

Spatial–Temporal Siamese Convolutional Neural Network for Subsurface Temperature Reconstruction

Shuyu Zhang¹, Yizhou Yang, Kangwen Xie, Jiahao Gao, Zhiyuan Zhang, Qianru Niu², Gongjie Wang, Zhihui Che, Lin Mu, and Sen Jia³, *Senior Member, IEEE*

Abstract—The reconstruction of subsurface ocean temperature using sea surface observations and in situ Argo measurements is an important yet challenging task. The availability of long-term and high-resolution sea surface remote sensing, combined with advancements in deep learning technology, has opened new opportunities for studying subsurface temperature (ST) reconstruction. In this study, a novel spatial–temporal Siamese convolutional neural network (SSCNN) is proposed to improve the accuracy of ST reconstruction in the Indian Ocean. First, considering the distinctions of temperature characteristics among different sea areas, a multiscale division scheme based on the correlation coefficient of integral ST is designed for refined reconstruction modeling. Second, since ocean heat is significantly affected by solar radiation, asymmetric convolutional operation with rectangular patches and kernels is designed to capture the information characteristics in longitude and latitude directions, respectively. Third, given the temporal changes and correlations of ocean temperature, an SSCNN with shared parameters is proposed for multiview feature mining and accurate temperature structure reconstruction. The reconstructed results provide a precise depiction of the subsurface Indian Ocean dipole (sub-IOD)’s evolution, including the spatial distribution of positive and negative anomaly signals and its temporal changes. It demonstrates that the subsurface dipole index series obtained from SSCNN reconstruction is consistent with that from International Pacific Research Center (IPRC) observation, remaining within a reasonable error range. Comparative experiments indicate that the SSCNN model surpasses other existing methods in terms of

higher accuracy and smaller error. Overall, this study provides a promising approach for effectively reconstructing the ST using deep learning methods and offers valuable insights for analyzing the evolution of subsurface positive dipole in Indian Ocean.

Index Terms—Indian ocean, remote sensing, Siamese convolutional neural network (CNN), spatial–temporal feature extraction, subsurface temperature anomaly (STA).

I. INTRODUCTION

THE issue of global warming has been widely discussed for years [1]. To gain a deeper scientific understanding of global warming, it is essential to focus on the ocean, which constitutes about 71% of the Earth’s surface area and plays a crucial role in the climate system [2]. It is imperative to enhance the research on physical properties of ocean subsurface, which is vital to comprehend the internal dynamic processes related to ocean. Studies have shown that temperature and salinity are the primary physical quantities to examine. The ocean is the global energy hub, with about 93% of the Earth’s excess heat stored in it [3]. The global warming has resulted in a significant increase in ocean heat content, particularly in the deep ocean [4], [5]. As heat continues to accumulate, the ocean heat content reached a record high in 2021 [6], [7], [8]. Therefore, obtaining extensive and long-term data on ocean subsurface temperature (ST) is crucial in improving our understanding of the ocean’s internal dynamic processes.

With the advancement of satellite remote sensing technology, ocean surface observation data have become increasingly available, offering high-resolution and long-time data [9], [10]. However, this technology is unable to observe the ocean subsurface, as electromagnetic radiation cannot penetrate deep enough [11]. Although the global subsurface observation data have gradually increased since the Argo program began in 2004 [12], the number of profiling floats is still inadequate for full ocean coverage. This limits the ability to capture small- and medium-scale signals in subsurface and the understanding of ocean’s internal mechanisms [13]. Studies have shown that ocean internal activities can manifest as surface phenomena, and vice versa [14]. Therefore, it is possible to use multisource sea surface remote sensing data and Argo subsurface observation data to reconstruct the vital subsurface dynamic information, such as the thermohaline structure [15].

Manuscript received 18 April 2023; revised 29 November 2023; accepted 23 December 2023. Date of publication 1 January 2024; date of current version 10 January 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62201354, Grant 42106007, and Grant 62271327; in part by the Shenzhen Science and Technology Program under Grant KCXFZ20211020164015024, Grant RCJC20221008092731042, Grant JCYJ20220818100206015, and Grant KQTD20200909113951005; in part by the Project of Department of Education of Guangdong Province under Grant 2023KCXTD029; and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515011290. (Corresponding authors: Zhihui Che; Qianru Niu.)

Shuyu Zhang, Yizhou Yang, Kangwen Xie, Jiahao Gao, and Sen Jia are with the College of Computer Science and Software Engineering, the Guangdong–Hong Kong–Macau Joint Laboratory for Smart Cities, and the Key Laboratory for Geo-Environmental Monitoring of Coastal Zone, Ministry of Natural Resources, Shenzhen University, Shenzhen 518060, China (e-mail: shuyu-zhang@szu.edu.cn; 847007730@qq.com; 2945045615@qq.com; 1305663244@qq.com; senjia@szu.edu.cn).

Zhiyuan Zhang is with the 91001 Unit of PLA, Beijing 100080, China (e-mail: zhangzhiyuan@meri.org.cn).

Qianru Niu, Zhihui Che, and Lin Mu are with the College of Life Sciences and Oceanography, Shenzhen University, Shenzhen 518060, China (e-mail: qniu@szu.edu.cn; chezh@szu.edu.cn; mulin@szu.edu.cn).

Gongjie Wang is with the PLA 31526 Troops, Beijing 100080, China (e-mail: wanggj_9015@sina.com).

Digital Object Identifier 10.1109/TGRS.2023.3348653

1558-0644 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

Most sea surface parameters have high spatial and temporal resolution, making it feasible to reconstruct a high-resolution ST structure, and this approach has emerged as a current research hotspot [16].

The reconstruction of ocean surface temperature based on traditional machine learning algorithms can usually obtain good results and stable model performance [17], [18], [19], [20]. Related works involve the methods of support vector machine (SVM) [21], random forest (RF) [22], [23], and extreme gradient boosting (XGBoost) [15], [24]. Nonetheless, due to the model limitations of these methods, they generally rely on the prior knowledge and lack the deep ability to extract spatial-temporal characteristics of sea surface information. Although the integrated algorithms such as RF [22], [23] and XGBoost [15] enhance the generalization ability by combining results of multiple weak learners, they still do not fully mine the semantic features of remote sensing input. With the rapid development of artificial intelligence, deep learning algorithms have shown strong adaptability to fit the complex relationships and learn the high-level features automatically [25]. Through the end-to-end learning paradigm, deep models can identify the patterns hidden in ocean big data for better reconstruction of ST [26]. Related works involve the methods of long short-term memory (LSTM) network [27], [28], [29], convolutional neural network (CNN) [30], [31], [32], and their integration [33].

First, Buongiorno Nardelli [27] proposed a technique using stacked LSTM network and Monte Carlo dropout to provide hydrographic vertical profiles over the North Atlantic Ocean. Su et al. [28], [29] built a bidirectional LSTM (Bi-LSTM) network for predicting global ocean ST and salinity anomalies, and then reconstructed a long time-series ocean heat content dataset and filled the pre-Argo data gaps. These studies established unified reconstruction models for large-scale ocean areas to conduct underwater predictions. However, the spatial and temporal characteristics of sea surface and subsurface variables are not the same among different regions, so direct modeling of the entire sea area would reduce the accuracy of feature extraction and subsurface reconstruction. Second, Han et al. [30] proposed a CNN method to estimate ST in the Pacific Ocean from a suite of satellite remote sensing measurements. Meng et al. [31], [32] retrieved ST and salinity anomalies in the Pacific Ocean at high resolution using CNN model, and also estimated ocean subsurface and interior variables for dynamic height fields and ocean geostrophic flows. Nonetheless, the size and shape of convolution kernels for CNN spatial feature learning are usually fixed in these studies, which do not correspondingly align with the contextual variation patterns of sea surface temperature (SST) and other variables in latitude (LAT) and longitude (LON) directions. Third, Zhang et al. [33] comprised the CNN, LSTM, and multiple-layer stacking to consider the horizontal and vertical temperature variations from sea surface to subsurface, which considered both temporal and spatial information for ST calculation. However, its extraction of complicated spatial and temporal characteristics of multisource remote sensing is insufficient, and the multiview and effective feature fusion of sea surface variables for ST reconstruction is lacked.

The Indian Ocean is the third largest ocean out of the world's oceans, and it holds a crucial position in the global climate system [34]. Due to its location primarily within the tropics, it is also the warmest ocean on the Earth [35], [36]. However, the Indian Ocean is significantly influenced by the monsoon climate, and anomaly changes in SSTs can have serious impacts on the surrounding. Indian Ocean dipole is a crucial phenomenon of sea surface temperature anomaly (SSTA), which is driven by the cooling of SST in the Southeast of tropical Indian Ocean [37], [38], [39]. Given the vital role of Indian Ocean, this study aims to reconstruct its ST structure within the areas North of 35°S above 1000-m underwater, using the proposed deep learning model that combines multi-source sea surface remote sensing and Argo observation data. The study area is restricted to a specific region, spanning from 25°E to 120°E and from 35°S to 30°N, to achieve a high-resolution and high-precision ST reconstruction.

In this study, we propose a novel spatial-temporal Siamese CNN (SSCNN) to improve the accuracy of ST reconstruction in the Indian Ocean. First, considering the observation limitations in coastal areas and various marine processes in noncoastal areas, the sea area is divided into multiple regions for spatial-temporal extraction and multiscale reconstruction, according to the correlation coefficient of integral ST. Second, the asymmetric convolutional calculation with rectangular patches and kernels is proposed to flexibly model the significant changes with LAT and slight changes with LON, as solar radiation mainly affects the spatial distribution of temperature. Third, due to the correlation between sea surface remote sensing variables of adjacent months, an SSCNN with shared parameters is designed to extract time-series features in parallel and fuse multiview features for decision-making, on the basis of contextual convolution branches. Finally, the reconstruction results of comparison experiments on different depths and seasons demonstrate that our SSCNN model outperforms other existing methods, with higher accuracy and smaller error.

The main contributions of this article are given as follows.

- 1) A multiscale scheme of sea area division based on the correlation coefficient of integral ST is proposed for refined reconstruction modeling. On the one hand, the modeling of coastal and noncoastal areas at different scales is conducive to learning the transitional region relationships and alleviating the land boundary influence. On the other hand, the division of noncoastal waters according to correlation coefficient can enhance the extraction of spatial-temporal characteristics and promote the fitting of Indian Ocean phenomena.
- 2) Since ocean heat is significantly affected by solar radiation, the asymmetric convolutional operation with rectangular patches and kernels is designed to capture the characteristics in LON and LAT directions, respectively. The zonal distribution is modeled to make CNN consider the obvious changes along with LAT and the slight changes along with LON. Inspired by this, the traditional square kernel is improved into the asymmetric rectangular kernel for better content interpretation of multisource remote sensing observations.

- 3) Given the temporal changes and correlations of temperature, an SSCNN with shared parameters is proposed for multiview feature mining and accurate temperature structure reconstruction. The Siamese branches have the advantages of consistent information processing and associated characteristic transmission. The reconstructed results provide a precise depiction of the sub-IOD's evolution, including the spatial distribution of positive and negative anomaly signals and its temporal changes.

The remainder of this article is structured as follows. Section II introduces the progress of temperature reconstruction using traditional machine learning and deep learning methods. In Section III, the detailed description of SSCNN model is illustrated. Section IV presents the experimental setup of dataset description and data preprocessing, and Section V presents the ablation experiments and comparison experiments for different depths and seasons. Finally, Section VI provides a summary and conclusion of this work.

II. RELATED WORKS

A. Traditional Machine Learning

In the early stage, traditional methods for ST reconstruction are mainly based on pattern assimilation and statistical reconstruction [10], [40]. The pattern assimilation methods produce reconstruction results that adhere to basic physical laws, but it is computationally expensive and requires a large amount of experience. On the other hand, statistical reconstruction methods such as linear regression [13], [41], [42], geographically weighted regression [43], and empirical orthogonal function have lower computational cost but rely heavily on the historical data.

