On Using the Finescale Parameterization and Thorpe Scales to Estimate Turbulence from Glider Data

TARA HOWATT, STEPHANIE WATERMAN, AND TETJANA ROSS

Abstract: Turbulence plays a key role in many oceanic processes, but a shortage of turbulence observations impedes its exploration. Parameterizations of turbulence applied to readily available CTD data can be useful in expanding our understanding of the space–time variability of turbulence. Typically tested and applied to shipboard data, these parameterizations have not been rigorously tested on data collected by underwater gliders, which show potential to observe turbulence in conditions that ships cannot. Using data from a 10-day glider survey in a coastal shelf environment, we compare estimates of turbulent dissipation from the finescale parameterization and Thorpe scale method to those estimated from microstructure observations collected on the same glider platform. We find that the finescale parameterization captures the magnitude and statistical distribution of dissipation, but cannot resolve spatiotemporal features in this relatively shallow water depth. In contrast, the Thorpe scale method more successfully characterizes the spatiotemporal distribution of turbulence; however, the magnitude of dissipation is overestimated, largely due to limitations on the detectable density overturn size imposed by the typical glider CTD sampling frequency of 0.5 Hz and CTD noise. Despite these limitations, turbulence parameterizations provide a viable opportunity to use CTD data collected by the multitude of gliders sampling the ocean to develop greater insight into the space–time variability of ocean turbulence and the role of turbulence in oceanic processes.

Keywords: Ocean; Continental shelf/slope; In situ oceanic observations; Turbulence; Parameterization

1. Introduction

Turbulence plays an important role in many oceanic processes. It drives the irreversible mixing of water properties which results in the water mass transformations necessary for closing the meridional overturning circulation (e.g., Munk 1966; Munk and Wunsch 1998; Hughes et al. 2009). It efficiently transports heat vertically in the water column (e.g., Munk 1966; Osborn 1980). It delivers nutrients from depth to surface waters, facilitating primary production and the sequestration of carbon to depth (e.g., Crawford and Dewey 1989; Hamme and Emerson 2006; Kunze et al. 2006a). It can also play an important role in controlling the depth distribution of plankton in the water column (e.g., Rothschild and Osborn 1988; Visser et al. 2001; Fuchs and Gerbi 2016). Directly observing turbulence intensity (using microstructure observations) currently requires extensive sampling effort with specialized instrumentation, and consequently our understanding of the space–time variability of turbulence is limited. This is especially true for ocean conditions where ship operations are difficult (e.g., in winter, during storms). Underwater gliders are becoming a popular platform for obtaining sustained, well-resolved finescale measurements. Gliders represent an untapped opportunity to explore variations in turbulence intensity on a range of time and length scales, some of which are difficult to achieve with shipboard observations.

Turbulence intensity is typically quantified by estimating the turbulent kinetic energy dissipation rate (hereafter turbulent dissipation $\epsilon$). Estimates of turbulent dissipation are ideally made from microstructure observations of shear (or temperature) variance that relate to the kinematics of turbulence (or its effect on a scalar field) (e.g., Gregg 1999; Lueck et al. 2002). Microstructure observations are considered direct observations of turbulence intensity in that they measure variance on scales within the turbulent inertial subrange [typically $O(1$ cm–1 m)] (Lueck and Huang 1999). Using gliders as a platform for microstructure measurements is fairly novel: the first published use of glider-collected microstructure data is the proof-of-concept study by Wolk et al. (2009). Gliders are low-vibration platforms capable of delivering microstructure measurements of comparable quality to those obtained from traditional free-falling profilers (Fer et al. 2014) and are useful for observing turbulence in varied contexts (e.g., Fer et al. 2014; Peterson and Fer 2014; Palmer et al. 2015; Schultze et al. 2017; Scheielfe et al. 2018; Schultze et al. 2020). They yield a spatial and temporal coverage in oceanic microstructure fields that is often unattainable from ship-based profiling, especially in inclement weather, and the high density and large number of measurements allows for the calculation of robust statistical measures of turbulent metrics critical to interpreting microstructure measurements. However, the need for specialized instrumentation, which has high power usage, has led to a scarcity of glider-collected microstructure data. Yet there is currently a plentitude of glider-collected finescale CTD data available to which turbulence parameterizations, such as the finescale parameterization and Thorpe scale method, can be applied. Estimates of
turbulent dissipation inferred from these methods are consid-
ered indirect observations of turbulence intensity in that they
are derived from measurements on larger scales involved in
the transfer of energy down to dissipative scales without knowl-
edge of velocity shear or scalar gradients on turbulent micro-
scales [see, e.g., Polzin et al. (2014) and Dillon (1982) for
descriptions]. As such, to estimate \( \epsilon \) these methods require
making more assumptions compared to using microstructure
measurements, but they have the key advantage of exploiting
more readily available observations to “fill in the map” of
turbulence intensity estimates.

The finescale parameterization (FP) method infers turbulent
dissipation, \( \epsilon_{FP} \), based on theory describing the downslope
spectral energy cascade between internal wave scales [typically
\( O(10–100\, \text{m}) \)] and turbulent microscales where dissipation and
mixing occur [see Polzin et al. (2014) for a detailed review].
Over the past few decades, many studies have demonstrated
the utility of FP in mapping the space and time variability of
internal wave-driven dissipation and mixing rates throughout
the world’s oceans (e.g., D’Asaro and Morison 1992; Mauri-
zen et al. 2002; Naveira Garabato et al. 2004; Sloyan 2005; Kunze
et al. 2006b; Fer et al. 2010; Whalen et al. 2012; Waterhouse et al.
2014; Whalen et al. 2015; Kunze 2017). Johnston and Rudnick
(2015) have also applied FP to glider-collected CTD data, suc-
cessfully producing estimates of the turbulent mixing rate of the
same order of magnitude as previous tracer release and micro-
structure observations in their study area. In contexts where the
fundamental assumptions inherent in the method were met, esti-
mates of turbulent dissipation using FP have been shown to
agree well with microstructure-based direct estimates: early
parameterization validation studies by Gregg (1989), Polzin
et al. (1995), and Gregg et al. (2003) all found agreement within
a factor of 2 between the parameterization predictions and direct
estimates, while Whalen et al. (2015) found similar levels of
agreement (between a factor of 2 and 3 for 96% of their com-
parisons) using Argo float data obtained from diverse regions of
varying topography and oceanographic flow conditions.

The Thorpe scale method (TM) also indirectly infers tur-
bulent dissipation, \( \epsilon_{TM} \), in this case by assuming that the ob-
served sizes of density inversions in the water column represent
the scales of turbulent overturns at the largest vertical scale of
the spectral energy cascade (see, e.g., Thorpe 1977; Dillon
These density overturns are typically \( O(1–10\, \text{m}) \), and thus still
multiple orders of magnitude larger than the scales of dissi-
pation. Studies have demonstrated the utility of TM in char-
acterizing turbulent dissipation in a variety of contexts (e.g.,
Dillon 1982; Crawford 1986; Ferron et al. 1998; Alford et al.
2011; Waterhouse et al. 2014; Alford et al. 2015). Additionally,
Thorpe (2012) used numerical simulations to show that TM can
be successfully applied to data collected by gliders, albeit with
limited performance when turbulent features are elongated
vertically and/or when the glider’s trajectory is too horizontal.
The performance of TM depends on the stratification of the
water column and environmental conditions affecting instru-
ment motion (Park et al. 2014). The success of TM in predicting
\( \epsilon \) in agreement with microstructure estimates varies, likely
due to how well the fundamental assumptions inherent in the
method are met (Mater et al. 2015). While good agreement
between TM and direct estimates of \( \epsilon \) have been reported (e.g.,
Dillon 1982; Ferron et al. 1998), other validation studies report
an overprediction of \( \epsilon_{TM} \) by one to two orders of magnitude
relative to estimates of \( \epsilon \) derived from microstructure data
(e.g., Klymak et al. 2008; Frants et al. 2013; Mater et al. 2015).

Some aspects of glider sampling present challenges to the
application of FP and TM to glider-collected CTD data. Most
notable are a profiling trajectory which is typically inclined at
angles of 20°–30° to the horizontal, and a low CTD sampling
frequency typically selected in order to conserve battery life
and extend glider missions. Thus, understanding the utility of
and best practices for implementing turbulence parameteri-
zations with glider-collected data is necessary and important in
order to tap into the opportunity presented by gliders to ex-
pose the variability of turbulence on a range of scales. While
insights into both the potential and limitations of these pa-
rameterization methods for glider data have been illustrated
using in situ glider observations in the case of FP (Johnston and
Rudnick 2015) and numerical simulations in the case of TM
(Thorpe 2012), directly comparing their inferences of \( \epsilon \) with
direct estimates derived from coincident and collocated glider-
collected microstructure data has yet to be attempted.

