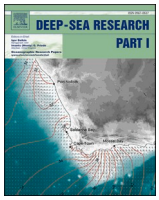
Deep–Sea Research I 202 (2023) 104179

Contents lists available at ScienceDirect

Deep-Sea Research Part I

journal homepage: www.elsevier.com/locate/dsri



Estimation of the barrier layer thickness in the Indian Ocean based on hybrid neural network model

Yizhi Zhao [a,c,d], Jifeng Qi [b], Shanliang Zhu [a,c,d,*], Wentao Jia [a,c,d], Xiang Gong [a,c,d], Wenming Yin [a,c,d], Baoshu Yin [b]

[a] School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, China
[b] CAS Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China
[c] Qingdao Innovation Center of Artificial Intelligence Ocean Technology, Qingdao, China
[d] Research Institute for Mathematics and Interdisciplinary Sciences, Qingdao University of Science and Technology, Qingdao, China

ARTICLE INFO

Keywords:
Barrier layer thickness
Particle swarm optimization
Artificial neural networks
Hybrid models

ABSTRACT

Accurately and effectively estimating of the barrier layer thickness (BLT) is essential for the research of ocean thermodynamics, ocean dynamics, and air-sea interaction. Artificial intelligence model provides an effective means for accurately estimating BLT from sea surface and gridded Argo data. The present study focuses on the application of a hybrid particle swarm optimization-based artificial neural networks model (PSO-ANN) for estimating the BLT in the Indian Ocean. The input variables of the hybrid model include sea surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH) and precipitation (P), and the output variable is the BLT value. Multivariate satellite and gridded Argo data collected from the Indian Ocean between January 2012 and December 2019 (i.e., a database consists 239,568 training datasets and 34,224 testing datasets) are provided to prepare the training and testing datasets for the model. The parameters of ANN model, such as network parameter, network weights, and dropout rates are optimized using the PSO algorithm to achieve the best estimation model. R-squared (R^2) and root mean square error (RMSE) are used to evaluate the performance of the model. Two groups of comparative experiments (Case 1 and Case 2) on the performance of the PSO-ANN model demonstrate that the model in Case 2 can better capture the complex features of the BLT in the ocean region. The performance of the PSO-ANN model in Case 2 is further compared with the data-driven estimation models such as the traditional ANN model and the known multilinear regression model (MRM), as well as the CESM2-WACCM dynamic model from CMIP6. The comparison results show that the dynamic model has the worst performance among the four models. Moreover, the annual average RMSE value for the PSO-ANN model is 1.83 m, which is 12% and 84% lower than that of the traditional ANN and MRM, respectively. The R^2 value for the model of 0.85 is improved by 4% and 40% compared to that of two models. Furthermore, three regions with significant seasonal fluctuations of the BLT in the Indian Ocean are selected to further evaluate the estimation accuracy of the hybrid model in 2019; the Southeast Arabian Sea (SEAS), the Bay of Bengal (BoB), and the Eastern Equatorial Indian Ocean (EEIO). As a result, the hybrid model is capable of reflecting seasonal variation trends in these regions, but there is room for improvement in the estimation accuracy. In addition, the correlation analysis between BLT and sea surface parameters indicates that there exist significant correlations between the BLT and SSS, P. The results of this study show that the proposed hybrid model can be used to estimate and analyze BLT in certain regions with complex ocean dynamics processes. Moreover, the model can be extended to estimate other key ocean parameters and provide effective technical support for studying the internal ocean parameters.

1. Introduction

The barrier layer is defined as the layer between the mixed layer and the isothermal layer, which is named for its ability to prevent heat and momentum exchange between the sea surface and below the isothermal layer (Vialard and Delecluse, 1998a,b). Accurately estimating of the barrier layer thickness (BLT) is helpful in understand the mechanisms of the upper ocean in thermal effects and dynamical action. On the basis of

* Corresponding author.School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, China.
E-mail address: zhushanliang@qust.edu.cn (S. Zhu).

https://doi.org/10.1016/j.dsr.2023.104179
Received 23 July 2023; Received in revised form 7 October 2023; Accepted 12 October 2023
Available online 18 October 2023

thermodynamics, the barrier layer maintains a vertically uniform temperature distribution, while the stable salinity layer inhibits the vertical transfer of heat flux from the sea surface to the ocean interior. As a result, the heat exchange between the mixed layer and the isothermal layer becomes inefficient (Godfrey and Lindstrom, 1989; Lukas and Lindstrom, 1987; Vialard and Delecluse, 1998a,b). Thus, accurate estimation of BLT contributes to understand the heat exchange mechanism in and around the ocean. From the perspective of dynamics, the momentum of wind disturbance is constrained within the mixed layer, which limits its vertical transmission. Furthermore, the velocity of wind-driven currents increases due to the energy accumulation (Roemmich et al., 1994). In addition, the barrier layer plays an important role in monitoring extreme weather and climate events, which can affect local air-sea interactions and further influences climate change in some locations and even larger regions, such as El Niño-Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) and Tropical Intra-Seasonal Oscillations (Drushka et al., 2014). Moreover, the vertical distribution of chlorophyll concentration in the upper ocean showed significant differences with and without barrier layer (Kawamiya and Oschlies, 2001). It can be concluded that accurate and effective estimation of BLT is of great significance for predicting climate change, the utilization and development of marine resources.

Due to the significant impact of BLT in the field of oceanography, accurate estimation of BLT has always been a critical research topic for oceanographers over the past few decades. In the earlier days, limited observational data greatly restricted the measurement of BLT. Early research mainly explored the temporal variability of BLT on relatively short time scales by using conductivity-temperature-depth (CTD) and buoy observation data (Cronin and Meghan, 2002; Murty et al., 1996; Sprintall and Tomczak, 1992). However, these observational data had drawbacks, i.e., the spatiotemporal distribution is uneven and discontinuous. With the increase of profile data collection in temperature and salinity, gridded Argo data has been widely used in various ocean data analyses (de Boyer Montégut et al., 2007; Mignot et al., 2007). The seasonal variation and formation mechanism of BLT have become a hot topic for oceanographers to study (Bosc et al., 2009; Drushka et al., 2014; Liu et al., 2009; Qu et al., 2014; Wang and Liu, 2016; Zheng et al., 2014). In fact, the horizontal resolution of these gridded datasets is not sufficient to observe the small to medium scale changes of BLT, therefore the accuracy issues caused by this drawback still need to be resolved.

Due to the difficulty of conducting ocean surveys, the temporal and spatial resolution of observational data is not high enough, which limits our understanding of ocean mechanisms and internal structure of the ocean. To solve such resolution problems, many researchers have tried to apply various methods to estimate subsurface temperature and salinity from multi-source sea surface data, such as parametric models (Chu et al., 2000), dynamical models (Courtois et al., 2017) and statistical models (Guinehut et al., 2012). In recent years, artificial intelligence methods have also been successfully used to estimate key parameters in the ocean interior, such as light gradient boosting machine (LightGBM) (Su et al., 2021), long short-term memory (LSTM) model (Zhang et al., 2023), and deep neural network (DNN) model (Wang et al., 2023). In particular, as surveying and remote sensing technologies develop continuously, satellite remote sensing data are becoming more and more abundant, which have become an important component of multi-source sea surface data information (Dong et al., 2011). These data can reveal the structure of many oceanic phenomena, such as tropical instability waves and seasonal variations in salt fluxes in the Indian Ocean (Nyadjro, 2013; Zheng and Zhang, 2012). Considering the advantages of high precision, broad coverage, and strong spatiotemporal continuity of satellite remote sensing data, the feasibility of using these data to estimate BLT is studied.

Previous researches have shown that the limitations of observational data and models have resulted in some remaining issues in the estimation of internal ocean parameters. For example, traditional statistical methods use only one or a few parameters and are lack of accuracy (Su et al., 2015), while dynamic simulations require significant computational resources (Su et al., 2019). Moreover, despite their success in the ocean field, each artificial intelligence method has its strengths and shortcomings (Qi et al., 2023). Therefore, there is still considerable room for improvement in the estimation model itself and its accuracy. In recent years, it has been shown that hybrid neural network models have the advantages of high accuracy and reliability in terms of predictive performance compared to most single neural network models (Vahedoost et al., 2020). In particular, some swarm intelligence algorithms have been used to optimize neural network parameters, which effectively improves the predictive performance of hybrid models (Elbeltagi et al., 2005; Koopialipoor et al., 2019; Moayedi et al., 2019). The particle swarm optimization (PSO), a typical swarm intelligence algorithm, has been well known and adopted by its simplicity, high solution quality, and excellent convergence properties (Sahraei and Codur, 2022). The research has indicated that the PSO uses less memory and has higher learning speed than genetic algorithm (GA), imperialist competition algorithm (ICA) in hybrid models (Moayedi et al., 2019). The comparative results have also shown that the PSO generally perform better than some algorithms such as GA, memetic algorithms (MA), ant-colony systems (ACO) and shuffled frog leaping (SFL) in terms of success rate and solution quality (Elbeltagi et al., 2005). These achievements are of great significance for the development of new BLT estimation models in this study.

Motivated by the aforementioned discussions, the primary objective of this study is to investigate ways to address the BLT estimation in a typical ocean region with complex ocean dynamics by developing a new artificial intelligence method based on sea surface parameter data. Another significant objective of this study is to attempt a comprehensive analysis of the estimation accuracy and investigate the relationship between BLT and various sea surface parameters. To achieve these goals, a hybrid model combining PSO algorithm and artificial neural networks (PSO-ANN) is proposed using multi-source sea surface satellite remote sensing data in the Indian Ocean as a case study. The significant contribution is that the findings of this study have implications for detecting and tracking the barrier layer in the Indian Ocean and provide insights for further scientific understanding of oceanographic processes. Moreover, the model used in this study can be extended to numerically estimate other subsurface variables in the global ocean.

This study is arranged as follows. Section 1 presents the overview and significance of this study. The literature review related to this study is presented in Section 2, such as data sources for BLT, the development of traditional BLT estimation and the application of artificial intelligence estimation methods in oceanography. Section 3 describes the methods used in this study, mainly including the data sources of the study area, the extraction and processing of the data, and the setup of the used model. The results and analysis of this research are presented in Section 4. Finally, Section 5 presents the conclusion and outlook of future research, and displays the shortcomings and limitations of this study.

2. Literature review

Over the past few decades, physical oceanographers have conducted numerous observational studies on the spatiotemporal distribution and physical influences of BLT in the world's oceans. Sprintall and Tomczak (1992) were pioneers in investigating the global distribution of BLT by utilizing temperature and salinity climatological data. They suggested that the salinity stratification at the upper level, which gives rise to BLT, is produced by monsoon-related rainfall and river runoff in the Indian Ocean. These efforts were significant in advancing our understanding of the distribution of BLT in the ocean. Subsequently, Tomczak and Godfrey (1994), as well as Monterey and Levitus (1997), produced seasonal distribution maps of BLT worldwide by using different global ocean datasets. Nonetheless, these studies share some common limitations. Due to the lack of sufficient observational data, the oceanographic data used in these studies underwent processes such as vertical averaging and

interpolation of original temperature, salinity, and density profiles. As a result, the determination of BLT had great errors.

