



Technical Note

Improving the Reconstruction of Vertical Temperature Profiles on Account of Oceanic Front Impacts

Xin Chen ¹, Chen Wang ^{1,2,*} , Huimin Li ^{1,2} and Yijun He ^{1,2} ¹ School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China² The Key Laboratory of Space Ocean Remote Sensing and Application, Ministry of Natural Resources, Beijing 100081, China

* Correspondence: cwang@nuist.edu.cn

Abstract: The application of remote sensing observations in estimating ocean sub-surface temperatures has been widely adopted. Machine learning-based methods in particular are gaining more and more interest. While there is promising relevant progress, most temperature profile reconstruction models are still built upon the gridded Argo data regardless of the impacts of mesoscale oceanic processes. As a follow-on to the previous study that demonstrates the influence of ocean fronts is negligible, we focus on the improvement of temperature profile reconstruction by introducing the sea surface temperature (SST) gradient into the neural network model. The model sensitivity assessments reveal that the normalization of the input variables achieves a higher estimation accuracy than the original scale. Five experiments are then designed to examine the model performances with or without the SST gradient input. Our results confirm that, for a given model configuration, the one with the input of the SST gradient has the lowest reconstruction bias in comparison to the in situ Argo measurements. Such improvement is particularly pronounced below 200 m depth. We also found that the non-linear activation functions and deeper network structures facilitate the performance of reconstruction models. Results of this work open new insights and challenges to refine the mapping of upper ocean temperature structures. While more relevant machine learning methods are worth further exploitation, how to better characterize the mesoscale oceanic processes from surface observations and bring them into the reconstruction models is the key and needs much attention.

Keywords: vertical temperature reconstruction; machine learning; neural network; the impact of ocean fronts



Citation: Chen, X.; Wang, C.; Li, H.; He, Y. Improving the Reconstruction of Vertical Temperature Profiles on Account of Oceanic Front Impacts. *Remote Sens.* **2022**, *14*, 4821. <https://doi.org/10.3390/rs14194821>

Academic Editor: Chung-Ru Ho

Received: 23 August 2022

Accepted: 23 September 2022

Published: 27 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Knowledge of the upper ocean thermal structure is significant for interpreting various oceanographic processes such as mesoscale eddies, internal waves and oceanic fronts, among others [1–4]. It is undoubtable that in situ measurements are the most straightforward and accurate way to obtain the vertical temperature profiles. The launch of the Argo program at the beginning of this century greatly increased the number of temperature profiles, with around 4000 operational floats across the global ocean [5]. They are still sparsely distributed in the ocean basin and have limited spatial coverage. By comparison, spaceborne remote sensing is able to continuously collect observations of multiple sea surface variables, including sea surface level anomaly (SLA), sea surface temperature (SST), sea surface salinity (SSS) and wind, waves and currents, etc. [6–8]. These variables can detail characteristics of the surface features that are found to be associated with the sub-surface thermal structure [9,10]. This drives the research progress into the reconstruction of temperature and/or salinity profiles from the surface observational data.

Numerous studies have been devoted to reconstructing the sub-surface temperature structure by combining satellite observations and in situ data. In general, the current approaches can be divided into two groups: physical and data-driven statistical. The physical

methods mainly involve data assimilation of numerical models [11] and dynamical theory models [12–14]. A data assimilation system obtains the a posteriori temperature profiles based on the given a priori information and the ocean state represented by the numerical model [15]. Though the ocean state constantly follows the governing physics in this way, uncertainties concerning the model and observations tend to add up and cause unexpected biases [16]. In addition, the parameterization in numerical models needs further improvements, particularly for the mesoscale and sub-mesoscale processes. Such methods thus can only work effectively and accurately under certain conditions or under some assumptions [17–20]. The dynamic models are mostly based on the surface quasi-geostrophic (SQG) theory to reconstruct the ocean interior structure. This theory resolves the upper ocean density field with constraints by surface variables, based on which the vertical temperature profiles can be inferred. By using only the sea surface height, the effective SQG is widely introduced to estimate the vertical profiles [21]. The interior plus SQG model was later proposed to extend such application scope and demonstrated good performance [12,13].

The data-driven statistical approach is based on the empirical relationships between the surface variables and the vertical profiles. Regression models are usually employed to obtain their relationship, including linear regression [10,17,22], empirical orthogonal function [23,24] and geographically weighted regression model [25]. With the advancement of machine learning techniques, many end-to-end models have been exploited to reconstruct the three-dimensional sub-surface thermal structure. For example, an artificial neural network model has been trained with inputs of SST and SSHA plus the in situ temperature profiles to map the vertical temperature profiles in the Indian Ocean [26]. The results are of high accuracy in that 50% of the data points have errors smaller than ± 0.5 °C. A self-organizing map neural network model and a support vector machine based model have also been developed to estimate the upper ocean temperature anomalies and then derive the vertical temperature structure in the North Atlantic and the Indian Ocean [27,28]. Validation relative to the Argo observations shows reasonable accuracy within the depth of 500 m. As a matter of fact, a lot of efforts have been dedicated to enhancing the reconstruction accuracy using the random forest method, extreme gradient boosting and light gradient boosting machine, etc. [29–33]. Recently, the deep learning models built with the convolutional neural networks [34], the long short-term memory (LSTM) neural network [35] as well as the convolutional LSTM model [36] have become popular for reconstructing the three-dimensional temperature structure.

Although the above methods have caused progress concerning the estimation of the upper ocean temperatures, the impact of various physical processes on the reconstruction performance has not been widely accounted for. An attempt in such a direction has been reported using a pre-clustering neural network and taking the surface wind observation as additional input to diagnose the dynamic difference [37]. Ocean fronts are the regions where the properties of seawater (temperature, salinity, nutrients, etc.) change rapidly in the interior. This mesoscale phenomenon has a major impact on the physical, chemical and biological environment in the ocean [38–40]. Convergent downwelling is usually present at the location of ocean fronts and is accompanied by strong internal vertical mixing, which is significant to the regional marine biodiversity and biochemical processes. Correspondingly, the three-dimensional temperature structure of the ocean front region is usually different from the surroundings. It is thus necessary to induce the surface variable variations associated with ocean fronts into the reconstruction model, which are expected to increase the estimation accuracy. Such an aspect has been evidenced in [41] using the polynomial regression approaches over the North Pacific Ocean. The mapping bias is found much larger in the regions with higher SST gradient, representative of the ocean front intensity. A quasi-linear relation is also observed between the reconstruction bias/standard deviation and the SST gradient, which demonstrates the necessity of including the impact of ocean fronts in the models.

In this study, we focus on incorporating the ocean front features into the reconstruction model to improve the mapping accuracy. The satellite measured SST and SLA, in com-

bination with the Argo temperature profiles collected during 1999–2018, are employed. The dataset is divided into training and validation subsets to build the reconstruction models, with and without the input of the SST gradient. The organization of this paper is as follows. In Section 2, we present the data and pre-processing applied in this work. Section 3 details the establishment of our reconstruction models and experimental design. In Section 4, we show the results in comparison with Argo profile measurements. A discussion and the conclusion are presented in Section 5.

2. Data and Pre-Processing

2.1. Sea Surface Variables

As previously demonstrated, the SST and SLA are two common surface variables for the inputs of the reconstruction models. SST links directly to the ocean sub-surface temperatures while SLA is associated with the sea surface dynamic changes induced by the upper ocean thermal structure. Our recent experiments confirm that input of both SST and SLA achieves better accuracy in the vertical temperature estimation than feeding only SST to the developed models [41]. In addition, given the non-negligible impacts of ocean fronts on the vertical temperature profile estimation, the SST gradient is promising taken as an input into the reconstruction models along with SST and SLA.

The reanalyzed SST and SLA products from the Copernicus Marine Environmental Monitoring Service (CMEMS) are employed in this study. Both are available on a daily basis and the spatial resolution is $0.05^\circ \times 0.05^\circ$ and $0.25^\circ \times 0.25^\circ$ for SST and SLA, respectively. The reanalysis products are selected because of their operational incorporation of satellite observations and in situ measurements, as well as their strict quality control and monitoring. Despite the satellite observations being available at a higher spatial resolution, spatial gaps between tracks are usually large, for instance, several hundred km between two consecutive altimeter tracks [42]. This renders them not suitable for consistent collocation in time and space with Argo temperature profiles. The SST and SLA data within 1999–2018 are utilized in the following experiments and analysis. Their product identifiers are SST_GLO_SST_L4_REP_OBSERVATIONS_010_011 and SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047 within the framework of CMEMS. This 20-year dataset is sufficient to establish the model training and validation dedicated to reconstruction of the vertical temperature profiles.

