



Article


Enhancing Few-Shot Prediction of Ocean Sound Speed Profiles through Hierarchical Long Short-Term Memory Transfer Learning

Jiajun Lu, Hao Zhang, Sijia Li, Pengfei Wu and Wei Huang



Article

Enhancing Few-Shot Prediction of Ocean Sound Speed Profiles through Hierarchical Long Short-Term Memory Transfer Learning

Jiajun Lu ¹, Hao Zhang ^{1,2}, Sijia Li ¹, Pengfei Wu ¹ and Wei Huang ^{1,*} 

¹ Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China; lujiacun@stu.ouc.edu.cn (J.L.); zhanghao@ouc.edu.cn (H.Z.); lisijia9430@stu.ouc.edu.cn (S.L.); wupengfei@stu.ouc.edu.cn (P.W.)

² Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8P 5C2, Canada

* Correspondence: hw@ouc.edu.cn

Abstract: The distribution of ocean sound speed profiles (SSPs) profoundly influences the design of underwater acoustic communication and positioning systems. Conventional methods for measuring sound speed by instruments entail high time costs, while sound speed inversion methods offer rapid estimation of SSPs. However, these methods heavily rely on sonar observational data and lack the capacity to swiftly estimate SSPs in arbitrary oceanic regions, particularly in scenarios with few-shot data. Precisely estimating non-cooperative maritime SSPs under such conditions poses a significant challenge. To explore temporal distribution patterns of sound speed and achieve precise SSP predictions with limited data, we propose a hierarchical long short-term memory transfer learning (H-LSTM-TL) framework. The core idea involves pre-training the base model on extensive public datasets, transferring the acquired knowledge to task models, and fine-tuning the task model on few-shot data to predict future SSPs. Through H-LSTM-TL, it accelerates model convergence, enhances sensitivity to few-shot input data, alleviates overfitting issues, and notably improves the accuracy of SSP predictions. Experimental results demonstrate that the H-LSTM-TL model exhibits strong generalization capabilities in few-shot data scenarios, effectively reducing overfitting problems and proving its applicability for rapid prediction of SSPs.

Keywords: hierarchical long short-term memory transfer learning (H-LSTM-TL); few-shot learning; ocean sound speed profile (SSP) prediction; overfitting effect; generalization capabilities



Citation: Lu, J.; Zhang, H.; Li, S.; Wu, P.; Huang, W. Enhancing Few-Shot Prediction of Ocean Sound Speed Profiles through Hierarchical Long Short-Term Memory Transfer Learning. *J. Mar. Sci. Eng.* **2024**, *12*, 1041. <https://doi.org/10.3390/jmse12071041>

Academic Editor: Sergei Chernyi

Received: 12 May 2024

Revised: 17 June 2024

Accepted: 20 June 2024

Published: 21 June 2024



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

1. Introduction

In the marine environment, the attenuation of acoustic signals is considerably smaller than that of radio and optical signals. This characteristic makes acoustic signals more suitable for long-distance information transmission, establishing them as the primary signal carrier for underwater sensing systems [1]. The underwater integrated positioning, navigation, timing, and communication (PNTC) system holds significant importance in various domains such as maritime disaster warning, underwater rescue operations, marine resource development, and national defense security. Acoustic signals, serving as the primary signal carrier for underwater PNTC systems, have garnered considerable research attention due to their propagation characteristics in the ocean environment. The acquisition of oceanic sound speed distributions plays a pivotal role in advancing the development of high-precision positioning, ranging, timing, and communication capabilities within PNTC systems because it has great influence on the signal propagation mode.

The sound speed distribution in the ocean is dynamic, influenced significantly by temperature, salinity, and pressure. These influences manifest in the change in spatial dimension with latitude and longitude and the change in temporal dimension with daily or seasonal periods [2]. According to [3], in shallow ocean regions, temperature predominantly

drives sound speed variation. However, as a certain sea depth is reached, pressure changes become the dominating factor. Due to the effects of sea winds, waves, and turbulent conditions, the temperature variation in these shallow waters is minimal, referred to as the surface isothermal layer [4]. In the deep-sea domain, the ocean temperature is notably lower and exhibits relative stability, forming the deep isothermal layer. Within this region, the impact of temperature on sound propagation velocity diminishes, becoming primarily influenced by pressure, displaying a linear increasing trend with pressure escalation. Between these two isothermal layers, a significant temperature decrease with depth is observed, resulting in a gradual reduction in sound speed, ultimately exhibiting a negative gradient distribution. This transitional region is identified as the thermocline.

The variation in sound speed in the vertical direction exceeds that in the horizontal direction, necessitating the common use of sound speed profiles (SSPs) to depict sound speed distribution [5]. Currently, the acquisition of ocean SSPs primarily involves measurement and inversion methods [6]. SSPs can be directly measured using a sound velocity profiler (SVP) [7] or indirectly determined by combining conductivity, temperature, and depth profilers (CTDs) [8,9], as well as expendable CTDs (XCTDs) [10], with empirical sound speed formulas. Sound speed profile (SSP) measurement methods offer precise measurements of sound speed distribution in specific regions, but they are time-consuming, and the measurement depth range of an XCTD is constrained by sensor pressure tolerance. Consequently, achieving a rapid and accurate construction of the sound speed field is crucial [11].

In recent years, a variety of methods utilizing acoustic field information to infer SSP have been proposed, aiming to establish the relationship between the marine environment (SSP distribution) and sound field measurements. Classical SSP inversion approaches primarily fall into categories such as matched field processing (MFP) [12,13], compressed sensing (CS) [14–16], and machine learning methods [17–19]. In comparison to traditional SSP measurement methods, SSP inversion methods are not only more time-efficient but also more intelligent [20], thus garnering extensive research attention. Both the SSP measurement and inversion methods focus on constructing spatially dimensional sound speed fields, with a high reliance on in situ observations.

In order to address the challenges posed by the excessive reliance on sonar observation data and the difficulty in rapidly estimating SSPs in non-cooperative maritime environments, we propose a hierarchical long short-term memory (H-LSTM) neural network for SSP prediction. To tackle the issue of inconsistent sound speed variations at different depths, we introduce the concept of depth-stratified processing. The method enables the rapid estimation of future time SSPs without the need for on-site data measurements. However, in scenarios where sound speed data are scarce, neural network models are prone to overfit the noise, leading to a decrease in predictive accuracy. Therefore, addressing the overfitting issue of the H-LSTM model in few-shot situations and enhancing its predictive capabilities warrant further investigation.

Currently, underwater reference SSP data are insufficient in quantity due to difficulties in collection. Despite initiatives like the Argo program promoting ocean data observation, SSP data remain deficient in global databases, lacking real-time measurement data. The scarcity of ocean observation data poses a significant challenge to the progress of marine science [21]. In situations where marine reference data are scarce, failing to conduct few-shot learning can lead to overfitting, thereby reducing the model's generalization ability. Therefore, enhancing the predictive capabilities of neural network models in few-shot situations is crucial.

Transfer learning (TL), as an emerging machine learning method, is well-suited for few-shot learning [22,23]. Transfer learning involves leveraging the experience and relevant parameters learned from a source task and transferring them to a new task model to ensure the model is not trained from scratch. Subsequently, the new task is used to fine-tune the task model, effectively mitigating the overfitting phenomenon. Transfer learning provides a model optimization approach for handling few-shot data, making it widely applicable

in various marine-related fields, such as wave characterization predictions [24] and ocean noise classification [25]. Given that historical ocean SSP data consist of vertically arranged sound speed values at different depths and the sound speed patterns at different depths exhibit inconsistent temporal variations, applying transfer learning directly to sound speed data using schemes from other domains can result in severe data feature loss. Therefore, for specific tasks related to predicting ocean SSP, it is essential to construct a more reasonable transfer learning model tailored to the characteristics of marine environments.

In order to address the overfitting issue in few-shot scenarios and to enhance the prediction efficiency and generalization ability of the model, we propose a transfer learning framework for sound speed prediction, and the internal model is adopted from our proposed H-LSTM, which combines the H-LSTM model and the TL framework to propose a novel method of transfer learning based on hierarchical long- and short-term memory neural networks (H-LSTM-TL). This method involves pre-training the base model with a substantial amount of SSP data from publicly available global datasets. Subsequently, the experiences and parameters learned by the converged base model are transferred to the task model. The task model is then fine-tuned using few-shot data relevant to the new task. Finally, the converged task model is utilized for SSP predictions in the new task. Through H-LSTM-TL, the convergence speed of the task model is accelerated, maintaining sensitivity to input data and reducing overfitting on limited new task samples, thus significantly improving SSP prediction accuracy.

The contributions of this paper can be summarized as follows:

- To obtain favorable prediction accuracy in few-shot data, we propose a transfer learning framework for SSP prediction. The internal model is not limited to deep neural network models, Gaussian process regression models, or Transformer models.
- To address the challenge of traditional SSP inversion methods relying heavily on sonar observation data and struggling with rapid estimation of non-cooperative maritime area SSPs, we propose a hierarchical long short-term memory (H-LSTM) neural network as the internal model within the TL framework.
- To validate the feasibility of H-LSTM-TL, we conducted ocean experiments to obtain few-shot SSP data and evaluated the accuracy of the model in predicting SSPs based on the measured data.

The rest of the paper is organized as follows: In Section 2, we briefly review related works about SSP inversion. In Section 3, we first propose a TL framework for the prediction of few-shot SSPs, followed by the presentation of the internal model H-LSTM for the TL framework, and finally, we propose the H-LSTM-TL model based on few-shot SSPs and give the implementation details. In Section 4, we give the experimental results and experimental analysis, which fully validate the feasibility and effectiveness of H-LSTM-TL for predicting SSPs. Finally, the conclusions are given in Section 5.