In this study we apply and evaluate the performance of the
strain-based finescale parameterization and Thorpe scale
method using CTD and microstructure data collected simul-
taneously on the same glider platform. We perform our vali-
dation study using data from a 10-day glider mission in a
coastal shelf environment that encompasses a range of turbu-
rence and stratification conditions. Section 2 discusses the
field campaign, microstructure processing methods, and
our implementations of the FP and TM to the glider CTD data.
Section 3 compares the statistical and spatiotemporal distrib-
utions of \( \epsilon \) inferred via the parameterizations to direct mi-
crostructure estimates. Section 4 summarizes the strengths and
weaknesses of each parameterization in this application, dis-
cusses how their assumptions combined with the glider sam-
ping capabilities underpin their limitations, and makes novel
recommendations for their implementation to glider-sampled
CTD data. Given the increasing availability of glider-collected
CTD datasets, recommendations on the application of turbu-
rence parameterizations to glider data are expected to have
widespread utility.

2. Methods

a. Field campaign

We deployed a Slocum underwater glider for 10 days be-
tween 2 and 12 August 2017 in Roseway Basin off the east coast
of Canada (Fig. 1a). The basin is approximately 30 km wide
and 150 m deep in the center. Roseway Basin is an energetic
coastal shelf environment forced predominantly by baroclinic
pressure gradients and tidal rectification (Han et al. 1997;
Hannah et al. 2001). Stratification is weaker on the south-
western portion of the continental shelf, where Roseway Basin
is located, due to fewer connections bringing dense water from
the deep ocean, being farther removed from the Gulf of Saint
Lawrence freshwater source, and having stronger tidal energy
Roseway Basin is a right whale critical habitat (Brown et al. 2009) where tidal rectification likely aggregates right whale prey (diapausing copepods) on the southeastern basin margin (Baumgartner et al. 2003; Davies et al. 2013). Motivated by interest in the factors that make Roseway Basin a preferred whale habitat, the glider traveled across the basin to its southeastern margin to make multiple repeat transects across its southeastern slope before completing a second full basin transect on its return leg (Fig. 1b).

The glider was equipped with a Sea-Bird G-1451 Slocum glider payload pumped CTD measuring in situ conductivity, temperature, and pressure at a sampling frequency of 0.5 Hz. It also carried an externally mounted Rockland Scientific microstructure turbulence package (MicroRider) carrying two airfoil velocity shear (SPM-38) and two fast-response temperature (FP07) probes measuring orthogonal components of velocity shear and temperature at a sampling frequency of 512 Hz. Initially the glider sampled continuously; partway through the mission (starting at profile 481) the MicroRider sampling was reduced to only downcast sampling in order to conserve battery charge and prolong the mission. In total, the glider collected 1046 CTD and 721 microstructure profiles. For these profiles, the mean (± one standard deviation) pitch was $-24° \pm 3°$ and $25° \pm 5°$, the angle of attack was $-2.8° \pm 0.4°$ and $2.8° \pm 0.8°$, the glide angle from the horizontal was $-27° \pm 3°$ and $28° \pm 5°$, and the glider speed through the water was $0.31 \pm 0.05$ m s$^{-1}$ and $0.40 \pm 0.10$ m s$^{-1}$ for downcasts and upcasts, respectively. The angle of attack was estimated using the hydrodynamic flight model developed by Merckelbach et al. (2010) and the measured glider pitch; the glide angle was computed from the estimated angle of attack and measured glider pitch. Glider speed was estimated from the glide angle and measured rate of change of pressure following Schultze et al. (2017). These flight characteristics are not remarkable, similar to ones previously reported (e.g., Fer et al. 2014; Schultze et al. 2017; Scheifele et al. 2018). The CTD downcast profiles were separated by a mean (± one standard deviation) horizontal spacing of $0.5 \pm 0.3$ km; and separated in time by $0.4 \pm 0.1$ h. The spacing of the profiles is primarily a function of the water depth, which ranged from 52 to 166 m.

b. Microstructure data processing

We estimated the turbulent kinetic energy dissipation rate from the shear microstructure data, $\varepsilon_{\mu U}$, using spectral analysis following Lueck (2016) and Scheifele et al. (2018). We assumed isotropic turbulence and calculated $\varepsilon_{\mu U}$ using

$$\varepsilon_{\mu U} = \frac{15}{2} \nu \left( \frac{\partial u'}{\partial x} \right)^2 = \frac{15}{2} \nu \int_0^\infty \Phi(k) \, dk,$$

where $\nu$ is the kinematic molecular viscosity, $\partial u'/\partial x$ is the along-path gradient of a perpendicular turbulent velocity component, and $\Phi(k)$ is the spectrum of observed shear variance per unit wavenumber $k$; $\Phi(k)$ was numerically integrated over $[k_1, k_u]$, where $k_1$ is the first nonzero wavenumber set by the spectral subsegment length (4 s, see below), and $k_u$ is an estimate of the wavenumber at which electronic noise starts to dominate the measurement (see Scheifele et al. 2018). We accounted for unresolved variance outside this integration range by using a correction procedure based on the nondimensionalized Nasmyth spectrum (see Scheifele et al. 2018).

We derived a second independent direct estimate of $\varepsilon$ from the temperature microstructure data using methodology following Scheifele et al. (2018). The turbulent kinetic energy dissipation rate from microstructure temperature-gradient variance, $\varepsilon_{\mu T}$, was calculated by assuming that the turbulence was isotropic, and that the temperature-gradient spectra followed the shape of a theoretical Batchelor spectrum (Batchelor 1959), such that

$$\varepsilon_{\mu T} = \nu \kappa_T \left( 2\pi k_B \right)^4,$$

where $\kappa_T$ is the molecular diffusivity of heat and $k_B$ is the Batchelor wavenumber. The Batchelor spectrum is a function of both $k_B$ and $\chi$, the dissipation of temperature variance (Osborn and Cox 1972); the latter we estimated from
the observed and Batchelor temperature-gradient spectra according to
\[ \chi = 6\kappa \int_0^{k_{IT}} \frac{\Phi(k) \Psi(k) dk}{\Psi(k) dk} + \int_{k_{IT}}^{k_B} \Psi(k) dk + \int_{k_B}^\infty \Psi(k) dk. \] (3)

Here \( \partial T/\partial x \) is the along-path gradient in the turbulent temperature fluctuations, and \( \Psi(k) \) and \( \Psi_B(k) \) are the spectrum of observed and Batchelor temperature-gradient variance, respectively. The calculation of \( \chi \) simplified the fitting of the Batchelor spectrum: \( k_B \) was found via an iterative procedure to choose the best fit from a family of Batchelor curves constructed using constant \( \chi \) but variable \( k_B \) over the range of wavenumbers \([k_{IT}, k_{a/T}]\) (see Ruddick et al. 2000). We chose the upper integration limit \( k_{IT} \) to avoid contamination due to noise at high \( k \) (it was defined as the intersection between the observed and 2 times a probe-specific empirical estimate of the noise spectrum). The lower limit \( k_{IT} \) reflects the top of the inertial convective subrange or the wavenumber bound of confidence in an iterative fitting procedure (Luketina and Imberger 2001; Steinbuck et al. 2009; Scheifele et al. 2018).

We prepared the microstructure shear (temperature) data for use in Eq. (1) [Eq. (3)] by transforming the shear (temperature) time series into power spectra, \( \Phi(f) \) [\( \Psi(f) \)] where \( f \) is the frequency, using a fast Fourier transform (FFT) of detrended and cosine-windowed segments of shear (temperature) data. FFTs of 19 half-overlapping 4 s subsegments of data were averaged to form one observed shear (temperature) power spectrum for each 40 s segment of data. Shear spectra were corrected by removing coherent acceleration signals measured by the MicroRider (Goodman et al. 2006). Temperature spectra were corrected using a transfer function with a response time of 3.3 ms to account for a thermal delay of the thermistor at high frequencies (Sommer et al. 2013; Bluteau et al. 2017). We converted the shear (temperature) frequency spectra into wavenumber spectra assuming Taylor’s frozen turbulence hypothesis (i.e., \( \Phi(k) = U\Phi(f) \) and \( k = fU \)), where \( U \), the mean along-path glider speed through the water over the 40 s interval, was estimated from the glider’s measured rate of change of pressure and an estimate of the glide angle (Merckelbach et al. 2010). Finally, we converted the temperature spectra to temperature-gradient spectra by multiplying them by the derivative operator, \((2\pi k)^2\).

