remote sensing

MDPI

Article

Reconstructing the Three-Dimensional Thermohaline Structure of Mesoscale Eddies in the South China Sea Using In Situ Measurements and Multi-Sensor Satellites

Zhiyuan Zhuang [1], Yanwei Zhang [1,*], Liuzhenyi Zhang [1], Weihan Ruan [1], Danni Lyu [1] and Jiancheng Yu [2]

[1] State Key Laboratory of Marine Geology, Tongji University, Shanghai 200092, China; zhuangzhiyuan@tongji.edu.cn (Z.Z.); 2111316@tongji.edu.cn (L.Z.); 2031680@tongji.edu.cn (W.R.); 2011469@tongji.edu.cn (D.L.)

[2] Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110169, China; yjc@sia.cn

* Correspondence: ywzhang@tongji.edu.cn; Tel.: +86-021-65981506

Abstract: The evolution of the three-dimensional thermohaline structure of mesoscale eddies is crucial for assessing energy and mass transfer during their long-distance propagation in the ocean. However, the understanding and quantitative evaluation of the role that mesoscale eddies play in driving variations of thermohaline in the deep sea remains constrained due to the scarcity of in situ observations, particularly in marginal seas such as the South China Sea (SCS). In this study, we propose an artificial intelligence (AI)–physics-based deep learning model that integrates satellite measurements and Argo data from 2003 to 2021 to reconstruct the three-dimensional thermohaline structure of mesoscale eddies in the SCS. Besides utilizing basic sea surface hydrodynamic parameters obtained from satellite data for model training, an additional branch incorporating eddy physical parameters was introduced to optimize the model. The results demonstrate that the model effectively reconstructs thermohaline properties within mesoscale eddies in the SCS. Compared to Argo observations, the average root mean square error (RMSE) for temperature (salinity) within anticyclonic eddies was 0.34 °C (0.036 PSU), while it was 0.36 °C (0.032 PSU) within cyclonic eddies in the upper 1500 m. Further validation using high-resolution glider observations tracking an anticyclonic eddy originating in the SCS confirms the model's efficiency, achieving an RMSE of 0.2962 °C (0.0138 PSU) for temperature (salinity). The accuracy of our proposed model significantly outperforms that of HYCOM and GLORYS simulations, with the RMSE reduced by 40% to 60%. The distinctive capabilities provide valuable insights into understanding the fine-scale structures of mesoscale eddies, especially in regions with limited in situ data.

Keywords: mesoscale eddies; deep learning model; thermohaline structure; satellite observations; Argo observations; glider measurements; the South China Sea

check for updates

Academic Editor: Mark Bourassa

Received: 28 October 2024
Revised: 16 December 2024
Accepted: 20 December 2024
Published: 25 December 2024

Citation: Zhuang, Z.; Zhang, Y.; Zhang, L.; Ruan, W.; Lyu, D.; Yu, J. Reconstructing the Three-Dimensional Thermohaline Structure of Mesoscale Eddies in the South China Sea Using In Situ Measurements and Multi-Sensor Satellites. *Remote Sens.* **2025**, *17*, 22. https://doi.org/10.3390/rs17010022

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mesoscale eddies, characterized by time scales varying from several days to hundreds of days [1,2], horizontal scales ranging from tens to hundreds of kilometers, and vertical extensions from tens to thousands of meters, are acknowledged as crucial for energy and mass transfer in the ocean [3–5]. The vertical process inside mesoscale eddies leads to geostrophic adjustments, such as depressions associated with anticyclonic eddies (AEs) and uplifts for cyclonic eddies (CEs), causing the exchange and mixing of the water mass between the sea surface and deep layers, thereby influencing the vertical stratification and

thermohaline structure [6,7]. The South China Sea (SCS), the largest marginal sea of the Northwest Pacific, is a region where mesoscale eddies are prevalent, influenced by the Kuroshio Current, Pacific intermediate and deep waters, monsoons, and typhoons [8,9]. On average, approximately 32.8 mesoscale eddies are generated annually in the SCS, significantly contributing to heat and mass transport [10,11]. Consequently, gaining insights into the three-dimensional (3D) structure of the mesoscale eddies is imperative for comprehending and quantitatively assessing their contributions to mass and energy transport within the SCS [12].

Underwater mooring buoys, Argo floats, moorings, and gliders, are commonly employed to measure thermohaline profiles in the ocean. For instance, an anticyclonic eddy in the SCS, with a radius of 150 km, was observed to extend to depths exceeding 2000 m through mooring deployments [4,13]. Additionally, Shu et al. [14] deployed 12 gliders as an observational network to investigate the fine subsurface structure of thermohaline properties and the temporal evolution of the eddies in the northern SCS. Argo observations are widely employed to study the composite 3D structure of mesoscale eddies. Sun et al. [15] conducted a synthetic study on the vertical structure of mesoscale eddies in the North Pacific Ocean, and the results show that the variation in physical parameters modulated by eddies is mainly restricted to above 800 m. Mesoscale eddies are 'bullet-shaped', with the upper part of the eddies having a large radius and the lower part having a small radius, and the vertical thickness is 800 m. Recent mooring observations have provided further validation that a considerable number of mesoscale eddies exert control over dynamic processes in the deep ocean, penetrating depths exceeding thousands of meters [16,17].

However, the state-of-the-art in situ observations are inadequate for investigating the vertical structural variability of eddies throughout their lifetimes. Previous studies primarily focused on case studies or provided a mean representation of the 3D structure of mesoscale eddies through composite analysis methods. Cumulatively, satellite remote sensing offers long-term data on various ocean surface variables, characterized by high spatial and temporal resolution. These datasets provide a comprehensive synoptic view of the surface expressions of mesoscale eddies, significantly advancing our comprehension of their global distribution, kinematic properties, and spatiotemporal dynamics. However, satellite observations alone are limited in their ability to capture the surface structure of mesoscale eddies. Given the strong nonlinear relationship and high correlation between surface and interior oceanic conditions [18], integrating subsurface measurements with satellite-derived surface data is essential for a thorough understanding of mesoscale eddies.

In recent years, increasing studies have focused on utilizing remote sensing observations with high temporal and spatial resolution to accurately map surface features to their corresponding subsurface layers. These approaches can be primarily categorized into three types: physics-driven dynamic methods, empirical statistical methods, and artificial intelligence (AI)-based techniques. The dynamic modeling approach involves applying a dynamic model to convert surface observations into control parameters for the model, thereby enabling the mapping of surface features to subsurface layers. For instance, a four-layer quasi-geostrophic model has been utilized to reconstruct the subsurface flow field, with surface flow represented by sea surface height (SSH) [19]. Additionally, the statistical dynamic approach, exemplified by the 'interior + surface quasi-geostrophy equation' (isQG), has also been used in an indirect way for 3D thermohaline reconstruction [20–22]. These dynamic methods are fundamentally grounded in simplified physical assumptions, which may limit their applicability in complex scenarios. In contrast, some studies have opted for empirical statistical models, particularly multiple linear regression (MLR). Pun et al. [23] proposed a new derivation method that increases the vertical resolution of the temperature structure to 26 isotherms. Jeong et al. [24] utilized MLR to reconstruct global

3D ocean temperature from SSTA, sea-level anomaly (SLA), and sea surface wind (SSW) stress anomaly data, using Argo data as label data.

However, these methods, which ignore the physical constraints and rely heavily on linear relationships, may encounter challenges in establishing robust statistical correlations, especially in regions with sparse observational data or where conditions undergo substantial variations. Recently, with the enhancement of remote sensing data resolution and spatial coverage, AI-based techniques have been widely applied to retrieve subsurface temperature and salinity fields from satellite observations [25–30]. Within the specific domain of mesoscale eddies, several studies have successfully utilized AI-based methods, such as convolutional neural networks (CNNs) [31], neural networks integrated with ResNet [32], and the U-net network [33], to invert subsurface temperature and salinity fields using satellite data. These studies indicate that surface observations can provide a reasonably accurate estimation of the hydrological parameters of the interior ocean. However, several challenges remain:

In contrast to the 3D vertical structure of thermohaline reconstructions, (1) the majority of studies on the reconstruction of mesoscale eddies are confined to the upper ocean. Meanwhile, there is a notable paucity of investigations specifically targeting the South China Sea (SCS) partly due to the sparsity of Argo data. (2) Existing studies overlook the constraints imposed by dynamic processes and are limited due to a lack of validation through supplementary in situ measurements.

In this study, we present a novel deep learning model to reconstruct the 3D thermohaline structure of mesoscale eddies in the SCS. Unlike traditional single-input models that rely solely on satellite data, a branch including eddy physical parameters was embedded into the model. This dual-input approach is expected to enhance the predictive capabilities. Furthermore, to validate the model, we conducted a case study focusing on an anticyclonic eddy measured by the glider.

This paper is structured as follows: Section 2 describes the study area, the data, and the preprocessing steps undertaken to prepare the datasets for the deep learning model. Section 3 details the methodology. Section 4 presents the specific results of the reconstruction and case study. Finally, Section 5 offers the discussion and conclusions.

2. Data

2.1. Study Area

Our study primarily focuses on the South China Sea (SCS), which features frequent mesoscale eddies originating from the Kuroshio in the Northwest Pacific. Considering the sparsity of Argo data and the diversity in the origin of mesoscale eddies in the SCS [34–36], we have expanded the scope of our study to include the Northwest Pacific Ocean (Figure 1a). This extended region exhibits various thermohaline properties, and previous studies have suggested that variations in background thermohaline fields significantly influence the 3D thermohaline structures within eddies [37–39]. Moreover, we incorporate Argo data between 2003 and 2021, which allows us to better assess the robustness of our model in a variety of oceanic environments.

2.2. Satellite Data

This study utilizes four remote sensing products: sea level anomaly (SLA), sea surface temperature (SST), sea surface salinity (SSS), and sea surface wind (SSW) data.

Sea level anomaly (SLA): the AVISO (Archiving, Validation, and Interpretation of Satellite Oceanographic) satellite altitude data provided by CLS-Collect Localization Satellites is used in this paper. The data are combined with observations from the Ocean Topography Experiment, including the Topex/Poseidon, Jason 1 and 2, ERS-1, and ERS-

2 satellites. All sea level heights are corrected by various methods, such as counter-pressure, electromagnetic bias, and tide, according to Mercator projection. The spatial resolution of the data is 0.25° × 0.25° across the global grid, and the time resolution is daily (SEALEVEL_GLO_PHY_L4_MY_008_047, https://doi.org/10.48670/moi-00148) (accessed on 22 March 2022).

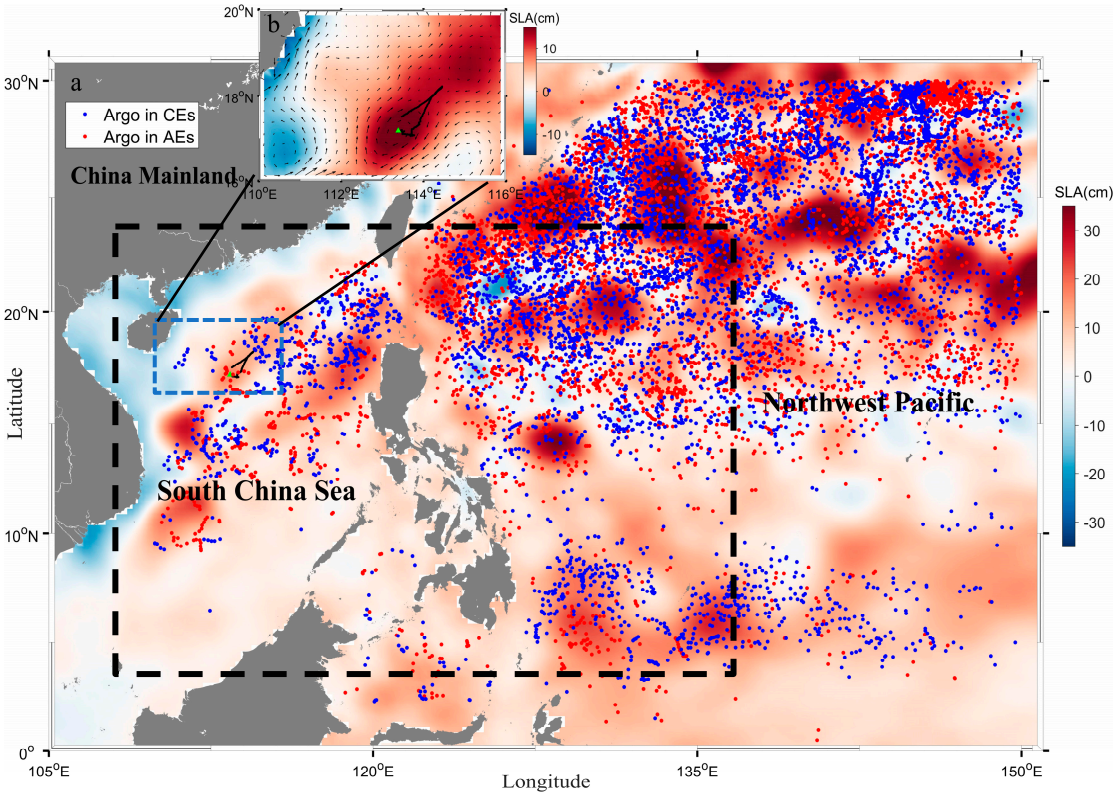

Figure 1. The spatial distribution of research datasets. (**a**) Argo profiles from 2003 to 2021 are used in the study region. Red dots indicate the Argo profiles within anticyclone eddies, while blue dots are the Argo profiles within cyclone eddies. The black dashed box shows the area of all the Argo profiles for validation. (**b**) Surface track of the glider. The map has combined SLA and surface geostrophic current velocity data distributed by AVISO (https://www.aviso.oceanobs.com) (accessed on 23 May 2021). The sampling trajectory of Glider j003 is represented by the black line, and the green triangle marks the initial position of Glider j003. The background map shows the sea level anomaly (SLA) on 8 July 2016.