2. Related Work

MFP, CS, and machine learning are three classical SSP inversion methods [26]. As early as 1991, Tolstoy et al. proposed an MFP framework combining empirical orthogonal function (EOF) decomposition, offering an effective solution for SSP inversion [12]. Subsequently, Taroudakis et al. integrated MFP with modal phase inversion for SSP inversion [27]. Yu et al. introduced an MFP method incorporating genetic algorithms for SSP inversion [28]. MFP remained a predominant algorithm for SSP inversion for a long time. However, its computational complexity prompted the introduction of CS methods for SSP inversion, aiming to expedite the inversion process [14–16]. In contrast to MFP, the CS framework reduces computational complexity, thereby improving SSP calculation efficiency. However, CS methods, by simplifying the mapping relationship from the acoustic field to sound speed distribution, sacrifice some inversion accuracy. Recently, several classical SSP inversion methods have been proposed, expanding on the MFP foundation. For instance, a method combining particle swarm optimization (PSO) with EOF for SSP inversion was introduced [29]. Another approach involves a single-benchmark assimilation method for

ocean SSPs [30]. Additionally, a single empirical orthogonal function-regression (sEOF-r) method [31] and a ray-gradient-enhanced surrogate model (RGE-SM) method [32] were developed to further enhance SSP inversion techniques.

With the continuous development of the global ocean observation network and the maturation of theories and practices in machine learning and deep learning, an increasing number of scholars are applying machine learning and deep learning methods to address relevant issues in the field of oceanography [33,34]. For example, Jain et al., aiming to demonstrate the feasibility of estimating SSPs using satellite-derived sea surface parameters, successfully inverted SSPs using artificial neural network (ANN) methods [35]. Hashem et al., utilizing ANNs, investigated the influence of temperature, electric fields, and magnetic fields on sound speed variations in the ocean [36]. Huang et al. proposed an SSP inversion method based on a combination of artificial neural networks and ray theory [17]. They later introduced a framework for autonomous underwater vehicles that assisted ocean sound speed inversion, combining ray tracing theory and artificial intelligence models [37], for the rapid acquisition of three-dimensional sound speed distributions. Yu et al. proposed a sound speed inversion method based on radial basis function (RBF) neural networks [38]. Bianco and Hua separately proposed ocean SSP estimation methods based on dictionary learning [39,40]. Li et al. introduced a nonlinear SSP inversion method based on a self-organizing map (SOM), utilizing abnormal data for sea surface temperature and sea surface height obtained from satellite observations, combined with EOF coefficients from Argo floats for SSP inversion [41]. Huang et al. introduced an auto-encoding feature-mapping neural network (AEFMNN) structure, effectively enhancing the robustness of the neural network model in constructing the sound speed field against interference [37]. Ou et al. proposed an SSP inversion algorithm based on a comprehensive learning model using random forest (RF), followed by a method reconstructing SSP using the extreme gradient boosting (XGBoost) model [42,43].

When inverting SSPs using methods like EOF, MFP, CS, or machine learning models, a substantial reliance on a large amount of empirical SSP data for reference is necessary. However, the time and economic costs associated with measuring SSP using conventional methods are often high. As a result, SSP data are extremely scarce in many maritime regions, severely hindering the progress of marine science. In situations of data scarcity, machine learning models can frequently face overfitting issues, leading to a decline in the accuracy of SSP inversion. To address overfitting with few-shot data, various methods have been proposed. These include cross-validation methods [44], early stopping techniques [45], generative adversarial networks for expanding datasets [46], regularization approaches [47], multi-task learning (MTL) [48,49], and transfer learning (TL) [22,23]. Transfer learning is the transfer from sample to sample. Usually, the source domain has a large number of samples and the target domain has only few-shot data, transferring the laws learned by the model on a large number of samples to the target domain.

Transfer learning has garnered widespread attention in various marine-related fields [50]. For instance, Miao et al. established a transfer learning model for predicting sea surface temperature anomalies and sea surface height anomalies based on satellite remote sensing observational data [51]. Kumar et al. constructed a TL model for wave characterization prediction using data from Mexico, Korea, and the United Kingdom for three regions [24]. Lu et al. proposed a TL method for ocean noise classification [25]. Rostami et al. proposed a new TL framework to solve the problem of classifying few-shot synthetic aperture radar (SAR) images [52]. Liang et al. proposed a TL method for wind speed prediction at multiple wind farms in wind farm control centers [53]. Chen et al. proposed a TL framework for target detection in remote sensing images [54]. Transfer learning provides a model optimization approach for handling few-shot data, making it widely applicable in many domains with limited data. TL is suitable for dealing with the overfitting problem of sound speed prediction with only a few reference samples.

However, two challenges persist in current research on sound speed prediction. One is the scarcity of ocean sound speed data. The other is the mismatch between the model

and the task due to the fact that the distribution of SSP time series data at different depths has significant differences. How to address these issues and enhancing the accuracy of predictive models for few-shot SSP tasks are crucial aspects of research in marine acoustics.

To tackle the few-shot problem in ocean sound speed prediction, we establish a transfer learning framework. For the phenomenon of mismatch between model and task caused by the difference in different sea depth data, we propose the idea of stratification and integrate it into the few-shot prediction model.

3. Methodology

Due to the complexity of the ocean environment, it takes quite a long time to measure SSPs with conventional sound speed sampling equipment, resulting in historical ocean SSP data being usually scarce. Therefore, we propose a transfer learning framework for few-shot sound speed prediction to solve the overfitting problem and to improve the prediction accuracy of the SSP. Considering the inconsistent patterns of sound speed changes with time at different depths in the same location, we propose the H-LSTM sound speed prediction model based on data hierarchical processing as the internal model of the transfer learning framework and then combine this model with the transfer learning framework to form the H-LSTM-TL. In this section, we will first give the transfer learning framework for SSP prediction, then give our chosen internal model H-LSTM, and finally describe our proposed H-LSTM-TL model as well as the specific workflow in detail.

3.1. Transfer Learning Framework for SSP Prediction

Traditional neural networks generally demand a substantial volume of data for effective learning and performance optimization on specific tasks. However, when faced with new tasks or datasets, these conventional models necessitate retraining and parameter adjustments. In scenarios involving few-shot training, ordinary models are prone to overfit the noise, resulting in diminished accuracy and difficulties in achieving the desired task outcomes.

To address the overfitting challenge in few-shot learning, various methods have been proposed [55]. In particular, transfer learning has emerged as an effective solution, garnering widespread attention. Transfer learning mitigates the issue of data scarcity in target tasks by transferring knowledge and experience from a source domain. It should be noted that the key to this approach lies in ensuring that different tasks or datasets follow the same distribution, referred to as task similarity. In transfer learning methods, we begin by defining a “base model” and extensively training it with a large amount of data closely resembling the target task. Subsequently, we transfer it to a “task model” which inherits the parameters and primary features learned during the pre-training phase of the base model. Finally, we fine-tune the task model using a few-shot dataset from the target task, effectively addressing the overfitting problem associated with training on limited data.

To more clearly compare traditional machine learning with transfer learning, we present a simplified flowchart of machine learning and transfer learning in Figure 1. Machine learning is an independent training process in which a model corresponds to only one specific task. Transfer learning, on the other hand, utilizes a large amount of data to pre-train on a base model and then transfers the learned knowledge to a task model with fine-tuning or further training on the target task.

In order to quickly obtain the real-time sound speed distributions and the future sound speed distributions in the case of sample scarcity, we propose a transfer learning framework for SSP prediction, as shown in Figure 2. Our goal is mainly to acquire the main features of a large amount of SSP data in global public datasets through transfer learning, that is, to adequately train the base model with a large amount of data, so that the base model has a certain amount of prior information. Then, the base model is transferred to the task model, and a few training iterations of the task model with few-shot data can enable the task model to reach a converged state and make accurate predictions of future SSPs. In the transfer learning framework, internal models can be selected according to the actual

task, such as deep neural network models [35], Gaussian process regression models [56], and Transformer models [57].

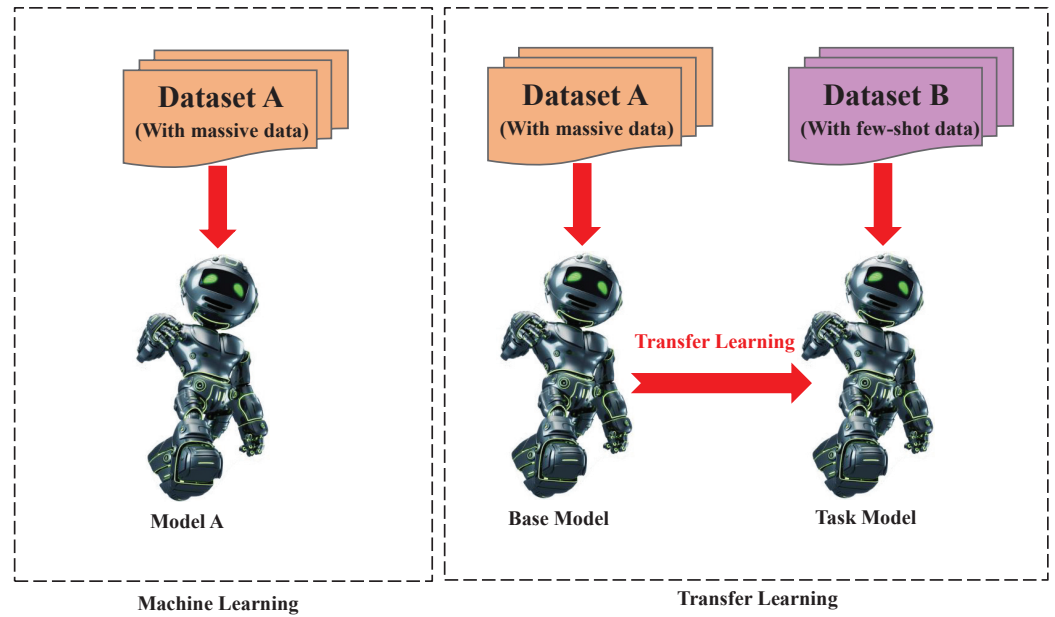


Figure 1. A simplified flowchart comparison between machine learning and transfer learning.

