# Converting Seismic Reflection Sections into Oceanic Temperature and Salinity: A New Approach

YAO MENG,[a] KATY L. SHEEN,[a] KATHRYN L. GUNN,[b] CLAIRE MÉNESGUEN,[c] TOBIAS EHMEN,[a] AND IAN G. C. ASHTON[d]

[a] *Centre for Geography and Environmental Science, Department of Earth and Environmental Sciences, Faculty of Environment, Science and Economy, University of Exeter, Penryn, United Kingdom*
[b] *School of Ocean and Earth Sciences, University of Southampton, Southampton, United Kingdom*
[c] *Laboratoire d'Océanographie Physique et Spatiale, CNRS, IRD, Ifremer, University Brest, Brest, France*
[d] *College of Engineering, Mathematics and Physical Sciences, University of Exeter, Penryn, United Kingdom*

(Manuscript received 6 March 2025, in final form 2 August 2025, accepted 30 August 2025)

ABSTRACT: Oceanic submesoscale dynamics are associated with horizontal scales between tens of meters and tens of kilometers and time scales of hours to weeks. Through impacting the transfer of energy and other fundamental ocean properties such as heat, salt, carbon, and nutrients, submesoscale processes are believed to play an important role in the climate system and marine biosphere. However, direct observations of these processes, especially in the ocean interior, remain limited due to their transient nature. Marine seismic reflection surveys offer a solution, resolving thermohaline structures on scales on the order of 10 m vertically and 100 m horizontally and capturing 100-km swathes in hours. While seismic data provide vertical temperature/salinity gradients, legacy datasets are often hindered by sparse hydrographic validation and uncertain inversions. Here, we present an improved inversion method combining root-mean-square sound velocity analysis and iterative Markov Chain Monte Carlo techniques to extract thermohaline fields with quantified uncertainties. The method is validated using Gulf of Cadiz seismic data with coincident hydrographic measurements and applied to a new Mozambique Channel dataset capturing mesoscale and submesoscale activities. Uncertainties for inverted temperature and salinity are 2.5°C (1.65°C) and 0.5 psu (0.08 psu) in the Gulf of Cadiz (Mozambique Channel), with Dix-equation-derived velocity conversion identified as the primary error source. This novel approach expands the use of legacy seismic reflection data as a tool for ocean fine scale to submesoscale analyses and will aid new, global insights into previously difficult-to-observe ocean dynamics.

KEYWORDS: Ocean; Eddies; Acoustic measurements/effects; Inverse methods

## 1. Introduction

Submesoscale currents, defined here as flows with spatial scales of 10 m–100 km and time scales of hours to weeks, are of paramount importance in oceanic processes. Taking the form of fronts, filaments, vortices, and topographic wakes, submesoscale flows can initiate the downscale transfer of energy from larger-scale current flows, intensified vertical motion, enhanced mixing, and boundary layer exchange (e.g., Capet et al. 2008; McWilliams 2016; Mahadevan 2016; Gula et al. 2019; Srinivasan et al. 2019; Gula et al. 2022; Cao et al. 2024). Submesoscale dynamics therefore likely play a critical role in the transfer of heat and momentum, air–sea interactions, the global overturning circulation, and biological productivity (de Ruijter et al. 2002; Chapman et al. 2003; Swart et al. 2010; Ternon et al. 2014; McWilliams 2016; Su et al. 2018; Strobach et al. 2022). However, a lack of submesoscale current observations, particularly within the ocean interior, has meant that this realm of ocean dynamics remains elusive. New approaches are required to adequately sample small, intermittent interior submesoscale currents in the field.

Traditional oceanographic tools, such as repeat hydrographic sections and mooring arrays, typically do not resolve horizontal length scales of less than several kilometers. As such, their horizontal resolutions are insufficient to adequately capture submesoscale phenomena. More recent autonomous instrumentation, while offering significant improvements, is still limited to sampling resolutions of hundreds or thousands of meters (Aulicino et al. 2021). Submesoscale observations below the upper-surface layers remain particularly rare.

Meanwhile, modeling capabilities continue to advance and are increasingly showing the importance of submesoscales (Chassignet and Xu 2021). However, the resolution of most global models is 1–5 km (Haarsma et al. 2016; Gutjahr et al. 2019; Li et al. 2020; Wang et al. 2021; Uchida et al. 2022), with nested simulations perhaps reaching resolutions of 0.5–2.5 km (Tedesco et al. 2024).

Most significantly, model runs are severely lacking observational validation. The need for novel observational methods that are able to capture submesoscale time and space scales throughout the water column is evident.

One solution is the use of marine seismic reflection survey datasets, also known as seismic oceanography (Ruddick et al. 2009). The technique relies on the reflection of active acoustic energy (a large bubble pulse) released from the stern of a

Denotes content that is immediately available upon publication as open access.

*Corresponding author*: Yao Meng, ym371@exeter.ac.uk

DOI: 10.1175/JTECH-D-25-0019.1

vessel, along temperature and/or salinity gradient horizons within the water column. Acoustic reflections are recorded by a long array of hydrophones called a "streamer" (i.e., several kilometers long) towed by the ship. It should be noted that most seismic ocean data are a by-product of surveys primarily conducted to map sub-seabed structures, e.g., by the hydrocarbon industry. Consequently, there is global and multidecadal data coverage (Dickinson and Gunn 2022). Reflections from seismic sections of the ocean have been shown to correlate well with thermohaline submesoscale structures and detect temperature changes of 0.03°C and salinity changes of 0.01 psu (Holbrook et al. 2003; Nandi et al. 2004; Ruddick et al. 2009). In general, temperature variability tends to dominate over salinity gradients in the seismic reflection signature (Sallarès et al. 2009).

Seismic reflection surveys are ideally tuned to capture submesoscale structures across hundreds of kilometers and throughout the water column; their final horizontal and vertical resolutions are on the order of tens of meters, and data are collected over a matter of hours. Furthermore, by analyzing sequentially collected seismic sections, these data can be used to give insights into dynamics of submesoscale processes over several days (e.g., Tsuji et al. 2005; Jones et al. 2010; Dickinson et al. 2020; Zou et al. 2021; Yang et al. 2022). Features such as fronts (Holbrook et al. 2003; Gunn et al. 2020), submesoscale coherent vortices (Sheen et al. 2012; Gunn et al. 2018; Xiao et al. 2021), and wave features (Eakin et al. 2011; Tang et al. 2014; Fortin et al. 2017; Geng et al. 2019; Fan et al. 2021) have been captured. Techniques have also been developed to extract estimates of internal wave energy and turbulent dissipation levels (Sheen et al. 2009; Holbrook et al. 2013; Tang et al. 2016; Fortin et al. 2017) and instability criterion (Ehmen et al. 2022).

One barrier to fully exploit the potential of most seismic ocean data however, is the difficulty of accurately extracting detailed temperature and salinity information: seismic reflections represent temperature and salinity vertical gradients convolved with the seismic source. Several inversion techniques have been developed to estimate temperature and salinity fields: see summary in Table D1 of Dickinson and Gunn (2022) and Xiao et al. (2021). An accurate inversion requires an accurate initial estimation of the background temperature and salinity field. Reliable sources of this information are simultaneously collected hydrographic data over small spatial scales (Hobbs 2007b). Ideally, hydrographic data need to be closely spaced (e.g., every few kilometers; Xiao et al. 2021) or the features of interest be horizontally coherent, i.e., internal solitary waves (i.e., Tang et al. 2016). Such coincident hydrographic data or oceanographic conditions are not typical for seismic surveys, limiting the detail and accuracy of inverted thermohaline fields and the hence interpretation of submesoscale flows.

Here, we develop and test a novel approach for extracting detailed thermohaline fields from marine seismic reflection survey that lacks high-resolution coincident hydrographic data, with quantified uncertainty. We combine two inversion techniques to develop a new approach: a velocity analysis (VA) approach based on and adapted from Gunn et al. (2018) and spatially iterative Markov chain Monte Carlo (MCMC) (Tang et al. 2016; Xiao et al. 2021) methods, hereafter called the "combined inversion."

The VA method makes use of the sound velocity information contained in seismic reflection data, itself a function of temperature and salinity, to estimate low wavelength or "background" thermohaline fields directly from the acoustic signal. The MCMC method provides iterative refinement of the VA estimated initial temperature and salinity fields, accounting for spatial variability at the scale of the seismic data resolution (i.e., tens of meters). A careful uncertainty analysis is conducted at both the VA and the MCMC stages. The result is a combined inversion that can be applied to almost any seismic reflection dataset (i.e., collected with or without simultaneous hydrographic data), thereby unlocking a global legacy dataset.

## 2. Data

The new combined inversion approach is applied to seismic reflection datasets from two different locations: the Gulf of Cadiz (GC) and the Mozambique Channel (MC) (Fig. 1). The Gulf of Cadiz dataset was obtained in 2007 as part of the Geophysical Oceanography (GO) project (Hobbs 2007a,b). The GO field campaign focused on capturing a subsurface eddy of warm, salty Mediterranean Outflow Water or "meddy" (Hobbs et al. 2009; Papenberg et al. 2010; Xiao et al. 2021). Being one of the few seismic reflection surveys collected primarily to image the water column, several in situ temperature and salinity measurements were made alongside seismic acquisition. This dataset thus presents a unique testbed to assess our new combined inversion methodology and uncertainty estimates. Second, the combined inversion methodology is applied to a 2016 legacy industrial seismic data from the Mozambique Channel. Collected primarily for geological exploration, the Mozambique dataset contains no coeval hydrographic data and thus provides a demonstration of the utility of the combined inversion method. Located in a region of high kinetic energy and water mass confluences along the Mozambique continental shelf break, the inverted thermohaline fields reveal a multitude of intricate submesoscale interior structures.

### *a. Gulf of Cadiz*

#### 1) Seismic reflection survey

The seismic reflection survey was acquired on 7 May 2007 onboard the RRS Discovery as part of the GO project (cruise number D318b—Hobbs 2007b). Here, we present data from a two-dimensional seismic section, GOLR12. Towed 9 m below the sea surface, the acoustic source fired every 20 s and consisted of six Bolt 1500LL air guns with a total volume of 2320 in.$^3$. The final bandwidth of the dataset is 5–70 Hz. Reflected acoustic waves were recorded along a single 2.4-km streamer that had 192 groups of hydrophones spaced every 12.5 m. The horizontal sampling interval, defined by the common midpoint (CMP) spacing, is 6.25 m, while the true horizontal resolution is fundamentally limited by the Fresnel zone and is likely to be coarser than this value under stacked processing. Although the theoretical minimum is half a seismic wavelength (10 m for

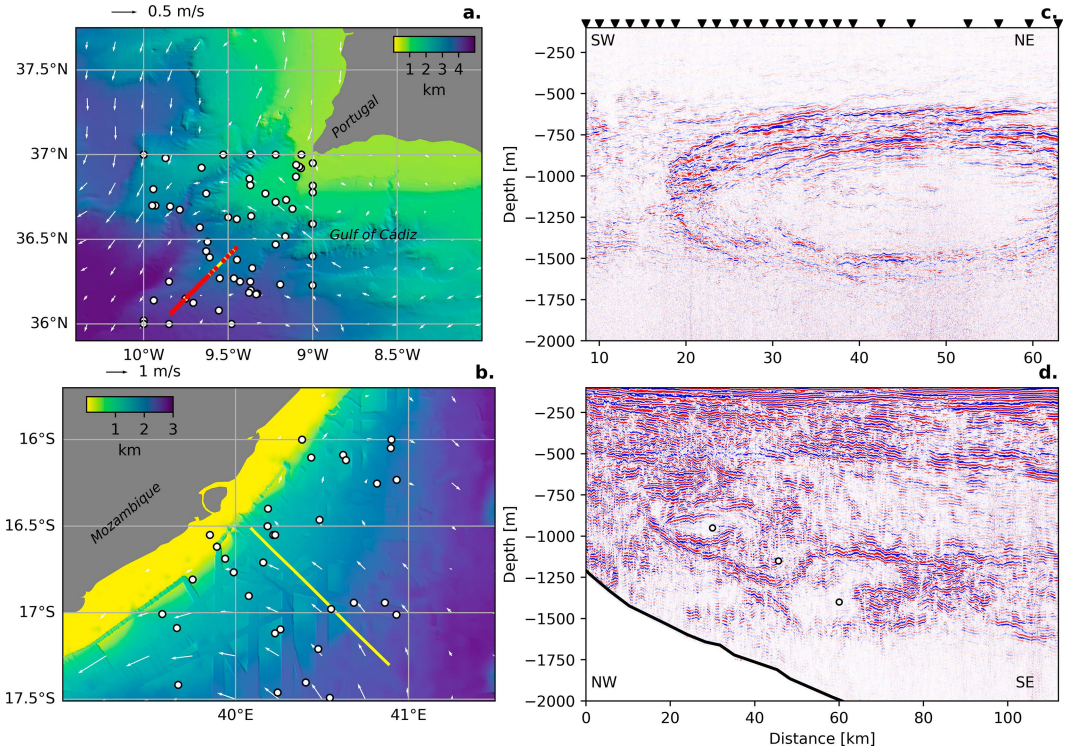


FIG. 1. Map of GC and MC datasets and seismic sections. (a) Bathymetry map. Yellow line denotes location of the GC seismic reflection profile collected on 7 May 2007, red circles denote XBT casts collected simultaneously with the seismic data, white circles denote historical hydrographic data from WOD acquired in May between 1930 and 2022, and white arrows denote surface geostrophic velocities on 7 May 2007 derived from Global Ocean Physics Reanalysis dataset. (c) Processed GC seismic image. Red and blue bands denote acoustic reflections due to temperature and salinity vertical gradients; black triangles denote XBT locations. (b) As in (a), but for MC dataset collected on 12 Jan 2016, WOD data collected in January–March between 1965 and 2022 and surface geostrophic velocity fields from Global Ocean Physics Reanalysis dataset. (d) As in (c), but for MC seismic section. Black line denotes seafloor; white circles denote three possible submesoscale lens-like structures. Data courtesy of SLB and INP.

a 70-Hz bandwidth), in practice, the horizontal resolution is likely closer to 43–107 m (2–5 times the seismic wavelength) since prestack migration was not applied. The vertical resolution ranges between 5.35 and 70 m. The final seismic section used in this study is 50 km long and traverses the continental slope covering water depths of 800–1600 m (Fig. 1c).

#### 2) HYDROGRAPHIC DATA

Profiles of in situ temperature, coincident in time and space with the seismic data, were achieved using expendable bathythermographs (XBTs) spaced approximately every 2.3 km (Hobbs 2007b; Fig. 1a). To obtain salinity information, conductivity–temperature–depth (CTD) data collected from the observation area were used to train a neural network that models the local temperature–salinity–depth relationship following Ballabrera-Poy et al. (2009) and Xiao et al. (2021) (three CTD casts were deployed along the same seismic transect just a few hours before or after the seismic acquisition). In this way, salinities corresponding to each XBT measurement were estimated with uncertainties of 0.01 psu (Xiao et al. 2021). In total, 24 profiles of temperature and salinity were available along the seismic section.

A set of historical hydrographic data was also acquired from the World Ocean Dataset (WOD) within the vicinity of the Gulf of Cadiz to facilitate the testing of the combined inversion method (white circles in Fig. 1a). These data collected during the month of May spanning from 1930 to 2022 and in the region 36°–37°N and 9°–10°W comprise 14 CTD profiles, 60 ocean station data (bottle, low-resolution CTD/XCTD), and profiling floats. These historical hydrographic data were utilized to establish the temperature–salinity–depth relationship necessary for the inversion process.