Sea surface temperature (SST): the SST dataset used in this study is the NOAA Optimum Interpolation (OI) SST product, which provides daily data at 0.25° resolution [40]. The OISST dataset is derived from infrared satellite observations of the Advanced Very High-Resolution Radiometer (AVHRR), with supplemental information provided by in situ observations and proxy SSTs calculated from sea ice concentrations (https://www.ncei.noaa.gov/products/optimum-interpolation-sst) (accessed on 22 March 2022).

Sea surface salinity (SSS): the SSS product is available with a spatial resolution of 0.25° and is provided by the Copernicus Marine Environment Monitoring Service (CMEMS). It is derived from a multivariate optimal interpolation algorithm that combines SSS images from various satellite sources, including the Soil Moisture Active Passive (SMAP) satellite from NASA and the Soil Moisture and Ocean Salinity (SMOS) satellite from the European Space Agency (ESA). Additionally, the algorithm incorporates in situ salinity measurements and satellite-derived sea surface temperature (SST) data to enhance the accuracy and reliability

(MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013, https://doi.org/10.48670/moi-00051) (accessed on 2 March 2024).

Sea surface wind (SSW): the SSW data are sourced from the Cross-Calibrated Multi-Platform (CCMP) wind dataset, which was produced using satellite, moored buoy, and numerical models provided by Remote Sensing Systems (RSSs) [41]. The dataset offers a temporal resolution of 6 h and a spatial resolution of 0.25° (available at www.remss.com, https://doi.org/10.56236/RSS-uv6h30) (accessed on 2 March 2024).

The SSTA, SSSA, SSWUA, and SSWVA data used in this study represent anomalies of SST, SSS, SSWU, and SSWV. The SSTA and SSSA data are calculated through the temporal and spatial filters, with the temporal filter retaining signals with periods over 7 days and the spatial filter using an average Hanning window to maintain spatial scales smaller than 600 km. The SSWUA and SSWVA data are extracted from SSW dataset following the same procedure.

2.3. Mesoscale Eddy Dataset and Physics Parameters

The mesoscale eddies dataset was extracted from the altimetric Mesoscale Eddy Trajectories Atlas (META3.2 DT all-sat version), which was produced by SSALTO/DUACS and distributed by AVISO+ (https://aviso.altimetry.fr) with support from CNES, in collaboration with IMEDEA (https://doi.org/10.24400/527896/a01-2022.005.220209) (accessed on 8 December 2024) [42]. The dataset covers from 2003 to 2021. Eddy radius (Rad), amplitude (Amp), vorticity (Vor), rotational speed (U), and kinetic energy (EKE) are determined to characterize the fundamental dynamics of eddies [43,44]. The U and EKE can be calculated as

$$\mathrm{U} = \sqrt{{U_g}^2 + {V_g}^2} \tag{1}$$

$$EKE = \frac{1}{2}\left({U_g}^2 + {V_g}^2\right) \tag{2}$$

where U_g and V_g are the east–west and north–south geostrophic velocity anomaly components, respectively.

2.4. In Situ Data

2.4.1. Argo Data

Thermohaline structures based on Argo data are used as the ground truth dataset to establish the reconstruction model. This dataset, provided by the French Research Institute for the Exploitation of the Seas (IFREMER), encompasses a global scope. As shown in Figure 1a, we downloaded Argo data spanning from 2003 to 2021. The profiles from 2003 to 2019 were utilized to develop the neural network training model, while the profiles from 2020 and 2021 were employed to validate the accuracy of the reconstruction method. The red dots in Figure 1a represent Argo profiles within anticyclonic eddies (the specified quantity is 10,656), while the blue dots denote those within cyclonic eddies (the specified quantity is 10,196).

A quality control process was conducted as follows:

(a) The minimum depth observed by Argo should be larger than 10 m, and each Argo profile must contain an effective shallow observation point between 0 and 50 m and a deep observation point below 1000 m;
(b) The minimum number of effective observation samples between 10 and 1500 m should be larger than 30;
(c) The temperature/salinity profiles are interpolated on the standard depth of WOA13.

2.4.2. Glider Data

The glider observations of an anticyclonic eddy in the South China Sea, conducted from 3 to 16 July 2016, were employed to evaluate the performance of the reconstruction model. This glider was equipped with a Slocum Glider Payload CTD sensor produced by Sea-Bird Electronic (SBE) in the Bellevue, WA, USA, which recorded conductivity (salinity), temperature, and depth from the sea surface to a 1000 m water depth with a vertical resolution of ~0.6 m at a 6 Hz interval. The observation accuracies of the CTD are ± 0.0003 S m^{-1} for electric conductivity, ± 0.002 °C for temperature, and $\pm 0.1\%$ for pressure, respectively. Each dive took about 3 h, and it ascended to the surface to fix its position with a GPS module via satellite communication. The glider's surface trajectory during the observation period is shown in Figure 1b.

2.4.3. WOA Data

We use the World Ocean Atlas 2013 (WOA13) climatological data from the U.S. National Oceanographic Data Center (NODC) as the background parameter to calculate anomalies in temperature and salinity induced by eddies according to Equations (1) and (2). The temporal resolution is monthly averages, and the horizontal spatial resolution is $0.25° \times 0.25°$.

$$T_{anomaly}(x, y, z) = T_{Argo}(x, y, z) - T_{WOA}(x, y, z) \quad (3)$$

$$S_{anomaly}(x, y, z) = S_{Argo}(x, y, z) - S_{WOA}(x, y, z) \quad (4)$$

where (x, y, z) are the dimensions of longitude, latitude, and depth, respectively.

2.5. HYCOM and GLORYS Data

Model data have the advantage of continuity both in time and in space, which could provide a supplementary way to analyze eddies' detailed structure. Here, we utilize HYCOM and GLORYS to further validate the accuracy of the 3D eddy structures inverted by the reconstruction model.

2.5.1. HYCOM

The Hybrid Coordinate Ocean Model (HYCOM) offers reasonably accurate simulations of three-dimensional dynamic parameters of ocean motions based on assimilating available satellite altimeter observations and satellite and in situ sea surface temperature as well as in situ vertical temperature and salinity profiles from XBTs, Argo floats, and moored buoys. It has been widely employed to investigate the detailed structure and temporal–spatial evolution of mesoscale eddies [45,46] (https://tds.hycom.org/thredds/catalog.html, accessed on 2 March 2024).

2.5.2. GLORYS

Global Ocean Reanalysis and Simulation (GLORYS) was generated using the 'Nucleus for European Modelling of the Ocean' (NEMO) ocean model and assimilated satellite altimetry data, sea surface temperature data, and temperature and salinity data from ARGO and CTD data [47]. It has a 1/12° horizontal resolution and 50 vertical levels, representing a significant ability to resolve mesoscale variability [48] (GLOBAL_MULTIYEAR_PHY_001_030, https://doi.org/10.48670/moi-00021, accessed on 8 December 2024).

3. Methodology

The reconstruction model comprises three branches: the first one is a convolutional neural network (CNN) with an attention mechanism (Figure 2a), the second one incorpo-

rates prior knowledge (Figure 2b), and the third one is a Back Propagation Neural Network (BPNN) (Figure 2c).

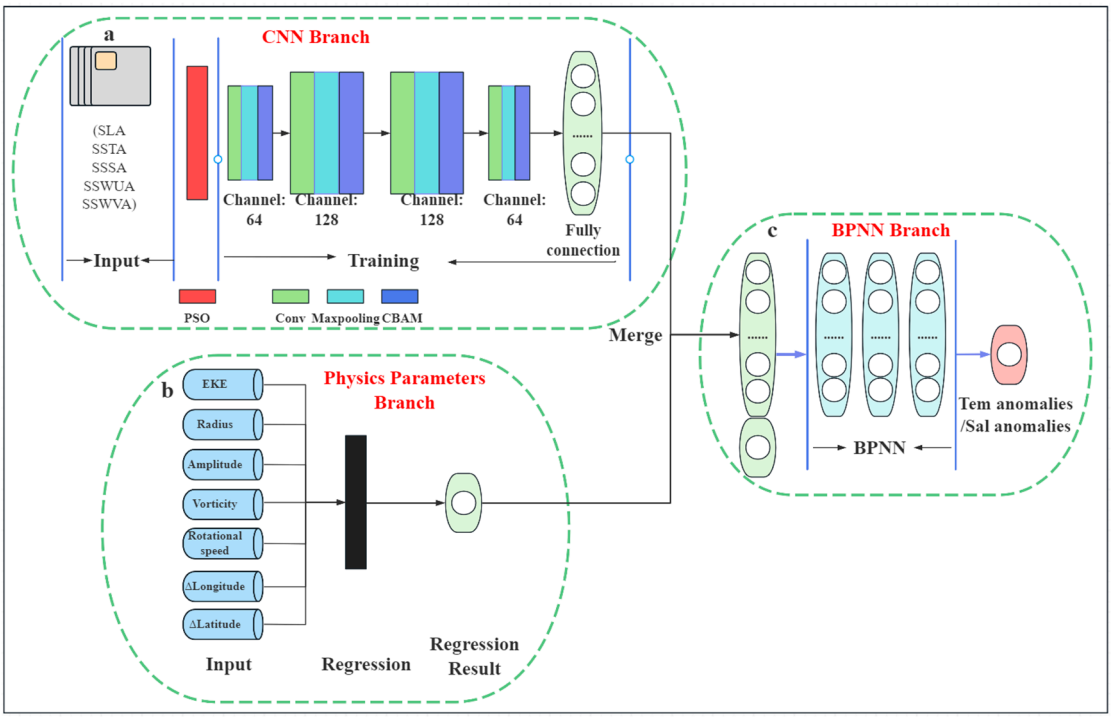

Figure 2. The schematic of the reconstruction model. (**a**) The convolutional neural network (CNN) branch: this branch integrates five types of satellite data as inputs, which are then processed through convolutional layers (Conv), feature Maxpooling layers (Maxpooling), and the Convolutional Block Attention Module (CBAM) during model training. The Particle Swarm Optimization (PSO) algorithm is employed to fine-tune the optimal channels of the CNN at various layers. (**b**) The physics parameters branch: this branch incorporates seven distinct physical parameters associated with the eddies into the reconstruction model using a regression algorithm. (**c**) The Backpropagation Neural Network (BPNN) branch: this branch merges the outputs from the CNN and physics parameters branches as inputs. It then processes these inputs through three hidden layers, utilizing a backpropagation mechanism to generate predictions for subsurface temperature and salinity anomalies.

3.1. CNN Branch

The conventional CNN network consists of a single input layer, training convolutional layers (Conv), feature pooling layers (Maxpooling), normalization layers, a feature connection layer, and an output layer [49]. In this study, we have enhanced this architecture by additionally incorporating an attention mechanism with input from the SLA, SSTA, SSSA, SSWUA, and SSWVA. Prior to embedding the inputs into the network, the Particle Swarm Optimization (PSO) algorithm, proposed by Kennedy and Eberhart [50,51], was employed to fine-tune the parameters and select the optimal output channels for the convolutional layers. The final set of the number of output channels is 64, 128, 128, and 64, with a filter size of 5 × 5, a stride of 1, and zero padding for all convolutional layers. Subsequently, a Batch Normalization (BN) layer stabilizes training by normalizing activations in each channel. The Maxpooling layers ensure that the outputs are invariant and less sensitive to the small translations in the input. Finally, the outputs of the Maxpooling layers then undergo the Convolutional Block Attention Module (CBAM), which selectively focuses on important features and can suppress unimportant ones [52]. The CBAM consists of a channel attention module (CAM) and spatial attention module (SAM) to make the model more capable of extracting features in both channel and spatial dimensions.

