Global Patterns of Bias in Ocean Mixing Parameterization Identified Through Unsupervised Machine Learning

Ratnaksha Lele[1], Sarah G. Purkey[2], Jennifer A MacKinnon[3], and Jonathan D Nash[4]

[1]Scripps Institution of Oceanography, UC San Diego
[2]Scripps Institution of Oceanography, UCSD
[3]Scripps Institution of Oceanography
[4]Oregon State University

November 8, 2023

Global Patterns of Bias in Ocean Mixing Parameterization Identified Through Unsupervised Machine Learning

Ratnaksha Lele [1], Sarah G. Purkey [1], Jennifer A. MacKinnon [1], Jonathan D. Nash [2]

[1]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA
[2]College of Earth, Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR

Key Points:

- Unsupervised machine learning identifies clusters with distinct shear and strain spectral energy transport across wavenumbers.
- Geographical distribution shows horizontal and vertical preferences in some of the identified clusters.
- Two clusters diverge significantly from GM spectra identify regions of biases in finescale parameterizations.

Corresponding author: Ratnaksha Lele, `rlele@ucsd.edu`

Abstract

Turbulent mixing in the ocean is often parameterized in terms of the downscale energy transfer by internal waves. Expressed in terms of the vertical wavenumber spectrum of oceanic velocity shear (V_z^2) and isopycnal strain (ζ_z^2), the "finescale parameterization" relies on several parameters, including key assumptions relating to the spectral properties. Here we use an unsupervised learning model to identify spatial correlations between embedded parameters of the finescale parameterization based upon data from 1875 full-depth hydrographic profiles from 15 sections traversing the global ocean. The clustered patterns along the sections have marked horizontal and vertical spatial dependence associated with distinct modes of spectral variation. Two clustered regions are identified where the underlying spectra deviate significantly from the canonical Garrett-Munk (GM) spectrum, suggesting potential departures from implicit assumptions about the downscale energy cascade. Spectral composites in these two regions show intensification of variance in the low and high wavenumber regimes respectively, as well as distinction in overall spectral levels and geographic prevalence. Furthermore, these clusters are found to be associated with regions where parameterized estimates of the turbulent dissipation rate ϵ differ significantly (exceeding a factor of 5) from co-located in-situ observations measured using χ-pod temperature microstructure. Extending the methodology to other hydrographic datasets has the potential to reveal reasons for this parameterization bias and to identify the dynamical underpinnings leading to more robust parameterizations of oceanic turbulent mixing.

Plain Language Summary

Turbulent mixing caused by breaking internal waves is the primary driver of the vertical heat transport and is critical for closing the ocean's energy budget. To circumvent the complexities in obtaining in-situ measurements of mixing, simplified parameterized models to estimate the rate of mixing are widely used by utilizing relatively easily collected oceanic properties such as temperature and velocity as inputs. However, inaccuracies in predictions by these simplified models arise when certain assumptions in the model are violated. In this study, by incorporating data collected from a global suite of ship based observations, we use a data-driven approach to identify the spatial distribution of two distinct regions in the ocean where large biases in the predictions by the simplified models are possible. Extending this approach, future studies could potentially identify the underlying causes of such disparities to further improve models of turbulent mixing in the ocean.

1 Introduction

Turbulent mixing plays a critical role in the overturning circulation of the global ocean, driving the vertical and horizontal transport of heat and tracers (Ganachaud & Wunsch, 2000; Wunsch & Ferrari, 2004). While mixing at the molecular level can be explicitly characterized by thermodynamic diffusion equations, the observed interior ocean stratification requires vigorous turbulent mixing that is 10-100 times stronger than that from molecular diffusion alone (e.g., W. H. Munk, 1966; Bryan, 1987; Talley, 2003; Cimoli et al., 2023), driven primarily through breaking internal waves (K. L. Polzin et al., 1997; Kunze et al., 2006; Whalen et al., 2012; Waterhouse et al., 2014; MacKinnon et al., 2017). The strength of this turbulent mixing is governed by distinct physical and dynamical processes which result in rich geographical patterns of mixing throughout the global ocean. (K. L. Polzin et al., 1997; Naveira Garabato et al., 2004; Whalen et al., 2012; Waterhouse et al., 2014; Whalen et al., 2018).

Resolving the spatiotemporal patterns of turbulent mixing in the ocean from observations is significantly challenging owing to the intermittent nature of mixing. At present, the most accurate estimates of turbulent mixing come from specialized microstructure instrumentation deployed from ships (K. L. Polzin et al., 1997; St. Laurent et al., 2012; Naveira Garabato et al., 2019; Lele et al., 2021), on moorings (Moum & Nash, 2009), and autonomous platforms (Rudnick et al., 2013; Johnston & Rudnick, 2015; Shroyer et al., 2016). These microstructure instruments allow for estimates of kinetic energy dissipation rate (ϵ) and temperature gradient variance (χ) by measuring high-frequency velocity and/or temperature gradients. The resolved turbulent gradient spectra in the inertial subrange of turbulence are then used to compute ϵ and χ (e.g., Oakey, 1982; Gregg, 1999; Itsweire et al., 1993). However, global microstructure observations have sparse global spatio-temporal coverage (Waterhouse et al., 2014).

In response to the low abundance of microstructure observations, the community has embraced a set of mixing parameterizations based upon internal wave-wave interaction theories called finescale parameterizations that allow for the estimate of ϵ from lower resolution temperature and salinity data (Henyey et al., 1986; Gregg, 1989; K. L. Polzin et al., 1995). These parameterizations estimate turbulent dissipation of energy by estimating the rate of downscale energy transfer through wave-wave interactions by combining the measured internal wave spectral level and theoretical and empirical models of wave interactions. The applicability of finescale parameterizations on the more widely available oceanographic ship-based and Argo-based Conductivity Temperature Depth (CTD) and Lowered Acoustic Doppler Current Profiler (LADCP) data has drastically increased the spatial coverage of mixing estimates as well as our understanding of the spatial geography of mixing in the ocean (e.g., Whalen et al., 2015; Kunze, 2017b) with overall broad agreement with measurements obtained from microstructure instrumentation (e.g., K. L. Polzin et al., 1995, 2014; Whalen et al., 2015; Whalen, 2021) Crucially for our work, the spectral energy level is estimated by comparing the *average* spectral level within a limited wavenumber band to the idealized Garrett-Munk (GM) model (Garrett & Munk, 1972; W. Munk, 1981). Since the finescale parameterizations are referenced to the GM model in their formulation of spectral energy transport through the internal wave vertical wave number space, large departures from the GM model are susceptible to engendering biased estimates (K. L. Polzin et al., 2014).

Dissipation rates ϵ and related eddy diffusivities κ obtained from finescale parameterizations show overall broad agreement with measurements obtained from microstructure instrumentation (e.g., K. L. Polzin et al., 1995, 2014; Whalen et al., 2015; Whalen, 2021), however, some discrepancies and biases have also been previously documented (e.g., MacKinnon & Gregg, 2003; Waterman, Polzin, Naveira Garabato, et al., 2014). The underlying assumptions of the parameterizations are violated in many regions of the ocean, such as in the surface mixed layer, or where turbulent mixing is controlled by double diffusion, hydraulic jumps and strong wave-mean flow interactions over rough topography (Waterman, Polzin, Naveira Garabato, et al., 2014; K. L. Polzin et al., 2014). There are also regions where the parameterized mixing rate does not match that observations from microstructure for unclear reasons. A hypothesis considered here is that deviations of the spectral shape or other properties of the internal wave spectrum from the assumed GM form may be relevant (Müller & Liu, 2000; K. L. Polzin & Lvov, 2011), or variability in other individual parameters of the parameterization themselves, based on the local geography, topographic conditions and the presence of external forcing to the local internal wave field (Waterman, Polzin, Naveira Garabato, et al., 2014; Chinn et al., 2016; Pollmann, 2020). Recently, both supervised and unsupervised learning approaches have been used across a variety of fluid mechanical applications to provide new insight into fundamental relationships and patterns of variability in our oceans (Giglio et al., 2018; Brunton et al., 2020; Callaham et al., 2021; Kaiser et al., 2022; Mashayek et al., 2022). In particular, clustering techniques have proven useful in generating insights and exploring existing oceanographic data such as categorizing datasets of temperature-salinity profiles (e.g., Rosso et al., 2020; Jones et al., 2019; Boehme & Rosso, 2021), classifying global ecological marine provinces (Sonnewald et al., 2020) and identifying dominant dynamical balances in global ocean circulation models (Sonnewald et al., 2019).

In this study, we employ unsupervised learning to characterize a parameter-space associated with large mismatches between finescale and microstructure observations of oceanic turbulent mixing. Drawing inspiration from unsupervised learning approaches in the spectral domain applied to earthquakes and astronomical observations (Johnson et al., 2020; Ivezic et al., 2014), we use latent features extracted from oceanic shear and strain spectra as well as other variables (features) used in the formulation of finescale parameterizations to identify regions of distinct co-variations connected to properties of turbulent mixing in the ocean and underlying dynamics of internal wave-wave interactions. The curated hydrographic dataset used in the study is described in Section 2, with the underlying principles of finescale parameterizations, feature development, dimensionality reduction and clustering model laid out in Section 3. Finally, we describe the geography and spatial characteristics of the clustering results and the interpretation of the results in the context of the underpinning finescale parameterization for estimating turbulent mixing in the ocean in Sections 4 and 5.

2 Data

2.1 Ship-based Hydrographic Data

The principal data used in this study are 1875 profiles of high-quality full-depth CTD and LADCP data collected along 15 hydrographic sections from around the globe as part of either the Climate and Ocean Variability, Predictability and Change (CLIVAR) or the Global Ocean Ship-based Hydrographic Investigations Program (GO-SHIP) programs, between the years 2000 and 2021 (Figure 1, Table 1). The horizontal station spacing between CTD casts is nominally 55 kilometers, with stations spaced closer in regions of interest (e.g. trenches, rough topography, boundary current regions). Vertically, the CTD-cast data used here include the profile from 500 m down to a maximum depth, usually 6000 m, or to within 10-20 m from the seafloor. The top 500 m of the profiles are not considered in order to remove the surface mixed layer. Conservative temperature (θ), squared buoyancy frequency (N^2), and potential density ρ_θ are calculated from the CTD instrumentation using the Gibbs-Seawater Oceanographic Toolbox (McDougall, 2011; Jackett & McDougall, 1997). The publicly available LADCP data product has an 8-meter vertical resolution pre-processed using procedures laid out for the GO-SHIP program (Visbeck, 2002; Thurnherr et al., 2010). The LADCP data product for all 15 sections contains data binned at a nominal 1-meter resolution and with horizontal (U,V) and vertical component (W) of ocean velocity from the ocean surface down to the maximum CTD depth. All LADCP data obtained are co-located with CTD data for each CTD cast along the sections.

2.2 Microstructure mixing estimates from CTD-mounted χ-pods

Estimates of ϵ from rosette-mounted microstructure χ-pods taken along the P06 section were obtained from cchdo.edu (see data availability statement). ϵ was estimated using the high wavenumber temperature gradient fluctuations dT'/dz measured by the 100Hz FP07 thermistor probe following the methods of Moum and Nash (2009) and Lele et al. (2021). The data have all been processed and cleaned including (1) removing any points with platform-induced noise, (2) calculating dissipation rates of the temperature variance, χ in 1-s bins, (3) any data in regions of very weak stratification where $\overline{dT/dz}$ is less than 10^{-4} K m^{-1} was removed, (4) ϵ was calculated from χ following Osborn and Cox (1972) and (5) data were binned into 200-m half overlapping segments, ensuring binned averages comparable binned finescale parameterization data (see Section 3.1).

3 Methods

3.1 Estimating Mixing from Finescale Parameterizations

Profiles of ϵ and κ are estimated from 1875 CTD stations containing a total of *64816* spectral estimates of internal wave shear $\langle V_z^2 \rangle$ and strain $\langle \zeta_z^2 \rangle$ variances using the finescale parameterization method following Gregg (1989); Henyey et al. (1986); K. L. Polzin et al. (1995); Kunze et al. (2006). Shear and strain variances are computed from CTD temperature and salinity and LADCP horizontal velocities profiles along the sections. Variance levels relative to the canonical Garrett-Munk (GM) spectra (W. Munk, 1981) are used to relate vertical turbulent eddy diffusivity (κ) to the turbulent kinetic energy dissipation rate ϵ via the Osborn (1980) relationship $\kappa = \Gamma \frac{\epsilon}{N^2}$, wherein mixing efficiency Γ considered to be nominally 0.2 (K. L. Polzin et al., 2014) and N is the buoyancy frequency. This relationship is further broken down as:

$$\kappa = \kappa_0 E_{v_z} h(R_\omega) J(f/N), \tag{1}$$

with

$$E_{v_z} = \frac{\langle V_z^2 \rangle^2}{\langle V_z^2 \rangle^2_{GM}} \tag{2}$$

$$h(R_\omega) = \frac{3(R_\omega + 1)}{2\sqrt{2} R_\omega \sqrt{R_\omega - 1}} \tag{3}$$

$$J(f/N) = \frac{f \cosh^{-1}(N/f)}{f_{30} \cosh^{-1}(N_0/f_{30})} \tag{4}$$

where $\cosh^{-1}$ is the inverse hyperbolic cosine function, and constant values $\kappa_o = 5{\times}10^{-6} \mathrm{m^2 s^{-1}}$, $f_{30} = 7.292{\times}10^{-5}$ rad s^{-1} and $N_0 = 5.2{\times}10^{-3}$ rad s^{-1}, where f$_{30}$, N_o and κ_o denote the Coriolis frequency at 30^{o}N latitude, the canonical GM buoyancy frequency and background diffusivity respectively.

The angle brackets in Equation 2 indicate integration of LADCD-derived shear spectra over a wavenumber band capturing finescale internal wave shear variance (Gregg, 1989; K. L. Polzin et al., 2014). The factor $J(f/N)$ in Equation 1 is a latitudinal correction applied to account for weaker turbulent dissipation rates found near equatorial regions (Henyey et al., 1986; Gregg et al., 2003), while the factor h(R_ω) in Equation 3 accounts for deviations from the GM spectrum based on the frequency content of the internal wave field given by R_ω, reducing to unity when R_ω is set to the canonical GM value of 3 (K. L. Polzin et al., 1995). The dependence on strain (ζ_z) is introduced in the parameterization through the shear to strain variance ratio $R_\omega = \frac{\langle V_z^2 \rangle}{N \langle \zeta_z^2 \rangle}$, a measure of the internal wave fields aspect ratio or frequency content. This, under a monochromatic wave assumption, can be summarized as:

$$\frac{\omega}{f} = \sqrt{\frac{R_\omega + 1}{R_\omega - 1}} \tag{5}$$

representing the contribution of near-inertial (ω/f ≈ 0) to non near-inertial internal waves in the domain.