In recent years, with the increasing abundance of Argo data, remote sensing data and reanalysis data, oceanographers have paid increasing attention to estimate the BLT by using multi-source ocean data. New methods based on numerical or statistical models have also been proposed to estimate the BLT and analyze its changes. By using eddy-resolving numerical simulation, Saha et al. (2021) investigated the mechanism of the northwest tropical Atlantic barrier layer formation, to highlight the relationship between BLT and freshwater distribution. Meanwhile, Valsala et al. (2018) explored the interannual variability of mixed layer in the Bay of Bengal by using an oceanic thermodynamic model, which estimated the seasonal cycle of the BLT in the tropical Indian Ocean via coupling mechanisms. In addition, Felton et al. (2015) used a multilinear regression model (MRM) to estimate the BLT and evaluated the relationship between satellite data and Argo data. They also analyzed the seasonal and interannual variations of BLT in the Indian Ocean.

In recent times, some advanced artificial intelligence methods have been used to estimate internal ocean parameters. Owing to the large-scale and spatiotemporal nature of ocean data, new artificial intelligence methods can effectively extract the intrinsic relationships between them, which brings new feasibility to the estimation and analysis of ocean data (Schott et al., 2009). Su et al. (2021) applied LightGBM algorithm to estimate subsurface temperature anomaly (STA) using multiple-source satellite observation data. Additionally, some deep learning models were employed. For example, Ali et al. (2004) used the neural network method to estimate the thermal structure of the subsurface layer by using parameters such as sea surface temperature (SST), sea surface height (SSH), wind stress, net radiation, and net heat flux. Pauthenet et al. (2022) presented a Multilayer Perceptron (MLP) method that predicted temperature, salinity and mixed layer depth (MLD) from the surface satellite data in the Gulf Stream region. Qi et al. (2023) estimated the ocean subsurface temperature and salinity in the tropical Indian Ocean down to 1000 m using a new convolutional neural network (CNN). Zhang et al. (2020) proposed a multilayer convolutional long short term memory (M-ConvLSTM) model to predict the horizontal and vertical temperature changes from the sea surface to about 2000 m underwater. By virtue of DNN model, Meng et al. (2020) estimated the high-resolution STA and salinity anomaly (SSA) in the Pacific Ocean by using high-resolution sea surface data. In order to further improve the estimation accuracy, some hybrid estimation models were also proposed. Combining the light gradient boosting machine (LightGBM) model and Deep Forest method, Dong et al. (2022) constructed a hybrid model to predict the subsurface salinity using the observational data of SST, sea surface salinity (SSS), SSH, and sea surface wind (SSW). Recently, Gu et al. (2022) combined the K-means clustering algorithm and the traditional ANN model to estimate the MLD using the SST, SSS, SSH and SSW data.

As mentioned above, the results of their studies indicated that the estimation accuracy of the hybrid estimation model surpasses the traditional single ANN model. However, within these applications, the network structure of the ANN model and the optimized network parameters were difficult to determine. Moreover, the learning algorithm was prone to local optimization and slow convergence, which restricted the estimation performance of ANN (Momeni et al., 2014). To solve these issues, a hybrid PSO-ANN model is proposed to estimate BLT in the Indian Ocean using SSS, SST, SSH, precipitation (P), longitude (LON) and latitude (LAT) as input variables. Further, the performance of this model is evaluated by comparing it with the traditional ANN model, MRM and CESM2-WACCM model. Finally, the Pearson correlation coefficient between BLT and sea surface parameters is applied to quantitatively analyze the correlation between these variables.

3. Methodology

This section aims to show the workflow of this study, which includes study design, study area, data sources and processing methods, and PSO-ANN models. This section describes the data and models in detail. The data section includes the collection and processing of input and output variables, and the model section includes ANN, PSO, and hybrid models.

3.1. Research design

The present study employs a hybrid neural network model, and augments the performance of the network through optimization algorithms. The aim of this research is to provide model support for deriving the BLT of various ocean regions based on satellite observation data. The research workflow consists of five stages. The first stage collects and processes raw satellite data from various public databases. Four sea surface parameters (SSS, SST, SSH, P) and geographic information (LAT, LON) for each data are used in this stage. In the second stage, the training and testing sets are established. BLT is computed by using Argo's temperature and salinity data. The third stage divides all the data into 12 months and builds hybrid models for the 12 months. The fourth stage trains the neural network model and performs PSO. Four sea surface satellite parameters and geographic information of each month are selected as input variables for the model. The optimized neural network structure can be obtained by the PSO. The fifth stage applies the model to estimate the BLT. The accuracy and reliability of this model are evaluated by using the root mean square error (RMSE) and the R-squared (R^2), where R^2, also known as a determination coefficient with a value range from 0 to 1, is obtained from the real and estimated values and represent a measure of the fit degree of the model. The model has higher fitting degree as the R^2 value approaches 1.

3.2. Study area

In this study, the Indian Ocean (in the range 30°E~120°E and 30°S~30°N) is selected as the study area for BLT estimation, as shown in Fig. 1. The Indian Ocean is the warmest and third-largest ocean in the world. The tropical Indian Ocean forms the majority of the largest warm pool on Earth, which has a large impact on shaping the climate on both regional and global scales (More et al., 2020). Changes in Indian Ocean temperature have a significant impact on summer climate change in the Indian-Western Pacific region and play a crucial role in global climate change (Schiller and Oke, 2015). Compared with the tropical Pacific and Atlantic Oceans, the Indian Ocean is characterized by shallow isothermal depth (ILD) in the west and strong interannual variability in upper ocean temperatures in the east, which provides a unique sea area for assessing the seasonal and interannual variability of BLT. Therefore, the Indian Ocean is selected as the study area of BLT in this study. In addition, three special sea areas with relatively large BLT variability are

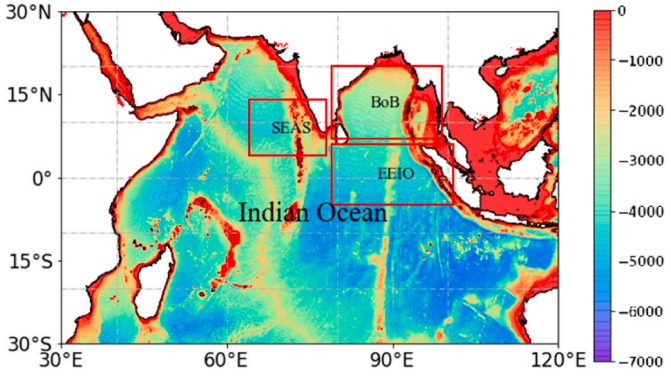


Fig. 1. Bathymetry (m) of the study area used in this research.

selected for analysis; the Southeast Arabian Sea (SEAS, 4°N-14°N, 64°E-77°E), the Bay of Bengal (BoB, 7°N-20°N, 79°E-96°E), and the Eastern Equatorial Indian Ocean (EEIO, 5°S-6°N, 79°E-101°E), which are showed in Fig. 1 with three red box.

3.3. Ocean data collection

The data used in this study are sea surface satellite observations (SSS, SST, SSH, and P), Argo temperature and salinity data (0-2000 m) used to calculate BLT. The ocean data presented in this article span from January 2012 to December 2019, covering a total of 96 months. The data information is summarized in Table 1. The SSS data are 0.25° × 0.25° gridded data acquired by Soil Moisture and Ocean Salinity (SMOS) (Boutin et al., 2018). The SST data are from the National Oceanic and Atmospheric Administration (NOAA) with a spatial resolution of 1° × 1° (Reynolds et al., 2001). The SSH data are obtained from the Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO) data project, with a spatial resolution of 0.25° × 0.25° (Morimoto, 2009). The P data are sourced from the Global Precipitation Climatology Project (GPCP) grid data, which provide satellite gridded data at high resolution (0.25° × 0.25°) (Behrangi and Song, 2020). The gridded Argo data are obtained from the Asia Pacific Data Research Center (APDRC) with a spatial resolution of 1° × 1°, which contain the data for 27 ocean layers from the surface to 2000 m below the seafloor (Roemmich and Gilson, 2009). Since the relationship between the sea surface data and BLT varies with the spatial location information, the geographic information of the study sea area, i.e., LON and LAT spatial location information, are considered in the model. In comparative experiments of Section 4.5, the CESM2-WACCM dynamic model-derived BLT values are calculated from its temperature and salinity profile data obtained from the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Liang et al., 2022).

3.4. Extraction and processing of the ocean data

Python and its software library are used in this study to extract and process the collected sea surface data. First, considering the differences in resolution and time periods between the variable and labeled data, all collected data are unified to the same 1° × 1° spatial resolution based on 2D interpolation method. All data are calculated as monthly averages in the time dimension so that all data have the same spatial and temporal coverage for the Indian Ocean. The mask data for coastal areas and islands are treated as null values. In addition, if any parameter variable of the spatial location point in the study area is null value, this point will be deleted.

Considering the effects of salinity and temperature, ILD and MLD are determined from gridded Argo data by using the criteria of (Sprintall and Tomczak, 1992). ILD is defined as the depth at which the difference with SST reached the temperature difference ΔT, which in this study is defined as $0.5°C$. MLD is defined as the depth at which the density difference is greater than $\Delta\sigma$, with the following standard for using density difference.

$$\Delta\sigma = \sigma_t(T_0 - \Delta T, S_0, P_0) - \sigma_t(T_0, S_0, P_0) \quad (1)$$

Where σ_t is the potential density (kg/m^{-3}) of the calculated layers and $\Delta\sigma$ represents the difference in potential density between two adjacent layers. T_0, S_0, P_0 denote the SST, SSS, and sea surface pressure, respectively. The MLD is then determined by searching each profile and calculating the $\Delta\sigma$ between the profiles. If the density value falls between two Argo levels, then linear interpolation method is used to calculate the MLD. BLT is defined as the difference between the ILD and the MLD (BLT=ILD-MLD). The processing flow of all data is shown in Fig. 2. The CESM2-WACCM dynamic model-derived BLT can be calculated using the similar methods mentioned above.

After the data preprocessing phase, a total of 96 months of valid datasets are obtained from January 2012 to December 2019, with 2724 valid data points for each month. The data from January 2012 to December 2018 are divided into 12 portions per year by month, each portion consists of the same month of data for each year and are used to train the neural network. The 2019 data are used as testing set to evaluate the model. Specifically, the BLT calculated from the gridded Argo temperature and salinity profile data are used as labels for training set and testing set. To eliminate the effect of different dimensions and orders of magnitude of the parameters on the model, the above data are normalized before training and testing.

Table 1
Data summary.

Index	Input Variable	Data Source	Output Variable	Data Source	Time Range	Time/ Spatial Resolution
Data	SSS SST SSH P	SMOS NOAA AVISO GPCP	BLT	Argo calculation	2012-2019	Monthly/ 1° × 1°

3.5. Artificial neural network

ANN is a powerful tool for dealing with regression and classification problems in nonlinear case, which consist of a number of interconnected neurons. Each neuron receives input signals from the input layer, multiplies the signal by its weight, sums the values, and then passes the calculated signal to the next layer. The signal propagates in the same way through the hidden layers before being finally passed to the output layer to get the output target. The ANN structure used in this study is shown in Fig. 3. Compared with traditional statistical methods, ANN is a data-driven adaptive method (Wang et al., 2019), which has excellent intelligent characteristics. In recent years, ANNs have successfully solved practical problems in many fields, such as traffic flow prediction (Olayode et al., 2021), subsurface temperature reconstruction (Lu et al., 2019; Wang et al., 2021), blasting problems (Dimitraki et al., 2018; Nguyen et al., 2020a,b) and some other aspects (Asteris et al., 2016; Cavaleri et al., 2017).