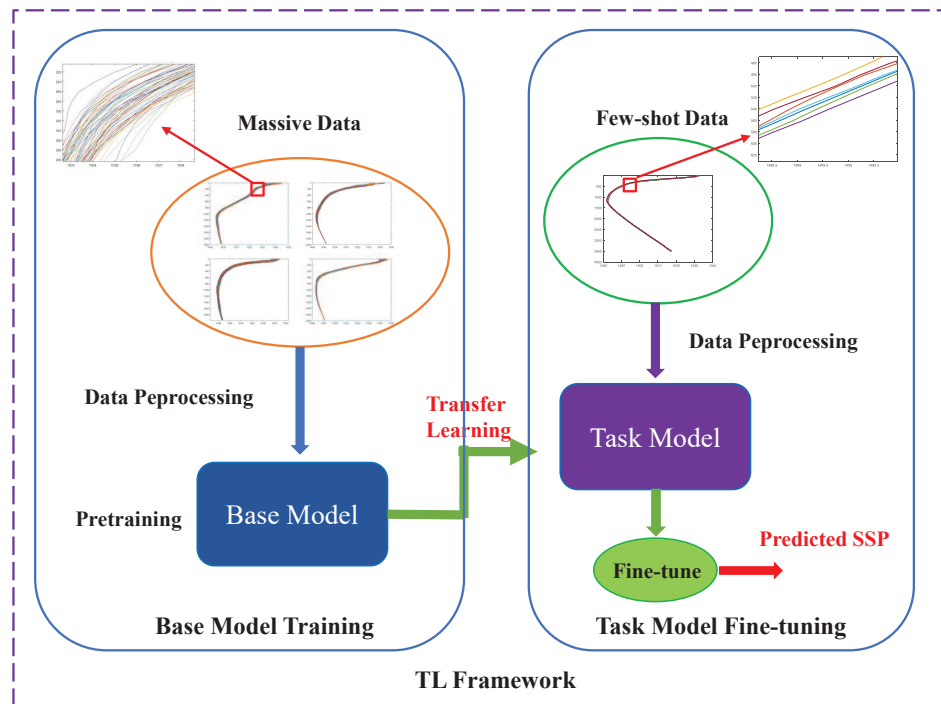


Figure 2. Transfer learning framework for SSP prediction.

3.2. H-LSTM Model

To ensure favorable prediction performance, we propose a hierarchical long short-term memory neural network as the base model and task model of the transfer learning framework. The core idea of the model is to first process the sound speed data in layers by depth and define the corresponding LSTM model for different depth layers. Then, the time series data of each depth layer are trained and predicted. The structure of the J -th layer of H-LSTM is shown in Figure 3, which mainly consists of an input layer, an H-LSTM layer, a

fully connected layer, and an output layer. The whole computational flow of the H-LSTM unit can be defined by a series of equations.

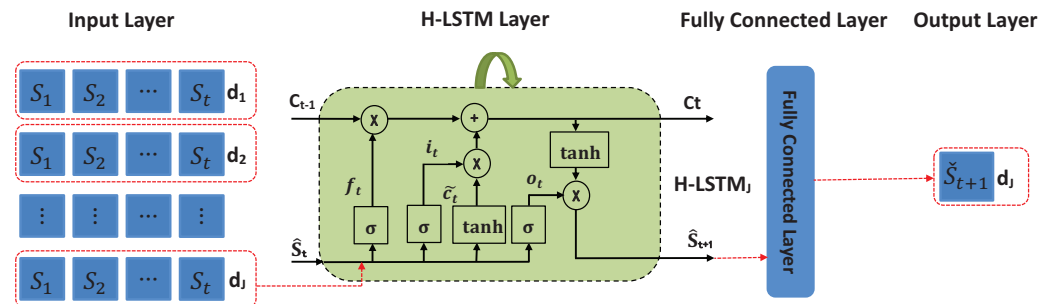


Figure 3. The structure of the J -th layer of H-LSTM.

$$f_t = \sigma(W_f \cdot [\hat{S}_t, S_t] + b_f), \tag{1}$$

$$i_t = \sigma(W_i \cdot [\hat{S}_t, S_t] + b_i), \tag{2}$$

$$\tilde{C}_t = \tanh(W_c \cdot [\hat{S}_t, S_t] + b_c), \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \tag{4}$$

$$o_t = \sigma(W_o \cdot [\hat{S}_t, S_t] + b_o), \tag{5}$$

$$\hat{S}_{t+1} = o_t \cdot \tanh(C_t), \tag{6}$$

where f_t , i_t , C_t , and o_t are the outputs of the forgetting, input, control, and output gates in the LSTM cell, W_f , W_i , W_c , and W_o are the weight matrices corresponding to the various gates, b_f , b_i , b_c , and b_o are the corresponding biases, and \hat{S}_{t+1} is the predicted SSP for the next moment of time as output from the hidden layer.

3.3. H-LSTM-TL Model

To enhance the predictive accuracy of future full-depth SSPs under few-shot conditions, we propose an H-LSTM-TL model. Through pre-training on tasks from other regions or other datasets within the same region, the model utilizes acquired knowledge to guide the training of new tasks. The aim is to address the overfitting issue on few-shot target tasks and improve the predictive accuracy of limited data. Firstly, the base model learns from other tasks, comprehensively capturing the primary features of the basic task data. Subsequently, the base model is transferred to the task model. The task model utilizes pre-learned experiences to rapidly adapt to the few-shot data with minimal fine-tuning to achieve the function of the target task. In the context of sound speed prediction, distinct base models can be established for different spatial and temporal regions of the ocean. These base models undergo pre-training using abundant remote sensing sound speed data or Argo sound speed data for the respective regions and they store the converged base models. When presented with a new task, the offline-trained base model is transferred to the task model by assessing the similarity between the study area or data of the new task and the basic task. Then, fine-tuning is performed to accomplish sound speed prediction for the new task.

The H-LSTM-TL model is mainly divided into the base model and the task model, and the working process is divided into the base model training phase, the task model fine-tuning phase, and the SSP prediction phase. Before the prediction task is assigned, the pending base models for multiple sea areas are prepared in advance, and the base model training phase is completed offline with a large amount of SSP data for the corresponding sea areas so that the trained base model can be directly used for the task model transfer when the actual sound speed prediction task is assigned. In the base model training phase, firstly, the sound speed data from the selected sea area are layered according to different depths and sorted according to the time dimension. Then, they are normalized into a

hierarchical normalized time series dataset. Secondly, the normalized dataset is input into the base model as a training set, and the hierarchical SSP data at the current moment S_t are inputted. The current time-step S_t is computed in the H-LSTM layer and will obtain the cell state C_t and the hidden layer output H_t , and the obtained C_t and H_t will also be involved in the computation of the next time-step S_{t+1} when it is used as training input. Finally, all weights are updated by calculating and back propagating the error between the predicted sound speed value and the actual sound speed value, and the training is repeated several times until the base model converges. In the task model fine-tuning phase, firstly, the important information learned during the pre-training of the base model is transferred to the task model. Second, the task model is fine-tuned with the target task few-shot data and it observes the convergence status of the task model. Finally, the task model is trained to a state of convergence after several fine-tunings. In the SSP prediction phase, the hierarchical SSP data $S_{t,dj}$ of the target task at the current moment (where dj represents the data of the j -th depth layer, $j = 1, 2, \dots, J$, and J is the maximum depth layer) are taken as the input to the task model, and the hierarchical SSP data $\hat{S}_{t+1,dj}$ of the future moment can be predicted by performing forward propagation once. Finally, the predicted sound speed data of different depth layers are combined and plugged into the SSP prediction phase. Finally, the predicted sound speed data in different depth layers are combined and interpolated to obtain the future full-ocean depth SSP.

To facilitate the intuitive understanding of the core idea of the H-LSTM-TL model, the basic structure of the model is shown in Figure 4. The figure shows the structure of the J -th layer, and the remaining depth layers are similar.

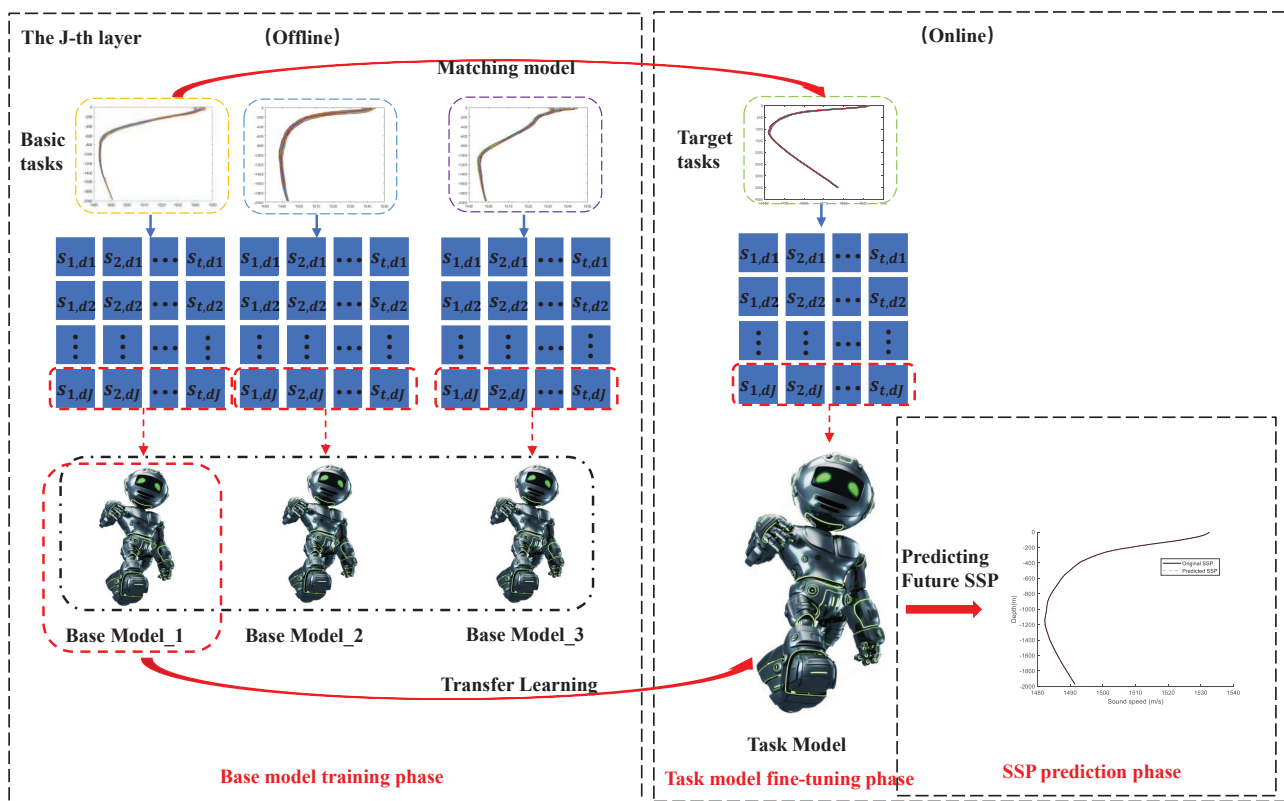


Figure 4. The basic structure of H-LSTM-TL (the J -th layer).