Nonlinear machine learning algorithms are becoming increasingly important in the research, and integrated learning algorithms such as RF and SVM are particularly popular due to their ease of use with large datasets [17], [18], [19], [20]. Su et al. [21] employed the SVM method to reconstruct the subsurface temperature anomaly (STA) in the Indian Ocean. Subsequently, Su et al. [22] proposed an RF machine learning method that used multisource satellite remote sensing data to estimate the global ocean ST. It reconstructed the STA based on the anomaly of sea surface height (SSH), SST, salinity (SSS), and wind (SSW) variables. The anomaly of variables are obtained by subtracting the climatology, which is the average data for approximately 30 years, including the sea surface remote sensing and ST. Yang et al. [23] also used the RF algorithm to analyze the reconstruction effect in different ocean basins around the world.

In the later studies, Shahdi et al. [44] adopted XGBoost and RF methods for ST prediction, and derived a geothermal gradient map for the whole region by fitting linear regression to the predicted temperature along the depth. Qi et al. [24] proposed an ensemble machine learning model to retrieve subsurface thermal structure in the South China Sea, with combination of XGBoost, RF, and light gradient boosting machine (LightGBM) as benchmark models. It found that the geographical information is essential to estimate the subsurface thermal structure, and the ensemble model is more accurate

and precise than the individual models. Traditional machine learning algorithms have made significant contributions to the reconstruction of ocean ST and salinity structure, but they rely heavily on the feature engineering. On the other hand, deep learning algorithms learn the weights and extract the features automatically without feature engineering, which are adaptable and easy to transform.

B. Deep-Learning-Based Reconstruction

With the rapid development of artificial intelligence technology and marine big data, there has been a growing trend in integrating the advantages of these two to improve the accuracy and efficiency of STA reconstruction [45], [46], [47]. CNNs and other deep learning models have shown great potential in improving the performance and overcoming the limitations in various fields [25], [48], [49], including speech recognition, natural language processing, and image classification [50], [51], [52], [53]. Although the application of deep learning in the field of Earth science is still in its early stage, there have been some successful exploration cases [26]. Moreover, self-organizing map (SOM) [54], [55], clustering neural network [56], multilayer perceptron (MLP) [57], artificial neural network (ANN) [9], backpropagation neural network (BPNN) [11], and other neural network methods are also applied to reconstruct the STA, demonstrating their ability to simulate the nonlinear relationships. Recent studies also explored the ability of LSTM to extract temporal characteristics from marine data, which has obtained the stable reconstruction results [27], [28], [29].

In the study of reconstructing SST and SSS using CNNs, several approaches have been proposed, including 2-D CNN [30], [31], [32], 1-D CNN [16], and CNN combined with LSTM [33]. Although these methods have shown promising results, there are still some limitations. For example, Han et al. [30] proposed to use the multisource sea surface parameters to establish a monthly CNN model to reconstruct the ST, but the CNN cannot utilize the correlation before and after the feature sequence time. Meng et al. [31] discarded more than half of the images with missing value grids when building the sea surface feature maps, which significantly reduces the utilization of marine data and the applicability for coastal areas. Additionally, Su et al. [16] used only 1-D CNN to model the relationships between sea surface variables and subsurface properties, which has deficiency in the spatial-temporal feature extraction of sea surface remote sensing information. There still exist the problems of imprecise nonregional modeling, symmetrical LON and LAT feature extraction, and insufficient spatial-temporal semantic integration in deep learning methods for ST reconstruction.

To overcome these limitations, further research is needed to explore an optimized deep model to capture the temporal and spatial dependencies in marine data and promote the ST reconstruction accuracy. Therefore, this article proposes a novel SSCNN for effective reconstruction. It comprehensively uses multiple remote sensing to achieve full extraction of spatial-temporal features, adaptive learning of regional patterns, and efficient reconstruction of ST.

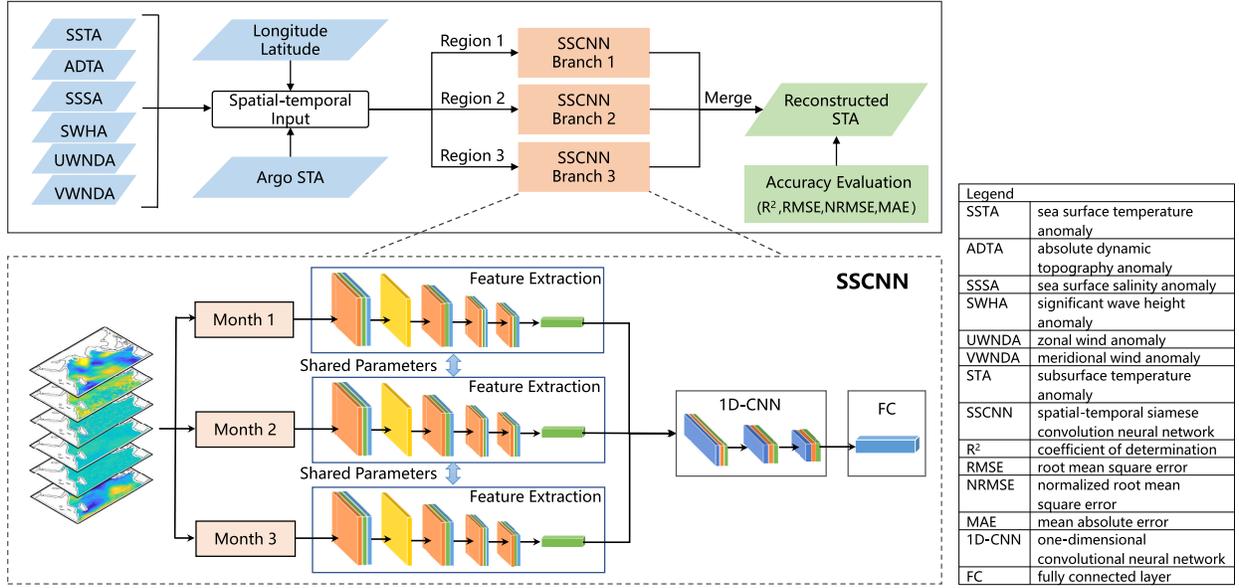


Fig. 1. Overall flowchart of the proposed SSCNN method for ST reconstruction.

III. METHODOLOGY

The SSCNN utilizes diverse sea surface remote sensing variables to reconstruct the STA above 1000 m in Indian Ocean, as shown in Fig. 1. On the one hand, input variables include SSTA, salinity anomaly (SSSA), significant wave height anomaly (SWHA), absolute dynamic topography anomaly (ADTA), zonal wind anomaly (UWNDA), and meridional wind anomaly (VWNDA). On the other hand, input variables also include geographic location information of LON and LAT. Specifically, the framework includes four main parts. First, the Indian Ocean is divided into three parts through temperature correlation coefficient, and different scales are created to improve the data extraction. Second, the rectangular convolution patches and kernels are designed to enhance the extraction of spatial features, achieving asymmetric expression of sea surface variables along the LAT and LON. Third, diverse remote sensing variables of sea surface in each month are organized into multichannel inputs and fed into Siamese CNN branches for spatial-temporal learning. Finally, feature vectors of each branch are fused to obtain the temperature output of various depths by reconstruction module.

A. Multiscale Division and Modeling

1) *Temperature Correlation Coefficient Calculation*: Considering that different sea areas have dissimilar distribution and variation characteristics of temperature and salinity, sea area division according to the temperature integral correlation coefficient is proposed to enhance the precise reconstruction modeling. The ocean thermocline experiences significant temperature variations, which is approximately at depth of 100–300 m below the sea surface. In order to accurately calculate the temperature correlation coefficient between grid points, it is necessary to consider both the SST and thermocline temperature simultaneously. The temperature changes below a depth of 300 m tend to be relatively stable, with less

variation in deeper water. Hence, the temperature information above 300 m is integrated and summed for the subsequent calculation of correlation coefficient. However, the function expression that describes how SST varies with depth is unknown. As a result, an integral approximation of the observation data is obtained through the gradient algorithm in this study, represented as

$$T_g = \text{trapz}(\mathbf{x}, \mathbf{y}) \quad (1)$$

where T_g represents the temperature of grid point after integration, $\mathbf{x} = [0, 5, \dots, 300]$ represents the depth vector from sea surface, and $\mathbf{y} = [T_1, T_2, \dots, T_{13}]$ represents the ST vector that corresponds to the depth vector. $\text{trapz}(\cdot)$ denotes the function to approximate the integral of ST using trapezoidal rule. The vector \mathbf{x} has nonuniform space between its different values, with an increasing interval as the depth expands. Considering the temperature of subsurface changes more slowly with depth, the upper levels are sampled densely, and the deep levels are sampled sparsely. Specifically, the sampling interval is set as 5, 10, 20, 25, and 50 m for the depth range of 0–10, 10–30, 30–50, 50–150, and 150–300 m, respectively. By using this sampling interval, the vector \mathbf{x} can effectively capture different temperature changes at various depths in the subsurface.

Then, each of grid points is processed point by point, and temperature data for each point are treated as a series with the length of 84, representing the 84 months from 2010 to 2016. For a given grid point a , its temperature sequence is denoted as $(a_1, a_2, \dots, a_{84})$, and the mean value \bar{a}_t of entire sequence is calculated. Similarly, for another grid point b , its temperature sequence is denoted as $(b_1, b_2, \dots, b_{84})$, and the mean value \bar{b}_t is also calculated. Subsequently, the correlation coefficient between these two temperature series can be expressed as

$$\bar{a}_t = \sum_{i=1}^n (a_{t,i}) / n \quad (2)$$

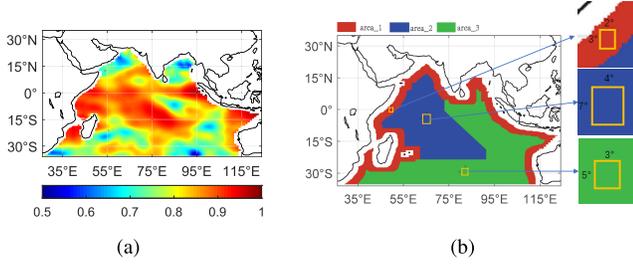


Fig. 2. Spatial distribution maps of: (a) correlation coefficient of ST at depth of 0–300 m and (b) sea area division according to temperature correlation coefficient in Indian Ocean.

$$\bar{b}_t = \sum_{j=1}^n (b_{t,j}) / n \quad (3)$$

$$R(a, b) = \frac{\sum_{k=1}^n (a_{t,k} - \bar{a}_t)(b_{t,k} - \bar{b}_t)}{\left(\sum_{i=1}^n (a_{t,i} - \bar{a}_t)^2 \sum_{j=1}^n (b_{t,j} - \bar{b}_t)^2 \right)^{1/2}} \quad (4)$$

where a_t and b_t represent the temperature of grid point a and b at time t , respectively, and $n = 84$ means the total number of months. The temperature correlation coefficient describes the distribution patterns of water mass, providing the basis for precise reconstruction modeling.

2) *Multiscale Sea Area Division*: Since the ocean heat is significantly influenced by the solar radiation and displays a zonal distribution, rectangular areas of varying sizes are established to calculate the local correlation of grid points. The length of rectangle is aligned with the LAT direction, and the width corresponds to the LON direction. By calculating the local correlation of grid points within these areas, the temperature integral correlation coefficient maps can be obtained. The results indicate that using a rectangular area in the size of $9^\circ \times 6^\circ$, the produced temperature correlation in the Indian Ocean is relatively continuous, which is proper to divide the regions and facilitate the subsequent modeling. Fig. 2(a) illustrates the temperature correlation coefficient of grid points calculated using $9^\circ \times 6^\circ$ size in the Indian Ocean.

Based on the correlation coefficient map, the noncoastal Indian Ocean has been divided into two regions. The division is based on a cutoff correlation value of 0.85, as shown in Fig. 2(b). This value is chosen to ensure that grid points with high correlation coefficients are grouped together as much as possible, while those with low correlation coefficients are grouped together as a separate region. This approach helps in creating relatively complete and continuous regions, which is convenient for model construction. Coastal areas were treated separately due to a relative lack of sea surface observation data, especially for annual salinity data where there are often missing values. The complex boundaries between land and sea in coastal areas also make it difficult to group them with noncoastal areas. Fig. 2(b) illustrates the final division result of Indian Ocean, with the blue region representing noncoastal area with higher correlation, the green region representing noncoastal area with lower correlation, and the red region representing coastal area.