Profiles of ϵ and κ are calculated at each CTD station along the section from 200-m half-overlapping segments in depth using the parameterization given by Equation 1. It is important to note, however, that these parameterized estimates of diffusivity κ and dissipation rate ϵ do not sufficiently resolve mixing processes in the boundary layer, hydraulic jumps, double diffusion or internal wave driven turbulence in regimes with significant wave-mean flow interaction (Waterman, Polzin, Naveira Garabato, et al., 2014) and they produce spatially averaged estimates of mixing over multiple wave periods.

3.2 Feature Development

Here, we define and extract features from various attributes of the parameterization, to examine and understand the patterns of their cross-covariances as they relate to internal-wave driven mixing in the global ocean using unsupervised machine learning.

Building upon the parameterization in Equation 1 as the basis for feature development, we focus on measured shear and strain spectra which are the primary components of the parameterization. Buoyancy frequency normalized shear [ϕ_{V_z}] and strain [ϕ_{ζ_z}] wavenumber spectra are calculated from the Fourier transforms of the vertical LADCP and CTD data for shear and strain respectively. To calculate shear variance $\langle V_z^2 \rangle$, segments are constructed starting from the bottom in 320 m half-overlapping windows, each tapered with a 10% sine^2 window function to obtain its vertical wavenumber spectra (Kunze et al., 2006), which are then integrated between wavelengths of 320 m and 150 m to avoid high wavenumber instrument noise contamination (Kunze et al., 2006).

Strain is calculated from the buoyancy frequency as $\zeta_z = (N^2 - \bar{N^2})/\bar{N^2}$, where the mean stratification $\bar{N^2}$ is determined from quadratic fits to the profile segments (Kunze et al., 2006). Further, the strain variance is calculated by integrating the strain power spectrum between wavelengths of 150 m and up to 10 m while also satisfying strain variance $\langle \xi_z^2 \rangle < 0.2$ to avoid underestimating the variance through oversaturation of the spectrum (Gargett, 1990). $\langle V_z^2 \rangle$ and $\langle \xi_z^2 \rangle$ values are then normalized by the integrated GM model spectrum over the same respective bandwidths to represent the energy density in the internal wave field in the units of the GM energy density (Gregg & Kunze, 1991; W. Munk, 1981).

Each GM-normalized shear and strain spectrum is further normalized with its respective shear and strain variances across the finescale integration band to de-emphasize the known relationship between internal wave spectral level and stratification (Gregg, 1989; Kunze, 2017a). Further, we consolidate the dominant types of spectral variability by reducing the dimensionality of the data using Non-Negative Matrix Factorization (NMF) decomposition (Figure 1; described further in Section 3.2.1).

3.2.1 NMF Decomposition of Shear and Strain Spectra

While unsupervised learning could in theory identify clusters in any N-dimensional space, the quality of the resultant clustering formulation is directly proportional to the number of data points in the N-dimensional space. It is therefore prudent to introduce a low-rank approximation of the input N-dimensional space to reduce redundant co-variances in the data. Factor analysis and principal component analysis (PCA) are two of the many classical methods used to accomplish the goal of dimensionality reduction and detecting structures among the variables. Often the data to be analyzed are non-negative, and the low-rank data are further required to be comprised of non-negative values in order to avoid contradicting physical realities. Therefore, we reduce the dimensionality of the input spectral data using non-negative matrix factorization (NMF) (Lee & Seung, 1999; Berry et al., 2007) to decompose high-dimensional spectra of shear and strain into lower-dimensional latent spectral representations (Figure 1b). These low-dimensional embeddings (Figure 1c (green box),2a-d) are further aggregated into a feature matrix along with other auxiliary features (Figure 2e-g) and are then used as feature inputs to the GMM model (Figure 1c, Section 3.2.2).

The decomposition aims to approximate the input data matrix $\mathbf{X}$, consisting of non-negative elements, comprised of n individual spectral data points each with m wavenumbers, into a low-rank non-negative approximation consisting of a latent feature matrix W and corresponding hidden coefficients H. This can be expressed as: $\mathbf{X}_{[n\times m]} \approx \mathbf{W}_{[n\times p]}\mathbf{H}_{[p\times m]}$ (Figure 1b). The matrix W can be regarded as spectral end-members whose linear combinations with the coefficient matrix H reconstruct the original data matrix X. The quality of the approximation of X is measured using the Frobenius norm $||X - WH||_F^2 = \sum_{ij}(X - WH)_{ij}^2$ and the optimization algorithm is carried out using the NMF implementation in the Python library *scikit-learn* (Pedregosa et al., 2011). In this study, the input data matrix $\mathbf{X}$ for both shear and strain spectra consists of $n = 67816$ total spectra respectively obtained along the 15 GO-SHIP hydrographic section descibed earlier in Section 2.1.

It is conceivable that the reconstruction of the original spectra gets progressively better with the increase in the number of NMF components (p) i.e. the addition of more latent dimensions. In theory, the number of latent dimensions is inversely proportional to the reconstruction error- hence $p = m$ would result in a perfect reconstruction as the additional latent dimensions could in theory encode more of the information present in the original input matrix $\mathbf{X}$. However, here we choose $p = 2$ i.e. *two* latent dimensions to represent high dimensional (m=10) shear and strain spectra, as it results in the greatest decrease in the reconstruction error with respect to the number of latent dimensions while still preserving relevant spectral characteristics (not shown). Although increasing the number of latent dimensions beyond two results in a better reconstruction of the original spectral matrix $\mathbf{X}$, it can be counter-productive from an unsupervised learning standpoint as it can lead to inconsistencies in the final solutions produced by the clustering model often referred to as the "curse of dimensionality" (Bishop, 2006).

3.2.2 Final Feature Matrix (F)

Two NMF components each of the shear and strain spectra respectively are aggregated into a "feature matrix" F (Figure 1c) and used as input to an unsupervised learning model (Section 3.3). The sensitivity of the final results (Section 4) to the introduction of additional relevant features in the feature matrix- including the shear variance $\langle V_z^2 \rangle$, buoyancy frequency [N] and internal wave aspect ratio R_ω, all derived from the parameterization in Equation (1) is explored in Section 4.3. *Note*: The primary results discussed hereafter other than those specifically noted, describe the results of using only the 4 NMF components, two derived from the shear spectra and two derived from the strain spectra (Figure 1c green box, 2a-d).

3.3 Unsupervised Learning of Turbulent Mixing Data

An unsupervised machine learning clustering technique is used to identify groups with similar shear and strain spectra characteristics by applying a Gaussian Mixture Model (GMM) framework (e.g., Maze et al., 2017). The algorithm assumes the dataset with D features can be explained as derived from a mixture of K Gaussian distributions in D dimensions, where each feature represents a new dimension describing the data. The GMM model computes the parameters mean μ_k, covariance Σ_k and weights λ_k using the Expectation-Maximization algorithm in order to maximize the likelihood of the data X belonging to cluster k, denoted by the conditional probability distribution $p(k|x)$. The probability that data X belongs to the k^{th} component of the mixture of Gaussian distributions is given by:

$$p(k|x) = \frac{\lambda_k \mathcal{N}(x; \mu_k, \Sigma_k)}{\Sigma_{k=1}^{K} \lambda_k \mathcal{N}(x; \mu_k, \Sigma_k)} \tag{6}$$

with the multivariate normal Gaussian distribution given by:

$$p(x; \mu_k, \Sigma_k) = \frac{1}{\sqrt{2\pi^D |\Sigma|}} \exp\Big[-\frac{1}{2}(x - \mu_k)^T \Sigma^{-1} (x - \mu_k) \Big] \tag{7}$$

The conditional probability $p(k|x)$ in Equation 6 over all clusters k equals 1. The GMM algorithm assigns the cluster label k to the component for which this conditional probability is maximum i.e. $k = \text{argmax}_x\, p\big(k|x\big)$. We further mask out data with a maximum conditional probability less than 70% i.e. $k = \text{argmax}_x[p\big(k|x\big) > 0.7]$ (Figure 4a, gray) to avoid the possibility of having cluster labels with similar probability densities potentially near strong eddy or frontal forcings (Jones et al., 2019).

The choice of the number of clusters is a subjective one, and depends on the desired application of the clustering problem. The number of optimal clusters can vary widely based on the criteria used for convergence, tuning and choice of hyperparameters used (such as type of covariances), as well as the amount of data and choice of feature inputs given to the clustering algorithm. Dimensionality reduction for shear and strain spectra using NMF decomposition and clustering with the GMM model in this study were implemented using open-source python machine learning library scikit-learn (Pedregosa et al., 2011). We validate the optimal number of clusters outputted from the GMM model initialized with a "full" covariance matrix based on Akaike and Bayesian information criterion (AIC and BIC) scores (Schwarz G, 1978; Konishi et al., 2004). The AIC and BIC scores were computed for the entire feature matrix F created with the entirety of the data collected from 15 sections (not shown) as well random subsets of it for K=2 to K=14. The scores computed from 50 bootstraps of the random feature matrix subsets show a minimum between K=7 and K=9 clusters (Figure 7, purple shading). This conclusion is consistent when using a different metric for optimal clustering, the silhouette coefficient (Rousseeuw, 1987) (not shown). Although we use K=7 as the optimal number of clusters, the final results described in Section 4 are quantitatively the same, regardless of the choice of the number of clusters between K=7 and K=9 (Section 4.3).

4 Results

Seven distinct clusters of data are identified using the GMM model, which we explore to gain insight into the physical and geographical patterns relevant to turbulent mixing. We also consider the spatial structure of clusters and their correspondence with patterns of mismatch between finescale and microstructure-derived estimates to further contextualize the results. The feature matrix F input to the GMM model is comprised of only the two NMF-components of the normalized shear spectra and two NMF-components of the normalized strain spectra (Figure 1c, green box) for approximately 70,000 data points, each representing a 100-m vertical segment of data collected from 1875 profiles along 15 GO-SHIP sections (Figure 2a-d, 3a). The GMM is constrained to 7 clusters, hereafter discussed and referred to by the arbitrarily assigned cluster number. In terms of relative proportions of the assigned cluster labels- Cluster 5 was the most prevalent, followed by Clusters 4, 3, 7, 2, 1, and finally 6 (Figure 3d).

It is insightful to disentangle and isolate the original raw input data associated with each cluster to identify patterns that could potentially be linked to underlying physical mechanisms. We use the final clustering assignments to construct a composite average of the original "raw" shear and strain spectra belonging to each of the 7 clusters prior to any normalization and NMF decomposition (Section 3.2). The spectral data are "raw" in the sense that these spectra in their original form are the basis of the shear and strain variance ($\langle V_z^2\rangle$, $\langle \zeta_z^2\rangle$) calculations in the finescale parameterization described in Equations 1 and 2. At the individual level, the spectral energy density of the raw spectra across all 15 sections span orders of magnitude and appear to have incoherent geographical distributions and spatial dependence. However, considering the individual spectra combined with their corresponding clustering labels, we find that the average composite spectra (Figure 4 d, e) have distinct spectral shapes and unique slope and roll-off characteristics in vertical wavenumber space. These perceptible spectral characteristics, combined with the cluster spatial distributions and dependence hint at the potentially differing underlying physical mechanisms responsible for the non-linear downscale energy transport and turbulent mixing in these regions.

4.1 Identification of Non-GM Spectral Conditions & Parameterized Mixing Bias

The finescale parameterization laid out in Equations 1-4 aims to represent nonlinear spectral energy transport in the vertical wavenumber domain based on arguments set forth by (Garrett & Munk, 1972, 1975, 1979), requiring careful treatment of deviations from this framework. The intent of parameterization is to encapsulate the non-linear internal wave-wave interaction within a finite amplitude and vertical length scales not only well resolved by CTD and LADCP instrumentation (used for shear and strain calculation) and relatively free from contamination from instrumental noise or background stratification, but also from the effects of competing physical and dynamical processes such as near-boundaries mixing, wave-mean interaction, shear-driven mixing, double diffusion which could potentially short-circuit the downscale energy transfer and the basis of the parameterization. In observations (e.g., Gregg et al., 1993; K. L. Polzin et al., 1995; Brink, 1995; Eriksen, 1998), the wavenumber shear spectra at smaller wavenumbers (<0.1 cpm) are relatively white (flat) with roughly equal distribution of shear variance in this regime. The transition to turbulence occurs at length scales greater than 0.1 cpm governed by non-linear dynamics and shear instability-driven non-local energy transport (Gargett et al., 1981; Gregg et al., 1993). The finescale parameterization is employed to predict the turbulent dissipation from energy transport calculated at the intermediate scales (<0.1cpm, Figure 4c, d grey vertical lines). Here, large deviations from GM-model prescriptions can induce biases in the estimates and are potentially emblematic of additional physical processes at play beyond wave-wave interactions (K. L. Polzin et al., 2014).

Composite averages of shear and strain spectra computed within each cluster across all 15 sections (Figure 4c,d, Supporting Information Figure S3,S4) reveal *two* clusters (Cluster 1 and 7) with spectral characteristics differing significantly from the other clusters and from GM model spectra. Averaged shear spectra in Cluster 1 shows spectral levels comparable to other clusters but are characterized with steep ("red") slope compared to GM, with spectral roll-off at much lower wavenumbers and larger vertical scales than the other composites. At approximately the same vertical scales, shear spectra belonging to Cluster 7 show an enhancement in shear spectral power where the spectra appear "blue" and roll-off quite steeply after shear-enhanced hump. The shear-to-strain ratios (R_ω) implied by the Cluster 1 composite suggest a decrease in R_ω at higher vertical wavenumbers which could be interpreted as an increased contribution of high frequency waves at the lower wavenumbers using linear wave approximation (Equation 4). Studies have suggested that this is also possible due to the presence of quasi-permanent finestructure from rotating stratified turbulence (K. Polzin et al., 2003; K. Polzin & Ferrari, 2004).

The deviation from the assumptions about downscale spectral energy transport across wavenumbers in the parameterization is explored by comparing the ratio of the finescale parameterized estimates of turbulent dissipation rate ϵ to the concurrent co-located in-situ microstructure measurements of ϵ from CTD-mounted χ-pods (Lele et al., 2021) along the 2017 occupation of the P06 line within each cluster. The ratio of the two different estimates $\log_{10}\left(\frac{\epsilon_{\text{fine}}}{\epsilon_{\chi\text{pod}}}\right)$ or the "mixing bias" along the P06 section where positive (negative) values indicate finescale over-prediction (under-prediction) compared to measurements from $\chi-$pods (Figure 4a). The clustering from the GMM model combined with the mixing bias along the P06 are combined to produce estimates of average bias for each cluster (Figure 4c). The averaged mixing bias and 95% confidence intervals for clusters 2-6 fall well within a factor 5 (Figure 3c,dashed black line). Clusters 1 and 7 however, show a high and a low bias respectively, with average disagreement between finescale and

χ-pod estimates as large as an order of magnitude along P06. Further, the averaged spectral properties of the clusters also reveal marked deviations from their respective canonical GM shear and strain counterparts (Figure 4d,e). The inconsistencies between the rate of downscale energy transfer as prescribed by the GM model (e.g. Cluster 1) and possible shear-enhancing high-wavenumber energy sources (e.g. Cluster 7), serve as useful indicators of potential physical-dynamical processes unresolved in the finescale parameterizations.