We applied a number of quality control criteria, largely following the methods detailed in Scheifele et al. (2018). Individual turbulent dissipation estimates were flagged and removed if they satisfied one or more of the following criteria: 1) a failed visual inspection test; this removed data after thermistor 2 was damaged (profile 590); 2) the magnitude of the glider’s acceleration \( |dU/dt| \) was above the tenth percentile \((|dU/dt| > 1.9 \times 10^{-3} \text{ m s}^{-2}); \) this criterion identified where the glider’s speed changed too much over the span of one \( \epsilon \) estimate; 3) the glider was within 1 m of an inflection point; this identified where the estimates of the glider’s angle of attack and speed were uncertain, and where the data were prone to contamination from mechanical vibrations; 4) estimates from the two shear (temperature) probes differed by greater than a factor of 10; in this case the larger estimate was removed to avoid the inclusion of contaminated spectra; 5) the ratio between the glider’s velocity and an estimate of the turbulent velocity scale, \( U(\epsilon fN)^{1/2} \), was less than 5; this identified violations of Taylor’s frozen turbulence hypothesis (Fer et al. 2014). In addition, estimates of \( \epsilon_{\mu U} \) were removed if 6) the mean absolute deviation (MAD) criterion (see Ruddick et al. 2000) was greater than 0.35; this removed shear spectra that deviated significantly from the Nasmyth spectrum (Nasmyth 1970). Estimates of \( \epsilon_{\mu T} \) were also removed if 7) the MAD criterion was greater than 2; this removed temperature-gradient spectra that deviated significantly from the Batchelor spectrum (Batchelor 1959); 8) the sum of the temperature-gradient spectrum; and 9) there were fewer than \( n = 16 \) distinct wavenumbers available in the closed interval \([k_{IT}, k_{a/T}]\); this removed spectra with an insufficient number of spectral points to ensure an appropriate Batchelor spectrum fit. Percentages of data removed by these quality control criteria are listed in Table 1. We removed a total of 25% (31%) of \( \epsilon_{\mu U} / \epsilon_{\mu T} \) estimates resulting in a total of 20 042 (20 059) remaining estimates. The quality-controlled shear and temperature-gradient spectra typically show reasonable agreement with the Nasmyth and Batchelor spectra, respectively, for all decades of \( \epsilon \) estimates (Figs. 2a,b).

COMPARISON OF \( \epsilon \) BETWEEN MICROSTRUCTURE METHODS

Given our goal of evaluating the performance of FP and TM, we first compared estimates of \( \epsilon_{\mu U} \) and \( \epsilon_{\mu T} \) to establish confidence in a ground truth estimate of \( \epsilon \) to use as the basis for comparison.

Independent estimates of \( \epsilon \) from shear and temperature microstructure data were in good agreement with some limitations. The geometric means of \( \epsilon_{\mu U} \) and \( \epsilon_{\mu T} \) agreed to within the 95% confidence interval, and \( \epsilon_{\mu U} \) and \( \epsilon_{\mu T} \) distributions had similar and overlapping interquartile ranges (Table 2). The skewness of the \( \epsilon \) distributions in logarithmic space was positive for both metrics (Table 2, Fig. 3). At the high end of the distributions, there were fewer \( \epsilon_{\mu T} \) relative to \( \epsilon_{\mu U} \) (illustrated by the taller and longer tail of the \( \epsilon_{\mu U} \) histogram at \( \epsilon > 10^{-7} \text{ W kg}^{-1} \), Fig. 3). This is likely a consequence of the limited response time of the microstructure thermistor (Sommer et al. 2013; Bluteau et al. 2017). At the low end of the distributions, \( \epsilon_{\mu T} \) extended to lower \( \epsilon \) magnitudes (illustrated by the longer tail of the \( \epsilon_{\mu T} \) histogram at \( \epsilon < 10^{-10} \text{ W kg}^{-1} \)). This is likely a consequence of the higher noise floor of the microstructure shear probes (Scheifele et al. 2018). This also manifests as a smaller positive skewness for \( \epsilon_{\mu T} \) relative to \( \epsilon_{\mu U} \) (Table 2).

While these differences at the extreme ends of the \( \epsilon \) distribution may be useful to keep in mind in the comparisons to follow, the good agreement for the bulk of midrange \( \epsilon \) estimates gave us confidence in the accuracy of both \( \epsilon_{\mu U} \) and \( \epsilon_{\mu T} \). To streamline our evaluation of FP and TM performance,
in what follows we visualize comparisons of parameterized estimates of $c$ to direct microstructure-based estimates using $c_{\mu U}$ only.

c. Finescale parameterization

The strain-based finescale parameterization for internal wave-driven dissipation (Polzin et al. 2014) was applied according to

$$c_{FP} = \frac{N^2}{\langle \xi_i^2 \rangle} \frac{\langle \xi_i^2 \rangle_G M}{h(R_w) L(f, N)}.$$  \hspace{1cm} (4)

Here $\xi_i$ represents the observed vertical strain, presumed to be due to internal wave displacements of neutral density surfaces. It is given by $\xi_i = (N^2 - N_{ref}^2)/N_{ref}^2$, where $N_{ref}$ represents the background stratification (i.e., lacking wave-induced perturbations). Here $\langle \xi_i^2 \rangle$ is the observed strain variance determined by integrating the spectral energy density of $\xi_i$ over specified vertical wavenumbers, and $\langle N \rangle$ is the average stratification over the vertical segment used in the spectral calculation. This formulation takes advantage of normalizing the observed strain variance to that in the Garrett–Munk (GM) model internal wave spectrum (Garrett and Munk 1979). $\langle \xi_i^2 \rangle_G M$, that can be characterized by an analytically known dissipation rate which, at 30° latitude and in a background reference stratification of $N_{ref} = 5.24 \times 10^{-3}$ s$^{-1}$, is given by $c_{\mu} = 6.73 \times 10^{-10}$ W kg$^{-1}$ (e.g., as used in Gregg 1989). The correction factor $h(R_w)$ adjusts the prediction to account for the dominant frequency content of the wave field, estimated via the shear-to-strain ratio $R_w$ [see Kunze et al. (2006b) for a description]. The function $L(f, N)$ corrects for a latitudinal dependence predicted by theory (see Henyey et al. 1986).

To implement FP here, we computed vertical profiles of strain by first smoothing each raw neutral density (Jackett and McDougall 1997) profile to remove spikes and small-scale (<2 m) variability due to CTD noise (see Fig. 4a, black line, for a representative example). We used these smoothed density profiles to compute corresponding profiles of $N$ (Fig. 4b, black line). To determine the background $N_{ref}$ profile, we first computed a background density profile $\rho_{ref}$ by time averaging vertically binned density profiles using a time window of 12 h and a vertical bin size of 10 m, then smoothing the resulting profile with a 5 m low-pass filter (Fig. 4a, green line); $N_{ref}$ (Fig. 4b, green line) was computed from $\rho_{ref}$. Estimates of $c_{FP}$ were not sensitive to reasonable choices of time and vertical averaging scales; additionally, the horizontal distance traveled by the glider in 12 h (~10 km) was less than the deformation radius (~20 km).

We computed vertical wavenumber spectra of strain via application of an FFT to 75% overlapping 64 m segment lengths, which were first detrended, Hann windowed, and normalized. This segment length was chosen in order to maximize the number of $c$ estimates in each profile, while ensuring internal wave scales could be resolved [required that the FFT segment length contained at least 2.5 wavelengths of the

<table>
<thead>
<tr>
<th>$\epsilon$ decay</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Total removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{\mu U}$</td>
<td>9%</td>
<td>6%</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
<td>18%</td>
<td>0%</td>
<td>2%</td>
<td>1%</td>
<td>25%</td>
</tr>
<tr>
<td>$c_{\mu T}$</td>
<td>13%</td>
<td>10%</td>
<td>6%</td>
<td>7%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>31%</td>
</tr>
</tbody>
</table>

Table 1. Percentage of data flagged by quality control criteria described in section 2a. An estimate of $c$ may be flagged by more than one criterion so percentages do not add up to the total data removed.