The grid points in the blue region exhibit a high coefficient and strong correlation, and the remote sensing observation data

are relatively complete. Hence, larger patches in the size of $7^\circ \times 4^\circ$ are utilized to construct the 2-D sea surface feature input in order to effectively capture the spatial correlation of marine elements. On the other hand, the correlation coefficient of grid points in the green region is lower than that in the blue region, so smaller patches in the size of $5^\circ \times 3^\circ$ are adopted to consider the appropriate spatial context and construct the sea surface input. Using too large patches in the region with lower correlation coefficient would introduce irrelevant information and then cause accuracy reduction, and moderate patches are better for spatial extraction. As for the red region, it contains numerous missing data and complex boundaries, so further smaller patches in the size of $3^\circ \times 2^\circ$ are utilized to cover as much of the near-shore area as possible. In the Indian Ocean, the sea area is divided into multiscale regions for precise reconstruction modeling, including $7^\circ \times 4^\circ$ as large scale, $5^\circ \times 3^\circ$ as medium scale, and $3^\circ \times 2^\circ$ as small scale.

B. Asymmetric Convolutional Operation

Inspired by the rectangular region modeling above, an asymmetric convolution kernel with the size of 5×3 is designed to extract the sea surface features for more accurate STA reconstruction. The process is beneficial to capturing the significant temperature variations along with LAT as well as the slight variations along with LON. It involves sliding a window with the size of 5×3 over the channels of sea surface remote sensing feature maps. The step size of each slide is 2, and the padding of LON and LAT directions is 1 and 2, respectively. By setting the sliding step size as 2, the output shape of feature maps, the number of convolutional operations, and the redundancy of entire architecture are effectively reduced. The LON and LAT directions are filled differently to make full use of the edge information, improve the utilization of sea surface features, and enhance the overall accuracy. The LAT direction is filled with one more row than the LON direction because the ocean temperature is mainly affected by solar radiation with a zonal distribution. The size of sea surface remote sensing feature maps before and after the rectangular convolution kernel processing is expressed as

$$o_h = \frac{i_h + 2p_h - k_h}{s_h} + 1 \quad (5)$$

$$o_w = \frac{i_w + 2p_w - k_w}{s_w} + 1 \quad (6)$$

where i_h and i_w represent the height and width of remote sensing input before operation, respectively. For the input layer, i_h and i_w correspond to the patch sizes of $3^\circ \times 2^\circ$, $5^\circ \times 3^\circ$, and $7^\circ \times 4^\circ$ in nearshore and nonnearshore areas with different correlation coefficients, respectively. For the intermediate feature extraction layers, i_h and i_w correspond to the size of input feature maps for each layer. o_h and o_w represent the height and width of remote sensing feature map after operation, respectively. k_h and k_w represent the height and width of rectangular convolution kernel, respectively. p_h and p_w denote the padding size of height and width, respectively. s_h and s_w denote the step size of height and width, respectively. In this study, $k_h = 5$, $k_w = 3$, $p_h = 2$, $p_w = 1$, and $s_h = s_w = 2$.

In general, using a rectangle-shaped convolution kernel can capture more variations of sea surface remote sensing variables in the LAT direction, thereby improving the accuracy of temperature construction. The output of convolutional operation using a rectangular kernel and the activation function of rectified linear unit (ReLU) is expressed as

$$\mathbf{x}_l = \sigma(\mathbf{z}_l) = \text{ReLU}(\mathbf{W}_l * \mathbf{x}_{l-1} + \mathbf{b}_l) \quad (7)$$

$$x_{l,p} = \sum_{i=1}^5 \sum_{j=1}^3 \sum_{c=1}^C w_{l,q}^{i,j,c} \times x_{l-1,p}^{i,j,c} + b_{l,q} \quad (8)$$

where \mathbf{x}_l and \mathbf{x}_{l-1} represent the output at current layer and previous layer, respectively. \mathbf{W}_l means the weights of rectangular convolution kernel, and \mathbf{b}_l means the bias. Taking an example of patch p , $x_{l,p}$ and $x_{l-1,p}$ represent the convolution output and input, respectively. $w_{l,q}$ and $b_{l,q}$ are the weights and bias of q th rectangular convolution, respectively. $x_{l-1,p}^{i,j,c}$ means the specific value at i th row, j th column, and c th channel of $x_{l-1,p}$, and $w_{l,q}^{i,j,c}$ means the corresponding convolution weight. The range of i and j depends on the height and width of rectangular convolution kernel, and the range of c depends on the number of channels. Specifically, c is equal to 8 at the first convolution layer since there are eight variables of sea surface and location input into the Siamese CNN. On the basis of forward propagation of asymmetric convolutional calculation, the model further updates the parameters through backpropagation. In summary, the common square convolution kernels are simpler and more general purpose, while the designed rectangular convolution kernels are more flexible and useful for complex ocean information processing tasks. Asymmetric convolutional operation is used for the feature extraction of sea surface, pattern recognition of spatial-temporal information, and temperature reconstruction of subsurface.

C. SSCNN

The SSCNN is derived from the Siamese network structure, which has branches connected by shared weights. Siamese networks are advantageous when dealing with the similar inputs, and the changes in sea surface and subsurface elements have continuity and correlation in time series. Therefore, in this study, an SSCNN with multiple branches of consecutive monthly information is proposed for the feature extraction from multisource remote sensing. The number of branches and how to fuse their outputs are critical factors. Increasing the monthly input of remote sensing is beneficial for improving the feature extraction and thus enhancing the reconstruction accuracy. However, too many network branches would increase the number of parameters and introduce the potential interference, thus reducing the training efficiency. Based on these considerations, the SSCNN is designed with three shared branches, which receive the multisource sea surface remote sensing of three adjacent months in sequential order for deep feature learning.

As shown in Fig. 1, the feature extraction branches primarily consist of 2-D convolutional layers based on the rectangular convolution kernels. In order to balance the learning capacity and network complexity, six layers are adopted in the structure

TABLE I
PARAMETER SETTING OF SSCNN BRANCHES AT EACH LAYER

| Layer | Input_C | Output_C | Kernel | Stride | Padding |
|---------|---------|----------|--------|--------|---------|
| Conv1 | 8 | 16 | (5, 3) | 2 | (2, 1) |
| Pooling | / | / | (2, 2) | 2 | / |
| Conv2 | 16 | 32 | (5, 3) | 2 | (2, 1) |
| Conv3 | 32 | 64 | (5, 3) | 2 | (2, 1) |
| Conv4 | 64 | 128 | (5, 3) | 1 | (2, 1) |
| GAP | 128 | 128 | / | / | / |

of each Siamese CNN branch, including four convolutional layers and two pooling layers, as shown in Table I. Input_C and Output_C denote the channel numbers of input and output feature maps, respectively. The stride is set as 2 for the effective dimension reduction and spatial-temporal semantic extraction layer by layer. The settings of kernel and padding size are related to the asymmetric convolutional operation. Following each convolutional layer, there are batch normalization and activation function. Let $\mathbf{x} \in R^{N \times C \times H \times W}$ denote the multichannel feature maps in each batch, where N represents the number of samples, C represents the number of channels, and H and W represent the height and width of feature maps, respectively. The calculation of batch normalization is expressed as

$$\mu_c = \frac{1}{N \times H \times W} \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W x_{n,c,h,w} \quad (9)$$

$$\theta_c = \sqrt{\frac{1}{N \times H \times W} \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (x_{n,c,h,w} - \mu_c)^2 + \varepsilon} \quad (10)$$

$$\tilde{\mathbf{x}}_c = \frac{\mathbf{x}_c - \mu_c}{\theta_c} \times \gamma_c + \beta_c \quad (11)$$

where \mathbf{x}_c and $\tilde{\mathbf{x}}_c$ represent the feature map at c th channel before and after batch normalization, respectively. μ_c and θ_c are the mean value and standard deviation of channel c , respectively. ε is a value added to the denominator to increase the stability of calculation. γ_c and β_c are two learnable parameters of batch normalization, which are used to scale and shift the normalized values and automatically learned by the network. A value of γ_c close to 1 means that the feature is preserved as is, while a value greater than 1 amplifies the feature and a value less than 1 attenuates it. A value of β_c close to 0 means that no shift is applied, while a nonzero value shifts the feature by a certain amount.

The ST reconstruction for a target month is achieved using the SSCNN, which involves several steps. First, multisource remote sensing variables for the target month and last two months are organized into eight channels, respectively, including SSTA, SSSA, SWHA, ADTA, UWDA, VWDA, LON, and LAT. Second, the sea surface variables are input into Siamese branches with asymmetric convolution and shared weights for deep feature extraction. Diverse rectangular convolution kernels are adopted to learn the local and global features of multisource sea surface information, and the numbers of kernels at four convolutional layers are set to 16, 32, 64, and 128, respectively. To reduce the number of parameters in the fusion of Siamese branches, the remote sensing feature maps

of each branch are aggregated as the overall sea surface vector using global average pooling (GAP), which is represented as

$$z_c = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W u_c(h, w) \quad (12)$$

where H and W represent the height and width of remote sensing feature map, respectively, $u_c(h, w)$ represents the value of channel c at position (h, w) in feature map, and z_c means the aggregation value of channel c .

Weight sharing is implemented among the three branches to ensure that each branch extracts sea surface remote sensing features corresponding to each month in the same pattern. The GAP feature vectors of three branches are merged in the channel dimension, and the resulting temporal features of sea surface are fed into the final reconstruction module. The design of Siamese branches for spatial-temporal feature extraction is conducive to coupling the individual changing trends and mining the complex relationships among various variables. Specifically, each branch outputs a remote sensing feature vector with the size of 128, and the output vectors of three branches are concatenated, represented by

$$z_{b1} = \text{Branch}([\mathbf{x}_{m1}^{\text{SSTA}}, \dots, \mathbf{x}_{m1}^{\text{LON}}, \mathbf{x}_{m1}^{\text{LAT}}]) \quad (13)$$

$$z_{b2} = \text{Branch}([\mathbf{x}_{m2}^{\text{SSTA}}, \dots, \mathbf{x}_{m2}^{\text{LON}}, \mathbf{x}_{m2}^{\text{LAT}}]) \quad (14)$$

$$z_{b3} = \text{Branch}([\mathbf{x}_{m3}^{\text{SSTA}}, \dots, \mathbf{x}_{m3}^{\text{LON}}, \mathbf{x}_{m3}^{\text{LAT}}]) \quad (15)$$

$$z_{\text{con}} = \text{Concat}(z_{b1}, z_{b2}, z_{b3}) \quad (16)$$

where \mathbf{x}_{m1} , \mathbf{x}_{m2} , and \mathbf{x}_{m3} denote the remote sensing input of last two months and target month, respectively, for the Siamese feature extraction. $[\mathbf{x}_{m1}^{\text{SSTA}}, \dots, \mathbf{x}_{m1}^{\text{LON}}, \mathbf{x}_{m1}^{\text{LAT}}]$ means the vector of various sea surface variables. $\text{Branch}(\cdot)$ means the sequential operation of convolution and pooling layers in each branch, which obtains the corresponding GAP vectors of z_{b1} , z_{b2} , and z_{b3} . $\text{Concat}(\cdot)$ means the channel concatenation of GAP feature vectors to produce the comprehensive information for final reconstruction module.

D. Reconstruction Module

The reconstruction module consists of 1-D convolution and fully connected (FC) layers. The 1-D CNN layer first performs the convolution operation through 1×3 kernels and then transforms the features through ReLU activation. It calculates the feature correlations of multichannel information and extracts the internal relationships of spatial-temporal characteristics from three adjacent months. After the 1-D convolution and ReLU activation, there is a maxpooling operation to reduce the length of features and extract the compressed semantics for ST reconstruction. The above process of convolution, activation, and maxpooling is repeated twice to deepen the temporal feature fusion and enhance the fitting between sea surface and subsurface variables. Subsequently, FC layer connects to the input of 1-D CNN vectors and the output of STA values, which are organized into a sequence of different standard depths. Each STA at a specific depth is denoted as the corresponding output neuron, generated through the

integration of 1-D CNN and FC layers. The reconstruction module is expressed by

$$z_{a1} = \text{ReLU}(W'_{a1} * z_{\text{textcon}} + b_{a1}) \quad (17)$$

$$z_{a2} = \text{ReLU}(W'_{a2} * z_{a1} + b_{a2}) \quad (18)$$

$$\hat{y} = \text{FC}(z_{a2}) \quad (19)$$

where W'_{a1} and W'_{a2} denote the convolution weights of two 1-D CNN layers, respectively, and b_{a1} and b_{a2} are the corresponding bias. The obtained feature z_{a2} is input into the FC layer to reconstruct the STA \hat{y} .