Following the approach in [41], we quantify the intensity of ocean fronts through the SST gradient. The SST gradient has been found highly related to the estimation performance of the vertical temperature profiles. The reconstruction biases have been demonstrated to increase linearly with respect to the collocated SST gradient, implying the necessity of including this variable in the mapping function [41]. As such, similarly in this study, the typical Sobel operator is employed to estimate the SST gradient with its two components along the horizontal (longitudinal) and vertical (latitudinal) axis, given as follows:

$$D_x = \frac{1}{\Delta x} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot SST \quad (1)$$

$$D_y = \frac{1}{\Delta y} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \cdot SST \quad (2)$$

where Δx and Δy denote the spatial distance between adjacent pixels along the longitude and latitude, respectively. The SST gradient at a given pixel is then calculated by $|G| = \sqrt{D_x^2 + D_y^2}$.

2.2. Argo Profiles

At present, the largest publicly available database of upper ocean temperature profiles is collected and organized by the international and national Argo programs. An Argo float usually descends down to below the depth of 1000 m and drifts there for about 9 days, then rises up to the surface while collecting the profile of temperature and salinity. This operation mode results in the fact that the Argo data are both spatially and temporally sparse. The Argo data we use in this work are its original measurements, downloaded from <https://argo.ucsd.edu> (accessed on 15 August 2022).

We focus on the Northwest Pacific Ocean, as shown in Figure 1, where ocean fronts are persistent throughout the year along the Kuroshio and over the western subarctic area. Specifically, the region of interest extends in the range of [20°N, 50°N] and [120°E, 160°E]. Over this area in 1999–2018, there were 121,564 Argo temperature profiles in total being collected with the nearest SST and SLA in time and space. All Argo profiles are pre-processed to remove those with more than five outliers and shallower than 500 m. Note all the Argo measurements over the Sea of Japan are filtered out as the water properties and dynamics are quite different from the open oceans. Each Argo temperature profile is then interpolated to regular depths every 10 m in the range of [0 m, 500 m]. These processes finally end up with 94,065 profiles for training and independent testing of the reconstruction models.

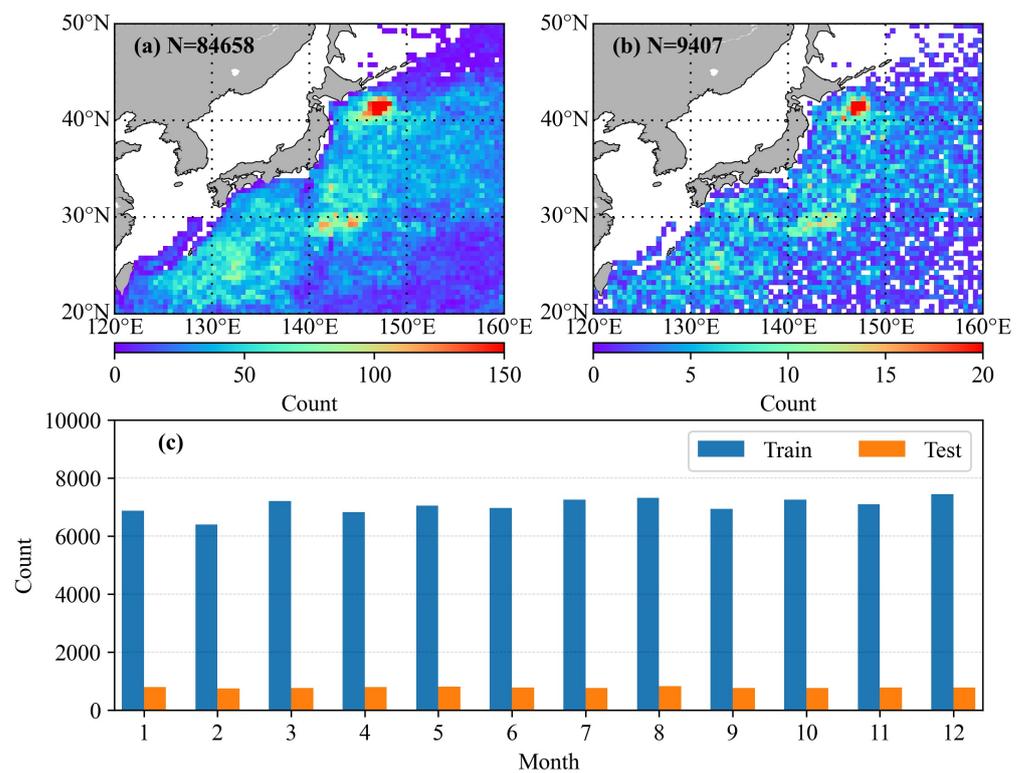


Figure 1. The data density of Argo profiles in 1999–2018 at a spatial bin of 0.5° in both latitude and longitude divided for (a) training and (b) independent testing. (c) The monthly data count for training and testing.

These Argo profiles are randomly divided into two parts: 90% (84,658 profiles) are fed into the model training and 10% (9,407 profiles) are used for model independent testing. Figure 1 illustrates the spatial data density for the two parts at a bin of 0.5° along both latitude and longitude. It is quite clear that the number of training profiles is not uniformly distributed; more than 150 profiles within one grid are accumulated off the Hokkaido islands of Japan, while one grid has approximately 20 profiles in the bottom right corner of the study region. This can be substantially attributed to the fact that Argo movements are

mostly controlled by ocean circulations. The Argo measurements are frequent over the area with strong ocean fronts. The uneven spatial distribution of Argo measurements in this area is expected to facilitate our model's development. In addition, the random dividing process brings no further data population issue as a similar Argo data distribution pattern is observed in Figure 1b for the validation part. Both training and validation datasets are uniform across the months, as presented in Figure 1c. There are roughly 7000 profiles per month for the training and 800 profiles for validation.

3. Reconstruction Model

3.1. Model Structure

Reconstruction of the vertical temperature profiles from the sea surface variables is basically a value-to-value mapping question. Regardless of the complex physics involved, the ocean water temperature at a certain depth is statically related, on the first order, to both SST and SLA. Our earlier exploitation verified that the relationship between SST and the ocean water temperature is nearly linear close to the sea surface and yet becomes non-linear as the depth increases. Therefore, reconstruction models are commonly developed based on the polynomial regressions, or more popular and efficient neural networks. In this study, we create our reconstruction models based on the fully connected neural networks.

A neural network is generally composed of one input and one output layer with several hidden layers in between. The neuron unit numbers of input and output layers are determined by the input and output variables. The hidden layers that contain an experimental number of units are key components to connect the dependent and independent variables. For a fully connected neural network, each layer unit is linked to all units of the neighbour layers through the so-called activation functions. The activation function defines how the variables are transformed within the neural network, where weights of the function are self-learning during the training procedure. There are many conceptual schematics available for the structure of neural network [43,44] and hence we would not repeat it in this paper. Yet, as a few experiments are conducted, we will detail the number of layers and units as well as the activation functions for each of our models.

In our case, the model inputs are SST, SLA and other variables plus with and without $|G|$ to examine whether the reconstruction of vertical temperature profiles can be improved. The model outputs are ocean water temperatures at regular depths, the same as processed Argo profile, i.e., every 10 m from 0 to 500 m. It is worth noting that, unlike the depthwise traditional regression methods, we establish our neural network reconstruction model to output the whole vertical temperature profile. For both training and model testing, we calculate two metrics to assess the model performance. They are root mean square error (RMSE) and mean absolute error (MAE), defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (4)$$

where y_i is the observed temperature profile and \hat{y}_i is the predicted profile.

