

## REVIEWER 2

*Estimates of TKE dissipation rate from inertial subrange fitting have not necessarily used consistent processing procedures from different groups, etc. The authors seek to publish a set of best practices for future research that follows. Although it does not include new scientific findings, I think this would be within the scope of EGU Ocean Science as a Technical Note and has potential to be useful to the community. I do have some minor comments and suggestions that I think would improve the paper.*

*1. The main point of this manuscript is to establish a set of “best practices”. While I agree that the methods in the manuscript are thorough and robust, I think it might be useful to discuss previous research in more detail and use this to guide justification of the new methods; i.e., how do the proposed methods differ from what has been used in the past and why are they superior? This, I think, would be important for 1) convincing readers that the proposed methods are indeed the best practice and 2) ultimately, having the community adopt them. This is probably best addressed in several different places throughout the text rather than in one place. To be fair, the authors have clearly thought about this issue and there are components of it already in the text. For example the justification on L218-224 is well referenced and is sort on in-line with what I think would improve the overall impact of the paper. However, I do not think the justification in other places is quite as thorough and convincing. A few places of these are pointed out below.*

*2. I suggest that the authors consider sharing relevant parts of the code used to prepare and process the sample datasets. I think the main intention of the authors is for their methods to ultimately be utilized by the community. While my suggestion is optional, sharing code would make it much easier for others to adopt the suggested practices.*

We thank Reviewer 2 for their feedback and respond to their comments below. Notably, we have expanded more on the choices of methods whenever feasible e.g., linear interpolation of quality-controlled data [L307-314]. However, we have not provided code. Prior to writing this manuscript, the development of the Working Group’s Terms of Reference examined the possibility of including code. Ultimately it was decided against distributing software (see [Terms of Reference](#)). The logic for this decision was the overarching objective being to document the best practices so that researchers can then validate different processing workflows against the benchmarks at any checkpoint (e.g., fitting algorithms at Level 3). Our chosen strategy avoids the pitfall of having the working group’s activities become obsolete because the code cannot be maintained and updated to keep up with new flavour of programming languages. It also intrinsically recognises that there may be other ways to arrive at the same answer.

### Line-by-line comments and suggestions:

*L32 – Include the full name of MAVS when first stated. I see it later on, but it should be included here*

Done on L34.

*L30-36 – I suggest to elaborate more on the reasons why the approaches are differently suited to high and low energy regimes.*

We have expanded more on L25-30 about the differences between shear probes and velocity sensors. Namely, shear probes can resolve the viscous subrange, and thus can sample in lower energy environments than velocity sensors.

*Fig 2 – I suggest to remove the gray shading inside the triangles. It suggest that the shading represents dissipation rate, but that is not the case.*

Done.

*L81-82 – Add additional clarification. That is, I think the authors are referring to the longer-term distribution of epsilon here, not that epsilon itself does not change over faster timescales than the sampling frequency (which as the authors are surely aware can happen and does not necessarily signify a violation of stationarity).*

*Reading on, I see the authors include a nice explanation of this at L259... maybe also briefly mention it here.*

As suggested, the text on L93-95 was modified by using elements of the explanation that had been given later on.

*L83-84 – Perhaps move this to the start of the paragraph?*

Done.

*L121 – I suggest to add a citation or additional explanation*

We added a reference to *Bluteau et al. (2011)* given their discussion on the use of the distance to the boundary (eq. 7) in lieu Corrsin length scale (eq. 6).

*L127-137 – Perhaps Fig 2 should be referred to at the start of the paragraph? It visually explains much of what is here, and I think may be easier to understand for readers new to the topic*

Figure 2 is now referenced at the beginning of this paragraph at L140.

*Table 1 title – U bar is 5th and 95th percentile, correct? Also, why not 2.5 and 97.5 to get the 95% confidence interval.*

No it was 50th (median) and 95th percentile. We now provide the 2.5th, 50th, and 97.5th percentile despite the 2.5th percentile being near zero when the tide turns.

*L154-159 – Although important for understanding the datasets, I feel this is a bit of an aside from the main text. If this is a suggestion as the format for data storage for future experiments, that should be clearly stated. From my experience, many (most?) published datasets are essentially just the “fourth level”, and if there is a suggestion for the inclusion of more information that should be discussed. In any case, I think the main objective is the methods for how to get epsilon, and perhaps this section can be pared down.*

*Reading on, this is necessary to explain for section 4, so maybe it can be incorporated into that section*

We have revised and shortened this paragraph and the beginning of this section to emphasize the purpose of the benchmarks and why we have four groups. We now also explicitly state which data should be submitted with peer-reviewed manuscripts and what should be summarised in the publications’ methods section [L663-667].

*L162 – I think linear interpolation needs to be justified here. I understand the necessity to fill in gaps for constructing spectra, and think further explanation would be useful.*

This section is just an overview of the steps, so we have removed the reference to linear interpolation; options for replacing missing data are now discussed more in Section 4.2.

*L164 – I don’t understand what “turbulence models” means in this context. I don’t think it’s the method of calculating epsilon, since that would still be the same within a dataset?*

It referred to the choice of turbulence equation (model) to fit the spectra. We have revisited the sentence accordingly on L175.

*Sec 4.1 (until L211) – I’m not sure how useful parts of this section are. That is, I don’t know if it is useful to state that QC should be done on a case-by-case basis or to follow manufacturer recommendations. I understand it is important to QC and thus why the authors have discussed this, but perhaps it could be shortened in places.*

The introduction to the processing method was shortened and re-organised so that basic QAQC and phase-wrapping appear in their own subsections before discussing despiking.