3.2. Physics Prior Knowledge-Embedded Branch

Polynomial regression [53] for retrieving three-dimensional thermohaline fields from sea surface information is directly employed in the physics-informed branch to calibrate and determine the weights of the various physical parameters. The physical parameters can be described as the dataset $X = (EKE, Vor, Amp, Rad, U)$. The inversion formula is as follows:

$$\begin{aligned} T_{i,k}(EKE, Vor, \quad & Amp, Rad, U) \\ & = a_{i,k}^{T_1}\left(EKE - \overline{EKE_{i,1}}\right) + a_{i,k}^{T_2}\left(Vor - \overline{Vor_{i,1}}\right) \\ & + a_{i,k}^{T_3}\left(Amp - \overline{Amp_{i,1}}\right) + a_{i,k}^{T_4}\left(Rad - \overline{Rad_{i,1}}\right) \\ & + a_{i,k}^{T_5}\left(U - \overline{U_{i,1}}\right) \end{aligned} \tag{5}$$

$$b_{i,j} = exp\left[-\left[\frac{x_i - x_j}{\Delta Lon_i}\right]^2 - \left[\frac{y_i - y_j}{\Delta Lat_i}\right]^2\right] \tag{6}$$

$$a_{i,k}^{T_1} = \sum_{j=1}^{N} b_{i,j}\left(T_{j,k}^{o} - \overline{T_{i,k}}\right)\left(EKE - \overline{EKE_{i,1}}\right) / \sum_{j=1}^{N} b_{i,j}\left(EKE - \overline{EKE_{i,1}}\right)^2 \tag{7}$$

where $T_{i,k}(EKE, Vor, Amp, Rad, U)$ are the reconstructed temperature anomalies at reconstructed point i and depth k, $b_{i,j}$ is the local correlation coefficient between points to be reconstructed and the measured data, i represents the points to be reconstructed, j denotes the points of measured data, x_i, x_j, y_i, y_j are longitudes and latitudes, ΔLon_i, ΔLat_i represent the zonal distance and the meridional distance between the Argo profile and the eddy center, and $a_{i,k}^{T_1}$, $a_{i,k}^{T_2}$, $a_{i,k}^{T_3}$, $a_{i,k}^{T_4}$, and $a_{i,k}^{T_5}$ are regression coefficients; they can be calculated by Equation (5). $T_{j,k}^{o}$ is the measured data, and $\overline{T_{i,k}}$ is the weighted average temperature at point i and depth k. The formula for $\overline{T_{i,k}}$ is as follows:

$$\overline{T_{i,k}} = \sum_{j=1}^{N} b_{i,j} T_{j,k}^{o} / \sum_{j=1}^{N} b_{i,j} \tag{8}$$

$\overline{EKE_{i,1}}$, $\overline{Vor_{i,1}}$, $\overline{Amp_{i,1}}$, $\overline{Rad_{i,1}}$, and $\overline{U_{i,1}}$ are weighted mean of the product at the ocean surface and can be calculated similarly to $\overline{T_{i,k}}$. The methodology for the calculation of salinity anomalies is analogous to that for the calculation of temperature anomalies.

3.3. BPNN Branch

The outputs from the aforementioned branches are concatenated and fed into the final output branch, which utilizes a Back Propagation Neural Network (BPNN) [54] as its core architecture. This branch consists of one fully connected (FC) layer and three hidden layers with (128, 256) neurons, respectively, to generate a scalar value that represents either the subsurface temperature or salinity (Figure 2c).

4. Results

To evaluate the accuracy and robustness of the reconstruction model within the mesoscale eddies, a specific anticyclonic eddy tracked by a glider is selected for analysis. This anticyclonic eddy was initially detected from altimeter data and originated in the SCS on 1 June 2016, at 16.75°N and 114.76°E, and disappeared on 22 July 2016, at 17.38°N and 113.17°E. It lasted for 52 days, moving northwestward by approximately 200 km with an average radius of 74 km and a propagating speed of 0.06 m/s (Figure 3a). The glider captured the eddy on July 3, 2016, and subsequently conducted observations around it. The track of the glider observation is shown in Figure 3a (black dots), with the SLA on 8 July, and the thin black line indicating the eddy boundary. The topographic depth of the eddy center and the pronounced temporal variability of U, Rad, Amp, EKE, and EI (eddy intensity, which also represents energy density, is the ratio of the mean EKE and the

area of AE and can be calculated using Formula (9)) are displayed in Figure 3b–d. In the initial phase of the eddy, these properties exhibited relatively low strength and magnitude. Following a brief decline, they showed a sustained increase. It is also noteworthy that the dynamic characteristics of the eddy exhibited relatively high values and experienced a decay period during the glider observation, providing a diverse range for validation.

$$EI = \frac{\overline{EKE}}{A} = \frac{\overline{EKE}}{\pi R^2} \tag{9}$$

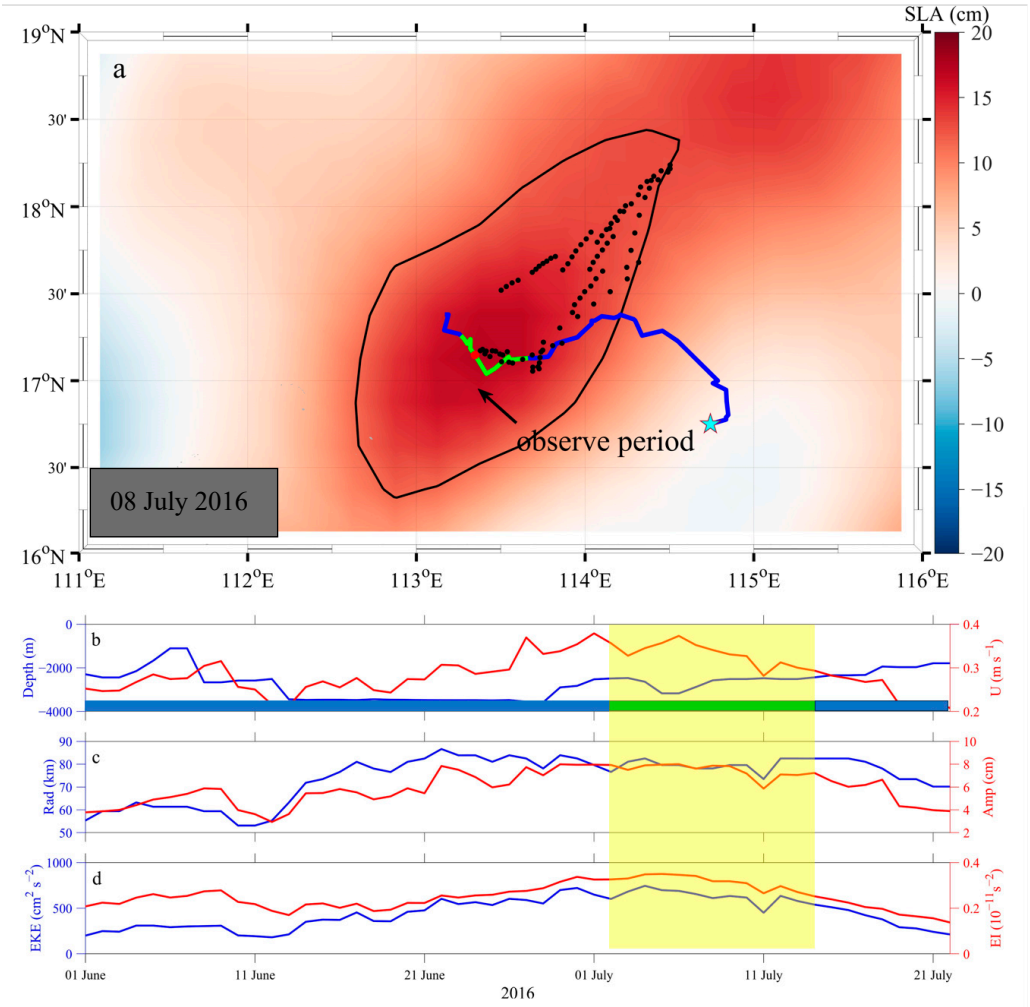

Figure 3. Surface track of the anticyclonic eddy and time-series of various variables. (**a**) The thick line represents the trajectory of the eddy center, with the green thick line indicating the trajectory during the observation period. The asterisk stands for the birth site of the eddy. The survey conducted by the glider took place from 3 to 16 July 2016. The black thin line represents the eddy edge on 8 July 2016, with the eddy center located at 17.07°N and 113.465°E (the red dot). The black dots are the observation sites of the glider. The background map is SLA for July 8. Several dynamic parameters are given: (**b**) topographic depth (blue line) and rotational speed (U, red line); (**c**) radius (Rad, blue line) and amplitude (Amp, red line); (**d**) eddy kinetic energy (EKE, blue line); and eddy intensity (EI, red line). The yellow shading indicates the period during which the glider captured the eddy (from 3 to 16 July 2016).

4.1. Validation of the Reconstruction Model Using Glider Observations

As illustrated in Figure 4, we reconstruct the temperature and the salinity of the eddy and assess the residual errors between the reconstruction model and the glider observations. Figure 4a–d present the results of temperature anomalies, while Figure 4e–h display the

corresponding salinity anomalies. Figure 4a shows the SLA at the center of the eddy and the distance from the eddy center to the observation sites of the glider, which is larger than the average radius over the entire period. Figure 4b,c depict the subsurface temperature anomalies from the glider observations and reconstruction model, respectively. Positive temperature anomalies induced by AEs primarily extend to 300 m, with the maximum primarily concentrated between 100 and 200 m. The subsurface salinity exhibits negative anomalies predominantly within the upper 400 m, with maxima anomaly signals extending to 600 m (Figure 4f,g). The comparison reveals a strong coupling between the model-reconstructed thermohaline structure of the anticyclonic eddy and its evolution with the measured data. The three dashed lines in Figure 4 represent three specific dates that the glider sampled the eddy center and eddy edge, respectively. These anomalies increase with proximity to the eddy's center, consistent with the expected diffusion of thermal anomalies from the eddy core to the surrounding regions [55]. Figure 4d presents the residual errors (peaking at 1 °C) between the reconstructed subsurface temperature and the glider observations within AEs. Similarly, Figure 4h presents the residual errors of subsurface salinity within AEs, reaching a maximum of 0.2 PSU and primarily concentrated in the upper ocean layers.

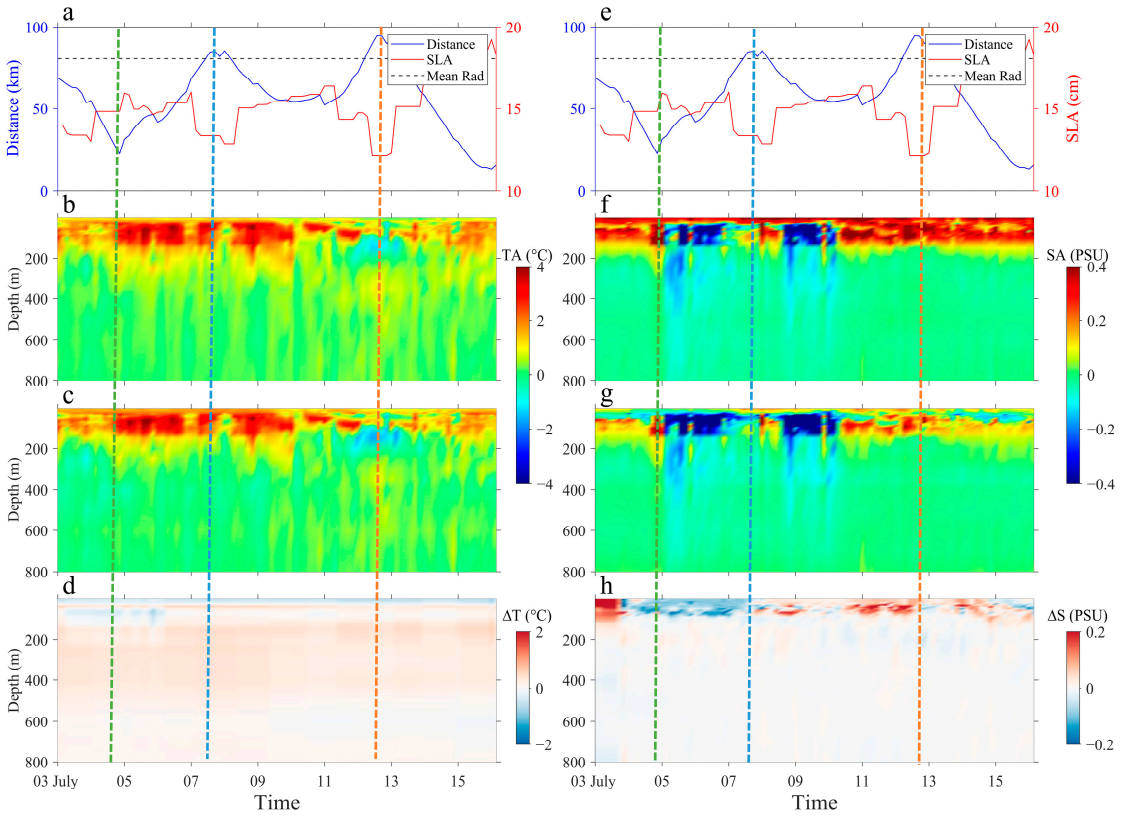

Figure 4. Comparison of subsurface temperature and salinity anomalies between glider observations and reconstruction model. Panels (**a**,**e**) depict the SLA and the distance between the eddy center and the glider observation station. Panels (**b**,**c**) show the temperature anomalies from the glider observations and the reconstruction model, respectively. Panel (**d**) illustrates the residual error, calculated as the difference between the glider observations and the data inverted by the reconstruction model. Panels (**f**–**h**) are the same as panels (**a**–**d**), respectively, but for the comparison of salinity anomalies. The three dashed lines represent three specific dates that the glider sampled at the eddy center and eddy edge, respectively.

