Gamma > 10$). In this particular example, the exact choice of threshold does not qualitatively alter the result although, quantitatively, we notice a sensitivity in Fig. 13a at pressures deeper than $3000db$. The reason for this is clear from inspection of Fig. 9a,b which show that this pressure separates the portion of water column in the “motion dominated” and “error dominated” regimes. In the latter regime, sampling size becomes an issue hence the sensitivity seen in
 580 Fig. 13a. In Fig. 13b, however, the averaging at fixed height above the seafloor samples many regions in the motion dominated regime, even at small value of this coordinate (see for example the high fractions in Fig. 9a in the abyssal parts of the Irminger Sea), and the sensitivity to the exact threshold used for Γ becomes negligible.

4.2 Usefulness of the scatter metric

Most studies of finescale structure in the ocean have focused on estimating TKE dissipation and ultimately vertical diffusivities,
 585 with either global budgets in mind (e.g., Kunze, 2017) or ocean GCMs parameterisations (e.g., Polzin, 2009; Oka, 2025). Here instead of asking how much mixing results from observed perturbed density profiles, we asked whether the density profiles are so disturbed, by either physical motions or instrument errors, that an estimate of the stratification is not meaningful. This question is usually not addressed in studies looking at long term changes in ocean stratification where averaging over large horizontal and vertical domains are considered (e.g., Li et al., 2020; Sallée et al., 2021; Tan and Thurneer, 2023). The
 590 consideration of instrument errors, and in particular the likelihood that they decrease with time, underscores the need to address this issue to detect secular changes robustly. As an illustration of how the scatter diagnostic developed here could be applied to this problem, we consider the Iberian Abyssal Plain below $4500db$ and between $16^\circ W - 10^\circ W$. As seen in Fig. 9b, this region is squarely in the “error dominated regime” because the stratification is so low there that η_{err} dominates the scatter and the parameter Γ remains close to unity. So for this region, it is sensible to assume that the measured stratification $\tilde{N}^2$ is the sum of



595 a true stratification N_o^2 and fluctuations due to instrument errors δN_{ins}^2 :

$$\tilde{N}^2 \approx N_o^2 + \delta N_{ins}^2 \quad (25)$$

Averaging a calculation with $(h = 25m, n_l = 5)$ over this region for each section provides a timeseries displayed in Supplementary Fig. S6 (orange curve). For each year the standard error of the mean was computed and is indicated as a vertical errorbar. There is suggestion of an increasing abyssal stratification with time (black line), although this turns out not to be
600 significant at the 95 % confidence level. The long term mean stratification, denoted with a $\langle . \rangle$ and obtained by averaging over all analysis layers and all sections (sample size $n = 7000$), is found to be $\langle \tilde{N}^2 \rangle = (0.66 \pm 0.01) \times 10^{-7} s^{-2}$ (continuous red line in Supplementary Fig. S6). This number is robust to changing the layer size, the result for $(h = 100m, n_l = 1)$ being $\langle \tilde{N}^2 \rangle = (0.67 \pm 0.01) \times 10^{-7} s^{-2}$ (sample size $n = 1595$). This suggests that the sample is large enough that a robust estimate of the long term mean has been obtained, i.e. $\langle \tilde{N}^2 \rangle \approx N_o^2$ and $\langle \delta N_{ins}^2 \rangle \approx 0$, and the region is best described, given the
605 level of instrument errors, as quiescent with no apparent variability over the 2002-2023 period. This is already a very useful result provided by consideration of the scatter on what controls the variability of the measured stratification in this region. With the sampling issues raised in section 4.1 in mind, we tentatively explore further if this description can be refined. Limiting ourselves now to layers with large Γ , we have:

$$\tilde{N}^2 \approx N_o^2 + \delta N_{phys}^2 \quad \text{for } \Gamma \gg 1 \quad (26)$$

610 where δN_{phys}^2 is the fluctuation in measured N^2 caused by physical motions. If we repeat the calculation above but only consider layers satisfying the condition $\Gamma > 3$, the sample size reduces dramatically to $n = 9$ for $(h = 25m, n_l = 5)$ and $n = 33$ for $(h = 100m, n_l = 1)$, consistent with the view that the low stratification and finite instrument errors allow only very rarely to detect the impact of motions on the density field. The calculation with $(h = 100m, n_l = 1)$ produces an average stratification higher than $N_o^2 = (0.67 \pm 0.01) \times 10^{-7} s^{-2}$, with a value of $(1.1 \pm 0.1) \times 10^{-7} s^{-2}$. This is not entirely a surprise since a large
615 value of $\Gamma = \eta/\eta^{err}$ is readily obtained in layers with large stratification, because for these the value of η^{err} in (21) is reduced. Further analysis of the scatter in these layers suggest that it is low (not shown) and so this sub-sampled population likely reflects layers below anticyclones with a large vertical scale (assuming quasi-geostrophic dynamics apply). The calculation with $(h = 25m, n_l = 5)$ is more surprising because it produces a lower stratification, with an average value of $(0.59 \pm 0.05) \times 10^{-7} s^{-2}$. Here the scatter in these layers is high (not shown), suggesting that they are affected by motions which distort significantly
620 and reduce the stratification (e.g., cyclones of small vertical extent centred near 4500db). It is particularly remarkable that for the $(h = 25m, n_l = 5)$ case the standard error of the mean is only a factor of 5 larger than that based on the raw calculation given the much lower sample size ($\sqrt{7000/9} = 28$). This weak spread between the sub-sampled layers suggests a systematic reduction of the stratification by motions with a vertical scale 12.5 – 50m in this region⁹, although these events can only be detected sporadically. It will be very interesting to apply this methodology to global hydrographic observations of the abyss
625 and analyse the time variability of their stratification through the lenses provided by the two limits expressed by (25) and (26).

We have focused on the application of the scatter metric to hydrographic sections because of the high vertical resolution and quality (calibration) they offer. Nevertheless, the large set of CTD observations provided by ARGO floats, including Deep

⁹In a similar way as, in the troposphere, cyclonic eddies developing on the jet stream set a lower bound on the midlatitudes stratification (Jukes, 2000).