*Fig 4 – I suggest to put a box around the legend in panel a) for clarity and consistency with panel b).*

The legend box is now added to panel a).

*L312-321 – See my general comment. This is a place where I think additional background information, and references that highlight deficiencies of other methods, would be useful. E.g. why is this spectral averaging procedure different and superior to others?*

Some elements of spectral analysis cross over between ATOMIX groups. The ATOMIX paper on shear probes discuss this particular point (Lueck *et al.*, 2024) so we have added a clearer pointer to that here.

*L340-356 – I think additional background information and referencing would be useful here, as well. Is there another study that has found this with observations, in addition to the study using synthetic spectra which is already cited?*

There aren't other papers with real observations. Besides the true  $\varepsilon$  is unknown with real observations making the assessment of fitting methods challenging. We have added a paragraph at L354-363 summarizing the six methods tested in the cited paper by Bluteau (2025a), which was written solely to address the lack of papers on testing of fitting methods.

*L346,355,357 – “loglad” and “ladlog” are the same thing, right? Please clarify. These are not clearly defined.*

Yes, this is a typo. They should all be labelled logLAD.

*Fig 6 – I don't understand the purpose of showing Fig 6. Is it just to show the values of “A”? These are already stated in the text, and I barely see any difference between the different colored lines (I think that is to show insensitivity to the sampling, perhaps?). Maybe I am missing something here.*

Figure 6 (now 8) is meant to visually illustrate that the functional form of the expected bounds for the calculated spectral slopes in equation 24. The figure has been moved much later in section 4.6.4 when discussing flagging  $\hat{\varepsilon}$ .

We modified the y-label to highlight that it represents the  $A$  thresholds that must be selected by users to flag their data. We have added the individual data points for each performed tests, which now highlights the maximum deviation from the known (prescribed) slope  $\beta_1 = -5/3$  occurs when  $A = 17$ .

Yes, the lines being flat means  $A$  remains consistent across the four series of numerical experiments conducted across a wide range degrees of freedom  $d$  (x-axis). Each series were conducted against 3200 synthetic spectra (100 per degrees of freedom tested) (Bluteau, 2025a,b).

*Sec 4.4.2 – Similar to my previous comments, are there any other studies that have tried different inertial ranges, especially with field observations? Or that have found deficiencies with a method or specifications that are not recommended in the present manuscript? While I agree the results of Bluteau 2025 for synthetic observations cited here and elsewhere are very relevant, I think additional background would strengthen some of the recommendations.*

We didn't find any literature that tested methods for finding the inertial subrange. Many studies fix the wavenumbers being fitted, and presume it won't include the low and high wavenumbers flattened by the impact of the mean flow (energy containing range) and noise, respectively.

Below, we illustrate the differences among the methods against our recommended “slope” method for synthetic spectra (Figure r1) and the *Tidal Shelf ADV* spectra (Figure r4 and r5). The tests against the synthetic

spectra show that the three other methods tend to situate the inertial subrange at low wavenumbers (Figure r2 a-c), albeit at higher wavenumbers than the slope method (Figure r2 d-f). Nevertheless, the returned  $\hat{\epsilon}$  generally agreed with the prescribed value  $\epsilon_p$  (Figure r2 d-f) since the inertial subrange is sufficiently broad and all methods correctly avoided the noise-dominated portion of the spectra when searching for the inertial subrange.

When noise levels were increased by an order of magnitude to  $\Phi_n = 10^{-6}$  (m/s)<sup>2</sup>/rad/m, the identified wavenumbers tended to sit at low wavenumbers for all four methods since the noise floor now impinges on the high wavenumbers of inertial subrange (Figure r3 a-c). There was still a tendency for our recommended method to situate the inertial subrange at lower wavenumbers than the other three methods (Figure r3 d-f). The estimated  $\hat{\epsilon}$  tends to deviate more when the identified wavenumbers from the other methods deviate from those determined by our recommended “slope” method (Figure r3 d-f). However, these deviations tend to be associated with spectra that have low degrees of freedom i.e., with minimal spectral averaging (not shown).

We also did some tests where both the low wavenumbers were flattened (not shown) to replicate another common problem with observed spectra deviating from the expected model form. The estimated  $\hat{\epsilon}$  agreed better with the prescribed value when the slope method was used to locate the inertial subrange, and this despite situating it at lower wavenumbers than the three other methods.

For illustration purposes, we also applied the four identification methods against our *Tidal Shelf ADV* benchmark dataset. The comparison between our recommended “slope” method with the median of  $|\log \hat{\Psi} - \log \Psi|$  is presented in Figure r4. The larger discrepancies tend to be associated with very poor spectra (Figure r5a) that end up being flagged later on for other reasons besides the poor slope.

We have added a bit more information about the testing in the manuscript in §4.4.2, although we do note that all methods for identifying the inertial subrange do fairly well.