Figure 5 shows the scatterplots of the temperature and salinity estimated from the reconstruction model ($T_{Reconstruct}/S_{Reconstruct}$) and observed by the glider (T_{Glider}/S_{Glider}).

Most of the dots lie in the 95% confidence interval, and there is a significant correlation of 0.9953 between $T_{Reconstruct}$ and T_{Glider} and 0.9493 between $S_{Reconstruct}$ and S_{Glider}, signifying a high level of consistency between these two datasets.

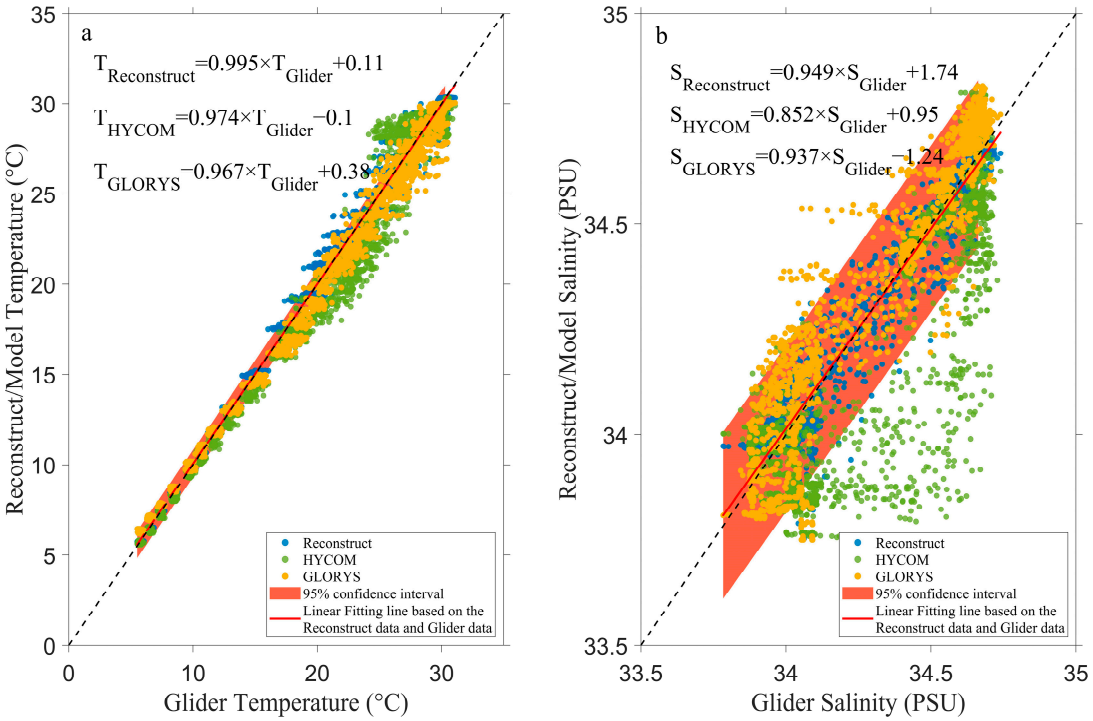

Figure 5. Linear fitting between glider data and reconstructed data (red line) with 95% confidence interval (pink shading). The green, orange, and blue dots represent the HYCOM, the GLOYRS, and the reconstructed results, respectively. (**a**) Temperature; (**b**) salinity.

We select three distinct profiles influenced by weak (center of the eddy) and strong (edge of the eddy) eddy effects to further compare the glider observation and the dataset reconstructed by the deep learning model and to focus on local variability (Figure 6a–c). The vertical trends of temperature and salinity from the reconstructed data generally exhibit a mirrored pattern with those obtained from the glider measurements (T in Figure 6a1,b1,c1 and S in Figure 6a4,b4,c4). The positive discrepancy in temperature ($T_{Glider} - T_{Reconstruct}$) mainly occurs above ~200 m (Figure 6a2,b2,c2). The discrepancy in temperature increases significantly with a slightly enlarged negative discrepancy recognized in the intermediate layers between 200 and 600 m when the eddy effect is strong (Figure 6c2). This phenomenon may be attributed to the potential missed detection of local temperature variability due to the glider's drifting motion during descent and ascent. A similar trend can also be observed regarding the salinity discrepancy ($S_{Glider} - S_{Reconstruct}$) (Figure 6a5,b5,c5). The vertical gradients show a higher consistency between the reconstructed dataset and the glider measurements, particularly in the upper mixed layer as well as deeper layers below 200 m (T (Figure 6a3,b3,c3), S (Figure 6a6,b6,c6)).

In conclusion, the above statistical and comparative analyses demonstrate a close alignment between the reconstruction results and glider measurements, offering convincing evidence for the utility of the reconstruction model in investigating the vertical thermohaline structure of mesoscale eddies. Furthermore, a detailed examination of the profiles across multiple locations intuitively reveals the influence of mesoscale eddy edge effects on the accuracy of the reconstruction model.

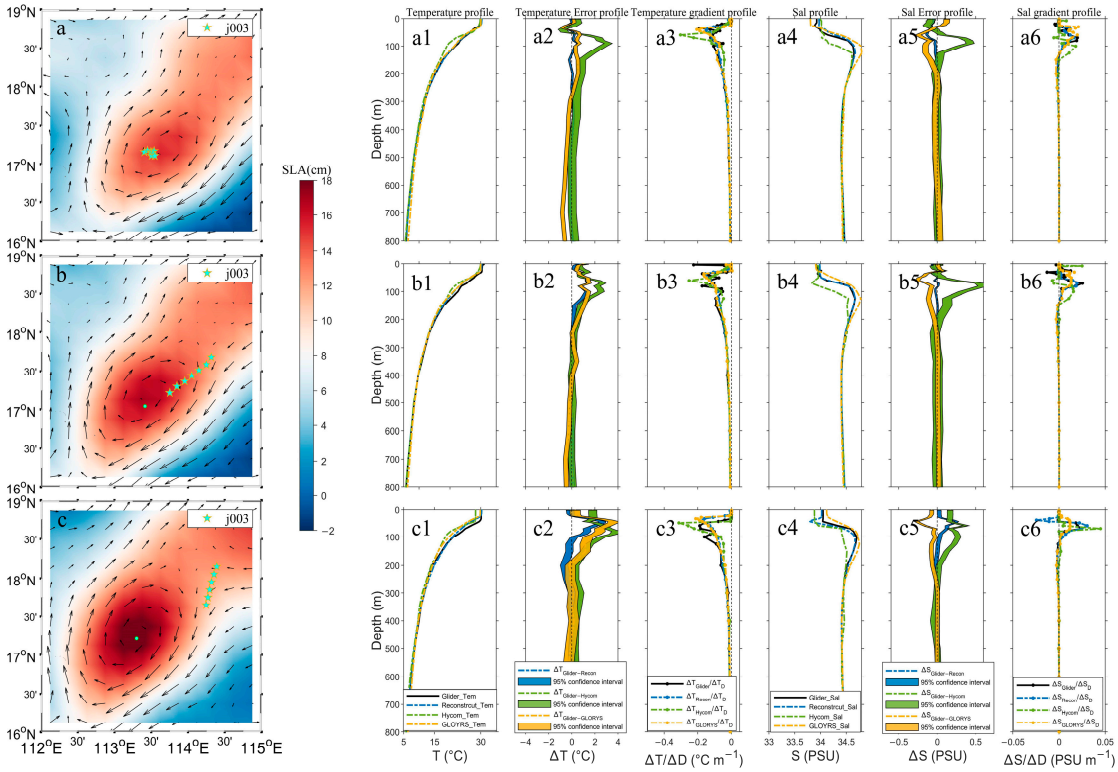

Figure 6. Detailed examination of glider observations, model outputs, and the reconstructed results. (**a**) The location of the glider (green stars) was colored by SLA on 4 July 2016. The arrows represent the ocean current field. Panels (**b**,**c**) present the same as panel (**a**), but for 7 July 2016, and 11 July 2016, respectively. Panels (**a1**,**b1**,**c1**) display the glider temperature (black line), reconstructed temperature (blue line), HYCOM temperature (green line), and GLOYRS temperature (orange line); panels (**a2**,**b2**,**c2**) illustrate the differences between the temperature profiles along with their 95% confidence interval where the blue intervals, green intervals, and orange intervals represent the differences between the glider data and reconstructed temperature, HYCOM temperature, and GLOYRS temperature, respectively. Panels (**a3**,**b3**,**c3**) show temperature gradients. Panels (**a4**,**b4**,**c4**), (**a5**,**b5**,**c5**), and (**a6**,**b6**,**c6**) are the same as panels (**a1**,**b1**,**c1**), (**a2**,**b2**,**c2**), and (**a3**,**b3**,**c3**), respectively, but for salinity.

4.2. Comparison Between HYCOM, GLORYS, and the Reconstruction Model

The scatter diagram in Figure 5 shows that the correlation between T_{HYCOM} and T_{Glider} (T_{GLORYS} and T_{Glider}) is 0.974 (0.967); the correlation between S_{HYCOM} and S_{Glider} (S_{GLORYS} and S_{Glider}) is 0.852 (0.937). Both correlations are lower than those between the reconstruction model and the glider. Furthermore, as depicted in Figure 6 from the profile-by-profile comparisons, the vertical trend of the reconstruction data exhibits better agreement with the glider data compared to the HYCOM simulation and GLORYS reanalysis data, particularly in the water depth range of 0–200 m. The width of the confidence intervals for the reconstruction data is significantly narrower than that of the HYCOM and GLORYS, suggesting higher accuracy and stability of the reconstruction model.

The accuracy of the reconstruction model is further validated by comparing the $RMSE_{HYCOM_glider}$ (the root mean square error (RMSE) between HYCOM and glider data), $RMSE_{GLORYS_glider}$ (the RMSE between GLORYS and glider data), and $RMSE_{reconstruct_glider}$ (the RMSE between reconstruction data and glider data). As depicted in Figure 7, $RMSE_{reconstruct_glider}$ varies with water depth for both temperature (Figure 7a) and salinity (Figure 7b). Within the upper layers (0–200 m), minimal RMSE values are observed at the top layer. It subsequently increases with water depth, reaching a peak at approximately 50 m before gradually declining. Below 200 m, the RMSE decreases generally as depth increases. A similar trend can be seen in the curve of $RMSE_{HYCOM_glider}$. However, we found

that the $RMSE_{reconstruct_glider}$ decreases across all layers by comparing $RMSE_{HYCOM_glider}$ and $RMSE_{GLORYS_glider}$, particularly at depths shallower than 200 m. This substantial decrease indicates a marked improvement in prediction accuracy in the shallower regions when using the reconstruction model. The analysis of RMSE not only validates the accuracy of the reconstruction model but also highlights its superiority over the HYCOM simulation and GLORYS in predicting temperature and salinity, especially in the upper layers of the water column. These findings suggest that the reconstruction of the three-dimensional structure of the mesoscale eddy in this study is highly reliable.

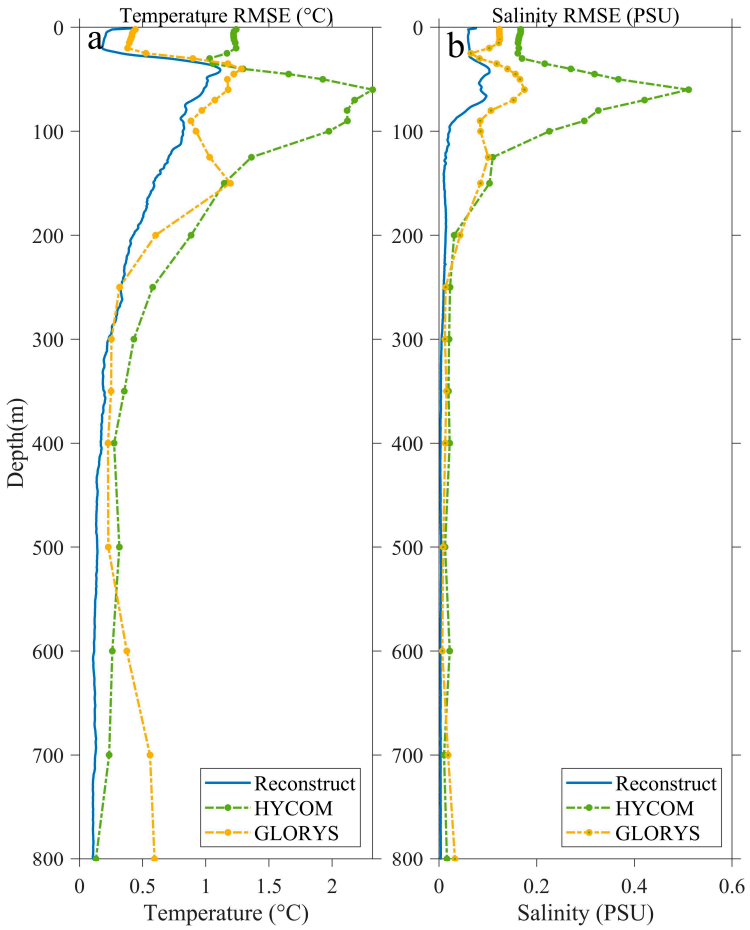

Figure 7. The comparison of the RMSE between reconstructed results, HYCOM simulation, GLORYS reanalysis data, and glider observations. (**a**) The root mean square error (RMSE) of temperature. (**b**) The RMSE of salinity. The blue lines represent the RMSE between the data from the reconstruction model and the glider observations, the green lines represent the RMSE between the HYCOM simulation and the glider observations, and the orange lines represent the RMSE between the GLORYS reanalysis and the glider observations.