The whole model is trained through the backpropagation of L2 loss function, and the final reconstruction outputs are K -dimensional vectors, where K is the number of standard layers in Indian Ocean. L2 loss function is adopted to minimize the sum of squares of difference between the observed and reconstructed STA. Let y_i represent the observed STA vector of various standard layers at grid point i , \hat{y}_i represent the corresponding reconstructed STA vector, and M represent the number of sample grid points. L2 loss function can be shown as

$$\mathcal{L}(y, \hat{y}) = \frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2 \quad (20)$$

where the iterative optimization of model is carried out by the stochastic gradient descent algorithm with backpropagation. In general, SSCNN method learns the spatial-temporal features from multisource sea surface remote sensing and reconstructs the ST through multiscale division, asymmetric convolution, and Siamese structure.

IV. EXPERIMENTAL SETUP

In this section, we present a detailed description of experimental setup to demonstrate the implementation of our proposed model in reconstructing the ST of Indian Ocean. The evaluation of research results is based on the mean absolute error (MAE), root-mean-square error (RMSE), normalized root-mean-square error (NRMSE), and coefficient of determination (R^2) as quantitative indicators.

A. Description of Datasets

The remote sensing data used in this study include multiple variables related to the sea surface, such as SST, SSS, sea wave height (SWH), absolute dynamic topography (ADT), and SSW (UWIND and VWIND). The SST data are obtained from the National Oceanic and Atmospheric Administration, Washington, DC, USA and acquired by the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) sensor. It covers a global extent and a time frame from 1981 to present with daily temporal resolution and 0.25° spatial resolution. The SSS data are obtained from the European Space Agency Paris, France and acquired by the Soil Moisture and Ocean Salinity (SMOS) sensor. It covers a global extent and a time frame from May 2010 to August 2017 with monthly temporal resolution and 0.25° spatial resolution. The SWH data are obtained from the Copernicus Marine Service, Brest, France, describing a global wave reanalysis of past sea conditions from 1993 to present,

with 3-hourly temporal resolution and 0.2° spatial resolution. The ADT data are also obtained from the Copernicus Marine Service and derived from the global ocean grid L4 SSH and postprocessing. It covers a time frame from 1993 to present with daily temporal resolution and 0.25° spatial resolution. The UWND and VWND data are obtained from the Cross-Calibrated Multi-Platform (CCMP) project. They cover a global extent and a time frame from 1987 to present with monthly temporal resolution and 0.25° spatial resolution. The ST data are obtained from the International Pacific Research Center (IPRC), Hawaii, HI, USA for the Argo gridded data. It covers a spatial extent of 180°W – 180°E and 60°S – 60°N , and a time frame from 2005 to present with monthly temporal resolution and 1° spatial resolution.

B. Data Preprocessing

The spatial and temporal resolutions of sea surface remote sensing observation data and IPRC Argo data are not consistent. To address this, the nearest neighbor interpolation algorithm is used to standardize the spatial resolution to $0.25^\circ \times 0.25^\circ$ and the temporal resolution to monthly intervals. Since the seasonal signal is stronger than other physical ocean signals, it is necessary to remove it from the data and focus on other important physical ocean signals. Each variable's climatic average is subtracted from the corresponding monthly data to produce the anomaly of each variable, including SSTA, SSSA, SWHA, ADTA, UWNDA, VWNDA, and STA. To facilitate the model training, maximum and minimum normalization is carried out on each variable, shown as

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (21)$$

where X_{norm} represents the normalized variable, and X represents the original value. X_{\min} and X_{\max} denote the minimum and maximum values among all training samples, respectively. When maximum and minimum normalization is applied to the testing data, X_{\min} and X_{\max} are still the minimum and maximum values of training samples.

The SSCNN method proposed in this article adopts three scales to reconstruct the STA of Indian Ocean, which are $3^\circ \times 2^\circ$, $5^\circ \times 3^\circ$, and $7^\circ \times 4^\circ$, respectively. For these multiple scales, the target area has a size of 12×8 , 20×12 , and 28×16 grid points, respectively. To collect the remote sensing information of sea surface around a target grid point, if any of the sea surface variables have missing data, skip the target grid point and continue to retrieve neighborhood information for the next target grid point. For the coastal area, small scale of 12×8 grid points is employed to extract the local characteristics and learn the interaction relationships. For the relatively low- and high-correlation area, medium scale of 20×12 grid points and large scale of 28×16 grid points are employed to fully extract the ocean information, respectively, since the surrounding remote sensing observation data are generally complete.

V. EXPERIMENTAL RESULTS

A. Analysis of Parameters

1) *Input Variables*: To begin with, the importance of different sea surface variables for STA reconstruction is analyzed

TABLE II
DIFFERENT COMBINATIONS OF INPUT VARIABLES

| Case | Variables |
|-------|--|
| Case1 | ADTA, SSTA, SWHA, UWNDA, VWNDA, SSSA |
| Case2 | SSTA, SWHA, UWNDA, VWNDA, SSSA, LON, LAT |
| Case3 | ADTA, SWHA, UWNDA, VWNDA, SSSA, LON, LAT |
| Case4 | ADTA, SSTA, UWNDA, VWNDA, SSSA, LON, LAT |
| Case5 | ADTA, SSTA, SWHA, SSSA, LON, LAT |
| Case6 | ADTA, SSTA, SWHA, UWNDA, VWNDA, LON, LAT |
| Case7 | ADTA, SSTA, SWHA, UWNDA, VWNDA, SSSA, LON, LAT |

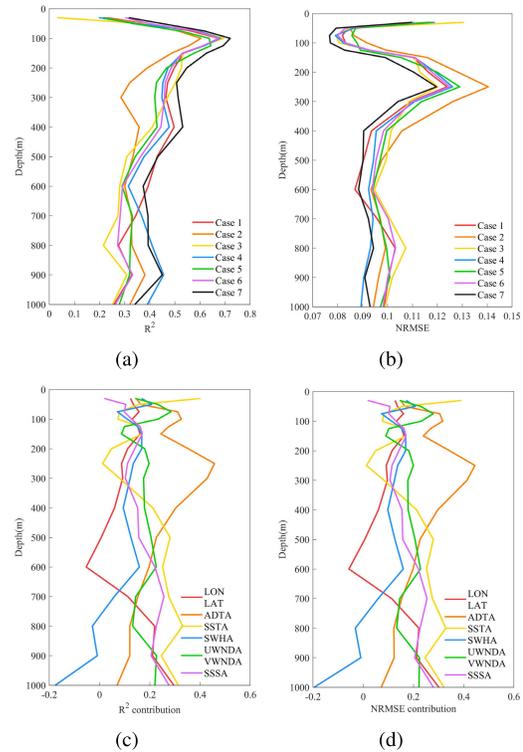


Fig. 3. Reconstruction accuracy of STA at different depths in terms of: (a) R^2 and (b) NRMSE under various cases, and the (c) R^2 contribution and (d) NRMSE contribution of input variables.

through the experiments of various combinations, as shown in Table II and Fig. 3. A total of seven combinations of input variables are designed to compare the differences in STA reconstruction by sequentially removing each sea surface variable. Among them, Case 1–Case 6 remove geographical location (LON and LAT), ADTA, SSTA, SWHA, SSW (UWNDA and VWNDA), and SSSA variables, respectively, while Case 7 includes all variables for comparison. The target month is set as October 2015, and the vertical layers include standard depths of 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, and 1000 m. Considering the months of data for various sea surface elements in common, the monthly dataset is available from January 2011. The samples before target month are used for training, and the samples of target month are used for testing. It can be seen that Case 7 exhibits the highest accuracy in terms of R^2 and the lowest error in terms of NRMSE among all the cases at most depths. This indicates that all the sea surface variables (SSTA, SSSA, ADTA, SWHA, UWNDA,

and VWNDA) and location variables (LON and LAT) are important for accurately reconstructing the STA of Indian Ocean. The absence of any variable would lead to a reduction in the accuracy of reconstruction, and similar conclusion is also mentioned in [28]. Overall, for depths from 30 to 1000 m, the R^2 value first increases obviously, reaching a maximum at about 100 m, and then gradually decreases and stabilizes. The average R^2 of 16 standard depths (30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, and 1000 m) is 0.435, 0.380, 0.399, 0.445, 0.409, 0.422, and 0.494 for Case 1–Case 7, respectively. The p -values of paired-sample t-test (two-sided) for the R^2 of 16 layers are 1.48E-05, 4.41E-07, 6.73E-05, 0.0006, 1.83E-10, and 1.30E-07 between Case 7 and Case 1, Case 2, Case 3, Case 4, Case 5, and Case 6, respectively. Correspondingly, the NRMSE value first decreases rapidly, reaching a minimum at about 100 m, and then gradually increases and decreases again. The average NRMSE of 16 standard depths is 0.099, 0.105, 0.102, 0.098, 0.101, 0.100, and 0.094 for Case 1–Case 7, respectively. The p -values of paired-sample t-test (two-sided) for the NRMSE of 16 layers are 8.43E-06, 7.12E-06, 1.55E-05, 0.0004, 2.43E-10, and 1.24E-07 between Case 7 and Case 1, Case 2, Case 3, Case 4, Case 5, and Case 6, respectively.

The contribution of different variables is evaluated in terms of R^2 and RMSE in Fig. 3, by comparing and normalizing the difference of R^2 and RMSE between other cases and Case 7. The accuracy of 300–700-m STA reconstruction results appears to be less affected by LAT and LON information, possibly because the spatial location information of sea surface elements has already been considered when organizing them into feature maps. The absence of ADTA significantly reduces the accuracy of STA reconstruction at depths of 100–500 m, causing a substantial increase in error. The absence of SSTA obviously reduces the accuracy of results at depths of 500–1000 m. ADTA and SSTA have significant contributions to the STA reconstruction due to their internal relationships and strong correlations. The SWHA, SSSA, UWNDA, and VWNDA have similar and certain importance for temperature anomaly information above 600 m, but the effectiveness of SWHA decreases with increasing depth. Besides the SSTA and ADTA, the results above 125 m exhibits more dependence on wind anomaly data (UWNDA and VWNDA), which is possibly due to the significant air–sea interaction in Indian Ocean and its prevailing monsoon climate. Overall, different variables have unequal contributions in different depth range, but the SSCNN model input with all variables achieves the highest accuracy for ST reconstruction.