3.4. Workflow of H-LSTM-TL for SSP Prediction

The specific workflow of the H-LSTM-TL-based SSP prediction method is shown in Figure 5, including data preprocessing, H-LSTM neural network construction, base model pre-training, model transfer, model matching, task model fine-tuning, and SSP prediction. We will give the specific implementation steps of the method in the following.

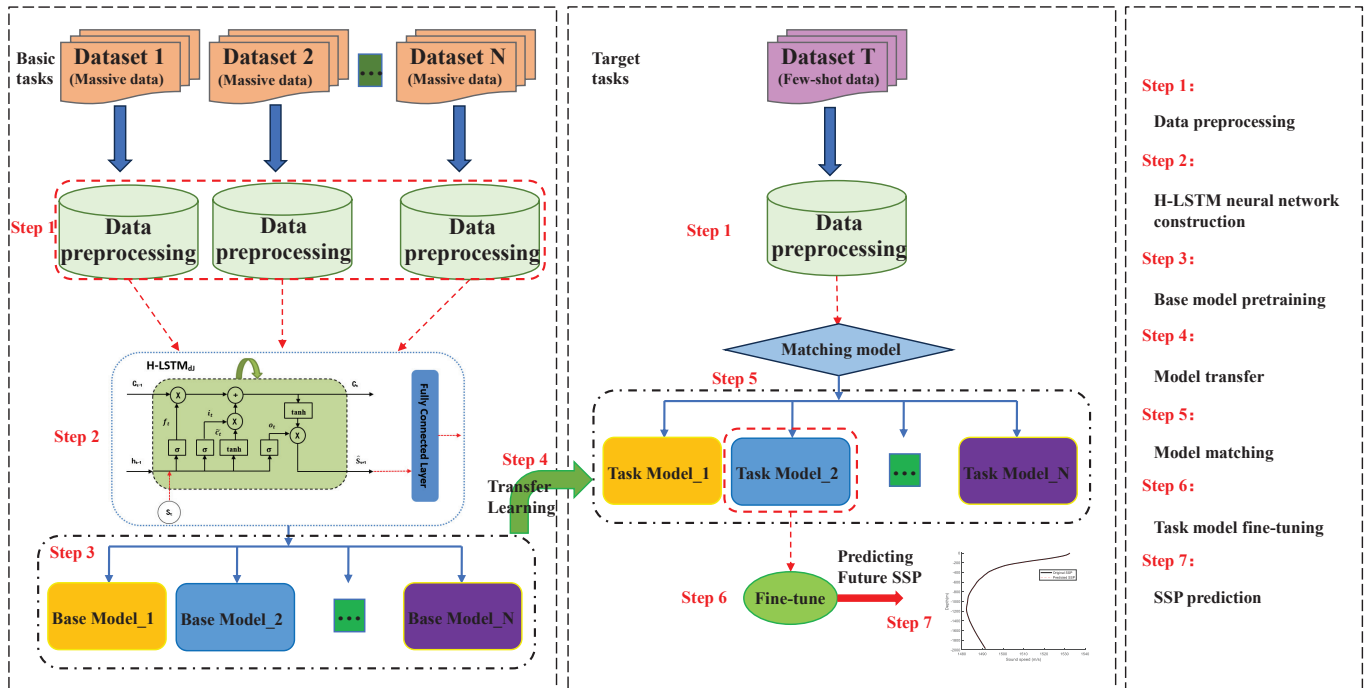


Figure 5. The workflow of the H-LSTM-TL.

Step 1. Data preprocessing:

We need to prepare a large amount of historical SSP data from multiple sea areas. These data are intended for the pre-training of the base model, so that the trained base model can be directly used for the transfer of the task model for SSP prediction in the relevant sea areas when given the actual sound speed prediction task. Since the H-LSTM-TL model is trained by hierarchical SSP data, we need to perform hierarchical processing on the SSP data of the different selected sea areas. The hierarchical prediction approach can effectively solve the problem that ordinary LSTM models lose the feature information of historical data when trained at different depths.

Taking the historical SSP data of one of the selected sea areas as an example, firstly, the I historical SSPs are sequentially standardized and linearly stratified in the depth direction, and a total of J layers are divided. Then, the historical data are processed to be sorted in the time dimension to form the J -layer SSP time series. The hierarchical standardized time series dataset of this sea domain is given in the form of a matrix in Equation (7).

$$S = \begin{bmatrix} s_{1,d_1} & s_{2,d_1} & \dots & s_{I,d_1} \\ s_{1,d_2} & s_{2,d_2} & \dots & s_{I,d_2} \\ \dots & \dots & \dots & \dots \\ s_{1,d_J} & s_{2,d_J} & \dots & s_{I,d_J} \end{bmatrix}, \tag{7}$$

where s_{I,d_J} is the sound speed value at the J -th depth layer for the I -th SSP sorted by time, with I being the time index and J being the depth layer index. Data preprocessing for the target task's few-shot SSPs is performed in a similar way, differing only in a smaller number of SSPs.

Step 2. H-LSTM neural network construction:

According to the hierarchical normalized time series data in Step 1, an H-LSTM network is constructed for each depth layer of the time series, the number of nodes in the input layer is 1, an H-LSTM layer is taken as the hidden layer, the number of hidden units is taken as N , and a fully connected layer with a linear activation function is added between the hidden layer and the output layer.

Step 3. Base model pre-training:

The H-LSTM models for each depth layer are fully pre-trained with hierarchical standardized time series datasets from different sea areas, which enables the base model to capture the relationship between the training inputs and outputs and to learn the changing patterns of multiple types of time series SSPs. After the model reaches convergence, the base models corresponding to the different oceanic data are saved for further use in transfer learning.

Step 4. Model transfer:

The base model saved in Step 3 is transferred to different task models that have prior knowledge related to the base task data. The purpose of transfer learning is to utilize the knowledge and experience of the base task to accelerate the learning process of the target task. Since these task models are not trained from scratch, they can quickly learn the distribution pattern of the new SSPs' data with less training on the few-shot data of the new task, maintain sensitivity to the changes in the sound speed distribution, and are less prone to the degradation of the prediction accuracy due to overfitting.

Step 5. Model matching:

When the target task is assigned, the data of the target task are compared with the base task data. Regional correlation or data similarity is considered to match the appropriate task model. The data similarity is measured by the mean square error MSE, with smaller values representing higher similarity.

$$MSE = \frac{1}{J} \sum_{j=1}^J (\hat{s}_{mean,dj}^T - s_{mean,dj}^B)^2, j = 1, 2, \dots, J, \tag{8}$$

where $\hat{s}_{mean,dj}^T$ and $s_{mean,dj}^B$ are the sound speed averages of the target task data and the base task data in the j -th layer, respectively, and J is the maximum co-depth layer.

Step 6. Task model fine-tuning:

In the task model fine-tuning phase, the matched task model is further trained and optimized using the target task few-shot data. The training in this phase is based on the pre-trained model, and then the parameters and structure of the model are fine-tuned to better adapt to the target task to achieve more accurate SSP prediction in the future.

Step 7. SSP prediction:

In the SSP prediction phase, firstly, the hierarchical SSP data at the current moment of the target task are taken as input, and the future SSP data in this depth layer can be predicted by performing one-time forward propagation. Secondly, the predicted sound speed data from different depth layers are combined in ascending order of depth layers, and the combined hierarchical SSPs are interpolated at full-ocean depth. Finally, the interpolated future full-ocean depth SSP is obtained. To verify the accuracy of the model for SSP prediction, we calculate the root mean square error (RMSE) between the predicted SSP and the actual SSP, and a smaller RMSE represents a higher prediction accuracy.

$$RMSE = \sqrt{\frac{\sum(\hat{s}_{p,dj} - s_{r,dj})^2}{J}}, j = 1, 2, \dots, J, \tag{9}$$

where $\hat{s}_{p,dj}$ and $s_{r,dj}$ are the sound speed values of the predicted SSP and the actual SSP in the j -th layer, respectively, and J is the maximum co-depth layer.

4. Results and Discussion

4.1. Data Preprocessing

4.1.1. Data Sources

To verify the performance of the proposed H-LSTM-TL in this paper for predicting the SSP, we conducted a deep-sea experiment in the South China Sea in mid-April 2023, with an area of $10 \text{ km} \times 10 \text{ km}$ and a depth of more than 3500 m, and the collected SSP data lasted for three days. Since the 14 SSP data collected were at different moments of the 3 days, covering almost a whole day in time, we approximated the 14 SSP data as the 24 h SSP changes in the area, with time intervals of approximately every two hours.

Due to the long time taken to collect sound speed data by SSP measurement devices such as CTDs and XCTDs, we collected relatively small numbers of SSPs in our ocean experiments. Training a model with a small amount of data may lead to overfitting problems because neural network models can easily memorize few-shot data, thus failing to learn the basic features and patterns of the data. When a model relies too greatly on a small amount of data, its generalization ability is also weakened and it is unable to accurately make predictions when faced with new task data.

To solve the overfitting, we selected historical South China Sea sound speed data from the global Argo dataset [58], which is strongly correlated with the Ocean Experiment collection. We also selected historical sound speed data from the Pacific, Atlantic, and Indian Oceans. The base model in H-LSTM-TL is pre-trained using Argo data, and the converged base model is transferred to a new task (Ocean Experiment data).