4.3. Evaluating Three Settings of the Model

It is noted that a substantial volume of Argo data from the Northwest Pacific Ocean was incorporated to train the 3D mesoscale eddy thermohaline model for the SCS. Considering the differences in geographical environment and water mass characteristics between the Northwest Pacific Ocean and the SCS, it is crucial to assess the feasibility of directly utilizing Argo data from the Northwest Pacific Ocean into training the model for the SCS. Thus, we trained three models using distinct datasets for validation:

(1) Model1: training the model based on the Argo profiles in SCS and extended area; the amount of data points in the dataset is 20,852;

(2) Model2: training the model based on the Argo profiles in the Northwest Pacific ocean; the amount of data points in the dataset is 12,202.

(3) Model3: training the model based on the Argo profiles in SCS (the area marked by the black box in Figure 1); the amount of data points in the dataset is 8650.

We applied the three models to reconstruct the temperature and salinity anomalies (TAs and SAs) within the anticyclonic eddy observed by the glider in the SCS. The results are presented in Figure 8, where panels (a,e) depict the TAs and SAs observed by the glider, respectively. Panels (b–d) and (f–h) represent the residual error, calculated from glider observations (Glider_TA and Glider_SA) minus the data reconstructed by three models, respectively. Notably, Model1, trained using the combined dataset from the SCS and the extended area of Northwest Pacific Ocean (10,656 Argo profiles within AEs), produces the most accurate temperature and salinity anomalies. The residual errors between reconstructed anomalies and the observations are minimal and primarily concentrated in the upper ocean (0–200 m), where the eddy-induced anomalies are most significant. Model2, trained using Argo profiles from the broader Northwest Pacific, demonstrates improved performance compared to Model3, which was trained exclusively on Argo profiles within the SCS. The residual errors in anomalies based on Model2 are smaller than those of Model3 but larger than those of Model1. The reconstructed anomalies capture the general vertical structure of the observed eddy but exhibit noticeable overestimation, particularly in the upper ocean. This overestimation is more pronounced in the salinity anomalies (Figure 8g), where positive biases are evident in the upper 100 m. Model3 exhibits the weakest performance among the three models. The residual errors closely resemble the observations (Figure 8d,h), indicating the reconstructed anomalies are overly smoothed and linear, and the model underestimates the magnitude of the anomalies. This may be attributed to the sparsity of the dataset in the SCS. Specifically, while the total dataset contains 8650 profiles, only 4420 profiles are associated with anticyclonic eddies, limiting the model's ability to capture the full variability of mesoscale eddy-induced anomalies.

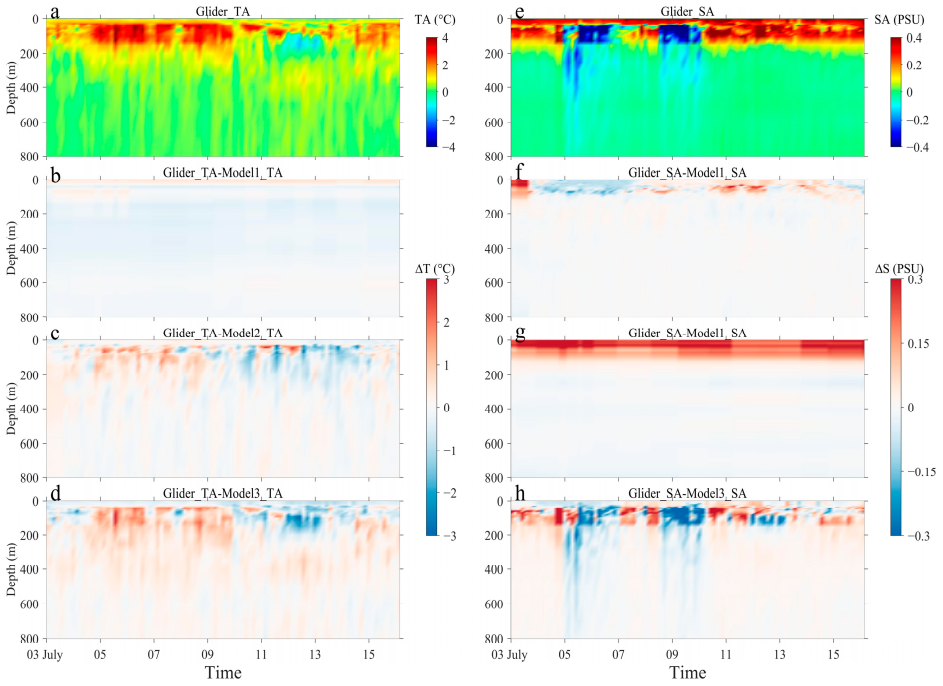

Figure 8. Comparison of the subsurface temperature and salinity anomalies reconstructed by three settings of the model. Panels (**a**,**e**): the temperature and salinity anomalies observed by the glider. Panels (**b**–**d**): the residual error calculated from Glider_TA minus the data reconstructed by the three models. Panels (**f**–**h**) are same as panels (**b**–**d**) but focus on the comparison of salinity anomalies.

4.4. Reconstruction of Three-Dimensional Thermohaline Structure of the Eddy

The vertical three-dimensional structure of the eddy is reconstructed by reconstructing the subsurface temperature and salinity for each point within a 61 × 61 grid centered on the eddy. Temperature and salinity anomalies induced by the eddy are visualized at depths of 50 m, 100 m, 300 m, 500 m, and 800 m for visualization. Figure 9 provides detailed depictions of the 3D temperature and salinity anomalies' structures within the AE, offering insights into both horizontal and vertical distributions. According to the results from the deep learning model, the anomalies induced by the AE primarily concentrated within the upper 400 m (Figure 9a,d). Horizontally, the anomaly contour distribution displays an elliptical shape with its long axis oriented in the northeast–southwest direction. Vertically, the temperature anomalies weaken with increasing depth above 400 m with the anomaly intensity decreasing radially from the eddy center toward the periphery. Below 400 m, the temperature anomaly intensity stabilizes, showing relatively smaller vertical variations. In the deeper layer, the horizontal anomaly distribution exhibits a distinct pattern: the temperature anomalies increase radially outward from the eddy center, in contrast to the weakening observed in the upper layers. This inversion may reflect the eddy's interaction with surrounding water masses and its impact on deeper ocean stratification. For further comparison, the 3D temperature and salinity anomaly structures obtained from the HYCOM (Figure 9b,e) and GLORYS (Figure 9c,f) are also presented. Specifically, both the HYCOM and GLORYS models tend to over-smooth the horizontal anomaly gradients, particularly near the center of the eddy.

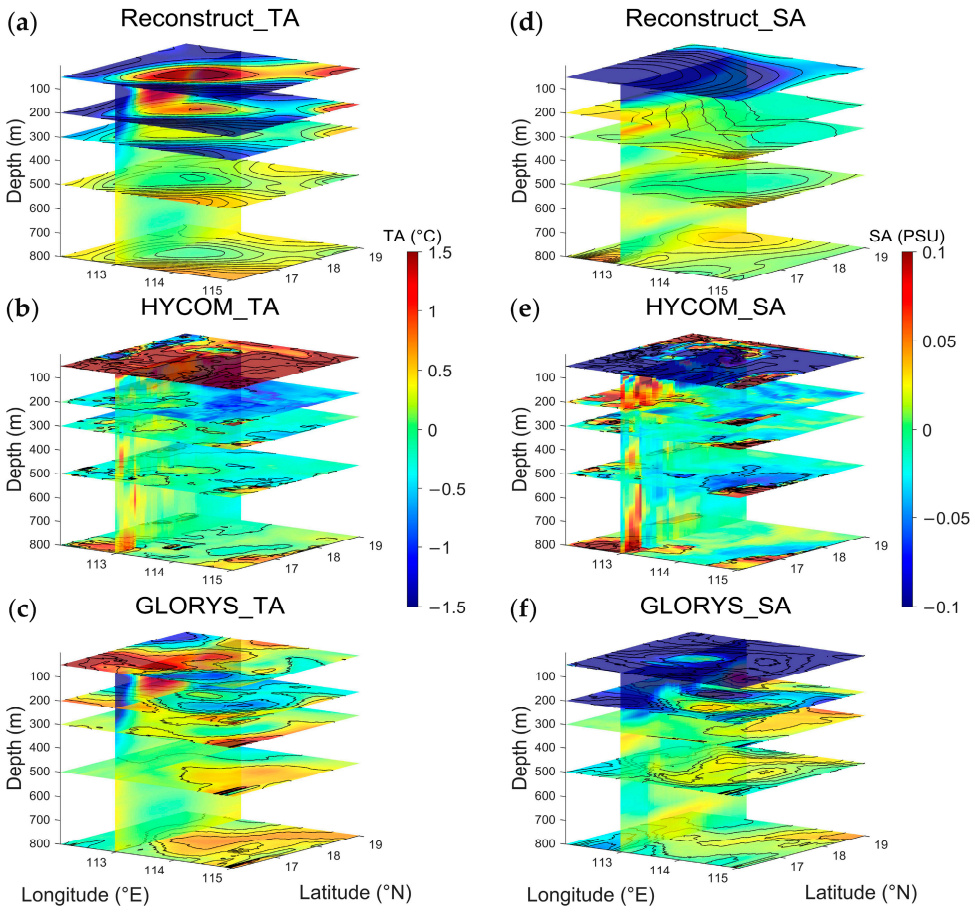

Figure 9. Comparison of the three-dimensional structure of the anticyclonic eddy according to the reconstruction model, HYCOM, and GLORYS outputs on 8 July 2016. Panels (**a**,**d**): 3D temperature and salinity structures reconstructed by the deep learning model. Panels (**b**,**e**): structures from HYCOM. Panels (**c**,**f**): structures from GLORYS.

4.5. Subsurface Temperature and Salinity Reconstruction for Anticyclonic/Cyclonic Eddies in Long Time Series

Sections 4.1 and 4.2 employed glider measurements, HYCOM simulation, and GLORYS reanalysis data to assess the model's accuracy in reconstructing subsurface temperature and salinity within mesoscale eddies, demonstrating its superior performance. Next, we utilized the model to reconstruct the subsurface temperature and salinity for AEs and CEs over a long-term period (time: 2020–2021; spatial: the SCS including the extended Western Pacific area marked by the black box in Figure 1), calculating the residual errors between the reconstructed results and corresponding Argo observations. The results shown in the figures are sorted chronologically based on the matching dataset of Argo data and the mesoscale eddies dataset. Figure 10 displays the subsurface temperature within AEs and CEs, and Figure 11 shows the salinity. We investigated the correlation between SLA with the temperature anomalies (TAs) and salinity anomalies (SAs) at different depths based on the Pearson correlation coefficient analysis. As shown in Figure 10a,e, the trends of the SLA, TA, and SA agree well, especially at 350–400 m depth. Figures 10b,c and 11b,c show the subsurface temperature and salinity anomalies within AEs observed by Argo floats and reconstructed data by the deep learning model. At the spatial scale, the positive anomalies induced by AEs exhibit varying depths, with some extending to approximately 200 m, while others reach around 600 m or even deeper. Figures 10f,g and 11f,g display the subsurface temperature and salinity anomalies associated with CEs. These anomalies are predominantly negative within the upper 600 m, with their maxima primarily distributed between 200 m and 400 m. The variation may be attributed to the distinct oceanic regions where these eddies are located. There is a clear seasonal variation in mesoscale eddies at seasonal time scales where the intensity of the AEs is stronger than the CEs in summer and the reverse in winter. The characteristics of subsurface temperature and salinity within these eddies are consistent with previously established research findings [56,57]. Figure 10d presents the residual errors (the difference between Argo observations and the data reconstructed by the model) for subsurface temperature within AEs, with a maximum error below 2 °C, primarily concentrated in the thermocline between 50 m and 200 m. A similar trend is observed in the subsurface temperature reconstruction within CEs, as shown in Figure 10h. On the other hand, the maximal residual errors for subsurface salinity within AEs are less than 0.2 PSU (Figure 11d,h). These results conclude that the deep learning model, which incorporates eddy physical parameters, is highly effective for reconstructing subsurface temperature and salinity within both anticyclonic and cyclonic eddies.