3.2. Data Normalization

The inputs of SST, SLA and other variables have different measurement scales, which might result in a weighted contribution to the trainable weights of reconstruction models. To handle this, a max-min normalization procedure is usually employed to scale the input variables within the range of [0, 1]. Such a process has been proven not only to improve the computational efficiency but also to help lower the prediction deviations. In addition, the deviation between the ocean water temperature and the sea surface temperature differs with depth as shown in Figure 2. For the sake of convenience of comparison, here we

only present the distributions of temperature difference between sea surface and depth of 100 m, 200 m and 400 m, respectively. While the water temperature at 100 m depth is almost equivalent to SST, a certain number of data points are lower than SST by about 9 °C (blue curve). At a depth of 200 m, the temperature difference is consistently prevalent between 0 °C and 12 °C (orange curve). With depth decreasing to 400 m, the temperature difference enlarges and largely ranges from 4 °C to 16 °C (green curve). Given the same SST, the water temperature overall decreases with depth and thus might gain biased weights at the deeper levels during the model training process. Though the neural network is able to constrain the different extents of input variables, it will take longer training times to reach the steady state.

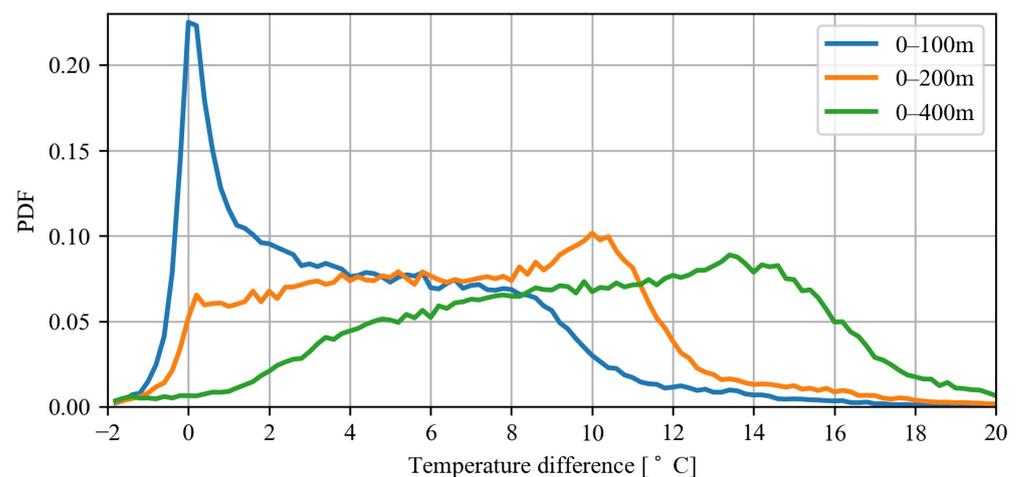


Figure 2. Probability density function (PDF) of temperature difference between the sea surface and at depth of 100 m, 200 m and 400 m, respectively.

To further assess the possible impact of normalization on the reconstruction model training, the following sensitivity experiment is carried out. The temperatures at the original scale and the normalized scale are designed as different model inputs for comparison. To exclude the influence of the model optimizer, ‘Sgd’ stands for stochastic gradient descent and ‘Adam’ for adaptive moment estimation that computes the data gradient to update the model parameter but this is tested in different ways. Note that the optimizer is an algorithm to change the model training attributes in order to reduce the model loss and speed up the training process (see [44] for details). These results in four models as listed in Table 1. All models are designed as four layers with two of them as hidden layers containing 12 and 6 units, respectively. The input layer takes the SST field and the output layer corresponding to the ocean water temperature at 150 m. As annotated in Table 1, we used a combination of ‘ReLU’ and ‘Linear’ as the activation function for all four models. Only the data within the range of 145°E–150°E and 35°N–40°N are used for training and testing of the models for computational efficiency. Such experiments are sufficient to determine whether the normalization is necessary or not.

Table 1. Model parameters of the sensitivity analysis.

Model	Normalization	Layers	Activation Function	Optimizer
Model 1	Not	1-12-6-1	R-R-L ¹	Sgd
Model 2	Not	1-12-6-1	R-R-L	Adam
Model 3	MinMaxScaler()	1-12-6-1	R-R-L	Sgd
Model 4	MinMaxScaler()	1-12-6-1	R-R-L	Adam

¹ R is ‘ReLU’ and L is ‘Linear’.

Figure 3a,b demonstrate the temperature profiles in their original scale and normalized scale to [0, 1] using the min-max scaling algorithm. It can be seen that the normalized

scale accounts well for the max-min contrast at each depth. For instance, the absolute temperature difference is comparable at the surface and 480 m for the two shown profiles, while their normalized scales deviate by 0.13. By inputting the original temperatures to model 1/2 and the normalized temperatures to model 3/4, the RMSE and MAE at each training epoch are given in the bottom panel of Figure 3. Overall, the trend of RMSE is quite similar to that of MAE. As we show values for every epoch, all four models exhibit comparable fluctuations. Both RMSE and MAE show that model 1 has the worst model performance, while models 2–4 reach to closing and relatively smaller values after the epoch of about 100. Yet, it is clear that the decreasing trend of both RMSE and MAE is the most rapid for model 4. Its turning point is at an epoch of 10, much shorter than 30 for model 2 and 70 for model 3. Thus, we can conclude that the normalization procedure and Adam optimizer would lead to a higher computational efficiency for training of the reconstruction models. Without particular statement, this configuration will be employed throughout the rest of this paper.

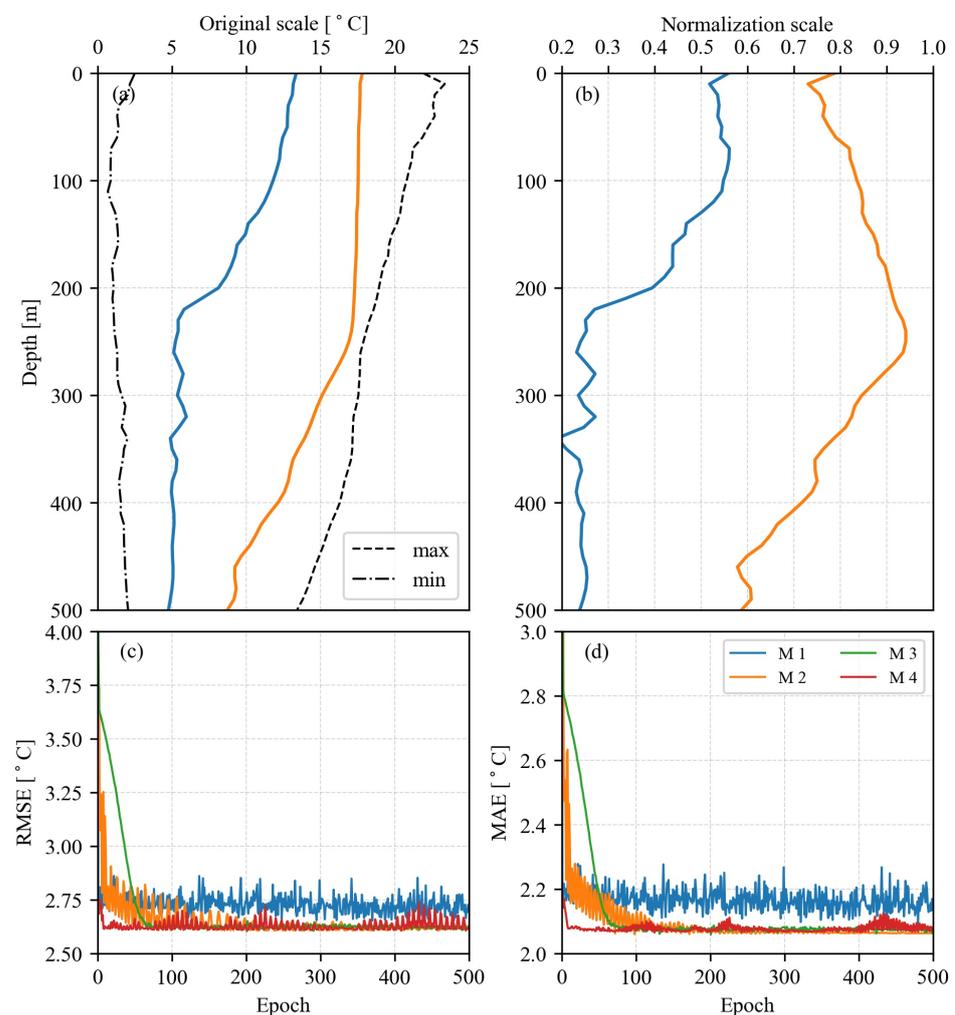


Figure 3. (Top) Two examples of vertical temperature profiles depicted by the blue and orange lines at (a) original scale and (b) normalized scale. Dashed and dotted lines in (a) represent the profile of maximum and minimum temperature, respectively. (Bottom) (c) RMSE and (d) MAE of training models relative to the epoch with the input of profiles at (a) original scale and (b) normalized scale.