Fig. 2. Spectral energy density plots for (a) shear, (b) temperature gradient, and (c) strain averaged for each $c$ decade (solid lines, shading lighter toward higher-$c$ decade) with the theoretical (a) Nasmyth, (b) Batchelor, and (c) Garrett–Munk spectra (dashed lines). The temperature gradient spectra in (b) are normalized by $\chi$. The notation $c$ indicates $c$ decades in units of W kg$^{-1}$, and $n$ indicates the number of spectra in each decade. The $c$ decades are defined as the average of all spectra with $c$ values within the bounding $c$ values (i.e., an $c$ decade of $10^{-8}$ W kg$^{-1}$ contains $c$ values in the range from $10^{-8.5}$ to $10^{-7.5}$ W kg$^{-1}$).
Table 2. Statistics of the \(\epsilon\) estimate distribution for microstructure shear, microstructure temperature, finescale parameterization, and Thorpe scale methods. The bootstrapped 95% confidence intervals are indicated in the square brackets below the value. Skewness metrics were calculated in logarithmic space so that the metrics are representative of the histograms in Figs. 3 and 6.

<table>
<thead>
<tr>
<th></th>
<th>Microstructure (\epsilon_U)</th>
<th>Microstructure (\epsilon_T)</th>
<th>Finescale parameterization (\epsilon_{FP})</th>
<th>Finescale parameterization (\epsilon_{FP}^{TM})</th>
<th>Thorpe scale (\epsilon_{TM})</th>
<th>Thorpe scale (\epsilon_{TM}^{TM})</th>
<th>Thorpe scale (\epsilon_{TM}^{PF})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometric mean ((\times 10^{-9} \text{ W kg}^{-1}))</td>
<td>3.7 (3.7, 3.9)</td>
<td>3.8 (3.7, 3.9)</td>
<td>5.1 (4.1, 6.0)</td>
<td>2.1 (1.8, 2.6)</td>
<td>2.7 (2.2, 3.1)</td>
<td>96 (90, 104)</td>
<td>4.1 (3.8, 4.5)</td>
</tr>
<tr>
<td>Arithmetic mean ((\times 10^{-9} \text{ W kg}^{-1}))</td>
<td>29 (27, 33)</td>
<td>20 (18, 25)</td>
<td>14 (12, 18)</td>
<td>5.5 (3.8, 8.9)</td>
<td>5.1 (4.5, 6.1)</td>
<td>705 (513, 1312)</td>
<td>37 (32, 46)</td>
</tr>
<tr>
<td>Interquartile range ((\times 10^{-9} \text{ W kg}^{-1}))</td>
<td>1.0–14</td>
<td>1.1–14</td>
<td>1.8–15</td>
<td>0.9–5.1</td>
<td>1.4–5.4</td>
<td>25–377</td>
<td>1.0–17</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.7 (1.6, 1.9)</td>
<td>0.4 (0.4, 0.5)</td>
<td>0.7 (0.4, 1.4)</td>
<td>5.2 (2.1, 17.1)</td>
<td>1.8 (0.9, 3.9)</td>
<td>1.4 (1.2, 1.8)</td>
<td>1.9 (1.6, 2.3)</td>
</tr>
<tr>
<td>Number of estimates</td>
<td>19123</td>
<td>197</td>
<td>197</td>
<td>2848</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

largest wavelengths of interest, as recommended by Emery and Thomson (1997). We integrated the strain variance over vertical wavenumbers corresponding to wavelengths of 10–25 m, chosen to be in a range that was well resolved given our segment length choice and over which the observed strain spectra were approximately flat. We assumed \(R_m = 7\), the global average (Kunze et al. 2006b), chosen in the absence of internal wave-scale shear data. Although average strain spectra do slope downward for lower \(\epsilon_{FP}\) decades, strain spectra typically show a similar shape to the GM spectrum (Fig. 2c), and were judged within an acceptable spectral density range when compared to Kunze et al. (2006b, see their Fig. 2).

In our implementation we made a number of choices to improve the robustness of \(\epsilon_{FP}\). These included 1) only using data from below the pycnocline, as the vertical length scales of internal waves and the pycnocline were similar, and the method cannot differentiate between them; by excluding these spectra, we have greater confidence that the observed variance at the wavelengths considered is an internal wave signal (Kunze et al. 2006b; Kunze 2017); and 2) only using segments containing the full 64 m of data (i.e., not including data within one segment length of the seafloor), so to robustly resolve the wavelengths of interest. Additionally, 3) we included a quality control criteria that removed spectra with MAD > 1.7; this removed strain spectra that deviated significantly from the corresponding normalized GM spectrum (this removed 3% of our \(\epsilon_{FP}\)).

d. Thorpe scale

The Thorpe scale estimate of turbulent dissipation, \(\epsilon_{TM}\), was calculated using vertical displacements from a stably sorted neutral density profile and assuming that the Ozmidov scale \(L_O\), the largest vertical scale on which turbulent eddies are isotropic, was approximately equal to the Thorpe scale \(L_T\). Under this assumption, following Thorpe (1977) and Dillon (1982),

\[
\epsilon_{TM} = (C_0 L_T)^2 N^3, \tag{5}
\]

where the constant \(C_0\) accounts for differences between \(L_O\) and \(L_T\) (we used \(C_0 = 0.8\); Dillon 1982) and \(N\) is the relevant buoyancy frequency. Individual overturns (see Fig. 5 for a representative example) were identified as the region of the profile where the cumulative sum of the vertical displacements of neutral density were near zero (i.e., \(\approx 1 \times 10^{-3}\) m). The length of this region is the overturn length, \(L\), which was used in setting the length over which to calculate \(L_T\). The value of \(L_T\) was calculated as the root-mean-square (rms) of the vertical displacements of the stably sorted density profiles \(\zeta_{sort}\) (Fig. 5, dashed gray line) compared to the measured profiles \(\zeta\) (Fig. 5, yellow line) i.e., \(L_T = \text{rms}(dz)\) with \(dz = \zeta - \zeta_{sort}\). The relevant \(N\) in Eq. (5) is specific to \(L\), calculated using the mean density and total density change over \(L\).

Quality control criteria were applied to improve the robustness of \(\epsilon_{TM}\). For each overturn, we required 1) \(L\) to be greater than 2 m, the expected minimum resolvable overturn size given the CTD sampling frequency (0.5 Hz), the glider vertical fall rate (\(-0.15 \text{ m s}^{-1}\)), and the smoothing applied to the

![Fig. 3. Distributions of \(\epsilon\) estimates derived from microstructure shear (purple) and temperature (blue) observations. Histograms include only points in space–time that have an \(\epsilon\) estimate for both methods.](image-url)
raw density data to eliminate noise in the CTD measurements (see section 2c); 2) $L$ to be greater than the expected minimum resolvable overturn size given an estimate of CTD noise levels, $L_{\text{noise}} = \sqrt{(2g\rho_{\text{noise}})(\rho_0 N_{\text{smooth}}^2)}$ m where $\rho_{\text{noise}} = 1 \times 10^{-4}$ kg m$^{-3}$ based on the CTD precision specifications of $1 \times 10^{-4}$ C and $1 \times 10^{-4}$ mS cm$^{-1}$, $\rho_o = 1027$ kg m$^{-3}$, and $N_{\text{smooth}}^2$ comes from a 10 m running mean; 3) the density change over $L$ to be greater than twice the noise level, $\Delta \rho > 2\rho_{\text{noise}}$; 4) the maximum density displacement magnitude within the overturn, $|\rho - \rho_{\text{ort}}|$, to be greater than twice the noise level, $\max(|\rho - \rho_{\text{ort}}|) > 2\rho_{\text{noise}}$; and 5) the overturn ratio, $R_0 = \min(L^+/L, L^-/L)$, to be greater than 0.2; here $L^+$ and $L^-$ indicate the vertical extent of the positive and negative Thorpe displacements, respectively. This latter criterion avoids including nonsymmetric overturns (Gargett and Garner 2008). We note that this approach to quality control for the Thorpe scale parameterization analysis, we first smooth the neutral density profile to remove noise with vertical scales $< 2$ m, thereby increasing the reliance on criterion 1 (based on minimum resolvable overturn size) in the quality control procedure. A consequence of this choice is that the CTD noise level quality control steps (criteria 2–4) are applied to an approximately electronic noise-free record. This justifies our choice of the CTD precision for the $\rho_{\text{noise}}$ level. Smoothing the raw density profiles before applying TM is not necessary, in which case, using both a higher $\rho_{\text{noise}}$ and a run-length quality control criteria to remove small overturns is preferred (Galbraith and Kelley 1996). Percentages of data removed by all quality control criteria are listed in Table 3.