While ANN models are widely used, they also have drawbacks such as difficulty in determining the network structure and parameters during the learning process, overfitting and poor prediction (Chen et al., 2017; Ghorbani et al., 2017). The back-propagation algorithm of the model also tends to fall into local optima, which has the problem of slow convergence (Chau, 2006). These inherent drawbacks greatly limit the performance of ANN models in practical applications. To overcome these problems, the optimization of ANN through intelligent algorithms has gradually become a research hot topic in the application of ANNs, and has yielded significant research results in various fields (Grisales-Norea et al., 2020; Seifi and Soroush, 2020; Yousaf et al., 2020). For complex multidimensional ocean data, an improved PSO algorithm is used in this study to optimize each parameter of ANN, and it is applied to the estimation model of the BLT in the Indian Ocean.

3.6. Particle swarm optimization

The PSO algorithm is a stochastic, parallel modern evolutionary algorithm proposed by Kennedy and Eberhart (Kennedy, 2011). The algorithm is widely used because of its simplicity, ease of implementation and efficient convergence (Sahraei and Codur, 2022; Nguyen et al., 2020a,b; Yang et al., 2019). The fundamental concept is to find the optimal solution through collaboration and information-sharing among individuals within a group. Particles continuously change its position in a multidirectional search area, record improved locations during their movements, and ultimately discover the optimal solution. The population particle-matrix with d-dimensional and n number of particles is as given in Eq. (2).

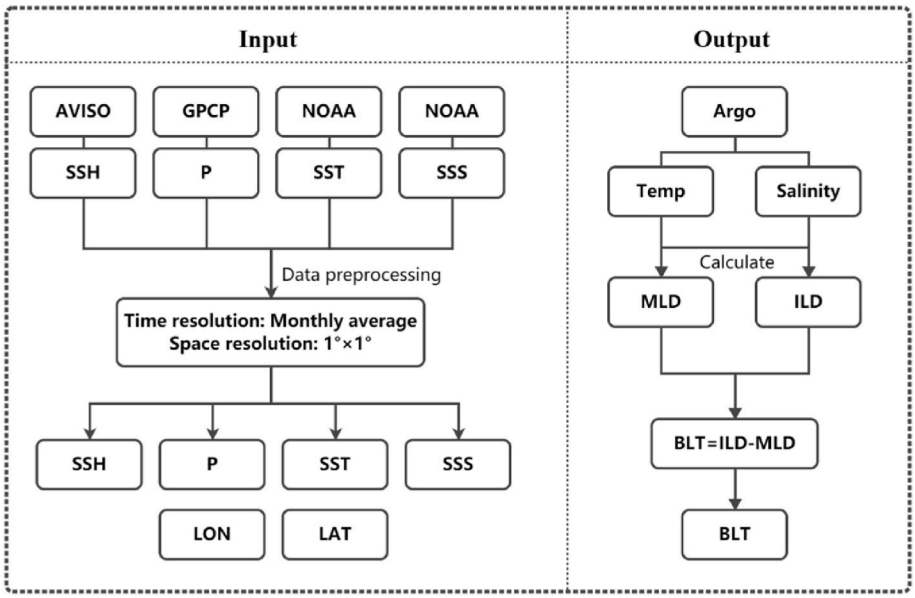


Fig. 2. Data processing flow chart.

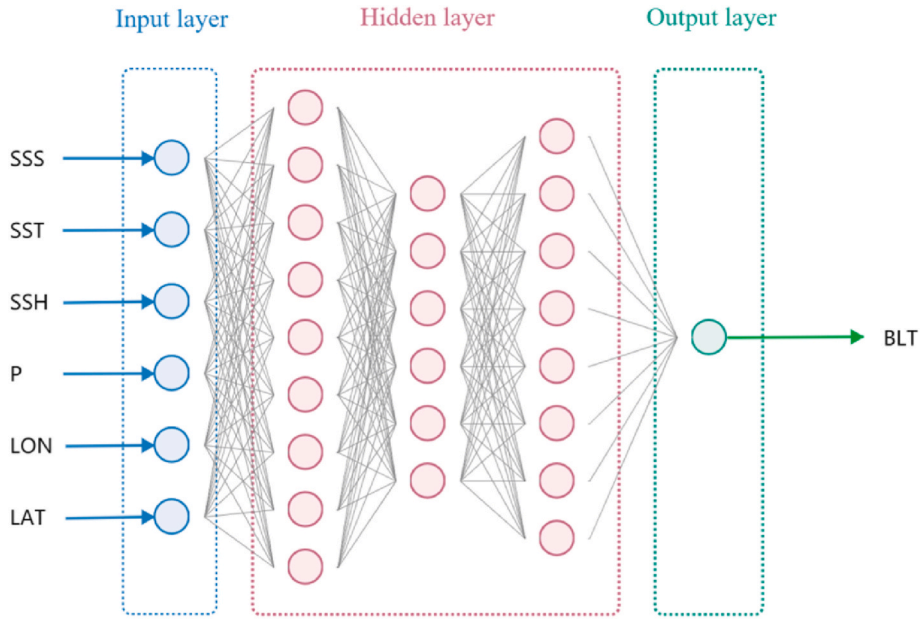


Fig. 3. The structural framework of ANN model.

$$X_i = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nd} \end{bmatrix} \quad (2)$$

Where the i-th particle is denoted as $x_i = [x_{i1}, x_{i2} \cdots x_{id}]$. A velocity parameter can be set for each solution, which the velocity of each particle is denoted as $v_i = [v_{i1}, v_{i2} \cdots v_{id}]$. The fitness value of each compound is determined by calculating the value of RMSE. The current local best position is recorded as the local best vector P_{best}, denoted as $P_{best} = [p_{i1}, p_{i2} \cdots p_{id}]$. The global optimal position of all particles is G_{best}, which represents the global best among the current local best. G_{best} is unique for all particles in the iteration and is denoted as $G_{best} = [g_{i1}, g_{i2} \cdots g_{id}]$.

After identifying these two optimal values, the particle velocity and position are updated throughout the iteration process. The position and velocity updates for the compound in the n-th iteration are given by the following two equations:

$$v_{ij}^{t+1} = \omega \times v_{ij} + c_1 \times rand1 \times \left(p_{ij} - x_{ij}^t\right) + c_2 \times rand2 \times \left(p_{gj} - x_{ij}^t\right) \quad (3)$$

$$x_{ij}^{t+1} = x_{ij}^t + v_{ij}^{t+1}, j = 1, 2, \ldots, d; i = 1, 2, \ldots, n \quad (4)$$

Among them, c_1 is called the individual learning factor and c_2 is called the social learning factor, which influence the individual and global learning ability. ω is called the inertia factor, which influences the individual and global search ability. The PSO search strategy is shown in Fig. 4.

3.7. Hybrid network model

In order to improve the estimation accuracy and reliability of the traditional ANN model, the PSO-ANN is used to build an estimation model of the BLT in the Indian Ocean. An ANN model consisting of an

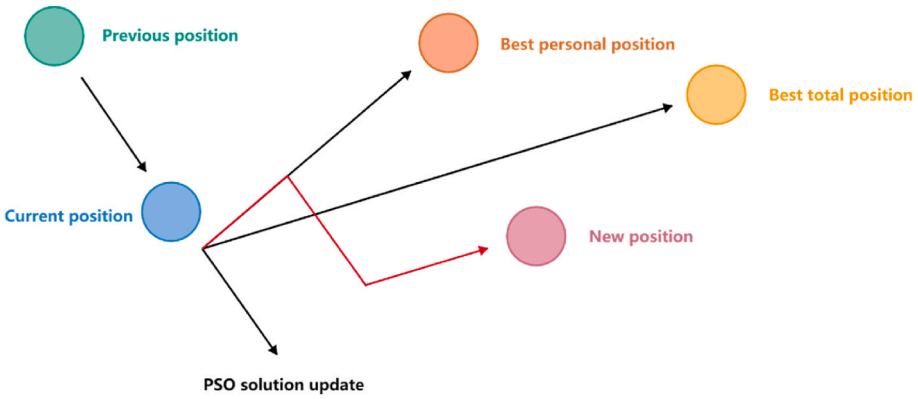


Fig. 4. PSO search strategy.

input layer, an output layer, one or several hidden layers and a dropout layer will be used, whose main parameters include the number of network layers, the number of nodes in each layer and the probability of dropout. Each hidden layer in these models comprises one or more neurons, with each neuron in the layer calculating the weighted sum of all the outputs from the previous layer, transforming it with a nonlinear activation function, and then outputting the result to the next layer. In this study, the rectified linear unit (ReLU) function is chosen as the activation function of the hidden layer. Compared to some traditional activation functions, the ReLU function not only saves computational

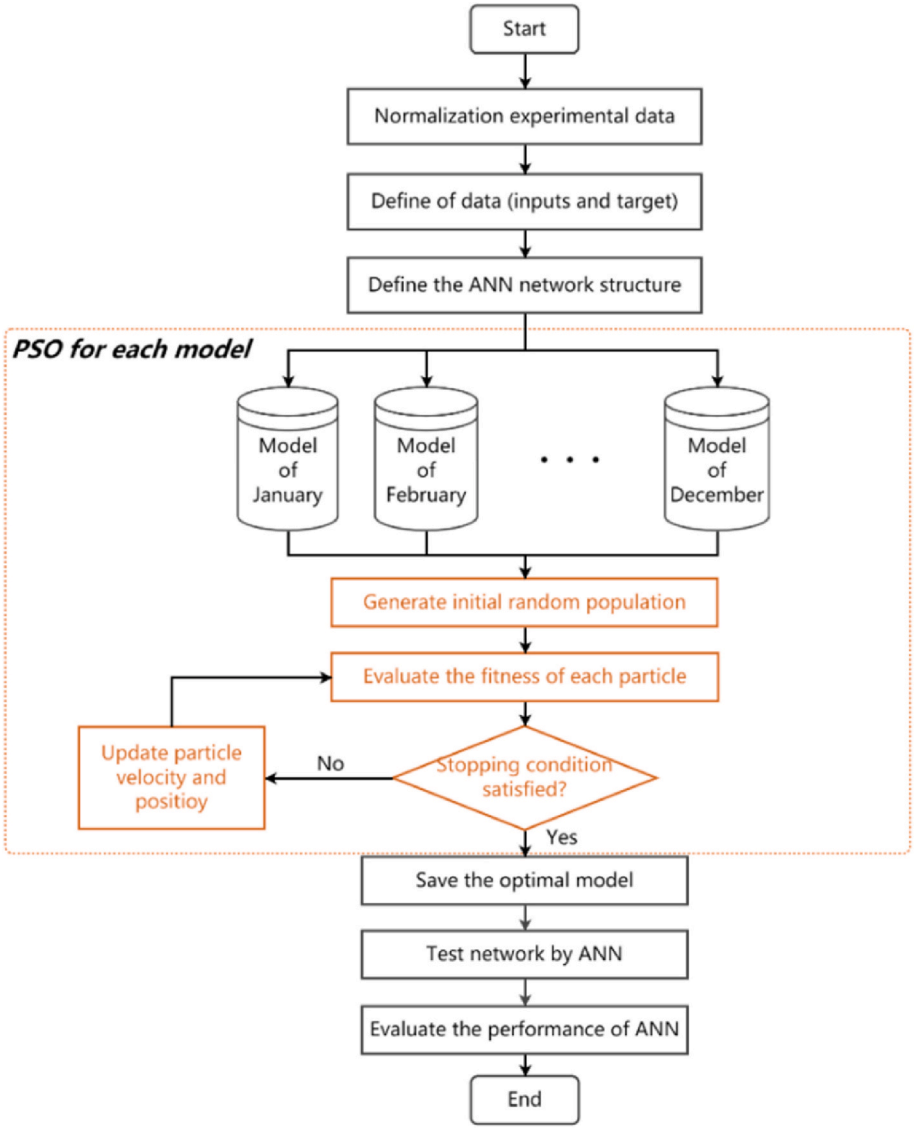


Fig. 5. Flowchart of the methodologies used in this research.

effort, but also avoids gradient disappearance. In the ANN model training process, 12 sets of data are input to the model separately, PSO is applied to optimize the network structure and parameters separately, and 12 optimal estimation models are obtained. Experiments show that the hybrid model can effectively take into account nonlinearity and seasonality to provide more accurate estimation results. The flow of the hybrid algorithm is shown in Fig. 5.