### *b. Mozambique Channel*

#### 1) SEISMIC REFLECTION SURVEY

The Mozambique Angoche survey was acquired between January and March 2016 by SLB. This three-dimensional seismic survey spans the western continental slope of the Mozambique Channel in the Indian Ocean. Water depths vary from 1000 to 2500 m. The acoustic source comprised an array of 18 air guns, with a combined volume of 5085 $in.^3$. The resulting bandwidth is 1–90 Hz. The air guns were fired every 10 s at a tow depth of 6 m. Reflected acoustic waves were recorded

along twelve streamers each of 8-km length. Each streamer had 2571 groups of hydrophones spaced every 3.125 m. On each traverse, the seismic experiment acquired a three-dimensional swath of data approximately 800 m wide and up to 120 km long. Here, we extract and analyze a single vertical cross section: L02. This seismic section was collected between 0325 and 1720 UTC 12 January 2016, extends 115 km (Fig. 1d).

#### 2) HYDROGRAPHIC DATA

As is typical with most seismic reflection surveys, coincident hydrographic data were not collected during the Mozambique Channel seismic reflection survey. Instead, we rely on historical hydrographic data to form a local temperature–salinity relationship as required for the inversion. All available hydrographic data within 160 km of the seismic survey (16°–17.5°S, 39.5°–41°E) and collected during the same months as the seismic survey (i.e., January–March) were downloaded from the World Ocean Database. In total, 16 ocean station data (bottle, low-resolution CTD/XCTD, plankton data) and 17 profiling floats fit these criteria, spanning depths of 0.5–2 km (see white circles in Fig. 1b).

## 3. Methods

### *a. Seismic processing to image oceanic structure*

Standard seismic processing techniques have been applied to both the Gulf of Cadiz and the Mozambique Channel data. Importantly for subsequent inversion work, care is taken to generate a true-amplitude seismic section. The processing flow includes 1) geometry setting, 2) frequency filtering using a 10–70-Hz (15–90 Hz) Butterworth filter for the Gulf of Cadiz (Mozambique Channel) data, 3) seabed mute, 4) direct-wave removal using an adaptive linear filter, 5) noise attenuation using a frequency–wavenumber filter (Wei et al. 2025), 6) sorting into CMP gathers, 7) velocity analysis: picked every 50 (200) CMPs for the Gulf of Cadiz (Mozambique Channel), 8) spherical divergence amplitude correction, and 9) deconvolution following the strategy of Sacchi (1997): the source wavelet is modeled based on the source array geometry and airgun volumes with amplitude calibration conducted using a normal calibration factor, also following Sacchi (1997). Note that the dataset was preprocessed with a mute applied below the seafloor, so seabed multiples were not available for absolute amplitude calibration using the seabed primary and multiple reflections. These techniques are adapted from those used to build seismic images of solid Earth (Yilmaz 2001). For more details about seismic reflection processing, see also Jones and Levy (1987), Warner (1990), Sacchi (1997), Warner (1990), and Biescas et al. (2014).

### *b. Combined inversion for fields of temperature and salinity*

The amplitude of the seismic water reflection is controlled by changes in sound speed and density within the water column, which are in turn a function of temperature and salinity (Ruddick et al. 2009; Sallarès et al. 2009). Therefore, the seismic reflectivity field contains information about water properties that can be recovered through inversion.

Two such established inversion techniques for seismic ocean data include the VA and MCMC methods. The VA method exploits the integrated impact of water properties on the velocity of the seismic sound pulse as it passes through the water column. This method provides a background, smoothly varying temperature and salinity field (Gunn et al. 2018, 2020). While the advantage of the VA approach is that it does not require coincident hydrographic data, it is limited in its ability to capture fine-scale variability (i.e., <1 km horizontally and <100 m vertically). Uncertainties in VA inversion results are also high, particularly at depth. The MCMC approach on the other hand produces inverted thermohaline fields (with uncertainties) by matching individual reflector amplitudes with the most likely temperature and salinity variations (Tang et al. 2016). MCMC outputs therefore capture the fine-scale detail, but an accurate prior low-wavelength temperature/salinity field is required: MCMC inversions have only been successful in datasets with sufficiently highly sampled coincident hydrographic data (Xiao et al. 2021) or where the thermohaline structures of interest are sufficiently homogeneous laterally (Tang et al. 2016).

Here, we combine the two techniques, such that the VA provides the starting model for the MCMC analysis (bypassing the need for a dense distribution of independent hydrographic measurements), while the MCMC iteratively improves the recovered spatial resolution of the VA output. These steps are detailed in Fig. 2. The result is an inversion approach that exploits the benefits and overcomes the limitations of both the VA and the MCMC methods and can be applied to any seismic dataset.

#### 1) VA TO EXTRACT INITIAL FIELD OF OCEAN PROPERTIES

The integrated sound speed, or root-mean-square sound speed (RMS velocity), $v_{\mathrm{rms}}$, at various depth intervals along the path between the air gun source and hydrophone receiver is directly estimated as part of any basic seismic processing sequence (i.e., the velocity analysis for normal move out (NMO) correction prior to stacking CMP gathers Yilmaz 2001). The velocity picking process is illustrated in Fig. A2, using semblance plots, which show coherence of reflected arrivals as a function of RMS velocities. The $v_{\mathrm{rms}}$ values are estimated by exploiting the variation in ray path travel times for different hydrophone receivers, for a given reflection depth. Semblance plots are utilized to find optimal $v_{\mathrm{rms}}$ values for depths associated with strong reflectors, at regular horizontal intervals across the seismic survey: known as velocity picking, this process is typically done by eye and is therefore somewhat subjective. Here, $v_{\mathrm{rms}}$ vertical profiles were estimated every 312.5 m/50 CMPs (316 m/200 CMPs) along the Gulf of Cadiz (Mozambique Channel) seismic section. The $v_{\mathrm{rms}}$ values (i.e., the integrated sound velocity between each picked point and the sea surface) are next converted to interval

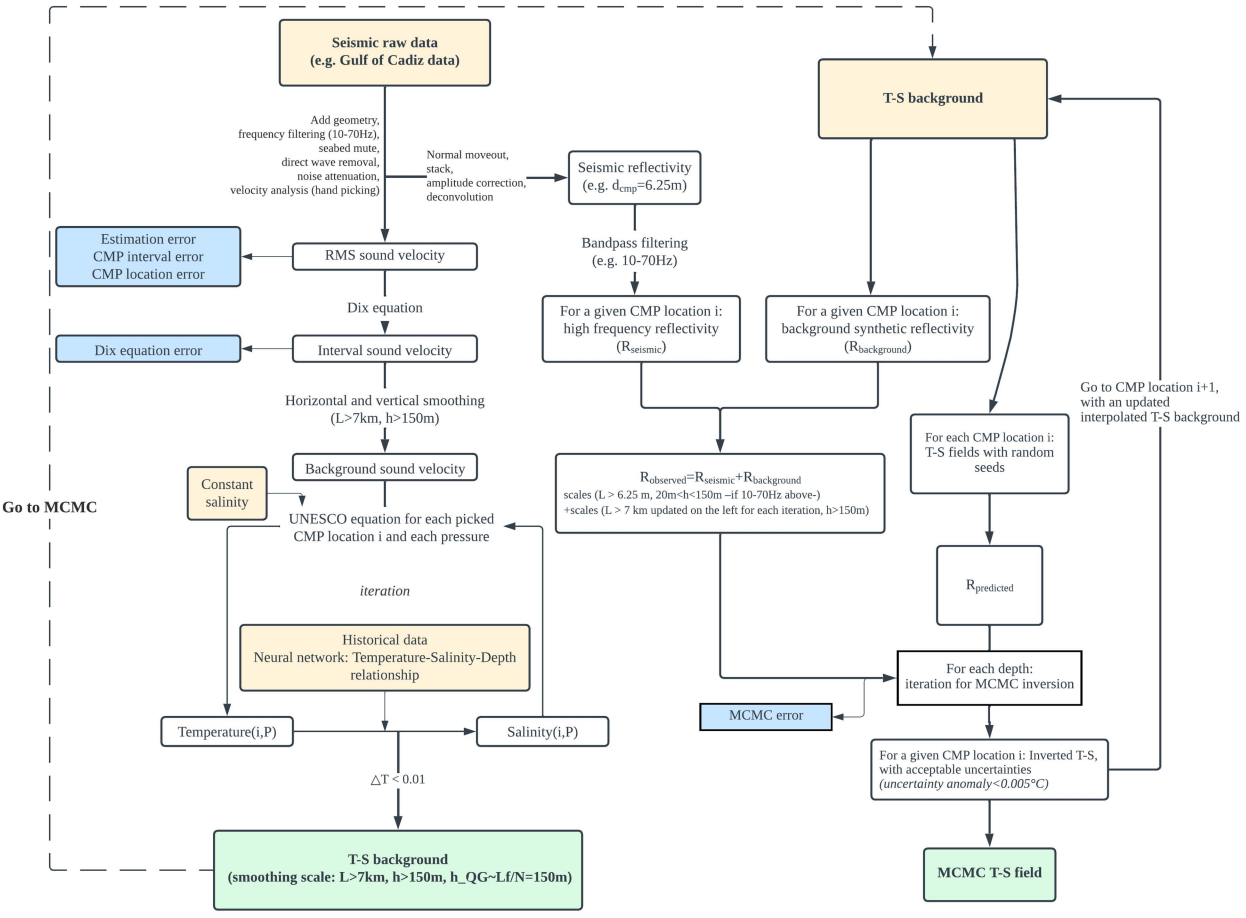


FIG. 2. Flowchart of the combined inversion steps. The inversion is divided into two stages: (left) extraction of the smoothed temperature–salinity fields from sound velocity inversion obtained from seismic processing and (right) utility of seismic resolution to add the fine-scale structure to inversion output fields. Yellow rectangles denote input data; green rectangles denote output data; blue rectangles denote error estimation.

velocities, $v_{\text{int}}$ (i.e., the average sound speed between any two picked reflectors), using the Dix equation (Dix 1955):

$$v_{\text{int},i} = \sqrt{\frac{v_{\text{rms},i}^2 t_i - v_{\text{rms},i-1}^2 t_{i-1}}{t_i - t_{i-1}}}, \tag{1}$$

where $v_{\text{int},i}$, $v_{\text{rms},i}$, and $t_i$ are the interval velocity, RMS velocity, and time at the $i$ depth layer, respectively. The uncertainty introduced by this conversion is addressed in section 3c(1). This dominant source of error is often overlooked in previous studies that applied the VA inversion method (e.g., Gunn et al. 2018; Chhun and Tsuji 2020). The $v_{\text{rms}}$ fields were not interpolated when calculating interval velocities, ensuring that errors related to the Dix equation could be accurately quantified.

After deriving interval velocities from the RMS velocity using the Dix equation, following Gunn et al. (2018), interval velocity fields were interpolated onto the same two-way travel time (TWTT) grid to match the data sampling rate and allow consistent smoothing. The velocity fields were then smoothed horizontally (by a length scale $L$) and vertically (over a TWTT interval equivalent to a physical scale $H$ using an assumed average sound speed 1500 m s$^{-1}$) to enhance the continuity of physical structures. Here, $L$ and $H$ were chosen such that $L/H \sim N/f$, where $N$ is the buoyancy frequency and $f$ is the Coriolis parameter: $L = 7$ km and $H = 150$ m in the Gulf of Cadiz and $L = 8$ km and $h = 100$ m in the Mozambique Channel. The chosen scales are dynamically consistent with the dominant physical processes in each region.

We then follow the iterative method as documented in Gunn et al. (2018), with the one key modification: a neural network approach was employed using MATLAB's Neural Fitting Tool to predict the temperature–salinity–depth relationship extracted from historical hydrographic data (Ballabrera-Poy et al. 2009). We used 70% of the data for training, 15% for validation, and 15% for testing. The network architecture consisted of two input nodes (predictors: temperature and depth), a single hidden layer with 10 neurons and an output layer (predicted variable: salinity). The Levenberg–Marquardt algorithm was selected for network training. The validation root-mean-square error is less than 0.04. Finally, a local temperature–salinity–depth relationship is used to convert interval velocities into temperature and salinity fields.

TABLE 1. Summary of inversion uncertainties for the GC data. Sound speed, inverted temperature, and salinity mean (maximum) uncertainties are computed to assess the error introduced by estimating the velocity picking, varying the CMP interval, varying the CMP location, the Dix equation conversion, and the MCMC posterior error. The overall error as well as the difference between inversion outputs and hydrographic observations (column titled "Bias") is also shown.

| | Estimation | CMP interval | CMP location | Dix equation | MCMC posterior | Overall | Bias |
|---|---|---|---|---|---|---|---|
| Sound speed (m $s^{-1}$) | 2.2 (3.6) | 1.5 (2.2) | 1.8 (2.4) | 9.5 (26.6) | — | 15 (34.8) | — |
| Temperature (°C) | 0.3 (0.5) | 0.2 (0.5) | 0.2 (0.5) | 1.8 (2.5) | 0.02 (0.3) | 2.5 (4.2) | 0.5 (2.3) |
| Salinity (psu) | 0.05 (0.11) | 0.04 (0.1) | 0.05 (0.13) | 0.36 (0.6) | 0.005 (0.07) | 0.5 (0.92) | 0.14 (0.64) |

#### 2) MARKOV CHAIN MONTE CARLO APPROACH TO IMPROVE FIELD OF OCEAN PROPERTIES

The MCMC inversion approach is based on the stacked seismic reflection amplitude analysis, first introduced by Tang et al. (2016) and then updated by Xiao et al. (2021). The method requires a prior background temperature and salinity field: here, we use the output from the VA analysis inversion (Fig. 2). Small, random perturbations are added to the prior temperature and salinity fields from which an initial reflectivity field, $R_{\text{predicted}}$, is calculated. A likelihood ratio evaluates the match between $R_{\text{predicted}}$ and the observed seismic reflectivity, $R_{\text{observed}}$. If a threshold likelihood ratio is reached, the input temperature–salinity pair is updated. The process is repeated iteratively for 1000 times, sufficient to reach a converged stable state (Fig. A4). Final temperature and salinity values are computed as the mean of the last two-thirds of perturbations that reach a stable state. Following Xiao et al. (2021), we divide the seismic section into horizontal units for the MCMC analysis. Within each unit, the MCMC inversion starts from the boundaries and sequentially moves toward the center. In Xiao et al. (2021), analysis unit boundaries were determined by XBT positions, where temperature–salinity values were best constrained. Here, we use the velocity picked CMP locations to set the analysis box boundaries, i.e., the Gulf of Cadiz seismic section was split into a total of 186 inversion units, each 312.5 m wide, and the Mozambique Channel dataset was split into 52 inversion units, each 2.2 km wide.

While the MCMC inversion was conducted on each CMP (sampling interval 6.25 m) for the Gulf of Cadiz data, due to computational time for the larger Mozambique Channel, the inversion was conducted on every 10th CMP (sampling interval ~ 15.6 m). For both datasets, the inversion depth spacing, or TWTT, was set to 1 m $s^{-1}$ (equivalent to 0.75 m).

Finally, we note that inverted results were excluded if the associated posterior error [see section 3c(5)] exceeded a threshold of 0.005°C. Any values exceeding this threshold were replaced by results linearly interpolated from the background field. This control mechanism helps prevent any significant bias of one MCMC-derived profile from affecting the inversion of the next profile.

### c. *Quantifying inversion uncertainty*

To assess how well the combined inversion resolves fields of temperature and salinity, we estimate the uncertainty associated with both the VA and MCMC methodologies described above. Errors are categorized into four groups: (i) the Dix equation error associated with transforming RMS sound velocity $v_{\text{rms}}$ to interval velocity $v_{\text{int}}$, (ii) the error associated with the subjective nature of velocity picking to determine the initial $v_{\text{rms}}$ model, (iii) procedural errors from choice of CMP intervals and CMP locations, and (iv) the posterior error associated with the MCMC method. To validate the error analysis, the bias between available hydrographic data and combined inversion results is also evaluated. The different errors are described in detail below and summarized in Table 1.

