Measuring Turbulent Dissipation with Acoustic Doppler Velocimeters in the Presence of Large, Intermittent, Infragravity Frequency Bores

DUNCAN C. WHEELER[a] AND SARAH N. GIDDINGS[a]

[a] *Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California*

(Manuscript received 27 August 2021, in final form 15 August 2022)

ABSTRACT: This manuscript presents several improvements to methods for despiking and measuring turbulent dissipation values with acoustic Doppler velocimeters (ADVs). This includes an improved inertial subrange fitting algorithm relevant for all experimental conditions as well as other modifications designed to address failures of existing methods in the presence of large infragravity (IG) frequency bores and other intermittent, nonlinear processes. We provide a modified despiking algorithm, wavenumber spectrum calculation algorithm, and inertial subrange fitting algorithm that together produce reliable dissipation measurements in the presence of IG frequency bores, representing turbulence over a 30 min interval. We use a semi-idealized model to show that our spectrum calculation approach works substantially better than existing wave correction equations that rely on Gaussian-based velocity distributions. We also find that our inertial subrange fitting algorithm provides more robust results than existing approaches that rely on identifying a single best fit and that this improvement is independent of environmental conditions. Finally, we perform a detailed error analysis to assist in future use of these algorithms and identify areas that need careful consideration. This error analysis uses error distribution widths to find, with 95% confidence, an average systematic uncertainty of ±15.2% and statistical uncertainty of ±7.8% for our final dissipation measurements. In addition, we find that small changes to ADV despiking approaches can lead to large uncertainties in turbulent dissipation and that further work is needed to ensure more reliable despiking algorithms.

SIGNIFICANCE STATEMENT: Turbulent mixing is a process where the random movement of water can lead to water with different properties irreversibly mixing. This process is important to understand in estuaries because the extent of mixing of freshwater and saltwater inside an estuary alters its overall circulation and thus affects ecosystem health and the distribution of pollution or larvae in an estuary, among other things. Existing approaches to measuring turbulent dissipation, an important parameter for evaluating turbulent mixing, make assumptions that fail in the presence of certain processes, such as long-period, breaking waves in shallow estuaries. We evaluate and improve data analysis techniques to account for such processes and accurately measure turbulent dissipation in shallow estuaries. Some of our improvements are also relevant to a broad array of coastal and oceanic conditions.

KEYWORDS: Coastlines; Estuaries; Turbulence; Oceanic waves; Data processing/distribution; Error analysis

1. Introduction

In recent years, acoustic Doppler velocimeters (ADVs) have proven valuable tools for measuring turbulent statistics in various environments. With their fast sample rates, relatively accurate measurements, and ability to measure in shallow water, ADVs have led to improvements in our understanding of surf zone and shallow estuarine turbulence (e.g., Feddersen 2012; Jones and Monismith 2008). Key to this success has been the development of reliable data processing techniques for despiking ADV data and calculating wavenumber spectra from ADV velocity data in the presence of surface gravity waves (e.g., Goring and Nikora 2002; Shaw and Trowbridge 2001).

At the same time, physical oceanographers are becoming increasingly aware of the importance of infragravity (IG) oscillations (25–250 s periods) for coastal, shallow reef, and estuarine processes (e.g., Bertin et al. 2018; Becker et al. 2016). IG waves have long been known to contribute to swash and run up on beaches (e.g., Guza and Thornton 1982; Ruessink et al. 1998). More recently, researchers have shown that IG oscillations are important in sediment transport and estuarine mouth closure processes (e.g., Behrens et al. 2013; Bertin et al. 2019; Bertin and Olabarrieta 2016). Some studies have also shown that shallow estuarine mouths can act as low-pass frequency filters, leading to oscillations within estuaries dominated by IG frequencies with velocity amplitudes as large as tidal velocities (e.g., Williams and Stacey 2016; Harvey et al. 2022; McSweeney et al. 2020). To be explicit, we are using oscillation and wave as distinct terms here, with oscillation implying repeated back and forth motion with a definable frequency range and wave implying a more regular process with consistent statistics that follows wave theory.

With our new understanding that IG frequencies are dynamically important, it is increasingly important to understand how IG motions affect velocity and turbulence. It is particularly important to tease out how IG velocity oscillations interact with other sediment transport mechanisms and how they affect mixing within estuaries. To address turbulence in the presence of IG motions, ADVs appear to be the

Corresponding author: Duncan C. Wheeler, dcwheele@ucsd.edu

DOI: 10.1175/JTECH-D-21-0144.1

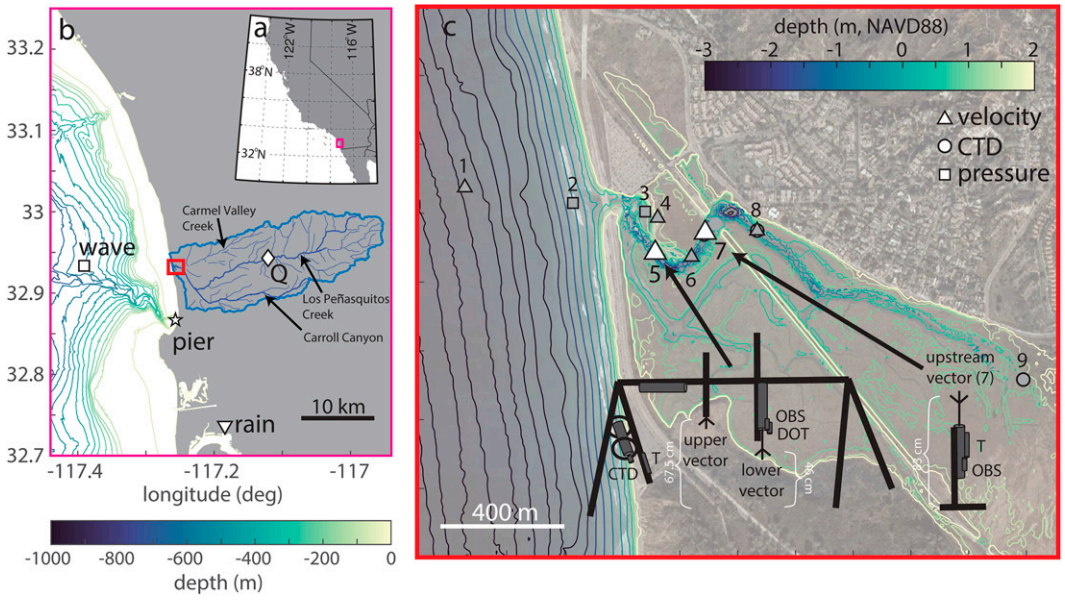

FIG. 1. Map of deployment at Los Peñasquitos Lagoon. (a) The California coast with the domain of (b) in the pink box. (b) The San Diego coastline with the lagoon watershed outlined in blue and the domain of (c) in the red box. (c) The lagoon bathymetry with the locations of various instruments deployed for the experiment as well as sketches of the moorings at locations 5 and 7 with distances between the bottom and the ADV heads. This paper only uses data from the ADVs at locations 5 and 7.

ideal tool, as IG waves are often most dominant in particularly shallow waters. However, many of the existing methods for analyzing ADV data in the presence of waves make assumptions that fail in the presence of IG motions. In particular, the irregular IG motions introduce nonstationary turbulence with unpredictable changes in velocity variance that interfere with despiking algorithms, and the nonlinear behavior of the oscillations lead to changes in spectral slope when using standard frozen turbulence assumptions. To be clear, we use the term oscillations instead of waves because the data presented here are nonlinear and asymmetric, frequently with broken fronts. Such oscillations match observations from other experiments in shallow estuaries (e.g., Williams and Stacey 2016; McSweeney et al. 2020) and are most accurately described as IG frequency bores. Therefore, we will use this terminology for the rest of the paper. Harvey et al. (2022) and Fig. 2e have raw ADV time series that provide visualizations for these oscillations.

This manuscript closely examines and adapts existing techniques for despiking ADV data and measuring turbulent dissipation in the presence of large IG frequency bores in a shallow estuary. We identify where current algorithms fail and suggest alterations in order to deal with the unique challenges presented by dominant IG oscillations. We note that the effects we correct for are not unique to IG frequency bores, meaning that many of the techniques presented here have the potential to be applied to a wider range of situations. In addition, our modifications to inertial subrange fitting methods are improvements even in the absence of nonlinear dynamics. We provide tolerance calculations and an error analysis framework with the goal that future researchers may use these techniques to more easily evaluate turbulence using ADVs. This error analysis examines variables not traditionally considered in final turbulence calculations, allowing us to identify areas where ADV processing for turbulence calculations needs careful consideration and further work to produce reliable results.

2. Data

The data used for the analyses presented were obtained during a short field deployment at the beginning of 2020 aimed at understanding IG frequency bores within Los Peñasquitos Lagoon. Los Peñasquitos lagoon, shown in Fig. 1, is a low-inflow, bar-built estuary in Southern California with a mouth that has been heavily modified due to the presence of a road. The shallow sill at the mouth is above sea level at low tide, while most of the lagoon interior is around 2 m deep or less. When submerged, the shallow sill acts as a low-pass frequency filter where larger sea and swell waves dissipate in the surfzone while IG oscillations propagate into the estuary (Harvey et al. 2022; Williams and Stacey 2016). The 1–2.5 m tides during our deployment drive a semidiurnal 0.3–1.5 m water-level change inside the lagoon, leading to water depths at station 5 (see Fig. 1) that range from 0.2 to 1.7 m.

As part of the deployment, two synced ADVs (Nortek Vectors) were mounted, along with other instruments, on a sawhorse frame that was secured with sand anchors into the bed at location 5 in Fig. 1c, 300 m upstream of the mouth. These two ADVs sampled the water 30 and 50 cm above the bottom. A third, bottom frame mounted, upward looking ADV

was deployed 250 m farther upstream at location 7 in Fig. 1c, sampling 1 m above the bottom. Diagrams for both moorings are shown in Fig. 1c. All three ADVs started sampling at midnight on 2 February and collected data continuously at 16 Hz, with 5 s breaks every hour to write the data to memory. One of the synced ADVs stopped sampling the afternoon of 11 February due to memory limitations. The other two ADVs included synced optical backscatter measurements and low batteries led to shortened and irregular bursts starting the morning of 14 February, making later data unusable.

After recovery, the data for each ADV were converted to a netCDF4 file and processed for quality control. The pressure was corrected for atmospheric pressure changes using a nearby National Estuarine Research Reserve meteorological station and converted to depth using the Python gsw package (https://pypi.org/project/gsw/3.0.3/; McDougall and Barker 2011). Initial cleaning of the velocity data used a correlation cutoff of 70% and signal-to-noise ratio (SNR) cutoff of 10 dB for any beam, based on initial data inspection as recommended by the manufacturer (Nortek 2018). Details on how to access these data and the code for the methods presented below can be found in the data availability statement.

3. Despiking

The first challenge when using ADV field measurements is despiking. Due to phase shift ambiguities caused by contamination from previous pulses reflecting off air bubbles and other reflective surfaces suspended in the water, spikes can appear in ADV velocity records (e.g., Goring and Nikora 2002). These spikes can contaminate otherwise valid data, leading to errors in later averaging or turbulence calculations.

a. *Existing methods*

Perhaps the most common method for eliminating these spikes is the Goring and Nikora (2002) method. This approach assumes that valid points collected by an ADV, after removing the low-frequency signal, are clustered in a dense ellipsoid in phase space. Here, phase space is defined by using the measured velocity (u), first time derivative (du/dt), and second time derivative (d^2u/dt^2) to form 3 independent axes. The outer limits of the valid data ellipsoid are determined using the universal threshold, $\sqrt{2\log n} \times \sigma$, which is an upper bound on the maximum expected value in a sample of size n taken from a Gaussian distribution with standard deviation σ. Goring and Nikora (2002) use the universal threshold to define an ellipse for each of the three 2D projections of phase space. Any points outside any one of those ellipses are replaced and the algorithm is repeated until no new spikes are detected. Goring and Nikora (2002) also note that due to the different spike numbers in vertical versus horizontal velocities, each velocity component should be despiked independently. Wahl (2003) showed that the standard deviation–based universal threshold could be biased by outliers and suggested the median of the absolute deviation from the median (MAD) as a robust alternative.

Further modifications to the phase space method came from Parsheh et al. (2010), who noticed that particularly large spikes could introduce a bias in the derivative, making normal data points appear as spikes in phase space. The proposed fix used the MAD-based threshold to perform an initial screening based only on velocity magnitude before continuing to the phase space method. In addition, Parsheh et al. (2010) suggested that replacing spikes with the last valid data point prevented the extended contamination of cleaned data from yet to be detected spikes found when using other interpolation methods.

Upon examination of our data, we find that the distributions for our velocities and derivatives are noticeably non-Gaussian, with long tails that extend beyond the universal threshold and MAD threshold cutoffs (see Fig. 2a). As a result, the use of unmodified versions of the methods proposed by Goring and Nikora (2002), Wahl (2003), and Parsheh et al. (2010) misidentifies too many valid data points as spikes. These tails are likely a direct result of nonstationary turbulence leading to changes in the variance of our data over our burst length (30 min). In addition, this nonstationary effect appears to be related to the passage of IG frequency bores, which come in at largely irregular and unpredictable intervals and can be seen as the low-pass signal in Figs. 2e and 2f. This short term increase in turbulent energy with passing bores is very similar to what Simpson et al. (2004) observe in the presence of a tidal bore propagating up a shallow estuary. As a result, simply changing the burst length for our despiking algorithm was not an option.