We use the clustering assignments from the GMM model along the P06 section (Figure 4b), to compute the mixing bias for individual clusters, i.e. the averaged mixing bias corresponding to each cluster label along the section. We indicate the mean bias for each clusters with 95% confidence interval (Figure 4c, error bars) as well as the kernel density estimate showing the overall distribution of the mixing bias for individual clusters (Figure 4c, violin plot). The averaged mixing bias and 95% confidence intervals for Clusters 2-6 fall well within a factor 5 (Figure 3c, (dashed black line)).

The formulation of the finescale parameterization in Equation 1 states that diffusivity κ and dissipation rate ϵ (through the Osborn relation, Section 3.1) are proportional to the total integrated shear variance from shear spectra $\langle V_z^2 \rangle$. Considering this relationship between $\langle V_z^2 \rangle$ and ϵ, the mixing biases between finescale parameterized observed along P06 between ϵ_{fine} and $\epsilon_{\chi\text{-pod}}$ likely occurs as a result of the overestimation (underestimation) of $\langle V_z^2 \rangle$ in locations where Cluster 1 (Cluster 7) occur (Figure 4f). In the case of Cluster 1, a "redder" than GM-like spectra (Figure 4e,f, pink line) results in an overestimation of shear variance due to the assumed spectral shape being GM-like or flat (Figure 4f, pink shading). The overestimated shear variance through the relationship described in Equation 1 engenders a highly inflated estimate of ϵ_{fine} by almost an order of magnitude (Figure 4c). A reverse mechanism occurs in the case of Cluster 7 in which an increasingly positive slope ("bluer") compared to the GM-like spectra leads to an underestimation of $\langle V_z^2 \rangle$ and consequently a depressed estimate of ϵ_{fine}.

Using limited-modes of spectral variation through the NMF decomposition along 15 sections as inputs to the GMM model, we isolated *two* regions where underlying shear and strain spectra have characteristics to induce biases in parameterized mixing estimates. Spectral properties obtained in other process-based studies and certain localized environments have shown similarities to spectral features we identify here using a global dataset. Several different physical mechanics have been proposed in which non-white gradient spectra are associated with physics unresolved or problematic for finescale estimation (Kunze et al., 2002; Klymak et al., 2008; K. L. Polzin & Lvov, 2011; Brink, 1995; Eriksen, 1998). For example, well resolved spectra from a study around the Kergulean Plateau region (Waterman et al., 2013; Waterman, Polzin, Garabato, et al., 2014) associated with finescale overestimation exhibit steeper and rapid roll-offs at lower wavenumber attributed to strong wave-mean interactions in the region, similar to spectra found in Cluster 1. Similarly, generation or reflection at boundaries can inject shear at higher wavenumbers with loss of low-wavenumber energy and gain in high-wavenumber energy (Eriksen, 1985), as seen in composites from Cluster 7. Although diagnosing and interpreting the plethora of possibilities in the physics driving such peculiarities in the spectral energy transports in wavenumber space is beyond the scope of the paper- we further aim to prognosticate the spatial structure distribution of regions of potential finescale mixing bias along these sections.

4.2 Geographical Distribution

The spatial distribution of the clusters shows a rich and varied geographical distribution along the 15 sections considered here (Figure 3a). From a high-level perspective, the clustering reveals a rough dependence on stratification as seen by the alignment in most sections with the contours of buoyancy frequency along those sections (Figure 3a, black lines). Even though the inputs to the GMM consist of buoyancy-normalized spectral data that have been standardized by their respective integrated variances in order to diminish the a priori stratification dependence, the clustering patterns nevertheless reveal an ostensible relationship with buoyancy frequency (N).

In addition to the geographical cluster assignments by individual sections, more insight into the distinguishing characteristics of the clustering patterns can be gained by looking at the spatial variations in probability densities of each clusters vertically (depth and height-above-bottom) and horizontally (along-section) for each section individually, as well as by computing composites encompassing all 15 sections (Figure 3 [b-c,e-f] 5, 6, Supporting Information Figure S1-S2).

The upper ocean was dominated by clusters 2, 3 and 4, each showing similar vertical distributions, with some differences in their zonal and meridional distributions. Cluster 2 forms the majority of clusters within the Southern Ocean, as seen by the increase in prevalence southward of 55^oS along S4P, I06 and P16S , and the peak of the latitudinal distribution from the zonal composite found around 62.5^oS (Figure 3e). No distinct patterns emerge in the zonal and meridional distribution of cluster 3 and 4, suggesting minimal geographical precedence (Figure 3e, f). In depth, all three clusters become more prevalent closer to the seafloor (Figure 3c), resulting in peaks around 3000 m, also reflecting the variations in bathymetry of the sections (Figure 3b). The presence of Cluster 2 in the Southern Ocean and other sections near the bottom bathymetry, for example, along the P06, P02 and A20 (Supporting Information Figure S2) is consistent with regions of low stratification and is seen clearly in the contours of buoyancy frequency along those sections (Figure 3a, black lines)

Above the ocean bottom, the analysis found an increase in the relative abundance of clusters 5 and 6. These two cluster are prominently found in upper ocean along most sections between 500 m and 2000 m. Cluster 6 is the least prominent of the assigned labels and forms only 7.8% of the total assigned clusters and is mostly found in the upper ocean, typically between 1500 m and 2000 m in the Atlantic ocean (e.g. a13, a16n 16s and a10 lines). Cluster 5 is the second most common upper ocean cluster other than Cluster 1, with no notable zonal preference.

Cluster 1, associated with “redder” shear spectra, is predominately found in the upper ocean along most sections, existing primarily between 500-1500 m depth (Figure 3 b), with the notable exception in the Ross Sea (S4P) and Gulf of Mexico (A20) which show a second mid-depth around roughly 4000 m (Figure 3a, Supporting Information Figure S1). In addition, the zonal section composite also reveals a strongly increased proportion of Cluster 1 along the equator, with a clear peak observed within 5 degrees of the equator. In addition, the cluster is found most often in the subtropics with it rarely observed at high latitudes (Figure 3e). Meridional variability in Cluster 1 is observed with a vast preponderance in the Southern Hemisphere’s subtropical Atlantic and Eastern Pacific (Figure 3f).

Cluster 7, associated with “bluer” shear spectra with enhanced energy at wavenumbers between 150-100 m,is distributed in the mid to deep oceans, forms roughly 10.4% of the total cluster labels along the 15 sections (Figure 3d) and is most prevalent above the bottom bathymetry with a peak around 500 m from the bottom bathymetry (Figure 3c). Zonally, an increased proportion of cluster 7 is found in the Southern Ocean, scattered vertically throughout the sections, with cluster 7s found from the surface all the way down to the bottom topography (e.g. S4P).

4.3 Sensitivity of GMM to Number of Clusters [K] and additional feature inputs[d]

To test the robustness of the findings discussed above, we explore the sensitivity of this study to two key analysis choices. First, the effect of constraining the number to cluster to 7 is tested, and second, the effect of adding additional features to the GMM model is explored.

The analysis was run with a range of fixed number of clusters (K) ranging from 2 to 14. The optimal number of class labels requires model hyperparameter tuning, and the results are shown by the BIC scores (Figure 7). The BIC score shows a minima at K=7, but with some ambiguity for K=7-9. Here, we discuss sensitivity of our final results and conclusions to the clustering produced by the GMM model with the same four spectral inputs inputs, but with K=8 (Labels 0-7) as the optimal number of clusters here. The clustering distribution using K=8 and four inputs is overall very similar to the distribution with K=7. The additional 8^{th} class label is assigned to regions in the upper ocean and seems to split regions assigned to Cluster 1 in Figure 3a into two regions with labels 4 and 6 (Supporting Information Figure S5a).

This is further supported by the mean strain and shear spectra calculated from composite averages of individual cluster labels across the whole dataset (Supporting Information Figure S5b,c). Clusters 4 and 6 are associated with shear spectra with negative slopes whereas Cluster 5 (similar to Cluster 7 in Figure 3a) has a positive slope. Further, we see a similar association of finescale biased ϵ with regions in Clusters 4 and 6 over-predicting ϵ, while regions within Cluster 5 under-predict ϵ compared to measurements from χ-pods along the P06 section. Similarly, we tested K=9 (not shown) and it did not change the key findings of this study. Thus, while minor qualitative differences are to be expected with clustering assignments for each of the clusters with the results discussed in Figure 3 and 4, we do not find any quantitative differences in the iteration of results discussed above with the final conclusions of the study.

Second, the sensitivity of the final results to the incorporation of additional features as inputs to the GMM model is explored. The decision boundaries delineating one cluster from the next in the GMM model is a function of the means and covariances that describe the multi-dimensional Gaussian distributions. In general, addition or subtraction of feature inputs to the clustering model, aside from varying the dimensionality of the clustering space, can greatly affect these means and covariances and as a consequence the delineation and distribution of individual clusters in space. In an effort to critique the final results as not merely serendipitous artifacts attributable to the choice of feature inputs, various permutations of feature inputs to the GMM are explored, all derived from parameters in the finescale parameterizations (Section 3.1, Equations 1-5). We compare our main results to a GMM run using seven feature inputs consisting of four shear and strain spectra NMF decompositions, shear variance $\langle V_z^2 \rangle$, buoyancy frequency (N) and the internal wave aspect-ratio R_ω (Figure 2a-g) with seven output clusters (Labels 0-6).

Compared to the four feature run presented in the main text, the seven feature run produces clusters that are highly correlated to buoyancy frequency as seen in the alignment with buoyancy frequency contours along most of the 15 sections (Supporting Information Figure S6a, black solid lines). The results also show an overall higher posterior probability of clustering assignment as seen in the reduction in probability mask applied for posterior probabilities less than 70% (Supporting Information Figure S6a, grey mask). However, computing averaged strain and shear spectra composites for each cluster as before shows two clusters associated with large deviations from the GM-model shape (Supporting Information Figure S6b,c, Clusters 2 and Cluster 4). Biases in finescale ϵ estimates also exist for the same two clusters with regions along Cluster 4 overpredicting and regions along Cluster 2 underpredicting ϵ compared to observations to χ-pods along the P06 section (Supporting Information Figure S6d). With no significant quantitative differences in the results relating to the finescale bias, we recentered the focus of the main text on describing and discussing the results from the four-feature GMM output (Figure 2a-d).

5 Conclusions

In this study, we use a novel unsupervised learning approach with a Gaussian Mixture Model (Jeff A. Bilmes, 1998; Bishop, 2006) to cluster and identify patterns of turbulent mixing-related features derived from fundamental constituents of finescale parameterizations of internal wave-driven turbulent mixing in the ocean using a global dataset of ship-based hydrographic CTD and LADCP data collected on 15 GO-SHIP lines. Using an NMF decomposition of oceanic shear and strain spectra, we extracted spectral features consisting of encoded information about spectral level, shapes and slopes (Section 3.2,Figure 1, 2). These features once aggregated into a feature matrix are clustered using the GMM model into seven different domains characterizing and delineating their collective variation in the N-dimensional space represented by the extracted features. The class labels roughly align with stratification in the ocean on average vary with depth and height-above-bottom across the global ocean. Latitudinal and longitudinal variations among the clusters are more convoluted.

Further, we explore the implications and potential effects of spectral deviations in wavenumber space from the canonical Garrett and Munk (GM) internal wave spectrum (Garrett & Munk, 1972, 1975, 1979), for application of the finescale parameterizations to global data collected along 15 GO-SHIP sections. We identify the average composite shear and strain spectra associated with each of the seven clusters revealing two clusters (Cluster 1 and Cluster 7) associated with distinct spectra differing significantly from both the other composites and GM model in their wavenumber distribution of shear and strain spectral energy. Since the wavenumber distribution dictates the rate of energy transport and downscale energy transfer from large to smaller scales and ultimately to wave-breaking scales, the spectral characteristics within each cluster are ultimately tied to underlying physical mechanisms at play for turbulent mixing to occur in those regions. While uncovering the underlying mechanisms at play driving each cluster's spectral distribution is beyond the scope of this paper, we explore the robustness of mixing estimates obtained from finescale parameterizations in these regions further.

Studies have previously found large biases in finescale parameterized estimates where physical and dynamical environments short-circuit the underlying assumptions of the parameterizations, for e.g. regions in the surface mixed layer, near boundaries or where turbulent mixing is controlled by double diffusion, hydraulic jumps and strong wave-mean flow interactions over rough topography (Waterman, Polzin, Naveira Garabato, et al., 2014; K. L. Polzin et al., 2014; MacKinnon & Gregg, 2003). Our analysis is consistent with prior studies regarding the broad agreement between finescale parameterized and microstructure estimates of mixing in the open ocean thermocline where the underlying assumptions made in the parameterizations apply (K. L. Polzin et al., 1995, 2014; Whalen et al., 2015; Waterman, Polzin, Naveira Garabato, et al., 2014). However, based on the wavenumber distribution of global oceanic shear and strain spectra, we provide a rationale behind large biases in finescale parameterized estimates as well as identify their potential global spatial distribution based on data along 15 GO-SHIP lines.

Two clusters associated with high and low-biased finescale ϵ estimates when compared to co-located temperature microstructure observations from χ-pods along the P06 section were identified. The clusters are distinct in their spatial distribution along the P06 section. Cluster 1 associated with regions of finescale overestimation is primarily found in the upper ocean between 500 m and 2000 m in depth, while Cluster 7 is linked to regions of finescale underestimation and is found mostly in the deeper ocean below 3000 m with an increased abundance roughly 500 m-1000 m above the bottom bathymetry (Figure 3b,c). Both along the P06 section and averaged globally, the two cluster regions consist of roughly 20% of the total clustered data.

More work is needed to further our understanding of the underlying dynamical processes and the geographical distribution of various flavors of internal wave-wave interactions found in the ocean. Regardless of the cause, this study has shown that caution must be used when applying finescale parameterizations ubiquitously throughout the ocean. We show there are regions of the ocean where the prevalence of more ” “red” or more “blue” spectra energy could lead to biases in estimates of mixing derived from finescale parameterization that assume a GM-like universal form. This study could serve as a template to apply unsupervised machine learning approaches to localized process-based hydrographic studies or in engineering innovative features derived from hydrographic observations in an effort to understand the geographical and spatial distribution of the underlying dynamics.