4. Results and discussions

4.1. Availability of satellite data

The accuracy of a machine-learning model is sensitive to the original input data (Jiang et al., 2021). Before utilizing the PSO-ANN model to estimate the BLT in the Indian Ocean, the satellite-derived SSS and SST data are briefly validated by comparing them with the gridded Argo data. In Fig. 6a, the monthly average SMOS-derived SSS shows good agreement with the Argo SSS data on a seasonal scale. It can be observed that there exists a certain level of discrepancy between remote sensing data and Argo data, which may be due to different depths of measurement and errors caused by gridded Argo. However, the differences are not significant and remain within 0.25 psμ for all months. As shown in Fig. 6b, the seasonal variation of the NOAA-derived SST averaged over the Indian Ocean has also good agreement with the Argo-derived SST. For example, both of them show that the maximum SST value (>28.40 °C) occurs in April, and the minimum SST value (<26.10 °C) occurs in August. The differences between the Argo SST and NOAA SST varies from -0.20 °C to 0.30 °C. In addition, the GPCP-derived P data and AVISO-derived SSH data have been widely applied in related studies (Behrangi and Song, 2020; Jeon et al., 2018; Morimoto, 2009; Yin and Gruber, 2010). These results demonstrate the availability of the four satellite data used in this study.

4.2. Development of the hybrid network model

In this study, a PSO-ANN model for estimating the BLT is developed by combining PSO algorithm and ANN model using multi-source remote sensing data in the Indian Ocean as a case study. A dataset of sea surface parameters (SST, SSH, SSS, P) covering 96 months and 2724 data points per month, as well as LON and LAT for each point, is collected from satellite and gridded Argo data. A PSO algorithm is utilized to determine the optimal number of layers, the number of neurons per layer, and the dropout rate for the ANN. A population fitness calculation method based on the RMSE is designed to optimize the parameters of each network model while keeping the training input variables constant. At each iteration, update the position of the particles and record the values of the best particles to obtain the best results.

Ocean datasets from 96 months of open data sources are used for PSO-ANN model training and testing in the Python environment. The monthly average data from January 2012 through December 2018 are taken as the training and validation sets to optimize the model architecture, with 80% of the data randomly selected for training and the remaining 20% for validation. Validation data are used to avoid overfitting by assessing the performance of the trained model after each epoch. The loss function is the minimization of the RMSE between our estimation and the real value. The data from the 12 months of 2019 are used as the testing set to evaluate the model's performance. Previous studies have shown that the BLT in the Indian Ocean is characterized by high fluctuations in seasonal variability (Felton et al., 2015). Therefore, two groups of comparison experiments are designed to verify the impact of seasonal fluctuations of BLT on the model performance, named Case 1 and Case 2, respectively. The training sets of Case 1 comprises all monthly data from 2012 to 2018. All training data are fed into a PSO-ANN estimation model, which yielded an optimized model. During the testing phase, the testing data for 12 months in 2019 are input into the model separately to obtain the estimated results per month. In research of Case 2, 84-month data from January 2012 to December 2018 are divided into 12 groups according to the same month. Each group of data is fed into a PSO-ANN estimation model for training respectively, resulting in 12 optimized models. The month data for 2019 is input into the model for the corresponding month to obtain the testing results for this month.

Based on the experimental design mentioned above, the range of parameters for optimizing the ANN model is set as follows: 2-4 hidden layers, 32-128 neurons per layer, and a dropout rate of 0.1-0.4. The PSO algorithm is used for 30 optimization iterations to obtain the optimal parameters of the model. The results indicate that the optimized PSO-ANN model had three hidden layers, with each layer containing 128 neurons and a dropout probability of 0.25 for Case 1. The model's RMSE on the testing set is shown in Fig. 7. For Case 2, 12 models are trained on each monthly dataset and the network parameters of each model are optimized using PSO algorithm. Table 2 shows the optimized parameters for the 12 ANN models. The RMSE of each model on the testing set is shown in Fig. 7.

As can be seen from Fig. 7, the RMSE values of the PSO-ANN models for Case 1 and 2 have similar seasonal variation tendency. However, the model in Case 1 cannot accurately capture the seasonal signals of BLT compared to the model in Case 2, as demonstrated by the larger RMSE. In fact, the monthly RMSE values on the testing set for Case 2 are all smaller than that of Case 1 (Fig. 7). For example, the minimum and maximum RMSE values of Case 2 are 3.08 m in April and 6.26 m in August, which are 0.14 m and 2.79 m smaller than that of Case 1, respectively. It is worth noting that both models have the maximum RMSEs in August, but the August RMSE of Case 2 is significantly reduced compared to that of Case 1. This may be because the model in Case 1 cannot reflect the special relationship in August due to its greater fluctuation in BLT compared to other months. These results show that the PSO-ANN model for Case 2 can better fit the nonlinear relationship between the BLT in the Indian Ocean and the input parameters.

As an example of the training effect of the PSO-ANN model in Case 2, Fig. 8a shows the variation in the loss values of training and validation data in January 2019. The training of the model is terminated when the loss value no longer decreases. As shown in Fig. 8a, the best performance of the PSO-ANN model for training and validation is done in epoch 7

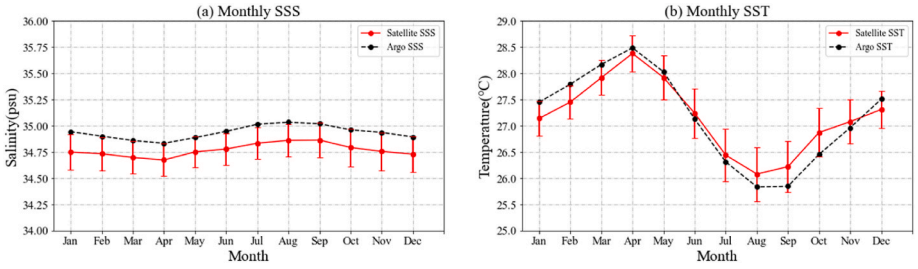


Fig. 6. Comparison of the gridded Argo (dashed black line) and satellite (solid red line) for (a) the monthly mean SSS and (b) SST in the Indian Ocean from January 2012 to December 2019. Error bars have been included for satellite data based upon the standard deviation.

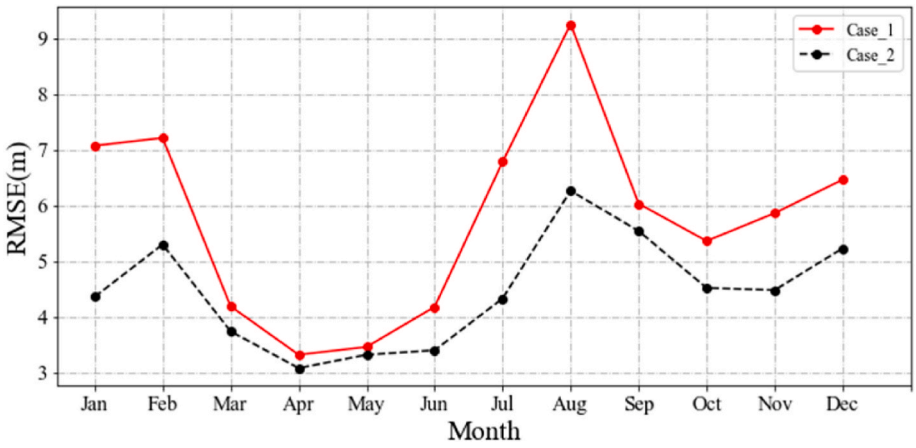


Fig. 7. Monthly RMSE for case 1 and case 2.

Table 2
The optimized network parameters values for Case 2.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Number of neurons per layer	63	65	60	76	60	99	88	100	60	75	100	71
Number of hidden layers	3	4	3	2	2	2	4	2	2	4	4	3
Dropout rate	0.3	0.21	0.23	0.3	0.3	0.3	0.24	0.27	0.22	0.26	0.3	0.3

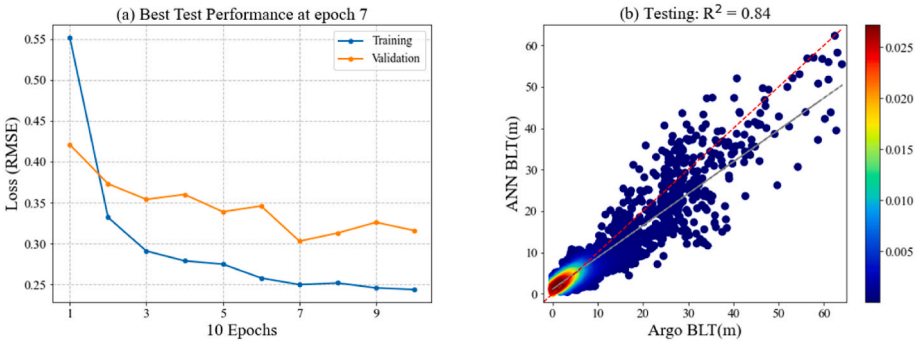


Fig. 8. (a) The performance of the ANN model when using the multisource observation datasets in January 2019. (b) Density scatter plots of the BLT from PSO-ANN estimation and gridded Argo data in January 2019.

with loss value of 0.31. The density scatter plots between the model estimated BLT and the Argo-derived BLT in January are depicted in Fig. 8b. The plots illustrate that although the estimated BLT values are somewhat undervalued, most data points are distributed along the equal value lines, indicating a good correlation between the estimated and observed BLT in the Indian Ocean. Moreover, the RMSE and R^2 values between the estimated and observed BLT are calculated, with values of 4.36 m and 0.84 in January, respectively. These results indicate that the PSO-ANN model in January has good performance in estimating BLT in the Indian Ocean.