*Also, I'm curious if there was a distinction or impact on performance when only short segments had a good-enough fit to be classified in the inertial range, as opposed to cases where longer segments fit  $k^{-5/3}$ .*

For completeness, we present equivalent plots for the tests done over much shorter subsets ( $\delta = 0.3$ , Figure r6 and r7) than those described above ( $\delta = 1$ , Figure r2 and r3). Compared to the tests with longer subsets ( $\delta = 1$ ), the identified wavenumbers tend to become more variable for each method, especially for the spectra with lower noise levels (Figure r6 a-c vs r7 a-c). The estimated  $\hat{\epsilon}$  also deviated more from the prescribed  $\epsilon_p$  across the three “non-slope” methods when the noise floor increased from  $\Phi_n = 10^{-7}$  to  $\Phi_n = 10^{-6}$  (m/s)<sup>2</sup>/rad/m (Figure r6 d-f and r7 d-f), albeit the deviations were less pronounced than for the tests with longer subsets (Figure r2 and r3).

Regardless of the performance difference when using short subsets, we recommend using fairly long subsets  $\delta$  to identify the location of the inertial subrange so that it matches the recommended  $\delta \geq 0.8$  noted at L424 for estimating  $\hat{\epsilon}$  from fitting velocity spectra.

*L504 – A=17 is dependent on the estimates you made, and not universal, correct? That is implied at L507, I think. I suggest to reword here (and in the other subsections of 4.6 where applicable) whether the suggested thresholds are expected to be reasonable in all cases.*

Most of the text surrounding  $A$  has now been moved into section 4.6.4 when discussing the flagging of  $\hat{\epsilon}$  estimates.  $A$  is user-defined and has been empirically determined by scaling the results in Figure 7 and 8 of

Bluteau (2025a) with the main parameters responsible for the variations in estimated  $\hat{\beta}_1$  across their series of numerical experiments. The user-selected  $A$  presumes that the functional form of the expected bounds for the calculated spectral slopes in equation 24 is correct. The estimated  $A$  for each individual spectra tested have been added in Figure 8 to highlight the maximum  $A = 17$  in addition to 99.7% bounds that were already illustrated.

*Sec 4.6.6 – Wouldn't anisotropy potentially also influence the spectral shape? Also, related to this, it might be worth mentioning that these flags are not exclusive at the start of section 4.6.*

Yes, it does impact the spectral shape which is why we have section 4.4.2 and Figure 6 to discuss how to identify the inertial subrange. This section emphasizes the need for identifying the inertial subrange to avoid the flatter spectra at low wavenumbers (anisotropy) and high wavenumbers (noise).

The text in §4.6.6 was also revised since this flag is mostly for identifying segments when the turbulence levels are too weak to be considered isotropic as opposed to flagging spectra with flatter shapes at low wavenumbers.

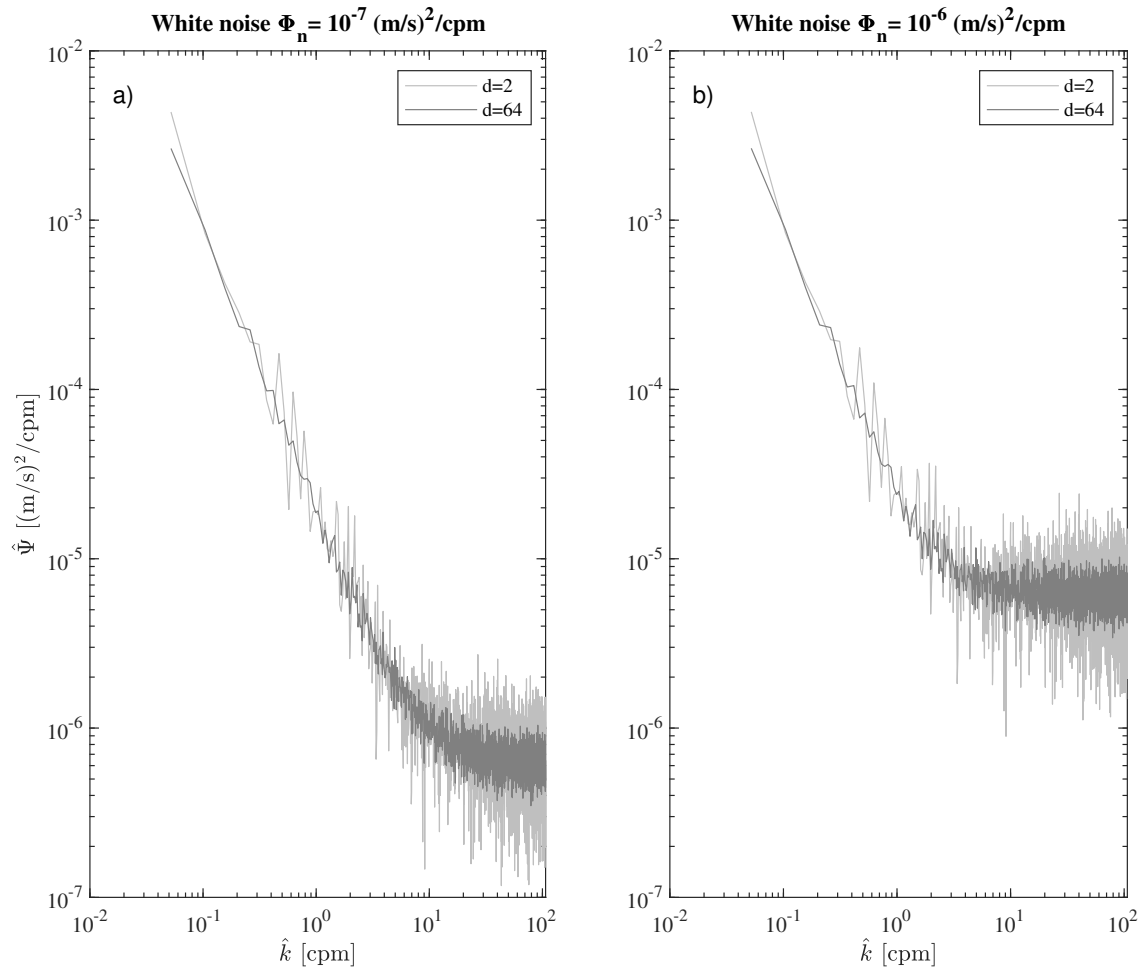
*L563 – I don't quite understand why this flag was not applied. Is it because phase unwrapping was easily rectified (as I think the earlier text might suggest)?*