Meanwhile, Table 1 and Figure 12 indicates that the root mean square error (RMSE) of temperature and salinity varies across different depths, with the error peaking at the depth of 50 m and maintaining a relatively high RMSE in the middle layer (50–600 m depth). The outliers in the RMSE trend might be associated with the complex dynamic processes of the upper ocean and the impact of the thermocline, where the temperature and salinity anomalies' characteristics are significant, and the variation range is large across these depths [29,58]. In contrast to previous results indicating a gradual decrease in estimation accuracy below 500 m, our deep learning model improves the accuracy of estimating thermohaline structure as depth increases. This highlights the potential of satellite measurements at the sea surface to interpret deeper ocean processes influenced by mesoscale eddies. To further validate the accuracy and robustness of the reconstruction model, the eddy convolutional network (ECN) proposed by Yu et al. [31] is employed to conduct a comparison analysis. The RMSE for two models is provided in Table 1. The RMSE values for the reconstruction model are significantly smaller than those for ECN both within AEs and CEs, demonstrating that the reconstruction model is more effective in capturing temperature and salinity anomalies across various depth layers within eddies than the ECN.

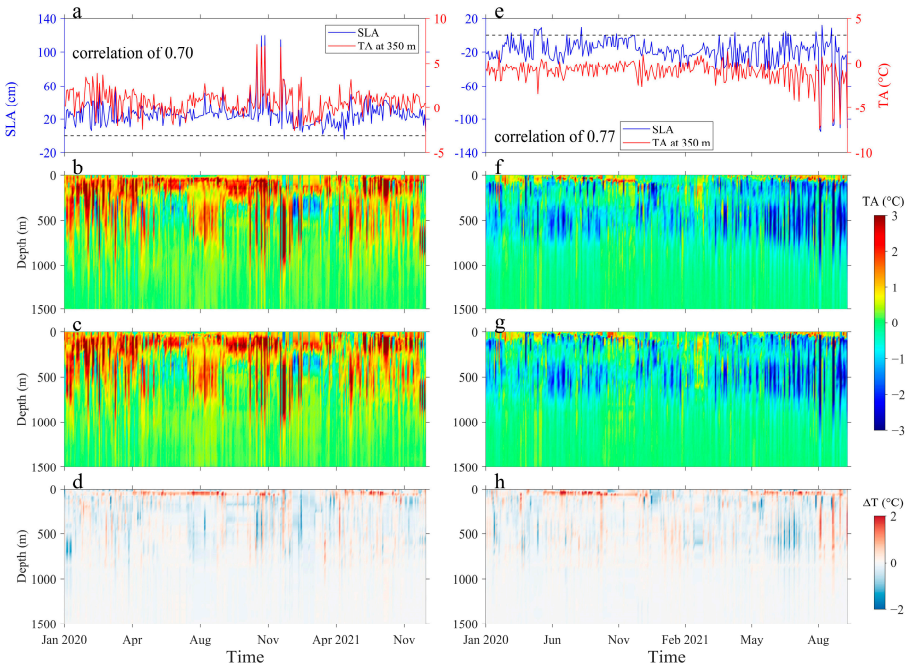

Figure 10. Comparison of temperature anomalies (TAs) between the Argo observations and the reconstruction model. Panels (**a**–**d**): the comparison of the temperature anomalies within AEs. (**a**) Times series of temperature anomalies (350 m, red solid line) is overplotted against SLA (blue solid line). (**b**,**c**) The temperature anomalies derived from Argo observations and the reconstruction model, respectively; (**d**) the residual error of subsurface temperature within AEs. The residual error is calculated as the difference between Argo observations and the data reconstructed by the model; panels (**e**–**h**) are the same as panels (**a**–**d**), but for the comparison of temperature anomalies within CEs.

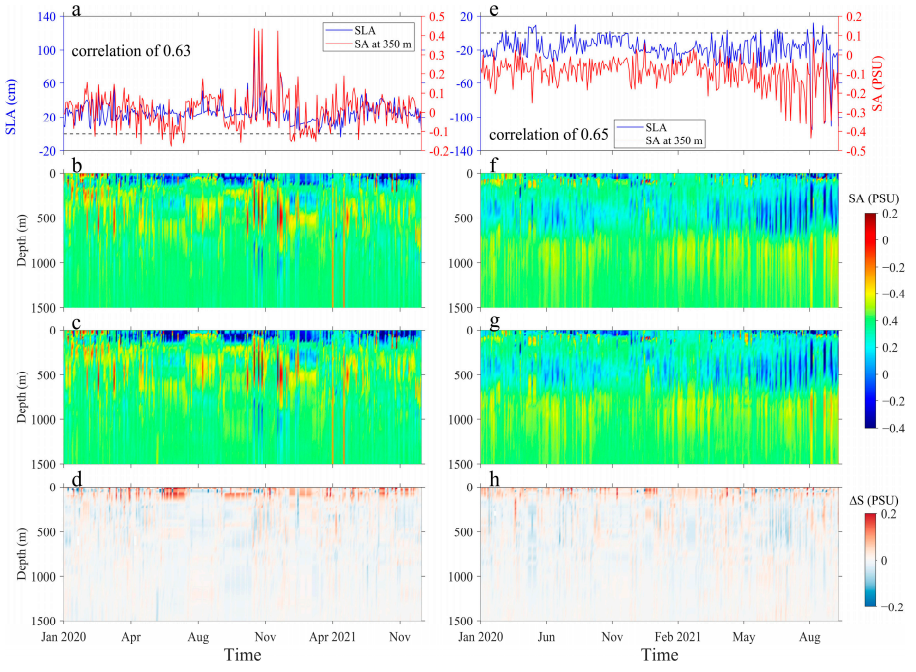

Figure 11. Comparison of salinity anomalies (SAs) between the Argo observations and the reconstruction model. Panels (**a**–**d**): the comparison of the salinity anomalies within AEs. (**a**) Times series of salinity anomalies (350 m, red solid line) is overplotted against SLA (blue solid line); (**b**,**c**) the salinity anomalies derived from Argo observations and the reconstruction model, respectively; (**d**) the residual error of subsurface salinity within AEs. The residual error is calculated as the difference between Argo observations and the data inverted by the model; panels (**e**–**h**) are the same as panels (**a**–**d**), respectively, but focus on the comparison of salinity anomalies within CEs.

Table 1. RMSE (root mean square error) of the reconstructed subsurface temperature and salinity within AEs and CEs at different depths.

Depths/Location	RMSE				RMSE (ECN)	
	Temp (°C)		Sal (PSU)		Temp (°C)	
	AEs	CEs	AEs	CEs	AEs	CEs
5 m	0.45	0.47	0.079	0.064	0.85	0.88
50 m	0.75	0.72	0.071	0.057	0.94	1.07
100 m	0.44	0.48	0.062	0.040	0.83	0.94
150 m	0.43	0.45	0.030	0.027	0.78	0.98
200 m	0.38	0.38	0.023	0.025	0.70	0.68
300 m	0.34	0.32	0.025	0.029	0.64	0.72
400 m	0.37	0.38	0.023	0.028	0.63	0.58
500 m	0.33	0.40	0.020	0.027	0.88	0.64
600 m	0.32	0.37	0.019	0.023	0.91	0.68
700 m	0.26	0.22	0.013	0.018	0.61	0.28
800 m	0.17	0.16	0.014	0.019	0.35	0.14
900 m	0.11	0.11	0.013	0.015	0.17	0.15
1000 m	0.09	0.08	0.012	0.014	0.18	0.08
1200 m	0.07	0.05	0.012	0.010	/	/
1500 m	0.04	0.03	0.008	0.008	/	/
Average	0.34	0.36	0.036	0.032	0.65	0.6
Eddy center	0.0686	0.05	0.01	0.0063	/	/
Eddy edge	0.12	0.1	0.013	0.015	/	/

Furthermore, we employed the coefficient of determination (R^2) and root mean square error (RMSE) to evaluate the model's capability to reconstruct temperature and salinity for cyclonic eddies (CEs) and anticyclonic eddies (AEs). As shown in Figure 12, the results demonstrate distinct differences in model performance between the two types of eddies. Overall, the RMSE (R^2) is smaller (larger) for AEs compared to CEs, indicating that the model performs more effectively in reconstructing the thermohaline structure of AEs. The maximum errors occur in the upper ocean within both AEs and CEs where mesoscale eddies typically exhibit the strongest thermohaline anomalies due to enhanced vertical mixing and surface forcing. These errors diminish with increasing depth. Specifically, at depths ranging from the surface to 200 m, the RMSE is at its peak due to larger temperature and salinity gradients, while the R^2 is relatively lower in this region. Between 200 m and 500 m, the RMSE decreases significantly, and the R^2 improves, as the temperature and salinity gradients become more uniform and easier for the model to predict. Below 500 m, the errors stabilize, and the differences in model performance between AEs and CEs become less pronounced.

4.6. Evaluating the Impact of the Edge Effects of Mesoscale Eddies

In the physics prior knowledge-embedded branch of the reconstruction model, we consider the distance between the Argo profile and the eddy center as a constraint to improve the reconstruction accuracy of the subsurface temperature and salinity at any point within the eddy. However, the profile-by-profile validation conducted in Section 4.1 reveals that the edge effects weaken the accuracy of the reconstruction model. We screened the Argo data into two categories to further assess this effect on the accuracy of the deep learning model: the Argo data at a distance equivalent to 1/4 eddy radius from the center is characterized as the center of the eddy, while the Argo data at a distance equivalent to 3/4

or even one eddy radius from the center is served as the edge of the eddy. Figure 10a,b,e,f show the residual errors of the temperature and salinity at the centers of the eddies, while Figure 13c,d,g,h represent those at the edges of the eddies. The residual errors are more pronounced at the edges of the eddies than at the centers of the eddies, with the maximum temperature error up to 2 °C and the maximum salinity error up to 0.2 PSU. We further calculate the average RMSE of the temperature and salinity across all depths at the centers and edges, respectively. The RMSE values for the eddy center and eddy edge are shown in Table 1.

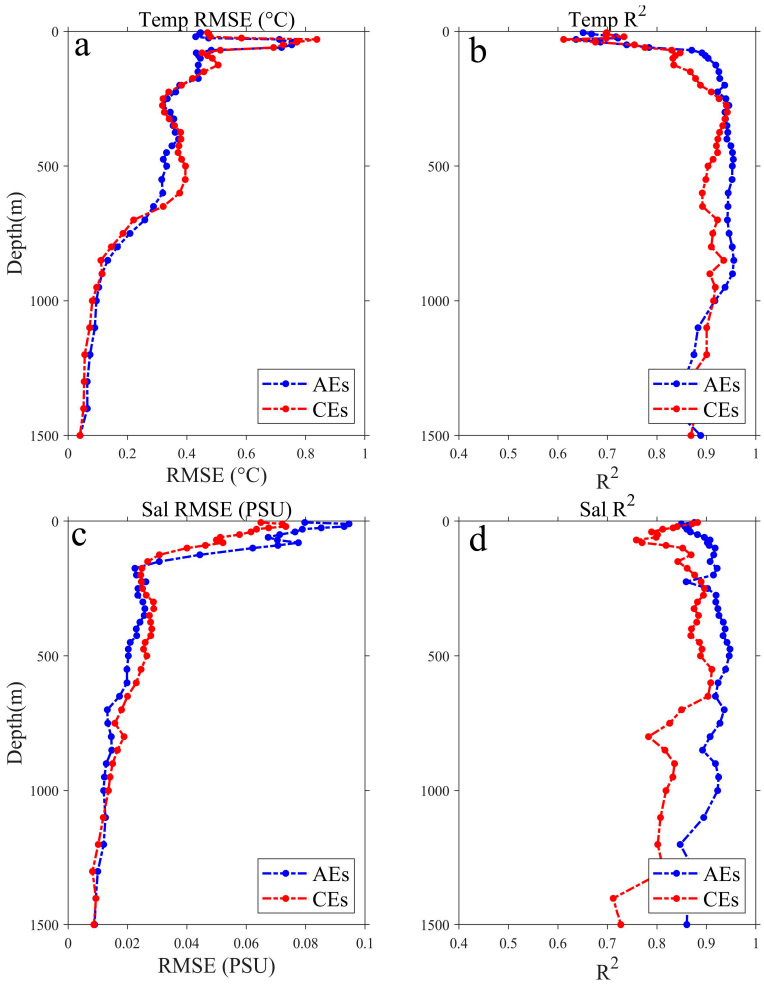

Figure 12. Comparison of RMSE and R^2 of the reconstruction model applied in AEs (blue lines) and CEs (red lines). Panels (**a**,**c**): the comparison of the RMSE of temperature and salinity. Panels (**b**,**d**): the comparison of the R^2 of temperature and salinity.

Despite the residual errors and the RMSE demonstrating that the reconstruction accuracy at the edges of the eddies is inferior to that near the centers of the eddies, the results at the edges of the eddies are still within a reasonable error range, which is consistent with previous studies [32]. This discrepancy may be due to the pronounced peripheral effect, where the substantial shear variance at the edges of the eddy is 1 to 3 orders of magnitude greater than that observed at the center [59,60]. The increased variance at the edges complicates the reconstruction process, increasing errors in temperature and salinity estimations. These metrics also demonstrate that relying solely on eddy surface parameters as prior knowledge may not be sufficient to fully improve the reconstruction accuracy of the deep learning model.