2) *Kernel Size*: To verify the effectiveness of the proposed asymmetric convolution kernel in extracting spatial–temporal features from multisource sea surface remote sensing data, a parameter analysis is conducted on the 5×3 , 3×3 , and 5×5 convolution kernels. The reconstruction results are compared between symmetric and asymmetric convolution kernels to assess the differences. For the 5×3 convolution kernel, the average R^2 of different depths are 0.320, 0.432, 0.399, and 0.494 for January, April, July, and October in 2015, respectively, and the average NRMSE are 0.096, 0.102, 0.093, and 0.094, respectively. For the 3×3 convolution

TABLE III
PERFORMANCE OF SSCNN AT DIFFERENT DEPTHS FOR STA
RECONSTRUCTION OF OCTOBER 2015

| Depth (16 layers) | R^2 | RMSE | NRMSE | MAE |
|----------------------|-------|-------|-------|-------|
| 30m | 0.316 | 0.218 | 0.110 | 0.173 |
| 50m | 0.448 | 0.279 | 0.079 | 0.220 |
| 75m | 0.616 | 0.476 | 0.077 | 0.347 |
| 100m | 0.720 | 0.448 | 0.077 | 0.334 |
| 125m | 0.692 | 0.335 | 0.083 | 0.264 |
| 150m | 0.626 | 0.294 | 0.099 | 0.228 |
| 200m | 0.547 | 0.240 | 0.110 | 0.187 |
| 250m | 0.507 | 0.217 | 0.120 | 0.158 |
| 300m | 0.513 | 0.186 | 0.104 | 0.134 |
| 400m | 0.531 | 0.121 | 0.090 | 0.094 |
| 500m | 0.432 | 0.105 | 0.090 | 0.078 |
| 600m | 0.374 | 0.113 | 0.088 | 0.083 |
| 700m | 0.393 | 0.111 | 0.092 | 0.083 |
| 800m | 0.393 | 0.113 | 0.094 | 0.089 |
| 900m | 0.450 | 0.096 | 0.091 | 0.074 |
| 1000m | 0.341 | 0.087 | 0.093 | 0.067 |

kernel, the average R^2 of different depths are 0.291, 0.425, 0.401, and 0.476 for January, April, July, and October in 2015, respectively, and the average NRMSE are 0.097, 0.102, 0.093, and 0.095, respectively. The p -value of paired-sample t-test (two-sided) for the R^2 of four months is 0.048 between 5×3 and 3×3 kernels, and the p -value for the NRMSE is 0.078. For the 5×5 convolution kernel, the average R^2 of different depths are 0.307, 0.392, 0.394, and 0.469 for January, April, July, and October in 2015, respectively, and the average NRMSE are 0.096, 0.106, 0.094, and 0.096, respectively. The p -value of paired-sample t-test (two-sided) for the R^2 of four months is 0.016 between 5×3 and 5×5 kernels, and the p -value for the NRMSE is 0.092. The designed 5×3 convolution kernel achieves the highest overall R^2 and the lowest overall NRMSE, indicating the advantage in STA reconstruction compared to the symmetric 3×3 and 5×5 convolution kernels. The asymmetric convolution kernel takes into account the different distribution of sea surface variables in the LON and LAT directions, resulting in better performance in extracting the sea surface remote sensing and reconstructing the ST structure.

B. Results of Different Depths

Table III presents the performance evaluation results of SSCNN for STA reconstruction at different depths (16 layers) using four indicators: R^2 , RMSE, NRMSE, and MAE for October 2015. The average R^2 , RMSE, NRMSE, and MAE of 16 depth layers are 0.494, 0.215, 0.094, and 0.163, respectively. Specifically, R^2 ranges from 0.316 to 0.720, RMSE ranges from 0.087 to 0.476, NRMSE ranges from 0.077 to 0.120, and MAE ranges from 0.067 to 0.347. The results show that RMSE remains at a low value throughout the depth range, with the highest value occurring at depth of 75 m. The trend of RMSE is unstable and fluctuating from 0.218 to the maximum of 0.476 in depth range from 30 to 75 m, and then, it gradually decreases from 0.476 to the minimum of 0.087 with increasing depth from 75 to 1000 m. Meanwhile, MAE shows a similar trend to RMSE, which rapidly raises from 0.173 to the

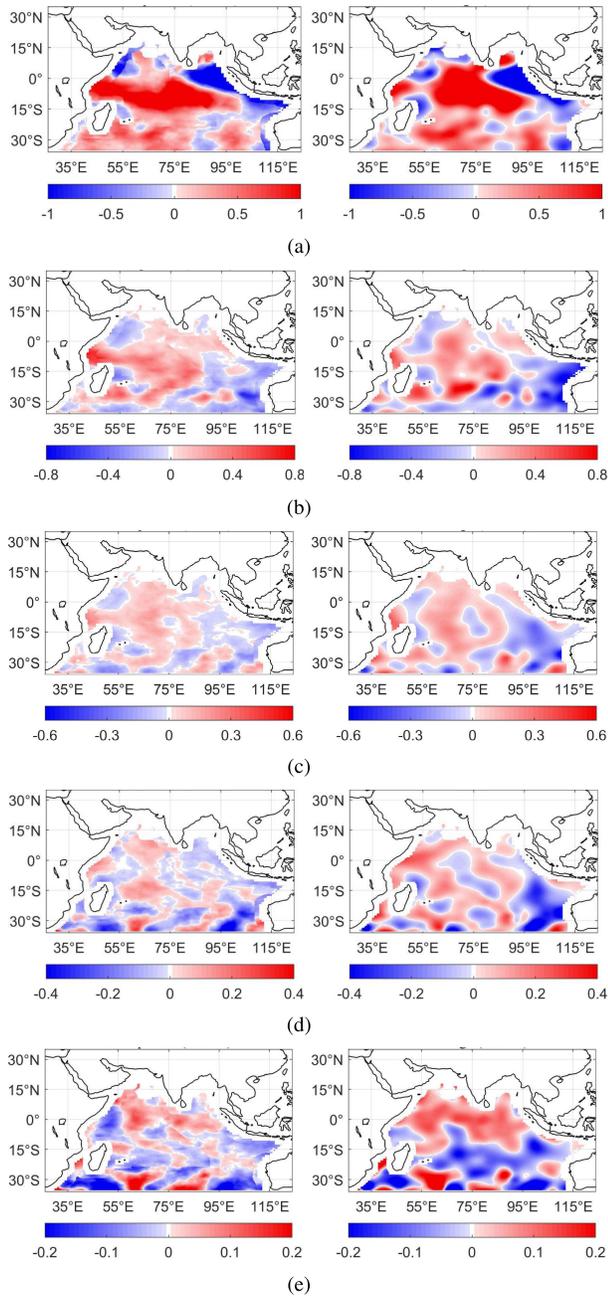


Fig. 4. STA reconstruction of (Left) SSCNN compared with the STA observation of (Right) IPRC for October 2015 at depths of: (a) 100 m; (b) 300 m; (c) 500 m; (d) 700 m; and (e) 1000 m.

maximum of 0.347 in depth range from 30 to 75 m, and then, it gradually reduces from 0.347 to the minimum of 0.067 with increasing depth from 75 to 1000 m. For R^2 , it reaches the maximum of 0.720 at depth of 100 m and then gradually decreases with some fluctuations. For NRMSE, it obtains the minimum of 0.077 at a depth of 75 and 100 m and reaches the maximum of 0.120 at a depth of 250 m. In general, STA and its reconstruction accuracy have significant changes in shallow layers with the maximum gradient in thermocline and gradually stabilize in deep layers.

Fig. 4 shows the comparison between STA reconstruction of SSCNN model and STA observation of IPRC dataset for October 2015 at depths of 100, 300, 500, 700, and 1000 m,

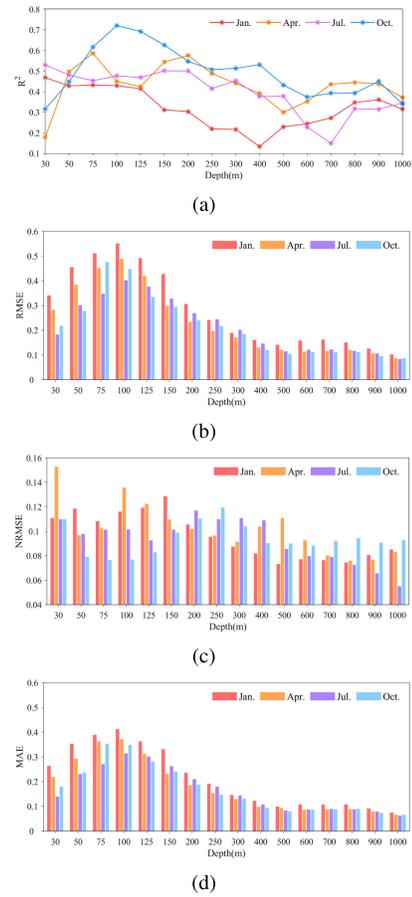


Fig. 5. Accuracy evaluation of STA reconstruction results at different depths of Indian Ocean for four months in terms of: (a) R^2 ; (b) RMSE; (c) NRMSE; and (d) MAE.

respectively. It demonstrates that the spatial distribution of SSCNN reconstruction is consistent with the IPRC Argo data. The reconstructed results of Indian Ocean at a depth of 100 m show the highest similarity with IPRC Argo data, particularly in the equatorial eastern coastal waters with strong negative anomaly signals and the equatorial central waters with strong positive anomaly signals. At a depth of 300 m, the negative anomaly signals in equatorial eastern coastal area and the positive anomaly signals in equatorial central area decrease, while the negative anomaly signals in 10°S–30°S eastern coastal area increase. As layer becomes deeper, the ST tends to be stable, resulting in weaker positive and negative anomaly signals. Accordingly, the RMSE of 700 and 1000 m is relatively low without obvious changes of STA. At different depths, the SSCNN model can reconstruct STA results with anomaly amplitude and spatial distribution close to the Argo observation.

C. Comparison of Seasonal Results

Fig. 5 illustrates how the SSCNN reconstruction error and correlation coefficient vary with depth across months of different seasons. Different types of graphs are used for R^2 , RMSE, NRMSE, and MAE because they represent different aspects of model performance and provide distinct insights. January, April, July, and October in 2015 are chosen as the

target months to demonstrate the SSCNN performance for winter, spring, summer, and autumn, respectively. In general, the R^2 , RMSE, and MAE first increase and then decrease with growing depth across different seasons, while the NRMSE fluctuates in the increasing and decreasing trends with growing depth. The overall reconstruction accuracy of SSCNN model is highest in October with average R^2 of 0.494 and average NRMSE of 0.094, followed by April with average R^2 of 0.432 and average NRMSE of 0.102, July with average R^2 of 0.399 and average NRMSE of 0.093, and January with average R^2 of 0.320 and average NRMSE of 0.096. The R^2 of October within depth range of 100–150 m reflects a relatively obvious correlation advantage with the highest value of 0.720, and its NRMSE shows a lower value within depth range of 50–150 m and a higher value within depth range of 700–1000 m. It is possibly related to the positive dipole in Indian Ocean and the corresponding high dipole mode index (DMI), which leads to an unstable upper ocean dynamic environment. The R^2 of January within depth range of 150–500 m presents a relatively weak correlation with the lowest value of 0.134, and its RMSE and MAE at most depths show a higher value. The R^2 of April within depth range of 30–150 m reflects a fluctuation trend of increasing and decreasing, and its NRMSE shows the highest value of 0.153 at a depth of 30 m and the second highest value of 0.136 at a depth of 100 m. The R^2 of July within a depth range of 600–800 m presents a relatively weak correlation, and its corresponding RMSE, NRMSE, and MAE are medium, which is possibly related to the changes of Asian summer monsoon. Overall, the SSCNN method is suitable for reconstructing STA for different seasons in Indian Ocean, as the model shows stable performance with good accuracy across different seasons.

Fig. 6 depicts the spatial distribution of STAs reconstructed by SSCNN at a depth of 100 m in Indian Ocean for four months. The results show that SSCNN reconstruction results are consistent with the STA of IPRC Argo data across different months, and the spatial distribution of STA varies significantly from January to October in 2015. Specifically, there exist obvious negative anomaly signals in the equatorial western Indian Ocean and positive anomaly signals in the equatorial eastern area during January. Subsequently, ST warming in the equatorial western Indian Ocean off the coast of eastern Africa starts around April, weakens briefly around July, and then reaches a high value of positive anomaly around October. Correspondingly, ST cooling in the equatorial eastern area starts around April and reaches a high value of negative anomaly around October. It indicates that the subsurface dipole of Indian Ocean is in a positive phase during this period. The aforementioned changes of positive and negative STA are well-presented in the SSCNN reconstruction results, which are consistent with the actual situation in Indian Ocean.