3. Results

a. Statistical comparison of $\epsilon$

To evaluate the performance of each parameterization, we first compare estimates of $\epsilon$ from the finescale parameterization analysis...
(FP) and Thorpe scale method (TM) to direct estimates from the shear and temperature microstructure data ($\epsilon_{\mu U}$ and $\epsilon_{\mu T}$) in terms of their statistical distributions (Table 2). We show comparisons to the shear-based microstructure estimates only (Figs. 6a,c); there is strong agreement between the microstructure estimates (section 2b) and conclusions are unchanged when temperature-based microstructure estimates are considered. The microstructure measurements resolve $\epsilon$ on much smaller spatiotemporal scales than either parameterization; to compare them we arithmetically average $\epsilon_{\mu U}$ and $\epsilon_{\mu T}$ on the space–time grid of the parameterized estimates, specifically on the 32 m resolved vertical grid of FP and over the vertical extent of each Thorpe overturn. We further include averaged $\epsilon_{\mu U}$ and $\epsilon_{\mu T}$ values only at the times/locations when/where there is an FP or TM estimate. Consequently, the distributions of $\epsilon_{\mu U}$ and $\epsilon_{\mu T}$ are different for the finescale parameterization versus Thorpe scale method comparisons, and further different from the fully resolved microstructure dataset. Differences relative to the latter reflect a combined influence of averaging on coarser spatiotemporal scales and subsetting the microstructure-derived estimates to select locations in space and time.

The finescale parameterization is relatively successful at capturing the average magnitude of the microstructure-derived $\epsilon$ estimates (Table 2, Fig. 6a). The arithmetic mean value of the FP distribution versus that of the corresponding and equivalently averaged set of $\epsilon_{\mu U}$, hereafter denoted $\epsilon_{\mu U}^{FP}$, differ by a factor of 2.5, while the geometric mean values differ by a factor of 2.4 (in both cases FP overpredicts on average). The interquartile ranges of the two $\epsilon$ distributions overlap. Collectively, these metrics suggest that FP can provide a useful measure of both the mean value and central tendency of the $\epsilon_{\mu U}^{FP}$ distribution. When compared to the microstructure estimates, $\epsilon_{FP}$ underestimates the arithmetic mean by approximately a factor of 2, primarily a result of the absence of $\epsilon_{FP}$ when/where $\epsilon_{\mu U}$ tends to be high (section 3c), and possibly the effect of averaging inherent in FP which removes high-end outliers from the distribution. Despite this, the FP distribution in logarithmic space does indicate positive skewness, consistent with the microstructure estimates, albeit to a lesser degree than both the fully resolved and subsetted-and-averaged distributions. A similar characterization emerges from the comparison of $\epsilon_{FP}$ to $\epsilon_{\mu T}$ (see Table 2). Overall, FP provides a useful estimate of the average magnitude and central tendency of the microstructure $\epsilon$ distributions to within a factor of 2–3, with a tendency to overpredict the magnitude and underpredict the skewness of the distribution when compared to direct estimates at corresponding times/locations and on equivalent scales.

In contrast, the Thorpe scale method systematically overestimates the magnitude of the microstructure-derived $\epsilon$ estimates (Table 2, Fig. 6c). The arithmetic mean values of the TM versus corresponding equivalently averaged $\epsilon_{\mu U}$ (hereafter denoted $\epsilon_{\mu U}^{TM}$) distributions differ by a factor of 19, the geometric mean values differ by a factor of 23, and the interquartile ranges do not overlap. The degree of overprediction is similar when the TM are compared to the fully resolved $\epsilon_{\mu U}$, and a similar characterization emerges from the comparison of $\epsilon_{TM}$ to $\epsilon_{\mu T}$ (see Table 2). The TM distribution does indicate similar positive skewness to that of both the fully resolved and subsetted $\epsilon_{\mu U}$ distributions (TM underpredicts the skewness of the $\epsilon_{\mu U}$ distribution by 18% and the $\epsilon_{\mu U}^{TM}$ distribution by 26%). Thus, overall, TM performs relatively well in capturing the positive skewness of the distributions of direct $\epsilon$ estimates, but overestimates the magnitude of $\epsilon$ by about a factor of 20 on average.

### b. Point-to-point comparison of $\epsilon$

We next consider the ability of the parameterizations to capture variability in $\epsilon$ by comparing parameterized estimates to their corresponding direct estimate at the same time and place (Figs. 6b,d). Thus, we compare $\epsilon_{FP}$ and $\epsilon_{TM}$ to their

---

**Table 3.** Percentage of data flagged by quality control criteria described in section 2d: 1) $L > 2$ m; 2) $L > L_{noise} = (2gq_{noise})(pP_{maxcoh})$; 3) $\Delta P > 2p_{noise}$; 4) $\max(|\Delta \rho - \rho_{sort}|) > 2p_{noise}$; and 5) $R_0 = \min(L'/L, L'/L) > 0.2$. An estimate of $\epsilon$ may be flagged by more than one criteria so percentages do not add up to the total data removed.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_{TM}$</td>
<td>56%</td>
<td>14%</td>
<td>10%</td>
<td>16%</td>
<td>0.2%</td>
<td>64%</td>
</tr>
</tbody>
</table>
corresponding equivalently averaged microstructure-derived estimates as described in section 3a. These point-to-point comparisons are a more stringent test of parameterization performance than the statistical comparisons hitherto discussed.

Point-to-point comparisons of $\epsilon_{FP}$ and $\epsilon_{MU}^{FP}$ confirm the relative success of FP in capturing the general magnitude of the direct equivalently averaged $\epsilon$ estimates, but also reveal a shortcoming of FP in representing the variability of those estimates (Fig. 6b). Points cluster around the one-to-one line, and the general agreement between the two estimates in terms of magnitude is further confirmed by the geometric mean value of the set of $\epsilon_{FP}/\epsilon_{MU}^{FP}$ ratios: 2.4. However, a negligible slope of the line of best fit through these points indicates negligible sensitivity of $\epsilon_{FP}$ to variations in $\epsilon_{MU}^{FP}$, and thus poor performance in distinguishing the times and locations when and where $\epsilon_{MU}^{FP}$ is relatively high versus when and where $\epsilon_{mu}^{FP}$ is relatively low. Additionally, FP yields $\epsilon$ estimates with a narrower range compared to the range of $\epsilon_{TM}$: three versus six orders of magnitude, respectively. These shortcomings in resolving variability are further compounded by the fact that our application of FP is restricted to a limited portion of the full water column and as a consequence fails to characterize regions with high $\epsilon_{MU}$ and $\epsilon_{MU}$ and above the pycnocline and close to the bottom; this limitation further restricts the ability of FP to represent the range and variability of dissipation resolved by the microstructure measurements. Overall, although FP captures the general magnitude of $\epsilon$ to within a factor of 2–3, it is limited in its ability to capture the full-range $\epsilon$ variability present in the environment, largely because it cannot be applied in certain high dissipation parts of the water column.

In contrast, point-to-point comparisons of $\epsilon_{TM}$ and $\epsilon_{TM}^{TM}$ estimates show that TM shows a degree of sensitivity to variability in direct estimates of $\epsilon$, despite its tendency to overpredict $\epsilon$ magnitude (Fig. 6d). The line of best fit indicates covariance between $\epsilon_{MU}^{TM}$ and $\epsilon_{TM}$ with a slope of $0.28 \pm 0.03$ to the 95% confidence level. However, the tendency of TM to overpredict direct estimates of $\epsilon$ is further illustrated by the
tendency for points in this comparison to cluster above the one-to-one line, and further by the geometric mean value of all $e_{TM}/e_{U}$ ratios 22. Thus, overall, TM shows relative utility in mapping variability in direct $e_{U}$ estimates, despite its tendency to grossly overestimate $e$ magnitude.

c. Spatiotemporal comparison of $e$

We finally consider the ability of the parameterizations to capture the spatiotemporal distribution of $e$ by comparing space–time views of $e$ delivered by each method. Figure 7 highlights the varied successes and limitations of the parameterizations in mapping space–time variability, resolving spatiotemporal features and, importantly, accessing the full extent of space–time described by the microstructure observations.