Line	Profiles	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
a20	74	2.9	9.4	17.3	28.8	22.8	6.2	12.6
a22	90	1.9	15.2	15.7	28.5	17.1	6.9	14.8
a13	128	10.2	1.5	13.1	19.3	36.5	12.9	6.4
a16n	119	8.5	1.6	12.9	18.5	35.5	15.3	7.8
a16s	111	9.5	2.2	15.0	24.7	27.4	10.6	10.6
i06	56	13.2	22.7	11.4	7.6	33.0	2.5	9.5
i08	114	5.7	12.8	13.7	16.0	34.6	5.2	12.0
i07	110	8.6	2.4	15.4	23.5	31.4	9.0	9.7
p18	209	2.6	2.9	17.5	34.9	19.2	10.5	12.4
p16s	86	8.6	20.5	11.6	14.0	29.8	2.8	12.6
p02	159	9.8	17.6	11.7	14.0	31.1	4.9	10.9
a10	116	16.5	1.6	11.6	17.0	35.7	10.9	6.7
s4p	72	0.5	36.1	11.9	19.6	9.1	3.6	19.2
p06	244	7.9	10.8	12.9	18.5	31.8	7.1	11.0
i05	187	25.3	10.8	9.7	8.2	35.7	3.0	6.3
TOTAL	1875	9.4	9.8	13.4	19.6	29.6	7.8	10.4

Table 1. Total number of full-depth profiles for each of the 15 GO-SHIP lines in the study along with the percentage distribution of each cluster from the GMM model output along a given line corresponding to results described in the main text and Figure 3.

Acknowledgments
We thank the WOCE, CLIVAR, and GO-SHIP programs for collection of the hydrographic and LADCP data used in this study. CLIVAR and GO-SHIP data were collected and made publicly available by the International Global Ship-based Hydrographic Investigations Program and the national programs that contribute to it. Lele was supported by the NASA FINESST program (Grant 80NSSC20K1609). Lele, Purkey, MacKinnon and Nash were supported by NSF (OCE-2023289). Purkey was also supported by NSF (OCE-2023545).

Open Research

All data used for this article are publicly available. Hydrographic and LADCP data for hydrographic lines used in this study (Table 1) can by searched by the line “number” through the CCHDO (http://cchdo.ucsd.edu) and

University of Hawaii (https://currents.soest.hawaii.edu/go-ship/ladcp/) website. The χ-pod data collected on P06 2017 Leg 1 and 2 are available from CCHDO (https://microstructure.ucsd.edu/#/cruise/320620170703) and (https://microstructure.ucsd.edu/#/cruise/320620170820) respectively.

References

Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., & Plemmons, R. J. (2007). Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics and Data Analysis*, *52*(1), 155–173. doi: 10.1016/j.csda.2006.11.006

Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York : Springer, [2006] ©2006.

Boehme, L., & Rosso, I. (2021). Classifying Oceanographic Structures in the Amundsen Sea, Antarctica. *Geophysical Research Letters*, *48*(5). doi: 10.1029/2020GL089412

Brink, K. H. (1995). Tidal and lower frequency currents above Fieberling Guyot. *Journal of Geophysical Research*, *100*. doi: 10.1029/95jc00998

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). *Machine Learning for Fluid Mechanics.* doi: 10.1146/annurev-fluid-010719-060214

Bryan, F. (1987). Parameter sensitivity of primitive equation ocean general circulation models. *Journal of Physical Oceanography*, *17*, 970–985. doi: 10.1175/1520-0485(1987)017⟨0970:psopeo⟩2.0.co;2

Callaham, J. L., Koch, J. V., Brunton, B. W., Kutz, J. N., & Brunton, S. L. (2021). Learning dominant physical processes with data-driven balance models. *Nature Communications*, *12*(1016). doi: 10.1038/s41467-021-21331-z

Chatfield, C., Bendat, J. S., & Piersol, A. G. (1987). Random Data: Analysis and Measurement Procedures. *Journal of the Royal Statistical Society. Series A (General)*. doi: 10.2307/2981634

Chinn, B. S., Girton, J. B., & Alford, M. H. (2016). The impact of observed variations in the shear-to-strain ratio of internal waves on inferred turbulent diffusivities. *Journal of Physical Oceanography*, *46*(11), 3299–3320. doi: 10.1175/JPO-D-15-0161.1

Cimoli, L., Mashayek, A., Johnson, H. L., Marshall, D. P., Naveira Garabato, A. C., Whalen, C. B., … MacKinnon, J. A. (2023). Significance of diapycnal mixing within the Atlantic meridional overturning circulation. *AGU Advances*, *4*(2), e2022AV000800.

Eriksen, C. C. (1985). Implications of Ocean Bottom Reflection for Internal Wave Spectra and Mixing. *Journal of Physical Oceanography*, *15*. doi: 10.1175/1520-0485(1985)015⟨1145:ioobrf⟩2.0.co;2

Eriksen, C. C. (1998). Internal wave reflection and mixing at Fieberling Guyot. *Journal of Geophysical Research: Oceans*, *103*. doi: 10.1029/97jc03205

Ganachaud, A., & Wunsch, C. (2000). Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*. doi: 10.1038/35044048

Gargett, A. E. (1990). Do we really know how to scale the turbulent kinetic energy dissipation rate ϵ due to breaking of oceanic internal waves? *Journal of Geophysical Research*, *95*(C9), 15971–15974. doi: 10.1029/jc095ic09p15971

Gargett, A. E., Hendricks, P. J., Sanford, T. B., Osborn, T. R., & Williams, A. J. (1981). A Composite Spectrum of Vertical Shear in the Upper Ocean. *Journal of Physical Oceanography*, *11*. doi: 10.1175/1520-0485(1981)011⟨1258:acsovs⟩2.0.co;2

Garrett, C., & Munk, W. (1972). Space-Time scales of internal waves. *Geophysical Fluid Dynamics*, *3*(3), 225–264. doi: 10.1080/03091927208236082

Garrett, C., & Munk, W. (1975). Space-time scales of internal waves: A progress report. *Journal of Geophysical Research*, *80*. doi: 10.1029/jc080i003p00291

Garrett, C., & Munk, W. (1979). Internal Waves in the Ocean. *Annual Review of Fluid Mechanics*, *11*. doi: 10.1146/annurev.fl.11.010179.002011

Giglio, D., Lyubchich, V., & Mazloff, M. R. (2018). Estimating Oxygen in the Southern Ocean Using Argo Temperature and Salinity. *Journal of Geophysical Research: Oceans*, *123*(6), 4280–4297. doi: 10.1029/2017JC013404

Gregg, M. C. (1989). Scaling turbulent dissipation in the thermocline. *Journal of Geophysical Research*, *94*(C7), 9686–9698. doi: 10.1029/jc094ic07p09686

Gregg, M. C. (1999). Uncertainties and limitations in measuring ϵ and χ(T). *Journal of Atmospheric and Oceanic Technology*, *16*, 1483–1490. doi: 10.1175/1520-0426(1999)016⟨1483:ualima⟩2.0.co;2

Gregg, M. C., & Kunze, E. (1991). Shear and strain in Santa Monica Basin. *Journal of Geophysical Research: Oceans*, *96*(C9), 16709–16719. Retrieved from `http://dx.doi.org/10.1029/91JC01385` doi: 10.1029/91JC01385

Gregg, M. C., Sanford, T. B., & Winkel, D. P. (2003). Reduced mixing from the breaking of internal waves in equatorial waters. *Nature*. doi: 10.1038/nature01507

Gregg, M. C., Winkel, D. P., & Sanford, T. B. (1993). Varieties of fully resolved spectra of vertical shear. *Journal of Physical Oceanography*, *23*. doi: 10.1175/1520-0485(1993)023⟨0124:VOFRSO⟩2.0.CO;2

Henyey, F. S., Wright, J., & Flatté, S. M. (1986). Energy and action flow through the internal wave field: An eikonal approach. *Journal of Geophysical Research*, *91*. doi: 10.1029/jc091ic07p08487

Itsweire, E. C., Koseff, J. R., Briggs, D. A., & Ferziger, J. H. (1993). Turbulence in stratified shear flows: implications for interpreting shear-induced mixing in the ocean. *Journal of Physical Oceanography*, *23*, 1508–1522. doi: 10.1175/1520-0485(1993)023⟨1508:TISSFI⟩2.0.CO;2

Ivezic, Z., Connolly, A., Vanderplas, J., & Gray, A. (2014). *Statistics, data mining, and machine learning in astronomy: a practical Python guide for the analysis of survey data.* doi: 10.5860/choice.52-0257

Jackett, D. R., & McDougall, T. J. (1997). A neutral density variable for the world's oceans. *Journal of Physical Oceanography*, *27*(2), 237–263. doi: 10.1175/1520-0485(1997)027⟨0237:ANDVFT⟩2.0.CO;2

Jeff A. Bilmes. (1998). A Gentle Tutorial of the EM Algorithm. *International Computer Science Institute*, *4*(510).

Johnson, C. W., Ben-Zion, Y., Meng, H., & Vernon, F. (2020). Identifying Different Classes of Seismic Noise Signals Using Unsupervised Learning. *Geophysical Research Letters*, *47*(15), 1–10. doi: 10.1029/2020GL088353

Johnston, T. M., & Rudnick, D. L. (2015). Trapped diurnal internal tides, propagating semidiurnal internal tides, and mixing estimates in the California Current System from sustained glider observations, 2006-2012. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *112*, 61–78. doi: 10.1016/j.dsr2.2014.03.009

Jones, D. C., Holt, H. J., Meijers, A. J., & Shuckburgh, E. (2019). Unsupervised Clustering of Southern Ocean Argo Float Temperature Profiles. *Journal of Geophysical Research: Oceans*, *124*(1), 390–402. doi: 10.1029/2018JC014629

Kaiser, B. E., Saenz, J. A., Sonnewald, M., & Livescu, D. (2022). Automated identification of dominant physical processes. *Engineering Applications of Artificial Intelligence*, *116*, 105496.

Klymak, J. M., Pinkel, R., & Rainville, L. (2008). Direct Breaking of the Internal Tide near Topography: Kaena Ridge, Hawaii. *Journal of Physical Oceanography*, *38*, 380–399. doi: 10.1175/2007JPO3728.1

Konishi, S., Ando, T., & Imoto, S. (2004). Bayesian information criteria and smoothing parameter selection in radial basis function networks. *Biometrika*, *91*(1). doi: 10.1093/biomet/91.1.27

Kunze, E. (2017a). The Internal-Wave-Driven Meridional Overturning Circulation. *Journal of Physical Oceanography*, *47*(11), 2673–2689. doi: 10.1175/JPO-D-16-0142.1

Kunze, E. (2017b). Internal-Wave-Driven Mixing: Global Geography and Budgets. *Journal of Physical Oceanography*, *47*(6), 1325–1345. Retrieved from `http://journals.ametsoc.org/doi/10.1175/JPO-D-16-0141.1` doi: 10.1175/JPO-D-16-0141.1

Kunze, E., Firing, E., Hummon, J. M., Chereskin, T. K., & Thurnherr, A. M. (2006). Global abyssal mixing inferred from lowered ADCP shear and CTD strain profiles. *Journal of Physical Oceanography*, *36*(8), 1553–1576. doi: 10.1175/JPO2926.1

Kunze, E., Rosenfeld, L. K., Carter, G. S., & Gregg, M. C. (2002). Internal waves in Monterey Submarine Canyon. *Journal of Physical Oceanography*, *32*. doi: 10.1175/1520-0485(2002)032⟨1890:IWIMSC⟩2.0.CO;2

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*, 788–791. doi: 10.1038/44565

Lele, R., Purkey, S. G., Nash, J. D., Mackinnon, J. A., Thurnherr, A. M., Whalen, C. B., ... Talley, L. D. (2021). Abyssal Heat Budget in the Southwest Pacific Basin. *Journal of Physical Oceanography*, *51*(11), 3317–3333. doi: 10.1175/JPO-D-21-0045.1

MacKinnon, J. A., & Gregg, M. C. (2003). Mixing on the Late-Summer New England Shelf—Solibores, Shear, and Stratification. *Journal of Physical Oceanography*, *35*(12), 2408–2424. doi: 10.1175/1520-0485(2003)033⟨1476:motlne⟩2.0.co;2

MacKinnon, J. A., Zhao, Z., Whalen, C. B., Waterhouse, A. F., Trossman, D. S., Sun, O. M., ... Alford, M. H. (2017). Climate process team on internal wave-driven ocean mixing. *Bulletin of the American Meteorological Society*, *98*(11), 2429–2454. doi: 10.1175/BAMS-D-16-0030.1

Mashayek, A., Reynard, N., Zhai, F., Srinivasan, K., Jelley, A., Naveira Garabato, A., & Caulfield, C. P. (2022). Deep ocean learning of small scale turbulence. *Geophysical Research Letters*, *49*(15), e2022GL098039.

Maze, G., Mercier, H., Fablet, R., Tandeo, P., Lopez Radcenco, M., Lenca, P., ... Le Goff, C. (2017). Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean. *Progress in Oceanography*, *151*. doi: 10.1016/j.pocean.2016.12.008

McDougall, P., Trevor J. ; Barker. (2011). Getting started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox. *Scor/Iapso Wg127*.

Moum, J. N., & Nash, J. D. (2009). Mixing Measurements on an Equatorial Ocean Mooring. *Journal of Atmospheric and Oceanic Technology*, *26*(2), 317–336. doi: 10.1175/2008JTECHO617.1

Müller, P., & Liu, X. (2000). Scattering of internal waves at finite topography in two dimensions. Part I: Theory and case studies. *Journal of Physical Oceanography*, *30*, 532–549. doi: 10.1175/1520-0485(2000)030⟨0532:SOIWAF⟩2.0.CO;2

Munk, W. (1981). Internal Waves and Small-Scale Processes. In *Evolution of physical oceanography: Scientific surveys in honor of henry stommel.*

Munk, W. H. (1966). Abyssal recipes. *Deep-Sea Research and Oceanographic Abstracts*. doi: 10.1016/0011-7471(66)90602-4

Naveira Garabato, A. C., Frajka-Williams, E. E., Spingys, C. P., Legg, S., Polzin, K. L., Forryan, A., ... Meredith, M. P. (2019). Rapid mixing and exchange of deep-ocean waters in an abyssal boundary current. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(27), 13233–13238. doi: 10.1073/pnas.1904087116

Naveira Garabato, A. C., Polzin, K. L., King, B. A., Heywood, K. J., & Visbeck, M.