The specific spatial locations of the ARGO SSP data (base model) and the Ocean Experiment SSP data (task model) as well as how the data are divided for the task model are given in Figure 6, and detailed data information is shown in Table 1.

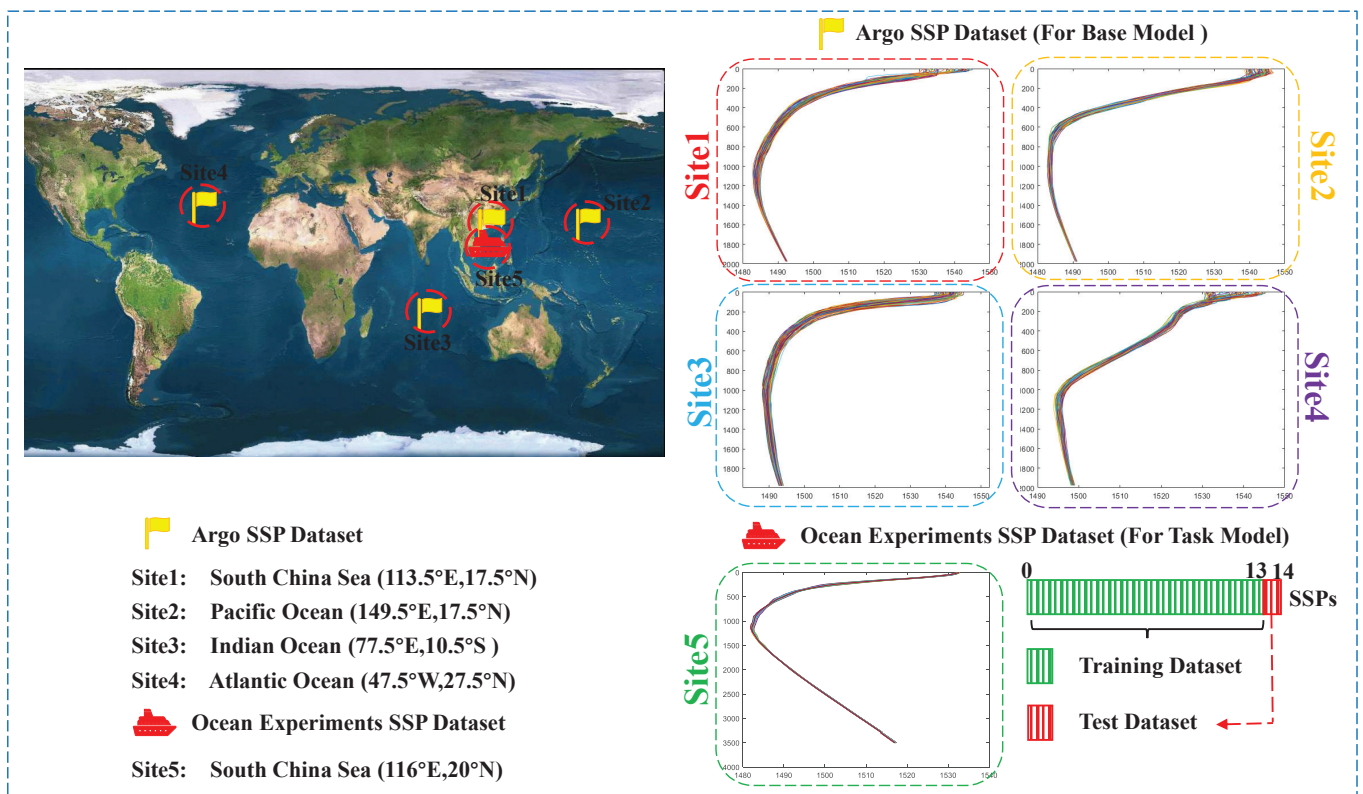


Figure 6. Data source.

Table 1. Detailed data information.

Argo Dataset (For Base Model)					
Study Area	Time Dimension	Temporal Resolution	Number of SSPs	Full Depth	Layered Processing
South China Sea (113.5° E, 17.5° N)	2017–2021 (60 months)	Month mean	60	0–1975 m	Unequal interval (58 layers)
Pacific Ocean (149.5° E, 17.5° N)					
Indian Ocean (77.5° E, 10.5° S)					
Atlantic Ocean (47.5° W, 27.5° N)					
Ocean Experiments Dataset (For Task Model)					
Study Area	Time Dimension	Temporal Resolution	Number of SSPs	Full Depth	Layered Processing
South China Sea (116° E, 20° N)	27 March 2023 (24 h)	Approximately two hours mean	14	0–1975 m	Unequal interval (58 layers)

4.1.2. Hierarchical Processing of Data

The depth of the SSP data acquired by the Ocean Experiment exceeded 3500 m, but the maximum depth of the Argo data for model pre-training was only 1975 m. To maintain consistency between the base model and the task model, the Ocean Experiment SSP data were interpreted as 0–1975 m. Although the maximum depth of the data is only 1975 m, the change in SSP in the deep-sea part is relatively small (a linear trend with depth), and 1975 m already contains the typical SSP structure including the surface layer, seasonal thermocline, main thermocline, and deep-sea isothermal layer of the ocean SSP. Thus, 0–1975 m can already approximately represent the main characteristics of the whole-ocean depth SSP data.

Due to the H-LSTM-TL needing to predict the SSPs in layers, if the number of SSP data layers is too large, it will seriously intensify the time consumption of model training. The full-ocean depth SSP is divided into 58 depth layers at unequal intervals in the depth direction, and the standardized layering is as follows: from 0 to 10 m, one layer is divided every 5 m; from 10 to 180 m, one layer is divided every 10 m; from 180 to 460 m, one layer is divided every 20 m; from 500 to 1250 m, one layer is divided every 50 m; from 1300 to 1900 m, one layer is divided every 100 m; the depth over 1900 m is divided into one layer, making a total of 58 layers. This method of layering can fully represent the main features of the full-ocean depth SSP. After the layering process, the SSP data are then sorted in the time dimension.

Taking the Ocean Experiment SSP data (14 SSPs) as an example, the hierarchical normalized dataset S after depth stratification and time-ordering processing can be represented as a matrix of sound speed values in 58 rows and 14 columns, and the Argo data are layered similarly.

$$S = \begin{bmatrix} s_{1,d_1} & s_{2,d_1} & \cdots & s_{14,d_1} \\ s_{1,d_2} & s_{2,d_2} & \cdots & s_{14,d_2} \\ \cdots & \cdots & \cdots & \cdots \\ s_{1,d_{58}} & s_{2,d_{58}} & \cdots & s_{14,d_{58}} \end{bmatrix}. \tag{10}$$

4.2. Parameter Settings of H-LSTM-TL

The experiments involved in this paper were implemented in MATLAB R2023b. To ensure that the base and task models in the hybrid model H-LSTM-TL can be trained to convergence, we performed several parameter adjustments and finally determined the

appropriate model parameters. The main parameter settings of H-LSTM-TL are shown in Table 2.

Table 2. Key parameter setting of H-LSTM-TL.

Key Parameters	Settings	
	Base Model	Task Model
Layers	4	4
Hidden size	150	150
Learning rate	0.01	0.01
Epoch	300	30
Loss function	MSE	MSE

4.3. Prediction Accuracy

In order to solve the problem of model overfitting or underfitting caused by too few Ocean Experiment data, we pre-trained the base model 300 times with the historical South China Sea Argo SSP data, which is strongly correlated with the Ocean Experiment data, then transferred the converged base model to the task model, then fine-tuned the task model with the Ocean Experiment data 30 times, and finally predicted the SSPs in future moments. Additionally, we performed accuracy comparison experiments by directly testing the H-LSTM model, back propagation (BP) neural network model [38], and polynomial fitting (PF) method [59] using Ocean Experiment data.

The base model in H-LSTM-TL is pre-trained with five years of South China Sea Argo data from 2017 to 2021. The task model and other models are trained with the first 13 SSP data of the measured approximate 24-hour South China Sea Ocean Experiment data, and they predict the SSP for the next two hours, respectively. A comparison of the predicted hierarchical SSP with the original hierarchical SSP for the H-LSTM-TL and other models is shown in Figure 7.

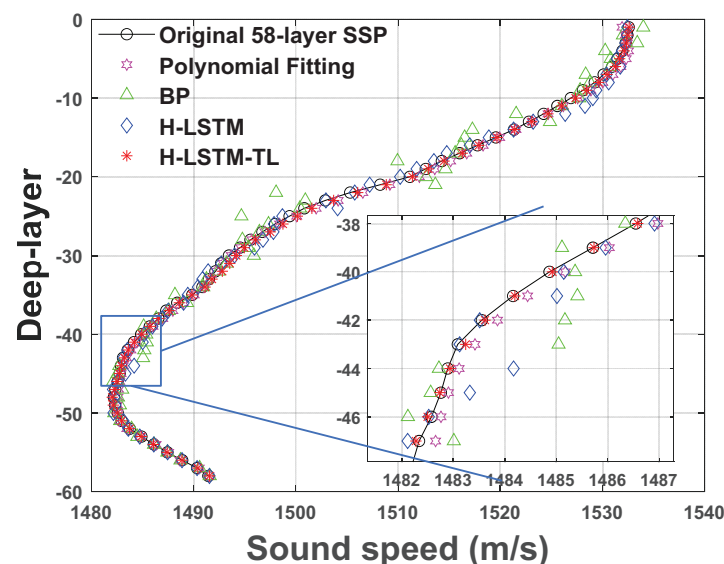


Figure 7. Comparison of predicted hierarchical SSP with original hierarchical SSP.

Figure 7 reveals that the hierarchical SSP predicted by H-LSTM-TL closely resembles the original profile, with minimal errors at each depth layer. In scenarios with limited data, the result of polynomial fitting methods provides predictions close to the original data, but its accuracy still falls short compared to H-LSTM-TL. However, both the BP model

and H-LSTM model exhibit certain shortcomings in handling few-shot data, as is evident from their poorer performance across most depth layers. It is evident that H-LSTM-TL outperforms other models in predicting layered SSP at each depth layer, demonstrating superior performance compared to other state-of-the-art methods.