The section views of the fully resolved shear microstructure (Fig. 7a) and TM estimates (Fig. 7c) show that $e_{U}$ and $e_{TM}$ exhibit many similar features in their space–time distributions. Broadly, both show enhanced $e$ in the upper 50 m of the water column, as well as near the seafloor over shallow bathymetry on the basin margin. In addition, a number of individual patches of elevated dissipation are captured by both methods, for example, as seen near the bottom (yellow patches in Fig. 7c) around profile number 70, 550, 830, and after profile number 980. However, this comparison also reveals limitations of the Thorpe scale method in capturing the full picture of space–time variability. TM delivers an estimate of $e$ only when an overturn is detected; at times/places when/where there is no $e_{TM}$ (the white space in Fig. 7c), the dissipation rate is assumed to be below the detectable limit for TM. This is a function not only of the local dissipation rate, but also the stratification (see section 4). We find that few Thorpe overturns are identified in the highly stratified waters of the upper water column (i.e., around and above ~50 m depth), while overturns are found only intermittently in the midwater column over deeper regions (where $e$ is small). Additionally, we see a systematic tendency for TM to overpredict the microstructure-derived dissipation rate in the pycnocline (around 50 m depth).

In contrast, features in the space–time distribution of $e$ are difficult to identify for FP (Fig. 7b). This is due to the coarse space–time resolution of $e_{FP}$. This resolution is imposed by the requirement to consider profile segments that resolve internal wave vertical scales, and is also due to the fact that, by its very nature, $e_{FP}$ presents a view of $e$ that is averaged on relatively large internal wave time and length scales. This averaging smears out the smaller-scale spatiotemporal features that in part define the space–time distribution of the direct estimates.

![Figure 7. Space–time section plot of $e$ from (a) microstructure shear, (b) FP, and (c) TM as a function of depth and profile number. The gray represents the bathymetry as the glider made multiple transects over the southwest margin of the basin.](image-url)
Most critically, it is due to the fact that the FP can only be applied in a limited region of the full water column (i.e., below the pycnocline and not within less than one vertical profile segment length from the bottom). Owing to the shallow water depth and limited water column extent between the pycnocline and seafloor at this location, this restriction severely limits the number of $\varepsilon_{\text{FP}}$ (there are typically only one to three estimates of $\varepsilon_{\text{FP}}$ per profile). As discussed above, it also prevents FP from resolving the most distinct contrasts in $\varepsilon_{\text{MU}}$ magnitude, i.e., strong dissipation around and above $\sim 50 \text{ m}$ depth and near the bottom relative to weak dissipation in the interior of the water column.

4. Discussion

In this study, we applied the strain-based FP and TM to glider-collected CTD data from a coastal shelf environment. These parameterizations were evaluated by comparing their output to estimates of $\varepsilon$ from microstructure data collected concurrently on the same glider platform. FP provides an accurate estimate of the average magnitude of turbulent dissipation at middepths in the water column, but a poor representation of its space–time variability. Conversely, TM represents the space–time variability of $\varepsilon$ reasonably well, but systematically overestimates its magnitude. In this way, we have shown that each of these parameterizations have utility, but also some important limitations. Differing strengths and weaknesses imply that they may be used in complementary ways. In what follows, we discuss the underpinnings of their limitations, which compromise parameterization performance. Next, we put our results into the context of those from validation studies using data collected by traditional sampling platforms. We conclude with some practical recommendations to aid in the successful implementation of turbulence parameterizations using glider data in future applications.

a. Limitations inherent to parameterization formulation and CTD sampling capabilities

To first order, the limitations of FP and TM derive from a combination of 1) constraints inherent in the method; 2) the violation of implicit assumptions; 3) constraints on where the parameterization can be applied in the water column; and 4) CTD capabilities.

1) FINESCALE PARAMETERIZATION

In this application, the main limitation of FP is its inability to resolve spatiotemporal features. This stems in part from the relatively coarse time and space scales resolved by the method: an important consequence of the physical basis of the FP formulation is that the time and space scales resolved by the method are limited to those of the internal wave field. As such, $\varepsilon_{\text{FP}}$ can only produce an estimate of the average dissipation associated with the steady-state internal wave environment over vertical scales of tens to hundreds of meters and time scales of hours to tens of hours. FP is not expected to deliver a picture of spatiotemporal variability comparable to that obtained with microstructure estimates (Whalen 2021). However, even on resolvable scales, we further find that FP fails to capture the spatiotemporal variability of equivalently averaged microstructure estimates (Fig. 6b). We attribute this primarily to the limited dynamic range of $\epsilon$ that can be accessed by FP in this particular setting. The coarse vertical resolution of $\varepsilon_{\text{FP}}$ (a feature inherent to the finescale spectral method), coupled with constraints that limit where the parameterization can be applied in the water column [above a minimum height of one segment length from the seafloor, below the pycnocline (see Kunze et al. 2006b; Kunze 2017) and away from boundaries (see Polzin et al. 2014)] impose significant restrictions on the fraction of physical space where FP can yield an estimate of $\epsilon$ in this shallow water environment. Further, these restrictions result in the systematic exclusion of segments of the water column where $\epsilon$ is enhanced (above the pycnocline and in shallow water close to bathymetry). Ultimately, the finescale parameterization yields $\epsilon$ estimates only in the low-$\epsilon$ waters of the middepths in the center of the basin, a place where the dynamic range in direct estimates of dissipation is small. This fact precludes the parameterization from effectively mapping the most meaningful horizontal variability in dissipation in this setting, which is derived from the contrast between the deep basin and shallow basin margin.

2) THORPE SCALE

In this study, the major limitation of TM is its overprediction of $\epsilon$ magnitude. This can arise both from constraints imposed by CTD sampling capabilities, and from the violation of assumptions inherent in the physical basis of the method. Motivated by discussion in the literature, here we consider the possible roles of CTD sampling characteristics (e.g., Ullman et al. 2003; Frants et al. 2013), overturn life stage (Seim and Gregg 1994; Mater et al. 2015), and the degree of turbulent anisotropy (Dillon 1982; Mater et al. 2013) in underpinning $\varepsilon_{\text{TM}}$ overestimation documented here.

(i) Constraints imposed by CTD capabilities

Constraints imposed by CTD capabilities, specifically via imposing a minimum detectable overturn size and hence a minimum Thorpe scale $L_{\text{TM}}$ can plausibly explain a majority of the overestimation we observe. This conclusion is based on the outcome of a Monte Carlo experiment (see appendix) that assesses how the minimum Thorpe scale detectable given CTD capabilities (specifically the sampling frequency and noise level necessitating density profile smoothing), $L_{\varepsilon_{\text{TM}}} = 1 \text{ m}$, affects TM’s characterization of the $\epsilon$ field described by the microstructure observations. We use the hypothetical $\varepsilon_{\text{TM}}$ distribution obtained in the experiment (dashed yellow histogram in Fig. 6c) to quantify a plausible influence of CTD capabilities on the overestimation of $\varepsilon_{\text{MU}}$ magnitude. We find that restricting $L_{\varepsilon_{\text{TM}}}$ to $1 \text{ m}$ can bias the distribution of $\varepsilon_{\text{TM}}$ high relative to that of $\varepsilon_{\text{MU}}$ by an order of magnitude, thus explaining approximately half of the observed difference between the geometric mean values of the $\varepsilon_{\text{MU}}$ and actual $\varepsilon_{\text{TM}}$ distributions. Further, the hypothetical distribution provides insight into the lower bound of the observed $\varepsilon_{\text{TM}}$ distribution, suggesting it is set by $L_{\varepsilon_{\text{TM}}}$. We note that this restriction on detectable Thorpe scale fails to explain the remaining half of the $\varepsilon_{\text{TM}}$ overestimation, and further the tendency for TM to overestimate $\varepsilon_{\text{MU}}$ at the high end of the $\epsilon$ distribution.
A restriction on detectable Thorpe scale sets a lower bound on the minimum resolvable buoyancy Reynolds number, \( \text{Re}_{B \text{min}} = \frac{\nu}{N^2} \), for a given stratification, \( N \), which is helpful to consider when interpreting TM’s performance in this particular application (Fig. 8). The \( \text{Re}_{B \text{min}} \) value is defined by the minimum resolvable dissipation rate from the Thorpe scale method, \( \dot{\epsilon}_{\text{Tmin}} \), imposed by \( L_{\text{Tmin}} \) [where \( \dot{\epsilon}_{\text{Tmin}} = (C_0 L_{\text{Tmin}})^2 N^3 \)]; this implies that for a given value of \( L_{\text{Tmin}} \), there is a distribution of \( \text{Re}_{B \text{min}} \) values resolvable by TM that depends on the distribution of \( N \). A comparison of this distribution for \( L_{\text{Tmin}} = 1 \) m (brown histogram in Fig. 8a) to that of \( \text{Re}_B \) values inferred from the microstructure estimates (assumed to represent the true \( \text{Re}_B \) distribution) for both \( \dot{\epsilon}_{\mu} \) and \( \dot{\epsilon}_{\mu}^\text{TM} \) estimates (purple and yellow histograms in Fig. 8a, respectively), reveals a severe mismatch in the minimum Reynolds numbers resolvable by the Thorpe scale method and the \( \text{Re}_B \) values that characterize this environment. We find that TM performs significantly better with respect to point-to-point agreement between \( \dot{\epsilon}_{\text{TM}} \) and \( \dot{\epsilon}_{\mu}^\text{TM} \) estimates when \( \text{Re}_B > 1000 \) relative to when \( \text{Re}_B < 10 \) (Fig. 8c vs Fig. 8b). We conclude that shortcomings in TM to accurately predict the magnitude of \( \dot{\epsilon} \) derives, in significant part, from insufficient CTD capabilities that precludes the detection of sufficiently small overturns given the \( \dot{\epsilon} \) magnitude and stratification that characterize this environment.