(2004). Widespread Intense Turbulent Mixing in the Southern Ocean. *Science*, *303*, 210–213. doi: 10.1126/science.1090929

Oakey, N. S. (1982). Determination of the Rate of Dissipation of Turbulent Energy from Simultaneous Temperature and Velocity Shear Microstructure Measurements. *Journal of Physical Oceanography*, *12*, 256–271. doi: 10.1175/1520-0485(1982)012⟨0256:dotrod⟩2.0.co;2

Osborn, T. R. (1980). Estimates of the Local Rate of Vertical Diffusion from Dissipation Measurements. *Journal of Physical Oceanography*, *10*(1), 83–89. doi: 10.1175/1520-0485(1980)010{\%}3C0083:EOTLRO{\%}3E2.0.CO;2

Osborn, T. R., & Cox, C. S. (1972). Oceanic fine structure. *Geophysical & Astrophysical Fluid Dynamics*, *3*(4), 321–345. doi: 10.1080/03091927208236085

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pollmann, F. (2020). Global characterization of the ocean's internal wave spectrum. *Journal of Physical Oceanography*, *50*(7), 1871–1891. doi: 10.1175/JPO-D-19-0185.1

Polzin, K., & Ferrari, R. (2004). Isopycnal dispersion in NATRE. *Journal of Physical Oceanography*, *34*. doi: 10.1175/1520-0485(2004)034⟨0247:IDIN⟩2.0.CO;2

Polzin, K., Kunze, E., Toole, J. M., & Schmitt, R. W. (2003). The partition of finescale energy into internal waves and subinertial motions. *Journal of Physical Oceanography*, *33*. doi: 10.1175/1520-0485(2003)033⟨0234:TPOFEI⟩2.0.CO;2

Polzin, K. L., & Lvov, Y. (2011). Towards Regional Characterization of the Oceanic Internal Wavefield. *Reviews of Geophysics*, *49*. doi: 10.1029/2010RG000329

Polzin, K. L., Naveira Garabato, A. C., Huussen, T. N., Sloyan, B. M., & Waterman, S. (2014). Finescale parameterizations of turbulent dissipation. *Journal of Geophysical Research: Oceans*, *119*(2), 1383–1419. doi: 10.1002/2013JC008979

Polzin, K. L., Toole, J. M., Ledwell, J. R., & Schmitt, R. W. (1997). Spatial variability of turbulent mixing in the abyssal ocean. *Science*, *276*(5309), 93–96. doi: 10.1126/science.276.5309.93

Polzin, K. L., Toole, J. M., & Schmitt, R. W. (1995). Finescale Parameterizations of Turbulent Dissipation. *Journal of Physical Oceanography*, *25*(3), 306–328. doi: https://doi.org/10.1175/1520-0485(1995)025⟨0306:FPOTD⟩2.0.CO;2

Rosso, I., Mazloff, M. R., Talley, L. D., Purkey, S. G., Freeman, N. M., & Maze, G. (2020). Water Mass and Biogeochemical Variability in the Kerguelen Sector of the Southern Ocean: A Machine Learning Approach for a Mixing Hot Spot. *Journal of Geophysical Research: Oceans*, *125*(3), 1–23. doi: 10.1029/2019JC015877

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. doi: 10.1016/0377-0427(87)90125-7

Rudnick, D. L., Johnston, T. M., & Sherman, J. T. (2013). High-frequency internal waves near the Luzon Strait observed by underwater gliders. *Journal of Geophysical Research: Oceans*, *118*. doi: 10.1002/jgrc.20083

Schwarz G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2).

Shroyer, E. L., Rudnick, D. L., Farrar, J. T., Lim, B., Venayagamoorthy, S. K., Laurent, L. C., ... Moum, J. N. (2016). Modification of upper-ocean temperature structure by subsurface mixing in the presence of strong salinity stratification. *Oceanography*, *29*(2), 62–71. doi: 10.5670/oceanog.2016.39

Sonnewald, M., Sonnewald, M., Dutkiewicz, S., Hill, C., & Forget, G. (2020). Elucidating ecological complexity: Unsupervised learning determines global marine eco-provinces. *Science Advances*, *6*(22). doi: 10.1126/sciadv.aay4740

Sonnewald, M., Wunsch, C., & Heimbach, P. (2019). Unsupervised Learning Reveals Geography of Global Ocean Dynamical Regions. *Earth and Space Science*, *6*. doi: 10.1029/2018EA000519

St. Laurent, L., Naveira Garabato, A. C., Ledwell, J. R., Thurnherr, A. M., Toole, J. M., & Watson, A. J. (2012). Turbulence and diapycnal mixing in drake passage. *Journal of Physical Oceanography*, *42*, 2143–2152. doi: 10.1175/JPO-D-12-027.1

Talley, L. D. (2003). Shallow, Intermediate, and Deep Overturning Components of the Global Heat Budget. *Journal of Physical Oceanography*. doi: 10.1175/1520-0485(2003)033⟨0530:siadoc⟩2.0.co;2

Thurnherr, A. M., Visbeck, M., Firing, E., King, B. A., Hummon, J., Krahmann, G., & Huber, B. A. (2010). A Manual For Acquiring Lowered Doppler Current Profiler Data. *The GO-SHIP Repeat Hydrographic Manual: A Collection of Expert Reports and Guidelines*, *14*(134).

Visbeck, M. (2002). Deep velocity profiling using lowered acoustic Doppler current profilers: Bottom track and inverse solutions. *Journal of Atmospheric and Oceanic Technology*, *19*(5), 794–807. doi: 10.1175/1520-0426(2002)019⟨0794:DVPULA⟩2.0.CO;2

Waterhouse, A. F., Mackinnon, J. A., Nash, J. D., Alford, M. H., Kunze, E., Simmons, H. L., ... Lee, C. M. (2014). Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *Journal of Physical Oceanography*, *44*(7), 1854–1872. doi: 10.1175/JPO-D-13-0104.1

Waterman, S., Naveira Garabato, A. C., & Polzin, K. L. (2013). Internal waves and turbulence in the antarctic circumpolar current. *Journal of Physical Oceanography*, *43*(2), 259–282. doi: 10.1175/JPO-D-11-0194.1

Waterman, S., Polzin, K. L., Garabato, A. C., Sheen, K. L., & Forryan, A. (2014). Suppression of internal wave breaking in the antarctic circumpolar current near topography. *Journal of Physical Oceanography*, *44*(5), 1466–1492. doi: 10.1175/JPO-D-12-0154.1

Waterman, S., Polzin, K. L., Naveira Garabato, A. C., Sheen, K. L., & Forryan, A. (2014). Suppression of Internal Wave Breaking in the Antarctic Circumpolar Current near Topography. *Journal of Physical Oceanography*, *44*(5), 1466–1492. doi: 10.1175/JPO-D-12-0154.1

Whalen, C. B. (2021). Best Practices for Comparing Ocean Turbulence Measurements Across Spatiotemporal Scales. *Journal of Atmospheric & Oceanic Technology*, *38*(4), 837–841. doi: https://doi.org/10.1175/JTECH-D-20-0175.1

Whalen, C. B., MacKinnon, J. A., & Talley, L. D. (2018). Large-scale impacts of the mesoscale environment on mixing from wind-driven internal waves. *Nature Geoscience*, *11*(11), 842–847. Retrieved from `http://dx.doi.org/10.1038/s41561-018-0213-6` doi: 10.1038/s41561-018-0213-6

Whalen, C. B., MacKinnon, J. A., Talley, L. D., & Waterhouse, A. F. (2015). Estimating the Mean Diapycnal Mixing Using a Finescale Strain Parameterization. *Journal of Physical Oceanography*, *45*(4), 1174–1188. doi: 10.1175/JPO-D-14-0167.1

Whalen, C. B., Talley, L. D., & MacKinnon, J. A. (2012). Spatial and temporal variability of global ocean mixing inferred from Argo profiles. *Geophysical Research Letters*, *39*(17), 1–6. doi: 10.1029/2012GL053196

Wunsch, C., & Ferrari, R. (2004). Vertical mixing, energy, and the general circulation of the oceans. *Annual Review of Fluid Mechanics*, *36*(1), 281–314. doi: 10.1146/annurev.fluid.36.050802.122121

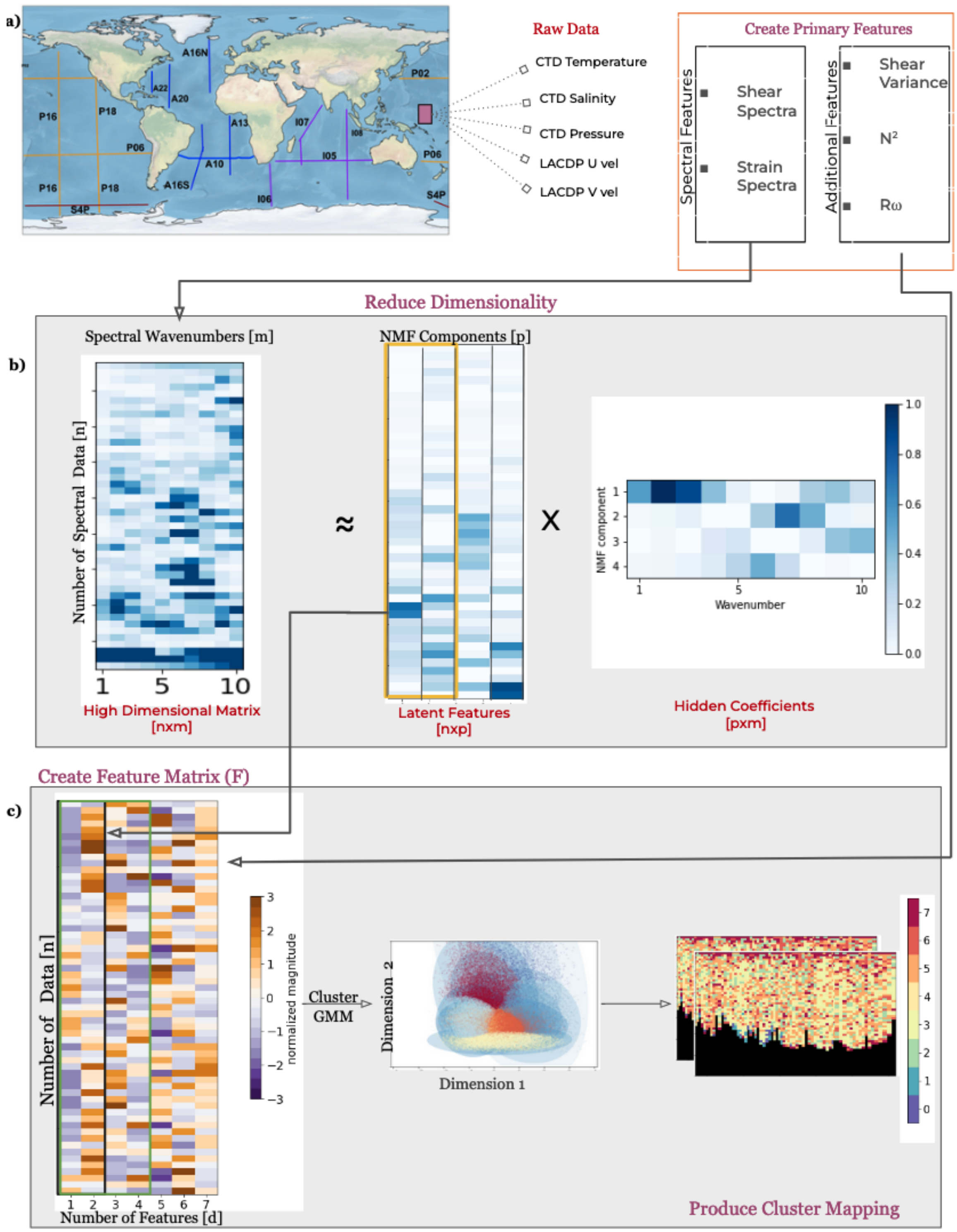


Figure 1. Method schematic shows the locations of the Raw CTD and LADCP data along the 15 GO-SHIP lines in the study which are used to create primary features consisting of spectral and non-spectral data (see Section 3.2). (b) An example showing dimensionality reduction through NMF decomposition for shear spectra, converting high m dimensional spectra in the input data matrix $\mathbf{X}$ into lower p dimensional spectral features in the form of a latent feature matrix W and a corresponding hidden coefficient matrix H respectively (see Section 3.2.1). (c) Two latent spectral features each (p=2) of shear and strain are aggregated into a feature matrix F (green box, see Section 3.2.2), with additional features (see Section 4.3) are used as inputs to the GMM model to generate cluster mappings for all GO-SHIP sections (Table 1).

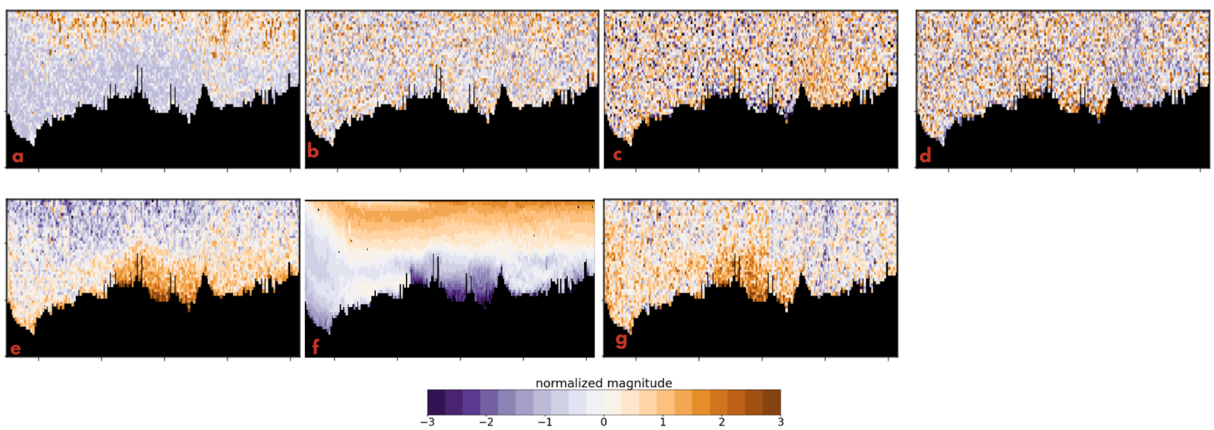


Figure 2. Example of geographical feature distribution along the P18 section for 7 different features with normalized magnitudes used for clustering using the GMM model including NMF-1 shear spectra (a), NMF-2 shear spectra (b), NMF-1 strain spectra (c), NMF-2 strain spectra (d), shear variance $\langle V_z^2 \rangle$ (e), buoyancy frequency [N] (f) and R_ω (g).