3.3. Experiment Setup

The complete procedures of data processing and model training are illustrated as the flowchart in Figure 4. Six variables, including surface parameters (SST, SLA and SST gradient $|G|$), month and location (latitude and longitude) are taken as model inputs.

In particular, $|G|$ is flexible in the experiments to quantify its improvements concerning the reconstruction of vertical temperatures. Each variable is individually normalized using the min-max scaling algorithm within $[0, 1]$ to balance their weights in the training process. The input vertical temperature profiles are also normalized using the same algorithm at each water depth. The entire dataset is then randomly divided into two parts: 90% for the training and 10% for the independent testing. The training part is further grouped as model training (80%) and internal model validation (20%). Note that the internal model validation could be considered as a model self-evaluation and update process by consistently reducing the RMSE and/or MAE. The testing data is then used for external model validation by predicting temperature profiles based on the input variables and the trained deep learning models. It is worth remembering here that the output of the training model is on the normalized scale and an inversion is needed to recover the prediction profiles for comparison with the Argo measurements. In addition to accuracy metrics of MAE and RMSE, ΔT between reconstructed temperature and the corresponding Argo measurement at a given depth is also calculated in the subsequent analyses.

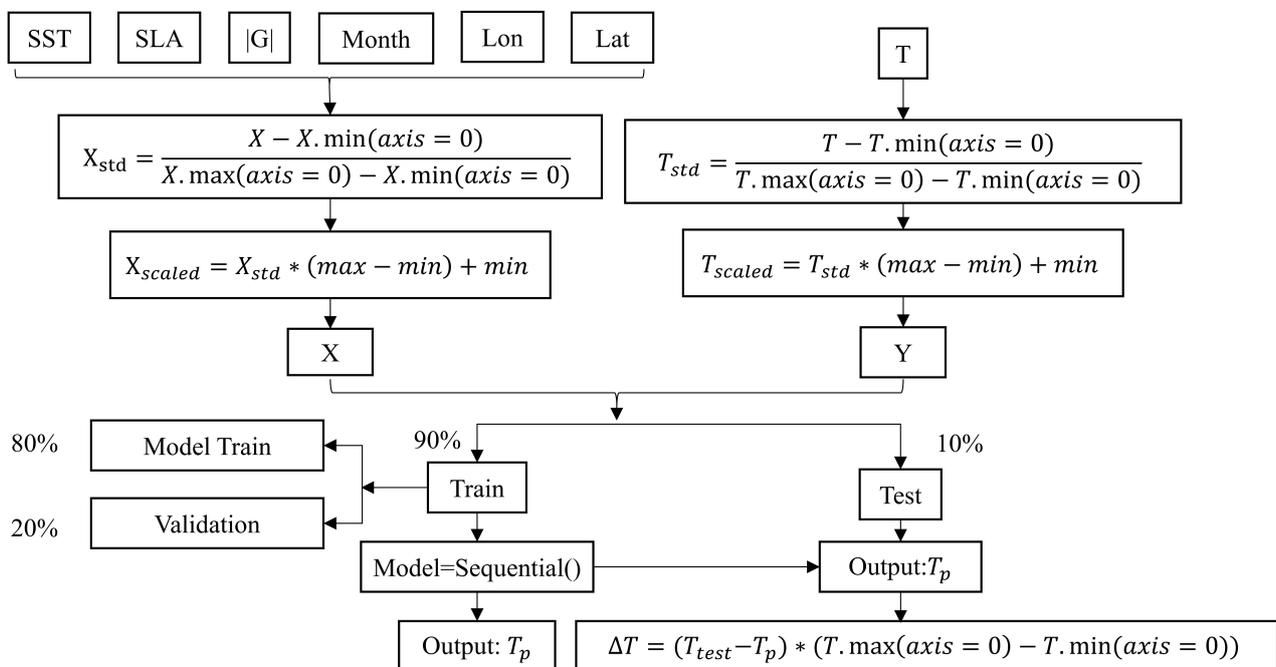


Figure 4. Flowchart of the data processing and reconstruction model training.

Table 2 lists the model details for our experiments examining whether including $|G|$ would improve the accuracy of the upper ocean temperature estimations or not. Moreover, to look for the most optimal model configuration with high accuracy, different model structures and activation functions are compared. The model learning rate is set as 0.0001, the batch size as 64 and the loss function as MAE after performing a couple of tests (not shown). For all the listed five models, SST, SLA, Month, Lon and Lat are taken as inputs along with with and without the SST gradient $|G|$. The output layer corresponds to the temperature profiles from the sea surface down to 500 m at a vertical resolution of 10 m, leading to 51 depths. Two representative model structures in particular are compared in this paper with two combinations of activation functions. As a matter of fact, we conducted hundreds of such similar model trainings and testings. Yet it is not necessary to document all these models as the selected one could fully represent our principal findings. In detail, the seven-layer model is designed to have neurons of 5/6, 8, 16, 32, 64, 128, 51 and the eight-layer model to have neurons of 5/6, 8, 16, 32, 64, 128, 512, 51. The linear and non-linear ReLU activation functions are mainly compared. First of all, the form of using ReLU

activation for all layers shows the lowest accuracy (not shown). For the regression problem, concerning a connection of the last two layers, it is suggested to use linear functions. As such, we propose two combinations of activation functions: linear activation function for all layers and ReLU activation function (ReLU) for the hidden layer and linear for the output layer.

Table 2. Reconstruction model details for experiments on improving the performance of upper ocean temperature estimations.

Model	Input	Structure	Activation Function
Model 5	SST, SLA, Month, Lon, Lat	$5 \times 8 \times 16 \times 32 \times 64 \times 128 \times 51$	L-L-L-L-L-L
Model 6	SST, SLA, Month, G , Lon, Lat	$6 \times 8 \times 16 \times 32 \times 64 \times 128 \times 51$	L-L-L-L-L-L
Model 7	SST, SLA, Month, G , Lon, Lat	$6 \times 8 \times 16 \times 32 \times 64 \times 128 \times 51$	R-R-R-R-R-L
Model 8	SST, SLA, Month, G , Lon, Lat	$6 \times 8 \times 16 \times 32 \times 64 \times 128 \times 512 \times 51$	R-R-R-R-R-L ¹
Model 9	SST, SLA, Month, Lon, Lat	$5 \times 8 \times 16 \times 32 \times 64 \times 128 \times 512 \times 51$	R-R-R-R-R-L

¹ R is 'ReLU' and L is 'Linear'.

4. Results

The five models 5–9 as listed in Table 2 are created and trained using the training dataset. Their RMSE and MAE calculated from the validation part at each training epoch are given in Figure 5. Note we only show the first 50 epochs by learning the fact in Figure 3 that these models can rapidly obtain a consistently low RMSE/MAE. While not shown, we would like to remind the reader that over-fitting appears with the training epoch. In general, it is clear that Model 5/6 with linear activation functions underperforms in comparison with the other three models that use ReLU activation function in the hidden layers. For the same model configuration, model 6 has relatively lower RMSE and MAE than model 5 since it takes the SST gradient |G| as input as well. At the 50th epoch, model 5 has RMSE of 0.0950 °C and MAE of 0.0726 °C, while the values for model 6 are 0.0919 °C and 0.0699 °C, respectively. This evidences the effect of |G| in improving the reconstruction model performance.