The flag was applied and is stored in the NetCDF. The sentence “For this dataset, the 8th flag, denoting phase-wrapped samples, was not used to discard velocity samples” was rewritten since the velocity samples were unwrapped for subsequent analysis rather than discarded, unlike samples flagged for any other reason.

*Fig 11e – Like Fig 4, I suggest to put a box around the legend as the legend symbol for “flagged” is hard to distinguish from the actual flagged values. This could also be done for other panels in this figure.*

Each legend in Fig 11 is now boxed.

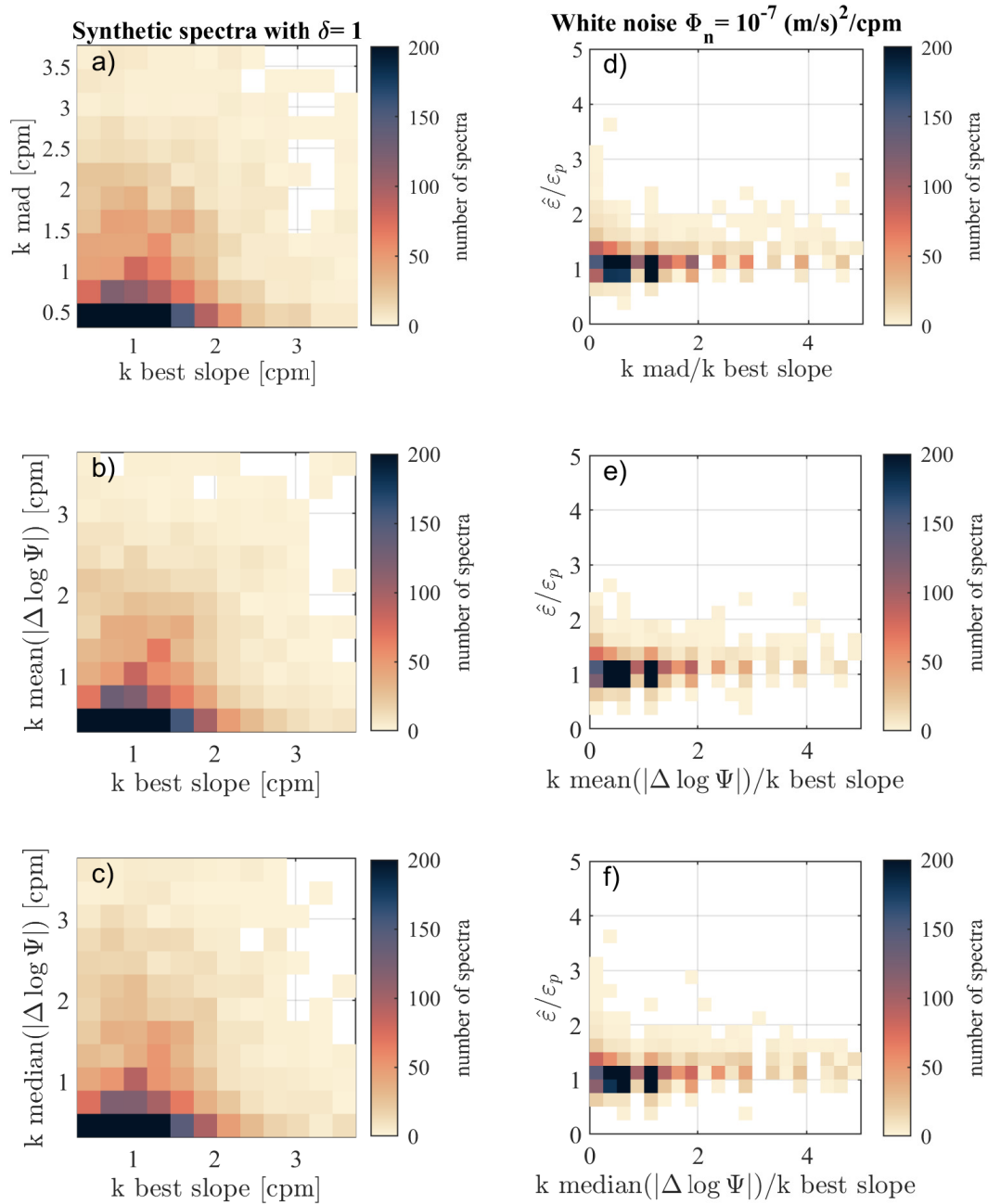
## Figures about identifying inertial subrange