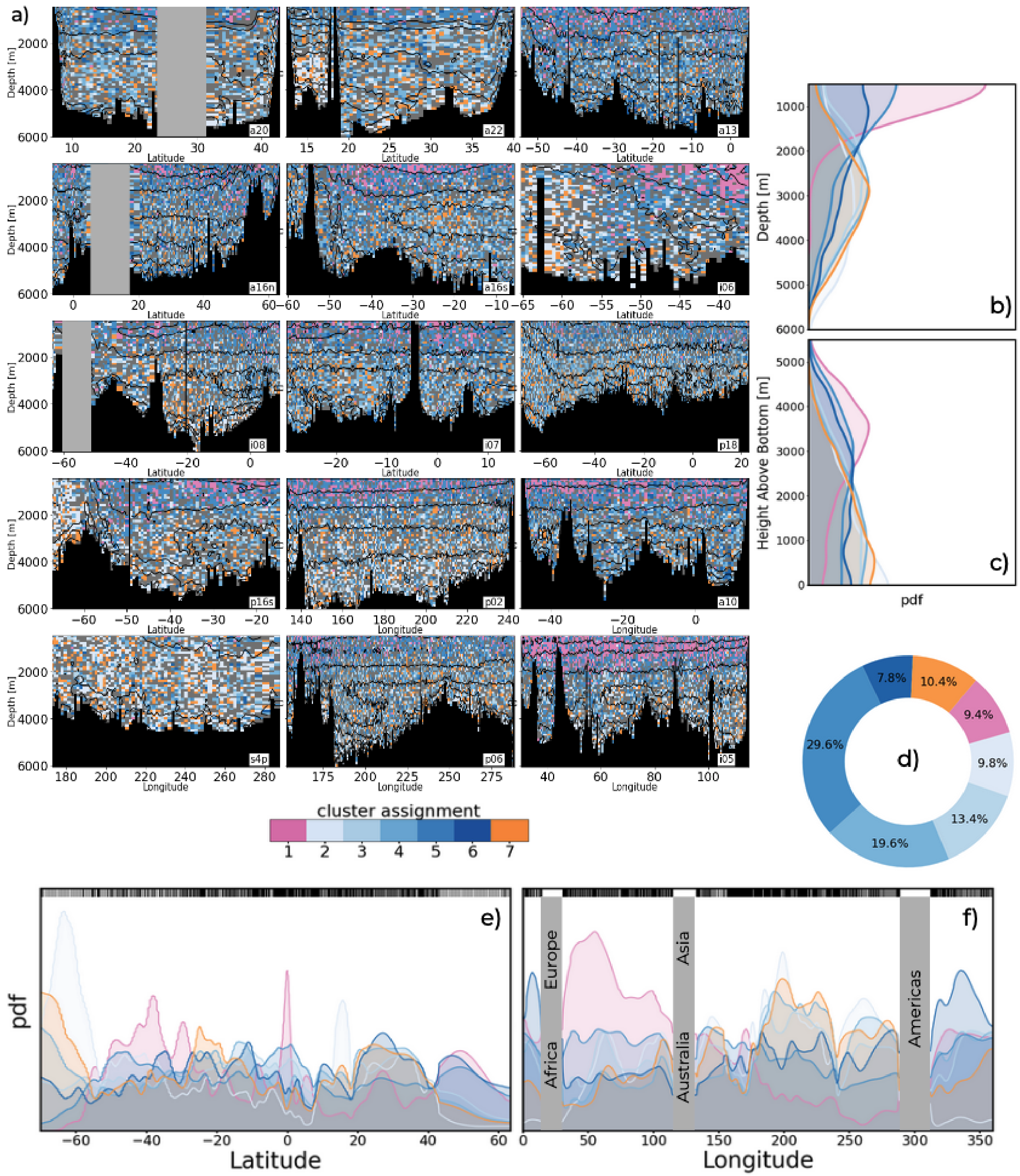


Figure 3. a) Resulting clustering along the 15 GO-SHIP lines produced by the GMM model with feature matrix based on the NMF features (Figure 2a-d). PDF showing the abundance and variation of individual clusters as a function of depth (b) and height above bottom bathymetry across all 15 GO-SHIP lines (c). Relative percentage distribution of the seven clusters from the GMM model with a posterior probability greater than 70% which are considered for the analysis (d). The zonal (e) and meridional (f) PDF computed from a composites of 10 zonal and 5 meridional sections respectively, with the location of CTD stations for the zonal and meridional sections are shown (black vertical lines in e and f).

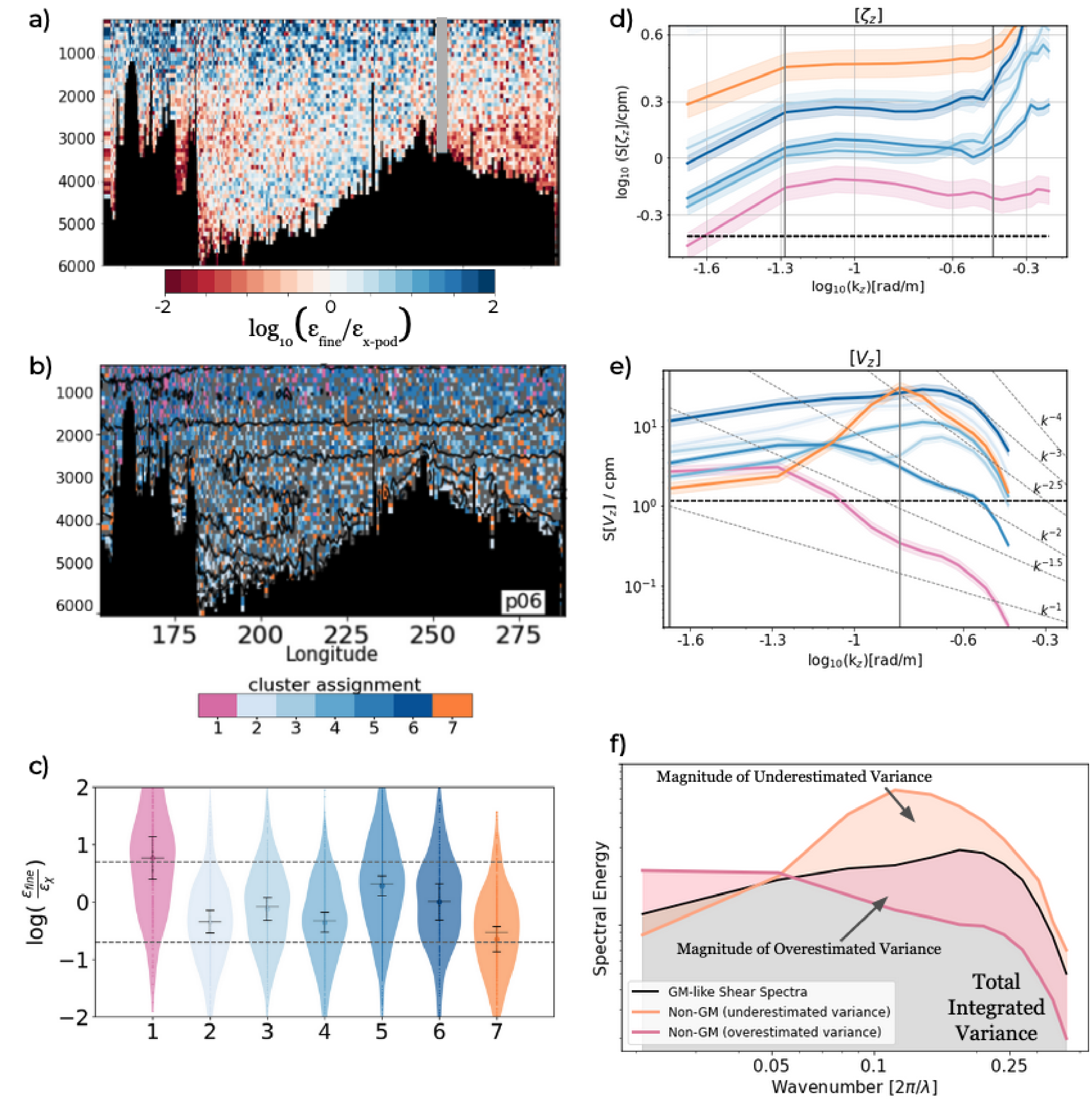


Figure 4. a) Ratio of estimates of turbulent dissipation rate from the finescale parameterization to measurements from CTD-mounted χ-pods taken concurrently along the P06 section expressed as $\log_{10}\left(\frac{\epsilon_{\text{fine}}}{\epsilon_{\chi\text{pod}}}\right)$, b) Cluster assignments from the GMM model along the P06 section (same as Figure 3a), c) Mean, 95% confidence intervals and violin plot computed for the ratio $\log_{10}\left(\frac{\epsilon_{\text{fine}}}{\epsilon_{\chi\text{pod}}}\right)$ from Figure 4a for seven clustered regions shown in Figure 4b, d-e). Mean strain and shear spectra computed as a composite average for the clusters computed using all 15 sections with 99% confidence intervals using computed using a χ^2 distribution (Chatfield et al., 1987) considering only $1/10^{th}$ degrees of freedom for better visibility (color shading). The average GM spectral levels are shown in the dashed black line, with the integration limits to calculate strain and shear variance shown by solid grey vertical lines. Slopes for shear spectra roll-offs between k^{-1} and k^{-4} are shown with high wavenumber asymptote k^{-2} representing inertial subrange in the GM model, f) Schematic outlining how biases in estimates of turbulent mixing could arise from spectra deviating from the assumed GM-like shape in the finescale parameterization (Equation 1) by either overestimating (Cluster 1) or underestimating (Cluster 7) shear $\langle V_z^2 \rangle$ and strain variance $\langle \zeta_z^2 \rangle$ calculated by integrating the respective spectra in the finescale wavenumber band.

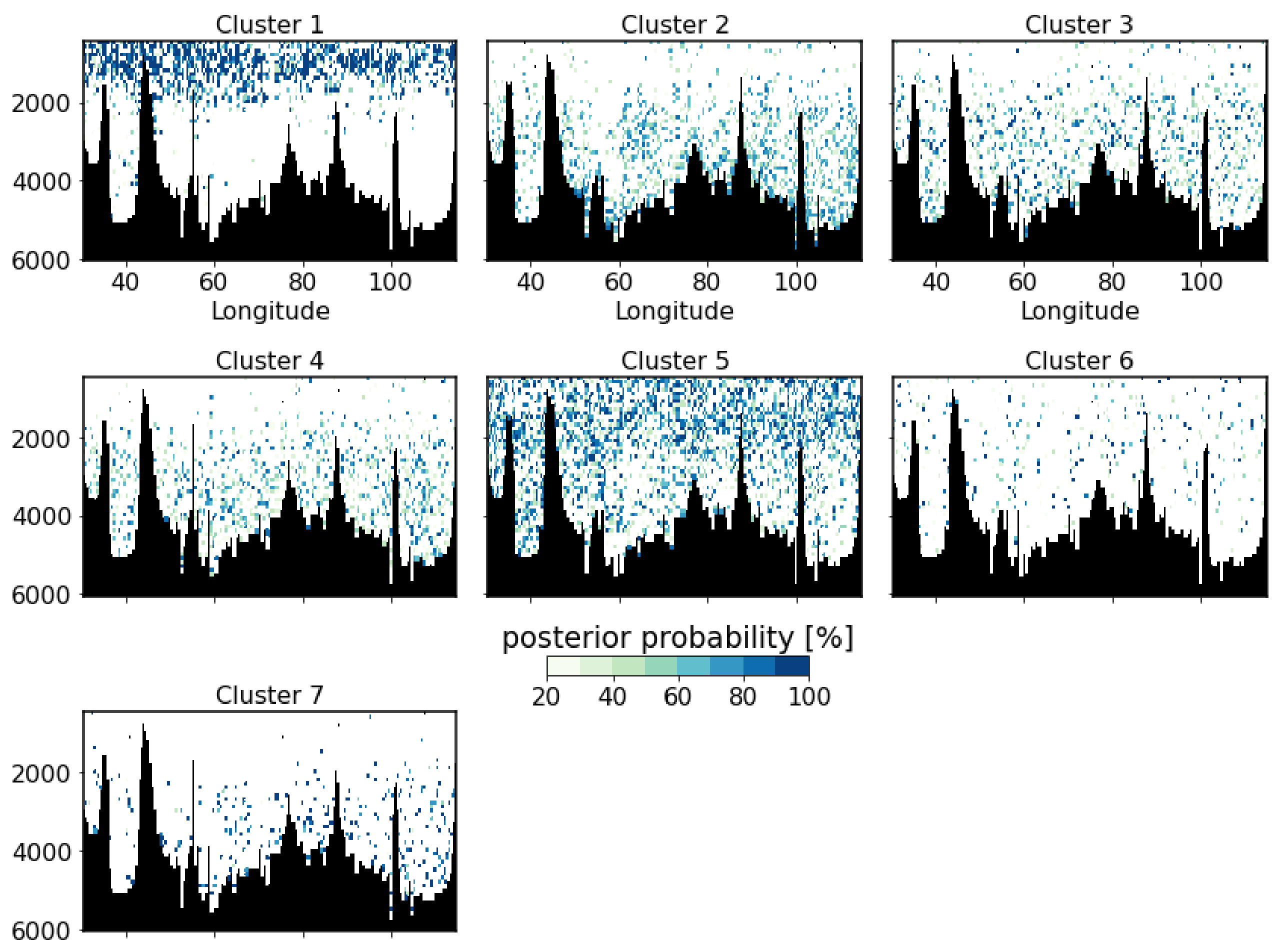


Figure 5. Posterior probabilities (%) $p(k|x)$ of data belonging to each of the clusters (1-7) as calculated with Equation 6 from the GMM model along the I05 section. Final cluster assignment of a data point belonging to a cluster k as shown in Figure 4 is made by computing $k = \text{argmax}_x\, p(k|x)$ as described in Section 3.3 .

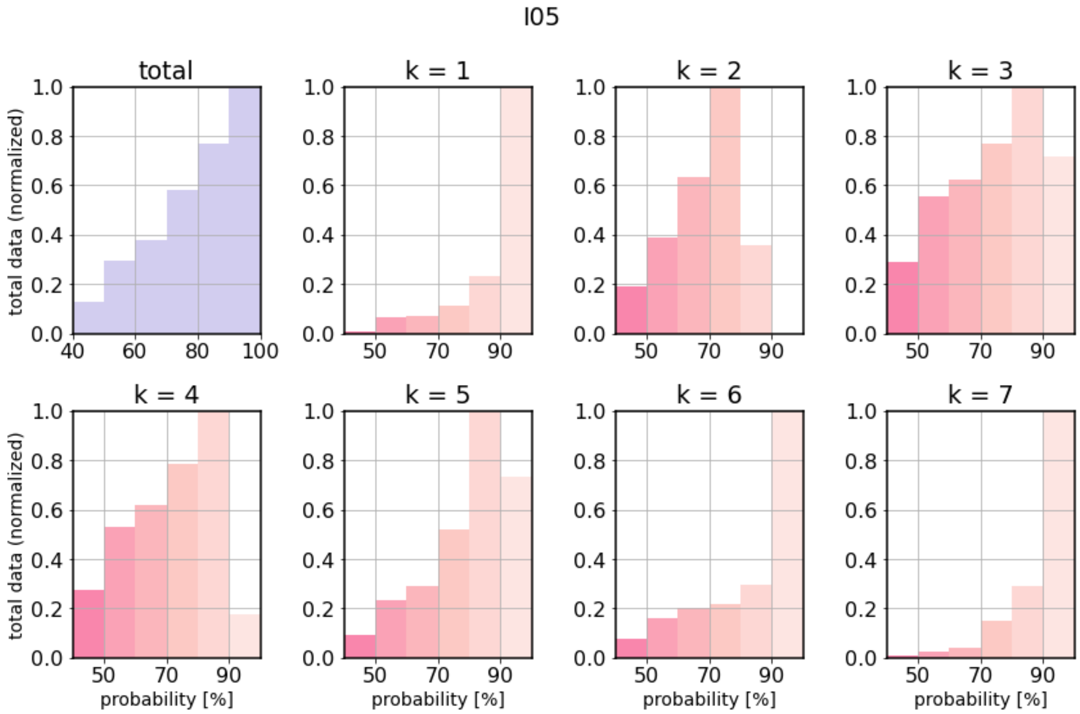


Figure 6. Histogram of the total percent posterior probabilities along the I05 section summed across all the clusters $\sum_{k=1}^{K} p\big(k|x\big)$in 10% bins between 40% to 100% (top left). Additionally, histogram of the percent posterior probabilities in each individual clusters k=1-7 corresponding to Figure 5 are displayed as well. Data displayed in each bin are normalized by number of data-points in the 10% bin with the most data.