Since Goring and Nikora (2002), several new methods of despiking have been proposed that we also find inappropriate for our situation. Cea et al. (2007) explore how correlations between different velocity components can be used to create a different 3D space for detecting spikes in a similar way to the phase space approach. However, the lack of temporal derivatives (e.g., du/dt) misses an essential element of spike detection for our situation. In data with varying velocity amplitudes, as in the presence of large IG frequency bores, some spikes have magnitudes that are no bigger than other points in the same burst of data. These spikes must then be identified by how much they stand out from their immediate neighbors, which is captured by a temporal derivative. In addition to Cea et al. (2007), Razaz and Kawanisi (2011) propose an improved wavelet-based despiking algorithm that decomposes data into wavelet packets. Sharma et al. (2018) suggest a modified singular spectrum analysis approach that identifies key eigenvectors which describe the data and can be used for reconstructing a spike-free signal. However, inherent in both the Razaz and Kawanisi (2011) and Sharma et al. (2018) approaches is a reliance on stationary data that are broken by the presence of IG frequency bores. To adapt either of these approaches would be very dependent on the individual characteristics of the IG frequency bores in each deployment, preventing the creation of a widely usable algorithm.

For replacing detected spikes, Jesson et al. (2013) tested a variety of proposed methods in combination with many of the above despiking methods. Their results show that the optimal replacement method depends on the detection method used, and verified the conclusions of Parsheh et al. (2010) that the

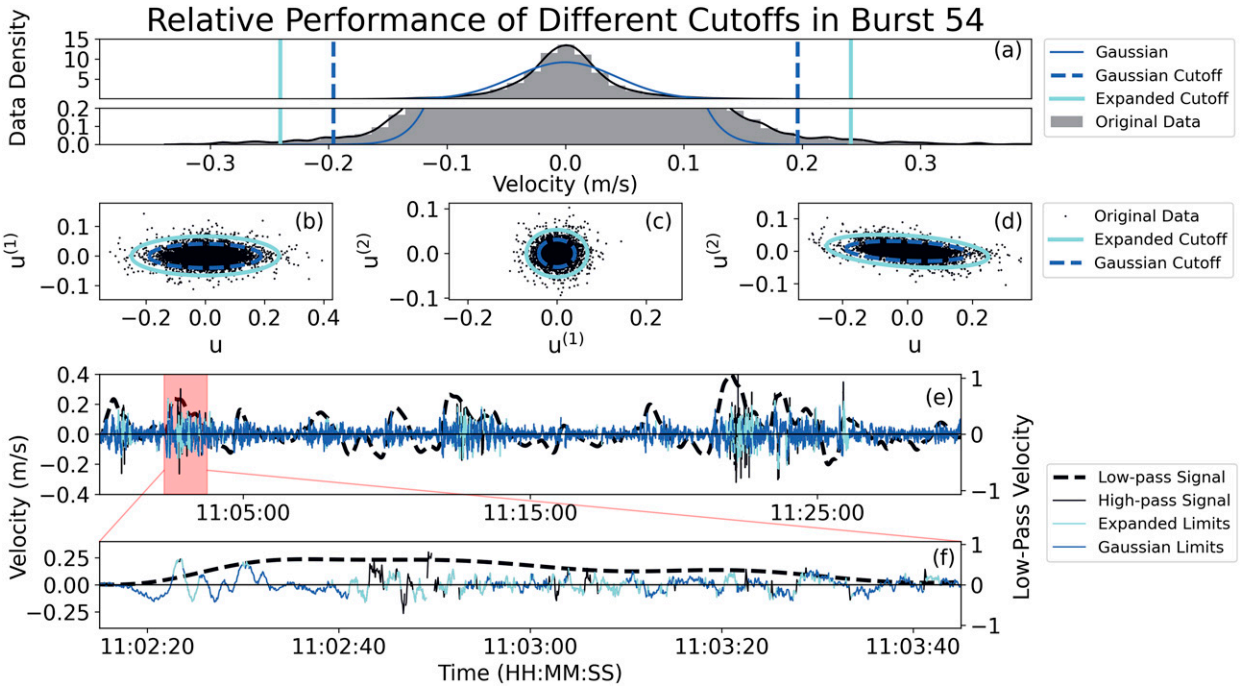

FIG. 2. Comparing the Gaussian-based cutoff with expanding cutoffs for a burst with IG energy. All data shown—except low-pass data in (e) and (f)—have the low-pass filtered signal (20 s) removed. In all plots, black corresponds to the original data, dark blue corresponds to the Gaussian-based cutoff, and cyan corresponds to the expanded cutoffs. (a) A histogram and corresponding kernel density estimate (black line) of the raw data along with a Gaussian distribution that has the same mean and standard deviation as the results of the expanded cutoff algorithm (dark blue). The dashed dark blue bars showing the universal threshold cutoffs encompass the Gaussian data but cut short the extended tails of the true data. The solid cyan bars showing the expanded cutoffs relax the restrictions, keeping more data. Note that the *y* axis has been expanded between 0 and 0.2 to highlight the extended tails of the data. (b)–(d) The 2D projections of phase space with the Gaussian- and expanded-cutoff-based ellipses. (e) The results of the two different despiking algorithms. The expanded limits appear to keep more data during periods with higher high-pass velocity (solid black line) variance that tend to correspond with large positive low-pass velocities (dashed black line). Because this particular time period has weak tides, the low-pass velocity largely corresponds to IG frequency oscillations. (f) A zoom-in on 1.5 min of (e) to highlight why we believe the Gaussian-based limits are throwing out real data. Note that the Gaussian limits are always stricter than the expanded limits, so anywhere light blue is seen represents time periods where the expanded limits keep data while the Gaussian limits throw out data. Similarly, black represents time periods where both limits throw out data.

last valid data point method was optimal for the modified phase space method.

b. *Expanding cutoff algorithm*

Based on the variety of approaches above, we determine that the best approach is to use the Goring and Nikora (2002) algorithm with a modified cutoff to account for our non-Gaussian data distribution. Because the Goring and Nikora (2002) universal threshold is too strict for our data, we modify the phase space approach with an expanding cutoff algorithm, similar to the one used by Islam and Zhu (2013). This approach uses the basic observation by Goring and Nikora (2002) that valid ADV data fall within a tight ellipse in phase space. However, rather than using a Gaussian distribution–based cutoff, we use the density of points in phase space to determine where the ellipse ends. Islam and Zhu (2013) use a bivariate kernel density function to create a density map of the data in each 2D phase space projection. The threshold is then determined by where the gradient in the density levels off (indicating an end to the dense cluster in phase space). However, Islam and Zhu (2013) found that the particular implementation of the bivariate kernel density function appears to heavily influence the mean velocity of their data. Because Islam and Zhu (2013) do not provide an explicit sensitivity analysis on the frequency spectrum of their data, the sensitivity of the mean currents suggests that this approach would not be a robust and easily adaptable algorithm for turbulence calculations; thus, here we use a modified expanding cutoff algorithm.

We begin by linearly interpolating any data already marked as bad and removing the 1/20 Hz low-pass filtered data so that we are only despiking the high-frequency signal (appendix A, steps 1 and 2). Then, for our modified cutoffs, we start with the universal threshold from Goring and Nikora (2002) as an initial estimate (appendix A, step 4). While these thresholds are inaccurate in magnitude, the relative magnitudes (and therefore the shape of the ellipsoid) are well represented by the standard deviation of the data, as can be seen in Figs. 2b–d. To determine the overall magnitude for a given 2D projection of phase space, we then increase the corresponding cutoff limits by

1% repeatedly, keeping the ellipse axis ratio constant throughout (appendix A, step 5). For each expansion, we then calculate an elliptical ring density of points as the number of data points that lie between the ellipses defined by the new cutoff and the previous increment divided by the corresponding area (appendix A, steps 6 and 7). The modified cutoff is then chosen as the point at which the elliptical ring density of points in the ring between sequential cutoffs decreases by more than 95% of the previous value (appendix A, steps 8 and 9) (see Figs. 2b–d). We do this expansion for each 2D projection of phase space independently. Figure 2 shows how these expanded cutoffs keep more valid data points than the original Goring and Nikora (2002) cutoffs, allowing us to produce a more complete despiked time series in the presence of nonstationarity.

Including these modified cutoffs, our full final despiking algorithm is as follows. We begin by linearly interpolating any NaNs in our data (appendix A, step 1) and then follow the Goring and Nikora (2002) algorithm exactly until the point of specifying cutoffs (appendix A, steps 2–4). We next use our expanding algorithm to define the modified cutoffs as described above (appendix A, steps 5–9). From here, we follow the modified phase space approach suggested by Parsheh et al. (2010) and Jesson et al. (2013) for replacing spikes. First, we check if there are any points that explicitly exceed the velocity threshold, marking those as spikes and replacing them with the last valid point (appendix A, step 10). We use the last valid point, because both Parsheh et al. (2010) and Jesson et al. (2013) show that this replacement method produces a more accurate spectrum than other interpolation methods when paired with the modified phase space approach introduced by Parsheh et al. (2010). If we identify spikes this way, we return to the beginning of the iteration process (appendix A, steps 11, 18, and 19), allowing us to remove particularly bad spikes that might lead to valid points being detected as spikes later on. If we do not identify spikes this way, we identify spikes using the full phase space ellipses, being careful to use the rotated ellipse for u–d^2u/dt^2 space as in Goring and Nikora (2002), and again replace identified spikes with the last valid point (appendix A, steps 12–18). Finally, we iterate until no new data points are identified as spikes, add back the low-frequency signal, and return any original NaN values that were interpolated at the beginning back to NaN (appendix A, step 20). This iteration occasionally gets stuck on repeatedly identifying the same points as spikes, so we impose a limit of 100 iterations to prevent an infinite loop. Similar to Goring and Nikora (2002), we despike each velocity component independently as shown in Figs. 2b–d. This phase space despiking algorithm with modified expanding cutoffs could be used on any dataset, but is particularly well suited when the data are non-Gaussian due to nonstationary turbulence over the burst length. Examples include irregular waves and bores such as these, but also the passage of fronts.

4. Spectral estimation

Once the velocity has been effectively despiked, the next major challenge in measuring turbulent dissipation with an ADV is calculating a wavenumber power spectrum of the vertical velocity while properly accounting for the presence of oscillating velocities.

a. *Existing methods*

The standard approach to calculating dissipation from velocity measurements is based off the inertial turbulence subrange. For wavenumbers small compared to the energy generation scales, but still much larger than the viscous dissipation scale, i.e., the Kolmogorov scale, the vertical velocity, horizontal wavenumber energy density spectrum follows (Pope 2000)

$$S_{ww}(\kappa_1) = C_1' \epsilon^{2/3} \kappa_1^{-5/3}, \tag{1}$$

with S representing a directional energy spectrum, the subscripts ww indicating a vertical velocity autospectrum, the functional dependency on κ indicating a spectrum in wavenumber space, the subscript 1 representing the along-flow direction, ϵ representing dissipation, and $C_1' = 4/3 \times 18/55 \times C \approx 0.65$, where $C = 1.5$ is Kolmogorov's constant, 18/55 is a conversion factor for the one dimensional spectrum in the direction of the mean current, and 4/3 is a further conversion factor for the spectrum in the direction perpendicular to the mean current (Pope 2000). Note that we are using vector notation whenever we refer to the velocity components used in calculating a spectrum and index notation whenever we refer to the direction of the wavenumbers calculated. By plotting the spectrum in log space, a line with slope −5/3 can be fit to observed data, and the y intercept of that line can be used to calculate the turbulent dissipation (e.g., Burchard et al. 2008; Bluteau et al. 2011; Rusello and Cowen 2011).

Since velocity measurements are often taken at one point in space over time, rather than over space at one point in time, initial spectra calculated are frequency spectra. Using Taylor's frozen turbulence hypothesis, the mean current during the observation period can then be used to convert observed frequencies to observed wavenumbers (e.g., Rusello and Cowen 2011; Lumley and Terray 1983).

To reduce spectral noise, spectra are calculated over minutes long batches of data, where windowing and segmentation can be used to produce high degrees of freedom for reduced errors. Because the turbulence is assumed constant over the averaging time scale, a robust fit to the low noise spectrum will accurately determine the turbulent dissipation. However, in the presence of waves with periods smaller than the observation period, the advection of turbulence past the sensor by wave orbital velocities can bias energy spectra high for frequencies higher than the wave frequency. This is because using the mean velocity to convert frequency to wavenumber is no longer an accurate assumption (e.g., Lumley and Terray 1983; Rosman and Gerbi 2017).

Past attempts to adjust for this wave bias have employed spectral-based corrections. By assuming a random, linear wave field that leads to roughly Gaussian distributed wave velocities, an analytical solution for the correlation of measured turbulent velocities can be found. Using this correlation, one can calculate the expected difference between the measured spectrum and the ideal inertial subrange. Such corrections

were first introduced by Lumley and Terray (1983), built on by Trowbridge and Elgar (2001) and Feddersen et al. (2007) among others, and have most recently been generalized to a wide range of cases by Rosman and Gerbi (2017).

b. Semi-idealized model

Because the above methods are based on linear surface gravity waves with Gaussian distributed velocities, we evaluate their effectiveness in the presence of IG frequency bores using a semi-idealized model based on the approach in Rosman and Gerbi (2017). We start with an idealized turbulent spectrum with a known dissipation, an inertial subrange, and rolloffs near the generation (L) and Kolmogorov (η) length scales, as specified by Rosman and Gerbi (2017) and Pope (2000):

$$\begin{aligned} S_{ww}(\kappa_1) &= \frac{1}{2}\left[S_{uu}(\kappa_1) - \kappa_1 \frac{dS_{uu}(\kappa_1)}{d\kappa_1}\right], \\ S_{uu}(\kappa_1) &= \int_{\kappa_1}^{\infty} \frac{E(\kappa)}{\kappa}\left(1 - \frac{\kappa_1^2}{\kappa^2}\right) d\kappa, \\ E(\kappa) &= C\epsilon^{2/3}\kappa^{-5/3} f_L(\kappa L) f_\eta(\kappa\eta), \\ f_L(\kappa L) &= \left[\frac{\kappa L}{\sqrt{(\kappa L)^2 + 4\pi^2}}\right]^{5/3+p_0}, \\ f_\eta(\kappa\eta) &= \exp(-c_\beta\{[(\kappa\eta)^4 + c_\eta^4]^{1/4} - c_\eta\}), \end{aligned} \tag{2}$$

with E representing a nondirectional energy spectrum, $f_L(\kappa L)$ representing the low wavenumber rolloff, $f_\eta(\kappa\eta)$ representing the high wavenumber rolloff, $C = 1.5$, $p_0 = 2$, $c_\beta = 5.2$, and $c_\eta = 0.40$. Here we use a generation length scale of 2 m, which reflects our shallow water, but is also large enough to resolve a significantly large inertial subrange. In addition, the Kolmogorov length scale is calculated as

$$\eta = \left(\frac{\nu^3}{\epsilon}\right)^{1/4}, \tag{3}$$

where ν is the temperature dependent kinematic viscosity of water.