4.3. Results analysis of the estimation model in Case 2

In this section, the performance of the PSO-ANN model for Case 2 is analyzed from different perspectives, as shown in Fig. 9. Fig. 9a and b displays the spatial distribution of the annual average BLTs based on the gridded Argo and the PSO-ANN model in 2019, respectively. And Fig. 9c shows the spatial differences between them (the difference between the Argo-derived BLT and PSO-ANN estimated BLT). Fig. 9d presents the correlation density scatterplots between the estimated and derived average BLTs in 2019. As can be seen from Fig. 9a–c, their spatial patterns agree well and the BLT discrepancies are mainly between −2 and 2 m in most regions. The annual average RMSE and R^2 values for the model are 1.83 m and 0.85, respectively. These results show that the PSO-ANN model can accurately estimate the BLT compared to the Argo data. For example, due to the influence of the monsoon in the sea area south of the equator, the BLT in this area from 15 °S to 30 °S is deeper and more zonal than that in other ocean regions, and the model captures this trend very well. The density scatter distribution also shows that most data points are basically evenly distributed along the isopleth (Fig. 9d), indicating a strong correlation between the estimated and observed BLTs. These results indicate that the PSO-ANN model in Case 2 has good estimation accuracy and can be used to estimate the BLT in the Indian Ocean with complex physical ocean motion. However, there is still room for improvement in the estimation accuracy of the model in certain regions. For example, the estimated BLT in the BoB is generally small, which may be caused by the complex change in salinity in the region. The salinity of the region is affected by the heavy precipitation of the southwest monsoon and a large amount of runoff input from the Brahmaputra River and the Ganges River at the top of the Bay (Hormann et al., 2019). The estimation effect in the SEAS also is poor compared to that of most other regions (Fig. 9c) due to low salinity waters transported to this region by the Kelvin wave along the coast of India and the currents from the BoB (Rao et al., 2010). These limitations indicate that

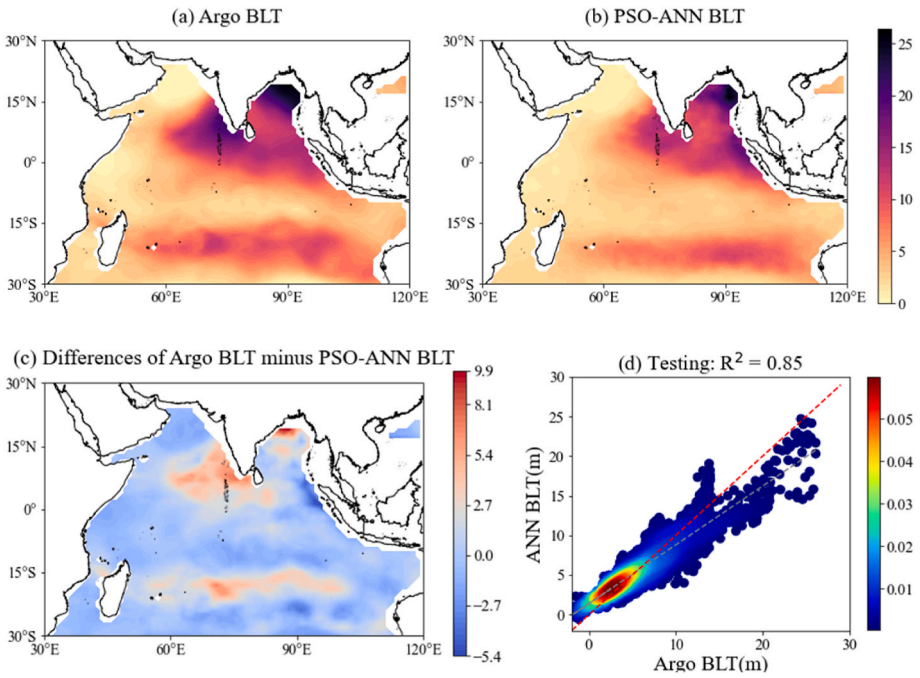


Fig. 9. Annual average BLT in the Indian Ocean from (a) gridded Argo data and (b) PSO-ANN model for Case 2 estimation in 2019 (unit: m). (c) Differences of the Argo-derived BLT minus the PSO-ANN estimated BLT. (d) Density scatter plots of the annual average BLT from gridded Argo data and PSO-ANN estimations in 2019.

the PSO-ANN model cannot completely capture the nonlinear signals in the regions with complex dynamic process, which may require the further improvement of ANN model in terms of its structure.

4.4. *Three typical ocean regions*

In order to further analyze the performance of PSO-ANN model for some regions with complex dynamic processes, three typical ocean regions with drastic changes in the BLT are selected to evaluate the performance of the model in Case 2. Fig. 10 shows the spatial distribution of the annual average RMSE and R^2 values for the model in 2019 in the SEAS, the BoB, and the EEIO. The average RMSE and R^2 values for each region are 5.02 m and 0.67 for the SEAS, 5.14 m and 0.65 for the BoB, and 6.20 m and 0.61 for the EEIO, respectively. Moreover, the values in many parts of the other two regions are less than 5 m for RMSE and greater than 0.80 for R^2 except for the values in the EEIO. The EEIO is situated off Sumatra-Java, which is influenced by more complex oceanographic processes due to the occurrence of the strongest positive IOD in 2019 (Du et al., 2020). This may make it is difficult for the model to accurately capture the variation characteristics of the BLT in the

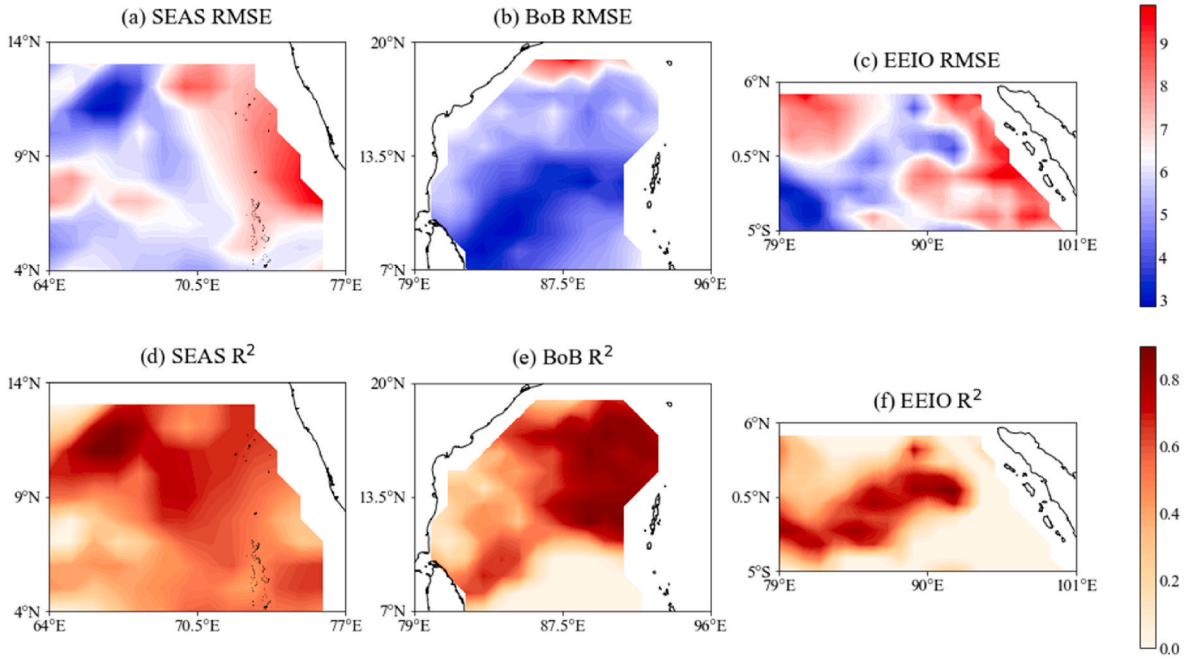


Fig. 10. Annual average RMSE (unit: m) (above) and R^2 (below) for the SEAS, the BoB, and the EEIO in 2019. The computation is based on the PSO-ANN model in Case 2 estimated and Argo-observed BLT.

region. On the whole, although the RMSE is slightly larger and the fitting effect is not so good, these findings also indicate that the PSO-ANN model can obtain the main information features of the BLT in these regions, and there is still room for improvement in the accuracy of the model.

4.5. Accuracy comparison with other models

In this study, the performance of the PSO-ANN model in Case 2 is compared with the data-driven models such as the traditional ANN model and MRM (Felton et al., 2015), as well as the CESM2-WACCM dynamic model from CMIP6. Fig. 11a, c, and e display the spatial distributions of the annual average BLTs estimated by the three models in 2019. The annual average RMSE and R^2 values are 2.04 m and 0.82 for the traditional ANN model, 3.38 m and 0.60 for the MRM, and 9.86 m and 0.26 for the CESM2-WACCM model, respectively. It is easy to see that the CESM2-WACCM dynamic model has the worst performance among the four models, as demonstrated by the maximum RMSE and minimum R^2 values. This may be due to the significant difference in temperature and salinity profiles data between the gridded Argo and CESM2-WACCM model. Meanwhile, the PSO-ANN model in Case 2 has the best estimation performance. For example, the PSO-ANN model can reduce the RMSE and R^2 values by 12% and 4%, 84% and 40% compared to the traditional ANN model and MRM (Figs. 9b and 11a and c), respectively. In addition, the differences between the Argo-derived BLT and three models estimated BLTs are also depicted in Fig. 11b, d and f. The red region indicates that the difference is more significant than zero, which means that the models underestimate the BLT values. The blue region means the models overestimate the BLT values. Figs. 9c and 11b, d and f show that the BLT estimated by these model are somewhat overestimated in BoB, SEAS, and the region from 15°S to 30°S, but the PSO-ANN model still has the smallest differences distribution. Moreover, the MRM performs poorly in general compared to the two data-driven models in estimating BLT through sea surface parameters. For example, the MRM and ANN model estimate the BLT values deeper than the PSO-ANN model in the EEIO and fails to effectively estimate the BLT in this region, as shown in Figs. 9c and 11b, d. The MRM also cannot estimate the spatial distribution of BLT in the SEAS (Fig. 11c). This may be because the MRM is only good at capturing smooth and linear dynamic processes. However, there may be complex nonlinear relationships between the BLT and various sea surface parameters due to the variable ocean-atmosphere interactions and regional ocean circulations in the Indian Ocean. Overall, these comparison results demonstrate the PSO-ANN model's capabilities in accurately estimating the BLT in the Indian Ocean with complex oceanographic processes.