ARGO, could also be used. In doing so it must be acknowledged that the vertical resolution (on the order of $\sim 10m$ above 1000m depth and $\sim 25m$ below – see Wong et al., 2020) will not allow a calculation of scatter using $h = 25m$. Nevertheless, the calculation with $h = 100m$ is entirely possible and would provide a three-dimensional view of this metric for the global ocean, as well as a better grasp of its time dependence and intermittence. The calculation of the associated instrument error will also be very useful to consider as, unlike ship-based CTD measures which are re-calibrated frequently, ARGO floats sensors can drift over time.

Besides its application to measurements of oceanic stratification and ARGO floats, the scatter metric introduced here is also complementary to techniques developed to study thermohaline intrusions (e.g., Shcherbina et al., 2009; Jakes et al., 2025). While these techniques are able to visualise where the intrusions occur and estimate their strength, the scatter offers a straightforward measure of how compensated they are in density. The analysis presented here suggests that non compensation is common and that it enhances the scatter above that expected from a canonical Garrett-Munk internal wave field ($\eta \ll 1$), to values $0.1 \leq \eta \leq 0.2$. It is unclear however why this compensation occurs in some places but not others. Detection of the impact of seamounts on deep and abyssal stratification (Mashayek et al., 2024) is another potentially fruitful area of application of the scatter diagnostic.

Our results suggest that double diffusive convection is able to release the potential energy stored in the temperature or salinity field and leads to even larger values of scatter $\eta > 0.5$, as was found below the Mediterranean Water (Fig. S5d,e,f). Western boundary current regions, such as captured here with the EGIC, also display the ability to generate scatter in excess of $\eta > 0.5$ (Fig. 12a,b,c). So, although it is tempting to use the scatter as a simple metric to separate types of motions based on their value of η , this approach would be too simplistic. Low values of scatter could as well result from geostrophic motions rather than internal waves, and high values of scatter can also result from very different dynamics as we’ve just highlighted. Ultimately, knowledge of density alone is too limited and horizontal velocity data needs to be considered to offer a more robust separation of the dynamics (e.g., Vladioiu et al., 2024).

Finally, we come back to our motivation (Section 1) of providing the high resolution modelling community with an observational benchmark for the submesoscale and internal wave dynamics. Because these models have a lower vertical resolution than that of the hydrographic data (Torres et al., 2018 for example use $1m$ vertical resolution near the surface and $30m$ all the way to $500m$ depth, with a total of 90 levels in the global MITgcm LLC4320 simulation, while Baker and Mashayek (2022) use a vertical resolution of $10m$ from the sea surface to $600m$ depth and $25m$ thereafter all the way to $4555m$ depth, with a total number of 225 vertical levels in a simulation of the circulation at Drake Passage), we do not expect that they could yet be compared to a calculation of scatter with $h = 25m$. Nevertheless, as with ARGO floats discussed above, the calculation of scatter with $h = 100m$ should be within the reach of these models in the near future and could readily be compared to diagnostics such as the PDF shown in Fig. 14b, computed from the relevant hydrographic sections or ARGO profiles.

4.3 Scatter metric and mixing

The EGIC at the western end of the OVIDE sections, the Reykjanes Ridge and the East Thulean Rise are locations where the scatter is enhanced in Fig. 10b. These regions also correspond to locations where the TKE dissipation rate ϵ estimated in



OVIDE sections from a finescale mixing parameterization by Ferron et al. (2016) is largest¹⁰. This correspondance is however not systematic, the most striking example of which being the Iberian Abyssal Plain where Ferron et al. (2016) found low values of ϵ and where we observe the highest scatter (Fig. 13a). To reinforce this point we note that, as emphasized by Ferron et al. (2016), were the finescale disturbances to follow exactly the GM spectrum, ϵ would vary in proportion to N_o while we showed in Section 2.2.2 that the scatter expected from a GML internal wavefield is essentially independent of N_o . This difference arises from the fact that the scatter is related to the mean of the vertical strain field but not to its variance (see Section 2.2.3), which is one of the variables informing the finescale parameterisation used by Ferron et al. (2016). The dependence of ϵ on stratification is very pronounced in Ferron et al. (2016) who use the $\log_{10} N = -2.8 s^{-1}$ isoline to demark the region of high TKE dissipation in their Fig. 2a. In contrast, this region corresponds to what we refer as the “wedge of Atlantic waters”, and is a region where analysis layers display very high occurrence of low scatter (Fig. 10a and Supplementary Fig. S7a).

Conversely, we could not find an obvious relationship between vertical diffusivity K_v , as estimated from microstructure data by Ferron et al. (2014 –see their Fig. 5b), and the scatter η . This can be understood from the fact most of the observations analysed by Ferron et al. (2014) were in the transitional regime where the Osborn relation $\epsilon \approx 0.2 K_v N^2$ is valid and that we do not find a systematic relation between η and ϵ .

In summary, the scatter measures how far from linear the distribution of buoyancy with depth is on a given vertical scale, but is not providing information about the intensity of mixing resulting from disturbances affecting this buoyancy field. Put simply, while we expect mixing events in a layer to alter drastically its density profile by folding irreversibly isopycnals and in so doing generate high values of scatter for that layer, i.e., high mixing is expected to lead to high scatter, we cannot systematically associate high scatter with high mixing. A striking illustration of this is the staircase-like structure of the buoyancy field produced below the Mediterranean water by double diffusive convection, which is conducive to large scatter (Supplementary Fig. S5d,e,f) but only produces moderate amount of turbulent kinetic energy dissipation (Figs. 5a and 2a in Ferron et al., 2014 and Ferron et al., 2016, respectively) and weak vertical diffusivity (Figs. 5b and Fig. 2b in Ferron et al., 2014 and Ferron et al., 2016, respectively).

5 Conclusion

A new method to quantify and visualize the “noise” present in hydrographic profiles of buoyancy was introduced. It is simply based on the linear regression of buoyancy onto depth over a layer of specified thickness h . This acts as a band pass filter on the buoyancy field, retaining fluctuations with vertical lengthscale between approximately $h/2$ and $2h$. We proposed to quantify the amount of “noise” by the scatter η present in the linear regression. This variable can also be interpreted physically as a measure of the potential energy (E_m) present in the density field on vertical scales between $h/2$ and $2h$ and, specifically, how it compares with the potential energy required to remove the stratification of the layer (E_m^*). The scatter η_{err} expected from instrument errors on buoyancy is also estimated and it is used to isolate “motion” and “error” dominated regimes in hydrographic data.