#### 1) DIX EQUATION ERROR

The conversion of RMS velocities to interval velocities is inherently complex and prone to uncertainties. Not only does it involve a derivative, with errors accumulating with depth, but the $v_{\text{rms}}$ function itself can be unreliable (Fig. A2 in the appendix).

For ease of explanation, we here provide a simplified version, the uncertainty estimate ($\Delta v_{\text{int}}$), [their Eq. (28); Hajnal and Sereda 1981] which highlights the dominant terms (Hajnal and Sereda 1981):

$$\Delta v_{\text{int},i} \approx \frac{2t_{i-1}}{t_i - t_{i-1}} \Delta v_{\text{rms},i-1}, \tag{2}$$

where the uncertainty is shown to be inversely proportional to the interval velocity layer thickness/travel time (i.e., the depth interval between two picked $v_{\text{rms}}$ points) and directly proportional to both the TWTT (i.e., depth) of the velocity pick and the RMS velocity error [computed as in section 3c(1)]. For the full version of the equation [their Eq. (27); Hajnal and Sereda 1981], as used in our analysis, the TWTT error is set to 0.001 s, which is the same order of magnitude as the seismic sampling interval. We note that the final interval velocity uncertainty is largely insensitive to the exact choice of TWTT error. For example, increasing the TWTT error to 3 ms [i.e., (1/4)$f$, where $f$ is the seismic bandwidth] was found to increase the resulting $\Delta v_{\text{int}}$ by only 0.1 m $s^{-1}$—two orders of magnitude smaller than the typical $\Delta v_{\text{int}}$ values reported in Table 1. Vertical distances between velocity picks/layer thicknesses of less than 0.1-s TWTT produce interval velocity errors of several tens of meters per second, even for weak RMS errors (i.e., 5 m $s^{-1}$). In addition, the errors grow cumulatively with depth, reaching 100 m $s^{-1}$ or more. Thus, while it is necessary to select a sufficient number of velocity picks to preserve the variability of the background velocity field, too densely picked velocities (in depth) generate unfeasible interval velocity errors. We found keeping picking travel time intervals to greater than 0.2-s TWTT (approximately 150 m) optimal.

The error associated with the interval velocity uncertainty on the final inverted fields was assessed by recomputing temperature and salinities using interval velocities with the Dix equation error added and subtracted. Final values quoted represent half the difference between the two fields.

#### 2) ROOT-MEAN-SQUARE VELOCITY ESTIMATION ERROR

The $v_{\mathrm{rms}}$ fields are estimated by selecting points of high coherence in semblance gathers—a process known as velocity picking (Fig. A2 in the appendix). To estimate uncertainty, we performed three rounds of RMS velocity picking on the same section. Each round was conducted by the same interpreter but separated in time and performed without reference to the previous picks. Velocity picks were found to be consistent within a standard deviation (std) of $\pm 9.2$ m s$^{-1}$. No systematic upward or downward change trends were found between different picks. Additionally, the spatial density of picks remained similar across attempts, guided by the same semblance amplitude. For the Gulf of Cadiz dataset, a section between 12.5 and 25 km from the start of the transect with strong submesoscale thermohaline structures was chosen. For the Mozambique Channel data, a section between 34 and 66 km was analyzed. Selected $v_{\mathrm{rms}}$ data were converted into three fields of interval velocities and their corresponding fields of temperature and salinity. The average and maximum standard deviation between these converted fields quantify the estimation error (Figs. A3a–c and A5a–c).

#### 3) VELOCITY ANALYSIS PROCEDURAL ERROR: CHOICE OF CMP INTERVAL

The procedural error associated with the choice of CMP interval for the generation of semblance plots and $v_{\mathrm{rms}}$ estimation was quantified. Using smaller CMP intervals can improve the positional accuracy resolved detail of thermohaline features, but this choice is labor intensive and often impractical. To test the effect of CMP interval, semblance gathers were generated at intervals of 50 (0.3 km), 100 (0.6 km), 150 (0.9 km), and 200 (1.2 km), across the same region of the seismic data (Figs. A3d–f and A5d–f). The average (maximum) standard deviation between inverted fields with different CMP intervals provides the associated error.

#### 4) VELOCITY ANALYSIS PROCEDURAL ERROR: CHOICE OF CMP LOCATION

The variability in derived thermohaline fields due to the exact choice of CMP analysis locations was quantified by comparing output fields using fixed CMP intervals of 100 and 150 CMP but in different locations (Figs. A3g–i and A5g–i). To estimate the uncertainty arising from CMP interval and location, we generated multiple RMS velocity fields using different CMP configurations and then inverted these to obtain corresponding temperature and salinity fields. At each spatial grid point, the standard deviation across these fields was computed to construct a 2D uncertainty field. This pointwise analysis avoids spatial averaging and preserves local variability in the uncertainty estimation. The final error is taken to be the average (maximum) standard deviation between outputs.

#### 5) MCMC PROCEDURAL ERROR: POSTERIOR

The accuracy of the MCMC inversion is highly dependent on the input temperature–salinity background field (Tang et al. 2016; Xiao et al. 2021). Here, the initial velocity field is derived using the VA method, for which the error has been quantified as described above. Therefore, we only consider the additional error introduced by the MCMC inversion process. The MCMC posterior uncertainty is calculated as the standard deviation of the final two-thirds of the iterations that reach a stable state (Fig. A4). Errors are computed at each inversion point, and thus, the spatial distribution of the recovered temperature and salinity uncertainty can be analyzed. Here, the section-mean and maximum uncertainties are used to represent the overall uncertainty of the MCMC inversion.

#### 6) FINAL INVERSION BIAS

Finally, we estimate a bias that quantifies the deviation of the inverted temperature and salinities from the true field. The bias is also evaluated in relation to our estimated uncertainties. For the Gulf of Cadiz data, hydrographic data (i.e., XBTs) collected coevally with the seismic data are used as the true field (see Fig. 1a). A neural network was employed to predict salinity based on the temperature–depth relationship, using a model trained on CTD data collected during the same research cruise (Ballabrera-Poy et al. 2009). For the Mozambique Channel, inversion outputs are compared with mean temperature and salinity profiles from available local historical data.

## 4. Results

First, temperature and salinity fields produced using the new combined inversion approach from the Gulf of Cadiz dataset are presented. Outputs are compared qualitatively and quantitatively with available coeval hydrographic measurements. Second, the inversion results from the seismic dataset collected in Mozambique Channel are discussed. This dataset, which is rich in submesoscale structures, provides a case study to examine how legacy seismic data without coincident hydrographic data and their inverted fields can be used to investigate the ocean.

### *a. Gulf of Cadiz*

The Gulf of Cadiz is a pathway for the exchange of waters between the Mediterranean Sea and the North Atlantic Ocean: surface Atlantic waters move toward the Straits of Gibraltar, while deeper Mediterranean Outflow Water crosses westward into the Atlantic Ocean. The westward Mediterranean Outflow Water flow is interspersed with subsurface coherent vortices or meddies that are typically 20–100 km wide with a vertical extent of 700–1200 m (McDowell and Rossby 1978; Armi and Zenk 1984; Richardson et al. 2000; Siedler et al. 2005; Pinheiro et al. 2010; Bashmachnikov et al. 2015). An estimated 15–20 meddies are produced annually forming a significant contribution of the transport of salty Mediterranean

Outflow Water into the Atlantic Ocean (Bower et al. 1997; Richardson et al. 2000).

Being significantly warmer (2°–4°C) and saltier (>0.4 psu) than the surrounding water, meddies are well imaged by seismic reflection surveys, e.g., Biescas et al. (2008) and Papenberg et al. (2010), and a meddy is clearly captured in the Gulf of Cadiz seismic section presented here (Fig. 1b). The meddy periphery is characterized by numerous interleaving layers which appear as bright reflections, for example, between depths of 600 and 1000 m (upper boundary) and around 1500 m (lower boundary). Comparatively, the meddy core is acoustically opaque, suggesting homogeneity in temperature and salinity. Several previous studies have utilized this section or similarly located seismic data to investigate meddy fine-scale structures and their associated dynamics and tracer transport e.g., (Pinheiro et al. 2010; Quentel et al. 2010; Ménesguen et al. 2012; Xiao et al. 2021; Song et al. 2011).

Although the Gulf of Cadiz data span the sea surface to the seabed, reflections within the upper 550 m (0–0.7-s TWTT) are relatively weak. This weak reflectivity may be due to a combination of factors, including the homogeneous water properties in the upper layer and potential presence of noise and attenuation from the seismic processing. In this study, we focus on the seismic reflections between 800- and 1600-m depth (1–2-s TWTT), containing the meddy.

#### 1) Velocity analysis inversion

The temperature and salinity field outputs from the VA inversion, which require no coincident hydrographic data, capture the overall shape and properties of the meddy well (Fig. 3). The anomalously warm and salty eddy core, with values that are close to the temperature and salinity measured by the hydrographic data, as well as the colder and fresher surrounding water, are reproduced. The average absolute differences between the VA-inverted temperature and salinity and hydrographic values are 0.4°C and 0.06 psu, respectively (within the depth range of 800–1500 m and at transect distances of 20–54 km). While the inverted temperature and salinity capture the rapid changes at the meddy boundaries, fine-scale structures with horizontal length scales less that tens of meters and vertical length scales less than a few hundred meters are missed.

One of the developments over previous studies which utilized the velocity analysis inversion (Gunn et al. 2018; Chhun and Tsuji 2020) is the estimation of the error introduced by computing Dix-derived interval velocities [section 3c(1); Figs. 3d,h]. The highest errors, reaching 2.5°C and 0.6 psu, are mainly located at the meddy boundaries: despite the higher signal-to-noise ratio in these regions, the rapid changes in sound velocity require small depth intervals between $v_{\mathrm{rms}}$ picks to capture the variability but which in turn amplify the Dix equation error [section 3c(1)]. Table 1 summarizes the different contributions to the VA-inverted temperature and salinity uncertainties: Errors introduced by the Dix conversion process dominate.

To evaluate the VA inversion results, coincident hydrographic (i.e., XBT) data were compared with inverted *T*–*S* fields. XBT deployment locations are shown in Figs. 3b and 3f. XBT data were smoothed vertically and horizontally on the same scale as the inverted data for comparison (Figs. 3e,f). Regions around the meddy boundary display some of the greatest differences, likely reflecting the large errors here in computing interval velocities. For example, the mean and maximum temperature bias at the lower meddy boundary reach 2.1°C and 2.5°C, while salinity bias are 0.4 and 0.5 psu (Figs. 3d,h). Inside the meddy, the inverted temperature and salinity are slightly warmer and saltier than the XBT data with a maximum difference of 1°C and 0.1 psu, respectively. Overall, differences between inverted fields and hydrographic data are less than the computed uncertainties.

#### 2) Combined inversion

The final combined VA–MCMC inversion results are shown in Fig. 4. Compared to the VA inversion alone, the MCMC adds detail to the recovered *T*–*S* at the resolution scale of the seismic data (order of 10 m in the horizontal and vertical). This improvement is particularly apparent around the meddy boundary, but increased thermohaline structure is also apparent within the meddy core. The fine-scale structure captured at the meddy boundary by the combined inversion illustrates the presence of double diffusion and water mass exchange processes, mixing the Mediterranean Outflow Water with the surrounding cold and fresh Atlantic waters (Xiao and Meng 2022). Overall, the final temperature and salinity inverted fields have a mean errors of 0.5°C and 0.14 psu, while maximum differences are 2.3°C and 0.64 psu, respectively, when compared to coincident hydrographic data, and are less than the overall expected error (Table 1). However, the MCMC is unable to recover regions with large temperature and salinity variations on intermediate scales (on the order of 1 km horizontally and 100 m vertically). This discrepancy is particularly apparent in the upper western edge of the meddy (depths 800–1200 m; distance 16–38 km) where temperatures and salinity can differ by 1.6°C and 0.5 psu, respectively, compared to hydrographic data. Comparatively, in the meddy core, temperature and salinity values differ from hydrographic data by only 0.04°C and 0.008 psu, respectively. Despite higher signal-to-noise ratios and strong reflectivity in the upper-left meddy periphery, essentially, a lack of input information at intermediate scales in the VA-inverted fields, or seismic reflection data, cannot be recovered (see Xiao et al. 2021). In addition, Tang et al. (2016) note that the uncertainty of the MCMC posterior distribution increases in regions of higher temperature and salinity variability. Having said this, the error introduced by the MCMC is an order of magnitude smaller than error sources associated with the VA inversion (Table 1).

To further investigate the combined inversion approach, inverted temperature and salinities are compared with data from an XBT profile which cuts through the western side of the meddy core (at CMP4092, distance of 25.5 km along the transect) (Figs. 5a and 5b). This XBT was chosen as it demonstrates both regions where the inversion performs poorly (mean bias of 1.04°C and 0.28 psu between 850- and 1050-m depth) and well (mean bias of 0.29°C and 0.08 psu between 1400- and 1600-m depth). The VA inversion results generally follow the trend of hydrographic profiles but display large

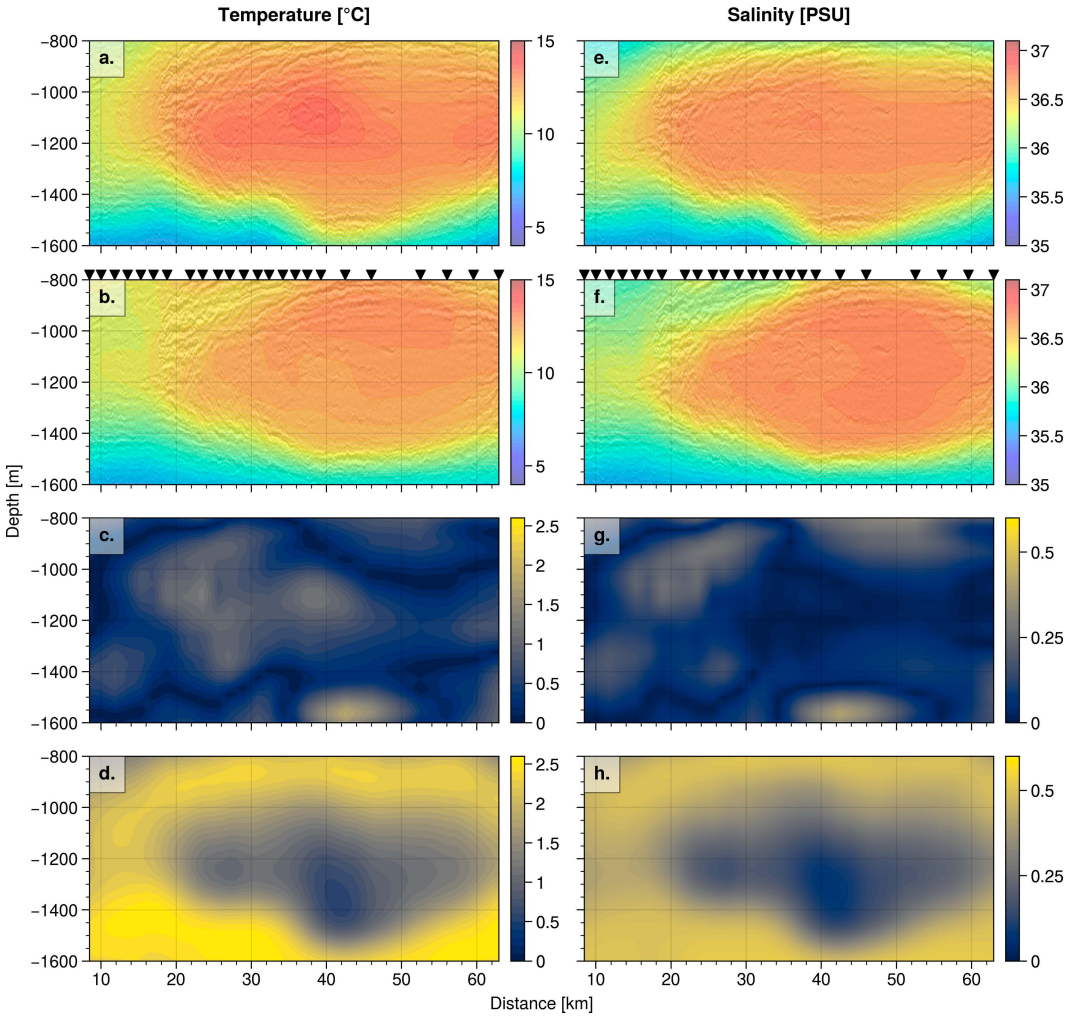


FIG. 3. Comparison of VA-inverted temperature and salinity fields with coincident hydrographic data for the GC seismic section. (a) Background colors denote the in situ temperature section from VA inversion. Overlayed wrinkles denote seismic reflection data. (b) Background colors denote in situ temperature from XBT hydrographic data collected at the same time as the seismic data with horizontal and vertical smoothing of 7 km and 150 m, respectively. Overlayed wrinkles denote seismic reflection data. Black triangles denote XBT locations. (c) The absolute difference between the (a) VA-inverted temperatures and (b) XBT-smoothed field. (d) VA-inversion temperature uncertainties due to Dix conversion. (e)–(g) As in (a)–(d), but for practical salinity, with (f) field estimated from XBT data using a neural network approach.

systematic differences—up to 1.8°C—in some regions. Furthermore, the VA analysis clearly fails to capture variability on vertical scales less than about 100 m. Comparatively, the combined inversion outputs capture smaller-scale thermohaline variations well, demonstrating the role of the MCMC inversion which exploits the fine-scale information in the seismic reflectivity, in refining the VA inversion. The MCMC, however, is unable to correct for errors on larger length scales introduced by the VA inversion.