We convert this spectrum to Fourier coefficients with random phases and perform an inverse Fourier transform to obtain an idealized spatial turbulent vertical velocity dataset for that chosen dissipation value. We then sample from these idealized data at a fixed location in space at 16 Hz, while advecting the data with our IG frequency bore containing measured horizontal velocities. The resulting semi-idealized (idealized turbulence field with a fixed dissipation value advected by a realistic IG bore containing velocity field) temporal dataset then represents the vertical velocities that we would measure with our field instruments if the vertical velocities followed the initial idealized turbulent spectrum.

From these semi-idealized data, we can test different approaches to reconstruct the initial idealized turbulent spectrum and the associated dissipation value. The following subsections describe the two main approaches that we test with this method.

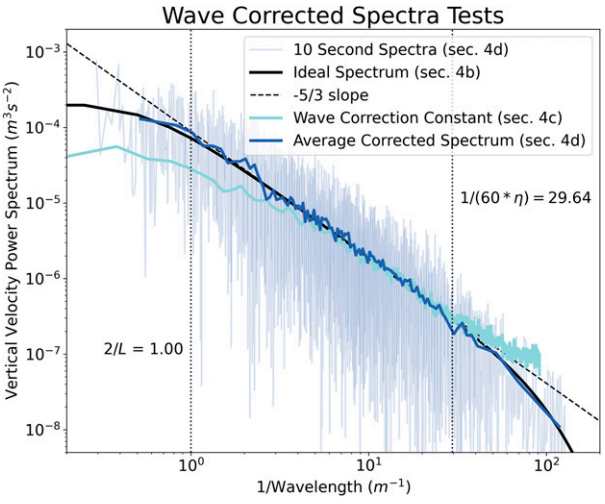

FIG. 3. The idealized turbulence spectrum used to produce our test data and the resulting reproduced spectra show that, in the presence of IG frequency bores, a segmented approach performs better than traditionally used wave correction methods. The semi-idealized spectrum (section 4b) is in black, the Rosman and Gerbi (2017) correction (section 4c) is in cyan, individual 10 s spectra from our method (section 4d) are in light blue, and the average corrected spectrum resulting from our method (section 4d) is in dark blue. Note that the 10 s spectra overlap a lot, so they are mostly transparent here to highlight where there is more overlap. The dashed vertical lines show the generation and Kolmogorov length scale derived boundaries of the inertial subrange.

c. Full Rosman and Gerbi correction

First, we are interested in testing the high-frequency limit of the correction originally presented by Lumley and Terray (1983) and most recently explored by Rosman and Gerbi (2017). This correction consists of dividing a single constant,

$$I\left(\frac{\sigma_{\text{wave}}}{u_c}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left|1 - \frac{\sigma_{\text{wave}}}{u_c}\zeta\right|^{2/3} e^{-(1/2)\zeta^2} d\zeta, \tag{4}$$

which we call the wave correction constant, into the average spectrum to correct the magnitude when converting from a temporal spectrum [$S_{ww}(\omega)$] to a spatial spectrum [$S_{ww}(\kappa)$] as

$$S_{ww}(\kappa) = \frac{S_{ww}(\omega)u_c}{I(\sigma_{\text{wave}}/u_c)}. \tag{5}$$

Note that the functionality of the wave correction constant I is explicitly included for clarity [see Eq. (4)] to demonstrate that the wave correction constant (I) depends only on the ratio of the standard deviation of the waves (σ_{wave}) to the mean current (u_c). In addition, the relation between wavenumber and frequency in Eq. (5) is $\kappa = \omega/u_c$.

Figure 3 shows that when we use I given by Eq. (4), the resulting spectrum appears to have the correct amplitude, but the wrong slope when compared to the initial ideal spectrum. This implies that the nonlinear aspects of the IG frequency bores introduce a slope variation in the frequency spectrum not predicted by, and therefore not corrected by, the linear

theory assumed by the above methods. Therefore, to avoid wave bias from IG frequency bores, we must find a different approach.

d. Segmented approach

Here, we adapt a method first introduced by George et al. (1994). By splitting the velocity data into short enough segments (as defined below), we ensure that the velocity varies linearly or not at all within any particular segment, thus enabling the use of Taylor's frozen turbulence hypothesis. Then, the resulting spectra can be averaged in wavenumber space, resulting in an accurately reproduced wavenumber spectrum. We use this approach over a 30 min burst of data to produce a final 30 min averaged spectrum that is minimally impacted by the phasing of individual IG oscillations. While 30 min is long enough for our 30 s to 2 min long oscillations, the averaging interval will need to be adjusted for datasets with different dominant frequencies. Also, because the inertial subrange scales as $\epsilon^{2/3}$, our final dissipation values will be composed of the 30 min average of instantaneous dissipation values raised to the two-thirds power, then raised to the three halves power to maintain proper units, e.g., $[\sum^{N}(\epsilon^{2/3}/N)]^{3/2}$. We will refer to this as the representative dissipation for the rest of the manuscript and discuss its implications in section 7.

For each 30 min burst of vertical and horizontal velocity, we first linearly interpolate any NaNs to produce consistently spaced data (appendix B, step 1) before splitting the horizontal velocity signal into low ($\tilde{u}$) and high (u') frequency components (appendix B, step 2) such that

$$u = \tilde{u} + u'. \tag{6}$$

We separate the frequencies by low-pass filtering with a cutoff frequency of 1/5 Hz, which we find is higher than the frequency of our IG bores while smaller than the frequency of our large-scale turbulence. This cutoff will generally work so long as there is a distinct separation between the bore frequencies and turbulent frequencies. After filtering the data, we split the low- and high-frequency horizontal and unfiltered vertical velocity data all into 10 s segments (appendix B, step 3) and calculate the mean advection speed for each segment as $c_{\text{seg}} = \sqrt{U_{\text{seg}}^2 + V_{\text{seg}}^2}$, where U_{seg} and V_{seg} are the average low-frequency horizontal velocity components for each segment. Next, we rotate the horizontal velocities into the direction of the mean velocity, as determined by the low-frequency signal for each segment independently, such that γ_{seg}, $\tilde{\gamma}_{\text{seg}}$, and γ'_{seg} represent the unfiltered, low-frequency, and high-frequency velocities in the direction of the mean current, respectively (appendix B, steps 4–7). We find that 10 s is sufficiently short compared to the minutes long IG frequency bore period while being long enough to calculate spectra that resolve the inertial subrange. For our data, the advection speeds of these 10 s segments range from 0.05 mm s^{-1} to 76 cm s^{-1}, giving resolved length scales ranging from 5 mm to 3 m.

Before we can calculate spectra for these 10 s segments, we must now use the low-frequency and high-frequency components of the horizontal velocity to filter out segments that break Taylor's frozen turbulence hypothesis or those where the advection velocity significantly changes over the 10 s segment. First, for Taylor's frozen turbulence hypothesis to hold, the turbulent fluctuations must be much smaller than the mean advection velocity (e.g., Bluteau et al. 2011; George et al. 1994). Following George et al. (1994), to satisfy this condition, we use the standard deviation of the rotated high-frequency velocity oscillations to characterize the turbulent fluctuations (appendix B, step 8). We then require the standard deviation of the high-frequency velocity component in the direction of the mean velocity to be less than 1/5 of the mean velocity magnitude (i.e., $5\sigma^{\gamma'}_{\text{seg}} \leq c_{\text{seg}}$), throwing out any segments that do not satisfy this condition (e.g., George et al. 1994) (appendix B, step 9).

For the second condition, the advection velocity cannot change too much over the 10 s segment. To satisfy this, George et al. (1994) modified a criterion first suggested by Lin (1953) to only keep segments where the variation of the advection velocity was much less than the mean. This criterion would lead us to throw out most of our segments, However, we can relax this condition using the results of an idealized test with the model introduced in section 4b. Noting that the low-frequency component of most of the 10 s segments is a roughly linearly changing advection velocity, we test the effects of a linear advection velocity on idealized data. Even though a linear advection velocity is not an oscillating velocity, we choose to test the correction used in Eq. (5) because it is designed for the high-frequency limit where the turbulent frequencies are much higher than the advection frequencies. Our low-frequency signal is defined as anything below 1/5 Hz, so we use the standard deviation of the linear advection velocity when applying the wave correction constant to our test cases (e.g., $\sigma_{\text{wave}} = \sigma^{\tilde{\gamma}}_{\text{seg}}$). In doing so, we find that the wave correction constant accurately corrects the magnitude of the resulting spectrum but fails to adjust for changes in the spectrum slope, much like the result of using the wave correction constant on IG oscillation velocities as shown in Fig. 3. However, we find that the slope error is the same order or smaller than the magnitude error and that by ensuring the change from using the wave correction constant is smaller than 5% (i.e., I between 0.95 and 1.05) we can guarantee a slope error under 5%. Plotting Eq. (4) for $\sigma^{\tilde{\gamma}}_{\text{seg}}/c_{\text{seg}} > 0$ shows that I is between 0.95 and 1.05 so long as $\sigma^{\tilde{\gamma}}_{\text{seg}}/c_{\text{seg}} \leq 1.025$. This allows us to relax the George et al. (1994) criterion and instead only require 10 s segments to satisfy $\sigma^{\tilde{\gamma}}_{\text{seg}}/c_{\text{seg}} \leq 1.025$ (appendix B, step 9).

Based on work by Feddersen (2010), we also eliminate segments that have too many unoriginal points. These unoriginal points can be due to interpolating NaNs or from replacing spikes as described above. Feddersen (2010) showed that if more than 1% of the data is unoriginal, the effects of individual spike replacement techniques start to become important. Therefore, we eliminate the 10 s segments with the largest number of unoriginal points until less than 1% of the total points in all remaining segments of the 30 min burst are unoriginal (appendix B, steps 10 and 11). It is important to note here that we do not place a limit on the number of consecutive unoriginal points or unoriginal points in a given segment. This is because including such cutoffs has little impact on the

final calculated dissipation values beyond the conditions we already use, and we choose to limit the number of subjective cutoffs used. Further discussion on the sensitivity of different subjective cutoffs (e.g., $5\sigma'_{\text{seg}} \leq c_{\text{seg}}$, $\sigma_{\text{seg}}^{\tilde{\gamma}}/c_{\text{seg}} \leq 1.025$, 1% of total points in all remaining segments) is given in section 6c.

Once we have eliminated bad segments, we detrend, apply a Hanning window, and calculate a power spectrum of the unfiltered vertical velocity for each segment (Thomson and Emery 2014b) (appendix B, step 12). We then use the local mean horizontal velocity with Taylor's frozen turbulence hypothesis to convert each segment to wavenumber space independently and correct the spectrum magnitude using I as in Eq. (5) (appendix B, step 13). The resulting individual wavenumber spectra can be seen as the thin blue lines in Fig. 3.

After converting each spectrum to wavenumber space, we remove the lowest and highest wavenumbers to avoid low-frequency contamination or aliasing of high-frequency energy. On the low end, the semi-idealized model indicates an apparent biasing of individual spectra high and that removing the 2 lowest wavenumbers is sufficient to eliminate this bias (appendix B, step 14). On the high end, we assume the presence of an inertial subrange. If the spectrum follows a −5/3 slope, we can calculate a wavenumber cutoff,

$$\kappa_{\text{cutoff}} = (\kappa_{\text{max}}^{-5/3} \times 4)^{-3/5}, \tag{7}$$

where we expect the spectrum to be 4 times larger than the spectrum at the Nyquist (κ_{max}) wavenumber (appendix B, step 15). Eliminating any points above this cutoff means that any remaining points are theoretically at least 4 times larger than the spectrum at the Nyquist frequency and can therefore never be biased by more than 25% due to aliasing, which the semi-idealized model indicates is sufficient to prevent substantial alteration of the inertial subrange slope. It is important to note that both of these cutoffs are based on our observations from looking at individual 10 s spectra and the resulting impacts on final average spectra. Therefore, we include these cutoffs in the sensitivity analysis in section 6. In addition, the bias resulting from this high wavenumber cutoff is detailed in our semi-idealized error analysis in section 6a.