¹⁰These authors used OVIDE sections over the 2002-2012 period and their Fig. 2a can be compared to a version of Fig. 10 for that time period in Supplementary Fig. S7.



The method does not require any spectral calculations and rely on elementary linear regression techniques. The maps of scatter are visually striking and highlight instantly phenomena of possible interest in a given hydrographic section. Application of the method to repeated OVIDE sections, with a focus on the oceanic finescale ($h = 25, 100m$), leads to the following conclusions:

- The bulk of the sections are characterized by low values of scatter $\eta < 0.1$. This indicates that even on small vertical scales a linear approximation of the buoyancy profile is successful. The ocean, at least in OVIDE, seems mostly quiescent. The canonical Garret-Munk spectrum, at least as modelled here using a Langevin equation, was shown to be relevant to describe this background of low scatter values (i.e., $E_m \approx E_m^{GM}$).
- The geographical distribution of the scatter however indicates the presence of rare events able to disturb the density field very significantly above this background of low values. These events were shown to be associated, with increasing values of scatter, with non density compensated thermohaline intrusions (e.g. in the Irminger Sea, $0.1 \leq \eta \leq 0.2$), energetic western boundary current regime (East Greenland - Irminger current), interaction of tidal (Reykjanes Ridge) and mesoscale flows (East Thulean Rise) with topographic features within $\sim 1000db$ of the seafloor ($\eta > 0.2$), and double diffusive convection (below the core of the Mediterranean Water, $\eta > 0.2$). All these phenomena are such that they inject potential energy on vertical scales within the $[h/2, 2h]$ interval, and bring E_m above the background E_m^{GM} and closer to E_m^* . The region of low abyssal stratification in the eastern portion of the section (Iberian Abyssal Plain) was also shown to be associated with high scatter, and this was suggested to reflect a different regime where E_m^* is so low that even a small amount of wave activity is able to create $E_m \sim E_m^*$.
- The statistical distribution of the scatter (probability density function, or PDF) highlights this rich dynamical structure, with no apparent gaps in the distribution and power law behaviour. The PDF has a distinct peak for $h = 25m$ but decreases monotonically for $h = 100m$.
- The method diagnosed larger instrument errors in the earlier (2002-2006) compared to the later (2008-2023) sections, a finding in agreement with a change of CTD instrument in 2007.

Because of its straightforward implementation we are hopeful that the calculation of the scatter η in the large and expanding set of in-situ observations (e.g., ship - based hydrography and ARGO floats) data will provide a useful benchmark for the submesoscale and internal wave dynamics simulated by high resolution OGCMs and regional ocean models. We also suggested that calculation of the scatter and its decomposition into “motion” and “error” dominated regimes could help in detecting and understanding changes in stratification in the global ocean, in quantifying the near density-compensation of thermohaline intrusions, and in isolating the impact of seamounts and other topographic features on deep and abyssal stratification.

Besides these important practical applications, we wish to emphasize the somewhat surprising “quiescence” of the density field on vertical scales of a few tens of meters revealed by the scatter maps (mostly dark with the choice of color palette used in this study). Preliminary analysis of a much larger set of hydrographic sections from the GO-SHIP easy ocean database (Katsumata et al., 2022) indicates that this finding also applies to many other sections and appears to be a defining feature of the global ocean. It seems as if the large scale stratification is modulated “elastically” by mesoscale eddies and low mode



internal waves (invisible here because their vertical scale is comparable to that of the large scale circulation) but also by motions with much smaller vertical scales. Besides bottom flows impinging on topographic features, it is noticeable that the other “disruptors” of the stratification are associated with the thermohaline composition of the ocean (intrusions and double diffusive convection), and one wonders whether the large scale circulation is in effect trying to minimise the energy put into these phenomena to preserve its own existence. This intriguing hypothesis deserves further investigations.

Appendix A: Potential energy calculation

Following Holiday and McIntyre (1980), we consider the energy E required to create, at a given depth z , a “buoyancy excess” $b_e(z)$ compared to a reference state $b_o(z)$. Because we focus on potential energy associated with motions of small vertical scales, for practical purposes we restrict ourselves to situations where $b_o(z) \approx N_o^2 z$ where N_o^2 is a constant and z is measured from the centre of a given layer. We also want to include non linear situations where the original density field present in the layer can become quite disturbed, and so we consider Holiday and McIntyre (1980)’s formula for constant stratification:

$$E = \frac{g^2}{N_o^2} \left((\rho_o + \rho_e) \ln(1 + \frac{\rho_e}{\rho_o}) - \rho_e \right) \quad (A1)$$

Here $\rho_o(z) = \rho_o(0)e^{-N_o^2 z/g}$ is the undisturbed density profile and $\rho_e = \rho_e(z)$ is the “excess density” estimated from the “excess buoyancy” $b_e(z) = b(z) - b_o(z)$ via the definition

$$\rho_e(z) \triangleq \rho_o(z) \left(e^{-b_e/g} - 1 \right). \quad (A2)$$

Note that because fixed volume displacements are considered in Holiday and McIntyre (1980), E is an energy density, accordingly expressed in Jm^{-3} .

Given that typically $N_o^2 \leq 10^{-4} s^{-2}$, the vertical scale $g/N_o^2 \geq 100km \gg h$ so we can safely approximate $\rho_o(z)$ by $\rho_o(0)$ for $|z| \leq h/2$ when it comes as a multiplicative factor. In addition, since for practical purposes $b_e/g \ll 1$, an asymptotic expansion of (A1) based on this small parameter leads to:

$$E(z) \approx \frac{1}{2} \rho_o(0) \frac{b_e(z)^2}{N_o^2} \quad (A3)$$

which is the form considered in Section. 2.1.2.