Synthetic seismograms show in general that the seismic reflectivity matches both the hydrographic data and MCMC outputs very well (Figs. 5c,d). The black line represents the synthetic seismogram computed using XBT- and neural network–derived temperature and salinity profiles. Temperature and salinity data were converted to sound speed and density profiles using the Gibbs Seawater (GSW) Oceanographic Toolbox. These profiles were used to calculate impedance contrasts, which were convolved with a source wavelet matching the seismic data's dominant frequency to generate the synthetic seismograms. To align the synthetic seismograms with the observed seismic data, the XBT depth profiles were converted to TWTT using sound speed derived from GSW and then scaled with a fixed reference sound speed to match the seismic depth. This adjustment minimizes misalignment and ensures consistency between the synthetic and observed seismic data. Some small differences between the XBT and seismic data, for example, fluctuations at depths of 1300 and 1520 m, are likely the result of the seismic data being a sum of

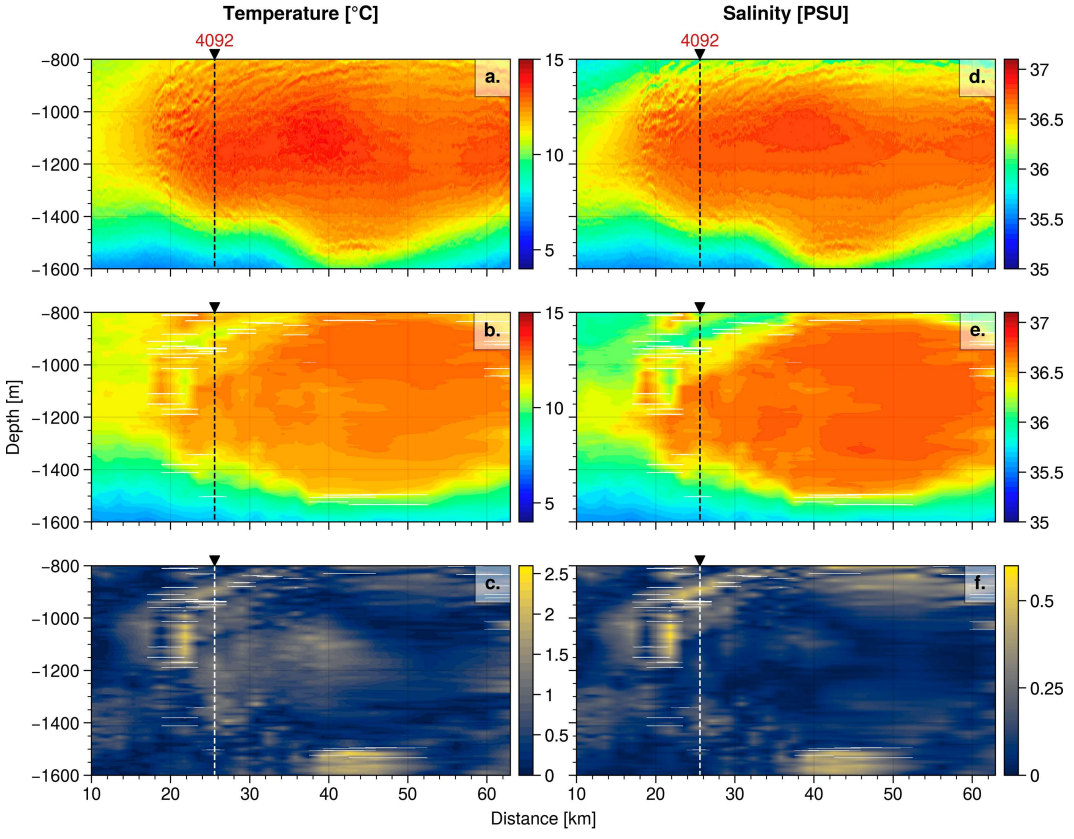


FIG. 4. Comparison of combined-inverted temperature and salinity fields with coincident hydrographic data for the GC seismic section. (a) Combined inversion (i.e., VA + MCMC) for in situ temperature. (b) In situ temperature from XBT data without smoothing. (c) Absolute difference between combined inversion temperatures and hydrographic data. (d)–(f) As in (a)–(c), but for practical salinity, with the field in (e) estimated from XBT data using a neural network approach. Dashed black vertical line denotes XBT used for detailed comparison in Fig. 5. The white strips in (b), (e), (c), and (f) represent missing values in the XBT dataset, where no temperature measurements were available.

reflectively collected over about an hour compared to the almost instantaneous XBT. These discrepancies will also contribute to some of the bias in the fine-scale variability of the inverted results and XBT data. We also considered the impact on MCMC results of noise in the seismic data, which was found to be negligible: inverted temperature and salinities were compared with results interpolated from neighboring CMPs, and differences did not exceed 0.2°C and 0.04 psu, respectively.

### b. *Mozambique Channel*

The Mozambique Channel funnels water from the northern Indian Ocean to its southern portions. The channel's circulation is characterized by a highly energetic western boundary current, the Mozambique Current. The current is made up of three major water mass layers: upper, intermediate, and deep water masses. The upper water masses include the warm, southward-flowing Subtropical Subsurface Water (STSW; $T > 15$°C, maximum $S > 35.2$ psu, center depths: 200 m) and South Indian Central Water (SICW; $T$: 7°–8°C, $S$: around 34.8–35.2 psu, depths 200–600 m) (Wyrtki 1973; Ullgren et al. 2012). The intermediate water masses consist of a mixture of cooler, northward-flowing Antarctic Intermediate Water (AAIW; $T$: 2.5°–7°C, minimum $S$: <34.5 psu, center depths 800–1200 m) (de Ruijter et al. 2002; Donohue and Toole 2003), and southward-spreading Red Sea Water (RSW) which reaches the Agulhas Current ($T$: 5°–10°C, maximum $S$: 34.7–35 psu, center depths 1000 m) (Swart et al. 2010; Ullgren et al. 2012): Interleaving between AAIW and RSW is common (Gründlingh 1985). The deep water mass is represented by the North Atlantic Deep water (NADW; $T$: 2°C, $S$: around 34.8 psu, depths > 2000 m) (Toole and Warren 1993; van Aken et al. 2004; Quartly and Srokosz 2004; Beal et al. 2000; Talley et al. 2011; Ullgren et al. 2012). These water masses and their interactions are visible in the seismic imagery, particularly in the southeast (SE), offshore end of the transect (Fig. 1d). Regions of higher reflectivity between 250–500 and 1000–1500 m represent highly stratified, interleaved waters, likely the SICW layer and interaction of the AAIW and RSW, respectively. An almost acoustically blank layer at depths greater than 1500 m is attributed to a more homogeneous, well-mixed water mass, possibly AAIW. Southward-propagating mesoscale eddies are commonplace in

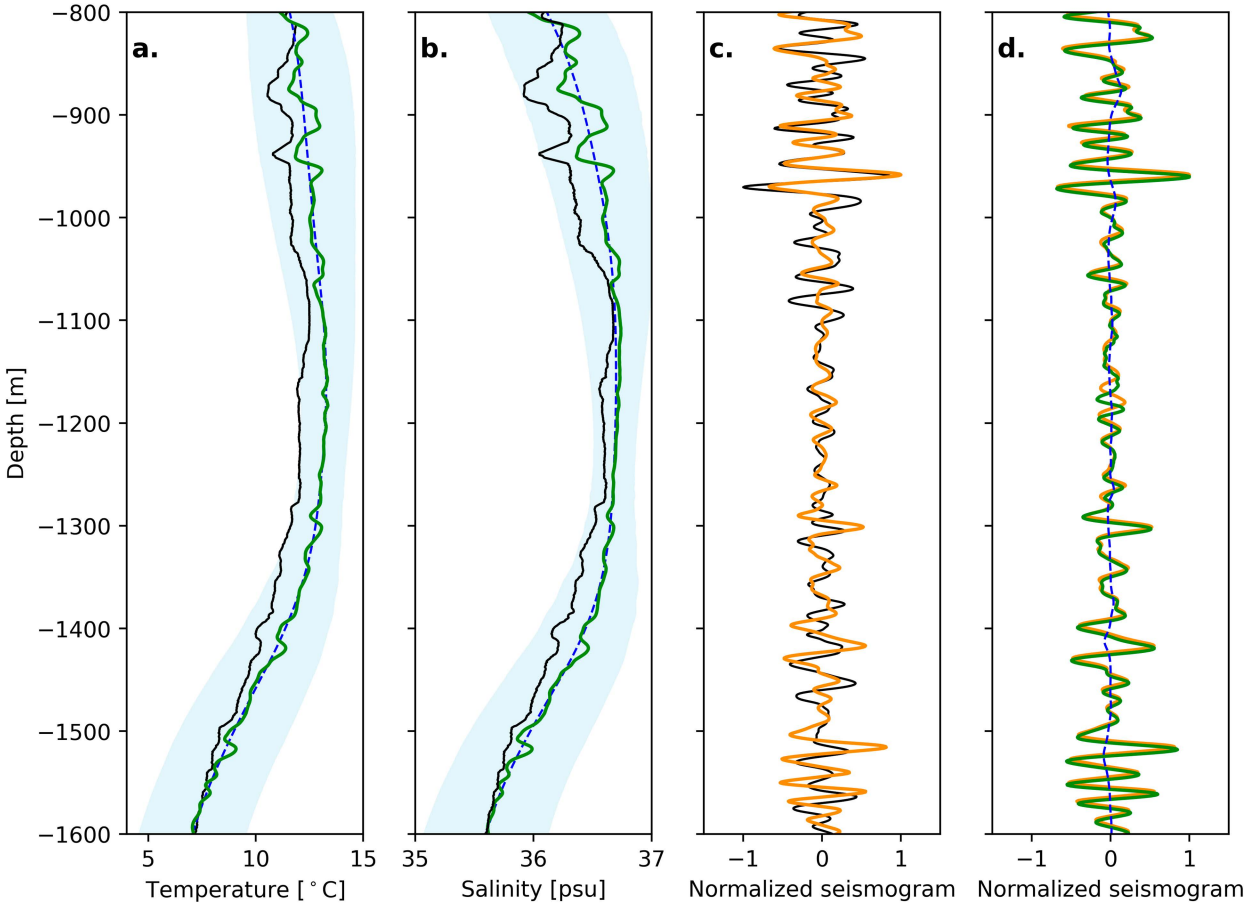


FIG. 5. Example depth profile comparing inverted data with hydrographic and seismic reflection data for GC. (a) Black line denotes temperature depth profile from XBT located at distance 25.5 m (CMP 4092), along section (see Fig. 4 for location); dashed blue line denotes temperature profile from the VA analysis only; green line denotes temperature profile from combined inversion (VA + MCMC); and shaded blue area denotes Dix equation error. (b) As in (a), but for salinity. (c) Black line denotes synthetic seismogram (i.e., modeled acoustic reflection response computed based on XBT-derived temperature and salinity profiles); orange line denotes seismogram from GC data. (d) Green line denotes synthetic seismogram from combined inversion (VA + MCMC), orange line denotes seismogram from GC data, and dashed blue line denotes synthetic seismogram computed from VA output only.

the Mozambique Channel (Halo et al. 2014), and surface geostrophic current fields indicate that the acoustic section transects the northeastern edge of one such anticyclonic mesoscale eddy (Fig. 1c). The seismic signature of this eddy appears as sloping reflections in the upper water column which dip from 250 m on the northwest (NW) end of the transect to 700 m at its SE extent. A more transparent region at distances greater than 40 km and depths of 800–1200 m may be a downwelling layer formed by the anticyclonic eddy transitioning the seismic survey at the time of acquisition (McGillicuddy 2016). Reflectivity generally increases toward the Mozambique coastline indicating the water masses become more stratified and interleaved in character here.

As commonly seen in ocean seismic images, submesoscale variability is apparent across the section as acoustically blank patches and swirling reflection boundaries (Quentel et al. 2010; Gunn et al. 2018; Tang et al. 2020). For example, three potential submesoscale lens-like structures, identified as transparent lenses with horizontal length scales of 20 km and vertical length scales of 200 m, respectively, are highlighted (Fig. 1d).