After cleaning the spectrum of each segment, we consider individual spectral density values as independent estimates of the true spectral value at their given wavenumber. We bin the N spectral values in increasing order by wavenumber into bins of at least $N_{\text{min}} = 50$ as evenly as possible, as explained in appendix B, steps 16–17. We choose 50 as a balance between providing sufficient degrees of freedom that we achieve a low noise spectrum and having enough bins that the final spectrum has enough data points for us to perform a robust fit. The exact number may need to be adjusted based on the noise level and size of each dataset and is included in the sensitivity analysis in section 6c. We then average the values in each bin together to obtain a low noise spectrum without wave bias (appendix B, step 18). Because we expect the spectral estimates to vary as $\kappa^{-5/3}$, we estimate the wavenumber of each average spectral density value (appendix B, step 19) as

$$\kappa_{\text{av}} = \left(\sum_{i=1}^{N_{\text{bin}}} \frac{\kappa_i^{-5/3}}{N_{\text{bin}}} \right)^{-3/5}, \tag{8}$$

where the index represents each individual estimate being averaged in the bin and N_{bin} represents the total number of spectral estimates in a bin. Each resulting averaged spectral estimate then has $2N_{\text{bin}}$ degrees of freedom. Note, we find that a maximum N for our dataset is 5760 (180 ten-second segments × 32 kept wavenumbers per spectrum) and that a minimum N for successful dissipation calculations in our dataset is 768. Generally, if grouping individual values into bins of at least N_{min}, N_{bin} will range from N_{min} to $2N_{\text{min}}$, but successful dissipation calculations will require many bins that will as a result be more evenly distributed and generally close to N_{min}. For our data, N_{bin} ranges from 50 to 96 due to how the remainder of $N/50$ distributes into the bins, but only ranges from 50 to 53 for bursts with successful dissipation calculations.

In summary, our spectrum calculation starts by interpolating NaNs and splitting the horizontal velocities into a low and high-frequency portion (using a 1/5 Hz cutoff) before splitting both the filtered horizontal and unfiltered vertical velocity data into 10 s segments (appendix B, step 2). We then rotate both the low- and high-frequency horizontal signals into the direction of the mean velocity for each 10 s segment independently (appendix B, steps 4–7). Next, we eliminate any segments where the standard deviation of the low-frequency horizontal signal is greater than 1.025 times the mean speed or the standard deviation of the high-frequency horizontal signal is greater than 1/5 times the mean speed, all calculated in the mean velocity direction (appendix B, steps 8 and 9). We also eliminate segments with the most unoriginal points until less than 1% of the total points in all remaining segments are unoriginal. After eliminating, we take a spectrum of the unfiltered vertical velocity of all the remaining segments, convert the spectra to wavenumber space individually, including the wave correction factor (I) in Eq. (4), and trim the two lowest wavenumbers along with the wavenumbers above κ_{cutoff} to avoid aliasing. Finally, we bin the N resulting spectral estimates as evenly as possible into floor(N/N_{min}) groups by wavenumber and average each bin to obtain an averaged spectrum. It is important to note that because we assume an inertial subrange at several steps in our spectrum calculation, we only expect the final result to be accurate within the inertial subrange. Figure 3 illustrates the method's success in the inertial subrange.

5. Dissipation calculations

To use Eq. (1) to calculate dissipation, we must fit a line to the logarithm of the wavenumber spectrum within the inertial subrange. This poses two distinct problems: 1) how to fit a line and 2) where to fit a line. The details of how to approach these two problems have varied slightly in the literature, with no consensus on what approach is best. The issues and modifications we explain in this section are not specific to nonlinear or intermittent dynamics such as IG frequency bores and are therefore helpful over a broad range of observational conditions.

a. How to fit

For fitting a line, a least squares error approach is commonly used (e.g., Feddersen 2010; Jones and Monismith 2008). However, the least squares error approach originates from the maximum likelihood estimation (MLE) method specific to data with Gaussian errors (e.g., Leo 1994). As Bluteau et al. (2011) point out, the errors on a spectrum follow a χ^2 distribution, meaning that an MLE approach (rather than a least squares error) would be more appropriate as it accounts for the χ^2 distribution. Unfortunately, the solution suggested by Bluteau et al. (2011) requires constant degrees of freedom over the entire spectrum (not guaranteed by our approach) and the use of computationally costly nonlinear minimization algorithms. Here, we choose to use a weighted least squares approach with some modifications to account for the effects of a χ^2 distribution.

Considering our spectral calculation method, every 10 s segment estimate of a spectrum comes from taking the square magnitude of a calculated Fourier coefficient. The real and imaginary portion of each Fourier coefficient are sampled from a Gaussian distribution with a mean of 0. Therefore, the spectrum estimate is sampled from $[S(\kappa)/d] \times \chi^2_d$, where S is the true spectrum value, $d = 2$ is the number of degrees of freedom, and χ^2_d is a standard χ^2 distribution with d degrees of freedom [see Thomson and Emery (2014b) for more details]. Because the mean of a χ^2 distribution is $\overline{\chi^2_d} = d$, we expect the mean of multiple spectrum estimates to approach the true spectrum value (Thomson and Emery 2014b). When we bin our spectrum estimates into groups of $N_{\text{bin}} \geq 50$ and average them together, we are adding $N_{\text{bin}}\,\chi^2_{d=2}$ variables. The resulting averaged spectrum is then sampled from the same distribution with higher degrees of freedom, $d_{\text{bin}} = 2N_{\text{bin}} \geq 100$.

When performing our dissipation fit, we work in log space with the observations

$$\hat{A}_i = \log(\hat{S}_i), \tag{9}$$

taken from the distribution

$$A^{\kappa_i}(d_{\text{bin},i}) = \log\left(\frac{C_1'\epsilon^{2/3}\kappa_i^{-5/3}}{d_{\text{bin},i}} \times \chi^2_{d_{\text{bin},i}}\right), \tag{10}$$

where C_1' is defined in Eq. (1), the hat $(\widehat{\cdots})$ indicates a single sample of a random value and the subscript i denotes a specific spectral observation being used in the fit. While a χ^2 distribution is skewed, large degrees of freedom lead to a more Gaussian distribution and logarithms tend to further suppress skewness. Therefore, we choose to model $A^{\kappa_i}(d_{\text{bin},i})$ as a Gaussian with expected value

$$\overline{A}^{\kappa_i}(d_{\text{bin},i}) = \log(C_1'\epsilon^{2/3}\kappa_i^{-5/3}) + \psi^{(0)}(d_{\text{bin},i}/2) - \log(d_{\text{bin},i}/2) \tag{11}$$

and variance

$$\sigma^2_A(d_{\text{bin},i}) = 2/d_{\text{bin},i}. \tag{12}$$

Note that $\psi^{(0)}(d_{\text{bin},i}/2) - \log(d_{\text{bin},i}/2)$ is negligible for high degrees of freedom and comes from the expectation of a log χ^2 distribution, with $\psi^{(0)}$ representing a digamma function (e.g., Lee 2012). Equation (12) comes from standard error propagation of the standard deviation of a χ^2 variable through Eq. (10) (Thomson and Emery 2014a). Modeling the log of our spectrum as a Gaussian allows us to use a weighted least squares fit with the inverse variances as the weights (e.g., Thomson and Emery 2014a; Leo 1994). We solve for a and b in $\hat{y}_i = ax_i + b$, where

$$\hat{y}_i = \hat{A}_i + \log(d_{\text{bin},i}/2) - \psi^{(0)}(d_{\text{bin},i}/2) = \log(\hat{S}_i) + \log(d_{\text{bin},i}/2) - \psi^{(0)}(d_{\text{bin},i}/2), \tag{13}$$

$$x_i = \log(\kappa_i), \tag{14}$$

and the errors in the fit are weighted by $\sigma^2_y(d_{\text{bin},i}) = 2/d_{\text{bin},i}$. Note that we have defined $\hat{y}_i$ in Eq. (13) with correction terms such that the expected value for y, $\overline{y}^{\kappa_i}(d_{\text{bin},i})$ is then

$$\overline{y}^{\kappa_i}(d_{\text{bin},i}) = \overline{A}^{\kappa_i}(d_{\text{bin},i}) + \log(d_{\text{bin},i}/2) - \psi^{(0)}(d_{\text{bin},i}/2), \tag{15}$$

which, plugging in Eq. (11), becomes

$$\overline{y}^{\kappa_i}(d_{\text{bin},i}) = -\frac{5}{3}\log(\kappa_i) + \log(C_1'\epsilon^{2/3}). \tag{16}$$

Therefore, if we require a and b in $\hat{y}_i = a\log(\kappa_i) + b$ to match the linear fit to theory, then $a = -5/3$ and $b = \log(C\epsilon^{2/3})$. Using these values, and applying a least squares fit to $\hat{y}_i$ (as in Leo 1994), we can use a to determine the location of the inertial subrange (as described in section 5b) and we can estimate the dissipation as

$$\epsilon = [\exp(b)/C_1']^{3/2}. \tag{17}$$

To confirm that our approach is satisfactory, we generate a test inertial spectrum, add random χ^2 noise, and calculate the dissipation from the spectrum. When repeating this many times, we find negligible differences between our approach and the full nonlinear approach of Bluteau et al. (2011), even for degrees of freedom as low as 10. We then proceed with confidence that our simplification to Gaussian uncertainties will not affect our final result.

b. Where to fit

When deciding what portion of the spectrum to fit as the inertial subrange, a variety of approaches have been suggested. Trowbridge and Elgar (2001) and Feddersen (2010) fit to a constant, prespecified frequency range, eliminating spectra that did not satisfy a misfit cutoff. Jones and Monismith (2008) and Bluteau et al. (2011) fit to all portions of the spectrum longer than a minimum frequency range, and select the fit that has the smallest misfit. Bluteau et al. (2011) use large-scale flow properties and the Kolmogorov length scale to define bounds to restrict where the final fit could be located, employing an iterative procedure to account for the Kolmogorov length scale being a function of the dissipation itself.

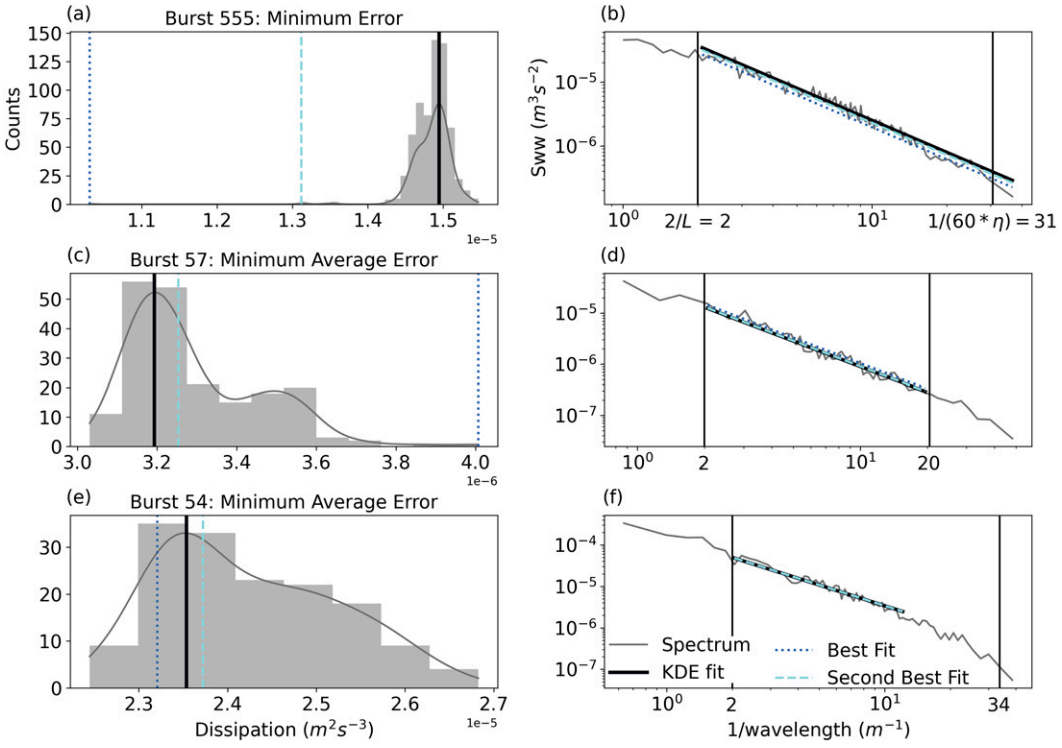

FIG. 4. Examples of inertial subrange fitting method. (left) Histograms from three different bursts of dissipation values from each individual fit. The thin gray lines show Gaussian kernel-density estimates. The vertical lines represent dissipation values from different fitting schemes: the most likely dissipation value selected as in appendix C, step 13 (black), the best fit according to metrics defined in row titles (dotted dark blue), and the second-best fit according to the same metric (dashed light blue). (right) Mean spectrum from each burst, with corresponding fits plotted over all wavenumbers used to produce the histograms (i.e., any wavenumbers over which there was a successful fit). The thin vertical lines denote theoretical wavenumber limits based on the generation and Kolmogorov scales. Note that the bottom row shows the same burst as in Fig. 2, all axes' limits are different and the x axis of the right column is nonradian wavenumber $[\kappa/(2\pi)]$.

To quantify the misfit, Jones and Monismith (2008) use the total squared error from the least squares fit of a −5/3 line to the spectrum. Feddersen (2010) uses two misfit tests. First, after performing a least squares fit of a line with unspecified slope, Feddersen (2010) uses the difference of the fitted slope from −5/3, normalized by the uncertainty of the fitted slope. Second, Feddersen (2010) uses the ratio of the horizontal and vertical spectra to test for isotropic turbulence. Bluteau et al. (2011) use the same condition as Jones and Monismith (2008), with an additional criterion based on the maximum absolute deviation of their fit and a χ^2 distribution.

After performing several tests with the semi-idealized model, we find a common problem with all of the methods suggested above. The statistical uncertainties from our fitting method consistently underestimate the error of our calculated dissipation values compared to the dissipation used to create the initial ideal spectrum. This is because slight changes in noise can change what portion of the spectrum is fit and dramatically change the final dissipation value. The first two rows of Fig. 4 show two bursts where the best and second best fits for two different measurements of best fit result in dramatically different dissipation values. The first row uses the total error as used by Jones and Monismith (2008) while the second uses the average error because we found that the total error preferentially selected shorter fits. We therefore conclude that specifying a specific portion of the spectrum as the best fit can introduce unnecessary uncertainty. To solve this problem, we instead calculate a kernel density estimate of the dissipation values from all possible valid fits to determine the most likely true dissipation value, as shown in Fig. 4 and described in the following paragraphs.