For the analysis of observations we use the full form (A1) and, to compute the mean “finescale potential energy” E_m defined in (9), we consider a discretized version based on the linear regression (4) applied to the “local” layer of thickness h ($n_l = 0$):

$$E_k = \frac{g^2}{s} \left((\rho_{o,k} + \rho_{e,k}) \ln(1 + \frac{\rho_{e,k}}{\rho_{o,k}}) - \rho_{e,k} \right) \quad (A4)$$

where $\rho_{o,k} = \rho_o(0)e^{-sz_k/g}$, the quantity $\rho_o(0)$ being taken as ρ_m (see Section 2.1.1), $\rho_{e,k} = \rho_{o,k} (e^{-b_{e,k}/g} - 1)$ and $b_{e,k} = b_k - (q + sz_k)$. The mean potential energy E_m of the layer can then be estimated as:

$$E_m = \sum E_k / (2K + 1), \quad (A5)$$



while the reference potential energy E_m^* is computed as:

$$755 \quad E_m^* = \rho_m s h^2 / 24. \quad (\text{A6})$$

Note that in this expression the parameter s is the slope of the linear regression applied to the “gross” layer of thickness $h(1 + 2n_l)$, a measure of N_o^2 in (10).

Appendix B: Scatter and band pass filter

We start with the non dimensional form of (12):

$$760 \quad b_k / (N_o^2 h) = z_k^* + \frac{\sigma}{\kappa} \cos(\kappa z_k^* + \phi) \quad \text{for } |z_k^*| \leq 1/2 \quad (\text{B1})$$

where $z_k^* \equiv z_k/h$ is a non dimensional depth and the non dimensional numbers $\sigma = \beta \delta b / N_o^2$ and $\kappa = \beta h$ were introduced in Section 2.2.1. The linear regression of b_k against z_k^* over the layer will solely depend on σ, κ and ϕ since the factor $N_o^2 h$ is just a multiplicative constant. For a given value of σ and κ , different values of the phase will lead to different values of the scatter η , of the offset q and slope s in the linear regression (4), i.e., we have $\eta = \eta(\sigma, \kappa, \phi)$ and $s = s(\sigma, \kappa, \phi)$.

765 The scatter calculation was applied to a discretisation of (B1) with $K = 12$ (the standard case in the OVIDE analysis in Section 3 for $h = 25m$ and a cast vertical resolution of $1db$). A population of phase values was constructed from 3600 bins, one-tenth of a degree each, while 25 equally spaced values of σ in the interval $0 \leq \sigma \leq 2$, and a linear distribution of κ in log-space (50 values spread between of $2\pi \times 25m/1m$ and $2\pi \times 25m/4000m$) were used. As in the observational analysis, the scatter was only computed when the linear regression was statistically significant at the 95 % level in a two-sided t-test, and

770 when the slope s of the linear regression was positive (corresponding to a physically possible undisturbed stratification).

Figure B1 (color field) summarizes the resulting statistics. The lower quartile of the distribution of $\eta(\sigma_o, \kappa_o, \phi)$ at fixed $(\sigma = \sigma_o, \kappa = \kappa_o)$ displayed in (a) remains below 0.2 even considering non linear waves ($\sigma > 1$, y-axis). The upper quartile, shown in (b), can however exceed values $\eta = 0.5$. The expected value in (c), i.e. the average over all values of the phase ϕ , is overall very small over the domain of parameters considered, with typically $\eta < 0.1$. An interesting aspect shared by all

775 panels is the maximum seen when $\kappa/2\pi = O(1)$, consistent with the discussion of Fig. 1 in Section 2.2.1, and indicating that the scatter acts as a band pass filter on the monochromatic perturbations. For future reference we have highlighted the range of values $\frac{1}{2} \leq \kappa/2\pi \leq 2$ by white dashed vertical lines in Fig. B1. These correspond to vertical wavelength between half and twice the size of the layer and capture roughly the location of the extrema in Fig. B1, or approximatively the bandwidth of the filter.

780 At a given value (σ, κ) , the potential energy E_k at each level k within the layer where the linear regression (4) is applied can be computed from (A4) as a function of the phase ϕ of the perturbation. Its lower quartile, upper quartile and expected value based on the phase population is plotted in Fig. B1a,b,c, respectively, in purple contours. The contouring is logarithmic, normalising E_k by the reference value E_m^* in (10). In agreement with (11) the contours follow closely those of the scatter shown by the colorfield. In particular, the bandwidth $1/2 \leq \kappa/2\pi \leq 2$ can also be seen to capture approximately the “ridge” of

785 high “finescale potential energy”.

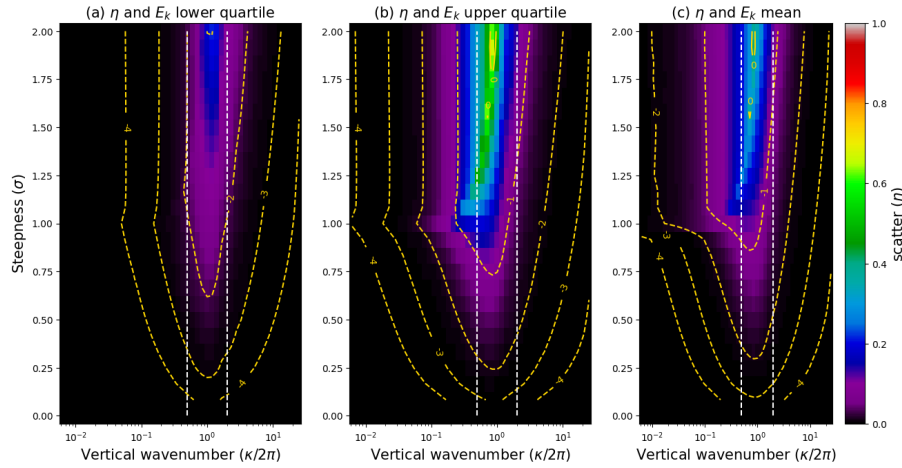


Figure B1. Scatter η and potential energy E_k found in the monochromatic wave model as a function of the non dimensional parameters σ (y-axis) and $\kappa/2\pi$ (x-axis). Both η and E_k are functions of (σ, κ, ϕ) and their (a) lower quartile (b) upper quartile and (c) mean value based on the population of values for ϕ is shown in color for η , and in gold contours for E_k . The latter display the value $\log_{10}(E_k/E_m^*)$, with dashed contours when negative and continuous contours when positive. The bandwidth $1/2 \leq \kappa/2\pi \leq 2$ is highlighted with two vertical white dashed lines.