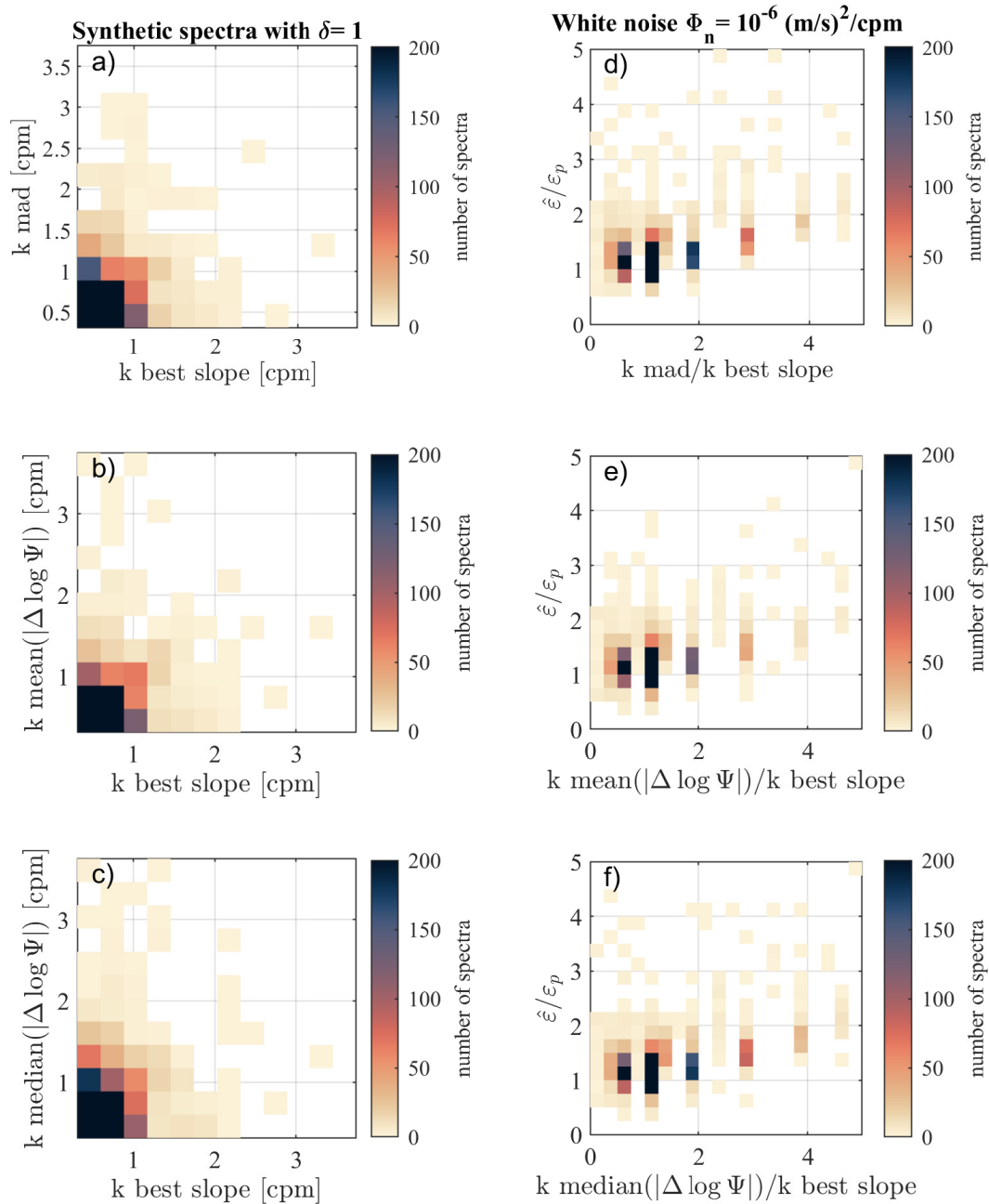
**Fig r1:** Example spectra used to test the four methods described in §4.4.2 of the manuscript for identifying the inertial subrange. These tests' results are shown in Fig r2, r3, r6, and r7. These tests were against the 3200 synthetic spectra with their uncertainty generated using the  $\chi_d$ -distribution (Bluteau, 2025b,a). White noise at a pre-specified energy level was added to each of the 3200 synthetic spectra to mimic noise floor impinging high wavenumbers. Two examples are shown for two different degrees of freedom  $d$  in with a (a) noise floor of  $\Phi_n = 10^{-7} \text{ (m/s)}^2/\text{rad/m}$  and (b) with a higher noise floor of  $\Phi_n = 10^{-6} \text{ (m/s)}^2/\text{rad/m}$ . The prescribed  $\varepsilon_p = 10^{-6} \text{ W kg}^{-1}$  and the inertial subrange ends approximately at  $\hat{k} \approx 13\text{cpm}$ .