We employ a similar iterative procedure as used by Bluteau et al. (2011), starting by identifying the theoretical bounds of the inertial subrange. As a lower bound, we set

$$\kappa_{\min} = 2 \times \frac{2\pi}{L}, \tag{18}$$

where L is the generation length scale, which we set to 1 m based on our shallow water depth. As an upper bound, we set

$$\kappa_{\max} = \frac{1}{60} \times \frac{2\pi}{\eta}, \tag{19}$$

where η is the Kolmogorov length scale defined in Eq. (3). These theoretical bounds can be seen by the vertical lines in Fig. 4b. The factor of 1/60 introduced in Eq. (19) is an

empirically derived value from laboratory experiments (e.g., Pope 2000); however, the factor of 2 in Eq. (18) is relaxed from the empirical value of 6. This relaxation is supported in part because the laboratory experiments covered low Reynolds numbers when compared to our measurement conditions and because the water column depth changes over the sampling period. This relaxation of the low wavenumber scaling factor can also be thought of as using a larger generation length scale and reflects the fact that we likely have longer horizontal turbulent length scales than implied by the depth of the water column. Such extended horizontal scales are common in shallow water, as discussed by Stacey et al. (1999), Kirincich et al. (2010), and Amador et al. (2022). We also verified that this relaxation was appropriate because roughly a quarter of our final dissipation measurements exhibit a well-defined inertial subrange well past the stricter cutoff.

For the iterative procedure, we begin by guessing a high dissipation of $\epsilon_{\text{test}} = 10^{-3}$ to start with a very high upper bound (appendix C, step 1). After setting the bounds based on Eqs. (18) and (19), we take every possible continuous subsection of the spectrum within those bounds that contains at least 10 data points and wavenumbers that span at least a quarter of a decade (appendix C, step 2). For each of these subsections, we fit a line as described above, first calculating the slope and slope error (Leo 1994) to determine what subsections to eliminate (as described below), then fixing the slope at $-5/3$ to obtain the y intercept and calculate the resulting dissipation [Eqs. (13)–(17) and appendix C, steps 3–9].

We do not use the total error of the fit to eliminate subsections because we find that the spectrum noise varies between bursts, requiring arbitrary cutoffs greater than the statistical errors suggested by our degrees of freedom. We also do not use the comparison of the vertical to horizontal spectra because it would require removing noise from the horizontal data. The noise is primarily Doppler white noise from the instrument, which varies with flow conditions, and is much larger for the horizontal velocity components than the vertical (e.g., Voulgaris and Trowbridge 1998). The most common method for evaluating noise levels requires our spectra to flatten, and alternatives, such as those suggested by Durgesh et al. (2014), require stationary data. Neither of these are true for our data, so we choose to focus on the lower noise vertical velocity data. Similar to the misfit used by Feddersen (2010), we choose to ignore spectrum subsections where $-5/3$ lies outside the 95% confidence interval of the fitted slope (appendix C, step 7).

In addition, some of our calculated spectra (~6%) appear to have a peak at around 1.3 Hz, which maps to a 1–10 cm wavelength with our frozen turbulence conversions (the exact location varied). These peaks tended to appear at periods when the instruments were measuring close to the surface, and the shallower instrument measured a much higher energy than the deeper instrument. These observations, along with the additional presence of the peaks in the pressure spectrum, leads us to believe that they were caused by local wind-generated waves. To avoid these peaks causing inaccurate fits, we take the residual of each fit,

$$y_i^{\text{res}} = \hat{y}_i - \left(-\frac{5}{3}x_i + b\right), \tag{20}$$

and calculate the prominence of each peak as well as the prominence of each minimum (or, equivalently, the prominence of each peak of the negative of the residual) using the SciPy peak_prominences function (Virtanen et al. 2020) (appendix C, step 10). If the largest prominence is greater than 0.8 (note, because we are working in log space, this corresponds to a little over a factor of 2), we assume that the corresponding spectrum subsection contains a peak and ignore that dissipation value (appendix C, step 11). We want to note that for datasets in deeper water, we expect that these high-frequency waves will not be present. However, most other datasets will not be in an estuary that filters out larger surface gravity waves. In such conditions, we recommend using the Shaw and Trowbridge (2001) or a similar method to remove the high-frequency waves from the vertical velocity data before calculating the spectrum with these or any other method.

After eliminating invalid subsections, we treat the remaining dissipation values as samples from a probability distribution. Using the SciPy gaussian_kde function (Virtanen et al. 2020) and Scott's rule for determining bandwidth (Scott 2015), we calculate a 1D Gaussian kernel-density estimate (KDE) of the $\log_{10}$ of the dissipation values (appendix C, step 12). Scott's rule sets a bandwidth, $h = m^{-1/5}$, based on the number of points used, m, that estimates the theoretical bandwidth that would produce a zero bias density estimate (Virtanen et al. 2020; Scott 2015). The KDE itself is a smoothed estimate of the continuous probability density function for the dissipation. Because it is continuous, we find the dissipation value corresponding to the peak probability of the KDE and select the calculated dissipation value closest to that peak. (appendix C, step 13). Three examples of a histogram and KDE can be seen in Fig. 4a, including the burst shown in Fig. 2, which can be seen in the last row.

This dissipation allows us to redefine our theoretical maximum wavenumber bound using Eq. (19) (appendix C, step 14). If any of the spectrum subsections used have more than one wavenumber outside of the new bounds, we then repeat the process until the subsections used to estimate the dissipation value satisfy the bounds set by that dissipation value (appendix C, step 15). We specify having more than one wavenumber outside of the bounds in order to allow fits slightly outside of theoretical bounds, similar to the approach taken by Bluteau et al. (2011). In addition, if an iteration leads to a higher dissipation value than the last, we force another iteration (appendix C, step 16) to include the new sections allowed by the relaxed upper bound (forcing a hard cutoff after five iterations with an increasing dissipation to avoid infinite loops).

6. Error analysis

Inherent in our dissipation calculation methods are several sources of error, both statistical and systematic. Tracking each individual source of error is difficult; however, here we present three different approaches to identify and quantify many of these errors and biases. This approach not only provides detailed error quantification; it also addresses the generalization of our methods by indicating which parameters may need

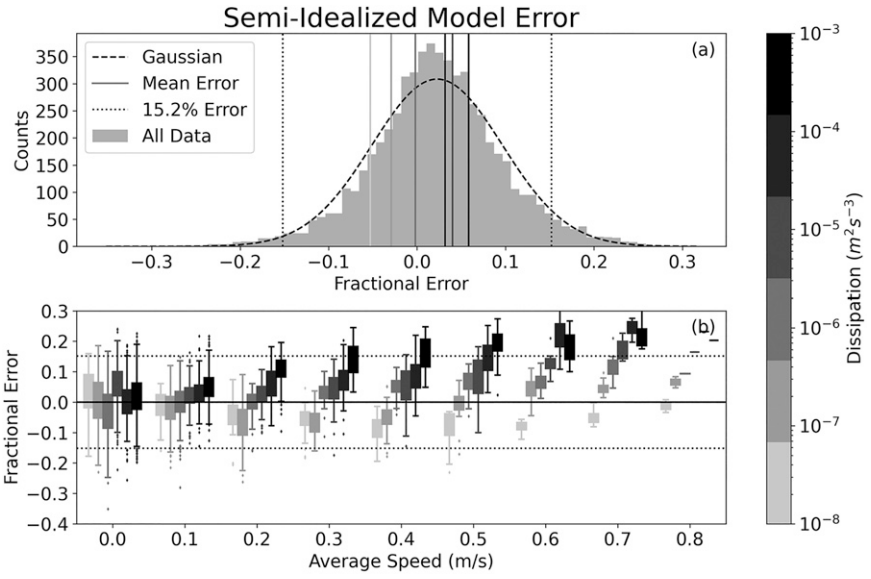

FIG. 5. Results of semi-idealized model–based error analysis. (a) Histogram of all errors, along with a Gaussian of the same mean and standard deviation. The vertical lines denote the mean errors of all tests derived from a single initial ideal dissipation value (by shading). The thin dotted vertical lines represent the 15.2% error when centered around 0. (b) Fractional errors scattered by the corresponding mean advection speed and colored by the initial ideal dissipation value with darker colors representing higher dissipation. The horizontal dotted black lines again represent the 15.2% error.

adjustment under different environmental conditions and identifying overall areas for future improvement.

a. Semi-idealized model–based errors

For testing our spectrum calculation and dissipation fitting methods, we can use our semi-idealized model to reveal general patterns and determine an estimate of the total error.

We start by establishing the idealized spectrum and associated spatial data for dissipation values of $\epsilon_{\text{ideal}} = 10^{-8}$, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, and 10^{-3} m^2 s^{-3}, using Eq. (2). For each dissipation value, we then use the observed along channel velocity of each 30 min burst from all 3 ADVs to generate semi-idealized temporal data as in section 4b, calculate a wave corrected spectrum as in section 4d, and calculate a dissipation value as in section 5. This gives us semi-idealized dissipation values with our methods (ϵ_{calc}) for every data burst and each dissipation level used for generating an ideal spectrum. To compare across dissipation levels, we calculate the fractional error as

$$\text{error} = \frac{\epsilon_{\text{calc}} - \epsilon_{\text{ideal}}}{\epsilon_{\text{calc}}}. \tag{21}$$

We choose to use the observed dissipation in the denominator rather than the ideal dissipation to better match what is obtained from real data. As Fig. 5a shows, a histogram of these errors from all dissipation values tested and all measured advection velocities collapse to close to a Gaussian distribution with a mean error of 2.3% and a standard deviation of 7.4%. To determine a statistical error on our measurements, we use the half-width of the middle 95% of the error distribution in Fig. 5a. This gives us an uncertainty on our dissipation measurements of 15.2%, or slightly more than twice the standard deviation. The 15.2% is our best estimate of the methodological error of our dissipation values.

Breaking out these test dissipation calculations by the ideal dissipation value used and the average advection speed, Fig. 5b further shows that our method introduces some systematic biases. Generally, higher advection speeds and higher dissipation values introduce a bias high. This bias also introduces the skewness of the distribution compared to the Gaussian in Fig. 5a. The source of these high biases is aliasing, which, because we convert each 10 s spectrum to wavenumber space with a different advection velocity, gets spread out across the entire spectrum. We do correct for aliasing in Eq. (7), but we only limit the aliasing bias to less than 25%. We do not limit the bias further because our tolerance test in Table 1 for the Nyquist threshold factor shows that using a stricter cutoff would lead to more noise. We also find that the strict cutoff would lead to 128 fewer final dissipation values. So, we instead note that a majority of these biases are within the 15.2% error and are accounted for in our data analysis except for particularly high dissipation values with high advection speeds. We also observe that while the bias high is consistent for very high dissipation values [$\mathcal{O}(10^{-3})$ m^2 s^{-3}], such high dissipation values are rarely observed in our data.

There is also a low bias at low dissipation values. Because of our shallow water, at particularly low dissipation values, the idealized inertial subrange practically does not exist. As a result, our fitting algorithm for the idealized spectrum is fitting

to the rolloff portions of the spectrum when our dissipation values are low, which biases these low dissipation values lower. This bias is generally smaller than the bias high at high dissipation, so further corrections are not required. In addition, these low biases occur for particularly low dissipation values (on the order of 10^{-8} m^2 s^{-3}), which are also rarely observed in our data.

b. Noise-based errors

Because of the ideal spectrum used, our semi-idealized model fails to capture two important sources of error. First, nonturbulent physical processes and errors in our despiking methods can introduce increased noise into the spectrum over the ideal χ^2 distribution. Second, white noise from the instrument can bias the entire spectrum high (e.g., Voulgaris and Trowbridge 1998).

To account for the former, our KDE dissipation fit approach provides a built in estimate of the variation introduced by spectrum noise. We take the half-width of the middle 95% of the final dissipation distribution used for estimating the KDE as an initial estimate of the uncertainty. This is effectively an estimate of the statistical uncertainty on our measurement. The average statistical uncertainty for our final dissipation estimates is around 7.8%, which is comparable the 15.2% methodological uncertainty, suggesting that these error estimates should be combined. However, these two uncertainties are not independent, because the semi-idealized model is already using the KDE dissipation fit approach. Calculating the corresponding statistical uncertainty of the semi-idealized model, we find an average contribution of 4.5% to the calculation of the methodological uncertainty, which is significant enough to make these uncertainty calculations dependent. Therefore, we record these two uncertainties separately and note that the summation in quadrature of the two as if they were independent provides an upper bound on the error while the maximum of the statistical or methodological uncertainty provides a lower bound on the error.

For the bias introduced by white noise, we first recall that we are using the vertical velocity component, which is the component with the lowest noise contribution for ADV's (e.g., Voulgaris and Trowbridge 1998). While we cannot identify an accurate noise level for many our bursts, about 20% of our bursts from all three ADVs do exhibit a spectrum flattening, equivalent to spectrum noise of typically 10^{-9}, and occasionally 10^{-8} m^2 s^{-3}. Subtracting this noise level from our spectra and recalculating dissipation values, we note that noticeable effects only occur for dissipation values close to the noise level and become negligible for dissipation values greater than 1.5 orders of magnitude larger than the noise value (dissipations around 10^{-7} m^2 s^{-3}). For the lower dissipation values, the bias can be on the same order of magnitude or larger than the upper bound of our error range, which is on average 18.7% for all of our data.

Because the white noise might bias our spectrum high while the semi-idealized model indicates a bias low at the same dissipation values, it is impossible to truly quantify the net bias affects on our low dissipation values. Ultimately the error on our low dissipation values [$\leq \mathcal{O}(10^{-7})$ m^2 s^{-3}] may be greater than the uncertainty we use. Therefore, we determine that care must be taken interpreting dissipation values below 10^{-7} m^2 s^{-3} in our particular dataset.