The VA-inversion temperature and salinity fields capture the general distribution of the water masses in the Mozambique Channel but are limited in resolving with intermediate scale, submesoscale details (Figs. 6a,e). Warmer, saltier waters are visible in the upper 300–500 m, consistent with STSW and SICW. Between 500 and 800 m, a layer of relatively warm, freshwater is likely AAIW. Waters below 1000 m are generally cooler and saltier but exhibit horizontal variability, particularly in salinity likely caused by RSW intruding into the AAIW layer (Gründlingh 1985; Ullgren et al. 2012). The lens-like structures identified in the seismic images are associated with regions of both warmer, saltier likely RSW-sourced water (distance 35–55 km, depth 1000–1250 m, and distance 20–32 km, 800–1100-m depth) and colder, fresher likely AAIW sourced water (52–65 km, 1200–1600-m depth): See white

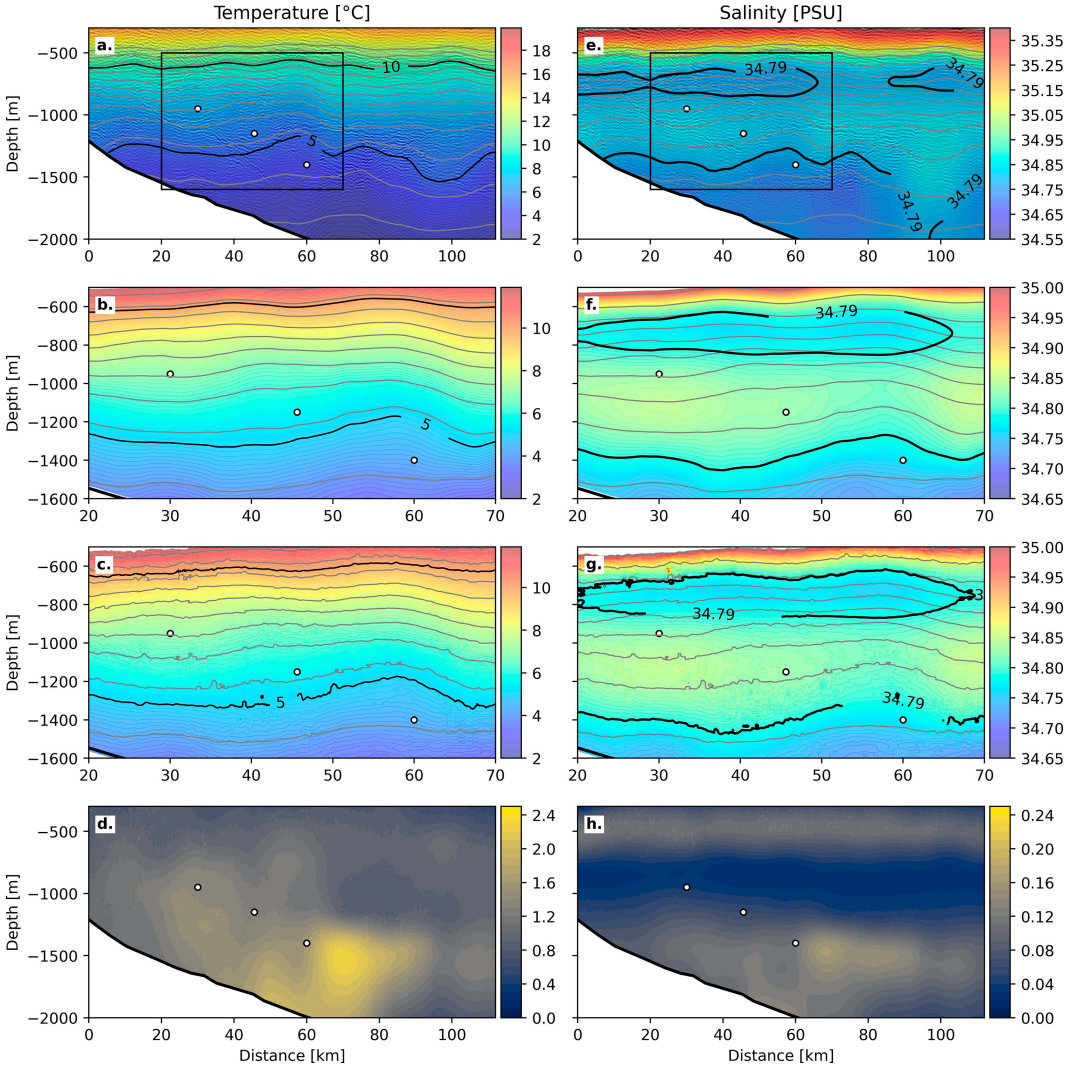


FIG. 6. Inverted temperature and salinity fields for the Mozambique seismic section, with uncertainties. (a) VA inversion for in situ temperature. Gray contours denote isopycnals ranging from 1026.6 to 1028.25 kg $m^{-3}$ with an interval of 0.1 kg $m^{-3}$, black line denotes RSW boundary characterized by potential temperatures above 5°C, and black box denotes demark region shown in (c). (b) Region indicated by box in (a). (c) As in (b), but for combined inversion (VA + MCMC) output. (d) Computed Dix equation inversion error, the dominant uncertainty source. (e)–(h) As in (a)–(d), but for salinity, and black lines identifying RSW by salinities between 34.79 and 35 psu. White dots denote possible submesoscale lens-like structures, as identified in Fig. 1d.

dots in Figs. 1d and 6 (Gorman et al. 2018; Ullgren et al. 2012).

The addition of MCMC processing step adds detail to the VA-inverted thermohaline fields (Figs. 6b,c,f,g). For example, density contours exhibit increased fine-scale structure and align better with the seismic reflections than in the VA-inverted fields alone (at least on horizontal scales $> 5$ km and vertical scales $> 50$ m). Fine-scale interleaving becomes apparent at the edges of lenses, providing a focused view of potential mixing and stirring processes.

With a lack of coincident hydrographic data, we utilize the error analysis methods verified in section 3c to compute uncertainties in the inverted thermohaline fields (Table 2). The overall mean error in temperature and salinity for the Mozambique Channel seismic section is 1.65°C and 0.08 psu, with a maximum error of 3.68°C and 0.23 psu, respectively. Similar to the Gulf of Cadiz dataset, the error associated with the Dix conversion of $v_{rms}$ to $v_{int}$ dominates. In terms of the spatial distribution, as in the Gulf of Cadiz data, higher temperatures uncertainties are present in regions of high reflectivity (e.g., at the boundaries of the submesoscale structures), with lower errors in more homogeneous areas (e.g., the core of the anticyclonic eddy). Interestingly, the salinity error distribution follows a different pattern, with higher errors in the upper 500 m and at the base of the layer of RSW intrusion (Figs. 6g,h). These arise from differences in how sound speed variations propagate through the

TABLE 2. Summary of inversion uncertainties for the MC data. As in Table 1, but without the bias column due to a lack of coincident hydrographic data.

| | Estimation | CMP interval | CMP location | Dix equation | MCMC posterior | Overall |
|---|---|---|---|---|---|---|
| Sound speed (m $s^{-1}$) | 0.63 (1.59) | 0.76 (1.43) | 0.5 (0.88) | 8.1 (36.6) | — | 9.99 (40.5) |
| Temperature (°C) | 0.2 (0.4) | 0.2 (0.3) | 0.1 (0.2) | 1.14 (2.26) | 0.01 (0.7) | 1.65 (3.68) |
| Salinity (psu) | 0.003 (0.009) | 0.004 (0.007) | 0.002 (0.009) | 0.07 (0.16) | 0.0008 (0.04) | 0.08 (0.23) |

inversion process. Notably, the sound speed models generated by adding or subtracting $v_{rms}$ errors to the original $v_{rms}$ are asymmetric. The high-error band in salinity at 500 m corresponds to the high-salinity SICW, where large salinity gradients increase sensitivity to sound speed errors.

Without coincident hydrographic data, the ability of the combined inversion method to capture mesoscale and smaller temporal and spatial variability cannot be assessed directly; however, averaged profiles align closely with historic WOD hydrographic data, and differences are well within computed errors (Figs. 7a,b).

The close alignment of inverted temperature and salinity profiles with key water mass characteristics (STSW, SICW, RSW, and AAIW) demonstrates the combined method's capacity to reproduce thermohaline structures (Fig. 7c). Additionally, green points represent four CTDs located at the southern edge of the Mozambique Channel showing the minimum salinity core of AAIW to highlight the dominant mixing between AAIW and RSW in the northern Mozambique Channel.

## 5. Discussion and conclusions

This work has developed a novel inversion approach that does not rely on coincident hydrographic data to recover temperature and salinity fields from seismic data, with uncertainty analysis. Inverted temperature and salinity uncertainties are on the order of 1°–2°C and 0.1 psu, respectively. The approach can therefore be applied to the wealth of seismic data available for which coeval in situ temperature and salinity measurements are typically not available.

The new inversion combines two established inversion approaches: velocity analysis inversion and spatially iterative Markov Chain Monte Carlo method. The VA inversion

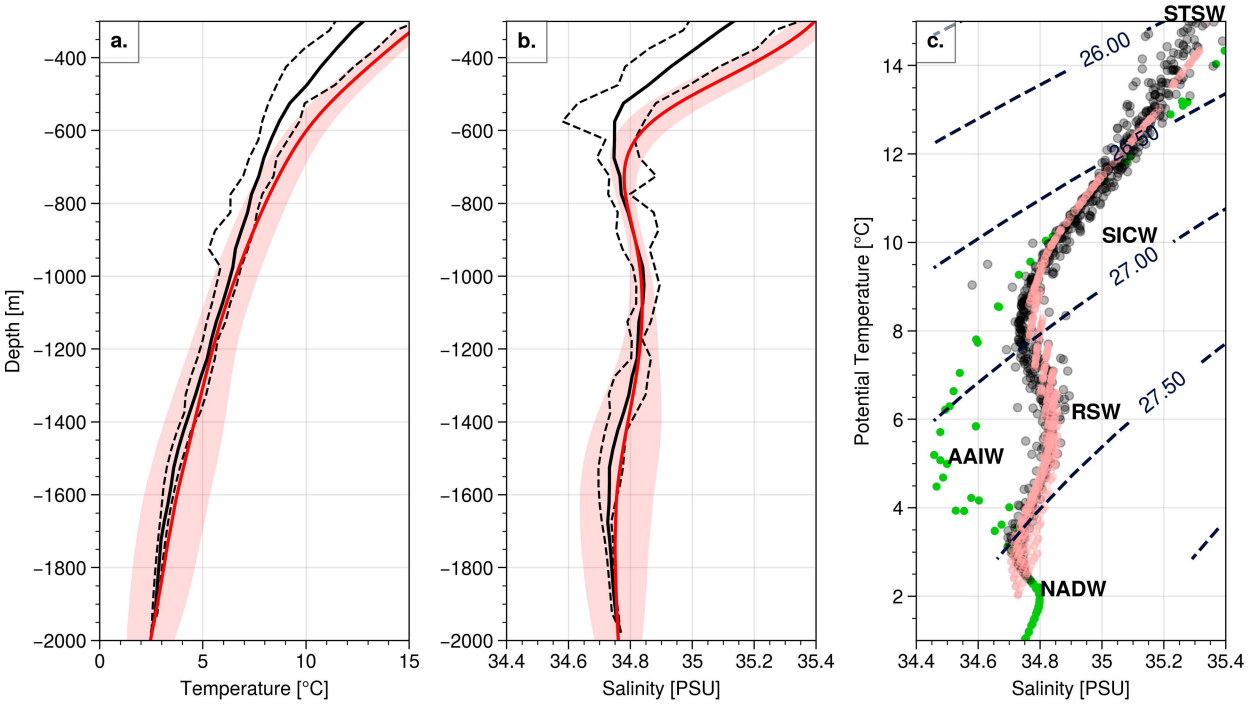


FIG. 7. Comparison of historical hydrographic and inverted data for MC. (a) Black solid line denotes mean temperature profile for all WOD historical data (see Fig. 1c), black dashed lines denote minimum and maximum ranges for WOD data, red solid line denote mean seismic section temperature profile from combined inversion, and red shaded region denotes Dix equation error. (b) As in (a), but for salinity depth profiles. (c) Temperature–salinity diagram with density contours (black dashed lines), black points denote WOD historical data, red transparent points denote inverted potential temperature and salinity data, and green points denote four CTDs from Global Ocean Data Analysis Project (GLODAPv2.2023) data at latitude of 24.6°S and longitude of 40.5°–41°E. Key water masses are labeled, including STSW, SICW, RSW, AAIW, and NADW.

provides the long-wavelength (on the order of 100 m–1 km horizontally) variation of temperature and salinity, while the MCMC adds fine-scale detail (on the order of 10 m horizontally and 10–100 m vertically).

Here, we apply the new combined inversion to two seismic datasets, one from the Gulf of Cadiz and the other from the Mozambique Channel. The Gulf of Cadiz dataset, which contains coeval hydrographic data, was used to validate the combined inversion approach and uncertainty analysis. The Mozambique Channel data, for which no coincident temperature or salinity data were available, demonstrated the utility of the method for legacy seismic data. Submesoscale thermohaline variability, for example, interleaving layers at the periphery of the submesoscale lenses, was resolved. One such lens-like structure in the Mozambique Channel data, located farthest offshore (52–65 km, 1200–1600-m depth), exhibited a diminished temperature and salinity signature, potentially indicating an "older" feature. This may reflect prolonged exposure to mixing with surrounding water, moving farther from its source region. These submesoscale vortex-type features are likely generated through interactions between RSW, AAIW, mesoscale eddies, and the continental slope (Roberts et al. 2014; Vic et al. 2015).

Compared to previous studies, our uncertainties in inverted temperature (1.65°–2.5°C) and salinity (0.08–0.5 psu) are notably higher. For example, Papenberg et al. (2010) report temperature (salinity) uncertainties of 0.1°C (0.1 psu), Dagnino et al. (2016) report temperature (salinity) uncertainties of 0.18°C (0.08 psu), and Xiao et al. (2021) report temperature (salinity) uncertainties of 0.16°C (0.055 psu). The difference reflects the lack of utility of coincident hydrographic data in our approach, confirming the conclusion from previous studies that numerous hydrographic measurements are key to reducing uncertainties in the inversion process (Xiao et al. 2021). Reflecting this, errors in our final inverted fields are dominated by the error in producing the background long-wavelength temperature and salinity field in the VA process (as opposed to the MCMC). In particular, the greatest contribution to the uncertainty is the Dix interval velocity estimation, which was not quantified in previous studies (e.g., Gunn et al. 2018; Chhun and Tsuji 2020). Minimizing Dix equation errors involves using large TWTT/depth intervals between velocity picks; however, this is at the expense of loosing detail in the final inverted field. Regions of rapidly varying temperature and salinity are therefore more prone to error than more homogeneous regions, despite the increased signal to noise in the seismic data. On the other hand, if no strong reflections are present over too large a region, accurate velocity picking is also comprised. The lack of reflectivity in the upper 500 m of the Gulf of Cadiz data made achieving a sensible RMS velocity field somewhat challenging and very careful picking was required (surface errors also propagated to picks deeper in the water column). Velocity picking in the water column is generally less subjective than in solid Earth, as the acoustic structure is simpler and not affected by interbed multiples or complex lithological layering. This results in clearer semblance coherence peaks, which facilitates more consistent velocity estimation. However, noise contamination can still introduce uncertainty by broadening the coherence energy distribution in semblance plots, especially in regions of weak thermohaline stratification or low signal-to-noise ratio. A notable example is found in shallow water, where residual energy from direct wave arrivals can interfere with early reflections and bias semblance amplitudes, particularly near the surface. This effect is evident in the Mozambique Channel dataset, where the semblance peaks in the upper 500 m are broad and poorly defined due to noise and weak stratification (Fig. A2). Automated methods for RMS velocity extraction to further reduce uncertainties and improve inversion resolution could reduce subjective biases in manual picking, improve efficiency in picking each CMP velocity, and minimize errors associated with CMP intervals and locations (e.g., Chhun and Tsuji 2020). Although the errors are high, this inversion is released from the requirements of coincident hydrographic data. Further work should go into increasing the accuracy and conversion of velocity fields.

The gap between the scales resolvable by VA inversion and MCMC methods poses a challenge for accurately extracting temperature and salinity from seismic ocean data. The VA inversion is suited to resolving features with horizontal scales greater than 10 km and vertical scales greater than 100 m, while MCMC excels at capturing fine-scale details with horizontal and vertical scales on the order of 10 m. For example, in the Gulf of Cadiz dataset, while the fine-scale variability was reproduced, the MCMC output was unable to adjust for systematic errors introduced in the previous VA inversion. Developing inversion methods that can bridge this spatial gap is a priority. Potentially, this could be achieved by integrating other data sources such as high-resolution satellite data or autonomous underwater vehicle observations or the use of full-waveform inversion techniques (Wood et al. 2008; Dagnino et al. 2016). Improving the input model needed for the MCMC analysis could also be achieved through an iterative, adaptive approach within the VA analysis itself or utilizing advances in machine learning (Padhi et al. 2010). A recent study by Mallick and Chakraborty (2022) proposed an approach for initial sound-speed model generation without manual horizon picking, coupled with genetic algorithm–based waveform inversion. This aligns with our motivation to reduce the dependence on manual velocity picking and improve the robustness of temperature and salinity inversion from seismic data. Data assimilation of seismic fields into high-resolution models is another exciting future avenue to both validate model outputs and improve understanding of submesoscale oceanic processes and their broader impacts (Sonnewald et al. 2021; Qi et al. 2022). For example, Sonnewald et al. (2021) emphasize the transformative role of machine learning in three key areas: observations, theory, and numerical modeling. Machine learning has shown promise in improving in situ sampling, satellite observations, theoretical exploration, and even data assimilation processes, accelerating research progress, and enhancing the accuracy of oceanographic studies. Similarly, Qi et al. (2022) demonstrate the capability of ensemble machine learning models to reconstruct subsurface thermal structures using sea surface and geographic data. Building on these advancements, integrating seismic data, machine learning–driven interior data, and high-resolution numerical models

could provide unprecedented insights into the dynamics and impacts of submesoscale and mesoscale processes in the ocean.