## REFERENCES

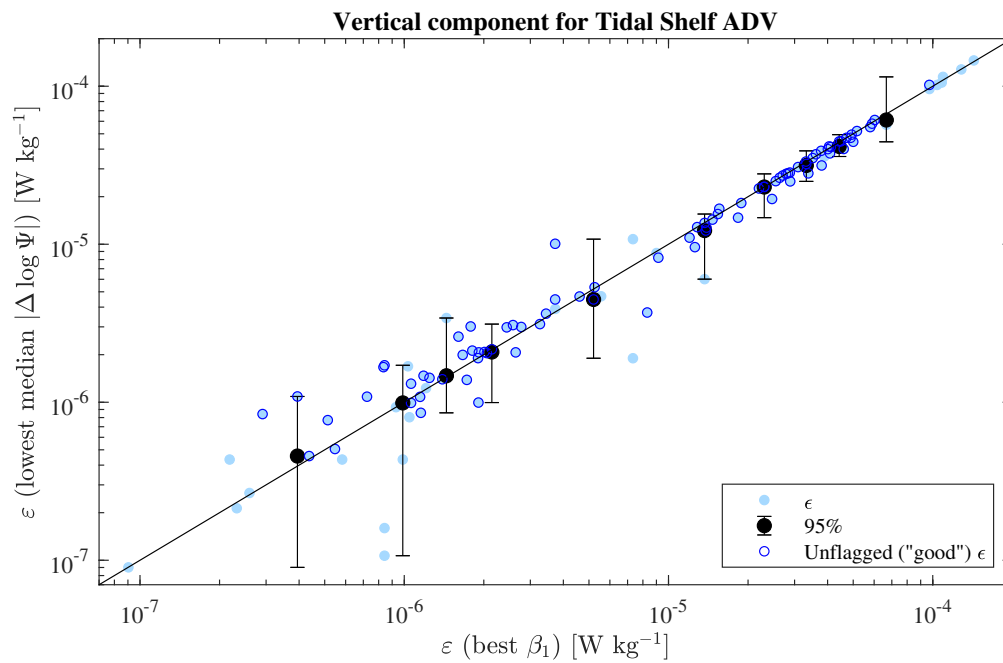
- Bluteau, C. (2025a), Assessing statistical fitting methods used for estimating turbulence parameters, *Limnol. Oceanogr.: Methods*, doi:10.1002/lom3.10729.
- Bluteau, C. (2025b), Synthetic spectra for assessing statistical fitting methods used to estimate ocean turbulence, *Zenodo*, doi:10.5281/zenodo.10576543, zenodo.
- 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. and Oceanogr.: Methods*, 9, 302–321, doi: 10.4319/lom.2011.9.302.



**Fig r2:** Left panels: comparison of the median wavenumber identified as being the most likely location for the inertial against those determined from our recommended “best slope” method for the three other methods tested: mad (a), mean of  $|\log \hat{\Psi} - \log \Psi|$  (b), median of  $|\log \hat{\Psi} - \log \Psi|$  (c). The right panels (d) to (f) show the deviations of the estimated  $\hat{\varepsilon}/\varepsilon_p$  for the three methods presented in y-axis of the left panels (a)-(c), respectively. The x-axis represents the ratio of the results presented in the left panels. The fitted decadal range for each subset was  $\delta = 1$  while the noise floor added was  $\Phi_n = 10^{-7} \text{ (m/s)}^2/\text{rad/m}$  (see example synthetic spectra in Figure r1 a).