The estimation performance of these data-driven models with seasonal variations is further compared. The RMSE and R^2 changes of the four models in different months are shown in Fig. 12. It can be seen that these models can all reveal the seasonal variability of BLT, which is consistent with previous research results (Felton et al., 2015). The RMSE is higher for all models during the summer months (July, August, and September), but the PSO-ANN model also outperforms other models. By comparing the results for each month, it can be found that there are different differences in the estimation effects for 12 months. For example, the minimum R^2 value is 0.29 in April, indicating that the PSO-ANN model has poor fitting performance, which may be due to insufficient existing surface parameters, making it difficult to estimate

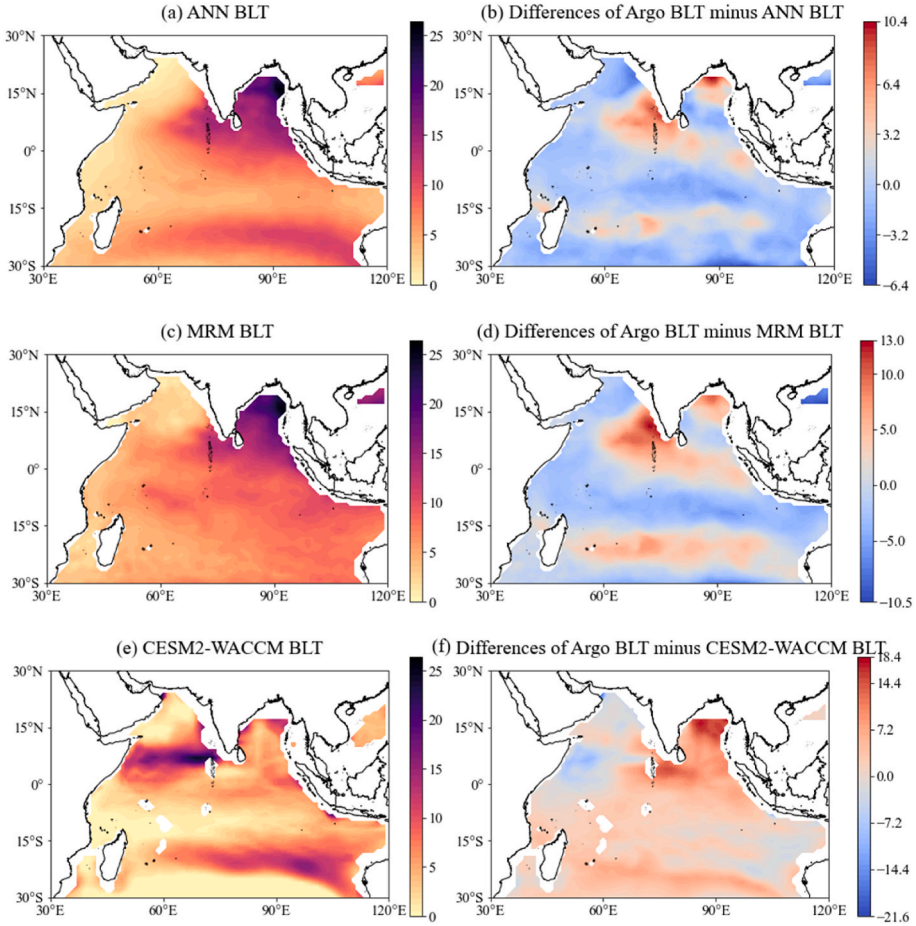


Fig. 11. Average annual BLT for the Indian Ocean in 2019 (unit: m). (a) ANN; (b) Differences of Argo-derived BLT minus ANN estimated BLT; (c) MRM; (d) Differences of Argo-derived BLT minus MRM estimated BLT; (e) CESM2-WACCM; (f) Differences of Argo-derived BLT minus CESM2-WACCM BLT.

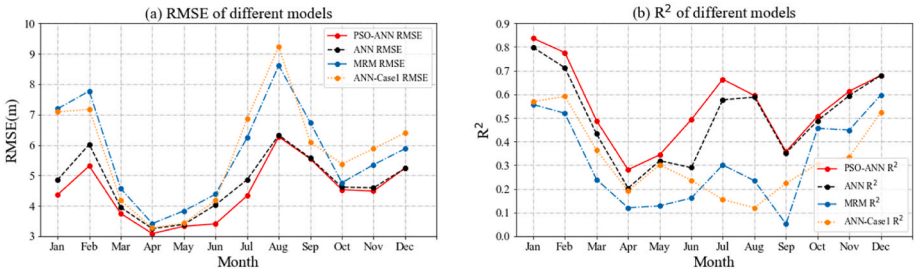


Fig. 12. The estimation accuracy of BLT in terms of RMSE (unit: m) and R^2 by four model in 2019; PSO-ANN (red line), traditional ANN (black line), MRM (blue line), and ANN-Case 1 (orange line).

the BLT in April. Additionally, the effectiveness of PSO algorithm in optimizing network parameters can also be seen in Fig. 12. Although the monthly RMSE and R^2 of the PSO-ANN model are better than that of the traditional ANN model, there are no significant differences in certain months. This indicates that whether the network parameters are optimized or not in these months, the ANN model can effectively capture the main features of the monthly average BLT. On the whole, the PSO-ANN model has higher accuracy, better fitting effect and can effectively capture the complex seasonal variation characteristics of the BLT in the Indian Ocean, as evidenced by the minimum RMSE and maximum R^2 values (Fig. 12). However, there is still room for improvement in the accuracy of the model.

4.6. Correlation analysis between the BLT and sea surface parameters

The previous analysis shows that the estimation effect of the proposed model is closely related to the input sea surface parameters. However, the existing models seldom further analyze the correlation between the BLT and these parameters. In order to study the relationship between the estimated BLT and each input variable of the model, the Pearson correlation coefficient is applied to quantitatively analyze the correlation between these variables. Fig. 13 shows the monthly average of the correlation coefficients in different months in 2019, respectively.

As seen in Fig. 13, there is a positive correlation between BLT and SST, SSH, and P in most months, while the correlation coefficient between BLT and SSS is negative throughout the year. The correlation coefficient between the BLT and SSS reaches its minimum value of -0.62 in December. This negative correlation shows that high salinity has a large influence on the BLT, which is consistent with the objective facts that an increase in salinity leads to a shallower BLT (Li et al., 2017). The correlation between BLT and SSH is positive throughout the year and is strongest in June with a correlation coefficient of 0.28. The correlation between BLT and P has a strong positive in most months, with a maximum correlation coefficient of 0.71 in November. This confirms that upwelling and freshwater fluxes (evaporation, precipitation and river runoff) all play a role in regulating subsurface variability in the Indian Ocean (Saha et al., 2021). The correlation analysis between the BLT and the sea surface parameters indicates that different sea surface parameters have different effects on the BLT in different months and there exist significant correlations between the BLT and SSS, P. In addition, since the Pearson correlation coefficient can only represent the linear correlation between different variables, other correlation coefficients can be selected to evaluate the nonlinear correlation between the BLT and sea surface parameters in future studies.

5. Conclusions

This study aims to provide an estimation method capable of determining and analyzing the BLT in some typical ocean regions with complex dynamic processes. The research goal is achieved by developing a hybrid PSO-ANN model and using multisource remote sensing and gridded Argo data in the Indian Ocean as a case study. The multisource datasets spanning 96 months from 2012 to 2019, with 2724 data points per month, are collected from satellite observations and Argo data in the Indian Ocean. The SST, SSS, SSH, P, LON, LAT and Argo-derived BLT are considered the model's input and output variables. The training data from January 2012 to December 2018 are input into the PSO-ANN model, with 80% of the data randomly selected for training and the remaining 20% for validation. Two groups of comparison experiments denoted as Case 1 and Case 2 are designed to verify the performance of the PSO-ANN model in estimating the BLT in the Indian Ocean. The comparative results demonstrate that the model in Case 2 can better capture the complex features of the monthly BLT in the ocean region. The annual average RMSE and R^2 values for Case 2 are 1.83 m and 0.85

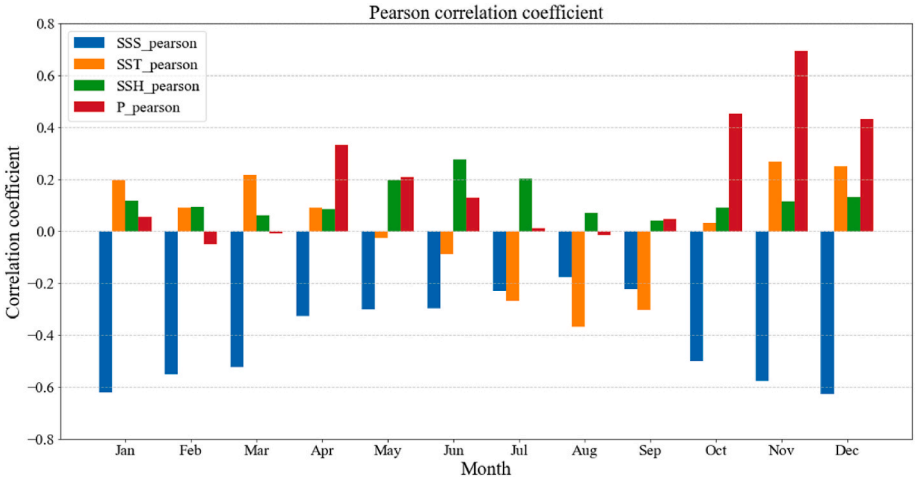


Fig. 13. Average monthly Pearson correlation coefficients between the estimated BLT and sea surface parameters (SST, SSS, SSH, and P) in 2019.

in 2019, respectively. This superior performance can be attributed to the good ability of the model for Case 2 to optimize network parameters and capture complex oceanographic processes.

The performance of the PSO-ANN model in Case 2 is also evaluated from multiple perspectives, including spatial distributions in different regions, comparison with existing models and seasonal variations. The comparison between the model estimated and Argo-derived BLTs indicates that the PSO-ANN model performs well and effectively reconstructs most of the observed BLT features using sea surface data. Furthermore, the spatial distributions of the annual average RMSE and R^2 values for the PSO-ANN model in Case 2 are analyzed in three typical ocean regions. The results demonstrate that the PSO-ANN model can capture the main information features of the BLT in these regions, but the estimation accuracy needs to be improved. The evaluation performance of the PSO-ANN model is further compared with the data-driven model such as the traditional ANN model and MRM, as well as the CESM2-WACCM dynamic model from CMIP6. The comparison results show that the dynamic model has the worst performance among the four models. The PSO-ANN model in Case 2 can reduce the RMSE and R^2 values by 12% and 4%, 84% and 40% compared to the traditional ANN model and MRM, respectively. Finally, the performance of these models with seasonal variations is further evaluated. The findings imply that the PSO-ANN model in Case 2 has a good seasonal applicability for the BLT estimation and outperforms other comparative models in all months, as demonstrated by the minimum RMSE and maximum R^2 values. Although, on the whole, the PSO-ANN model has superior performance compared to other models, the difference distributions shown in Figs. 9 and 11 also indicate that there is still significant room for improvement in the model's accuracy in the ocean regions with drastic changes in the BLT, such as the SEAS and the northern coast of the BoB. Therefore, future efforts may be directed towards decreasing the estimation error of the BLT in these regions. In addition, the quantitative correlation analysis results between the BLT and the sea surface parameters indicate that different sea surface parameters have different effects on the BLT in different months. The significant findings are that there exist significant correlations between the BLT and SSS, P.