Furthermore, the climatology values are added back to the STA reconstructed by SSCNN, to obtain the ST and further verify the precision with an independent dataset of previous study. It is compared with the ARMOR3D data (https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012), which includes 3-D multiobservation product of the ocean

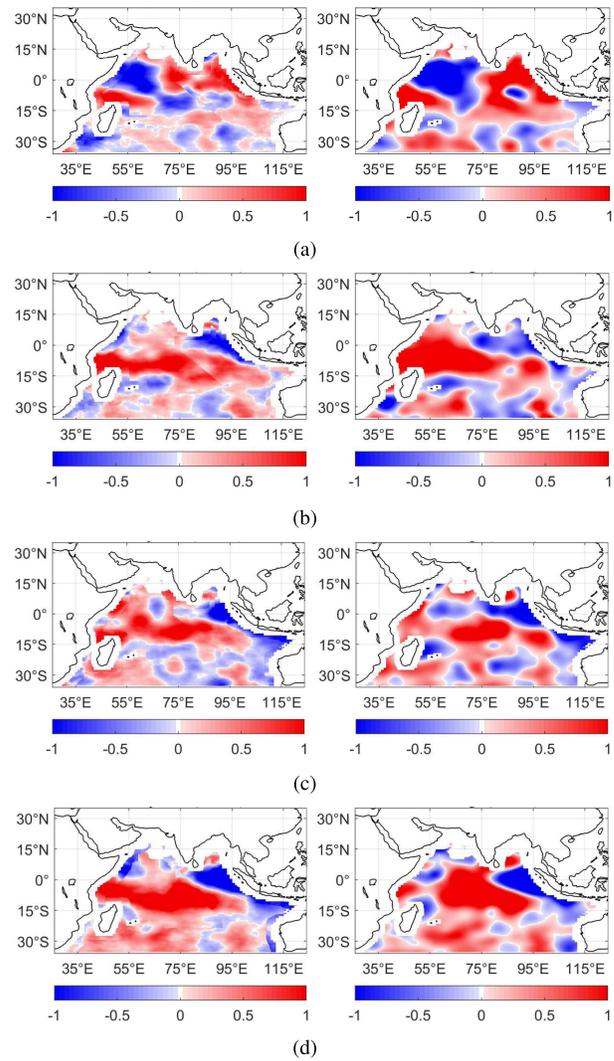


Fig. 6. Spatial distribution of (Left) STA reconstruction results and (Right) IPRC Argo data at depth of 100 m for: (a) January; (b) April; (c) July; and (d) October in 2015.

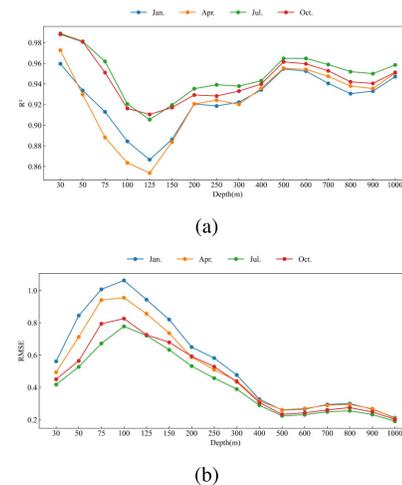


Fig. 7. Comparison of ST between SSCNN reconstruction and ARMOR3D dataset in terms of: (a) R^2 and (b) RMSE.

temperature [13], [58], [59]. Fig. 7 shows the R^2 and RMSE of ST between SSCNN reconstruction and ARMOR3D dataset, which are obtained within the range of 25°E–120°E

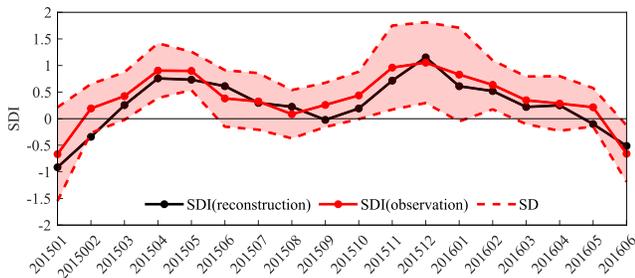


Fig. 8. Time series of SDI from January 2015 to June 2016 in Indian Ocean.

and 35°S–30°N for January, April, July, and October in 2015, respectively. The average R^2 values of 16 standard depths are 0.925, 0.923, 0.949, and 0.944 from January to October, respectively, and the average RMSE values are 0.555, 0.509, 0.425, and 0.460, respectively. The results illustrate that the reconstructed ST by the SSCNN model is consistent with the ARMOR3D dataset. The R^2 gradually decreases starting from a depth of 30 m and reaches its lowest point at a depth of 125 m, after which it gradually fluctuates and raises with increasing depth. Correspondingly, the RMSE gradually increases starting from a depth of 30 m, reaches its lowest value at a depth of 100 m, and then gradually decreases as the depth increases.

D. Analysis of Subsurface Positive Indian Ocean Dipole

In order to further analyze the subsurface positive dipole in Indian Ocean based on the STA reconstruction results by SSCNN model, the subsurface dipole index (SDI) is calculated according to [60]. The first step is to calculate the density of seawater using temperature, salinity, and pressure data. Then, the second step is to compute the heat content anomaly (HCA) in upper ocean through multiplying the mean STA by density and specific heat of seawater. Third, calculate the normalized average HCA in western (5°S–5°N, 50°–70°E) and eastern (5°S–5°N, 75°–95°E) regions of equatorial Indian Ocean. Finally, subtract the normalized average HCA in eastern region from that in western region to obtain SDI. It is important to note that SDI is a measure of gradient in HCA across the equatorial Indian Ocean, which has significant impacts on climate variability.

Fig. 8 displays the time series of SDI in Indian Ocean calculated using the STA reconstructions of SSCNN model and STA observations of IPRC Argo data. The red solid line represents the SDI calculated by STA observations of IPRC data, the black solid line represents the SDI calculated by STA reconstructions of SSCNN model, and the red dotted line along with enclosed red shaded area corresponds to one standard deviation of IPRC-SDI. The results demonstrate that the SDI series obtained from STA reconstructions of SSCNN model is consistent with that from STA observations of IPRC data, remaining within a reasonable error range. Moreover, the SDI series of STA reconstructions captures the same dipole phenomenon as described in [60]. In 2015, the sub-IOD lags behind the sea surface Indian Ocean dipole (ss-IOD) in occurrence time. It emerges at the end of autumn, reaches its peak during winter, and gradually ends in spring. The

SDI values of STA reconstructions for June, July, and August 2015 have a relatively small error, which can be attributed to the slow change of sea temperature in Indian Ocean during this period. In contrast, the SDI values for September, October, and November have a relatively large error, which is possibly related to the occurrence of positive Indian Ocean dipole (pIOD).

E. Comparison Experiments

To verify the superiority of proposed SSCNN method, XGBoost (XGB) [15], [61], RF [22], [62], fully connected neural network (FCNN) [63], Informer [64], ConvNeXt [65], and temporal convolutional network (TCN) [66] are used as comparison methods for analysis. In addition, to demonstrate the effectiveness of Siamese CNN structure, a single-branch spatial-temporal CNN (SCNN) is implemented for ablation analysis. In order to comprehensively compare the reconstruction performance between SSCNN and other methods, all 18 months of six seasons from December 2014 to May 2016 are considered, as shown in Table IV. Concretely, the indicators of winter in 2014 (2014-Win.), spring in 2015 (2015-Spr.), summer in 2015 (2015-Sum.), autumn in 2015 (2015-Aut.), winter in 2015 (2015-Win.), and spring in 2016 (2016-Spr.) are the average results from December 2014 to February 2015, from March 2015 to May 2015, from June 2015 to August 2015, from September 2015 to November 2015, from December 2015 to February 2016, and from March 2016 to May 2016, respectively. Overall, the SSCNN method basically achieves the highest R^2 and the lowest RMSE, NRMSE, and MAE in the STA reconstruction results of Indian Ocean from 30 to 1000 m in most seasons.

Specifically, the R^2 of SSCNN is 0.006–0.092, 0.030–0.182, 0.013–0.163, 0.044–0.193, 0.024–0.206, and 0.026–0.218 higher than that of other methods from 2014-Win. to 2016-Spr., respectively. The p -values of paired-sample t-test (two-sided) for the R^2 of 18 months are 3.51E-11, 3.51E-12, 2.75E-12, 1.98E-09, 2.24E-07, 5.65E-10, and 0.0006 between SSCNN and XGB, RF, FCNN, Informer, ConvNeXt, TCN, and SCNN methods, respectively. In terms of reconstruction error, the RMSE of SSCNN is averagely 0.020, 0.020, 0.018, 0.029, 0.036, and 0.022 lower than that of other methods from 2014-Win. to 2016-Spr., respectively. The p -values of paired-sample t-test (two-sided) for the RMSE of 18 months are 2.77E-08, 4.75E-09, 3.44E-09, 2.63E-06, 9.28E-05, 2.03E-07, and 0.225 between SSCNN and other methods, respectively. In order to eliminate the influence of STA outlier ranges on error evaluation, NRMSE is adopted to make the comparison more objective. The NRMSE of SSCNN is averagely 0.003, 0.008, 0.005, 0.008, 0.009, and 0.008 lower than that of other methods from 2014-Win. to 2016-Spr., respectively. The p -values of paired-sample t-test (two-sided) for the NRMSE of 18 months are 8.34E-10, 1.86E-10, 1.07E-10, 1.53E-08, 1.33E-06, 4.00E-09, and 0.005 between SSCNN and other methods, respectively. In terms of absolute error, the MAE of SSCNN is averagely 0.010, 0.014, 0.012, 0.018, 0.022, and 0.016 lower than that of other methods from 2014-Win. to 2016-Spr., respectively. The p -values of paired-sample t-test (two-sided) for the MAE of 18 months are 4.22E-09, 3.15E-09,

TABLE IV
RECONSTRUCTION PERFORMANCE EVALUATION FOR SIX SEASONS BY COMPARISON METHODS

| Indicator | Season | XGB | RF | FCNN | Informer | ConvNeXt | TCN | SCNN | SSCNN |
|-----------|-----------|-------|-------|-------|----------|--------------|-------|--------------|--------------|
| R^2 | 2014-Win. | 0.235 | 0.209 | 0.193 | 0.230 | 0.279 | 0.238 | 0.275 | 0.285 |
| | 2015-Spr. | 0.299 | 0.292 | 0.241 | 0.383 | 0.387 | 0.357 | 0.393 | 0.423 |
| | 2015-Sum. | 0.267 | 0.250 | 0.207 | 0.319 | 0.342 | 0.334 | 0.357 | 0.370 |
| | 2015-Aut. | 0.317 | 0.332 | 0.266 | 0.415 | 0.412 | 0.393 | 0.410 | 0.459 |
| | 2015-Win. | 0.393 | 0.386 | 0.321 | 0.455 | 0.469 | 0.446 | 0.503 | 0.527 |
| | 2016-Spr. | 0.420 | 0.415 | 0.314 | 0.462 | 0.474 | 0.474 | 0.506 | 0.532 |
| RMSE | 2014-Win. | 0.353 | 0.362 | 0.362 | 0.357 | 0.341 | 0.363 | 0.341 | 0.334 |
| | 2015-Spr. | 0.294 | 0.296 | 0.307 | 0.279 | 0.265 | 0.286 | 0.269 | 0.265 |
| | 2015-Sum. | 0.280 | 0.282 | 0.292 | 0.271 | 0.264 | 0.270 | 0.261 | 0.256 |
| | 2015-Aut. | 0.302 | 0.305 | 0.324 | 0.283 | 0.278 | 0.293 | 0.271 | 0.265 |
| | 2015-Win. | 0.358 | 0.364 | 0.381 | 0.340 | 0.339 | 0.350 | 0.312 | 0.313 |
| | 2016-Spr. | 0.297 | 0.300 | 0.321 | 0.284 | 0.280 | 0.282 | 0.268 | 0.268 |
| NRMSE | 2014-Win. | 0.108 | 0.109 | 0.110 | 0.108 | 0.104 | 0.108 | 0.104 | 0.104 |
| | 2015-Spr. | 0.112 | 0.113 | 0.117 | 0.105 | 0.104 | 0.108 | 0.104 | 0.101 |
| | 2015-Sum. | 0.107 | 0.108 | 0.111 | 0.103 | 0.101 | 0.102 | 0.099 | 0.099 |
| | 2015-Aut. | 0.109 | 0.108 | 0.114 | 0.101 | 0.101 | 0.103 | 0.100 | 0.097 |
| | 2015-Win. | 0.099 | 0.100 | 0.104 | 0.094 | 0.093 | 0.095 | 0.089 | 0.087 |
| | 2016-Spr. | 0.097 | 0.097 | 0.105 | 0.093 | 0.092 | 0.092 | 0.089 | 0.087 |
| MAE | 2014-Win. | 0.228 | 0.232 | 0.236 | 0.226 | 0.220 | 0.230 | 0.220 | 0.217 |
| | 2015-Spr. | 0.197 | 0.199 | 0.207 | 0.187 | 0.182 | 0.191 | 0.183 | 0.178 |
| | 2015-Sum. | 0.184 | 0.186 | 0.193 | 0.177 | 0.172 | 0.175 | 0.172 | 0.168 |
| | 2015-Aut. | 0.197 | 0.198 | 0.210 | 0.184 | 0.182 | 0.189 | 0.178 | 0.173 |
| | 2015-Win. | 0.227 | 0.232 | 0.242 | 0.216 | 0.214 | 0.220 | 0.204 | 0.200 |
| | 2016-Spr. | 0.200 | 0.202 | 0.218 | 0.193 | 0.191 | 0.191 | 0.183 | 0.181 |