To more clearly compare the accuracy performance of several models when performing layered prediction, the errors of some randomly selected depth layers from the 58 predicted layers of SSP are given in Table 3.

Table 3. RMSE of different depth layers.

Depth Layer	RMSE (m/s)				Depth Layer	RMSE (m/s)			
	H-LSTM-TL	H-LSTM	PF	BP		H-LSTM-TL	H-LSTM	PF	BP
8	0.4030	1.3641	0.7925	0.4753	15	0.1063	0.6856	0.3661	3.0507
9	0.3707	1.3053	0.7895	0.2721	16	0.1336	0.7599	0.5481	1.5823
10	0.4578	2.1526	0.6754	0.8159	17	0.2615	1.2522	0.6993	0.4375
11	0.4308	2.7047	0.5897	0.4375	18	0.2163	0.7768	0.8578	4.4335
12	0.4242	2.1136	0.4372	2.6509	19	0.2534	0.8888	0.8827	1.9334
13	0.2746	0.4040	0.3230	2.6500	20	0.2606	0.9427	0.9290	1.5949
14	0.0704	0.0426	0.1983	4.0691	21	0.5941	1.0218	0.9545	5.3949

From the RMSE of different depth layers, it can be seen that the prediction performance of our proposed H-LSTM-TL model is much superior to other models without transfer learning with training on the few-shot data. In the errors given for the 14 depth layers, the prediction errors of the H-LSTM-TL model are much smaller than those of other models in most of the depth layers, and the error in the 14th layer is only slightly larger than that of H-LSTM, but there is only a difference of less than 0.03 m/s, which is negligible. The results of the layered prediction fully show that the H-LSTM-TL model can accurately predict SSPs in future moments and the predicted SSP data can accurately represent the main features of the actual SSP.

Due to the fact that the training data for several models are not full-ocean depth data, but rather data from 58 unequally spaced depth layers divided over the range of 0–1975 m ocean depth, the predictions are also stratified. In order to test the prediction performance of H-LSTM-TL for future full-ocean depth SSPs, the predicted stratified sound speed distributions are interpolated at full-ocean depth by linear interpolation, and the full-ocean depth range is taken as 0–1975 m. A comparison of the full-ocean depth SSPs predicted by several models with the actual full-ocean depth SSP is shown in Figure 8.

It can be seen from the comparison between the full-ocean depth SSPs predicted by several models and the actual full-ocean depth SSP that the full-ocean depth SSPs predicted by the H-LSTM-TL model almost fit perfectly with the actual full-ocean depth SSP, with very high similarity. Under the condition of the same number of training times, the full-ocean depth SSP predicted by the H-LSTM model only fits better below 1200 m depth, and the prediction effect is generally poor in the range of 0–1200 m depth, although it can also roughly represent the main features of the actual full-ocean depth SSP. In the comparison of the prediction effects of both models, the H-LSTM model without transfer learning is much less effective in prediction than H-LSTM-TL. For the polynomial fitting method, the overall errors in predicting the full-ocean depth SSP are relatively uniform, providing a rough representation of the future SSP across the entire ocean depth. However, the BP model exhibits poor capability in handling small sample data to the extent that the SSP predicted by the BP model is not suitable for representing the future full-ocean depth SSP.

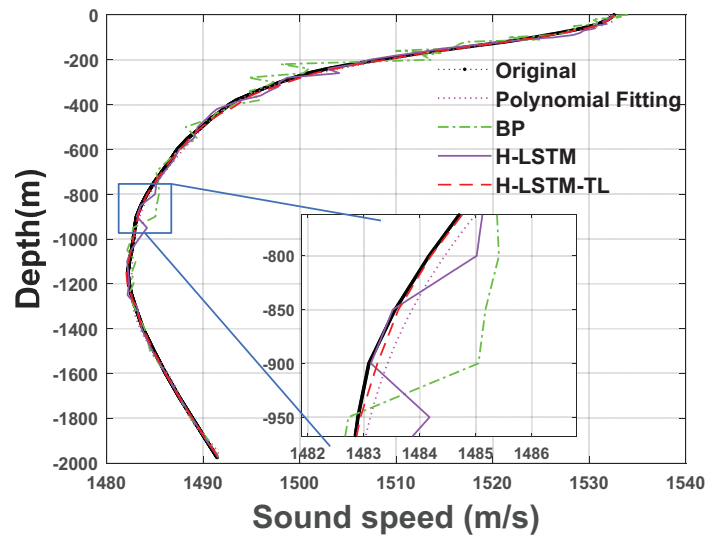


Figure 8. Comparison of predicted SSP with original SSP.

The sound speed disturbances of several models predicting the SSP with the actual SSP as the reference are given in Figure 9. The red line is the sound speed disturbance of the SSP predicted by H-LSTM-TL, and it can be seen that below the 400 m ocean depth, the sound speed disturbances are all within 0.5 m/s, and the sound speed disturbances in the shallow portion of the ocean, where the marine environment is complex, do not exceed 1 m/s, which is far more effective than that of other models. The full-ocean depth prediction errors of training the task model and the H-LSTM model 30 times, respectively, are given in Table 4, and the full-ocean depth prediction error of the H-LSTM-TL model is 0.2736 m/s, while that of the H-LSTM model is 0.5788 m/s, which is more than twice as large as that of the H-LSTM-TL. Because training H-LSTM only 30 times is far from convergence, we trained the H-LSTM model alone 150 times with Ocean Experiment data, and the model reached convergence with a full-ocean depth prediction error of 0.3858 m/s, which is still not as accurate as that of H-LSTM-TL. Additionally, Table 4 presents the overall prediction errors for other models. The RMSE of the polynomial fitting method and BP model are 0.4013 m/s and 1.1328 m/s, respectively. The experimental results show that our proposed H-LSTM-TL can accurately capture the internal features within few-shot data and make accurate predictions of the full-ocean depth SSP at future moments.

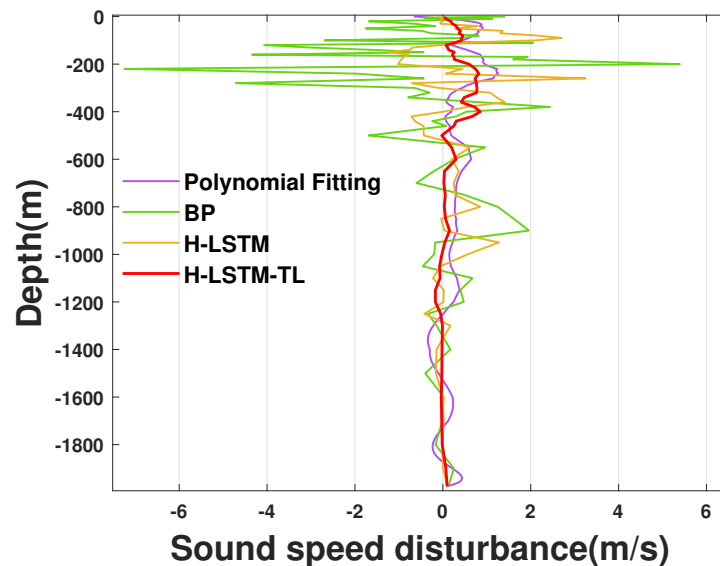


Figure 9. Sound speed disturbance of several models.

Moreover, we also tested the number of fine-tunings of the H-LSTM-TL model on the few-shot ocean SSP data. The amount of base model pre-training was fixed at 300 times to ensure that the base model could be trained to convergence. The fine-tuning times of the task model after transfer learning were taken as 30 times, 20 times, 15 times, 10 times, and 5 times to test the H-LSTM-TL. The corresponding full-ocean depth prediction errors are given in Table 4.

Table 4. RMSE for different numbers of training times.

H-LSTM-TL						H-LSTM		PF	BP
Dataset						Dataset		Dataset	
Base Model (South China Sea Argo)			Task Model (Ocean Experiments)			Ocean Experiments		Ocean Experiments	
Epochs (Task Model)						Epochs			
30	25	20	15	10	5	30	150	RMSE (m/s)	
RMSE (m/s)						RMSE (m/s)			
0.2736	0.2353	0.2133	0.1624	0.1416	0.1057	0.5788	0.3858	0.4013	1.1328

It can be seen from the full-ocean depth prediction errors with different numbers of fine-tuning times given in the table that the RMSE between the SSP predicted by the H-LSTM-TL model and the actual SSP also decreases gradually as the number of fine-tuning times of the task model decreases from 30 times to 5 times. Under the condition of training with few-shot data, this may have occurred because the task model with transfer learning had already reached a fit with only a few fine-tuning training times. If the task model is trained several times with few-shot data, the features captured by the model will become closer and closer to the case when the H-LSTM model is trained directly with the few-shot data, so that the transferred task model also reaches overfitting, and the accuracy becomes more and more worse as the number of training times increases. In general, it is only necessary to fine-tune the task model in our proposed H-LSTM-TL model 5 times to achieve accurate prediction accuracy. To enhance comparability with the non-transfer learning model, the task model underwent 30 fine-tuning iterations in the comparative experiments. This is because, for the ordinary model without transfer learning, no meaningful features can be learned in just 5 iterations.

4.4. Generalization Ability

Generalization ability, within the context of machine learning, refers to the capacity of an algorithm to effectively accommodate new samples. Its primary objective lies in discerning the underlying patterns within the data, such that the trained network can also accurately predict data outside the learning set with the same law. However, conventional neural network models often encounter overfitting issues, particularly when confronted with limited training data, consequently diminishing their generalization prowess.

In previous experiments, we selected data for the pre-training of the base model with a strong correlation with the new tasks, and both datasets were derived from the relevant regions of the South China Sea. In order to verify that our proposed H-LSTM-TL has good generalization ability, we selected Argo SSP data from the Pacific, Indian, and Atlantic Oceans, which are quite different from the sound speed distribution characteristics of the South China Sea, as the pre-training data for multiple base models. After pre-training on data from various marine regions, the base model was transferred to the task model. The training and testing of the task model use limited samples, preventing excessive learning of sample features. This approach preserves sensitivity to input data, ensuring the reliability of H-LSTM-TL for accurate prediction of few-shot ocean SSPs. The prediction results of the three regional models are shown in Figure 10.