Finally, we discuss briefly the presence of “overturning layers” (regression slope $s < 0$ and statistically significant) and “no- N^2 layers” (s not statistically significant) introduced in Section 2.3. The occurrence of overturnings (Fig. B2a) increases with σ but remains relatively low, 30 % at most, and is confined to the low wavenumber portion ($\kappa/2\pi \leq 1/2$) of the domain. This is because such large vertical wavelengths appear as straight line on the scale h of the layer and, for some values of the phase, the slope of these lines is negative. As expected from the large non linearity when overturns occur, their occurrence coincide with the region of parameter where the approximate and full potential energy calculations differ most strongly (see previous paragraph). Turning now our attention to the occurrence of “no- N^2 ” layers in Fig. B2b, we see that it is non zero only when $\sigma \geq 1$, consistent with the fact that the vertical derivative of the buoyancy profiles are bounded by $N_o^2(1 \pm \sigma)$ in the model (12). The presence of “no- N^2 layers” is even more rare than that of “overturning layers”, with the highest occurrence found on the low wavenumber side of the bandwidth $1/2 \leq \kappa/2\pi \leq 2$ when $\sigma > 1$. This region of parameter is where there is more likelihood of E_k approaching E_m^* (see Fig. B1, purple contours) and where pronounced “S-shape” profiles of buoyancy develop (not shown). We see here that it encompasses highest occurrence of both “overturning” and “no- N^2 ” layers.

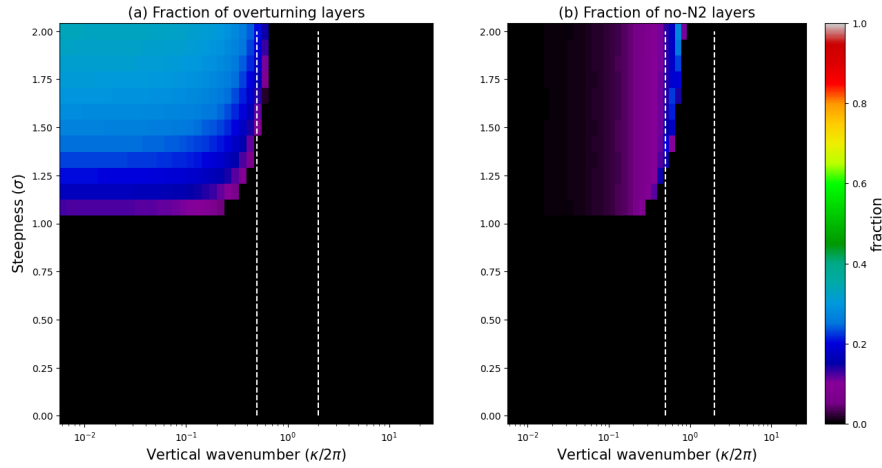


Figure B2. As in Fig. B1 but for the properties of the stratification found in the monochromatic wave model. The occurrence of (a) “overturning layers” and (b) “no- N^2 ” layers, based on the population of phase values, are shown in colors.

Appendix C: Synthetic buoyancy profiles satisfying the GM buoyancy spectrum

We consider buoyancy perturbations at depth z to be realisations of a random variable B obeying the Langevin equation (e.g.,
 800 Lemons, 1949):

$$B(z + dz) - B(z) = -\gamma B(z) + (\mu^2 dz)^{1/2} \mathcal{N}_z^{z+dz}(0, 1) \quad (C1)$$

Here the term $\mathcal{N}_z^{z+dz}(0, 1)$ is a Gaussian random variable over the depth interval $[z, z + dz]$ with zero mean and variance of unity. The two parameters of the model are the “damping” γ and the strength of the noise μ . If the model is numerically solved from a random initial condition at a depth $z = 0$, the buoyancy field will evolve towards a normal distribution with zero mean
 805 and variance $\mu^2/(2\gamma)$ (Lemons, 1949). The spectrum F_B of the buoyancy perturbations as a function of wavenumber β is red with a spectral slope -2 for wavenumbers $\beta \gg \gamma$, and given by:

$$F_B(\beta) = \frac{(\mu^2/\pi)}{\beta^2 + \gamma^2} \quad (C2)$$

In order to solve (C1) we need to specify the two parameters γ and μ . By inspection of (16) and (C2) we set

$$\gamma = \beta_\star = \frac{\pi j_\star}{D} \left(\frac{N_o}{N_{surf}} \right) \quad (C3)$$

810 and

$$\mu^2 = N_o^4 E D j_\star \pi. \quad (C4)$$

Equation (C1) was solved numerically for a given value of stratification N_o , and thus for a given layer centred on the depth with this value of the Brunt-Väisälä frequency, since the latter determines the values of γ and μ from eqs. (C3)-(C4).



The ‘time stepping’ was set to $\Delta z = 1m$ to provide a straightforward comparison with observed profiles and produce the
815 $b'_{k,GM}$ ($k = 1, \dots, 2K + 1$) in (15). Values of the parameters N_{surf} , D , E and j_* were set to those in Munk (1981), namely
 $N_{surf} = 5.2 \times 10^{-3} s^{-1}$, $D = 1300m$, $E = 6.3 \times 10^{-5}$ and $j_* = 3$. To produce the statistics used in Figs. 2, 3 and 14, an
ensemble of 1000 realisations of buoyancy perturbations within the layer of thickness h were produced over a range of depth
between $4000m$ and $500m$, every $500m$. The number of layers thus produced is 8000, which correspond roughly to the number
of “analysis layers” obtained with one or two OVIDE sections (See Table 1, 2nd column). The PDFs labelled GML in Figs. 3
820 and 14 were built using all these layers.

To compute the potential energies E_m^{GM} and E^{GM} at different depths (or equivalently here, different values of stratification),
a density $\rho(z)$ is needed (see Appendix A). This was achieved by first converting the depths z to pressure $p(z)$ using a standard
routine from the GSW package (McDougall and Barker, 2020), and then using the equation of state $\rho(z) = \rho\{S^{ref}, \Theta^{ref}, p(z)\}$
where S^{ref} and Θ^{ref} are defined in section 3.1.