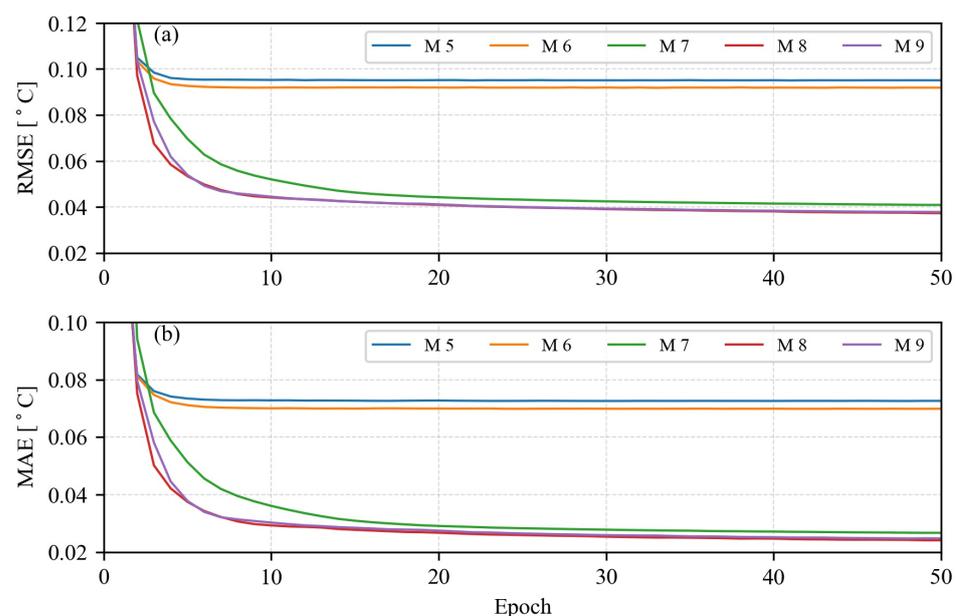


Figure 5. (a) RMSE and (b) MAE at each epoch for the five reconstruction models during training.

In comparison, model 7 differs from model 6 by using ReLU as activation functions for the hidden layers. Yet the enhancement of the model performance is significant. The RMSE and MAE of model 7 are much lower than that of model 5/6. For instance, RMSE and MAE

of model 7 reach $0.0373\text{ }^{\circ}\text{C}$ and $0.0242\text{ }^{\circ}\text{C}$ at the 50th epoch, respectively, which is almost an order of magnitude smaller than model 5/6. The improved reconstruction performance from model 5 to model 6 and then model 7 is further demonstrated in Figure 6, where the reconstructed and Argo observed temperature at two depths of (a–e) 50 m and (f–g) 250 m are scattered. The data density is counted at the bin of $0.5\text{ }^{\circ}\text{C}$ for both axes. At a depth of 50 m in the top panel, the data points distribute around the one-to-one line for all the three models. However, it is clear that the deviation of these data points from the diagonal line is larger for model 5/6 than model 7, particularly when the temperature is lower than $20\text{ }^{\circ}\text{C}$. Such features are more obvious at the depth of 250 m shown in the bottom panel of Figure 6. While the data points are clustered around a one-to-one line too, they look much tenser for model 7 than model 5/6. Another interesting aspect worth mentioning is that most of the temperature observations fall within the range of $[15\text{ }^{\circ}\text{C}, 30\text{ }^{\circ}\text{C}]$ at 50 m, while at 250 m, the temperature is either over the high or the low temperature end. This implies a great variability of deeper temperatures, possibly due to the appearance of mesoscale ocean dynamics that modify the thermocline structure. This analysis reveals that including $|G|$ and a non-linear activation function in the neural network advances the model performance in reconstructing the vertical temperatures.

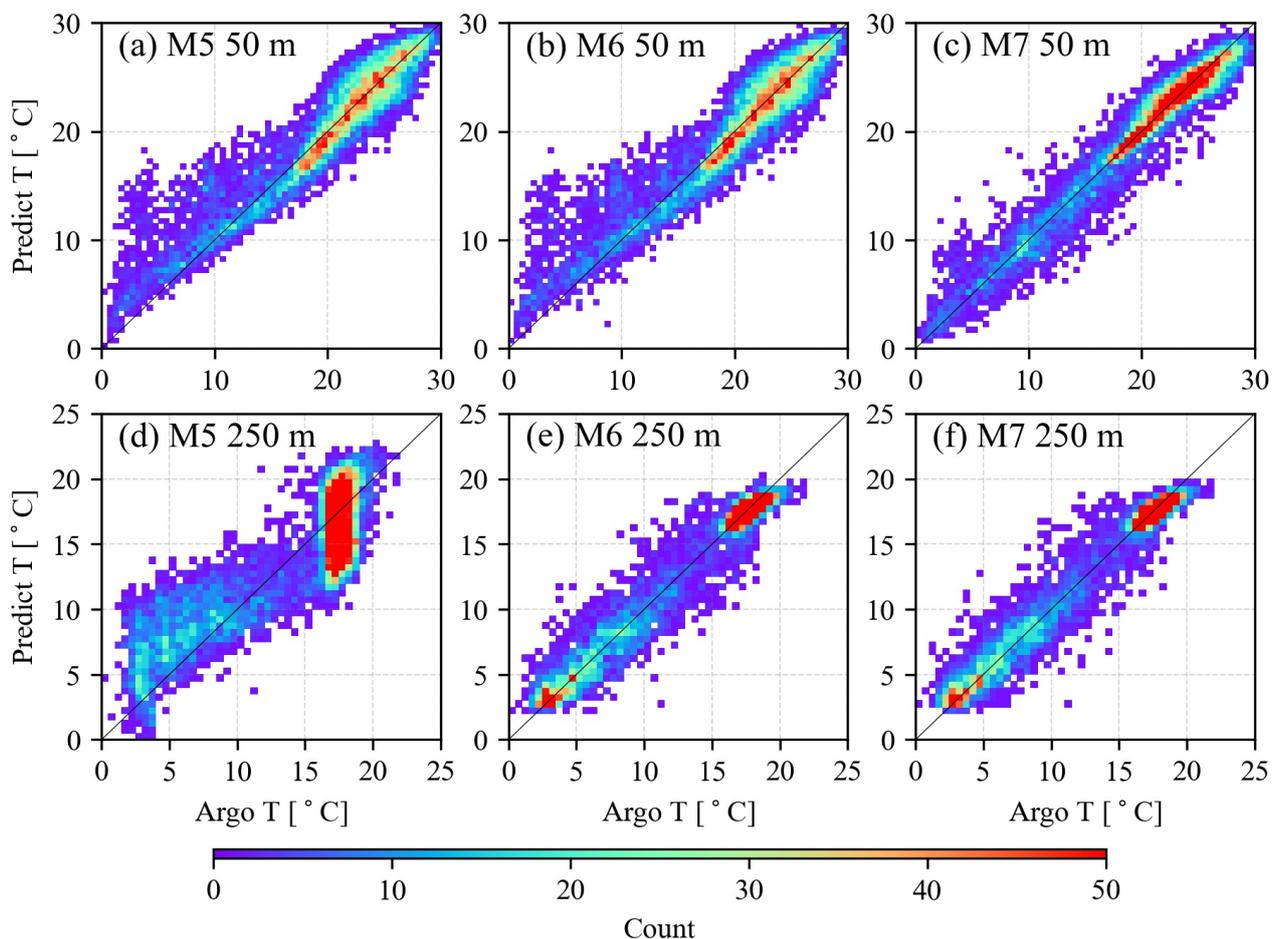


Figure 6. Q-Q plot of reconstructed temperature relative to Argo-measured temperature at two depths of (top) 50 m and (bottom) 250 m for (a,d) model 5, (b,e) model 6 and (c,f) model 7, respectively. Color represents the data density at a temperature bin of $0.5\text{ }^{\circ}\text{C}$ for both the x- and the y-axis.

Furthermore, the overall mean values of ΔT , RMSE and MAE at depths of 50 m and 250 m calculated between the Argo-measured and model-predicted temperatures for the

five models are given in Table 3, together with the correlation coefficient and normalized RMSE (also called scatter index, SI). The SI is defined as

$$SI = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i)^2} \quad (5)$$

For all parameters, model 5/6 is clearly larger than that of model 7/8/9 due to the linear activation function. While the ΔT of -0.0077 °C for model 7 at 50 m depth is the closest to 0 °C, model 8 attains the lowest RMSE and MAE. At a depth of 250 m, model 8 performs better than other models according to all the five accuracy metrics. These indicate that model 8 has the best performance among our experimental models. This model is designed with one more layer than model 7, leading to a relative improvement of reconstruction in ΔT of roughly 75% at a depth of 250 m. It is worth mentioning that we have trained models of more than eight layers and the accuracy does not dramatically increase. The model structure of eight layers is utilized in the following analysis for the sake of the computation time balance.

Table 3. The accuracy metrics calculated between the Argo-measured and model-predicted temperatures at depths of 50 m and 250 m for the five models listed in Table 2. The best performance metrics are highlighted by the underlined values.