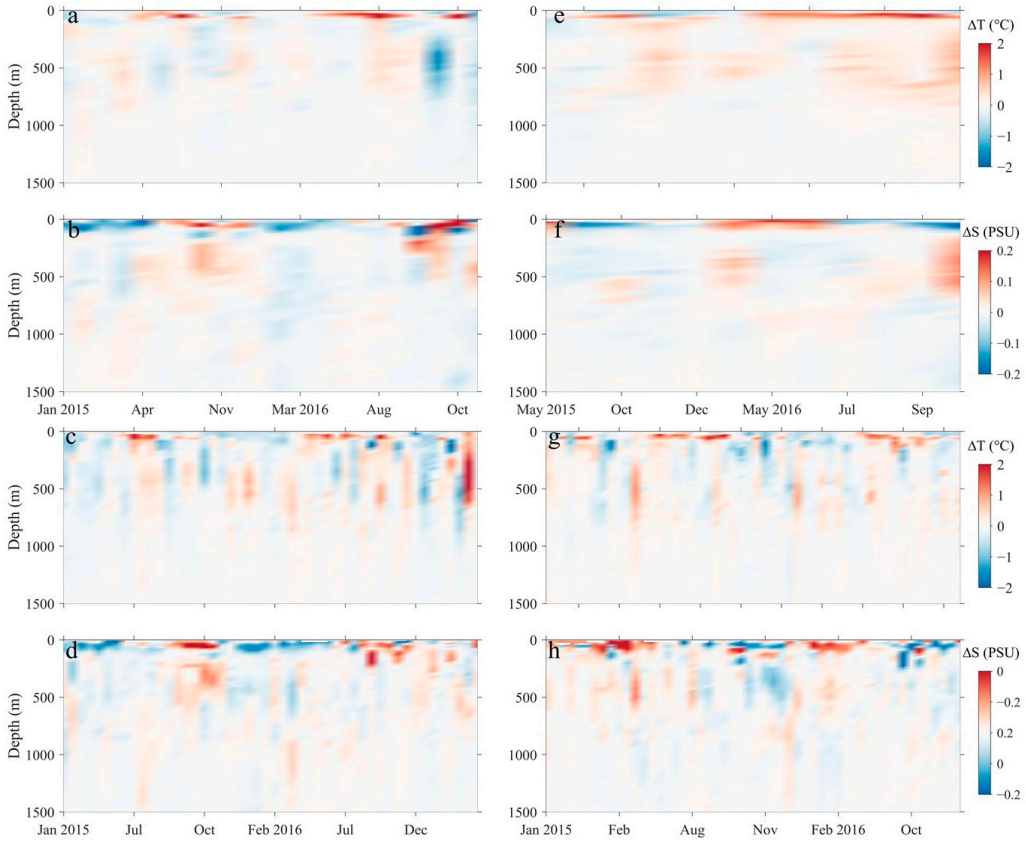

Figure 13. The residual error of subsurface temperature and salinity at the centers and edges of the eddies. Panels (**a**–**d**) represent the AEs (eddy centers in panels a-b and eddy edges in panels (**c**,**d**). Panels (**e**–**h**) are the same as panels (**a**–**d**), respectively, but focus on the CEs.

5. Discussion

In this study, we develop and validate a three-dimensional thermohaline reconstruction model of mesoscale eddies based on deep learning techniques that integrate satellite, Argo, glider, HYCOM simulation, and GLORYS reanalysis data. In light of the distinctive and intricate characteristics of the South China Sea, the mesoscale eddies within this region are identified as a research focus to provide an effective technique for studying the 3D thermohaline structure of mesoscale eddies in the marginal sea. Deviating from previous deep learning-based models that relied solely on sea surface remote sensing data to infer subsurface temperature and salinity [22,61], our study divides the model into three branches, including the satellite data branch, the prior knowledge of the eddies branch, and the reconstruction branch that integrate physical prior knowledge of eddies, such as eddy radius (Rad), amplitude (Amp), vorticity (Vor), rotational speed (U), and eddy kinetic energy (EKE) into the model. Compared to the traditional deep learning-based model, the reconstruction model achieves a reduction in RMSE ranging from 25% to 50% across all depths.

In addition to employing the Argo observations to verify the model, we conducted a comparison analysis between the reconstruction results and glider observations of an anticyclonic eddy (AE) observed in the South China Sea. Altimetric measurements reveal that this AE persisted for 52 days, with an average radius of 74 km and amplitude of 5.8 cm. Utilizing the model, we reconstructed the 3D thermohaline structure of the eddy. A profile-by-profile comparative analysis reveals two notable details of the vertical structure. First, there are larger reconstruction errors near the thermocline (100–200 m) compared to other depths. Second, a pronounced edge effect was observed, with greater reconstruction

accuracy at the centers of the eddy relative to the edges. These findings underscore the significance of internal physical mechanisms in reconstructing the 3D structure of eddies. Further validation through comparisons with HYCOM simulation and GLORYS reanalysis data confirmed the model's effectiveness, particularly in the thermocline region.

A thorough analysis and comparison of eddies were conducted on spatial and temporal scales utilizing the established model. The model-predicted subsurface temperature and salinity structures are highly consistent with Argo observations. In addition to the robust performance in accurately capturing the depth-dependent changes in subsurface temperature and salinity within eddies, the reconstruction model also effectively models the variations of the temperature and salinity horizontally. We categorize the Argo data into the data at the centers and the edges of the eddies to demonstrate this ability. Through comparison of the residual errors and RMSE between the reconstruction at the centers and the edges of the eddies, the model demonstrates excellent generalization.

Given the critical role of dataset size, regional specificity, and diversity in oceanographic applications, we trained three models with distinct datasets. Model1 is trained using the combined dataset from the SCS and the extended Northwest Pacific area, Model2 is trained using Argo profiles from the broader West Pacific, and Model3 is trained exclusively on Argo profiles within the SCS. The results demonstrate that Model1 has improved performance compared to Model2 and Model3. This highlights the importance of balancing dataset size and regional specificity to achieve robust reconstructions of mesoscale eddy-induced temperature and salinity anomalies. Future efforts should focus on integrating diverse in situ and remote sensing observations to enhance model performance.

In summary, this study provides a novel model for reconstructing the 3D thermohaline structure of mesoscale eddies by adapting satellite data and physical parameters of the mesoscale eddies as dual-input. This approach enables a more accurate assessment of the role of mesoscale eddies in transporting mass and energy and offers an efficient perspective for data reconstruction and parameter inversion related to various marine phenomena. Despite achieving success in reconstructing 3D eddy structures, it is important to acknowledge the limitations of the model. The uncertainties of reconstruction accuracy at the thermocline and eddy edges suggest that relying solely on eddy surface parameters as prior knowledge may not be sufficient to comprehensively capture the thermohaline characteristics of mesoscale eddies. To enhance model accuracy, it is essential to incorporate the interior physical mechanisms. Currently, several AI-based models with physical constraints have been proposed and have demonstrated impressive performance, such as Physics-Informed Neural Networks (PINNs) [62–64]. Further investigation should incorporate physical constraints at the model development stages to improve the inversion accuracy of 3D eddy structures.

Author Contributions: Conceptualization, Y.Z. and Z.Z.; methodology, Y.Z. and Z.Z.; software, Z.Z.; validation, Y.Z. and Z.Z.; formal analysis, Z.Z. and Y.Z.; investigation, Y.Z.; resources, Y.Z.; data curation, J.Y., Z.Z., L.Z., W.R. and D.L.; writing—original draft preparation, Z.Z.; writing—review and editing, Y.Z. and Z.Z.; visualization, Z.Z.; supervision, Y.Z.; project administration, Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (41876048) and Interdisciplinary Joint Research Projects of Tongji University. Yanwei Zhang was also supported by the Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai).

Data Availability Statement: The data sets used in this work are available in online repositories. AVISO SLA data can be downloaded from (SEALEVEL_GLO_PHY_L4_MY_008_047, https://doi.org/10.48670/moi-00148, accessed on 22 March 2022). SST data can be downloaded from (https://www.ncei.noaa.gov/products/optimum-interpolation-sst, accessed on 22 March 2022). SSS data

can be downloaded from (MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013, https://doi.org/10.48670/moi-00051, accessed on 2 March 2024). SSW data can be downloaded from (Available at www.remss.com, https://doi.org/10.56236/RSS-uv6h30, accessed on 2 March 2024). WOA13 data can be downloaded from https://www.ncei.noaa.gov/products/world-ocean-atlas/, accessed on 2 February 2022. Argo data were made available by the International Argo Program and the national programs that contribute to it (http://www.argo.ucsd.edu, accessed on 15 and 20 February 2022). The Argo Program is part of the Global Ocean Observing System. Simulated and reanalyzed products were distributed by HYCOM (https://tds.hycom.org/thredds/catalog.html, accessed on 2 March 2024). The GLORYS products can be downloaded from (GLOBAL_MULTIYEAR_PHY_001_030, https://doi.org/10.48670/moi-00021, accessed on 8 December 2024). Mesoscale eddy dataset can be downloaded from (https://doi.org/10.24400/527896/a01-2022.005.220209, accessed on 8 December 2024). The glider dataset presented in this article can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nan, F.; He, Z.; Zhou, H.; Wang, D. Three long-lived anticyclonic eddies in the northern South China Sea. *J. Geophys. Res.-Ocean.* **2011**, *116*, C05002. [CrossRef]
2. Chen, G.; Yang, J.; Han, G. Eddy morphology: Egg-like shape, overall spinning, and oceanographic implications. *Remote Sens. Environ.* **2021**, *257*, 112348. [CrossRef]
3. Sun, B.; Liu, C.; Wang, F. Global meridional eddy heat transport inferred from Argo and altimetry observations. *Sci. Rep.* **2019**, *9*, 1345. [CrossRef] [PubMed]
4. Zhang, Z.; Zhang, Y.; Wang, W.; Huang, R.X. Universal structure of mesoscale eddies in the ocean. *Geophys. Res. Lett.* **2013**, *40*, 3677–3681. [CrossRef]
5. Liu, F.; Tang, S.; Chen, C. Impact of nonlinear mesoscale eddy on phytoplankton distribution in the northern South China Sea. *J. Mar. Syst.* **2013**, *123*, 33–40. [CrossRef]
6. Xiu, P.; Chai, F. Eddies Affect Subsurface Phytoplankton and Oxygen Distributions in the North Pacific Subtropical Gyre. *Geophys. Res. Lett.* **2020**, *47*, e2020GL087037. [CrossRef]
7. Xu, W.; Wang, G.; Cheng, X.; Xing, X.; Qin, J.; Zhou, G.; Jiang, L.; Chen, B. Mesoscale Eddy Modulation of Subsurface Chlorophyll Maximum Layers in the South China Sea. *J. Geophys. Res.-Biogeosci.* **2023**, *128*, e2023JG007648. [CrossRef]
8. Shu, Y.; Wang, Q.; Zu, T. Progress on shelf and slope circulation in the northern South China Sea. *Sci. China-Earth Sci.* **2018**, *61*, 560–571. [CrossRef]
9. Liu, Q.; Kaneko, A.; Su, J. Recent progress in studies of the South China Sea circulation. *J. Oceanogr.* **2008**, *64*, 753–762. [CrossRef]
10. He, Q.; Zhan, H.; Cai, S.; He, Y.; Huang, G.; Zhan, W. A New Assessment of Mesoscale Eddies in the South China Sea: Surface Features, Three-Dimensional Structures, and Thermohaline Transports. *J. Geophys. Res.-Ocean.* **2018**, *123*, 4906–4929. [CrossRef]
11. Xiu, P.; Chai, F.; Shi, L.; Xue, H.; Chao, Y. A census of eddy activities in the South China Sea during 1993–2007. *J. Geophys. Res.-Ocean.* **2010**, *115*. [CrossRef]
12. Lin, X.; Wang, G.; Han, G.; Liu, Y.; Zhang, H.; Liao, X.; Ji, Q. Cross-Slope Transport by a Mesoscale Anticyclone in the Northern South China Sea. *J. Mar. Sci. Eng.* **2023**, *11*, 305. [CrossRef]
13. Zhang, Y.; Liu, Z.; Zhao, Y.; Wang, W.; Li, J.; Xu, J. Mesoscale eddies transport deep-sea sediments. *Sci. Rep.* **2014**, *4*, 5937. [CrossRef] [PubMed]
14. Shu, Y.; Chen, J.; Li, S.; Wang, Q.; Yu, J.; Wang, D. Field-observation for an anticyclonic mesoscale eddy consisted of twelve gliders and sixty-two expendable probes in the northern South China Sea during summer 2017. *Sci. China-Earth Sci.* **2019**, *62*, 451–458. [CrossRef]
15. Sun, W.; Dong, C.; Wang, R.; Liu, Y.; Yu, K. Vertical structure anomalies of oceanic eddies in the Kuroshio Extension region. *J. Geophys. Res.-Ocean.* **2017**, *122*, 1476–1496. [CrossRef]
16. Zhang, Y.; Xu, H.; Qin, R.; Xu, C.; Fan, D. Characteristics of turbulent kinetic energy dissipation rate and turbidity near the coast of East China Sea. *Chin. J. Oceanol. Limnol.* **2016**, *34*, 1134–1142. [CrossRef]
17. Zhang, Y.; Liu, Z.; Zhao, Y.; Li, J.; Liang, X. Effect of surface mesoscale eddies on deep-sea currents and mixing in the northeastern South China Sea. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2015**, *122*, 6–14. [CrossRef]
18. Klemas, V.; Yan, X.-H. Subsurface and deeper ocean remote sensing from satellites: An overview and new results. *Prog. Oceanogr.* **2014**, *122*, 1–9. [CrossRef]
19. Haines, K. A Direct Method for Assimilating Sea-Surface Height Data into Ocean Models with Adjustments to the Deep Circulation. *J. Phys. Oceanogr.* **1991**, *21*, 843–868. [CrossRef]