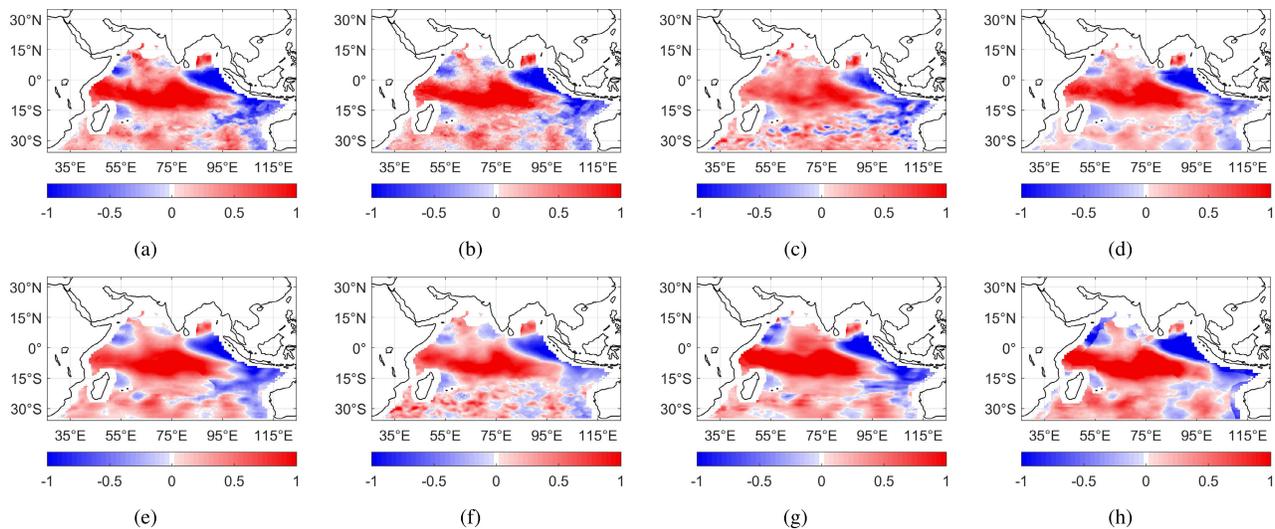


Fig. 9. Spatial distribution of STA reconstruction results at depth of 100 m for October 2015 using: (a) XGB; (b) RF; (c) FCNN; (d) Informer; (e) ConvNeXt; (f) TCN; (g) SCNN; and (h) SSCNN.

2.96E-10, 2.54E-07, 7.59E-06, 3.83E-08, and 0.012 between SSCNN and other methods, respectively. The XGB, RF, and FCNN methods have relatively worse accuracy for STA reconstruction since they lack the ability to automatically extract the remote sensing features or effectively express the ST relationships. The Informer, ConvNeXt, and TCN methods have better reconstruction accuracy than traditional machine learning methods, but they do not consider the distribution differences of information in LON and LAT directions or in various sea areas, thereby limiting the STA reconstruction performance. The results of SCNN indicate that single branch is insufficient for the parallel extraction of features at each time and the comprehensive coupling of features from multiple time series.

The spatial distribution of STA reconstruction results at a depth of 100 m for October 2015 using XGB, RF, FCNN, Informer, ConvNeXt, TCN, SCNN, and SSCNN methods is shown in Fig. 9. These methods can reconstruct the main positive anomaly signals in the equatorial central waters and the main negative anomaly signals in the equatorial eastern coastal waters, but the positive anomaly signals of FCNN are obviously weaker than that of other methods. The negative anomaly signals of SSCNN are more complete and continuous than that of other methods, considering its cooperative information extraction of spatial context and temporal sequence. The results of XGB, RF, FCNN, and TCN methods are more local and discrete than that of other methods in the LAT range from 15°S to 30°S, showing a worse temperature distribution

pattern. Overall, SSCNN method can improve the extraction and fusion of multisource spatial–temporal features through the design of multiscale regional division, asymmetric convolutional operation, and Siamese network structure, thereby improving the accuracy of STA reconstruction.

VI. CONCLUSION

In this study, a novel SSCNN is proposed to improve the accuracy of ST reconstruction in the Indian Ocean. The input feature maps are constructed using remote sensing data of SST and SSS, SWH, ADT, and SSW, combined with geographic location information. Multiscale regional division for accurate modeling, asymmetric convolutional operation for enhanced extraction, and Siamese network structure for spatial–temporal learning are designed in the SSCNN. The spatial distribution of SSCNN reconstruction is consistent with the IPRC Argo data. The R^2 , RMSE, and MAE of SSCNN method first increase and then decrease with growing depth across different seasons, while the NRMSE fluctuates in the increasing and decreasing trends with growing depth. The comparison experiments show that SSCNN method outperforms other methods of machine learning and deep learning in reconstructing the STA of Indian Ocean from 30 to 1000 m. The R^2 of SSCNN is 0.006–0.092, 0.030–0.182, 0.013–0.163, 0.044–0.193, 0.024–0.206, and 0.026–0.218 higher than that of other methods from 2014–Win. to 2016–Spr., respectively. The time series of SDI in Indian Ocean is calculated, and it demonstrates that the SDI series obtained from reconstructions of SSCNN model is consistent with that from observations of IPRC data, remaining within a reasonable error range.

Overall, this study provides a promising approach for effectively reconstructing the STA using remote sensing data and deep learning methods, which also offers valuable insights for analyzing the evolution of subsurface positive dipole in Indian Ocean. The SSCNN method is suitable for open ocean areas with sufficient Argo observational data. Further improvement is needed in deep learning’s ability and accuracy for reconstructing temperature from small samples in other high-LAT areas. Additionally, further optimization is required by integrating physical knowledge and oceanic equations for regions such as straits and bays. In the future, more high-resolution sea surface remote sensing data will be collected and employed to improve the accuracy and resolution of STA reconstruction. Moreover, deep learning methods will consider the remote sensing information of sea surface over longer time series and wider spatial contexts, and attention mechanism, self-supervised learning, and prior knowledge will be integrated to enhance the performance of deep learning model.

REFERENCES

- [1] R. S. Dhillon and G. von Wuehlisch, “Mitigation of global warming through renewable biomass,” *Biomass Bioenergy*, vol. 48, pp. 75–89, Jan. 2013.
- [2] G. C. Johnson and J. M. Lyman, “Warming trends increasingly dominate global ocean,” *Nature Climate Change*, vol. 10, no. 8, pp. 757–761, Aug. 2020.
- [3] B. Meyssignac et al., “Measuring global ocean heat content to estimate the Earth energy imbalance,” *Frontiers Mar. Sci.*, vol. 6, p. 432, Aug. 2019.
- [4] L. C. Allison et al., “Towards quantifying uncertainty in ocean heat content changes using synthetic profiles,” *Environ. Res. Lett.*, vol. 14, no. 8, Aug. 2019, Art. no. 084037.
- [5] X. Yan et al., “The global warming hiatus: Slowdown or redistribution?” *Earth’s Future*, vol. 4, no. 11, pp. 472–482, Nov. 2016.
- [6] L. Cheng et al., “Record-setting ocean warmth continued in 2019,” *Adv. Atmos. Sci.*, vol. 37, no. 2, pp. 137–142, Feb. 2020.
- [7] S. Qin, Q. Zhang, and B. Yin, “Seasonal variability in the thermohaline structure of the Western Pacific Warm Pool,” *Acta Oceanol. Sinica*, vol. 34, no. 7, pp. 44–53, Jul. 2015.
- [8] H. Su, Y. Wei, W. Lu, X.-H. Yan, and H. Zhang, “Unabated global ocean warming revealed by ocean heat content from remote sensing reconstruction,” *Remote Sens.*, vol. 15, no. 3, p. 566, Jan. 2023.
- [9] H. Wang, T. Song, S. Zhu, S. Yang, and L. Feng, “Subsurface temperature estimation from sea surface data using neural network models in the western Pacific ocean,” *Mathematics*, vol. 9, no. 8, p. 852, Apr. 2021.
- [10] B. Huang et al., “Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1,” *J. Climate*, vol. 34, no. 8, pp. 2923–2939, Apr. 2021.
- [11] H. Cheng, L. Sun, and J. Li, “Neural network approach to retrieving ocean subsurface temperatures from surface parameters observed by satellites,” *Water*, vol. 13, no. 3, p. 388, Feb. 2021.
- [12] D. Roemmich and J. Gilson, “The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo program,” *Prog. Oceanogr.*, vol. 82, no. 2, pp. 81–100, Aug. 2009.
- [13] S. Guinehut, A.-L. Dhomp, G. Larnicol, and P.-Y. Le Traon, “High resolution 3-D temperature and salinity fields derived from in situ and satellite observations,” *Ocean Sci.*, vol. 8, no. 5, pp. 845–857, Oct. 2012.
- [14] H.-J. Wang, H. Cheng, K.-W. Xie, F.-Y. Lu, and Y.-C. Du, “Applications of laser precisely processing technology in solar cells,” *Optoelectron. Lett.*, vol. 3, no. 5, pp. 385–387, Sep. 2007.
- [15] H. Su, X. Yang, W. Lu, and X.-H. Yan, “Estimating subsurface thermohaline structure of the global ocean using surface remote sensing observations,” *Remote Sens.*, vol. 11, no. 13, p. 1598, Jul. 2019.
- [16] H. Su, A. Wang, T. Zhang, T. Qin, X. Du, and X.-H. Yan, “Super-resolution of subsurface temperature field from remote sensing observations based on machine learning,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, Oct. 2021, Art. no. 102440.
- [17] A. Samat, E. Li, W. Wang, S. Liu, C. Lin, and J. Abuduwaili, “Meta-XGBoost for hyperspectral image classification using extended MSER-guided morphological profiles,” *Remote Sens.*, vol. 12, no. 12, p. 1973, Jun. 2020.
- [18] J. Miao et al., “Mapping seasonal leaf nutrients of mangrove with Sentinel-2 images and XGBoost method,” *Remote Sens.*, vol. 14, no. 15, p. 3679, Aug. 2022.
- [19] I. A. Girach, M. Ponmalar, S. Murugan, P. A. Rahman, S. S. Babu, and R. Ramachandran, “Applicability of machine learning model to simulate atmospheric CO₂ variability,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4107306.
- [20] S. M. Krishna Moorthy, K. Calders, M. B. Vicari, and H. Verbeeck, “Improved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3057–3070, May 2020.
- [21] H. Su, X. Wu, X.-H. Yan, and A. Kidwell, “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.*, vol. 160, pp. 63–71, Apr. 2015.
- [22] H. Su, W. Li, and X. Yan, “Retrieving temperature anomaly in the global subsurface and deeper ocean from satellite observations,” *J. Geophys. Res. Oceans*, vol. 123, no. 1, pp. 399–410, Jan. 2018.
- [23] X. Yang, H. Su, W. Li, L. Huang, X. Wang, and X. Yan, “Seasonal-spatial variations in satellite-derived global subsurface temperature anomalies,” *Nat. Remote Sens. Bull.*, vol. 23, no. 5, pp. 997–1010, 2019.
- [24] J. Qi, C. Liu, J. Chi, D. Li, L. Gao, and B. Yin, “An ensemble-based machine learning model for estimation of subsurface thermal structure in the South China Sea,” *Remote Sens.*, vol. 14, no. 13, p. 3207, Jul. 2022.
- [25] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, “Deep learning for single image super-resolution: A brief review,” *IEEE Trans. Multimedia*, vol. 21, no. 12, pp. 3106–3121, Dec. 2019.
- [26] N. Sun, Z. Zhou, Q. Li, and X. Zhou, “Spatiotemporal prediction of monthly sea subsurface temperature fields using a 3D U-Net-based model,” *Remote Sens.*, vol. 14, no. 19, p. 4890, Sep. 2022.