(ii) Violating assumptions of the Thorpe scale formulation

A second possible source of overestimation in \( \dot{\epsilon}_{\text{TM}} \) derives from the violation of fundamental assumptions of TM about the turbulence energy budget. TM assumes that the Ozmidov and Thorpe scales are approximately equivalent, which is appropriate for turbulent overturns associated with a well-developed inertial subrange (Dillon 1982; Mater et al. 2015). This assumption can be problematic for early-stage turbulent overturns, when density gradients may not yet be well mixed, and available potential energy exceeds turbulent
kinetic energy. In young over turns, the $L_T/L_O$ ratio should exceed 1, resulting in an over estimation of $\epsilon_{TM}$ (Mater et al. 2015; Scotti 2015). Noting the tendency for $L_T/L_O$ to increase with overturn size (Mater et al. 2015), the limitation on detectable overturn size implies that TM may be systematically prone to observe young over turns.

To assess the role of overturn age in $\epsilon_{TM}$ over estimation not accounted for by CTD capabilities, we visualize the agreement of $\epsilon_{TM}$ and $\tau_{\mu U}^{TM}$, expressed via their implied $L_T/L_O$ ratio, as a function of the nondimensional Thorpe scale, $\bar{L}_T = L_T/(\nu/N)^{1/2}$, a metric of inverse overturn age (Fig. 9; following Mater et al. 2015). We restrict our analysis to estimates of $\epsilon$ that TM can resolve with our CTD capabilities. Consistent with Mater et al. (2015), the degree of TM agreement with $\tau_{\mu U}^{TM}$ varies systematically with $\bar{L}_T$, and TM over-prediction ($L_T/L_O >$ unity) occurs when $\bar{L}_T$ is large. This overestimation is most severe when $Re_B$ is small, dissipation is weak and, to a lesser degree, stratification is strong. We conclude that a subset of the detectable over turns ($Re_B > 1000$) we observe are in the earlier stages of overturning; these violate the assumption that $L_T/L_O \sim 1$ leading to overestimation of $\epsilon$. Unlike our assessment of the influence of the minimum detectable overturn size, we are unable to quantify how much this may be contributing to the overall overestimation.

Violating another fundamental assumption of TM, isotropy, can also cause inaccuracy in $\epsilon_{TM}$ predictions (Dillon 1982; Mater et al. 2013, 2015). While we do see better point-to-point agreement with microstructure observations when $Re_B$ is large (Fig. 8c vs Fig. 8b), we attribute this to the aforementioned restriction on TM-resolvable $Re_B$ due to CTD capabilities. The effects of anisotropy on $\epsilon_{TM}$ magnitude cannot be assessed in this study because all resolvable over turns have $Re_B$ in the range of isotropic turbulence. Assessing whether the violation of the isotropy assumption is important to TM performance for gliders would require a higher CTD sampling rate and reduced CTD noise levels that would permit the resolution of $Re_B$ values in the anisotropic regime.

b. Comparison to past implementations and validation studies

To assess whether parameterization performance is uniquely influenced by the sampling capabilities of gliders, we compare our results to validation studies using data collected by traditional sampling platforms.

1) FINESCALE PARAMETERIZATION

Our key finding that FP and microstructure estimates agree to within a factor of 2–3 is consistent with past validation studies using data collected by a variety of sampling platforms (e.g., Gregg 1989; Polzin et al. 1995; Frants et al. 2013; Waterman et al. 2014; Whalen et al. 2015). It is further in line with community consensus that finescale methods are accurate to within a factor of 2–4 at best (MacKinnon et al. 2015). Our results also concur with the general order-of-magnitude level agreement reported by Johnston and Rudnick (2015), who compared glider-based FP estimates to equivalent metrics derived from disparate tracer release and microstructure observations in the California Current System. Our more exacting test provides no indication that the accuracy of glider-derived $\epsilon_{FP}$ differs in any significant way from that of estimates inferred using more traditional sampling methods.

In contrast, our finding that FP fails to resolve meaningful space–time variability is inconsistent with the study of Johnston and Rudnick (2015), which successfully uses FP and sustained glider observations to map the seasonal and cross-shore structure of mixing along repeat lines in the California Current System. Key differences between these applications are the water column depth, the scales of space–time variability targeted, and that Johnston and Rudnick (2015) were able to estimate $\epsilon_{FP}$ using shear from glider-mounted ADCP data.
The extensive dataset used in the Johnston and Rudnick (2015) study permits the characterization of space–time variability on much larger horizontal space and time scales that are better matched to the resolution of the finescale method.

2) Thorpe scale

Our key finding is that TM overestimates the magnitude of $\epsilon$ by a factor of $\sim 20$ is also within the range of agreement reported by validation studies using nonglider sampling platforms. For example, Dillon (1982), Ferron et al. (1998), Frants et al. (2013), and Mater et al. (2015) all report estimates of $\epsilon_{TM}$ exceeding estimates of $\epsilon$ derived from microstructure observations by one to two orders of magnitude. Thorpe (2012) suggested that, in the presence of internal waves, the sloping sampling trajectory of gliders could bias $\epsilon_{TM}$ high due to false overturn detection, an issue that is also of concern when estimating $\epsilon_{TM}$ using data from towed undulating platforms (Ullman et al. 2003; Ott et al. 2004). False overturn detection is a concern when the internal wave slope exceeds the slope of the instrument trajectory. Here the mean (± one standard deviation) glider trajectory slope determined from the glide angle was $0.52 \pm 0.07$, which exceeds the upper limit on internal wave slopes ($\sim 0.3$; see Thorpe 1978). Based on this, we presume that internal waves are unlikely to contribute significantly to $\epsilon_{TM}$ overestimation.

c. Recommendations for implementation of $\epsilon$ parameterizations on gliders

FP and TM have different strengths and weaknesses; making them better for different applications or oceanic environments. Additionally, there is the opportunity to use these parameterizations in complementary or synergistic ways.

1) FINESCALE PARAMETERIZATION

FP is best applied when determining the magnitude of $\epsilon$ accurately is important, and when the targeted space–time scales of turbulence variability are sufficiently large. FP provides estimates of dissipation at a relatively coarse vertical resolution; a limitation inherent to its formulation and specifically the need to choose a window length for spectral analysis. Thus, in theory, FP is better suited to mapping horizontal variability, which also tends to be larger scale compared to that in the vertical. Glider-mounted CTDs are promising for the application of FP due to the sustained, high-horizontal-resolution datasets that they can provide. The resolvable horizontal spatiotemporal scales of gliders is primarily dependent on their dive depth given that the glider travels at a near constant speed. In this application, the mean (± one standard deviation) horizontal spacing of subsequent upcast/downcast profiles in space and time were $0.25 \pm 0.15$ km and $0.2 \pm 0.05$ h, respectively.

Unfortunately, in this particular shallow water application, the ability of the glider to effectively map horizontal variability in dissipation was compromised by the FP’s coarse vertical resolution and the requirement to apply the method below the pycnocline; these limitations resulted in very few $\epsilon_{FP}$ estimates and, importantly, the failure to capture important regions of enhanced dissipation in shallow water. As discussed above, coarse vertical resolution is inherent to the FP formulation as it is set by the choice of spectral window length; here $64$ m was the minimum window length we could use to resolve a reasonable range of internal waves scales (10–25 m) over which to integrate. The summary figure (Fig. 11) highlights the implications of this and the below-the- pycnocline restriction on where in the water column FP can be applied (dotted green outline), illustrating that in this relatively shallow basin $\epsilon_{FP}$ estimates are constrained to the low-$\epsilon$ environment of the middepth waters in the center of the basin, and unable to capture regions of elevated $\epsilon$ over the shallow bathymetry. We conclude that the FP is best applied in deeper environments, i.e., where there is sufficient depth beneath the pycnocline to allow for a sufficient resolution of $\epsilon_{FP}$ estimates in both the horizontal and vertical to permit the visualization of spatiotemporal patterns, and where the targeted space–time variability exists in regions of the water column FP can access.