20. Yan, H.; Wang, H.; Zhang, R.; Chen, J.; Bao, S.; Wang, G. A Dynamical-Statistical Approach to Retrieve the Ocean Interior Structure From Surface Data: SQG-mEOF-R. *J. Geophys. Res.-Ocean.* **2020**, *125*, e2019JC015840. [CrossRef]
21. LaCasce, J.H.; Wang, J. Estimating Subsurface Velocities from Surface Fields with Idealized Stratification. *J. Phys. Oceanogr.* **2015**, *45*, 2424–2435. [CrossRef]
22. Liu, L.; Peng, S.; Huang, R.X. Reconstruction of ocean's interior from observed sea surface information. *J. Geophys. Res.-Ocean.* **2017**, *122*, 1042–1056. [CrossRef]
23. Pun, I.-F.; Price, J.F.; Jayne, S.R. Satellite-Derived Ocean Thermal Structure for the North Atlantic Hurricane Season. *Mon. Weather Rev.* **2016**, *144*, 877–896. [CrossRef]
24. Jeong, Y.; Hwang, J.; Park, J.; Jang, C.J.; Jo, Y.-H. Reconstructed 3-D Ocean Temperature Derived from Remotely Sensed Sea Surface Measurements for Mixed Layer Depth Analysis. *Remote Sens.* **2019**, *11*, 3018. [CrossRef]
25. Pauthenet, E.; Bachelot, L.; Balem, K.; Maze, G.; Tréguier, A.-M.; Roquet, F.; Fablet, R.; Tandeo, P. Four-dimensional temperature, salinity and mixed-layer depth in the Gulf Stream, reconstructed from remote-sensing and in situ observations with neural networks. *Ocean Sci.* **2022**, *18*, 1221–1244. [CrossRef]
26. Chen, Z.; Wang, P.; Bao, S.; Zhang, W. Rapid reconstruction of temperature and salinity fields based on machine learning and the assimilation application. *Front. Mar. Sci.* **2022**, *9*, 985048. [CrossRef]
27. Meng, L.; Yan, C.; Zhuang, W.; Zhang, W.; Yan, X.H. Reconstruction of Three-Dimensional Temperature and Salinity Fields From Satellite Observations. *J. Geophys. Res. Ocean.* **2021**, *126*, e2021JC017605. [CrossRef]
28. Bao, S.; Zhang, R.; Wang, H.; Yan, H.; Yu, Y.; Chen, J. Salinity Profile Estimation in the Pacific Ocean from Satellite Surface Salinity Observations. *J. Atmos. Ocean. Technol.* **2019**, *36*, 53–68. [CrossRef]
29. Su, H.; Li, W.; Yan, X.-H. Retrieving Temperature Anomaly in the Global Subsurface and Deeper Ocean From Satellite Observations. *J. Geophys. Res.-Ocean.* **2018**, *123*, 399–410. [CrossRef]
30. Han, M.; Feng, Y.; Zhao, X.; Sun, C.; Hong, F.; Liu, C. A Convolutional Neural Network Using Surface Data to Predict Subsurface Temperatures in the Pacific Ocean. *IEEE Access* **2019**, *7*, 172816–172829. [CrossRef]
31. Yu, F.; Wang, Z.; Liu, S.; Chen, G. Inversion of the three-dimensional temperature structure of mesoscale eddies in the Northwest Pacific based on deep learning. *Acta Oceanol. Sin.* **2021**, *40*, 176–186. [CrossRef]
32. Liu, Y.; Wang, H.; Li, X. A Deep Learning-Based Mesoscale Eddy Subsurface Temperature Inversion Model. In Proceedings of the IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 6884–6886.
33. Huo, J.D.; Yang, J.A.; Geng, L.T.; Liu, G.L.; Zhang, J.; Wang, J.C.; Cui, W. Temperature Structure Inversion of Mesoscale Eddies in the South China Sea Based on Deep Learning. *J. Mar. Sci. Eng.* **2024**, *12*, 723. [CrossRef]
34. Wu, M.; He, H.; Fan, D.; Wang, T.; Fu, B. Identification and Analysis of Mesoscale Eddies in the South China Sea. In Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 11–14 December 2020; pp. 2463–2469.
35. Chen, G.; Hou, Y.; Chu, X. Mesoscale eddies in the South China Sea: Mean properties, spatiotemporal variability, and impact on thermohaline structure. *J. Geophys. Res.-Ocean.* **2011**, *116*. [CrossRef]
36. Dong, D.; Brandt, P.; Chang, P.; Schuette, F.; Yang, X.; Yan, J.; Zeng, J. Mesoscale Eddies in the Northwestern Pacific Ocean: Three-Dimensional Eddy Structures and Heat/Salt Transports. *J. Geophys. Res.-Ocean.* **2017**, *122*, 9795–9813. [CrossRef]
37. Chen, G.; Chen, X.Y.; Huang, B.X. Independent Eddy Identification With Profiling Argo as Calibrated by Altimetry. *J. Geophys. Res.-Ocean.* **2021**, *126*, e2020JC016729. [CrossRef]
38. Liu, Z.; Chen, X.; Yu, J.; Xu, D.; Sun, C. Kuroshio intrusion into the South China Sea with an anticyclonic eddy: Evidence from underwater glider observation. *J. Oceanol. Limnol.* **2019**, *37*, 1469–1480. [CrossRef]
39. Wang, P.; Mao, K.; Chen, X.; Liu, K. The Three-Dimensional Structure of the Mesoscale Eddy in the Kuroshio Extension Region Obtained from Three Datasets. *J. Mar. Sci. Eng.* **2022**, *10*, 1754. [CrossRef]
40. Reynolds, R.W.; Smith, T.M.; Liu, C.; Chelton, D.B.; Casey, K.S.; Schlax, M.G. Daily high-resolution-blended analyses for sea surface temperature. *J. Clim.* **2007**, *20*, 5473–5496. [CrossRef]
41. Atlas, R.; Hoffman, R.N.; Ardizzone, J.; Leidner, S.M.; Jusem, J.C.; Smith, D.K.; Gombos, D. A Cross-Calibrated Multiplatform Ocean Surface Wind Velocity Product for Meteorological and Oceanographic Applications. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 157–174. [CrossRef]
42. Mason, E.; Pascual, A.; McWilliams, J.C. A New Sea Surface Height–Based Code for Oceanic Mesoscale Eddy Tracking. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1181–1188. [CrossRef]
43. Dai, J.; Wang, H.; Zhang, W.; An, Y.; Zhang, R. Observed spatiotemporal variation of three-dimensional structure and heat/salt transport of anticyclonic mesoscale eddy in Northwest Pacific. *J. Oceanol. Limnol.* **2020**, *38*, 1654–1675. [CrossRef]
44. Ding, M.; Lin, P.; Liu, H.; Hu, A.; Liu, C. Lagrangian eddy kinetic energy of ocean mesoscale eddies and its application to the Northwestern Pacific. *Sci. Rep.* **2020**, *10*, 12791. [CrossRef]

45. Lan, J.; Zhang, N.; Wang, Y. On the dynamics of the South China Sea deep circulation. *J. Geophys. Res.-Ocean.* **2013**, *118*, 1206–1210. [CrossRef]
46. Chassignet, E.P.; Hurlburt, H.E.; Smedstad, O.M.; Halliwell, G.R.; Hogan, P.J.; Wallcraft, A.J.; Baraille, R.; Bleck, R. The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *J. Mar. Syst.* **2007**, *65*, 60–83. [CrossRef]
47. Wang, Z.; Yashayaev, I.; Greenan, B. Seasonality of the inshore Labrador current over the Newfoundland shelf. *Cont. Shelf Res.* **2015**, *100*, 1–10. [CrossRef]
48. de Souza, J.M.A.C.; Couto, P.; Soutelino, R.; Roughan, M. Evaluation of four global ocean reanalysis products for New Zealand waters–A guide for regional ocean modelling. *N. Z. J. Mar. Freshw. Res.* **2021**, *55*, 132–155. [CrossRef]
49. Mao, K.; Liu, C.; Zhang, S.; Gao, F. Reconstructing Ocean Subsurface Temperature and Salinity from Sea Surface Information Based on Dual Path Convolutional Neural Networks. *J. Mar. Sci. Eng.* **2023**, *11*, 1030. [CrossRef]
50. Jain, M.; Saihjpal, V.; Singh, N.; Singh, S.B. An Overview of Variants and Advancements of PSO Algorithm. *Appl. Sci. -Basel* **2022**, *12*, 8392. [CrossRef]
51. Yu, F.; Zhuang, Z.; Yang, J.; Chen, G. A Glider Simulation Model Based on Optimized Support Vector Regression for Efficient Coordinated Observation. *Front. Mar. Sci.* **2021**, *8*, 671791. [CrossRef]
52. Chaudhari, S.; Mithal, V.; Polatkan, G.; Ramanath, R. An Attentive Survey of Attention Models. *Acm Trans. Intell. Syst. Technol.* **2021**, *12*, 1–32. [CrossRef]
53. Tang, B.; Zhao, D.; Cui, C.; Zhao, X. Reconstruction of ocean temperature and salinity profiles in the Northern South China Sea using satellite observations. *Front. Mar. Sci.* **2022**, *9*, 945835. [CrossRef]
54. Zhao, N.; Han, Z. A Simulation Model of Seawater Vertical Temperature by Using Back-Propagation Neural Network. *Pol. Marit. Res.* **2015**, *22*, 82–88. [CrossRef]
55. Zhang, Z.; Zhang, Y.; Wang, W. Three-compartment structure of subsurface-intensified mesoscale eddies in the ocean. *J. Geophys. Res.-Ocean.* **2017**, *122*, 1653–1664. [CrossRef]
56. Yang, G.; Wang, F.; Li, Y.; Lin, P. Mesoscale eddies in the northwestern subtropical Pacific Ocean: Statistical characteristics and three-dimensional structures. *J. Geophys. Res.-Ocean.* **2013**, *118*, 1906–1925. [CrossRef]
57. Yang, G.; Lin, X.; Han, G.; Liu, Y.; Chen, G.; Wang, J. Three-dimensional characteristics of mesoscale eddies simulated by a regional model in the northwestern Pacific Ocean during 2000–2008. *Acta Oceanol. Sin.* **2022**, *41*, 74–93. [CrossRef]
58. Su, H.; Wu, X.; Yan, X.-H.; Kidwell, A. Estimation of subsurface temperature anomaly in the Indian Ocean during recent global surface warming hiatus from satellite measurements: A support vector machine approach. *Remote Sens. Environ.* **2015**, *160*, 63–71. [CrossRef]
59. Wunsch, C. The Oceanic Variability Spectrum and Transport Trends. *Atmos.-Ocean* **2009**, *47*, 281–291. [CrossRef]
60. Wang, M.; Zhang, Y.; Liu, Z.; Zhao, Y.; Li, J. Temporal and spatial evolution of a deep-reaching anticyclonic eddy in the South China Sea. *Sci. China Earth Sci.* **2019**, *62*, 1002–1023. [CrossRef]
61. Xie, H.; Xu, Q.; Cheng, Y.; Yin, X.; Fan, K. Reconstructing three-dimensional salinity field of the South China Sea from satellite observations. *Front. Mar. Sci.* **2023**, *10*, 1168486. [CrossRef]
62. Wang, N.; Chen, Q.; Chen, Z. Reconstruction of nearshore wave fields based on physics-informed neural networks. *Coast. Eng.* **2022**, *176*, 104167. [CrossRef]
63. Rasht-Behesht, M.; Huber, C.; Shukla, K.; Karniadakis, G.E. Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *J. Geophys. Res. Solid Earth* **2022**, *127*, e2021JB023120. [CrossRef]
64. Xiao, M.-J.; Yu, T.-C.; Zhang, Y.-S.; Yong, H. Physics-informed neural networks for the Reynolds-Averaged Navier–Stokes modeling of Rayleigh–Taylor turbulent mixing. *Comput. Fluids* **2023**, *266*, 106025. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.