- [27] B. B. Nardelli, "A deep learning network to retrieve ocean hydrographic profiles from combined satellite and in situ measurements," *Remote Sens.*, vol. 12, no. 19, p. 3151, Sep. 2020.
- [28] H. Su, T. Zhang, M. Lin, W. Lu, and X.-H. Yan, "Predicting subsurface thermohaline structure from remote sensing data based on long short-term memory neural networks," *Remote Sens. Environ.*, vol. 260, Jul. 2021, Art. no. 112465.
- [29] H. Su, T. Qin, A. Wang, and W. Lu, "Reconstructing ocean heat content for revisiting global ocean warming from remote sensing perspectives," *Remote Sens.*, vol. 13, no. 19, p. 3799, Sep. 2021.
- [30] M. Han, Y. Feng, X. Zhao, C. Sun, F. Hong, and C. Liu, "A convolutional neural network using surface data to predict subsurface temperatures in the Pacific ocean," *IEEE Access*, vol. 7, pp. 172816–172829, 2019.
- [31] L. Meng, C. Yan, W. Zhuang, W. Zhang, X. Geng, and X.-H. Yan, "Reconstructing high-resolution ocean subsurface and interior temperature and salinity anomalies from satellite observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 3109979.
- [32] L. Meng, C. Yan, W. Zhuang, W. Zhang, and X. Yan, "Reconstruction of three-dimensional temperature and salinity fields from satellite observations," *J. Geophys. Res. Oceans*, vol. 126, no. 11, Nov. 2021, Art. no. e2021JC017605.
- [33] K. Zhang, X. Geng, and X.-H. Yan, "Prediction of 3-D ocean temperature by multilayer convolutional LSTM," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1303–1307, Aug. 2020.
- [34] B. Murali, M. S. G. Kumar, M. Ravichandran, and G. Bharathi, "Role of equatorial Indian ocean convection on the Indian summer monsoon," *MAUSAM*, vol. 72, no. 2, pp. 457–462, Oct. 2021.
- [35] M. Zhang, D. Jin, X. Wang, L. Chen, J. Luo, and Z. Wang, "Seasonal transition of precedent Indian ocean basin mode and subsequent Indian ocean dipole without el Niño–Southern oscillation impact," *Int. J. Climatol.*, vol. 42, no. 16, pp. 9023–9031, Dec. 2022.
- [36] V. K. Singh and M. K. Roxy, "A review of ocean-atmosphere interactions during tropical cyclones in the north Indian Ocean," *Earth-Sci. Rev.*, vol. 226, Mar. 2022, Art. no. 103967.
- [37] P. A. Utari, M. Y. N. Khakim, D. Setiabudidaya, and I. Iskandar, "Dynamics of 2015 positive Indian Ocean dipole," *J. Southern Hemisphere Earth Syst. Sci.*, vol. 69, no. 1, p. 75, 2019.
- [38] Y. Zhang and Y. Du, "Extreme IOD induced tropical Indian Ocean warming in 2020," *Geosci. Lett.*, vol. 8, no. 1, pp. 1–9, Dec. 2021.
- [39] L. Huang et al., "Decadal cooling events in the south Indian ocean during the Argo era," *J. Geophys. Res. Oceans*, vol. 127, no. 9, Sep. 2022, Art. no. e2021JC017949.
- [40] T. Song, W. Wei, F. Meng, J. Wang, R. Han, and D. Xu, "Inversion of ocean subsurface temperature and salinity fields based on spatio-temporal correlation," *Remote Sens.*, vol. 14, no. 11, p. 2587, May 2022.
- [41] B. B. Nardelli and R. Santoleri, "Reconstructing synthetic profiles from surface data," *J. Atmos. Ocean. Technol.*, vol. 21, no. 4, pp. 693–703, Apr. 2004.
- [42] Y. Jeong, J. Hwang, J. Park, C. J. Jang, and Y.-H. Jo, "Reconstructed 3-D ocean temperature derived from remotely sensed sea surface measurements for mixed layer depth analysis," *Remote Sens.*, vol. 11, no. 24, p. 3018, Dec. 2019.
- [43] H. Su, L. Huang, W. Li, X. Yang, and X. Yan, "Retrieving ocean subsurface temperature using a satellite-based geographically weighted regression model," *J. Geophys. Res. Oceans*, vol. 123, no. 8, pp. 5180–5193, Aug. 2018.
- [44] A. Shahdi, S. Lee, A. Karpatne, and B. Nojabaei, "Exploratory analysis of machine learning methods in predicting subsurface temperature and geothermal gradient of northeastern United States," *Geothermal Energy*, vol. 9, no. 1, pp. 1–22, Dec. 2021.
- [45] A. Barth, A. A. Azcárate, M. Licer, and J.-M. Beckers, "A convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations (DINCAE)," in *Proc. EGU Gen. Assem. Conf. Abstr.*, 2020, p. 9414.
- [46] W. Li, H. Su, X. Wang, and X. Yan, "Estimation of global subsurface temperature anomaly based on multisource satellite observations," *J. Remote Sens.*, vol. 21, pp. 881–891, Jan. 2017.
- [47] L. Zhong, L. Hu, and H. Zhou, "Deep learning based multi-temporal crop classification," *Remote Sens. Environ.*, vol. 221, pp. 430–443, Feb. 2019.
- [48] D. Bhatt et al., "CNN variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, Oct. 2021.
- [49] A. Rostamian and J. G. O'Hara, "Event prediction within directional change framework using a CNN-LSTM model," *Neural Comput. Appl.*, vol. 34, no. 20, pp. 17193–17205, Oct. 2022.
- [50] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021.
- [51] X. Liu, Z. Zhang, F. Meng, and Y. Zhang, "Fault diagnosis of wind turbine bearings based on CNN and SSA-ELM," *J. Vibrot. Eng. Technol.*, vol. 11, pp. 3929–3945, Nov. 2022.
- [52] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 22, no. 10, pp. 1533–1545, Oct. 2014.
- [53] J. P. A. Vieira and R. S. Moura, "An analysis of convolutional neural networks for sentence classification," in *Proc. 43th Latin Amer. Comput. Conf. (CLEI)*, Sep. 2017, pp. 1–5.
- [54] X. Wu, X.-H. Yan, Y.-H. Jo, and W. T. Liu, "Estimation of subsurface temperature anomaly in the North Atlantic using a self-organizing map neural network," *J. Atmos. Ocean. Technol.*, vol. 29, no. 11, pp. 1675–1688, Nov. 2012.
- [55] C. Chen, K. Yang, Y. Ma, and Y. Wang, "Reconstructing the subsurface temperature field by using sea surface data through self-organizing map method," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1812–1816, Dec. 2018.
- [56] W. Lu, H. Su, X. Yang, and X.-H. Yan, "Subsurface temperature estimation from remote sensing data using a clustering-neural network method," *Remote Sens. Environ.*, vol. 229, pp. 213–222, Aug. 2019.
- [57] M. Sannarino, B. B. Nardelli, S. Marullo, and R. Santoleri, "An artificial neural network to infer the Mediterranean 3D chlorophyll-A and temperature fields from remote sensing observations," *Remote Sens.*, vol. 12, no. 24, p. 4123, Dec. 2020.
- [58] N. Verbrugge, S. Mulet, S. Guinehut, and B. Buongiorno-Nardelli, "ARMOR3D: A 3D multi-observations T, S, U, V, product of the ocean," in *Proc. 19th EGU Gen. Assem. EGU From Conf. Held*, Apr. 2017, p. 17579.
- [59] S. Mulet, M.-H. Rio, A. Mignot, S. Guinehut, and R. Morrow, "A new estimate of the global 3D geostrophic ocean circulation based on satellite data and in-situ measurements," *Deep Sea Res. II, Top. Stud. Oceanogr.*, vols. 77–80, pp. 70–81, Nov. 2012.
- [60] G. Song and R. C. Ren, "Linking the subsurface Indian Ocean dipole to central Pacific ENSO," *Geophys. Res. Lett.*, vol. 49, no. 5, Mar. 2022, Art. no. e2021GL096263.
- [61] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
- [62] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [63] A. Borovykh, C. W. Oosterlee, and S. M. Bohté, "Generalization in fully-connected neural networks for time series forecasting," *J. Comput. Sci.*, vol. 36, Sep. 2019, Art. no. 101020.
- [64] H. Zhou et al., "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proc. AAAI Conf. Artif. Intell.*, May 2021, vol. 35, no. 12, pp. 11106–11115.
- [65] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11966–11976.
- [66] P. Hewage et al., "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station," *Soft Comput.*, vol. 24, no. 21, pp. 16453–16482, Nov. 2020.



Shuyu Zhang received the B.S. and Ph.D. degrees from the College of Earth Sciences, Zhejiang University, Hangzhou, China, in 2015 and 2020, respectively.

She is currently an Assistant Professor with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. Her research interests include remote sensing image analysis, ocean data processing, and deep learning.



Yizhou Yang received the B.E. degree in software engineering from the Guangdong University of Foreign Studies, Guangzhou, China, in 2022. He is currently pursuing the M.E. degree with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

His research interests include the intelligent processing of ocean data and deep learning.



Gongjie Wang received the B.S. degree in marine science from the Ocean University of China, Qingdao, China, in 2012, and the Ph.D. degree in atmospheric science from the National University of Defense Technology, Changsha, China, in 2018.

He is currently a Senior Engineer with PLA 31526 Troops, Beijing, China. His research interests include oceanic subsurface temperature/salinity reconstruction and short-term climate prediction.



Kangwen Xie received the B.E. degree in computer science and technology from Guangzhou University, Guangzhou, China, in 2020, and the M.E. degree in electronic information from the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, in 2023.

His research interests include ocean subsurface analysis and machine learning.



Zhihui Che received the B.S. degree from the Ocean University of China, Qingdao, China, in 2016, and the Ph.D. degree from the University of Chinese Academy of Sciences, Beijing, China, in 2021.

He is currently a Post-Doctoral Researcher with the College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China. His research interests include ocean analysis, climate change, and deep learning.



Jiahao Gao received the B.E. degree in computer science and technology from Henan Polytechnic University, Jiaozuo, China, in 2021. He is currently pursuing the M.E. degree with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

His research interests include thermohaline structure reconstruction and remote sensing data analysis.



Lin Mu received the B.S., M.S., and Ph.D. degrees in physical oceanography from the Ocean University of China, Qingdao, China, in 2000, 2002, and 2007, respectively.

He is currently a Full Professor with the College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China. His research interests include physical oceanography, prevention and mitigation of marine disasters, maritime search and rescue, and emergency response management of offshore oil spills.



Zhiyuan Zhang received the B.E. degree from the National University of Defense Technology, Changsha, China, in 2003, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2015.

Since 2003, he has been with the PLA, Beijing, where he is currently a Senior Engineer. His research interests include the numerical forecasting of marine environment, artificial intelligence, and high-performance computing.



Qianru Niu received the B.S. degree in physical oceanography from the Ocean University of China, Qingdao, China, in 2012, and the M.S. and Ph.D. degrees from the University of Maryland, College Park, MD, USA, in 2015 and 2017, respectively.

She is currently an Associate Professor with the College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China. Her research interests include physical oceanography, ocean modeling, wave-current interactions, and turbulent mixing.



Sen Jia (Senior Member, IEEE) received the B.E. and Ph.D. degrees from the College of Computer Science, Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively.

Since 2008, he has been with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, where he is currently a Full Professor. His research interests include remote sensing image processing, signal and image processing, and machine learning.