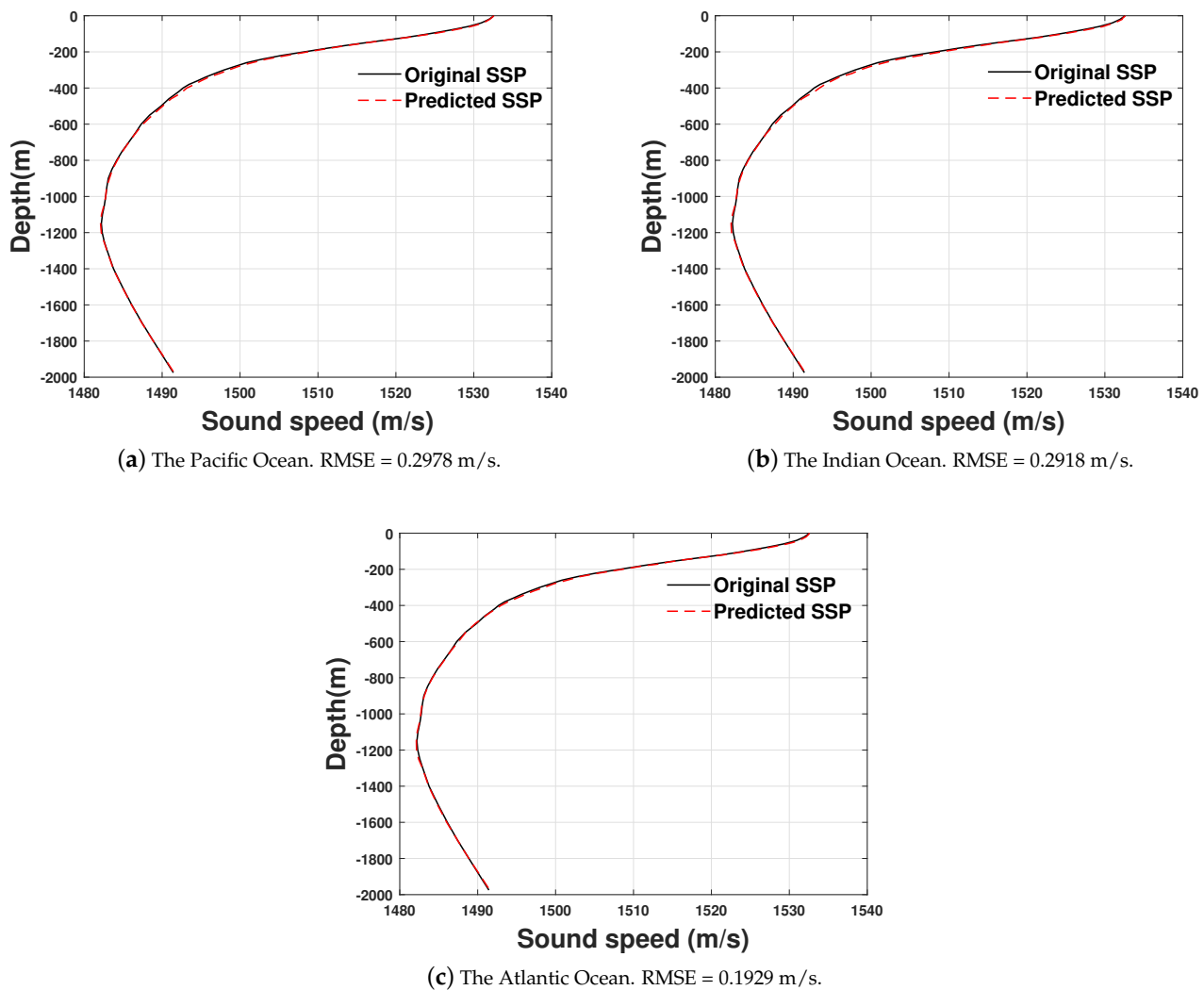


Figure 10. Prediction results of three regional models.

As shown in Figure 10, when the Argo data of the three regions are used as the pre-training data for the base model, the full-ocean depth SSPs predicted by H-LSTM-TL at future moments all perform well, and the predicted full-ocean depth SSPs can accurately represent the main features of the actual SSPs. For base model 1 (Pacific Ocean), the RMSE of task model 1 after transfer learning is 0.2978 m/s; for task model 2 (Indian Ocean), the RMSE is 0.2918 m/s; and for task model 3 (Atlantic Ocean), the RMSE is 0.1929 m/s. The RMSE values of all three models are less than 0.3 m/s, which can well represent the future full-ocean depth SSP.

Meanwhile, we can also notice that the prediction effect of the Atlantic Ocean Argo data is superior to that of other models when it is used as the pre-training data for the base model. This is because the original Argo data in the Atlantic Ocean have a great correlation with the distribution of the target task data, and the sound fixing and ranging (SOFAR) channel of the Atlantic Ocean SSP and the South China Sea experiments' SSP are both near 1000 m depth, while the SOFAR channel of the SSPs of the other two sea areas are near 600 m depth.

The experimental results fully showed that our proposed H-LSTM-TL is good at learning the generic features of the few-shot task data and capturing the general patterns in the data, to better and accurately predict future full-ocean depth SSPs. Meanwhile, the

model also has good generalization ability and general applicability for predicting SSP distribution in different regions with different features.

4.5. Convergence Comparison

To observe more intuitively the convergence of our proposed H-LSTM-TL model with the H-LSTM model without transfer learning under the training of few-shot data, the convergence of both models is shown in Figure 11, where the two lines in the lower half correspond to the trend of the Loss value during the training process and the two lines in the upper half correspond to the trend of the RMSE during the training process. It can be noticed that with H-LSTM-TL, the model only needs to be trained 25 times for the Loss value to drop below 0.05, reaching a converged state. When the model is trained up to 50 times, the Loss value tends to 0, reaching a state of complete convergence. When using H-LSTM, the model has to be trained at least 150 times if it is to reach the fully converged situation. It can also be noticed that after 50 iterations of training, the RMSE of the H-LSTM-TL model drops below 0.2 m/s, while the RMSE of the H-LSTM model only drops to about 0.4 m/s. The experimental results show that the H-LSTM-TL model converges much faster than the H-LSTM before improvement when learning with few-shot data. This means that the H-LSTM-TL model can learn internal dependencies in data more quickly, enabling it to enter a stable prediction phase earlier and thereby improving the time efficiency of predicting future SSP.

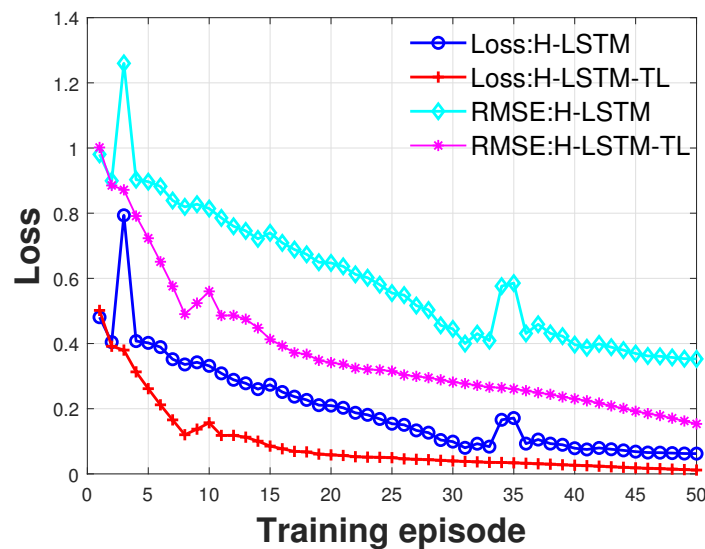


Figure 11. Convergence comparison of both models.

4.6. Time Efficiency Comparison

The time efficiency of future sound speed prediction is crucial for our further expansion of work on large-scale ocean SSP prediction. In this paper, our proposed H-LSTM-TL model is divided into a base model and a task model. Due to the fact that the pre-training process of the base model can be accomplished offline before transfer learning, the corresponding base model can be used for multiple transfers for different prediction tasks by simply saving it after pre-training. Therefore, for the H-LSTM-TL model, the only time overhead that needs to be considered is the task model fine-tuning phase.

In previous experiments, we analyzed the convergence performance of both models. The task model of H-LSTM-TL can reach a converged state after 25 iterations of training and can reach a fully converged state after 50 iterations of training, whereas it takes 150 iterations of training for the H-LSTM model to reach a converged state similar to that of H-LSTM-TL. The comparison of the time efficiency of both models for training to convergence is shown in Figure 12, where the H-LSTM takes longer to reach convergence, requiring 160 s, while the H-LSTM-TL takes only 58 s to reach full convergence, which is much more

time-efficient. The shorter training time not only enhances the efficiency of SSP prediction but also signifies reduced computational resources consumed during the training process. This is particularly crucial for large-scale applications of SSP prediction.

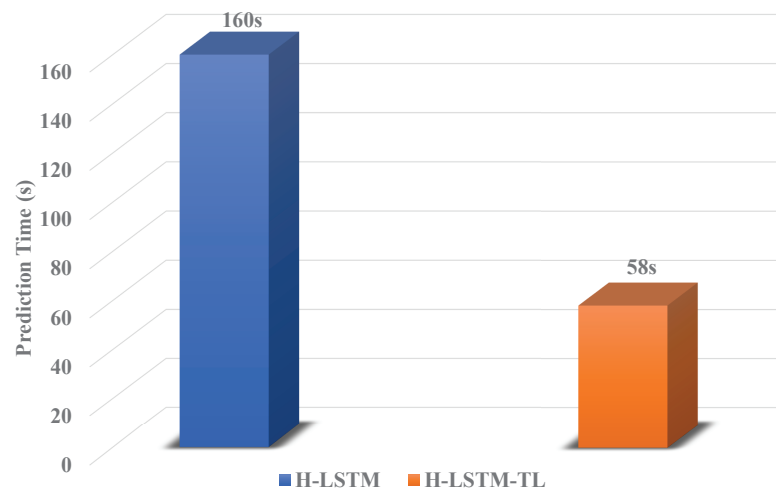


Figure 12. Time overhead of SSP predictions.