Depth		50 m					250 m				
Model	ΔT [°C]	RMSE [°C]	MAE [°C]	R^2	SI [%]	ΔT [°C]	RMSE [°C]	MAE [°C]	R^2	SI [%]	
M5	-0.345	2.271	1.617	0.879	1.2	0.053	2.677	2.167	0.754	3.1	
M6	-0.222	2.232	1.600	0.883	1.2	0.069	2.565	2.055	0.774	2.9	
M7	<u>0.008</u>	1.525	1.091	0.945	0.5	0.027	1.048	0.669	0.962	0.5	
M8	-0.045	<u>1.337</u>	<u>0.892</u>	<u>0.958</u>	<u>0.4</u>	<u>0.015</u>	<u>0.993</u>	<u>0.623</u>	<u>0.966</u>	<u>0.4</u>	
M9	-0.059	1.344	0.920	0.957	0.4	-0.054	1.006	0.626	0.965	0.4	

After removing $|G|$ from the inputs, the ΔT , RMSE and MAE of model 9 increase by 0.014 °C, 0.007 °C and 0.028 °C relative to model 8 at 50 m depth. A similar trend is observed for 250 m depth. The performance difference between these two models is further compared at each depth for the ΔT , RMSE and MAE in Figure 7. The result of model 7 is also included as a reference and shows clearly larger MAE and RMSE as well as variations in ΔT . The metrics of RMSE and MAE between model 8 and model 9 are very close as shown in Figure 5. As an alternative, we look into the ΔT (reconstructed Argo) as in Figure 7a, which illustrates noticeable distinctions. At depths between 0 and 100 m, the ΔT curve of model 8 is comparable to that of model 9. Both decrease from 0.1 °C at sea surface to about -0.08 °C at 30 m and then increase to 0 °C at 100 m. Yet, at increasing depths, model 9 ΔT becomes positive and gradually decreases around -0.1 °C below 300 m. This is in contrast to model 8 ΔT which stays at around 0 °C below 100 m. A detailed comparison between model 8 and model 9 for the ΔT between the Argo-measured and model-predicted temperatures is given in Figure 8 for the depths of 50 m, 250 m and 400 m, respectively. At 50 m depth, the ΔT distribution of model 8 is nearly the same as model 9 with both mean values of around -0.05 °C. The negative biases are expected given the curves in Figure 7a. With increasing depth, model 9 ΔT tends to be more negative while model 8 ΔT stays around 0 °C. All the above results and analyses prove that it is necessary to include the SST gradient $|G|$ on account of the ocean front impact to the reconstruction of vertical temperature profiles.

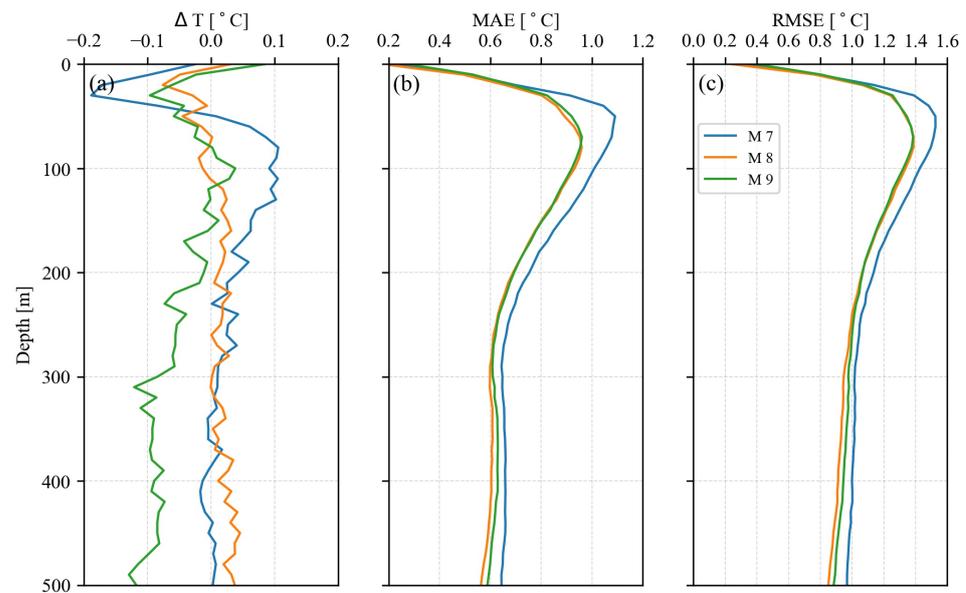


Figure 7. The averaged (a) ΔT , (c) RMSE and (b) MAE between the Argo-measured and model-predicted temperatures at all depths for model 7/8/9.

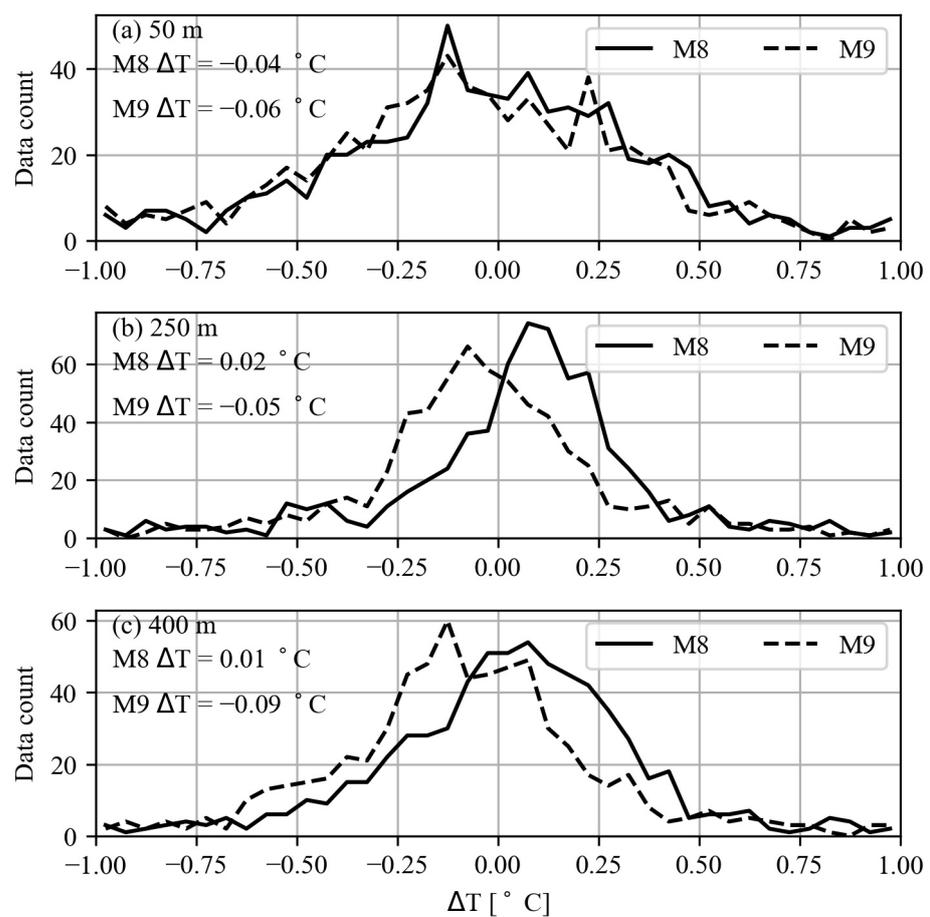


Figure 8. Distributions of ΔT between the Argo-measured and model-predicted temperatures at three depths of (a) 50 m, (b) 250 m and (c) 400 m for model 8/9, respectively. The temperature bin is $0.05\text{ }^{\circ}\text{C}$ and the mean ΔT is annotated in each subplot.

5. Discussion and Conclusions

The experiments conducted in this study prove that the accuracy of reconstruction of vertical temperature profiles from sea surface variables can be well refined by adding the SST gradient $|G|$ as input. To a certain extent, $|G|$ quantifies the intensity of ocean fronts and thus accounts for their impact on the upper ocean temperature structures [41]. Yet, we must acknowledge that $|G|$ is not the only parameter to characterize the ocean fronts [45]. A full description of ocean fronts with multiple features is likely to advance the mapping skill from sea surface observations to depthwise temperatures. In addition, the mesoscale ocean processes that bring about different sea surface traces are more than ocean fronts [1,4]. Analysis of these phenomena, such as ocean eddies, internal waves, upwellings, etc., in vertical profile mapping needs additional efforts.