Despite these future challenges, the novel combination of VA analysis and MCMC methods represents a significant step forward in extracting temperature and salinity with representative uncertainties from seismic ocean data that lack coincident hydrographic information. Given the global coverage of legacy seismic data in key oceanographic regions, the techniques presented here will aid new insights into fine-scale and submesoscale ocean processes for which observations are scarce.

*Acknowledgments.* All authors contributed to the writing and research of this article. Yao Meng acknowledges support from the China Scholarship Council and the University of Exeter. We sincerely thank Wuxin Xiao for sharing the MCMC inversion code and making this work possible. The Mozambique Channel seismic data were generously provided by Schlumberger Ltd. We are grateful to Richard Hobbs for modeling the simulated source wavelet using the Nucleus + airgun modeling software and Robert J. W. Brewin for the discussion of this work at the early stage. The seismic sections were processed using OMEGA/Western-Schlumberger and MATLAB, with final visualization completed using Python. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

*Data availability statement.* Bathymetry was downloaded from GEBCO 2022 dataset (GEBCO Compilation Group 2022). Altimeter products (i.e., surface geostrophic currents) were generated using the Global Ocean Physics Reanalysis product with a resolution of 1/12° (https://doi.org/10.48670/moi-00021). The historical hydrographic data were downloaded from World Ocean Database (https://www.ncei.noaa.gov/products/world-ocean-database). GC data can be made available on request through contacting the corresponding author or Katy Sheen on k.l.sheen@exeter.ac.uk. Confidential Data: Mozambique data courtesy of SLB and Instituto Nacional de Petróleo (INP) cannot be shared without gaining authorization from SLB. Please contact corresponding author or Katy Sheen on k.l.sheen@exeter.ac.uk for enquiry.

## APPENDIX

### Supplementary Figures of Temperature–Salinity Inversion and Error Analysis

Figures A1–A5 show additional figures that complement the methodology and uncertainty analysis described in the main text. Figure A1 shows WOD temperature–salinity scatter plots for the Gulf of Cadiz and Mozambique Channel datasets, together with the neural-network fitting approach used to establish the local temperature–salinity–depth relationship. Figure A2 presents velocity analysis results from both datasets. Semblance plots are displayed alongside derived RMS and interval velocities, as well as associated errors. Figures A3 and A5 demonstrate the uncertainty in GC and MC dataset associated with the estimation procedure, choice of CMP interval, and CMP location during velocity analysis. Figure A4 displays the evolution of the MCMC temperature–salinity chain for one depth in each dataset, showing how the inversion converges to stable values and uncertainties.

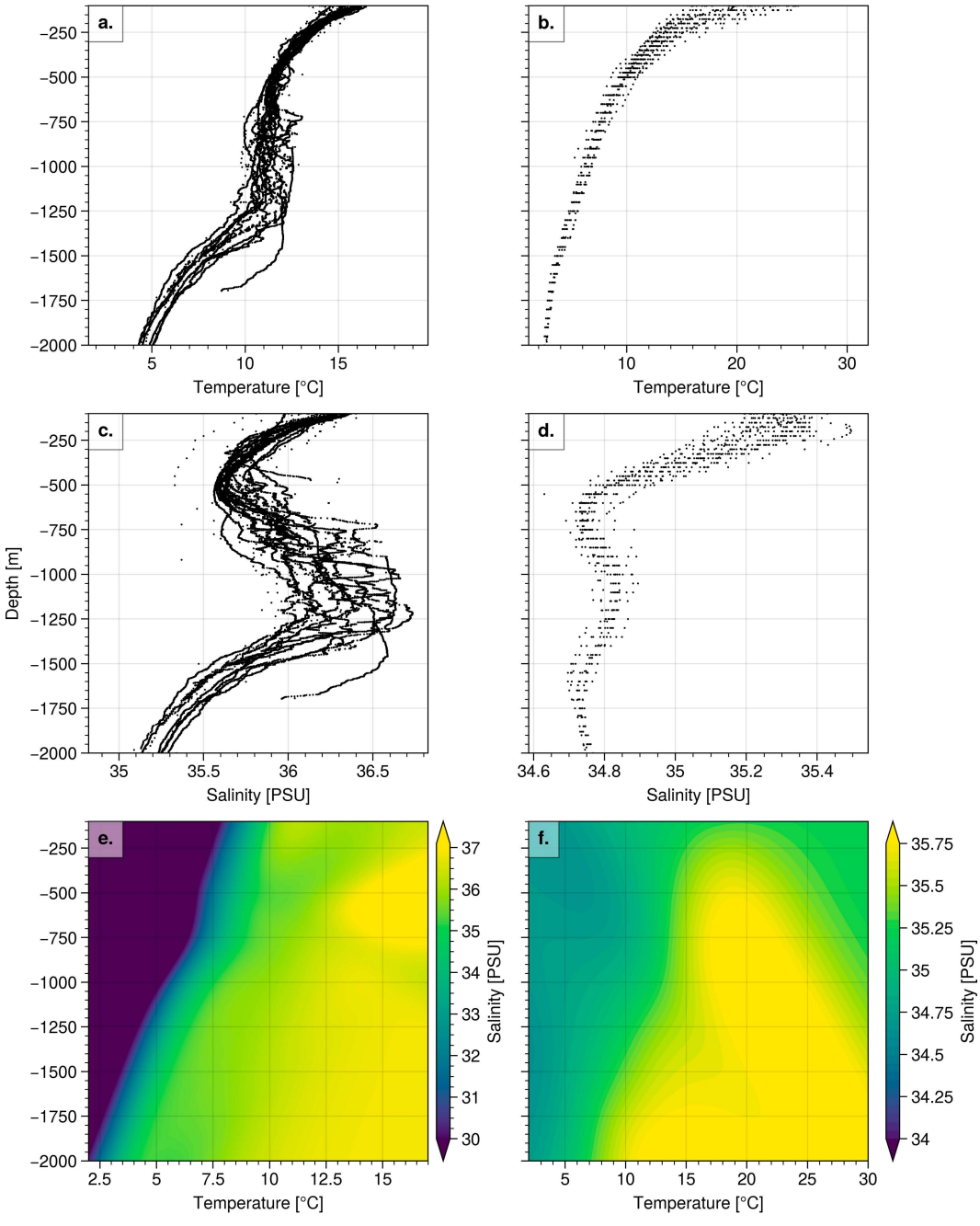


FIG. A1. (a)–(d) Map of WOD data with temperature and salinity scatters. (e),(f) The relationship between temperature, salinity, and depth. The neural network fitting approach was employed to train the salinity function of temperature and depth. The salinity is corresponding to a pair of temperature and salinity. A neural network was trained using coincident CTD data to establish the temperature–salinity–depth relationship. (a),(c),(e) GC and (b),(d),(f) MC.

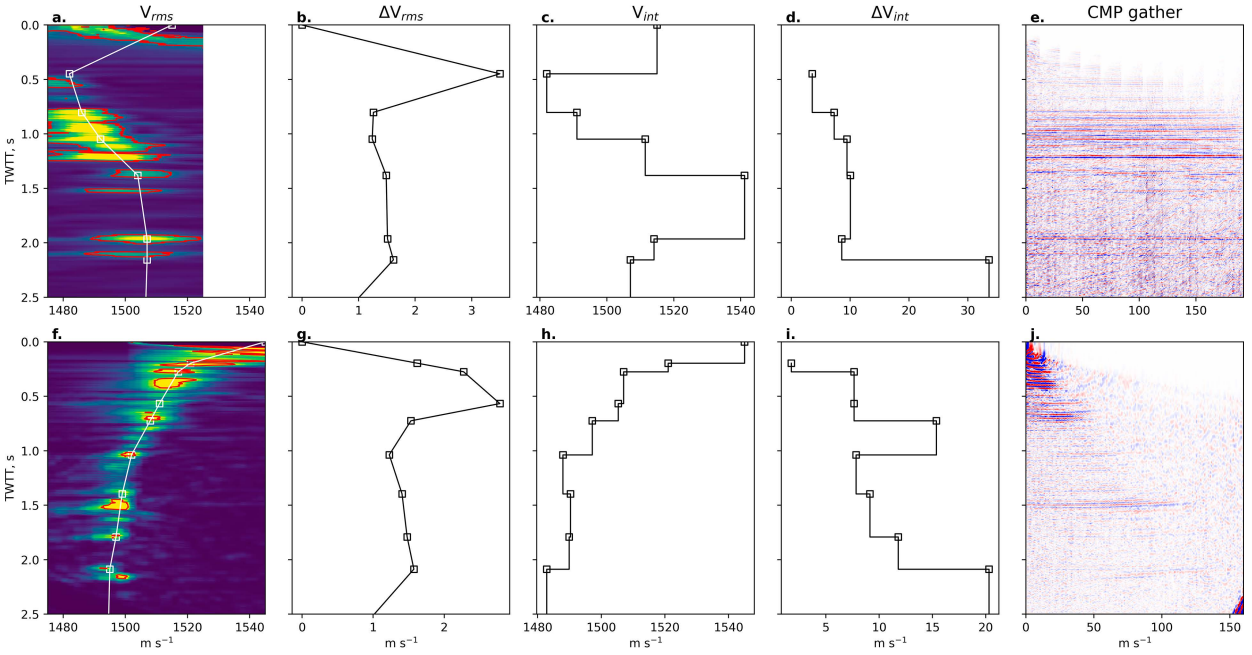


FIG. A2. VA from two seismic datasets. The VA estimates signal coherency for a range of stacking velocities over TWTT, generating semblance plots that reveal how ocean sound speed varies with depth and across the seismic survey. By selecting the peaks of maximum coherency, the sound speed at each reflection point can be determined. (a),(f) Semblance plots for the GC at CMP5650 and Mozambique datasets at CMP35800 that show picked RMS velocity as a function of TWTT. The TWTT is the total time required for seismic waves emitted from the source to reach an impedance and to be reflected or refracted back to the receiver. The warm colors represent the optimal RMS velocity. Black squares are the chosen RMS velocity. The same colormap range is shared in (a) and (f), allowing for a direct comparison that highlights the extensive range of velocity values in (a). Red contour level is 0.2. (e),(j) The NMO-corrected CMP gather for picked RMS velocity. The final seismic image is created by a series of stacked NMO corrected CMP gathers. (b)–(d) RMS velocity error, interval velocity, and interval velocity error for the GC data, respectively. (g)–(i) As in (b)–(d), but for MC data.

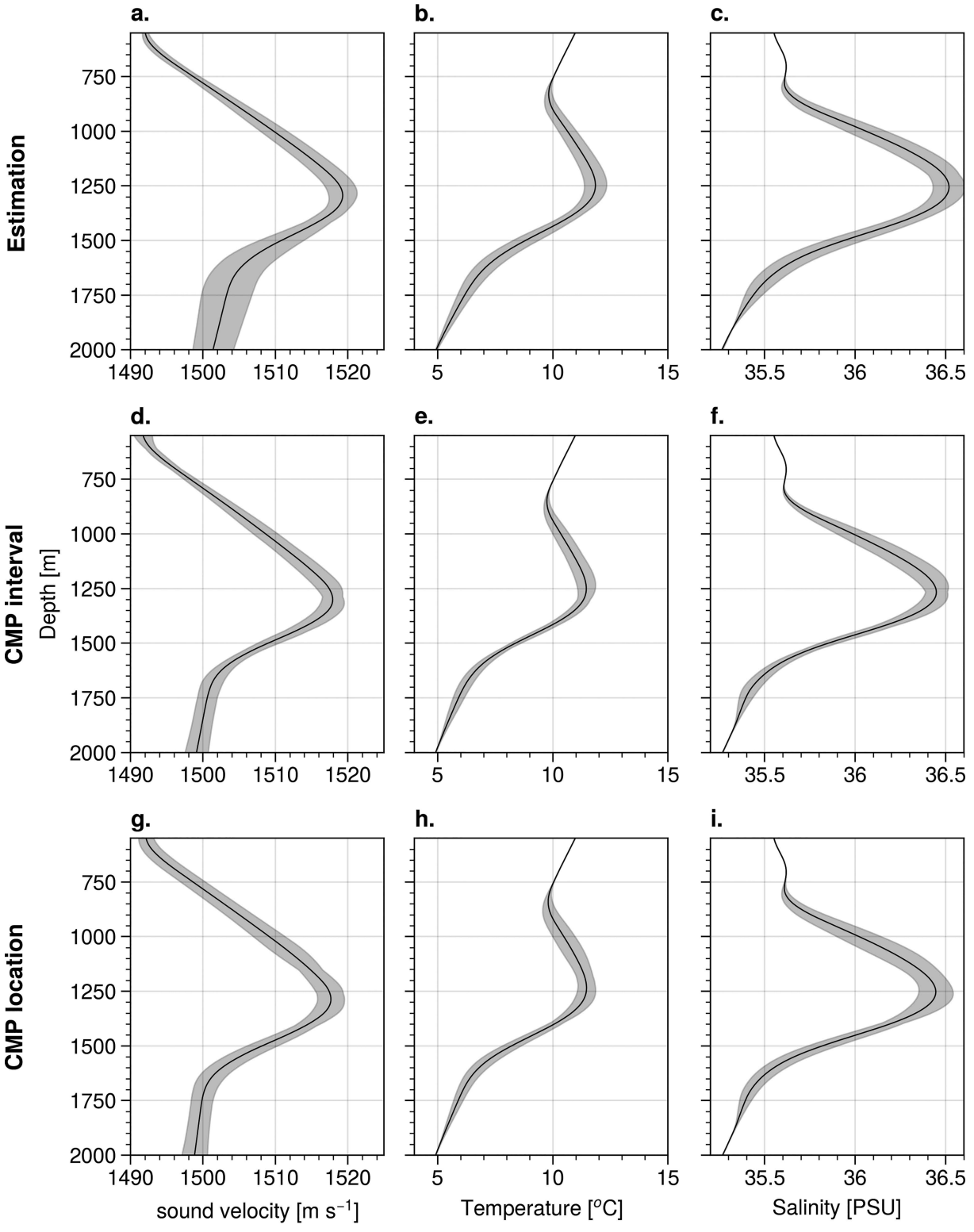


FIG. A3. Error caused by the estimation, choice of CMP interval, and location during VA, GC. (left) The horizontally averaged sound speed (black solid line) converted from (a) three separate VAs performed at the same CMP locations, (b) VAs performed at CMP intervals of 50 (0.3 km), 100 (0.6 km), 150 (0.9 km), and 200 (1.2 km), and (c) VAs performed at CMP intervals of 100 (0.6 km) and 150 (0.9 km) with different CMP locations, respectively. The shaded area is calculated from the std at each depth. (b),(e),(h) As in (a), (d), and (g), but for the temperature. (c),(f),(i) As in (a), (d), and (g), but for the salinity. The maximum of depth-averaged std from panels is taken as the error associated with the estimation, choice of CMP interval, and location during VA.