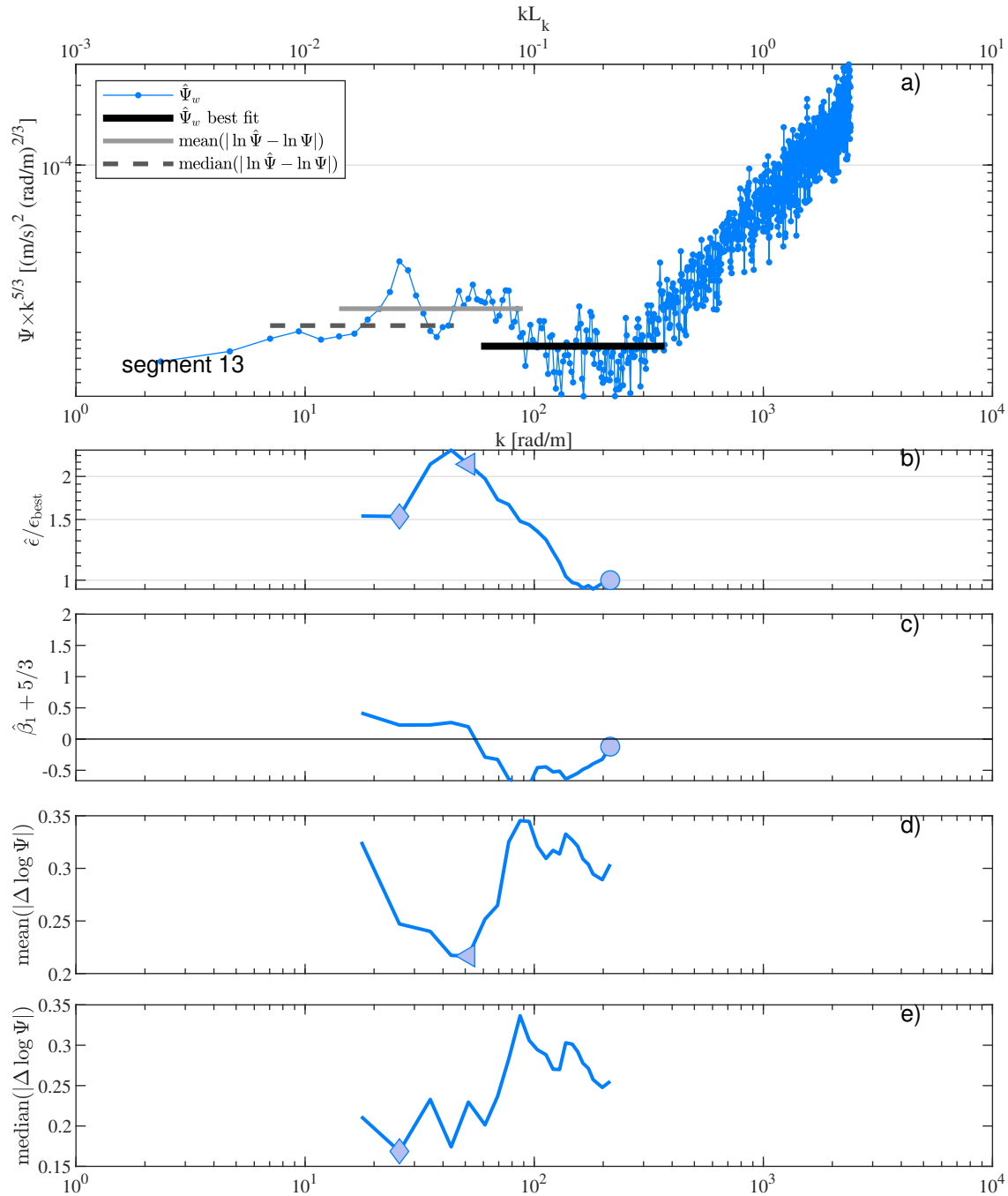


**Fig r3:** Same as Figure r2 but adding a noise floor of  $\Phi_n = 10^{-6} \text{ (m/s)}^2/\text{rad/m}$  (see example synthetic spectra in Figure r1 b).

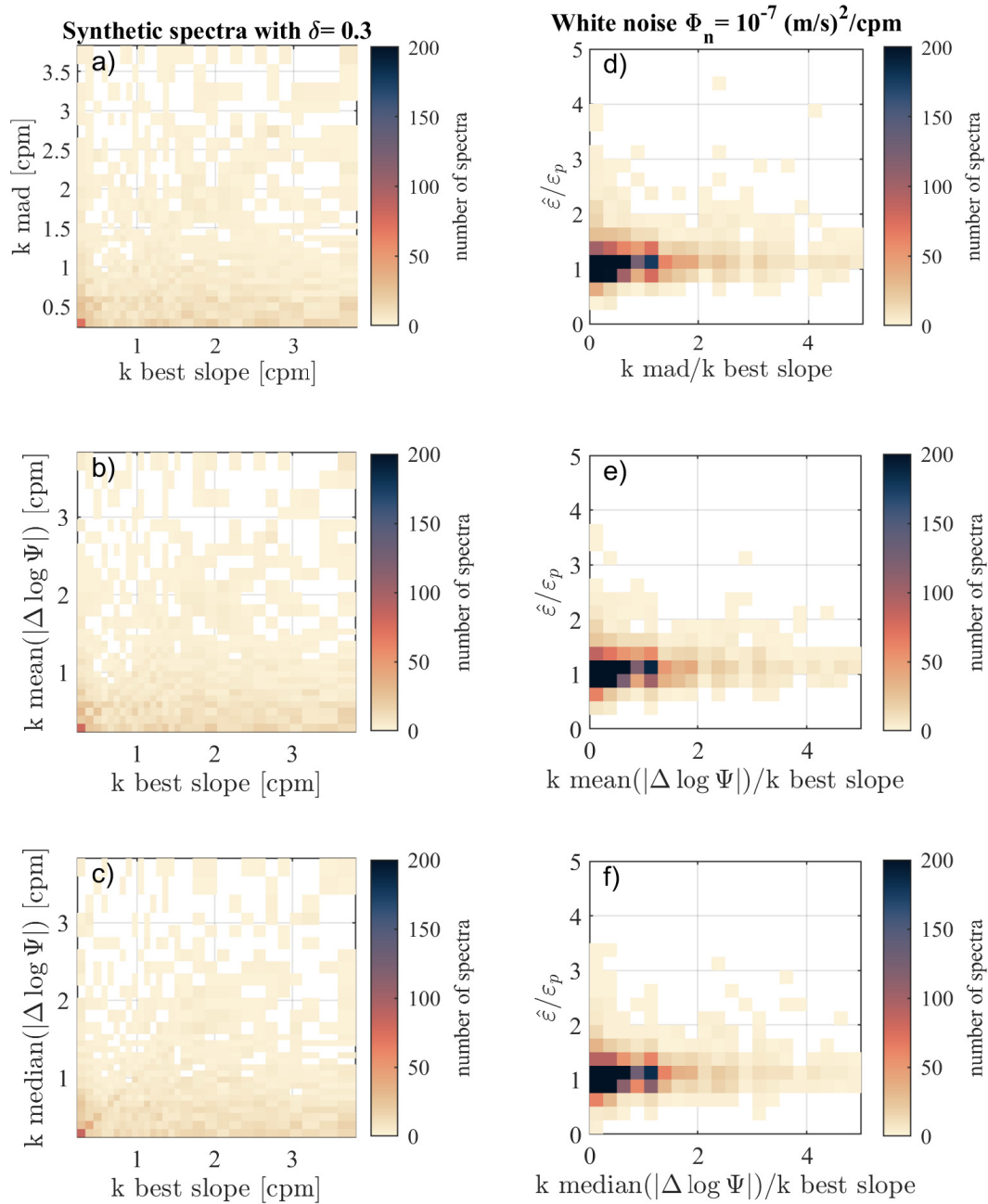


**Fig r4:** The returned  $\epsilon$  pending which method is used to determining the location of the inertial subrange. Using the minimum from the median of  $\Delta \log \Psi = |\log \hat{\Psi} - \log \Psi|$  returns generally similar results as the  $\epsilon$  corresponding to the spectra with the slope closest to  $-5/3$ . The observations are from the *Tidal Shelf ADV* benchmark.