825 *Code availability.* All analyses were conducted using Python. The codes are available upon request to AC.

Data availability. The OVIDE sections used in this study are available at: <https://www.seanoe.org/data/00353/46448/> and
<https://www.seanoe.org/data/00844/95607/>

Author contributions. AC conducted the analysis and wrote the text. PL and HM advised on the data analysis and the writing of the
manuscript. All authors reviewed and approved the final version.

830 *Competing interests.* The authors have no competing interests to declare.

Acknowledgements. Discussions with, and encouragements from, Jon Murray, Paulo Ceppi, Costanza Rodda, Andrea Wäschchenbach, Anda
Vladoiu, Helen Czerski and the London Ocean Group were deeply appreciated by AC. All authors also wish to thank the crews and cap-
tains operating during the many OVIDE campaigns whose dataset is invaluable to our study, and acknowledge support from the French
Oceanographic Fleet to OVIDE (<https://doi.org/10.18142/140>; OVIDE Group, 2026) and the Spanish Oceanographic Fleet. This work re-
835 ceived funding from the French national programme Les Enveloppes Fluides et l’Environnement (LEFE) and Ifremer. HM was supported by
the Centre National de La Recherche Scientifique (CNRS). PL was funded by Ifremer and by the EuroGO-SHIP project (Horizon Europe
101094690).



References

- Baker, L.E. and Mashayek, A., 2022. The impact of representations of realistic topography on parameterized oceanic lee wave energy flux. *Journal of Geophysical Research: Oceans*, 127(10), p.e2022JC018995. <https://doi.org/10.1029/2022JC018995>
- 840 Branellec, P., Lherminier, P., Reynaud, T., Le Bihan, C., Reynaud, T., and Le Bihan, C.: OVIDE 2018. CTD-O2 Data Report, <https://doi.org/10.13155/86588>, 2019.)
- D'Asaro, E.A. and Morehead, M.D., 1991. Internal waves and velocity fine structure in the Arctic Ocean. *Journal of Geophysical Research: Oceans*, 96(C7), pp.12725-12738. <https://doi.org/10.1029/91JC01071>
- 845 Daniault, N., Lherminier, P. and Mercier, H., 2011. Circulation and transport at the southeast tip of Greenland. *Journal of Physical Oceanography*, 41(3), pp.437-457. <https://doi.org/10.1175/2010JPO4428.1>
- Dematteis, G., Le Boyer, A., Pollmann, F., Polzin, K.L., Alford, M.H., Whalen, C.B. and Lvov, Y.V., 2024. Interacting internal waves explain global patterns of interior ocean mixing. *Nature Communications*, 15(1), p.7468. <https://doi.org/10.1038/s41467-024-51503-6>
- Feucher, C., Maze, G. and Mercier, H., 2019. Subtropical mode water and permanent pycnocline properties in the world ocean. *Journal of Geophysical Research: Oceans*, 124(2), pp.1139-1154. <https://doi.org/10.1029/2018JC014526>
- 850 Fernandez Castro, B., Fernandez Roman, D., Ferron, B., Fontela, M., Lherminier, P., Naveira Garabato, A., Perez, F.F., Spingys, C., Polzin, K. and Velo, A., 2024. Isopycnal eddy stirring dominates thermohaline mixing in the upper subpolar North Atlantic. *Journal of Geophysical Research: Oceans*, 129(9), p.e2023JC020817. <https://doi.org/10.1029/2023JC020817>
- Ferron, B., Kokoszka, F., Mercier, H. and Lherminier, P., 2014. Dissipation rate estimates from microstructure and finescale internal wave observations along the A25 Greenland-Portugal OVIDE line. *Journal of Atmospheric and Oceanic Technology*, 31(11), pp.2530-2543. <https://doi.org/10.1175/JTECH-D-14-00036.1>
- 855 Ferron, B., Kokoszka, F., Mercier, H., Lherminier, P., Huck, T., Rios, A. and Thierry, V., 2016. Variability of the turbulent kinetic energy dissipation along the A25 Greenland-Portugal transect repeated from 2002 to 2012. *Journal of Physical Oceanography*, 46(7), pp.1989-2003. <https://doi.org/10.1175/JPO-D-15-0186.1>
- 860 Garrett, C. and Munk, W., 1972. Space-time scales of internal waves. *Geophysical Fluid Dynamics*, 3(3), pp.225-264. <https://doi.org/10.1080/03091927208236082>
- Gargett, A. and Garner, T., 2008. Determining Thorpe scales from ship-lowered CTD density profiles. *Journal of Atmospheric and Oceanic Technology*, 25(9), pp.1657-1670. <https://doi.org/10.1175/2008JTECHO541.1>
- Gregg, M.C. and Kunze, E., 1991. Shear and strain in Santa Monica basin. *Journal of Geophysical Research: Oceans*, 96(C9), pp.16709-16719. <https://doi.org/10.1029/91JC01385>Digital Object Identifier (DOI)
- 865 Holliday, D. and McIntyre, M.E., 1981. On potential energy density in an incompressible, stratified fluid. *Journal of Fluid Mechanics*, 107, pp.221-225. doi:10.1017/S0022112081001742
- Jakes, M.I., Phillips, H.E., Cyriac, A., Foppert, A., Bindoff, N.L., Rintoul, S.R. and Polzin, K.L., 2025. Thermohaline interleaving in the Antarctic Circumpolar Current. *Journal of Physical Oceanography*, 55(3), pp.205-227. <https://doi.org/10.1175/JPO-D-24-0048.1>
- 870 Jukes, M.N., 2000. The static stability of the midlatitude troposphere: The relevance of moisture. *Journal of the Atmospheric Sciences*, 57(18), pp.3050-3057. [https://doi.org/10.1175/1520-0469\(2000\)057<3050:TSSOTM>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<3050:TSSOTM>2.0.CO;2)
- Kokoszka, F., 2012: Estimations du mélange vertical le long de sections hydrologiques en Atlantique Nord. *Sciences de la Terre. Université de Bretagne occidentale– Brest*. <NNT: 2012BRES0097>. <tel-02071038>. 10.70675/df39e236zf117z4a1az8f55z693c83b98dab