In summary, the total errors due to methodology and noise place an absolute lower bound on the error of 15.2% and an average upper bound on the error of 18.7%. Because the statistical error is sometimes greater than the systematic error, the lower bound for any given dissipation measurement is the maximum of systematic and statistical error, which on average for our dataset is 17.5%. We also find an occasional burst with significant statistical error that slightly increases our average errors. Thus, the average error could be reduced if a maximum statistical error limit were imposed. Finally, particularly high dissipation values [$\geq \mathcal{O}(10^{-3})$ m^2 s^{-3}] can be biased high by up to 25% due to aliasing and particularly low dissipation values [$\geq \mathcal{O}(10^{-7})$ m^2 s^{-3}] can have larger errors than calculated here due to a mixture of white noise and the effects of fitting to the rolloff of the inertial subrange.

c. Sensitivity analysis

In addition to the sources of errors mentioned above, our methods also introduce several changeable variables whose impact on the final results need to be understood. For each of these variables, we test them by selecting a high and low value with which to redo our analysis to understand how our final results on all 3 ADVs change. A summary of these changes can be seen in Table 1.

For the expanding cutoff despiking algorithm (section 3b), we use 3 predetermined constants. First, we choose to filter out the low-pass signal with a cutoff frequency of 1/20 Hz. Second, we incrementally expand the phase space cutoffs by 1% of the universal threshold. Third, we choose the stopping point of the expansion where the density of points between sequential cutoffs decreases by more than 95% of the previous value.

For our segmented spectrum calculation algorithm (section 4d), we use 8 predetermined constants. For separating turbulent from advection velocities, we use a frequency cutoff of 1/5 Hz and segment the data into 10-s-long segments to avoid wave contamination. When eliminating segments, we require the mean current to be larger than 5 times the turbulent standard deviation and the advection standard deviation to be less than 1.025 times the mean current. We further require that less than 1% of the points used for calculating a spectrum be replaced spikes. We then eliminate the two lowest wavenumbers and any wavenumbers where we expect the spectral values to be less than 4 times the spectral value at the Nyquist frequency assuming a $-5/3$ slope. Finally, we bin average with bin sizes of at least 50 points.

For the KDE-based inertial subrange fitting (section 5), we further introduce six predetermined constants. For determining the bounds on the inertial subrange, we multiply the generation length scale by a factor of 1/2. When identifying possible spectrum subsections to fit an inertial subrange to, we require the subsection to contain at least 10 data points and have wavenumbers that span at least a factor of 2.5

TABLE 1. Results of tolerance tests. Average change refers to the average of $(\epsilon_{new} - \epsilon_{original})/\epsilon_{original}$, allowing us to estimate any biases introduced by modifying the corresponding variable. Average absolute change refers to the same, but taking the absolute value before calculating the average. This allows us to identify potential changes, even if no average bias is introduced.

Methods section	Variable	Original value	New value	Average change (%)	Average absolute change (%)
Despiking	Low-pass signal filter frequency	1/20 Hz	1/30 Hz	0.3	7.5
			1/10 Hz	3.6	8.5
	Expansion step size factor	1%	0.5%	−10.3	13.4
			2%	21.9	24.1
	Expansion density change end	95%	90%	−1.8	5.0
			99%	1.2	2.5
Spectrum calculation	High/low separation frequency	1/5 Hz	1/10 Hz	−1.3	3.9
			2/5 Hz	2.9	4.9
	Segment size	10 s	5 s	12.5	18.23
			15 s	1.9	16.0
	Turbulence variance cutoff factor	5	2.5	3.7	6.9
			10	−8.8	12.0
	Wave variance cutoff	1.025	0.8	0.03	0.2
			1.2	0.02	0.05
	Unoriginal points allowed	1%	0.5%	−4.5	7.8
			2%	6.1	8.1
	No. of low wavenumbers removed	2	0	1.9	5.7
			4	2.7	6.8
	Nyquist threshold factor	4	2	0.5	6.7
			8	5.4	11
	Bin size	50 points	25 points	5.7	16.9
			75 points	2.2	5.4
Inertial subrange fitting	Generation length scale factor	0.5	0.25	1.2	3.5
			1	0.2	0.8
	Min data points in subrange	10	5	0.3	1.2
			15	1.2	1.9
	Min subrange wavenumber span	2.5	1.25	8.1	20.5
			5	−0.01	5.0
	Max peak prominence	0.8	0.7	0.6	2.2
			0.9	−0.1	1.5
	Slope confidence interval	0.95%	0.9%	0.6	1.8
			0.99%	0.2	2.0
	KDE bandwith scale factor	1.0	0.8	0.02	0.2
			1.2	0.02	0.2

(a quarter of a decade). For determining valid fits, we require −5/3 to be within the 95% confidence interval of the slope and that the maximum peak prominence of the data in the fit be smaller than 0.8. When determining the final dissipation value, we use Scott's rule with no alterations to determine the bandwidth for calculating the KDE (Scott 2015).

Reviewing the results in Table 1 shows that the dissipation is not very sensitive to a majority of the variables when compared to the 15.2% uncertainties we have already identified. One notable exception is the expansion step size factor in the despiking algorithm. For this particular variable, we tested smaller incremental changes than shown in Table 1 and found that there is a roughly consistent linear relationship between the change in expansion step size factor and the average dissipation change. This consistent and relatively strong relationship between expansion step size and dissipation bias indicates that careful consideration must be given to initial despiking to properly evaluate final dissipation uncertainty. We chose our particular step size by closely examining plots similar to Figs. 2a–d for many of our bursts, which does not guarantee the selection is appropriate for other datasets. Such sensitivity also points to an area for potential future methodological improvements, which we elaborate on in the discussion.

A few other variables introduce large average absolute changes. Altering segment size when calculating spectra seems to generally increase uncertainty in dissipation value. This reflects the fact that too short of a segment reduces the resolution of the inertial subrange while too long of a segment introduces stronger IG frequency bore driven bias. Similarly, reducing the bin size during the spectrum calculation introduces large dissipation changes, likely because of the presence of stronger noise in the final spectrum values due to lower degrees of freedom. Finally, reducing the minimum range between the minimum and maximum wavenumbers of the

spectrum subsections when fitting the inertial subrange introduces changes due to the fitting algorithm becoming much more sensitive to the high-frequency portion of the spectrum, which contains a large density of data points.

Two other variables with substantial sensitivities that might be of interest to readers are the turbulence variance cutoff and number of unoriginal points allowed. Both of these cutoffs were chosen to match cutoffs used in other methods. The turbulence variance cutoff is chosen to match that used by George et al. (1994), and shows sensitivity because relaxing this cutoff increases the number of segments with large variance included in calculating the spectrum. This increase in dissipation also likely includes segments where the frozen turbulence hypothesis is invalid, which is why we keep the George et al. (1994) cutoff. The number of unoriginal points allowed matches the fraction determined by Feddersen (2010) as the point past which the specific interpolation method used begins to matter. This sensitivity then tells us that our particular interpolation method would start to bias our dissipation high if we had more unoriginal points included in the spectrum calculation.

7. Discussion

In this section, we highlight a few key results that we believe should be considered in future work measuring turbulence using ADV data, including applicability of our methods to other types of datasets and important directions for future work.

For our despiking algorithm, we adapted the Islam and Zhu (2013) expanding cutoff phase space method to reduce the number of tunable parameters while still preventing the elimination of real data as spikes during time periods of intermittent variance. While our final algorithm is simplified relative to Islam and Zhu (2013), our tolerance tests show significant sensitivity to one parameter. Islam and Zhu (2013) also show significant sensitivity where the mean velocity changes by over 15% with their bandwidth tests. This is not equivalent to our sensitivity tests on the final dissipation value. For a more direct comparison, we find that our mean velocities change by less than 1% for all of our despiking sensitivity tests, albeit over a smaller parameter change. With this in mind, our despiking algorithm is a successful simplification of the Islam and Zhu (2013) method that results in a relatively robust algorithm for use in complex environmental conditions.

While our mean velocities are relatively stable, the sensitivity of our final dissipation values highlights the inherent subjectivity of despiking ADV data and how that subjectivity can introduce hidden biases in turbulence calculations. The subjectivity can be seen more clearly when looking at Fig. 2f, where our expanded cutoffs appear to eliminate some valid data for this particular data burst. So, while we find the presented algorithm to be the most reliable in the presence of large IG frequency bores, or any other process that leads to unpredictably intermittent velocity variance, we recommend extra caution when determining the uncertainties of turbulence calculations when using any despiking algorithm. Further, we look forward to future improvements in ADV despiking by other researchers. We believe that the expanding cutoff approach has inherent limitations and that another promising approach would be an adaptive algorithm that can identify varying length bursts of data with consistent variance over which to apply despiking rather than sticking with a predefined burst length.

The segmented spectrum calculation approach here is very specific to the presence of low-frequency velocity oscillations. However, the result that our nonlinear velocity oscillations can change the shape of a frequency spectrum, even at high frequencies, suggests that careful consideration should be given to the calculation of the spectrum before converting to wavenumber space in situations with nonlinear processes. Therefore, we recommend that future surf-zone turbulence measurements, or measurements in the presence of other nonlinear processes, carefully examine the assumptions of the spectrum calculation methods used and consider applying our segmented spectrum calculation in the case of nonlinear and nonstationary velocity oscillations.

Our method for fitting an inertial subrange to a wavenumber spectrum was not specific to the presence of IG frequency bores. As a result, it can apply to any situation where dissipation estimates from spectra are performed, and is likely the most widely applicable result of this paper. While the χ^2 adjusted, weighted least squares approach helps better account for non-Gaussian errors, we believe the more important improvement is the use of all possible fits to collectively determine the final dissipation value through a KDE estimation. This avoids any reliance on a single goodness of fit measurement, which could be more easily affected by random variations not captured by standard error analysis.

We also want to briefly mention that the presence of horizontal length scales in the inertial subrange that are longer that the water depth theoretically implies a level of anisotropy. To our knowledge, any field deployment observing turbulent dissipation generated by bottom boundary layers in shallow water would find similar levels of anisotropy. However, we are not aware of any existing work to understand the impacts that this has on inertial subrange derived dissipation values. Therefore, while we believe our measurements are accurate and comparable to past methods for measuring turbulent dissipation, we also believe that future work examining the impacts of anisotropic length scales on the inertial subrange would be valuable.

When we put all three methods together, we successfully produce representative dissipation values that are consistent between collocated instruments that are vertically separated by 20 cm, lower and upper, and can be seen in Fig. 6. These values display expected high values during strong tidal flows, and even show elevated levels in the presence of large IG wave oscillations. The typical 15.2% uncertainty gives us confidence in our results. We note the caveat that dissipation levels around 10^{-7} m^2 s^{-3} and below possibly have higher errors; however, these make up a relatively small fraction of our data. Specifically, 1.4% of the dissipation values from the two collocated ADVs shown in Fig. 6 and 15.1% of all the dissipation values when including the third ADV. The percentage

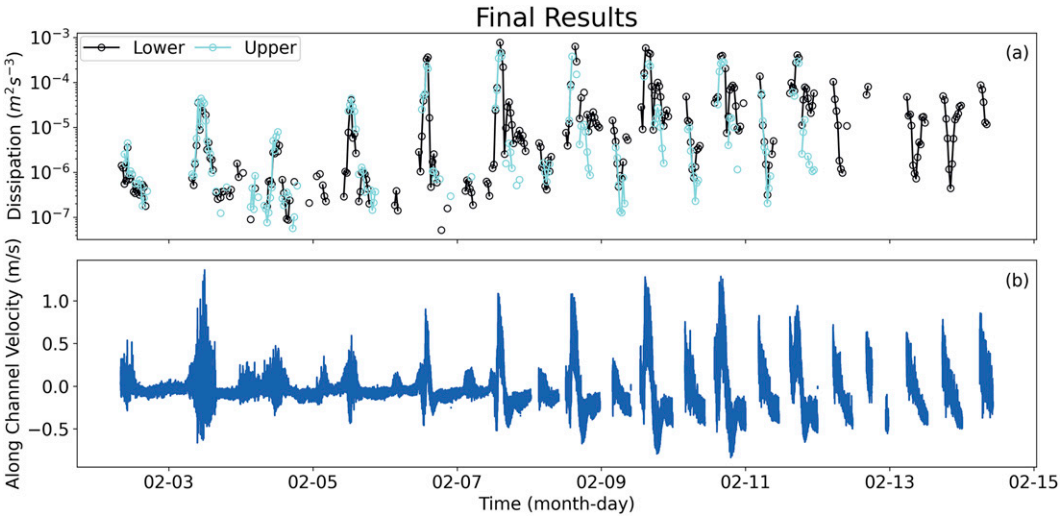

FIG. 6. Final results of despiking, spectrum calculation, and dissipation fitting algorithms combined. (a) The dissipation values from each of the two collocated instruments that are separated vertically by 20 cm. (b) The despiked, unfiltered, along-channel velocity from the lower ADV.

increases when we include the third ADV because it is located farther upstream in deeper water, where the turbulence is generally weaker.

Finally, our goal for this work was to measure turbulent dissipation in the presence of large IG frequency bores. However, by necessity, our calculations were performed in 30 min bursts, resulting in representative dissipation values that are the average of instantaneous dissipation values raised to the two-thirds power, then raised to the three-halves power as in $[\sum^{N}(\epsilon^{2/3}/N)]^{3/2}$. When considering how the two-thirds power affects an average, the representative dissipation will consistently slightly underestimate the average dissipation. Therefore, it is important to note that one of our hypotheses, and a focus of future work, is that the IG frequency bores are increasing turbulence in bursts as the bore passes the measurement location. If this is true, then we are measuring a representative dissipation that is likely a lower bound on the turbulent energy produced by the IG bores.