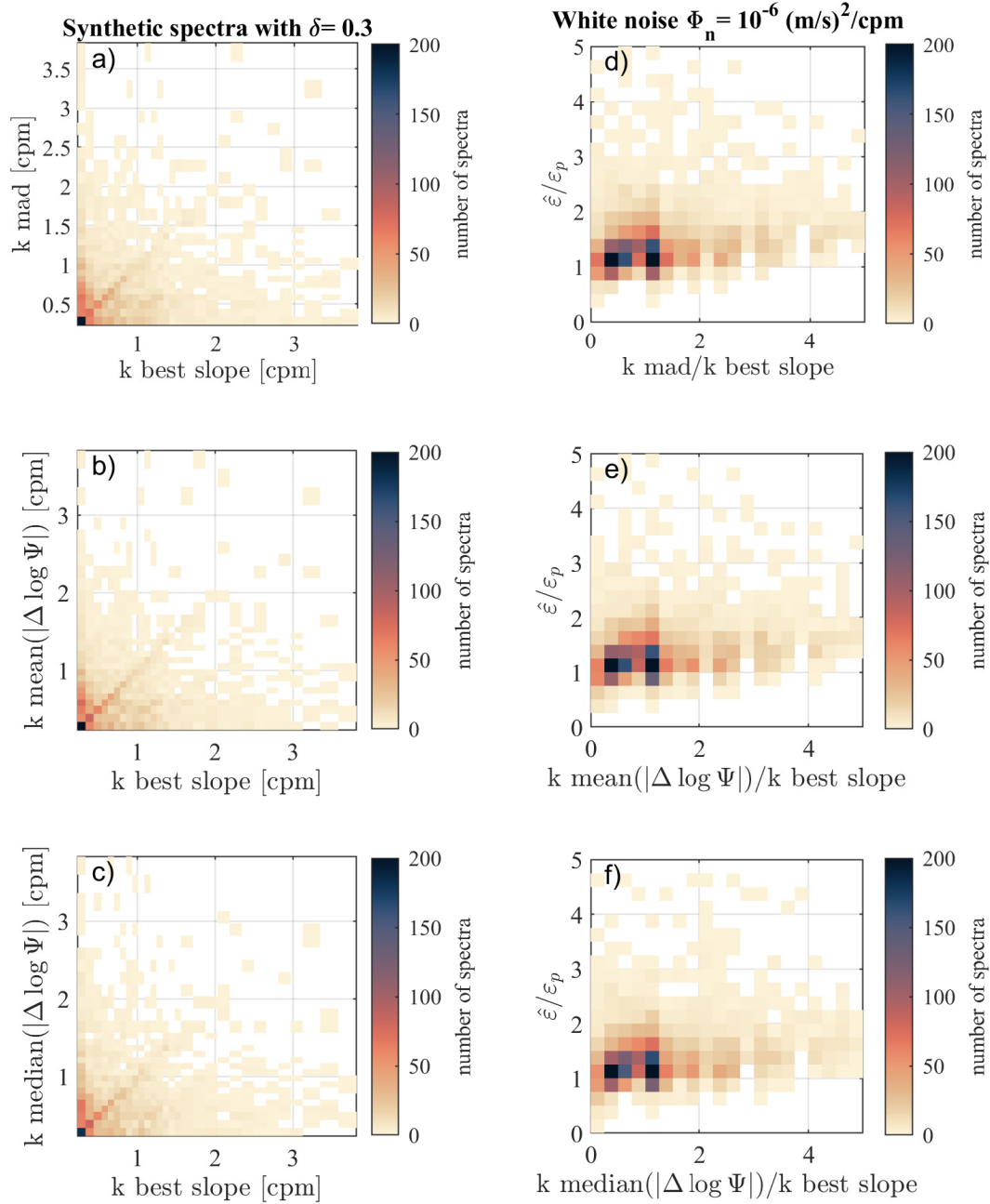
Lueck, R., et al. (2024), Best practices recommendations for estimating dissipation rates from shear probes, *Frontiers in Marine Science*, 11, doi:10.3389/fmars.2024.1334327.



**Fig r5:** Similar to Figure 6 in the manuscript. This segment had the biggest discrepancy in Figure r4 for where the inertial subrange was placed in the spectra (segment 13 from Tidal Shelf ADV experiment).



**Fig r6:** Left panels: comparison of the median wavenumber identified as being the most likely location for the inertial against those determined from our recommended “best slope” method for the three other methods tested: mad (a), mean of  $|\log \hat{\Psi} - \log \Psi|$  (b), median of  $|\log \hat{\Psi} - \log \Psi|$  (c). The right panels (d) to (f) show the deviations of the estimated  $\hat{\epsilon}/\epsilon_p$  for the three methods presented in y-axis of the left panels (a)-(c), respectively. The x-axis represents the ratio of the results presented in the left panels. The fitted decadal range for each subset was  $\delta = 0.3$  while the noise floor added was  $\Phi_n = 10^{-7} \text{ (m/s)}^2/\text{rad/m}$  (see example synthetic spectra in Figure r1 a).



**Fig r7:** Same as Figure r6 but adding a noise floor of  $\Phi_n = 10^{-6} \text{ (m/s)}^2/\text{rad/m}$  (see example synthetic spectra in Figure r1 b).