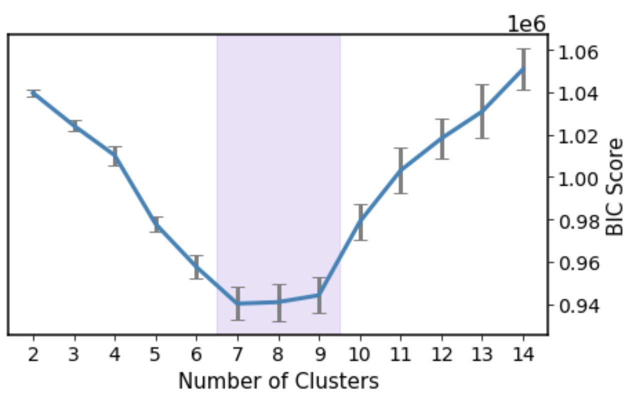
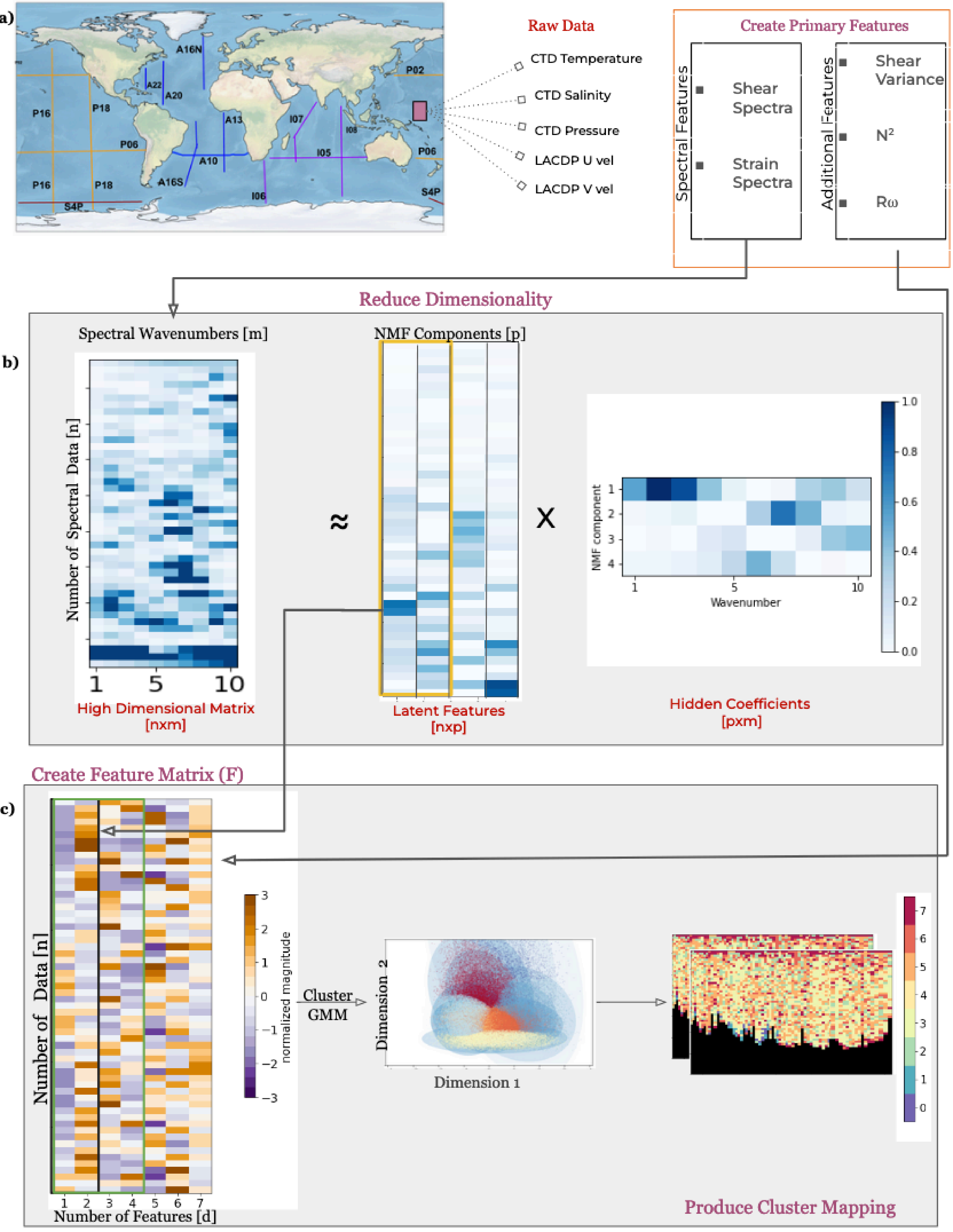


Figure 7. The BIC scores versus the specified number of clusters, with the means (solid blue line) and standard deviations (error bars) calculated from 50 random subsets of the data is also shown with the range of the smallest BIC values (between k=7 and k=9) is indicated (purple shading)

Figure 1.



a)
A16N
A22
A20
P02
P16
P18
A13
I07
I08
P06
I05
A10
A16S
I06
S4P
Raw Data
CTD Temperature
CTD Salinity
CTD Pressure
LACDP U vel
LACDP V vel
Create Primary Features
Spectral Features
Shear Spectra
Strain Spectra
Additional Features
Shear Variance
N²
Rω
Reduce Dimensionality
b)
Spectral Wavenumbers [m]
Number of Spectral Data [n]
1
5
10
High Dimensional Matrix [nxm]
≈
NMF Components [p]
Latent Features [nxp]
X
NMF component
Wavenumber
Hidden Coefficients [pxm]
Create Feature Matrix (F)
c)
Number of Data [n]
Number of Features [d]
normalized magnitude
Cluster GMM
Dimension 2
Dimension 1
Produce Cluster Mapping

Figure 2.

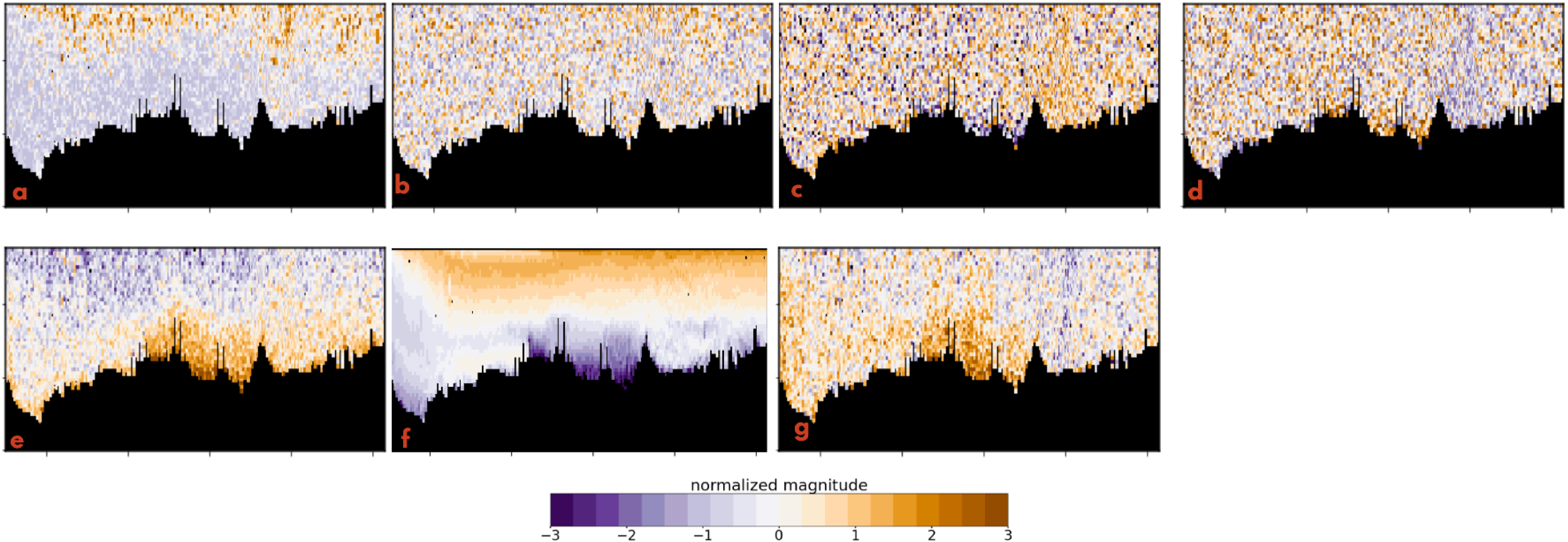

a
b
c
d
e
f
g
normalized magnitude
−3
−2
−1
0
1
2
3

Figure 3.