8. Conclusions

Using data from a deployment in Los Peñasquitos Lagoon in February 2020, we develop several new algorithms to calculate reliable turbulent dissipation values in the presence of large IG frequency bores. We implement an expanded phase space cutoff to remove spikes from nonstationary data. Then we use a segmented spectrum approach to account for nonlinear bore biases in producing wavenumber spectra from a stationary instrument. Finally, we use a Gaussian KDE approach for fitting an inertial spectrum to avoid unnecessary errors arising from selecting a single portion of the spectrum to fit to for calculating dissipation values. The final 30 min representative dissipation values for two vertically separated, collocated ADVs, using these methods, can be seen in Fig. 6. The code and data presented in this paper have also been published to the University of California, San Diego, Library Digital Collections (Wheeler et al. 2023) and to GitHub.

The despiking and spectral estimation methods introduced here can be applied in other situations where nonlinear oscillations and nonstationary data limit the application of more standard approaches. The KDE dissipation calculation approach can be applied to any dataset where the inertial subrange must be located and fit to and provides a robust methodology to finding an optimal fit. The error analysis presented provides a rigorous approach to assessing dissipation calculations and highlights the need for further careful consideration of despiking when measuring turbulent dissipation with ADVs. Finally, successful observations of representative dissipation values under large IG frequency bores opens the door for future studies to understand turbulent processes and the potential importance of these bores in nearshore and estuarine systems.

Acknowledgments. We thank members and volunteers of the Giddings and Merrifield labs and the Center for Coastal Studies field crew at UCSD for their help with data collection. We thank the Los Peñasquitos Lagoon Foundation (especially Mike Hastings), California State Parks (especially Cara Stafford, Darren Smith, and Reinhard Flick), and the Tijuana River National Estuarine Research Reserve (especially Jeff Crooks) for helping to facilitate this research and providing permits. We also thank Geno Pawlak, Pascal Polonik, and Yunyi Zhang at UCSD and Johanna Rosman at UNC for their valuable feedback during discussions about this project. Finally, we thank two anonymous reviewers whose comments substantially helped the presentation of this methodology. This material is based upon work supported by NOAA's National Centers for Coastal Ocean Science (NCCOS) Competitive Research Program under Award NA18NOS4780172 to SIO and the National Science Foundation Graduate Research

Fellowship Program under Grant DGE-2038238. This work was also partially funded by California (CA) Department of Parks and Recreation Division of Boating and Waterways Oceanography Program under Contract C1670005 with SIO. This is CHRP Contribution Number 258. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Data availability statement. The data for the two ADVs presented in Fig. 6, as well as the code for the algorithms and graphs presented here, are available through the UCSD Library Digital Collections at the DOI https://doi.org/10.6075/J0J67H27. A GitHub repository for the code used has also been created at https://github.com/dcwheeler/IG-Dissipation-Processing-Code.

APPENDIX A

Expanding Cutoff Despiking Algorithm

1) Use linear interpolation to fill any NaNs in data (from correlation and SNR-based filtering or time periods where the instrument was out of water), and store the location of those NaNs.
2) Remove low-pass filtered signal to avoid contamination from large nonturbulent signals. In our case, we found that using a 1/20 Hz, fourth-order Butterworth filter effectively eliminated contamination from IG waves without introducing negative effects from filtering particularly large spikes.
3) Use central differences to calculate the first ($u^{(1)}$) and second ($u^{(2)}$) derivatives for locating each point in phase space as in Goring and Nikora (2002), their Eqs. (7) and (8), where in their notation the first and second derivatives are Δu and $\Delta^2 u$, respectively.
4) Calculate the universal threshold cutoffs c^0 using the total number of data points n and the standard deviations σ for the corresponding derivative as in Goring and Nikora (2002), their Eq. (2):

$$c^0_{u^{(k)}} = \sqrt{2\ln n} \times \sigma_{u^{(k)}}, \tag{A1}$$

where k is denoting the number of derivatives.
5) Define the ith expanded cutoff for each dimension as

$$c^i_{u^{(k)}} = (1 + 0.01i) \times c^0_{u^{(k)}}. \tag{A2}$$

6) Starting with u–$u^{(1)}$ space, as a first step before calculating the elliptical ring density, define the number of points within the ith expanded ellipse as

$$P^i_{u-u^{(1)}} = \sum_{j=1}^{n} \frac{u_j^2}{c_u^{i\,2}} + \frac{u^{(1)2}_j}{c^{i\,2}_{u^{(1)}}} \le 1, \tag{A3}$$

where j is an index that iterates over all points.
7) Then, for the ith expanded ellipse, define the elliptical ring density of points in the ring between sequential cutoffs as

$$\rho^i_{u-u^{(1)}} = \frac{P^i_{u-u^{(1)}} - P^{i-1}_{u-u^{(1)}}}{\pi \times (c^i_u c^i_{u^{(1)}} - c^{i-1}_u c^{i-1}_{u^{(1)}})}. \tag{A4}$$

8) To determine when to stop expanding the cutoff, first calculate the fractional change in elliptical ring density between sequential cutoffs as

$$\Delta\rho^i_{u-u^{(1)}} = \frac{\rho^{i-1}_{u-u^{(1)}} - \rho^i_{u-u^{(1)}}}{\rho^{i-1}_{u-u^{(1)}}}. \tag{A5}$$

9) Then, determine the last expanded cutoffs for u–$u^{(1)}$ space, $C^u_{u-u^{(1)}}$ and $C^{u^{(1)}}_{u-u^{(1)}}$, as the cutoffs before the fractional ring density change exceeds 0.95:

$$C^{u^{(k)}}_{u-u^{(1)}} = c^i_{u^{(k)}} | \Delta\rho^{i+1}_{u-u^{(1)}} \ge 0.95. \tag{A6}$$

10) Mark any points where the velocity exceeds the final $C^u_{u-u^{(1)}}$ cutoff as spikes. This is a first pass, as described by Parsheh et al. (2010), to avoid large spikes influencing the derivatives at nearby data points and causing valid data to look like more spikes.
11) If spikes are found, skip to step 18 for replacing those spikes. If no spikes are found, proceed to step 12 to use the full phase space detection method.
12) Mark any points that are outside the ellipse with major and minor radii defined by $C^u_{u-u^{(1)}}$ and $C^{u^{(1)}}_{u-u^{(1)}}$ as spikes.
13) Repeat steps 6–9 and 12 for $u^{(1)}$–$u^{(2)}$ space. Note, steps 10 and 11 are skipped because the initial, velocity-only-based spike detection is only done for the u–$u^{(1)}$-based cutoffs.
14) Because of the correlation between the velocity and the second derivative, calculate the principal axis rotation angle θ between the velocity and second derivative as in Goring and Nikora (2002), their Eq. (9).
15) Rotate u–$u^{(2)}$ space to α–β space with the principal axis rotation angle, where

$$z = (u + i \times u^{(2)}) \times e^{-i\theta}, \tag{A7}$$

$$\alpha = \mathrm{Real}(z), \tag{A8}$$

and

$$\beta = \mathrm{Imag}(z). \tag{A9}$$

16) Calculate the initial cutoffs for the rotated ellipse using the solutions to the equations found in Goring and Nikora (2002), their Eqs. (10) and (11), as

$$c_\alpha = \sqrt{\frac{(c_u \cos\theta)^2 - (c_{u^{(2)}} \sin\theta)^2}{\cos^4\theta - \sin^4\theta}}, \tag{A10}$$

and

$$c_\beta = \sqrt{\frac{(c_{u^{(2)}})^2 - (c_\alpha \sin\theta)^2}{\cos^2\theta}}. \tag{A11}$$

17) Repeat step 5 for the c_α^i and c_β^i cutoffs, then repeat steps 6–9 and 12 for α–β space to complete the full phase space spike detection.
18) Replace each spike with the value of the last valid point.
19) If spikes were identified, return to step 3 in order to repeat the phase space spike detection and spike replacement until no more spikes are detected. On rare occasions, this loop will continue for a very long time, so we recommend a maximum number of iterations allowed (we use 100).
20) Record the location of any spikes that were replaced through all iterations, add back the low-frequency signal, return NaNs removed in step 1, and finish.

APPENDIX B

Segmented Spectrum Algorithm

1) Linearly interpolate any NaNs in the vertical and horizontal velocity data.
2) Separate the high-frequency turbulent velocities (u') from all other lower-frequency advection velocities ($\tilde{u}$) with an appropriate filter. In our case, we use a 1/5 Hz, fourth-order Butterworth filter that separates low-frequency advection velocities $\tilde{u}$, which include IG and lower-frequency oscillations, from the high-frequency turbulent velocities ($u' = u - \tilde{u}$). These separated signals will be used for identifying bad segments and correcting the magnitude of the spectrum.
3) Split the vertical velocity data along with both the low and high-frequency horizontal velocity data into short segments with N_{seg} data points each such that the velocity varies little over each segment. We found that 10 s worked well for our data, which, with our 16 Hz sampling rate, gives $N_{\text{seg}} = 160$.
4) Calculate the average low-frequency horizontal advection velocities, $U_{\text{seg}} = \sum_{i=1}^{N_{\text{seg}}} \tilde{u}^i/N_{\text{seg}}$ and $V_{\text{seg}} = \sum_{i=1}^{N_{\text{seg}}} \tilde{v}^i/N_{\text{seg}}$, for each segment.
5) Calculate the advection speed for each segment as $c_{\text{seg}} = \sqrt{U_{\text{seg}}^2 + V_{\text{seg}}^2}$.
6) Calculate the advection direction for each segment as $\theta_{\text{seg}} = \tan^{-1}(V_{\text{seg}}/U_{\text{seg}})$.
7) Rotate the low- and high-frequency horizontal velocities of each segment into the mean advection direction, keeping the primary components as $\tilde{\gamma}$ and γ'.
8) Calculate the standard deviations of the low- and high-frequency rotated primary velocities for each segment ($\sigma_{\text{seg}}^{\tilde{\gamma}}$ and $\sigma_{\text{seg}}^{\gamma'}$).
9) Eliminate any segments that fail to satisfy either of the relations

$$\sigma_{\text{seg}}^{\tilde{\gamma}} \le 1.025 c_{\text{seg}} \tag{B1}$$

or

$$\sigma_{\text{seg}}^{\gamma'} \le c_{\text{seg}}/5. \tag{B2}$$

10) Count the number of unoriginal data points in each segment, $n_{\text{seg}}^{\text{bad}}$.
11) Eliminate the segment with the largest $n_{\text{seg}}^{\text{bad}}$ until the total number of bad points is less than 1% of the total remaining points:

$$\sum_{i=1}^{k} n_i^{\text{bad}} < 0.01 N_{\text{seg}} l, \tag{B3}$$

where l is the number of segments left.
12) For each remaining segment, detrend, multiply by a Hanning window, Fourier transform, and calculate a power spectrum of the vertical velocity in frequency space $S_{ww}(\omega)$, where ω is the radian frequency. Thomson and Emery (2014b) provides a good discussion for the effects of detrending and windowing on spectra.
13) Convert each remaining spectrum to wavenumber space and correct magnitude as

$$S_{ww}(\kappa) = \frac{S_{ww}(\omega) c_{\text{seg}}}{I(\sigma_{\text{seg}}^{\tilde{\gamma}}/c_{\text{seg}})} \tag{B4}$$

and

$$\kappa = \frac{\omega}{c_{\text{seg}}}, \tag{B5}$$

where I is calculated according to Eq. (4).
14) Eliminate the two lowest wavenumber spectral values for each segment to avoid low-frequency bias.
15) For each segment, eliminate any spectral values for wavenumbers that are higher than

$$\kappa_{\text{cutoff}} = (\kappa_{\text{max}}^{-5/3} \times 4)^{-3/5}, \tag{B6}$$

where κ_{max} is the highest wavenumber estimated for that spectrum. This avoids significant aliasing from frequencies above the Nyquist frequency.
16) Sort each remaining spectral value from all remaining segments in increasing order by wavenumber.
17) Group sorted spectral values as evenly as possible into $M = \text{floor}(N/50)$ bins, where N is the total number of spectral values left.
18) Calculate the mean spectral value for each bin. The degrees of freedom for these average estimates is then $2N_{\text{bin}}$, where N_{bin} is the number of spectral values in the bin.
19) Calculate the corresponding wavenumbers of the averaged spectral values, assuming an inertial subrange, using Eq. (8).

APPENDIX C

KDE Dissipation Fit Algorithm

1) Calculate the minimum and maximum wavenumber cutoffs using Eqs. (18) and (19), assuming a dissipation of 10^{-3} m^2 s^{-3}. This gives an initial estimate of where the inertial subrange can be.
2) Find all possible continuous subsections of the spectrum with wavenumbers (κ_i, $i \in [1, N_{\text{sec}}]$) inside the wavenumber cutoffs ($\kappa_{\text{min}} < \kappa_i < \kappa_{\text{max}}$), with at least 10 spectral

values ($N_{sec} \geq 10$), and for which the wavenumbers span at least a quarter of a decade ($k_{N_{sec}}/k_1 \geq 2.5$). Here, N_{sec} refers to the total number of spectral values in a given subsection and i refers to a specific spectral value in the subsection.