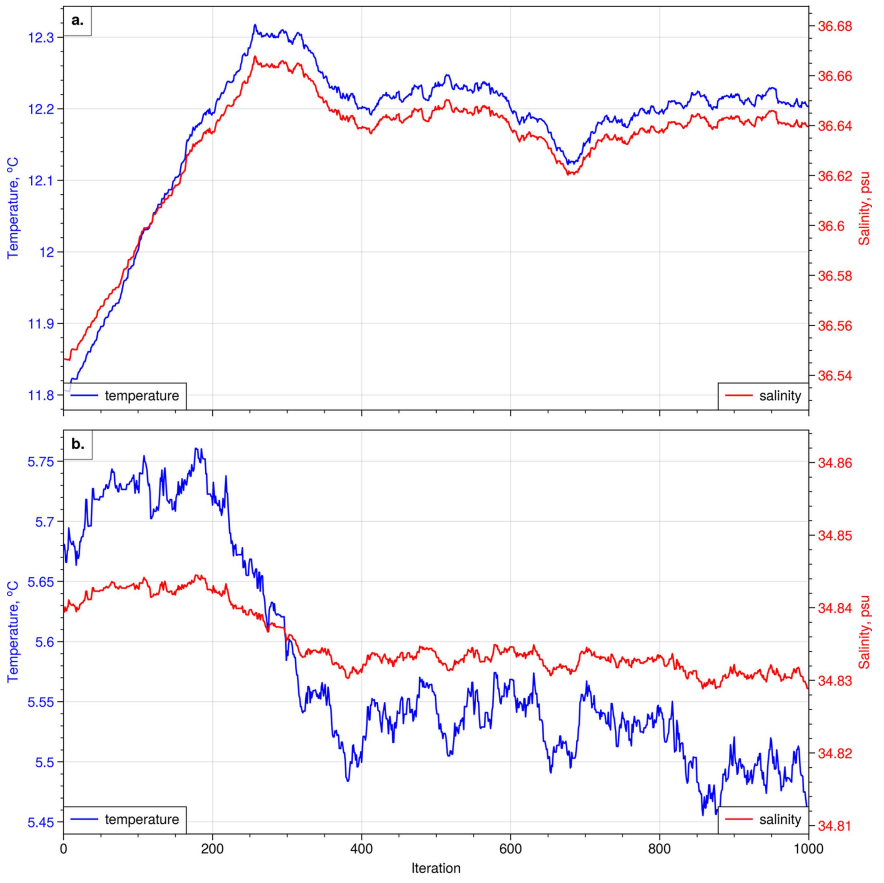


FIG. A4. The Markov temperature and salinity chain. (a) The GC *T*–*S* chain at a depth of 1245 m, CMP 3497. (b) The *T*–*S* chain of MC at depth of 1200 m, CMP 43401. The first iteration value was obtained from the *T*–*S* background. Once a sufficient number of iterations have been completed, the temperature and salinity values converge and exhibit fluctuations within a defined range. The mean and std derived from the final 20% of iterations are utilized to determine the MCMC *T*–*S* field and its associated uncertainty.

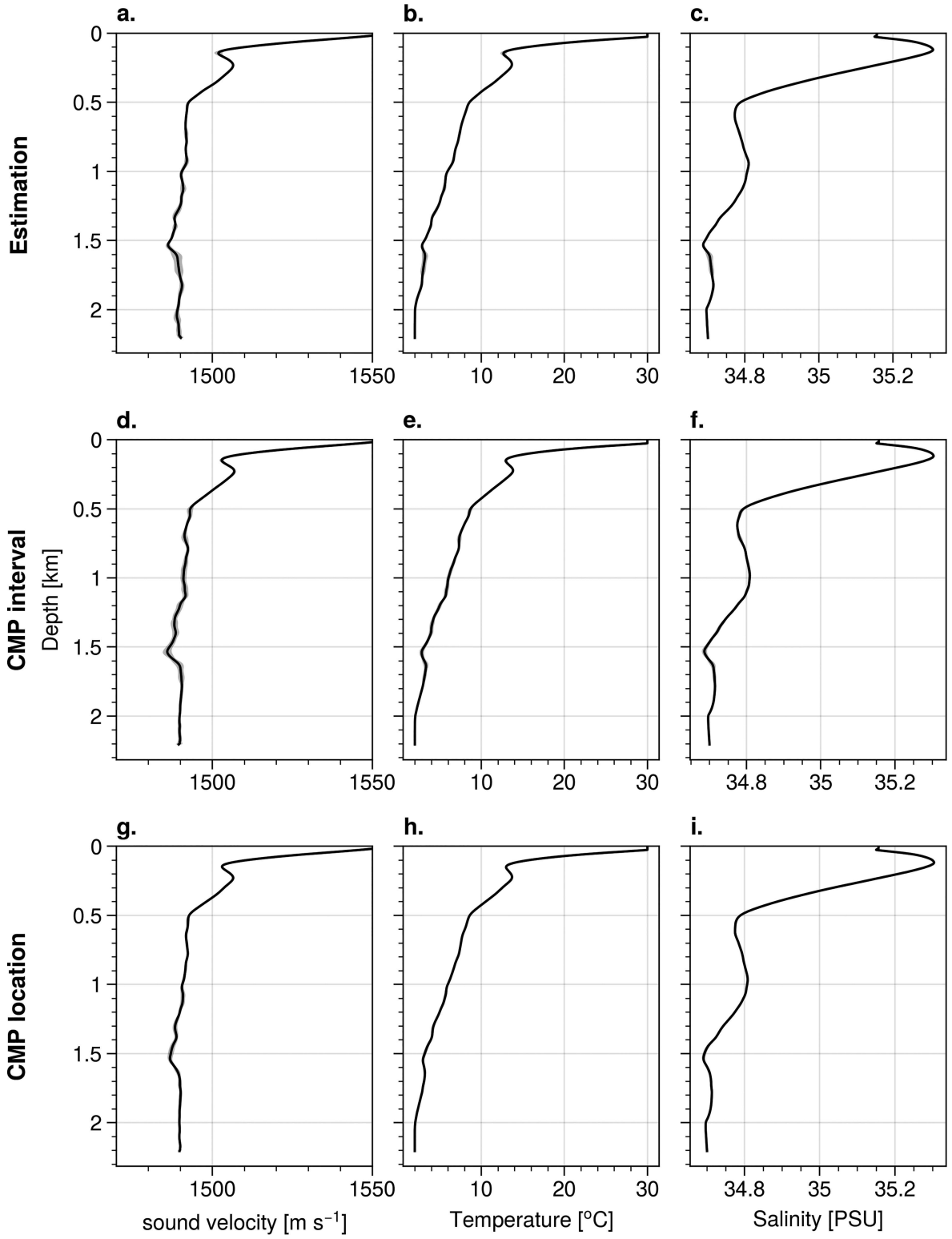


FIG. A5. As in Fig. A3, but for the MC dataset.

## REFERENCES


Armi, L., and W. Zenk, 1984: Large lenses of highly saline Mediterranean water. *J. Phys. Oceanogr.*, **14**, 1560–1576, https://doi.org/10.1175/1520-0485(1984)014<1560:LLOHSM>2.0.CO;2.

Aulicino, G., C. Cesarano, M. Zerrouki, S. Ruiz, G. Budillon, and Y. Cotroneo, 2021: On the use of ABACUS high resolution glider observations for the assessment of phytoplankton ocean biomass from CMEMS model products. *Ecol. Modell.*, **455**, 109619, https://doi.org/10.1016/j.ecolmodel.2021.109619.

Ballabrera-Poy, J., B. Mourre, E. Garcia-Ladona, A. Turiel, and J. Font, 2009: Linear and non-linear *T*–*S* models for the eastern North Atlantic from Argo data: Role of surface salinity observations. *Deep-Sea Res. I*, **56**, 1605–1614, https://doi.org/10.1016/j.dsr.2009.05.017.

Bashmachnikov, I., F. Neves, T. Calheiros, and X. Carton, 2015: Properties and pathways of Mediterranean water eddies in the Atlantic. *Prog. Oceanogr.*, **137**, 149–172, https://doi.org/10.1016/j.pocean.2015.06.001.

Beal, L. M., A. Ffield, and A. L. Gordon, 2000: Spreading of Red Sea Overflow waters in the Indian Ocean. *J. Geophys. Res.*, **105**, 8549–8564, https://doi.org/10.1029/1999JC900306.

Biescas, B., V. Sallarès, J. L. Pelegrí, F. Machín, R. Carbonell, G. Buffett, J. J. Dañobeitia, and A. Calahorrano, 2008: Imaging meddy finestructure using multichannel seismic reflection data. *Geophys. Res. Lett.*, **35**, L11609, https://doi.org/10.1029/2008GL033971.

——, B. R. Ruddick, M. R. Nedimovic, V. Sallarès, G. Bornstein, and J. F. Mojica, 2014: Recovery of temperature, salinity, and potential density from ocean reflectivity. *J. Geophys. Res. Oceans*, **119**, 3171–3184, https://doi.org/10.1002/2013JC009662.

Bower, A. S., L. Armi, and I. Ambar, 1997: Lagrangian observations of meddy formation during a Mediterranean undercurrent seeding experiment. *J. Phys. Oceanogr.*, **27**, 2545–2575, https://doi.org/10.1175/1520-0485(1997)027<2545:LOOMFD>2.0.CO;2.

Cao, H., and Coauthors, 2024: Isopycnal submesoscale stirring crucially sustaining subsurface chlorophyll maximum in ocean cyclonic eddies. *Geophys. Res. Lett.*, **51**, e2023GL105793, https://doi.org/10.1029/2023GL105793.

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin, 2008: Mesoscale to submesoscale transition in the California Current system. Part I: Flow structure, eddy flux, and observational tests. *J. Phys. Oceanogr.*, **38**, 29–43, https://doi.org/10.1175/2007JPO3671.1.

Chapman, P., S. F. Di Marco, R. E. Davis, and A. C. Coward, 2003: Flow at intermediate depths around Madagascar based on ALACE float trajectories. *Deep-Sea Res. II*, **50**, 1957–1986, https://doi.org/10.1016/S0967-0645(03)00040-7.

Chassignet, E. P., and X. Xu, 2021: On the importance of high-resolution in large-scale ocean models. *Adv. Atmos. Sci.*, **38**, 1621–1634, https://doi.org/10.1007/s00376-021-0385-7.

Chhun, C., and T. Tsuji, 2020: Sound speed of thermohaline fine structure in the Kuroshio Current inferred from automatic sound speed analysis. *Explor. Geophys.*, **51**, 581–590, https://doi.org/10.1080/08123985.2020.1736548.

Dagnino, D., V. Sallarès, B. Biescas, and C. R. Ranero, 2016: Fine-scale thermohaline ocean structure retrieved with 2-D prestack full-waveform inversion of multichannel seismic data: Application to the Gulf of Cadiz (SW Iberia). *J. Geophys. Res. Oceans*, **121**, 5452–5469, https://doi.org/10.1002/2016JC011844.

de Ruijter, W. P. M., H. Ridderinkhof, J. R. E. Lutjeharms, M. W. Schouten, and C. Veth, 2002: Observations of the flow in the Mozambique Channel. *Geophys. Res. Lett.*, **29**, 1502, https://doi.org/10.1029/2001GL013714.

Dickinson, A., and K. L. Gunn, 2022: The next decade of seismic oceanography: Possibilities, challenges and solutions. *Front. Mar. Sci.*, **9**, 736693, https://doi.org/10.3389/fmars.2022.736693.

——, N. J. White, and C. P. Caulfield, 2020: Time-lapse acoustic imaging of mesoscale and fine-scale variability within the Faroe-Shetland Channel. *J. Geophys. Res. Oceans*, **125**, e2019JC015861, https://doi.org/10.1029/2019JC015861.

Dix, C. H., 1955: Seismic velocities from surface measurements. *Geophysics*, **20**, 68–86, https://doi.org/10.1190/1.1438126.

Donohue, K. A., and J. M. Toole, 2003: A near-synoptic survey of the southwest Indian Ocean. *Deep-Sea Res. II*, **50**, 1893–1931, https://doi.org/10.1016/S0967-0645(03)00039-0.

Eakin, D., W. S. Holbrook, and I. Fer, 2011: Seismic reflection imaging of large-amplitude lee waves in the Caribbean Sea. *Geophys. Res. Lett.*, **38**, L21601, https://doi.org/10.1029/2011GL049157.

Ehmen, T., K. Sheen, A. Watson, A. Brearley, M. Palmer, D. Roper, and A. Thompson, 2022: High resolution acoustic imaging of frontal dynamics and thermohaline finestructure in the southwest Atlantic sector of the Southern Ocean. *EGU General Assembly 2022*, Vienna, Austria, European Geosciences Union, EGU22-411, https://doi.org/10.5194/egusphere-egu22-411.

Fan, W., H. Song, Y. Gong, S. Sun, K. Zhang, D. Wu, Y. Kuang, and S. Yang, 2021: The shoaling mode-2 internal solitary waves in the Pacific coast of Central America investigated by marine seismic survey data. *Cont. Shelf Res.*, **212**, 104318, https://doi.org/10.1016/j.csr.2020.104318.

Fortin, W. F. J., W. S. Holbrook, and R. W. Schmitt, 2017: Seismic estimates of turbulent diffusivity and evidence of nonlinear internal wave forcing by geometric resonance in the South China Sea. *J. Geophys. Res. Oceans*, **122**, 8063–8078, https://doi.org/10.1002/2017JC012690.

GEBCO Compilation Group, 2022: GEBCO_2022 Grid. https://doi.org/10.5285/e0f0bb80-ab44-2739-e053-6c86abc0289c.

Geng, M., H. Song, Y. Guan, and Y. Bai, 2019: Analyzing amplitudes of internal solitary waves in the northern South China Sea by use of seismic oceanography data. *Deep-Sea Res. I*, **146**, 1–10, https://doi.org/10.1016/j.dsr.2019.02.005.

Gorman, A. R., M. W. Smillie, J. K. Cooper, M. H. Bowman, R. Vennell, W. S. Holbrook, and R. Frew, 2018: Seismic characterization of oceanic water masses, water mass boundaries, and mesoscale eddies SE of New Zealand. *J. Geophys. Res. Oceans*, **123**, 1519–1532, https://doi.org/10.1002/2017JC013459.

Gründlingh, M. L., 1985: Occurrence of red sea water in the southwestern Indian Ocean, 1981. *J. Phys. Oceanogr.*, **15**, 207–212, https://doi.org/10.1175/1520-0485(1985)015<0207:OORSWI>2.0.CO;2.

Gula, J., T. M. Blacic, and R. E. Todd, 2019: Submesoscale coherent vortices in the Gulf Stream. *Geophys. Res. Lett.*, **46**, 2704–2714, https://doi.org/10.1029/2019GL081919.

——, J. Taylor, A. Shcherbina, and A. Mahadevan, 2022: Submesoscale processes and mixing. *Ocean Mixing: Drivers, Mechanisms and Impacts*, M. Meredith and A. Naveira Garabato, Eds., Elsevier, 181–214, https://doi.org/10.1016/B978-0-12-821512-8.00015-3.

Gunn, K. L., N. J. White, R. D. Larter, and C. P. Caulfield, 2018: Calibrated seismic imaging of eddy-dominated warm-water transport across the Bellingshausen Sea, Southern Ocean. *J. Geophys. Res. Oceans*, **123**, 3072–3099, https://doi.org/10.1029/2018JC013833.

——, N. White, and C.-C. P. Caulfield, 2020: Time-lapse seismic imaging of oceanic fronts and transient lenses within South Atlantic Ocean. *J. Geophys. Res. Oceans*, **125**, e2020JC016293, https://doi.org/10.1029/2020JC016293.

Gutjahr, O., D. Putrasahan, K. Lohmann, J. H. Jungclaus, J.-S. von Storch, N. Brüggemann, H. Haak, and A. Stössel, 2019: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP). *Geosci. Model Dev.*, **12**, 3241–3281, https://doi.org/10.5194/gmd-12-3241-2019.

Haarsma, R. J., and Coauthors, 2016: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.*, **9**, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016.

Hajnal, Z., and I. T. Sereda, 1981: Maximum uncertainty of interval velocity estimates. *Geophysics*, **46**, 1543–1547, https://doi.org/10.1190/1.1441160.