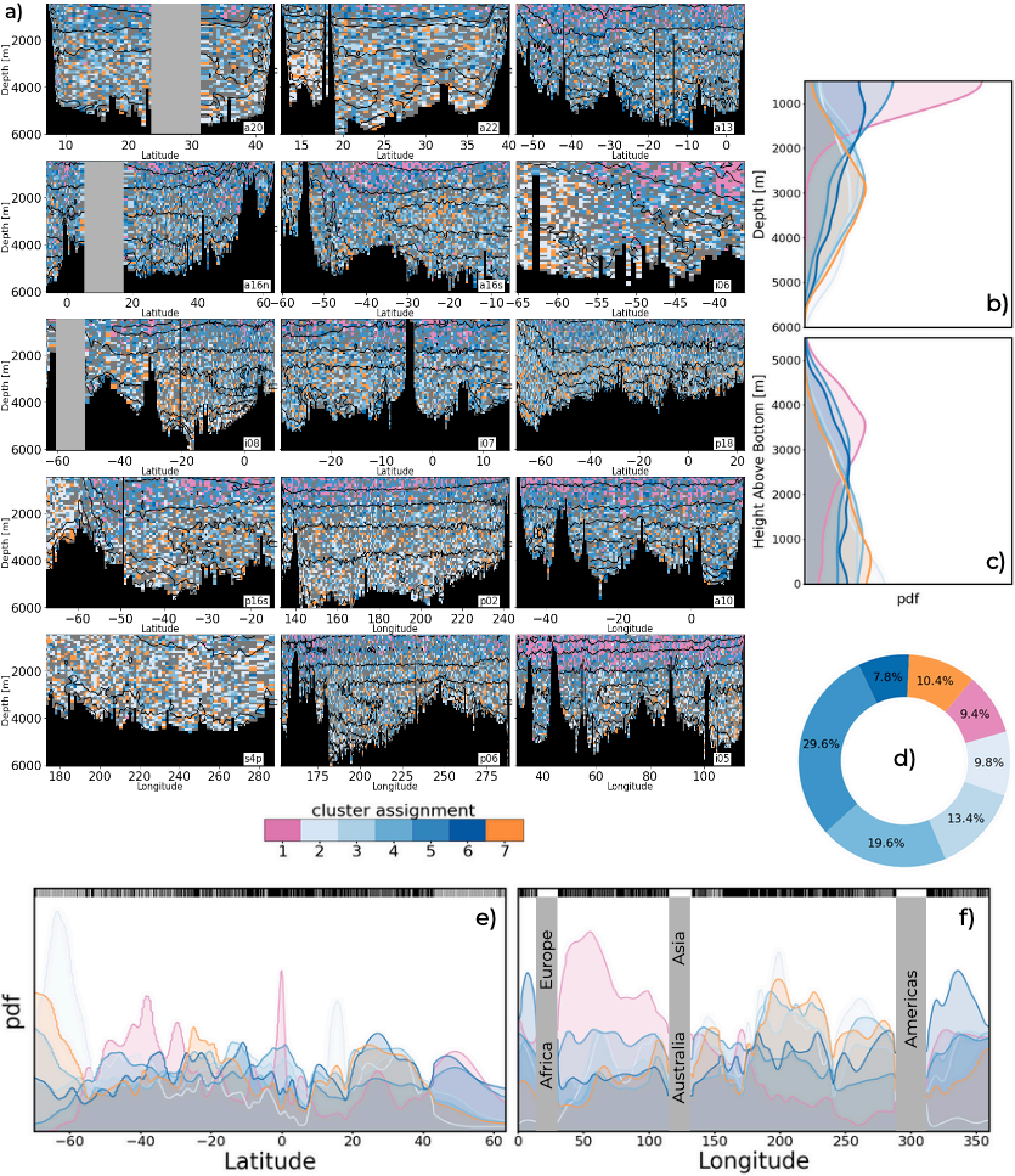
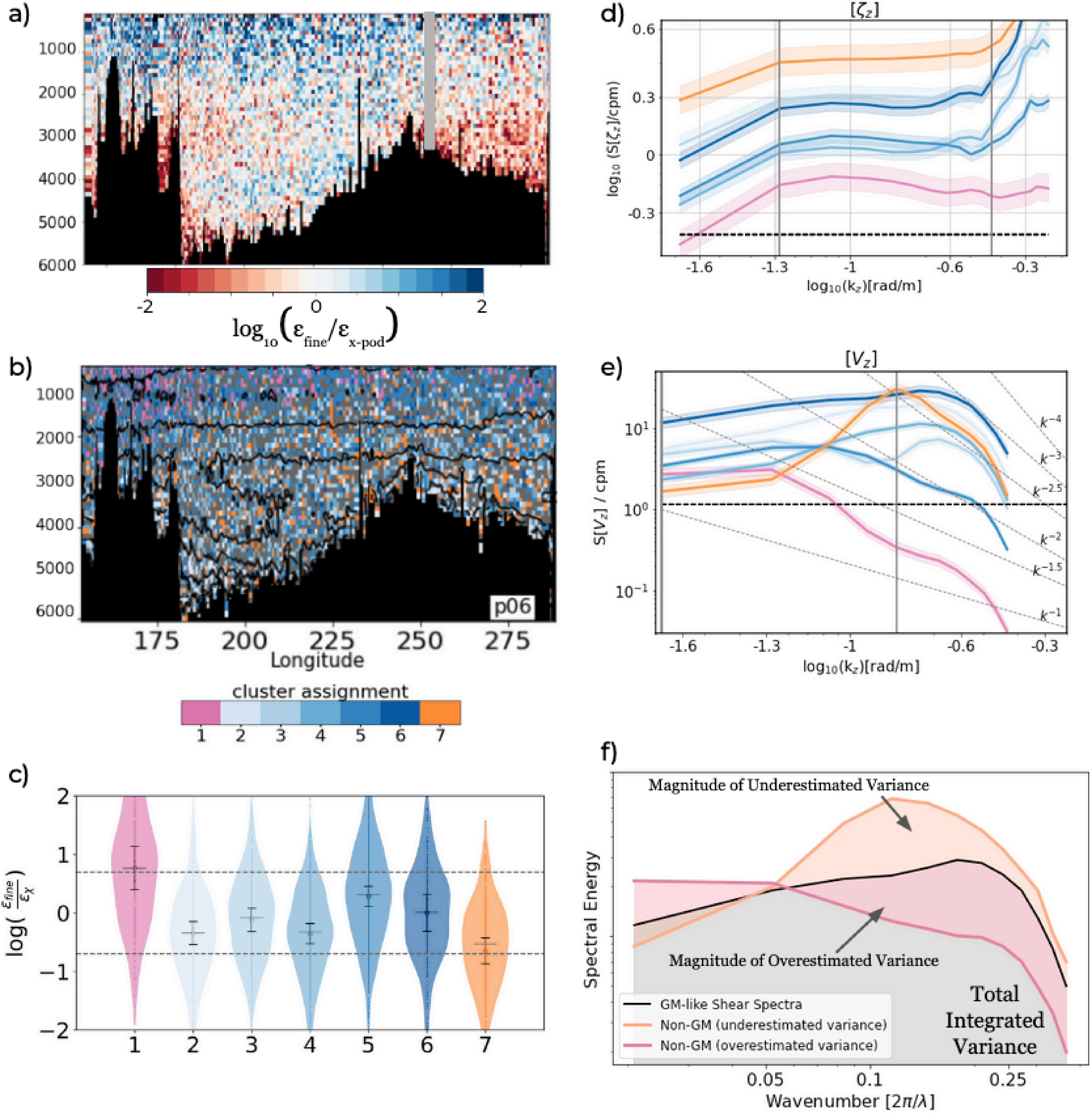
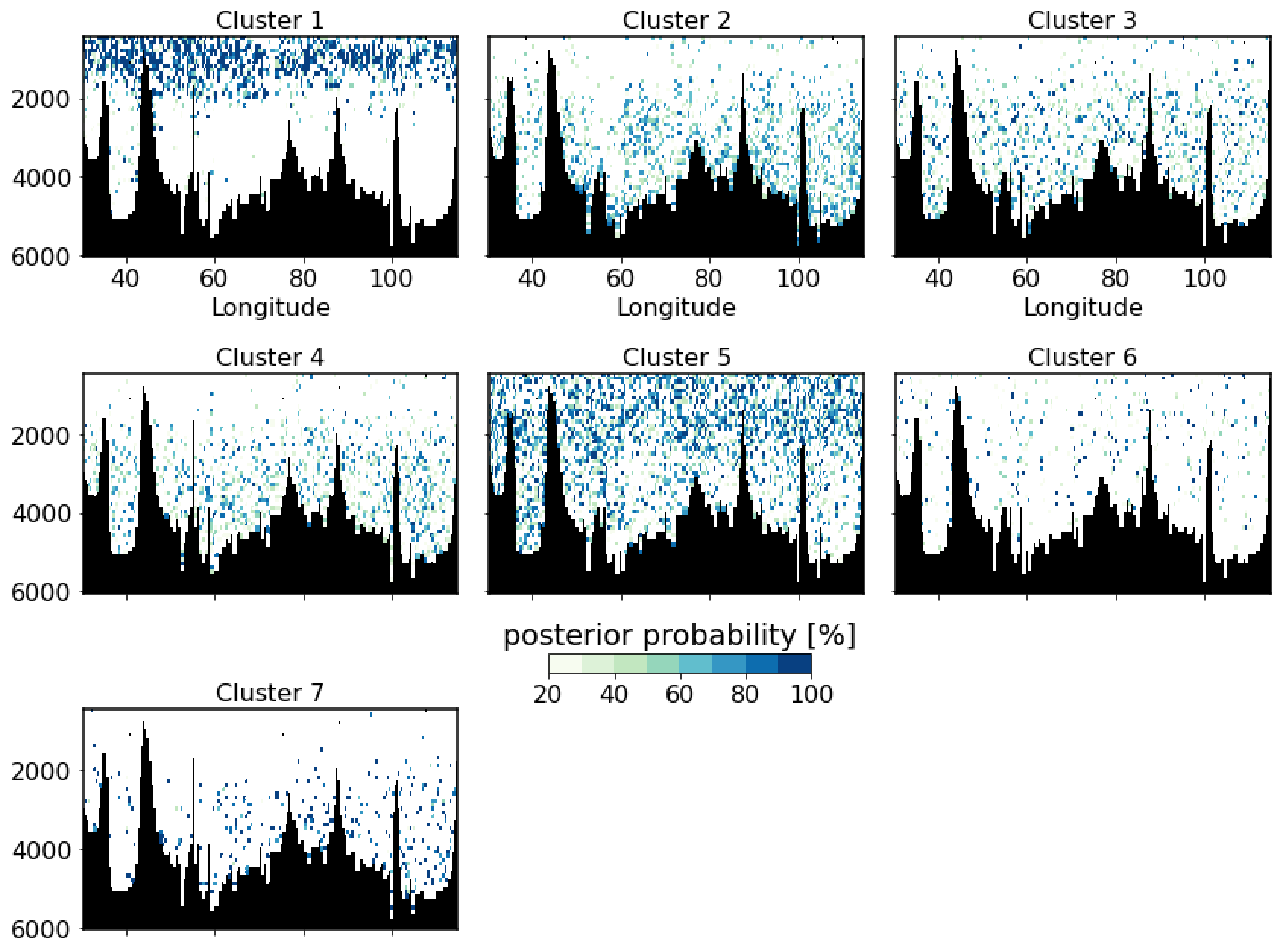

a)
b)
c)
d)
e)
f)
a20
a22
a13
a16n
a16s
i06
i08
i07
p18
p16s
p02
a10
s4p
p06
i05
Depth [m]
Latitude
Longitude
Height Above Bottom [m]
pdf
cluster assignment
1 2 3 4 5 6 7
7.8%
10.4%
9.4%
9.8%
13.4%
19.6%
29.6%
Europe
Africa
Asia
Australia
Americas

Figure 4.



a)
1000
2000
3000
4000
5000
6000
-2
0
2
$\log_{10}(\varepsilon_{fine}/\varepsilon_{x\text{-pod}})$
b)
p06
175
200
225
250
275
Longitude
cluster assignment
1
2
3
4
5
6
7
c)
$\log(\frac{\varepsilon_{fine}}{\varepsilon_\chi})$
d)
$[\zeta_z]$
$\log_{10}(S[\zeta_z]/\mathrm{cpm})$
$\log_{10}(k_z)$[rad/m]
e)
$[V_z]$
$S[V_z]$ / cpm
k^{-4}
k^{-3}
$k^{-2.5}$
k^{-2}
$k^{-1.5}$
k^{-1}
$\log_{10}(k_z)$[rad/m]
f)
Magnitude of Underestimated Variance
Magnitude of Overestimated Variance
Spectral Energy
GM-like Shear Spectra
Non-GM (underestimated variance)
Non-GM (overestimated variance)
Total Integrated Variance
0.05
0.1
0.25
Wavenumber $[2\pi/\lambda]$

Figure 5.



Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
2000
4000
6000
40
60
80
100
Longitude
posterior probability [%]
20
40
60
80
100

Figure 6.

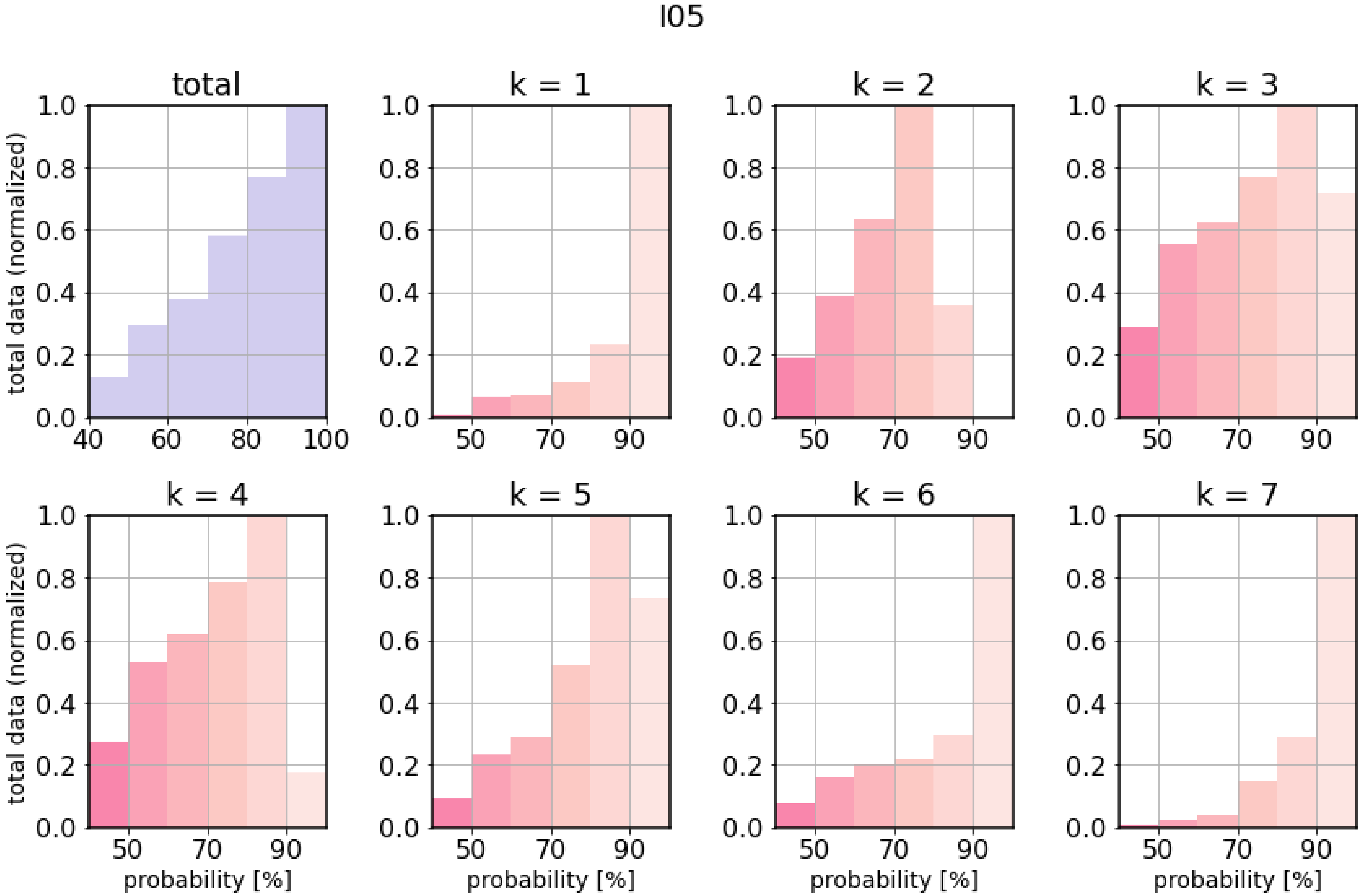
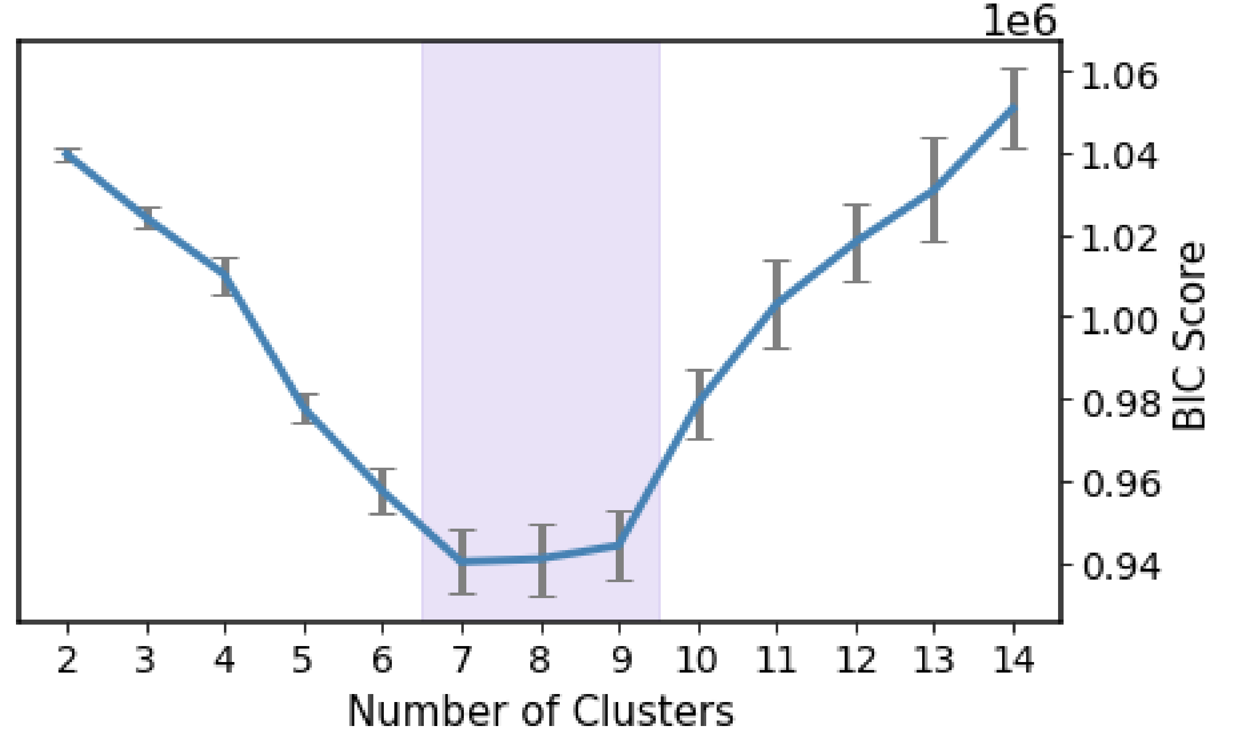

I05
total
k = 1
k = 2
k = 3
k = 4
k = 5
k = 6
k = 7
total data (normalized)
probability [%]
0.0
0.2
0.4
0.6
0.8
1.0
40
60
80
100
50
70
90

Figure 7.


1e6
1.06
1.04
1.02
1.00
0.98
0.96
0.94
BIC Score
2
3
4
5
6
7
8
9
10
11
12
13
14
Number of Clusters