In conclusion, the proposed PSO-ANN model exhibits superior performance in estimating the BLT in the Indian Ocean with complex oceanographic processes. These study results will assist in detecting and tracking the seasonal variation of the barrier layer in the Indian Ocean and provide insights for further scientific understanding of oceanographic processes in the ocean region. This study can be further expanded to estimate other internal parameters for typical ocean regions and can provide effective technical support for ocean researchers who are studying the variability of these parameters. Nevertheless, as a black box model, the PSO-ANN model has limitations in some aspects such as estimating extreme anomaly events and interpreting the physical mechanisms of the results. Moreover, there is still room for improvement in the accuracy of the model due to the impact of the gridded Argo data and the inherent shortcomings of the ANN model such as partial loss of input information during propagation and easily falling into local minimal value. In future studies, more advanced machine learning methods combining with oceanic dynamic mechanisms can be explored to further improve the estimation accuracy. Furthermore, suitable sea surface parameters should be selected to improve the evaluation accuracy and efficiency of the model in view of the important influence of sea surface parameters on the model. Additionally, since the Pearson correlation coefficient can only represent the linear correlation between different variables, other correlation coefficients should be selected to evaluate the nonlinear correlation between BLT and sea surface parameters in future studies.

Author contributions

Conceptualization, J.F.Q., S.L.Z., X.G.; Methodology, Y.Z.Z., S.L.Z., W.T.J; Formal analysis: Y.Z.Z., S.L.Z., W.T.J, W.M.Y., B.S.Y; Writing-Original draft, Y.Z.Z., S.L.Z., W.T.J; Writing-Reviewing & Editing, J.F.Q., S.L.Z., Y.Z.Z., W.T.J, X.G., W.M.Y. and B.S.Y.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This study was supported by the Youth Tutor Visiting Study and Training Project of Shandong Province.

References

Ali, M.M., Swain, D., Weller, R.A., 2004. Estimation of ocean subsurface thermal structure from surface parameters: a neural network approach. Geophys. Res. Lett. 31 (20), L20308 https://doi.org/10.1029/2004GL021192.

Asteris, P.G., Tsaris, A.K., Cavaleri, L., Repapis, C.C., Karypidis, D.F., 2016. Prediction of the fundamental period of infilled RC frame structures using artificial neural networks. Comput. Intell. Neurosci. 2016 (1), 20. https://doi.org/10.1155/2016/5104907.

Behrangi, A., Song, Y., 2020. A new estimate for oceanic precipitation amount and distribution using complementary precipitation observations from space and comparison with GPCP. Environ. Res. Lett. 15 (12), 124042–124049. https://doi.org/10.1088/1748-9326/abc6d1.

Bosc, C., D, T., Maes, C., 2009. Barrier layer variability in the western Pacific warm pool from 2000 to 2007. J. Geophys. Res. Oceans 114 (6), C6023. https://doi.org/10.1029/2008JC005187.

Boutin, J., Vergely, J.L., Marchand, S., D'Amico, F., Hasson, A., Kolodziejczyk, N., Reul, N., Reverdin, G., Vialard, J., 2018. New SMOS Sea Surface Salinity with reduced systematic errors and improved variability. Remote Sens. Environ. 214, 115–134. https://doi.org/10.1016/j.rse.2018.05.022.

Cavaleri, L., Chatzarakis, G.E., Trapani, F.D., Douvika, M.G., Asteris, P.G., 2017. Modeling of surface roughness in electro-discharge machining using artificial neural networks. Adv. Mater. Res. 6 (2), 169–184. https://doi.org/10.12989/amr.2017.6.2.000.

Chau, K.W., 2006. Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. J. Hydrol. 329 (3–4), 363–367. https://doi.org/10.1016/j.jhydrol.2006.02.025.

Chen, W., Panahi, M., Pourghasemi, H.R., 2017. Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. Catena 157, 310–324. https://doi.org/10.1016/j.catena.2017.05.034.

Chu, P.C., Fan, C., Liu, W.T., 2000. Determination of vertical thermal structure from sea surface temperature. J. Atmos. Ocean. Technol. 17 (7), 971–979. https://doi.org/10.1175/1520-0426(2000)017<0971:DOVTSF>2.0.CO;2.

Courtois, P., Hu, X., Pennelly, C., Spence, P., Myers, P.G., 2017. Mixed layer depth calculation in deep convection regions in ocean numerical models. Ocean Model. 120, 60–78. https://doi.org/10.1016/j.ocemod.2017.10.007.

Cronin, Meghan, F., 2002. Barrier layer formation during westerly wind bursts. J. Geophys. Res. Oceans 107 (C12), 1–21. https://doi.org/10.1029/2001JC001171.

de Boyer Montégut, C., Mignot, J., Lazar, A., Cravatte, S., 2007. Control of salinity on the mixed layer depth in the world ocean : 1. General description. J. Geophys. Res. 112 (C6), C6011. https://doi.org/10.1029/2006jc003953.

Dimitraki, L., Christaras, B., Marinos, V., Vlahavas, I., Arampelos, N., 2018. Predicting the average size of blasted rocks in aggregate quarries using artificial neural networks. Bull. Eng. Geol. Environ. 78 (4), 2717–2729. https://doi.org/10.1007/s10064-018-1270-1.

Dong, L., Qi, J., Yin, B., Zhi, H., Li, D., Yang, S., Wang, W., Cai, H., Xie, B., 2022. Reconstruction of subsurface salinity structure in the south China sea using satellite observations: a LightGBM-based deep forest method. Remote Sens. 14 (14), 3494. https://doi.org/10.3390/rs14143494.

Dong, W.X., Jun, H.G., Wei, L.I., Quan, Q.Y., 2011. Reconstruction of ocean temperature profile using satellite observations. J. Trop. Oceanogr. 30 (6), 10–17. https://doi.org/10.3969/j.issn.1009-5470.2011.06.002.

Drushka, K., Sprintall, J., Gille, S.T., 2014. Subseasonal variations in salinity and barrier-layer thickness in the eastern equatorial Indian Ocean. J. Geophys. Res. Oceans 119 (2), 805–823. https://doi.org/10.1002/2013jc009422.

Du, Y., Zhang, Y., Zhang, L.Y., Tozuka, T., Ng, B., Cai, W., 2020. Thermocline warming induced extreme Indian Ocean Dipole in 2019. Geophys. Res. Lett. 47 (18), e2020GL-e90079G https://doi.org/10.1029/2020GL090079.

Elbeltagi, E., Hegazy, T., Grierson, D., 2005. Comparison among five evolutionary-based optimization algorithms. Adv. Eng. Inf. 19 (1), 43–53. https://doi.org/10.1016/j.aei.2005.01.004.

Felton, C.S., Subrahmanyam, B., Murty, V.S.N., Shriver, J.F., 2015. Estimation of the barrier layer thickness in the Indian Ocean using Aquarius Salinity. J. Geophys. Res. Oceans 119 (7), 4200–4213. https://doi.org/10.1002/2013JC009759.

Ghorbani, M.A., Shamshirband, S., Haghi, D.Z., Azani, A., Bonakdari, H., Ebtehaj, I., 2017. Application of firefly algorithm-based support vector machines for prediction of field capacity and permanent wilting point. Soil Tillage Res. 172, 32–38. https://doi.org/10.1016/j.still.2017.04.009.

Godfrey, J.S., Lindstrom, E.J., 1989. The heat budget of the equatorial western Pacific surface mixed layer. J. Geophys. Res. Oceans 94 (C6), 8007–8017. https://doi.org/10.1029/JC094iC06p08007.

Grisales-Norea, L.F., Montoya, O.D., Ramos-Paja, C.A., 2020. An energy management system for optimal operation of BSS in DC distributed generation environments based on a parallel PSO algorithm. J. Energy Storage 29, 101488. https://doi.org/10.1016/j.est.2020.101488.

Gu, C., Qi, J., Zhao, Y., Yin, W., 2022. Estimation of the mixed layer depth in the Indian ocean from surface parameters: a clustering-neural network method. Sensors 22 (15), 5600. https://doi.org/10.3390/s22155600.

Guinehut, S., Dhomps, A.L., Larnicol, G., Le Traon, P.Y., 2012. High resolution 3-D temperature and salinity fields derived from in situ and satellite observations. Ocean Sci. 9 (2), 1313–1347. https://doi.org/10.5194/os-8-845-2012.

Hormann, V., Centurioni, L.R., Gordon, A.L., 2019. Freshwater export pathways from the Bay of Bengal. Deep-Sea Res. Pt. II 168, 104645. https://doi.org/10.1016/j.dsr2.2019.104645.

Jeon, C., Park, J.H., Kim, D.G., Kim, E., Jeon, D., 2018. Comparison of measurements from pressure-recording inverted echo sounders and satellite altimetry in the north equatorial current region of the western Pacific. Ocean Sci. J. 53 (2), 207–213. https://doi.org/10.1007/s12601-018-0012-4.

Jiang, F., Ma, J., Wang, B., Yuan, L., 2021. Ocean observation data prediction for Argo data quality control using deep bidirectional LSTM network. Hindawi. https://doi.org/10.1155/2021/5665386.

Kawamiya, M., Oschlies, A., 2001. Formation of a basin-scale surface chlorophyll pattern by Rossby waves. Geophys. Res. Lett. 28 (21), 4139–4142. https://doi.org/10.1029/2001GL013347.

Kennedy, J., 2011. Particle swarm optimization. Proc. of 1995 IEEE Int. Conf. Neural Networks 4 (8), 1942–1948. https://doi.org/10.1007/978-0-387-30164-8_630 (Perth, Australia), Nov. 27-Dec.

Koopialipoor, M., Fallah, A., Armaghani, D.J., Azizi, A., Mohamad, E.T., 2019. Three hybrid intelligent models in estimating flyrock distance resulting from blasting. Eng. Comput. 35 (1), 243–256. https://doi.org/10.1007/S00366-018-0596-4.

Li, Y., Han, W., Wang, W., Ravichandran, M., Shinoda, T., 2017. Bay of Bengal salinity stratification and Indian summer monsoon intraseasonal oscillation: 2. Impact on SST and convection. J. Geophys. Res. Oceans 122 (5), 4312–4328. https://doi.org/10.1002/2017JC012692.

Liang, Z., Rao, J., Guo, D., Lu, Q., 2022. Simulation and projection of the sudden stratospheric warming events in different scenarios by CESM2-WACCM. Clim. Dynam. 59 (11), 3741–3761. https://doi.org/10.1007/s00382-022-06293-2.

Liu, H., Grodsky, S.A., Carton, J.A., 2009. Observed subseasonal variability of oceanic barrier and compensated layers. J. Clim. 22 (22), 6104–6119. https://doi.org/10.1175/2009JCLI2974.1.

Lu, W., Su, H., Yang, X., Yan, X., 2019. Subsurface temperature estimation from remote sensing data using a clustering-neural network method. Remote Sens. Environ. 229, 213–222. https://doi.org/10.1016/j.rse.2019.04.009.

Lukas, R., Lindstrom, E., 1987. The mixed layer of the western equatorial Pacific Ocean. J. Geophys. Res. Oceans 96 (S01), 3343–3357. https://doi.org/10.1029/90JC01951.

Meng, L., Yan, C., Zhuang, W., Zhang, W., Geng, X., Yan, X., 2020. Reconstructing high-resolution ocean subsurface and interior temperature and salinity anomalies from satellite observations. IEEE Trans. Geosci. Remote 2022 (60), 4104114. https://doi.org/10.1109/TGRS.2021.3109979.