Halo, I., B. Backeberg, P. Penven, I. Ansorge, C. Reason, and J. E. Ullgren, 2014: Eddy properties in the Mozambique Channel: A comparison between observations and two numerical ocean circulation models. *Deep-Sea Res. II*, **100**, 38–53, https://doi.org/10.1016/j.dsr2.2013.10.015.

Hobbs, R., 2007a: Geophysical Oceanography (GO): Anewtool to understand the thermal structure and dynamics of oceans. *AAPG Eur. Reg. News.,* Vol. 2 (7).

——, 2007b: Rrs discovery cruise 318: Hydrographic, chemical and biological studies in the North Atlantic Ocean. https://www.bodc.ac.uk/data/documents/cruise/8182/#cii.

Hobbs, R. W., D. Klaeschen, V. Sallarès, E. Vsemirnova, and C. Papenberg, 2009: Effect of seismic source bandwidth on reflection sections to image water structure. *Geophys. Res. Lett.*, **36**, L00D08, https://doi.org/10.1029/2009GL040215.

Holbrook, W. S., P. Páramo, S. Pearse, and R. W. Schmitt, 2003: Thermohaline fine structure in an oceanographic front from seismic reflection profiling. *Science*, **301**, 821–824, https://doi.org/10.1126/science.1085116.

——, I. Fer, R. W. Schmitt, D. Lizarralde, J. M. Klymak, L. C. Helfrich, and R. Kubichek, 2013: Estimating oceanic turbulence dissipation from seismic images. *J. Atmos. Oceanic Technol.*, **30**, 1767–1788, https://doi.org/10.1175/JTECH-D-12-00140.1.

Jones, I. F., and S. Levy, 1987: Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loeve transform. *Geophys. Prospect.*, **35**, 12–32, https://doi.org/10.1111/j.1365-2478.1987.tb00800.x.

Jones, S. M., C. Sutton, R. J. J. Hardy, and D. Hardy, 2010: Seismic imaging of variable water layer sound speed in Rockall Trough, NE Atlantic and implications for seismic surveying in deep water. *Geol. Soc. London Pet. Geol. Conf. Ser.*, **7**, 549–558, https://doi.org/10.1144/0070549.

Li, Y., and Coauthors, 2020: Eddy-resolving simulation of cas-LICOM3 for phase 2 of the Ocean Model Intercomparison Project. *Adv. Atmos. Sci.*, **37**, 1067–1080, https://doi.org/10.1007/s00376-020-0057-z.

Mahadevan, A., 2016: The impact of submesoscale physics on primary productivity of plankton. *Annu. Rev. Mar. Sci.*, **8**, 161–184, https://doi.org/10.1146/annurev-marine-010814-015912.

Mallick, S., and D. Chakraborty, 2022: Prediction of the ocean water sound speeds via attribute-guided seismic waveform inversion. *Geophysics*, **87**, U67–U79, https://doi.org/10.1190/geo2021-0442.1.

McDowell, S. E., and H. T. Rossby, 1978: Mediterranean water: An intense mesoscale eddy off the Bahamas. *Science*, **202**, 1085–1087, https://doi.org/10.1126/science.202.4372.1085.

McGillicuddy, D. J., Jr., 2016: Mechanisms of physical-biological-biogeochemical interaction at the oceanic mesoscale. *Annu. Rev. Mar. Sci.*, **8**, 125–159, https://doi.org/10.1146/annurev-marine-010814-015606.

McWilliams, J. C., 2016: Submesoscale currents in the ocean. *Proc. Roy. Soc.*, **472A**, 20160117, https://doi.org/10.1098/rspa.2016.0117.

Ménesguen, C., B. L. Hua, X. Carton, F. Klingelhoefer, P. Schnürle, and C. Reichert, 2012: Arms winding around a meddy seen in seismic reflection data close to the Morocco coastline. *Geophys. Res. Lett.*, **39**, L05604, https://doi.org/10.1029/2011GL050798.

Nandi, P., W. S. Holbrook, S. Pearse, P. Páramo, and R. W. Schmitt, 2004: Seismic reflection imaging of water mass boundaries in the Norwegian Sea. *Geophys. Res. Lett.*, **31**, L23311, https://doi.org/10.1029/2004GL021325.

Padhi, A., P. Mukhopadhyay, T. Blacic, W. Fortin, W. S. Holbrook, and S. Mallick, 2010: Prestack waveform inversion for the water-column velocity structure- the present state and the road ahead. *SEG Technical Program*, Denver, Colorado, SEG Library, 2845–2849, https://doi.org/10.1190/1.3513435.

Papenberg, C., D. Klaeschen, G. Krahmann, and R. W. Hobbs, 2010: Ocean temperature and salinity inverted from combined hydrographic and seismic data. *Geophys. Res. Lett.*, **37**, L04601, https://doi.org/10.1029/2009GL042115.

Pinheiro, L. M., H. Song, B. Ruddick, J. Dubert, I. Ambar, K. Mustafa, and R. Bezerra, 2010: Detailed 2-D imaging of the Mediterranean outflow and meddies off W Iberia from multichannel seismic data. *J. Mar. Syst.*, **79**, 89–100, https://doi.org/10.1016/j.jmarsys.2009.07.004.

Qi, J., C. Liu, J. Chi, D. Li, L. Gao, and B. Yin, 2022: An ensemble-based machine learning model for estimation of subsurface thermal structure in the South China Sea. *Remote Sens.*, **14**, 3207, https://doi.org/10.3390/rs14133207.

Quartly, G. D., and M. A. Srokosz, 2004: Eddies in the southern Mozambique Channel. *Deep-Sea Res. II*, **51**, 69–83, https://doi.org/10.1016/j.dsr2.2003.03.001.

Quentel, E., X. Carton, M.-A. Gutscher, and R. Hobbs, 2010: Detecting and characterizing mesoscale and submesoscale structures of Mediterranean water from joint seismic and hydrographic measurements in the Gulf of Cadiz. *Geophys. Res. Lett.*, **37**, L06604, https://doi.org/10.1029/2010GL042766.

Richardson, P. L., A. S. Bower, and W. Zenk, 2000: A census of Meddies tracked by floats. *Prog. Oceanogr.*, **45**, 209–250, https://doi.org/10.1016/S0079-6611(99)00053-1.

Roberts, M. J., J.-F. Ternon, and T. Morris, 2014: Interaction of dipole eddies with the western continental slope of the Mozambique Channel. *Deep-Sea Res. II*, **100**, 54–67, https://doi.org/10.1016/j.dsr2.2013.10.016.

Ruddick, B., H. Song, C. Dong, and L. Pinheiro, 2009: Water column seismic images as maps of temperature gradient. *Oceanography*, **22** (1), 192–205, https://doi.org/10.5670/oceanog.2009.19.

Sacchi, M. D., 1997: Reweighting strategies in seismic deconvolution. *Geophys. J. Int.*, **129**, 651–656, https://doi.org/10.1111/j.1365-246X.1997.tb04500.x.

Sallarès, V., B. Biescas, G. Buffett, R. Carbonell, J. J. Dañobeitia, and J. L. Pelegrí, 2009: Relative contribution of temperature and salinity to ocean acoustic reflectivity. *Geophys. Res. Lett.*, **36**, L00D06, https://doi.org/10.1029/2009GL040187.

Sheen, K. L., N. J. White, and R. W. Hobbs, 2009: Estimating mixing rates from seismic images of oceanic structure. *Geophys. Res. Lett.*, **36**, L00D04, https://doi.org/10.1029/2009GL040106.

——, ——, C. P. Caulfield, and R. W. Hobbs, 2012: Seismic imaging of a large horizontal vortex at abyssal depths beneath the

Sub-Antarctic front. *Nat. Geosci.*, **5**, 542–546, https://doi.org/10.1038/ngeo1502.

Siedler, G., L. Armi, and T. J. Müller, 2005: Meddies and decadal changes at the Azores front from 1980 to 2000. *Deep-Sea Res. II*, **52**, 583–604, https://doi.org/10.1016/j.dsr2.2004.12.010.

Song, H., L. M. Pinheiro, B. Ruddick, and F. C. Teixeira, 2011: Meddy, spiral arms, and mixing mechanisms viewed by seismic imaging in the Tagus Abyssal Plain (SW Iberia). *J. Mar. Res.*, **69**, 827–842, https://doi.org/10.1357/002224011799849309.

Sonnewald, M., R. Lguensat, D. C. Jones, P. D. Dueben, J. Brajard, and V. Balaji, 2021: Bridging observations, theory and numerical simulation of the ocean using machine learning. *Environ. Res. Lett.*, **16**, 073008, https://doi.org/10.1088/1748-9326/ac0eb0.

Srinivasan, K., J. C. McWilliams, M. J. Molemaker, and R. Barkan, 2019: Submesoscale vortical wakes in the lee of topography. *J. Phys. Oceanogr.*, **49**, 1949–1971, https://doi.org/10.1175/JPO-D-18-0042.1.

Strobach, E., P. Klein, A. Molod, A. A. Fahad, A. Trayanov, D. Menemenlis, and H. Torres, 2022: Local air-sea interactions at ocean mesoscale and submesoscale in a Western Boundary Current. *Geophys. Res. Lett.*, **49**, e2021GL097003, https://doi.org/10.1029/2021GL097003.

Su, Z., J. Wang, P. Klein, A. F. Thompson, and D. Menemenlis, 2018: Ocean submesoscales as a key component of the global heat budget. *Nat. Commun.*, **9**, 775, https://doi.org/10.1038/s41467-018-02983-w.

Swart, N. C., J. R. E. Lutjeharms, H. Ridderinkhof, and W. P. M. de Ruijter, 2010: Observed characteristics of Mozambique Channel eddies. *J. Geophys. Res.*, **115**, C09006, https://doi.org/10.1029/2009JC005875.

Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift, 2011: Indian Ocean. *Descriptive Physical Oceanography*, 6th ed. L. D. Talley et al., Eds., Academic Press, 363–399, https://doi.org/10.1016/B978-0-7506-4552-2.10011-3.

Tang, Q., C. Wang, D. Wang, and R. Pawlowicz, 2014: Seismic, satellite and site observations of internal solitary waves in the NE South China Sea. *Sci. Rep.*, **4**, 5374, https://doi.org/10.1038/srep05374.

——, R. Hobbs, C. Zheng, B. Biescas, and C. Caiado, 2016: Markov Chain Monte Carlo inversion of temperature and salinity structure of an internal solitary wave packet from marine seismic data. *J. Geophys. Res. Oceans*, **121**, 3692–3709, https://doi.org/10.1002/2016JC011810.

——, S. P. S. Gulick, J. Sun, L. Sun, and Z. Jing, 2020: Submesoscale features and turbulent mixing of an oblique anticyclonic eddy in the Gulf of Alaska investigated by marine seismic survey data. *J. Geophys. Res. Oceans*, **125**, e2019JC015393, https://doi.org/10.1029/2019JC015393.

Tedesco, P., J. Gula, P. Penven, C. Ménesguen, Q. Jamet, and C. Vic, 2024: Can mesoscale eddy kinetic energy sources and sinks be inferred from sea surface height in the Agulhas current region? *J. Geophys. Res. Oceans*, **129**, e2023JC020833, https://doi.org/10.1029/2023JC020833.

Ternon, J. F., M. J. Roberts, T. Morris, L. Hancke, and B. Backeberg, 2014: In situ measured current structures of the eddy field in the Mozambique Channel. *Deep-Sea Res. II*, **100**, 10–26, https://doi.org/10.1016/j.dsr2.2013.10.013.

Toole, J. M., and B. A. Warren, 1993: A hydrographic section across the subtropical South Indian Ocean. *Deep-Sea Res. I*, **40**, 1973–2019, https://doi.org/10.1016/0967-0637(93)90042-2.

Tsuji, T., T. Noguchi, H. Niino, T. Matsuoka, Y. Nakamura, H. Tokuyama, S. Kuramoto, and N. Bangs, 2005: Two-dimensional mapping of fine structures in the Kuroshio Current using seismic reflection data. *Geophys. Res. Lett.*, **32**, L14609, https://doi.org/10.1029/2005GL023095.

Uchida, T., and Coauthors, 2022: Cloud-based framework for inter-comparing submesoscale-permitting realistic ocean models. *Geosci. Model Dev.*, **15**, 5829–5856, https://doi.org/10.5194/gmd-15-5829-2022.

Ullgren, J. E., H. M. van Aken, H. Ridderinkhof, and W. P. M. de Ruijter, 2012: The hydrography of the Mozambique Channel from six years of continuous temperature, salinity, and velocity observations. *Deep-Sea Res. I*, **69**, 36–50, https://doi.org/10.1016/j.dsr.2012.07.003.

van Aken, H. M., H. Ridderinkhof, and W. P. de Ruijter, 2004: North Atlantic deep water in the South-western Indian Ocean. *Deep-Sea Res. I*, **51**, 755–776, https://doi.org/10.1016/j.dsr.2004.01.008.

Vic, C., G. Roullet, X. Capet, X. Carton, M. J. Molemaker, and J. Gula, 2015: Eddy-topography interactions and the fate of the Persian Gulf Outflow. *J. Geophys. Res. Oceans*, **120**, 6700–6717, https://doi.org/10.1002/2015JC011033.

Wang, P., and Coauthors, 2021: The GPU version of LASG/IAP Climate System Ocean Model version 3 (LICOM3) under the Heterogeneous-Compute Interface for Portability (HIP) framework and its large-scale application. *Geosci. Model Dev.*, **14**, 2781–2799, https://doi.org/10.5194/gmd-14-2781-2021.

Warner, M., 1990: Absolute reflection coefficients from deep seismic reflections. *Tectonophysics*, **173**, 15–23, https://doi.org/10.1016/0040-1951(90)90199-I.

Wei, J., Z. Zhao, K. L. Gunn, S. P. S. Gulick, D. J. Shillington, and C. M. Lowery, 2025: Unlocking a global ocean mixing dataset: Toward standardization of seismic-derived ocean mixing rates. *J. Atmos. Oceanic Technol.*, **42**, 675–697, https://doi.org/10.1175/JTECH-D-24-0076.1.

Wood, W. T., W. S. Holbrook, M. K. Sen, and P. L. Stoffa, 2008: Full waveform inversion of reflection seismic data for ocean temperature profiles. *Geophys. Res. Lett.*, **35**, L04608, https://doi.org/10.1029/2007GL032359.

Wyrtki, K., 1973: Physical oceanography of the Indian Ocean. *The Biology of the Indian Ocean*, B. Zeitzschel and S. A. Gerlach, Eds., Ecological Studies, Vol 3, Springer, 18–36, https://doi.org/10.1007/978-3-642-65468-8_3.

Xiao, W., and Z. Meng, 2022: Turbulent heat fluxes in a Mediterranean eddy quantified using seismic and hydrographic observations. *J. Mar. Sci. Eng.*, **10**, 720, https://doi.org/10.3390/jmse10060720.

——, K. L. Sheen, Q. Tang, J. Shutler, R. Hobbs, and T. Ehmen, 2021: Temperature and salinity inverted for a Mediterranean eddy captured with seismic data, using a spatially iterative Markov chain Monte Carlo approach. *Front. Mar. Sci.*, **8**, 734125, https://doi.org/10.3389/fmars.2021.734125.

Yang, S., H. Song, B. Coakley, K. Zhang, and W. Fan, 2022: A mesoscale eddy with submesoscale spiral bands observed from seismic reflection sections in the Northwind Basin, Arctic Ocean. *J. Geophys. Res. Oceans*, **127**, e2021JC017984, https://doi.org/10.1029/2021JC017984.

Yilmaz, Ö., 2001: *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data.* Society of Exploration Geophysicists, 2027 pp.

Zou, Z., P. Bakhtiari Rad, L. Macelloni, and L. Zhang, 2021: Temporal and spatial variations in three-dimensional seismic oceanography. *Ocean Sci.*, **17**, 1053–1066, https://doi.org/10.5194/os-17-1053-2021.