- Kunze, E., 1993. Submesoscale dynamics near a seamount. Part II: The partition of energy between internal waves and geostrophy. *Journal of physical oceanography*, 23(12), pp.2589-2601. [https://doi.org/10.1175/1520-0485\(1993\)023<2589:SDNASP>2.0.CO;2](https://doi.org/10.1175/1520-0485(1993)023<2589:SDNASP>2.0.CO;2)
- Kunze, E., 2017. Internal-wave-driven mixing: Global geography and budgets. *Journal of Physical Oceanography*, 47(6), pp.1325-1345. <https://doi.org/10.1175/JPO-D-16-0141.1>
- Lemon, D. S., 1949: An introduction to stochastic processes in Physics, John Hopkins University Press, 110p. <https://doi.org/10.1119/1.1526134>
- Li, G., Cheng, L., Zhu, J., Trenberth, K.E., Mann, M.E. and Abraham, J.P., 2020. Increasing ocean stratification over the past half-century. *Nature Climate Change*, 10(12), pp.1116-1123. <https://doi.org/10.1038/s41558-020-00918-2>
- Lien, R.C. and Muller, P., 1992. Consistency relations for gravity and vortical modes in the ocean. *Deep Sea Research Part A. Oceanographic Research Papers*, 39(9), pp.1595-1612. [https://doi.org/10.1016/0198-0149\(92\)90050-4](https://doi.org/10.1016/0198-0149(92)90050-4).
- Lien, R.C. and Sanford, T.B., 2019. Small-scale potential vorticity in the upper-ocean thermocline. *Journal of Physical Oceanography*, 49(7), pp.1845-1872. <https://doi.org/10.1175/JPO-D-18-0052.1>
- Mashayek, A., Gula, J., Baker, L.E., Naveira Garabato, A.C., Cimoli, L., Riley, J.J. and De Lavergne, C., 2024. On the role of seamounts in upwelling deep-ocean waters through turbulent mixing. *Proceedings of the National Academy of Sciences*, 121(27), p.e2322163121. <https://doi.org/10.1073/pnas.2322163121>
- Mazloff, M.R., Cornuelle, B., Gille, S.T. and Wang, J., 2020. The importance of remote forcing for regional modeling of internal waves. *Journal of Geophysical Research: Oceans*, 125(2), p.e2019JC015623. <https://doi.org/10.1029/2019JC015623>
- McComas, C.H. and Müller, P., 1981. The dynamic balance of internal waves. *Journal of Physical Oceanography*, 11(7), pp.970-986. [https://doi.org/10.1175/1520-0485\(1981\)011<0970:TDBOIW>2.0.CO;2](https://doi.org/10.1175/1520-0485(1981)011<0970:TDBOIW>2.0.CO;2)
- McDougall, T.J. and Barker, P.M., 2014. Comment on “Buoyancy frequency profiles and internal semidiurnal tide turning depths in the oceans” by B. King et al. *Journal of Geophysical Research: Oceans*, 119(12), pp.9026-9032. <https://doi.org/10.1002/2014JC010222>
- McDougall, T. J., and P. M. Barker, 2020. Getting Started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox Version 3.06.12. SCOR - IAPSO WG127. <https://www.teos-10.org/software.htm>
- Millard, R.C., Owens, W.B. and Fofonoff, N.P., 1990. On the calculation of the Brunt-Vaisala frequency. *Deep Sea Research Part A. Oceanographic Research Papers*, 37(1), pp.167-181. [https://doi.org/10.1016/0198-0149\(90\)90035-T](https://doi.org/10.1016/0198-0149(90)90035-T).
- Müller, P., 1988: Vortical motions. Elsevier oceanography series, 46, pp.285-301. [https://doi.org/10.1016/S0422-9894\(08\)70553-1](https://doi.org/10.1016/S0422-9894(08)70553-1).
- Munk, W.H., 1981. Internal waves and small-scale processes. *Evolution of Physical Oceanography*, 264-291, MIT Press. <https://mitocw.ups.edu.ec/resources/res-12-000-evolution-of-physical-oceanography-spring-2007/part-2/>
- Oka, A., 2025. Deep ocean mixing mismatch between model and observational estimates. *Communications Earth & Environment*, 6(1), p.108. <https://doi.org/10.1038/s43247-025-02027-4>
- Pickart, R.S., Torres, D.J. and Fratantoni, P.S., 2005. The east Greenland spill jet. *Journal of Physical Oceanography*, 35(6), pp.1037-1053. <https://doi.org/10.1175/JPO2734.1>
- Pinkel, R., 2014. Vortical and internal wave shear and strain. *Journal of Physical Oceanography*, 44(8), pp.2070-2092. <https://doi.org/10.1175/JPO-D-13-090.1>
- Pollmann, F., 2020. Global characterization of the ocean's internal wave spectrum. *Journal of Physical Oceanography*, 50(7), pp.1871-1891. <https://doi.org/10.1175/JPO-D-19-0185.1>
- Polzin, K.L., Toole, J.M. and Schmitt, R.W., 1995. Finescale parameterizations of turbulent dissipation. *Journal of physical oceanography*, 25(3), pp.306-328. [https://doi.org/10.1175/1520-0485\(1995\)025<0306:FPOTD>2.0.CO;2](https://doi.org/10.1175/1520-0485(1995)025<0306:FPOTD>2.0.CO;2)