In this study, the necessity of taking ocean dynamics into account to link the sea surface fields with sub-surface structures is evidenced. By including $|G|$, the improvement of temperature estimation at 250 m depth is about 10% compared to the values obtained using the regression method in [41]. Such progress is expected to be further refined with more ocean processes included in the reconstruction models in the future. However, more importantly, this technique leads to a practical way of introducing the physics of oceanic phenomena into data-driven statistical methods. As a matter of fact, the combination of data-driven and physic-driven methods has seen great advantages [14]. The consistent mining into the data to extract more geophysical mechanisms is becoming increasingly significant.

In our case, to establish connections between the surface variables and temperatures profiles, we conducted many experiments and detailed several representative models in this paper. When using the linear activation functions throughout all the layers, the neural network is equal to high-order polynomial regressions, which show obviously coarser performance compared to the non-linear models (see Figure 5 and Table 3). A deeper network would facilitate the model performance as well, which sacrifices the efficiency with the increasing number of layers. The neural network structure depends on the complexity of problems and the size of the training dataset. For the reconstruction of vertical temperature profiles in this study, an eight-layer model is sufficient to attain solid model performance. Machine learning methods develop quickly and considerable possibilities in improving the reconstruction of vertical temperature profiles are therefore foreseen in the near future. The motivation of this work is to exploit new thoughts and more machine learning methods such as the convolutional neural network shall be further employed.

As a follow-on to our previous study which demonstrates the dependence of reconstruction model performance on the associated $|G|$, the accuracy of vertical temperature reconstruction models using SST, SLA, $|G|$, Month, LON and LAT as inputs is higher than that of models without $|G|$. The obtained improvements are more pronounced under the depth of 200 m. Figure 9 presents an example of the three-dimensional reconstructed temperature on 1 January 2010 over the area of interest predicted by model 8 of the best performance. On a daily basis, SST exhibits highly dynamic variability with rich sub-mesoscale and mesoscale processes. It has to be admitted that the front impact can not be fully resolved by only taking the SST gradient as a single additional input. As shown, the vertical temperature structure is a field of wide extent and subject to the impact of surrounding neighbors. Further explorations considering the environmental conditions are one way to improve the reconstruction accuracy. Now the complicated and deep convolutional neural network (CNN) is ready to use by extracting the temporal and spatial features of input parameters for higher mapping performance.

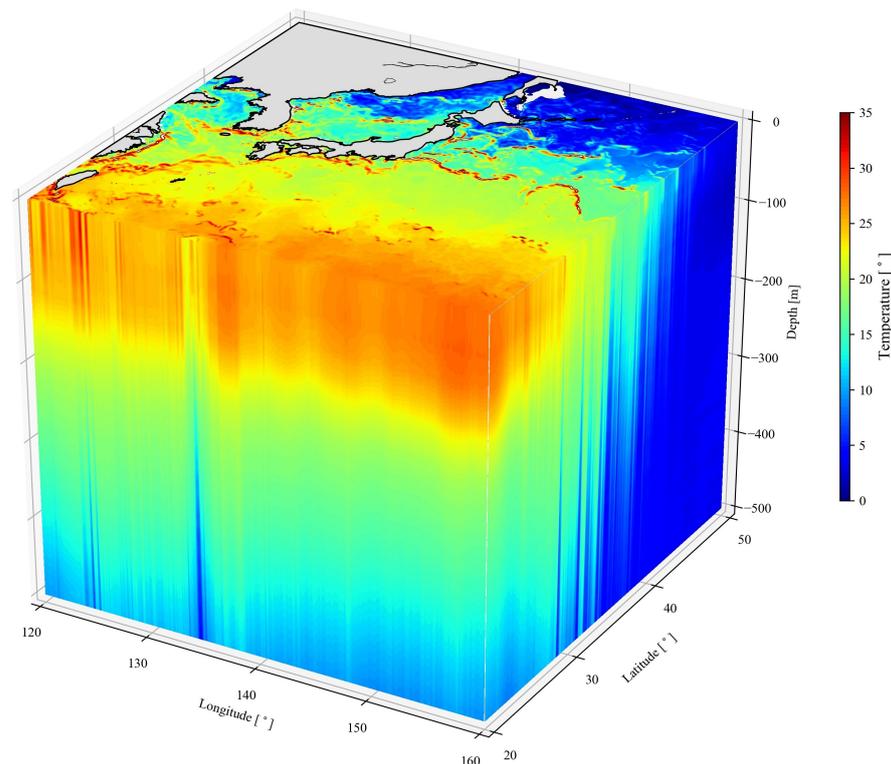


Figure 9. Three-dimensional representation of the reconstructed temperature using model 8 with input of daily SST and SLA on 1 January 2010.

Author Contributions: Conceptualization, C.W. and H.L.; methodology, X.C.; validation, C.W. and X.C.; formal analysis, X.C.; writing—original draft preparation, X.C. and H.L.; writing—review and editing, C.W. and Y.H.; supervision, Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Jiangsu Province under grant of BK20210666, the National Natural Science Foundation of China under Grant of 42006163 and 42027805, the Startup Foundation for Introducing Talent of Nanjing University of Information Science & Technology (NUIST).

Data Availability Statement: Sea surface temperature data are publicly available on E.U. CMEMS via <https://marine.copernicus.eu/> (accessed on 15 August 2022) and Argo profiles are available from <http://www.argodatamgt.org/> (accessed on 15 August 2022).

Acknowledgments: The authors would like to thank E.U. CMEMS for providing the SST and SLA products and the International and National Argo programs for providing the vertical temperature profiles. We also acknowledge the High Performance Computing Center of Nanjing University of Information Science & Technology for their support of this work.

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

References

- Garrett, C.; Munk, W. Internal Waves in the Ocean. *Annu. Rev. Fluid Mech.* **1979**, *11*, 339–369. [\[CrossRef\]](#)
- Sabinin, K.; Serebryany, A. Intense short-period internal waves in the ocean. *J. Mar. Res.* **2005**, *63*, 227–261. [\[CrossRef\]](#)
- Siegelman, L.; Klein, P.; Thompson, A.F.; Torres, H.S.; Menemenlis, D. Altimetry-Based Diagnosis of Deep-Reaching Sub-Mesoscale Ocean Fronts. *Fluids* **2020**, *5*, 145. [\[CrossRef\]](#)
- Wang, J.; Bethel, B.J.; Dong, C.; Li, C.; Cao, Y. Numerical Simulation and Observational Data Analysis of Mesoscale Eddy Effects on Surface Waves in the South China Sea. *Remote Sens.* **2022**, *14*, 1463. [\[CrossRef\]](#)
- Jayne, S.R.; Roemmich, D.; Zilberman, N.; Riser, S.C.; Johnson, K.S.; Johnson, G.C.; Piotrowicz, S.R. Present and Future. *Oceanography* **2017**, *30*, 18–28. [\[CrossRef\]](#)
- Merchant, C.J.; Embury, O.; Bulgin, C.E.; Block, T.; Corlett, G.K.; Fiedler, E.; Good, S.A.; Mittaz, J.; Rayner, N.A.; Berry, D.; et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* **2019**, *6*, 223. [\[CrossRef\]](#)