2) Thorpe scale

TM is best applied when characterizing the fine details of space–time variability is of interest, and when determining $\epsilon$ magnitude within an order of magnitude is sufficient. In our application, the overprediction of the magnitude of $\epsilon$ is predominantly a product of the low CTD sampling frequency and the need to remove noise in the CTD data that imposes a limitation on the minimum resolvable buoyancy Reynolds number. This limitation implies that accurate predictions of $\epsilon_{TM}$ are restricted to where $Re_B$ is large; which, in this setting, implies a restriction to over the shallow bathymetry. In these locations, TM agrees well with microstructure estimates, with $L_T/L_O$ ratios close to 1 (Fig. 10). Therefore, while TM can be applied in a greater fraction of physical space than FP (Fig. 11, yellow dot–dashed outline), accurate quantitative assessments of dissipation are restricted to regions where there is high $\epsilon$, weak stratification, and high values of $Re_B$. In future applications of TM to glider data, we recommend assessing the minimum resolvable buoyancy Reynolds number based on the minimum detectable Thorpe scale and using it to assess confidence in the accuracy of $\epsilon_{TM}$ estimates. If $Re_B$ information is available, it can be used as an additional quality control criterion to exclude estimates for which $L_O$ is not expected to be well approximated by $L_T$. This may be helpful in generating more accurate $\epsilon_{TM}$ predictions.

To extend the range of resolvable $Re_B$ and consequently expand where TM can be meaningfully applied, the CTD sampling frequency should be increased to resolve smaller Thorpe scales. By calculating $Re_B$ distributions using smaller values of $L_{Tmin}$, we find that $L_{Tmin} = 0.05$ m is sufficient to resolve the majority of the microstructure observations (dark gray dashed line in Fig. 8a). Assuming noise-free data, resolving a minimum Thorpe scale of this size would require a CTD sampling frequency of 1.5 Hz. Therefore, to remove the biggest limitation on TM, the CTD sampling frequency should be 1.5 Hz or greater in order to resolve small enough overturns, while keeping in mind that the glide slope must remain $> 0.3$ to avoid false overturn detection. This is faster sampling than is currently standard for glider CTDs, which generally sample slowly to conserve battery life and extend glider missions.
However, even with a higher CTD sampling rate, TM may still overestimate $\epsilon$ due to other limiting factors. For instance, electronic noise played a role in our relatively large $L_{T\text{min}}$; due to the need to smooth out noise in our dataset, $L_{T\text{min}}$ was higher than would be estimated from sampling frequency alone. Smoothing the raw density profiles before applying TM is not necessary, just desirable in our case for consistency with the FP analysis. For unsmoothed data, using both a higher $\rho_{\text{noise}}$ and a run length quality control criteria to remove small overturns is preferred (Galbraith and Kelley 1996). Nevertheless, in this case noise levels in the CTD data will likely impose similar limitations on the detectable Thorpe scale. Additional limiting factors such as violating assumptions that the turbulence is in steady state (i.e., not young) and/or isotropic, may also contribute to the overestimation of $\epsilon$.

It is reasonable to ask whether the choice of a constant value for $C_0$ of 0.8 compromises the performance of TM. The availability here of concurrent microstructure shear and microstructure temperature data permit investigation of this using Smyth et al.’s (2001) empirical relationship, $\Gamma = 0.33(C_0)^{-0.63}$, where $\Gamma$ is the local, instantaneous, microstructure-derived estimate of mixing coefficient (Osborn and Cox 1972). First, we consider whether use of this instantaneous, microstructure-derived value of $C_0$ improves TM’s performance. Agreement between TM and microstructure estimates of dissipation were quantified by $\log_{10}(\epsilon_{\text{TM}}/\epsilon_{\text{m}})$ versus $\log_{10}(\epsilon_{\text{TM}}/\epsilon_{\text{m}})$, where $\epsilon_{\text{TM}}$ is the TM estimate using the variable value of $C_0$ based on $\Gamma$. The variable $C_0$ marginally improved the mean agreement of these ratios, 1.4 versus 1.3, respectively, but significantly increased the range of deviation (the interquartile ranges of these ratios are 1.3 versus 1.9). Thus, we concluded that the use of instantaneous $C_0$ does not significantly improve TM performance. Further, the microstructure-derived $C_0$ distribution supported the use of a constant value of $C_0 = 0.8$: its median value was 0.76 with a 95% bootstrapped confidence interval of [0.73, 0.80].

![Figure 10](image1.png)

FIG. 10. (a) Space–time section plot of $Re_B$ using microstructure shear $\epsilon_{\mu U}$ as a function of depth and profile number. The area where $Re_B < 1000$ is indicated in light gray. The darker gray represents the bathymetry as the glider made multiple transects over the southwest margin of the basin. (b) A similar view of $L_T/L_O$ using microstructure shear $\epsilon_{\mu U}$ in the calculation of the Ozmidov scale $L_O$. Data are plotted only where $Re_B > 1000$.

![Figure 11](image2.png)

FIG. 11. A schematic summarizing the regions where the finescale parameterization and Thorpe scale method can be applied (green and yellow boxes, respectively) given a typical density profile (pink). The region of enhanced turbulence (black curls) and low stratification, and thus high $Re_B$, lead to the best oceanographic conditions for the application of these two parameterization methods (teal).
5. Conclusions

Both the finescale parameterization and Thorpe scale method have their strengths and limitations. FP more accurately estimates the magnitude of \( \epsilon \), but its coarse vertical resolution and the restricted range in the water column where it can be successfully applied make it a poor fit to coastal shelf environments. TM tends to overestimate \( \epsilon \), but does a better job of resolving spatiotemporal patterns. The application of TM can be improved by increasing the CTD sampling frequency and decreasing CTD noise. Together, these parameterizations have promise for estimating turbulent dissipation rates from CTD data when microstructure instrumentation is not available. Correctly and knowledgeably implementing FP and TM to CTD data collected by the multitude of gliders sampling the ocean can allow greater insight into the space–time variability of ocean turbulence and the role of turbulence in oceanic processes.

Acknowledgments. We thank Adam Comeau, Jude Van der Meer, and Chris Payne for help with glider preparations, deployment, and recovery; Benjamin Scheifele, Melanie Chanona, and Ruth Musgrave for microstructure, FP, and TM code, respectively; and Justine McMillan for assistance with microstructure processing. This work was supported by the MEOPAR WHaLE project; NSERC through a Canada Graduate Scholarship—Doctoral program; and UBC through a 4-yr fellowship.

Data availability statement. Data will be available at the University of British Columbia library with a DOI.

APPENDIX

Monte Carlo Experiment

This numerical experiment is based on the assumptions that both TM and microstructure observations sample the same overall distribution of \( \epsilon \) (given by \( \epsilon_{\mu/L} \), assumed to represent the true \( \epsilon \) distribution), but at any given time do not measure the same dissipation due to a spatial offset (e.g., owing to sensor separation) and/or temporal biases (e.g., owing to overturn age; see Mater et al. 2015). To execute the experiment, we randomly select pairs of \( \epsilon \) estimates from the \( \epsilon_{\mu/L} \) distribution, assigning one to TM and the other to the microstructure method. For the \( \epsilon \) estimate assigned to TM, we compute the \( L_T \) implied by its value, and compare it to the minimum Thorpe scale we deem detectable by our CTD sampling capabilities. The minimum overturn size, \( L_{\text{min}} = 2 \) m, is set by the 0.5 Hz CTD sampling frequency and our density processing procedures that remove variability on length scales of less than 2 m. Depending on symmetry, \( 0.5L \geq L_T > 0.1L \), so we assume a minimum detectable Thorpe scale of \( L_{\text{min}} = 1 \) m to be conservative. If the computed \( L_T \) is greater than \( L_{\text{min}} \), we keep the pair in our two hypothetical characterizations of the \( \epsilon \) field by TM and microstructure methods, respectively. Otherwise, we discard it. We continue to randomly resample the \( \epsilon_{\mu/L} \) distribution in this way until we accumulate the same number of estimates as in our actual Thorpe scale method characterization.

REFERENCES