- Polzin, K.L., 2009. An abyssal recipe. *Ocean Modelling*, 30(4), pp.298-309. <https://doi.org/10.1016/j.ocemod.2009.07.006>.
- Qiu, B., Chen, S., Klein, P., Wang, J., Torres, H., Fu, L.L. and Menemenlis, D., 2018. Seasonality in transition scale from balanced to unbalanced motions in the world ocean. *Journal of Physical Oceanography*, 48(3), pp.591-605. <https://doi.org/10.1175/JPO-D-17-0169.1>
- 915 Radko, T., 2013. Double-diffusive convection. Cambridge University Press. <https://doi.org/10.1017/CBO9781139034173>
- Ruddick, B., 1983. A practical indicator of the stability of the water column to double-diffusive activity. *Deep Sea Research Part A. Oceanographic Research Papers*, 30(10), pp.1105-1107. [https://doi.org/10.1016/0198-0149\(83\)90063-8](https://doi.org/10.1016/0198-0149(83)90063-8).
- Sallée, J.-B., Pellichero, V., Akhoudas, C., Pauthenet, E., Vignes, L., Schmidtke, S., Garabato, A. N., Sutherland, P., and Kuusela, M., 2021: Summertime increases in upper-ocean stratification and mixed-layer depth, *Nature*, 591, 592–598. [https://doi.org/10.1038/s41586-021-](https://doi.org/10.1038/s41586-021-03303-x)
- 920 03303-x
- Shcherbina, A.Y., Gregg, M.C., Alford, M.H. and Harcourt, R.R., 2009. Characterizing thermohaline intrusions in the North Pacific subtropical frontal zone. *Journal of Physical Oceanography*, 39(11), pp.2735-2756. <https://doi.org/10.1175/2010JPO4373.1>
- Tan, S. and Thurnherr, A. M., 2023: On the Global Decrease in the Deep and Abyssal Density Stratification Along the Spreading Pathways of Antarctic Bottom Water Since the 1990s, *Geophysical Research Letters*, 50, e2022GL102422. <https://doi.org/10.1029/2022GL102422>
- 925 Talley, L.D., Feely, R.A., Sloyan, B.M., Wanninkhof, R., Baringer, M.O., Bullister, J.L., Carlson, C.A., Doney, S.C., Fine, R.A., Firing, E. and Gruber, N., 2016. Changes in ocean heat, carbon content, and ventilation: A review of the first decade of GO-SHIP global repeat hydrography. *Annual review of marine science*, 8, pp.185-215. <https://doi.org/10.1146/annurev-marine-052915-100829>
- Vladoiu, A., Lien, R.C. and Kunze, E., 2024. Energy partition between submesoscale internal waves and quasigeostrophic vortical motion in the pycnocline. *Journal of Physical Oceanography*, 54(6), pp.1285-1307. <https://doi.org/10.1175/JPO-D-23-0090.1>
- 930 Wong, A. P. S., Wijffels, S. E., Riser, S. C., Pouliquen, S., Hosoda, S., Roemmich, D., Gilson, J., Johnson, G. C., Martini, K., Murphy, D. J., Scanderbeg, M., Bhaskar, T. V. S. U., Buck, J. J. H., Merceur, F., Carval, T., Maze, G., Cabanes, C., André, X., Poffa, N., Yashayaev, I., Barker, P. M., Guinehut, S., Belbéoch, M., Ignaszewski, M., Baringer, M. O., Schmid, C., Lyman, J. M., McTaggart, K. E., Purkey, S. G., Zilberman, N., Alkire, M. B., Swift, D., Owens, W. B., Jayne, S. R., Hersh, C., Robbins, P., West-Mack, D., Bahr, F., Yoshida, S., Sutton, P. J. H., Cancouët, R., Coatanoan, C., Dobbler, D., Juan, A. G., Gourrion, J., Kolodziejczyk, N., Bernard, V., Bourlès, B., Claustre, H.,
- 935 D'Ortenzio, F., Le Reste, S., Le Traon, P.-Y., Rannou, J.-P., Saout-Grit, C., Speich, S., Thierry, V., Verbrugge, N., Angel-Benavides, I. M., Klein, B., Notarstefano, G., Poulain, P.-M., Vélez-Belchi, P., Suga, T., Ando, K., Iwasaka, N., Kobayashi, T., Masuda, S., Oka, E., Sato, K., Nakamura, T., Sato, K., Takatsuki, Y., Yoshida, T., Cowley, R., Lovell, J. L., Oke, P. R., Van Wijk, E. M., Carse, F., Donnelly, M., Gould, W. J., Gowers, K., King, B. A., Loch, S. G., Mowat, M., Turton, J., Rama Rao, E. P., Ravichandran, M., Freeland, H. J., Gaboury, I., Gilbert, D., Greenan, B. J. W., Ouellet, M., Ross, T., Tran, A., Dong, M., Liu, Z., Xu, J., Kang, K., Jo, H., Kim, S.-D., and Park, H.-M.,
- 940 2020: Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations From a Global Array of Profiling Floats, *Front. Mar. Sci.*, 7, 700, <https://doi.org/10.3389/fmars.2020.00700>.



Time	n_a	n_s	n_o	$\sigma_\epsilon / \sigma_\epsilon^{ref}$
06/07 2002	22825 (5870)	1971 (38)	144 (5)	0.35 (0.3)
06/07 2004	25568 (6545)	2160 (63)	155 (3)	0.42 (0.51)
05/06 2006	24012 (6122)	1905 (54)	192 (17)	0.4 (0.44)
06/07 2008	24304 (5875)	409 (7)	49 (1)	0.06 (0.07)
06 2010	23336 (5668)	540 (6)	40 (4)	0.07 (0.08)
06/07 2012	22477 (5456)	527 (9)	25 (0)	0.08 (0.08)
05/06 2014	13133 (3186)	340 (8)	23 (2)	0.1 (0.11)
06/07 2016	23260 (5635)	435 (6)	27 (1)	0.07 (0.07)
06/07 2018	23422 (5678)	499 (7)	43 (2)	0.07 (0.08)
05/06 2021	23437 (5674)	468 (6)	18 (0)	0.08 (0.07)
06/07 2023	23503 (5685)	444 (13)	34 (4)	0.06 (0.06)

Table 1. Summary of instrument error calculations for the OVIDE sections used in this study. The numbers in the last four columns on the right, namely n_a (=number of “analysis layers”), n_s (= number of “no- N^2 layers”), n_o (= number of “overturning layers”), and $\sigma_\epsilon / \sigma_\epsilon^{ref}$ =(normalised measured of instrument error on buoyancy –see eq. (24)) refer to the calculation with $h = 25m$ while those in parenthesis correspond to that with $h = 100m$. The number of “analysis layers” was estimated based on calculations with $n_l = 5$ ($h = 25m$) and $n_l = 1$ ($h = 100m$).