7. Nerem, R.S.; Beckley, B.D.; Fasullo, J.T.; Hamlington, B.D.; Masters, D.; Mitchum, G.T. Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2022–2025. [[CrossRef](#)]
8. Reul, N.; Grodsky, S.; Arias, M.; Boutin, J.; Catany, R.; Chapron, B.; D’Amico, F.; Dinnat, E.; Donlon, C.; Fore, A.; et al. Sea surface salinity estimates from spaceborne L-band radiometers: An overview of the first decade of observation (2010–2019). *Remote Sens. Environ.* **2020**, *242*, 111769. [[CrossRef](#)]
9. Carnes, M.R.; Mitchell, J.L.; de Witt, P.W. Synthetic temperature profiles derived from Geosat altimetry: Comparison with air-dropped expendable bathythermograph profiles. *J. Geophys. Res. Ocean.* **1990**, *95*, 17979–17992. [[CrossRef](#)]
10. Carnes, M.R.; Teague, W.J.; Mitchell, J.L. Inference of Subsurface Thermohaline Structure from Fields Measurable by Satellite. *J. Atmos. Ocean. Technol.* **1994**, *11*, 551–566. [[CrossRef](#)]
11. 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](#)]
12. Wang, J.; Flierl, G.R.; LaCasce, J.H.; McClean, J.L.; Mahadevan, A. Reconstructing the Ocean’s Interior from Surface Data. *J. Phys. Oceanogr.* **2013**, *43*, 1611–1626. [[CrossRef](#)]
13. Liu, L.; Xue, H.; Sasaki, H. Reconstructing the Ocean Interior from High-Resolution Sea Surface Information. *J. Phys. Oceanogr.* **2019**, *49*, 3245–3262. [[CrossRef](#)]
14. 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](#)]
15. Moore, A.M.; Martin, M.J.; Akella, S.; Arango, H.G.; Balmaseda, M.; Bertino, L.; Ciavatta, S.; Cornuelle, B.; Cummings, J.; Frolov, S.; et al. Synthesis of Ocean Observations Using Data Assimilation for Operational, Real-Time and Reanalysis Systems: A More Complete Picture of the State of the Ocean. *Front. Mar. Sci.* **2019**, *6*, 90. [[CrossRef](#)]
16. Su, H.; Zhang, T.; Lin, M.; Lu, W.; Yan, X.H. Predicting subsurface thermohaline structure from remote sensing data based on long short-term memory neural networks. *Remote Sens. Environ.* **2021**, *260*, 112465. [[CrossRef](#)]
17. Guinehut, S.; Dhomp, A.L.; Larnicol, G.; Traon, P.Y.L. High resolution 3-D temperature and salinity fields derived from in situ and satellite observations. *Ocean Sci.* **2012**, *8*, 845–857. [[CrossRef](#)]
18. Liu, L.; Peng, S.; Wang, J.; Huang, R.X. Retrieving density and velocity fields of the ocean’s interior from surface data. *J. Geophys. Res. Ocean.* **2014**, *119*, 8512–8529. [[CrossRef](#)]
19. Yan, X.H.; Jo, Y.H.; Liu, W.T.; He, M.X. A New Study of the Mediterranean Outflow, Air–Sea Interactions and Meddies Using Multisensor Data. *J. Phys. Oceanogr.* **2006**, *36*, 691–710. [[CrossRef](#)]
20. Yan, X.H.; Okubo, A. Three-dimensional analytical model for the mixed layer depth. *J. Geophys. Res. Ocean.* **1992**, *97*, 20201–20226. [[CrossRef](#)]
21. Lapeyre, G.; Klein, P. Dynamics of the Upper Oceanic Layers in Terms of Surface Quasigeostrophy Theory. *J. Phys. Oceanogr.* **2006**, *36*, 165–176. [[CrossRef](#)]
22. Fox, D.N.; Teague, W.J.; Barron, C.N.; Carnes, M.R.; Lee, C.M. The Modular Ocean Data Assimilation System (MODAS). *J. Atmos. Ocean. Technol.* **2002**, *19*, 240–252. [[CrossRef](#)]
23. Maes, C.; Behringer, D.; Reynolds, R.W.; Ji, M. Retrospective Analysis of the Salinity Variability in the Western Tropical Pacific Ocean Using an Indirect Minimization Approach. *J. Atmos. Ocean. Technol.* **2000**, *17*, 512–524. [[CrossRef](#)]
24. Nardelli, B.B.; Santoleri, R. Methods for the Reconstruction of Vertical Profiles from Surface Data: Multivariate Analyses, Residual GEM and Variable Temporal Signals in the North Pacific Ocean. *J. Atmos. Ocean. Technol.* **2005**, *22*, 1762–1781. [[CrossRef](#)]
25. Su, H.; Huang, L.; Li, W.; Yang, X.; Yan, X.H. Retrieving Ocean Subsurface Temperature Using a Satellite-Based Geographically Weighted Regression Model. *J. Geophys. Res. Ocean.* **2018**, *123*, 5180–5193. [[CrossRef](#)]
26. Ali, M.M.; Swain, D.; Weller, R.A. Estimation of ocean subsurface thermal structure from surface parameters: A neural network approach. *Geophys. Res. Lett.* **2004**, *31*. [[CrossRef](#)]
27. Wu, X.; Yan, X.H.; Jo, Y.H.; Liu, W.T. Estimation of Subsurface Temperature Anomaly in the North Atlantic Using a Self-Organizing Map Neural Network. *J. Atmos. Ocean. Technol.* **2012**, *29*, 1675–1688. [[CrossRef](#)]
28. 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](#)]
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. Su, H.; Yang, X.; Lu, W.; Yan, X.H. Estimating Subsurface Thermohaline Structure of the Global Ocean Using Surface Remote Sensing Observations. *Remote Sens.* **2019**, *11*, 1598. [[CrossRef](#)]
31. Su, H.; Wang, A.; Zhang, T.; Qin, T.; Du, X.; Yan, X.H. Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102440. [[CrossRef](#)]
32. Meng, L.; Yan, C.; Zhuang, W.; Zhang, W.; Geng, X.; Yan, X.H. Reconstructing High-Resolution Ocean Subsurface and Interior Temperature and Salinity Anomalies From Satellite Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [[CrossRef](#)]
33. Zuo, X.; Zhou, X.; Guo, D.; Li, S.; Liu, S.; Xu, C. Ocean Temperature Prediction Based on Stereo Spatial and Temporal 4-D Convolution Model. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
34. Bolton, T.; Zanna, L. Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization. *J. Adv. Model. Earth Syst.* **2019**, *11*, 376–399. [[CrossRef](#)]

35. Nardelli, B.B. A Deep Learning Network to Retrieve Ocean Hydrographic Profiles from Combined Satellite and In Situ Measurements. *Remote Sens.* **2020**, *12*, 3151. [[CrossRef](#)]
36. Su, H.; Jiang, J.; Wang, A.; Zhuang, W.; Yan, X.H. Subsurface Temperature Reconstruction for the Global Ocean from 1993 to 2020 Using Satellite Observations and Deep Learning. *Remote Sens.* **2022**, *14*, 3198. [[CrossRef](#)]
37. Lu, W.; Su, H.; Yang, X.; Yan, X.H. Subsurface temperature estimation from remote sensing data using a clustering-neural network method. *Remote Sens. Environ.* **2019**, *229*, 213–222. [[CrossRef](#)]
38. Fevre, J.L. Aspects of the Biology of Frontal Systems. *Adv. Mar. Biol.* **1987**, *23*, 163–299. [[CrossRef](#)]
39. Olson, D.; Hitchcock, G.; Mariano, A.; Ashjian, C.; Peng, G.; Nero, R.; Podest, G.; Hjort, J. Life on the Edge: Marine Life and Fronts. *Oceanography* **1994**, *7*, 52–60. [[CrossRef](#)]
40. Belkin, I.M.; Cornillon, P.C.; Sherman, K. Fronts in Large Marine Ecosystems. *Prog. Oceanogr.* **2009**, *81*, 223–236. [[CrossRef](#)]
41. Chen, X.; Wang, C.; Li, H.; Hu, D.; Chen, C.; He, Y. Impact of ocean fronts on the reconstruction of vertical temperature profiles from sea surface measurements. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **2022**, *187*, 103833. [[CrossRef](#)]
42. Badulin, S.; Kostianoy, A.; Shabanov, P.; Sharmar, V.; Grigorieva, V.; Lebedev, S. Self- and Inter-Crossover Points of Jasons' Missions as New Essential Add-on of Satellite Altimetry in the Sub-Arctic Seas and the Southern Ocean. *Remote Sens.* **2021**, *13*, 658. [[CrossRef](#)]
43. Leung, F.; Lam, H.; Ling, S.; Tam, P. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Trans. Neural Netw.* **2003**, *14*, 79–88. [[CrossRef](#)] [[PubMed](#)]
44. Chen, J.C.; Zhou, D.Z.; Tang, Y.; Yang, Z.; Cao, Y.; Gu, Q. Closing the Generalization Gap of Adaptive Gradient Methods in Training Deep Neural Networks. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), Yokohama, Japan, 7–15 July 2021; pp. 3267–3275.
45. Xi, J.; Wang, Y.; Feng, Z.; Liu, Y.; Guo, X. Variability and Intensity of the Sea Surface Temperature Front Associated with the Kuroshio Extension. *Front. Mar. Sci.* **2022**, *9*. [[CrossRef](#)]