Mignot, J., de Boyer Montégut, C., Lazar, A., Cravatte, S., 2007. Control of salinity on the mixed layer depth in the world ocean: 2. Tropical areas. J. Geophys. Res. 112 (C10), C10010 https://doi.org/10.1029/2006jc003954.

Moayedi, H., Mehrabi, M., Mosallanezhad, M., Rashid, A.S.A., Ahmad, S.A., Pradhan, B., 2019. Modification of landslide susceptibility mapping using optimized PSO-ANN technique. Eng. Comput. 35 (3), 967–984. https://doi.org/10.1007/s00366-018-0644-0.

Momeni, E., Nazir, R., Jahed Armaghani, D., Maizir, H., 2014. Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. Measurement 57, 122–131. https://doi.org/10.1016/j.measurement.2014.08.007.

Monterey, G.I., Levitus, S., 1997. Seasonal variability of mixed layer depth for the world ocean.

More, T., Gnanaseelan, C., Kakatkar, R.A., S, D.J., 2020. Indian ocean warming trends and forcing mechanism with emphasis on Northeastern tropical Indian ocean. J. Coast Res. 89, 15–19. https://doi.org/10.2112/SI89-003.1.

Morimoto, A., 2009. Evaluation of tidal error in altimetry data in the Asian Marginal Seas. J. Oceanogr. 65 (4), 477–485. https://doi.org/10.1007/s10872-009-0041-9.

Murty, V.S.N., Sarma, Y.V.B., Rao, D.P., 1996. Variability of the oceanic boundary layer characteristics in the northern Bay of Bengal during MONTBLEX-90. Proc. Indian Acad. Sci. Earth Planet. Sci. Lett. 105 (1), 41–61. https://doi.org/10.1007/BF02880758.

Nguyen, H., Drebenstedt, C., Bui, X.N., Bui, D.T., 2020a. Prediction of blast-induced ground vibration in an Open-Pit mine by a novel hybrid model based on clustering and artificial neural network. Springer 29 (2), 691–709. https://doi.org/10.1007/s11053-019-09470-z.

Nguyen, H., Moayedi, H., Foong, L.K., Al Najjar, H.A.H., Jusoh, W.A.W., Rashid, A.S.A., Jamali, J., 2020b. Optimizing ANN models with PSO for predicting short building seismic response. Eng. Comput. 36 (3), 823–837. https://doi.org/10.1007/s00366-019-00733-0.

Nyadjro, E.S.S.B., 2013. Variability of salt flux in the Indian Ocean during 1960-2008. Remote Sens. Environ. 134, 175–193. https://doi.org/10.1016/j.rse.2013.03.005.

Olayode, I.O., Tartibu, L.K., Okwu, M.O., Severino, A., 2021. Comparative traffic flow prediction of a heuristic ANN model and a hybrid ANN-PSO model in the traffic flow modelling of vehicles at a four-way signalized road intersection. Sustain. Times 13 (19), 10704. https://doi.org/10.3390/su131910704.

Pauthenet, E., Bachelot, L., Balem, K., Maze, G., Tréguier, A., Roquet, F., Fablet, R., Tandeo, P., 2022. Four-dimensional temperature, salinity and mixed-layer depth in the Gulf Stream, reconstructed from remote-sensing and in situ observations with neural networks. Ocean Sci. 18 (4), 1221–1244. https://doi.org/10.5194/os-18-1221-2022.

Qi, J., Xie, B., Li, D., Chi, J., Yin, B., Sun, G., 2023. Estimating thermohaline structures in the tropical Indian Ocean from surface parameters using an improved CNN model. Front. Mar. Sci. 10, 1181182 https://doi.org/10.3389/fmars.2023.1181182.

Qu, T., Song, Y.T., Maes, C., 2014. Sea surface salinity and barrier layer variability in the equatorial Pacific as seen from Aquarius and Argo. J. Geophys. Res. Oceans 119 (1), 15–29. https://doi.org/10.1002/2013JC009375.

Rao, R.R., Kumar, M.S.G., Ravichandran, M., Rao, A.R., Gopalakrishna, V.V., Thadathil, P., 2010. Interannual variability of Kelvin wave propagation in the wave guides of the equatorial Indian Ocean, the coastal Bay of Bengal and the southeastern Arabian Sea during 1993–2006. Deep-Sea Res. Pt. I 57 (1), 13.

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., Wang, W.Q., 2001. An improved in situ and satellite SST analysis for climate. J. Clim. 15 (13), 1609–1625. https://doi.org/10.1175/1520-0442(2002)015.

Roemmich, D., Gilson, J., 2009. The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the argo program. Prog. Oceanogr. 82 (2), 81–100. https://doi.org/10.1016/j.pocean.2009.03.004.

Roemmich, D., Morris, M., Young, W.R., Jean-Rene, D., 1994. Fresh equatorial jets. J. Phys. Oceanogr. 24 (3), 540–558. https://doi.org/10.1175/1520-0485(1994)024<0540:FEJ>2.0.CO;2.

Saha, A., Serra, N., Stammer, D., 2021. Growth and decay of northwestern tropical atlantic barrier layers. J. Geophys. Res. Oceans 126 (5), e2020J-e16956J. https://doi.org/10.1029/2020JC016956.

Sahraei, M.A., Codur, M.K., 2022. Prediction of transportation energy demand by novel hybrid meta-heuristic ANN. Energy 249, 123735. https://doi.org/10.1016/j.energy.2022.123735.

Schiller, A., Oke, P.R., 2015. Dynamics of ocean surface mixed layer variability in the Indian Ocean. J. Geophys. Res. Oceans 120 (6), 4162–4186. https://doi.org/10.1002/2014JC010538.

Schott, F.A., Xie, S.P., Mccreary, J.P., 2009. Indian Ocean circulation and climate variability. Rev. Geophys. 47 (1), G1002–G1046. https://doi.org/10.1029/2007RG000245.

Seifi, A., Soroush, F., 2020. Pan evaporation estimation and derivation of explicit optimized equations by novel hybrid meta-heuristic ANN based methods in different climates of Iran. Comput. Electron. Agric. 173 (10), 105418 https://doi.org/10.1016/j.compag.2020.105418.

Sprintall, J., Tomczak, M., 1992. Evidence of the barrier layer in the surface layer of the tropics. J. Geophys. Res. 97 (C5), 7305–7316. https://doi.org/10.1029/92JC00407.

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., Yan, X.H., 2021. Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. Int. J. Appl. Earth Obs. 102, 102440 https://doi.org/10.1016/j.jag.2021.102440.

Su, H., Wu, X., Yan, X.H., Kidwell, A., 2015. 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. 160, 63–71. https://doi.org/10.1016/j.rse.2015.01.001.

Su, H., Yang, X., Lu, W., Yan, X.H., 2019. Estimating subsurface thermohaline structure of the global ocean using surface remote sensing observations. Rem. Sens. 11 (13), 1598. https://doi.org/10.3390/rs11131598.

Tomczak, M., Godfrey, J.S., 1994. Regional Oceanography: an Introduction. Pergamon. https://doi.org/10.1016/B978-0-08-041021-0.50010-2.

Vaheddoost, B., Guan, Y., Mohammadi, B., 2020. Application of hybrid ANN-whale optimization model in evaluation of the field capacity and the permanent wilting point of the soils. Environ. Sci. Pollut. Res. 27 (3–4), 13131–13141. https://doi.org/10.1007/s11356-020-07868-4.

Valsala, V., Singh, S., Balasubramanian, S., 2018. A modeling study of interannual variability of Bay of Bengal mixing and barrier layer formation. J. Geophys. Res. Oceans 123 (6), 3962–3981. https://doi.org/10.1029/2017JC013637.

Vialard, J., Delecluse, P., 1998a. An OGCM study for the TOGA decade. Part II: barrier-layer formation and variability. J. Phys. Oceanogr. 28 (6), 1089–1106. https://doi.org/10.1175/1520-0485(1998)0282.0.CO;2.

Vialard, J., Delecluse, P., 1998b. An OGCM study for the TOGA decade. Part l: role of salinity in the physics of the western Pacific fresh pool. J. Phys. Oceanogr. 28 (6), 1071–1088. https://doi.org/10.1175/1520-0485(1998)0282.0.CO;2.

Wang, D.L., Sun, Q.Y., Li, Y.Y., Liu, X.R., 2019. Optimal energy routing design in energy internet with multiple energy routing centers using artificial neural network-based reinforcement learning method. Appl. Sci. 9 (3), 520. https://doi.org/10.3390/app9030520.

Wang, H., Song, T., Zhu, S., Yang, S., Feng, L., 2021. Subsurface temperature estimation from sea surface data using neural network models in the western Pacific Ocean. Mathematics 9 (8), 852. https://doi.org/10.3390/math9080852.

Wang, L., Xu, Z., Gong, X., Zhang, P., Hao, Z., You, J., Zhao, X., Guo, X., 2023. Estimation of nitrate concentration and its distribution in the northwestern Pacific Ocean by a deep neural network model. Deep-Sea Res. Part I Oceanogr. Res. Pap. 195, 104005 https://doi.org/10.1016/j.dsr.2023.104005.

Wang, X., Liu, H., 2016. Seasonal-to-interannual variability of the barrier layer in the western Pacific warm pool associated with ENSO. Clim. Dyn. 47 (1–2), 375–392. https://doi.org/10.1007/s00382-015-2842-4.

Yang, X., Zhang, Y., Yang, Y., Lv, W., 2019. Deterministic and probabilistic wind power forecasting based on bi-level convolutional neural network and particle swarm optimization. Appl. Sci. 9 (9), 1794. https://doi.org/10.3390/app9091794.

Yin, X., Gruber, A., 2010. Validation of the abrupt change in GPCP precipitation in the Congo River Basin. Int. J. Climatol. 30 https://doi.org/10.1002/joc.1875.

Yousaf, S., Mughees, A., Khan, M.G., Amin, A.A., Adnan, M., 2020. A comparative analysis of various controller techniques for optimal control of smart nano-grid using GA and PSO algorithms. IEEE Access 8, 205696–205711. https://doi.org/10.1109/ACCESS.2020.3038021.

Zhang, H., Jing, M., Zhang, H., Li, L., Zheng, Y., Tang, J., Tian, D., Zhu, Y., 2023. Deep learning approach for forecasting sea surface temperature response to tropical cyclones in the western north Pacific. Deep-Sea Res. Part I Oceanogr. Res. Pap. 197, 104042 https://doi.org/10.1016/j.dsr.2023.104042.

Zhang, K., Geng, X., Yan, X., 2020. Prediction of 3-D ocean temperature by multilayer convolutional LSTM. IEEE Geosci. Remote Sci. 17 (8), 1303–1307. https://doi.org/10.1109/LGRS.2019.2947170.

Zheng, F., Zhang, R., Zhu, J., 2014. Effects of interannual salinity variability on the barrier layer in the western-central equatorial Pacific: a diagnostic analysis from Argo. Adv. Atmos. Sci. 31 (3), 532–542. CNKI:SUN:DQJZ.0.2014-03-004.

Zheng, F., Zhang, R.H., 2012. Effects of interannual salinity variability and freshwater flux forcing on the development of the 2007/08 La Nia event diagnosed from Argo and satellite data. Dynam. Atmos. Oceans 57, 45–57. https://doi.org/10.1016/j.dynatmoce.2012.06.002.