3) For each subsection, convert spectrum to log space (x–y space) using Eqs. (13) and (14).
4) For each subsection, calculate the variance, $\sigma_y^2(d_{bin,i}) = 2/d_{bin,i}$, from the degrees of freedom $d_{bin,i}$ of each converted spectral value ($\hat{y}_i$).
5) For each subsection, calculate a weighted least squares fitted slope (a_{sec}, where the subscript refers to the specific subsection) to x_i and $\hat{y}_i$, using the inverse of the variances from step 4 as the weights. We follow the approach used by Leo (1994).
6) Calculate the error on the fitted slope ($\sigma_{a_{sec}}$). Again, we follow Leo (1994).
7) Discard any subsection where $|a_{sec} + 5/3|/\sigma_{a_{sec}} \geq 1.960$, with 1.960 coming from the 97.5th percentile of the standard normal distribution.
8) For each remaining subsection, assume a slope of −5/3, and use the weighted least squares approach to calculate the y intercept (b_{sec}).
9) Using those y intercepts, calculate dissipation values for each remaining subsection (ϵ_{sec}) following Eq. (17).
10) Calculate the largest peak or valley prominence, using the SciPy peak_prominences function (Virtanen et al. 2020), of the residual from Eq. (20) for each subsection.
11) Further discard any dissipation values from subsections where the maximum prominence is ≥0.8. The remaining dissipation values are now referred to as valid dissipation values.
12) Using the valid dissipation values (ϵ_{sec}^{valid}), calculate a Gaussian KDE of $\log_{10}\epsilon_{sec}^{valid}$.
13) Select the valid dissipation value that is closest to the peak of the Gaussian KDE as the calculated dissipation value (ϵ_{calc}).
14) Repeat step 1 using the new dissipation value and Eq. (19) to obtain a new maximum wavenumber cutoff.
15) If any of the subsections used to calculate the dissipation values used for the Gaussian KDE have more than one wavenumber outside of the new wavenumber cutoffs, select the subset of the valid dissipation values coming from subsections entirely within the new cutoffs and repeat steps 12–14 until none of the subsections used have more than one wavenumber outside of the final cutoffs. Note, we do not need to repeat any of the other steps, because we have already calculated all possible potential dissipation values with the original nonstrict cutoffs. The repeat of steps 12–14 merely changes the number of valid subsections included in the calculations due to the refined cutoff values.
16) If the final dissipation value is larger than the previous dissipation value calculated, again repeat steps 12–15 until the final dissipation value is smaller than the previous dissipation value calculated. Note that this step has the potential to loop infinitely, so we stop if this condition is used five times.

REFERENCES

Amador, A., S. N. Giddings, and G. Pawlak, 2022: ADCP-based estimates of lateral turbulent Reynolds stresses in wavy coastal environments. *Limnol. Oceanogr. Methods*, **20**, 260–280, https://doi.org/10.1002/lom3.10485.

Becker, J. M., M. A. Merrifield, and H. Yoon, 2016: Infragravity waves on fringing reefs in the tropical Pacific: Dynamic setup. *J. Geophys. Res. Oceans*, **121**, 3010–3028, https://doi.org/10.1002/2015JC011516.

Behrens, D. K., F. A. Bombardelli, J. L. Largier, and E. Twohy, 2013: Episodic closure of the tidal inlet at the mouth of the Russian River—A small bar-built estuary in California. *Geomorphology*, **189**, 66–80, https://doi.org/10.1016/j.geomorph.2013.01.017.

Bertin, X., and M. Olabarrieta, 2016: Relevance of infragravity waves in a wave-dominated inlet. *J. Geophys. Res. Oceans*, **121**, 5418–5435, https://doi.org/10.1002/2015JC011444.

——, and Coauthors, 2018: Infragravity waves: From driving mechanisms to impacts. *Earth-Sci. Rev.*, **177**, 774–799, https://doi.org/10.1016/j.earscirev.2018.01.002.

——, D. Mendes, K. Martins, A. B. Fortunato, and L. Lavaud, 2019: The closure of a shallow tidal inlet promoted by infragravity waves. *Geophys. Res. Lett.*, **46**, 6804–6810, https://doi.org/10.1029/2019GL083527.

Bluteau, C. E., N. L. Jones, and G. N. Ivey, 2011: Estimating turbulent kinetic energy dissipation using the inertial subrange method in environmental flows. *Limnol. Oceanogr. Methods*, **9**, 302–321, https://doi.org/10.4319/lom.2011.9.302.

Burchard, H., and Coauthors, 2008: Observational and numerical modeling methods for quantifying coastal ocean turbulence and mixing. *Prog. Oceanogr.*, **76**, 399–442, https://doi.org/10.1016/j.pocean.2007.09.005.

Cea, L., J. Puertas, and L. Pena, 2007: Velocity measurements on highly turbulent free surface flow using ADV. *Exp. Fluids*, **42**, 333–348, https://doi.org/10.1007/s00348-006-0237-3.

Durgesh, V., J. Thomson, M. C. Richmond, and B. L. Polagye, 2014: Noise correction of turbulent spectra obtained from acoustic Doppler velocimeters. *Flow Meas. Instrum.*, **37**, 29–41, https://doi.org/10.1016/j.flowmeasinst.2014.03.001.

Feddersen, F., 2010: Quality controlling surf zone acoustic Doppler velocimeter observations to estimate the turbulent dissipation rate. *J. Atmos. Oceanic Technol.*, **27**, 2039–2055, https://doi.org/10.1175/2010JTECHO783.1.

——, 2012: Observations of the surf-zone turbulent dissipation rate. *J. Phys. Oceanogr.*, **42**, 386–399, https://doi.org/10.1175/JPO-D-11-082.1.

——, J. H. Trowbridge, and A. Williams III, 2007: Vertical structure of dissipation in the nearshore. *J. Phys. Oceanogr.*, **37**, 1764–1777, https://doi.org/10.1175/JPO3098.1.

George, R., R. E. Flick, and R. T. Guza, 1994: Observations of turbulence in the surf zone. *J. Geophys. Res.*, **99**, 801–810, https://doi.org/10.1029/93JC02717.

Goring, D. G., and V. I. Nikora, 2002: Despiking acoustic Doppler velocimeter data. *J. Hydraul. Eng.*, **128**, 117–126, https://doi.org/10.1061/(ASCE)0733-9429(2002)128:1(117).

Guza, R. T., and E. B. Thornton, 1982: Swash oscillations on a natural beach. *J. Geophys. Res.*, **87**, 483–491, https://doi.org/10.1029/JC087iC01p00483.

Harvey, M. E., S. N. Giddings, G. Pawlak, and J. A. Crooks, 2022: Hydrodynamic variability of an intermittently closed estuary over interannual, seasonal, fortnightly and tidal timescales.

Estuaries Coasts, **46**, 84–108, https://doi.org/10.1007/s12237-021-01014-0.

Islam, M. R., and D. Z. Zhu, 2013: Kernel density–based algorithm for despiking ADV data. *J. Hydraul. Eng.*, **139**, 785–793, https://doi.org/10.1061/(ASCE)HY.1943-7900.0000734.

Jesson, M., M. Sterling, and J. Bridgeman, 2013: Despiking velocity time-series—Optimisation through the combination of spike detection and replacement methods. *Flow Meas. Instrum.*, **30**, 45–51, https://doi.org/10.1016/j.flowmeasinst.2013.01.007.

Jones, N. L., and S. G. Monismith, 2008: The influence of white-capping waves on the vertical structure of turbulence in a shallow estuarine embayment. *J. Phys. Oceanogr.*, **38**, 1563–1580, https://doi.org/10.1175/2007JPO3766.1.

Kirincich, A. R., S. J. Lentz, and G. P. Gerbi, 2010: Calculating Reynolds stresses from ADCP measurements in the presence of surface gravity waves using the cospectra-fit method. *J. Atmos. Oceanic Technol.*, **27**, 889–907, https://doi.org/10.1175/2009JTECHO682.1.

Lee, P. M., 2012: *Bayesian Statistics: An Introduction*. 4th ed. John Wiley and Sons, 496 pp.

Leo, W. R., 1994: Statistics and the treatment of experimental data. *Techniques for Nuclear and Particle Physics Experiments*, Springer-Verlag, 81–113, https://doi.org/10.1007/978-3-642-57920-2_4.

Lin, C. C., 1953: On Taylor's hypothesis and the acceleration terms in the Navier-Stokes equation. *Quart. Appl. Math.*, **10**, 295–306, https://doi.org/10.1090/qam/51649.

Lumley, J. L., and E. A. Terray, 1983: Kinematics of turbulence convected by a random wave field. *J. Phys. Oceanogr.*, **13**, 2000–2007, https://doi.org/10.1175/1520-0485(1983)013<2000:KOTCBA>2.0.CO;2.

McDougall, T. J., and P. M. Barker, 2011: Getting started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox. SCOR/IAPSO Working Group 127 Rep., 28 pp.

McSweeney, S. L., J. C. Stout, and D. M. Kennedy, 2020: Variability in infragravity wave processes during estuary artificial entrance openings. *Earth Surf. Processes Landforms*, **45**, 3414–3428, https://doi.org/10.1002/esp.4974.

Nortek, 2018: The comprehensive manual for velocimeters. Nortek Support Doc., 139 pp., https://support.nortekgroup.com/hc/en-us/article_attachments/4551609336220/N3015-030_Comprehensive_Manual_-_Velocimeters_0322-compressed.pdf.

Parsheh, M., F. Sotiropoulos, and F. Porté-Agel, 2010: Estimation of power spectra of acoustic-Doppler velocimetry data contaminated with intermittent spikes. *J. Hydraul. Eng.*, **136**, 368–378, https://doi.org/10.1061/(ASCE)HY.1943-7900.0000202.

Pope, S. B., 2000: The scales of turbulent motion. *Turbulent Flows*, Cambridge University Press, 182–263, https://doi.org/10.1017/CBO9780511840531.008.

Razaz, M., and K. Kawanisi, 2011: Signal post-processing for acoustic velocimeters: Detecting and replacing spikes. *Meas. Sci. Technol.*, **22**, 125404, https://doi.org/10.1088/0957-0233/22/12/125404.

Rosman, J. H., and G. P. Gerbi, 2017: Interpreting fixed-location observations of turbulence advected by waves: Insights from spectral models. *J. Phys. Oceanogr.*, **47**, 909–931, https://doi.org/10.1175/JPO-D-15-0249.1.

Ruessink, B. G., M. G. Kleinhans, and P. G. L. van den Beukel, 1998: Observations of swash under highly dissipative conditions. *J. Geophys. Res.*, **103**, 3111–3118, https://doi.org/10.1029/97JC02791.

Rusello, P. J., and E. A. Cowen, 2011: Turbulent dissipation estimates from pulse coherent Doppler instruments. *2011 IEEE/OES 10th Current, Waves and Turbulence Measurements*, Monterey, CA, IEEE, 167–172, https://doi.org/10.1109/CWTM.2011.5759546.

Scott, D. W., 2015: Kernel density estimators. *Multivariate Density Estimation: Theory, Practice, and Visualization*, D. J. Balding et al., Eds., Wiley Series in Probability and Statistics, John Wiley and Sons, 137–216, https://doi.org/10.1002/9781118575574.ch6.

Sharma, A., A. K. Maddirala, and B. Kumar, 2018: Modified singular spectrum analysis for despiking acoustic Doppler velocimeter (ADV) data. *Measurement*, **117**, 339–346, https://doi.org/10.1016/j.measurement.2017.12.025.

Shaw, W. J., and J. H. Trowbridge, 2001: The direct estimation of near-bottom turbulent fluxes in the presence of energetic wave motions. *J. Atmos. Oceanic Technol.*, **18**, 1540–1557, https://doi.org/10.1175/1520-0426(2001)018<1540:TDEONB>2.0.CO;2.

Simpson, J. H., N. R. Fisher, and P. Wiles, 2004: Reynolds stress and TKE production in an estuary with a tidal bore. *Estuarine Coastal Shelf Sci.*, **60**, 619–627, https://doi.org/10.1016/j.ecss.2004.03.006.

Stacey, M. T., S. G. Monismith, and J. R. Burau, 1999: Observations of turbulence in a partially stratified estuary. *J. Phys. Oceanogr.*, **29**, 1950–1970, https://doi.org/10.1175/1520-0485(1999)029<1950:OOTIAP>2.0.CO;2.

Thomson, R. E., and W. J. Emery, 2014a: Statistical methods and error handling. *Data Analysis Methods in Physical Oceanography*, 3rd ed. R. E. Thomson and W. J. Emery, Eds., Elsevier, 219–311, https://doi.org/10.1016/B978-0-12-387782-6.00003-X.

——, and ——, 2014b: Time series analysis methods. *Data Analysis Methods in Physical Oceanography*, 3rd ed. R. E. Thomson and W. J. Emery, Eds., Elsevier, 425–591, https://doi.org/10.1016/B978-0-12-387782-6.00005-3.

Trowbridge, J., and S. Elgar, 2001: Turbulence measurements in the surf zone. *J. Phys. Oceanogr.*, **31**, 2403–2417, https://doi.org/10.1175/1520-0485(2001)031<2403:TMITSZ>2.0.CO;2.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272, https://doi.org/10.1038/s41592-019-0686-2.

Voulgaris, G., and J. H. Trowbridge, 1998: Evaluation of the acoustic Doppler velocimeter (ADV) for turbulence measurements. *J. Atmos. Oceanic Technol.*, **15**, 272–289, https://doi.org/10.1175/1520-0426(1998)015<0272:EOTADV>2.0.CO;2.

Wahl, T. L., 2003: Discussion of "Despiking acoustic Doppler velocimeter data" by Derek G. Goring and Vladimir I. Nikora. *J. Hydraul. Eng.*, **129**, 484–487, https://doi.org/10.1061/(ASCE)0733-9429(2003)129:6(484).

Wheeler, D. C., S. N. Gidings, and J. McCullough, 2023: Data and code from measuring turbulent dissipation with acoustic Doppler velocimeters in the presence of large, intermittent, infragravity frequency bores. University of California, San Diego, Library Digital Collections, accessed 6 February 2023, https://doi.org/10.6075/J0J67H27.

Williams, M. E., and M. T. Stacey, 2016: Tidally discontinuous ocean forcing in bar-built estuaries: The interaction of tides, infragravity motions, and frictional control. *J. Geophys. Res. Oceans*, **121**, 571–585, https://doi.org/10.1002/2015JC011166.