5. Conclusions

To realize the accurate prediction of future full-ocean depth ocean SSPs and to improve the real-time availability of ocean engineering sound speed data, it is particularly important to reasonably utilize historical underwater sound speed data to predict SSPs. In this paper, we proposed an H-LSTM-TL method for full-ocean depth SSP prediction. The method established the transfer learning framework H-LSTM-TL based on the deep learning LSTM neural network combined with the improved H-LSTM model with the idea of layering, which successfully reduced the number of training times for the model to reach convergence with few-shot data and improved the prediction efficiency of SSP.

To verify the feasibility and validity of the model, we performed deep-sea experiments to obtain few-shot SSP data. Through the verification of the measured few-shot SSP data, it was shown that the proposed H-LSTM-TL model can obtain better prediction accuracy, faster convergence speed, and better time efficiency. Moreover, we also tested the model on three typical sea areas and verified that the model has excellent generalization ability. The experimental results showed that the proposed H-LSTM-TL model in this paper is suitable for underwater sound speed prediction in the case of few-shot data, which can balance the accuracy and time efficiency of prediction and also has a strong generalization ability. Given the notable performance of H-LSTM-TL in few-shot time series modeling, it is widely applicable to various time series forecasting tasks. Particularly in marine-related fields, the potential application value of H-LSTM-TL is significant. For instance, it can be extended to achieve precise predictions in areas such as sea temperature, salinity, and turbulence.

The accurate prediction of SSP is crucial for underwater acoustic communication as it effectively aids in beam forming design, ensuring the rationality of signal transmission angles and thereby enhancing energy utilization efficiency. Furthermore, in the field of localization, precise SSP prediction combined with ray tracing techniques can significantly improve positioning accuracy, providing robust support for performance optimization and precise localization in underwater communication systems.

In future research, we will focus on integrating real-time satellite remote sensing data of ocean surface conditions with Argo data to achieve precise SSP predictions in various specific marine regions, such as areas with mixed currents like the Gulf Stream and regions with frequent internal solitary waves like the South China Sea. Additionally, we plan to combine the H-LSTM-TL model with advanced machine learning architectures, such as convolutional neural networks (CNNs) and attention mechanisms. This integration aims to

overcome the limitations of H-LSTM-TL in handling spatial data and to further enhance the model's ability to capture complex dependencies within spatiotemporal data.

Author Contributions: Conceptualization, J.L., W.H. and H.Z.; methodology, J.L. and W.H.; software, J.L. and W.H.; validation, J.L., S.L. and P.W.; formal analysis, J.L., W.H. and H.Z.; investigation, J.L., W.H. and H.Z.; resources, W.H. and H.Z.; writing—original draft preparation, J.L.; writing—review and editing, W.H., H.Z., S.L. and P.W.; funding acquisition, W.H. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Shandong Province (ZR2023QF128), Laoshan Laboratory (LSKJ202205104), China Postdoctoral Science Foundation (2022M722990), Qingdao Postdoctoral Science Foundation (QDBSH20220202061), National Natural Science Foundation of China (NSFC:62271459), National Defense Science and Technology Innovation Special Zone Project: Marine Science and Technology Collaborative Innovation Center (22-05-CXZX-04-01-02), and the Fundamental Research Funds for the Central Universities, Ocean University of China (202313036).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Argo data were made available by the China Argo real-time data center on the web <http://www.argo.org.cn/index.php?m=content&c=index&a=lists&catid=27> (accessed on 17 June 2024).

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

References

- Erol-Kantarci, M.; Mouftah, H.T.; Oktug, S. A survey of architectures and localization techniques for underwater acoustic sensor networks. *IEEE Commun. Surv. Tutorials* **2011**, *13*, 487–502. [[CrossRef](#)]
- Wang, D.; Shang, E. *Marine Acoustic*; China Science Publishing and Media Ltd. (CSPM): Beijing, China, 2013. (In Chinese)
- Liu, B.; Huang, Y.; Chen, W.; Lei, J. *Principles of Underwater Acoustics*; China Science Publishing and Media Ltd. (CSPM): Beijing, China, 2019. (In Chinese)
- Chu, P.C.; Fan, C. World Ocean Thermocline Weakening and Isothermal Layer Warming. *Appl. Sci.* **2020**, *10*, 8185. [[CrossRef](#)]
- Jensen, F.B.; Kuperman, W.A.; Porter, M.B.; Schmidt, H. *Computational Ocean Acoustics: Chapter 1*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011; pp. 3–4. [[CrossRef](#)]
- Huang, W.; Zhou, J.; Gao, F.; Lu, J.; Li, S.; Wu, P.; Wang, J.; Zhang, H.; Xu, T. Underwater Sound Speed Profile Construction: A Review. *arXiv* **2023** arXiv:2310.08251v1. <https://doi.org/10.48550/arXiv:2310.08251>.
- Zhang, S.; Xu, X.; Xu, D.; Long, K.; Shen, C.; Tian, C. The design and calibration of a low-cost underwater sound velocity profiler. *Front. Mar. Sci.* **2022**, *9*, 996299. [[CrossRef](#)]
- SST GmbH. Sea & Sun Technology CTD Probes. Available online: <https://www.sea-sun-tech.com/> (accessed on 17 June 2024).
- Williams, A. CTD (conductivity, temperature, depth) profiler. In *Encyclopedia of Ocean Sciences: Measurement Techniques, Sensors and Platforms*; Steele, J.H., Thorpe, S.A., Turekian, K.K., Eds.; Elsevier: Boston, MA, USA, 2009; pp. 25–34.
- Tsurumi-Seiki Co., Ltd. eXpendable Conductivity, Temperature and Depth Product Code: XCTD. Available online: <https://tsurumi-seiki.co.jp/en/product/e-sku-2/> (accessed on 17 June 2024).
- Chen, C.; Ma, Y.; Liu, Y. Reconstructing Sound speed profiles worldwide with Sea surface data. *Appl. Ocean Res.* **2018**, *77*, 26–33. [[CrossRef](#)]
- Tolstoy, A.; Diachok, O.; Frazer, L. Acoustic tomography via matched field processing. *J. Acoust. Soc. Am.* **1991**, *89*, 1119–1127. [[CrossRef](#)]
- Taroudakis, M.I.; Markaki, M.G. Matched field ocean acoustic tomography using genetic algorithms. In *Acoustical Imaging*; Springer: Berlin/Heidelberg, Germany, 1995; pp. 601–606.
- Choo, Y.; Seong, W. Compressive sound speed profile inversion using beamforming results. *Remote Sens.* **2018**, *10*, 704. [[CrossRef](#)]
- Bianco, M.; Gerstoft, P. Compressive Acoustic Sound Speed Profile Estimation. *J. Acoust. Soc. Am.* **2016**, *139*, EL90–EL94. [[CrossRef](#)] [[PubMed](#)]
- Li, Q.; Shi, J.; Zhenglin, L.; Yu, L.; Zhang, K. Acoustic sound speed profile inversion based on orthogonal matching pursuit. *Acta Oceanol. Sin.* **2019**, *38*, 149–157. [[CrossRef](#)]
- Huang, W.; Li, D.; Jiang, P. Underwater Sound Speed Inversion by Joint Artificial Neural Network and Ray Theory. In Proceedings of the Thirteenth ACM International Conference on Underwater Networks & Systems (WUWNet'18), Shenzhen, China, 3–5 December 2018; ACM: New York, NY, USA, 2018; pp. 1–8. [[CrossRef](#)]

18. Stephan, Y.; Thiria, S.; Badran, F. Inverting tomographic data with neural nets. In Proceedings of the 'Challenges of Our Changing Global Environment'. Conference Proceedings, OCEANS'95 MTS/IEEE, San Diego, CA, USA, 9–12 October 1995; IEEE: Piscataway, NJ, USA, 1995; Volume 3, pp. 1501–1504. [[CrossRef](#)]
19. Lu, J.; Huang, W.; Zhang, H. Dynamic Prediction of Full-Ocean Depth SSP by a Hierarchical LSTM: An Experimental Result. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 1–5. [[CrossRef](#)]
20. Huang, W.; Lu, J.; Li, S.; Xu, T.; Wang, J.; Zhang, H. Fast Estimation of Full Depth Sound Speed Profile Based on Partial Prior Information. In Proceedings of the 2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT), Qingdao, China, 21–24 July 2023; pp. 479–484. [[CrossRef](#)]
21. Lin, M.; Yang, C. Ocean observation technologies: A review. *Chin. J. Mech. Eng.* **2020**, *33*, 32. [[CrossRef](#)]
22. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
23. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [[CrossRef](#)]
24. Kumar, N.K.; Savitha, R.; Al Mamun, A. Ocean wave characteristics prediction and its load estimation on marine structures: A transfer learning approach. *Mar. Struct.* **2018**, *61*, 202–219. [[CrossRef](#)]
25. Lu, T.; Han, B.; Yu, F. Detection and classification of marine mammal sounds using AlexNet with transfer learning. *Ecol. Inform.* **2021**, *62*, 101277. [[CrossRef](#)]
26. Huang, W.; Li, D.; Zhang, H.; Xu, T.; Yin, F. A meta-deep-learning framework for spatio-temporal underwater SSP inversion. *Front. Mar. Sci.* **2023**, *10*. [[CrossRef](#)]
27. Taroudakis, M.I.; Markaki, M.G. On the use of matched-field processing and hybrid algorithms for vertical slice tomography. *J. Acoust. Soc. Am.* **1997**, *102*, 885–895. [[CrossRef](#)]
28. Yu, Y.; Li, Z.; He, L. Matched-field inversion of sound speed profile in shallow water using a parallel genetic algorithm. *Chin. J. Oceanol. Limnol.* **2010**, *28*, 1080–1085. [[CrossRef](#)]
29. Dai, M.; Li, Y.; Ye, J.; Yang, K. An improved particle filtering technique for source localization and sound speed field inversion in shallow water. *IEEE Access* **2020**, *8*, 177921–177931. [[CrossRef](#)]
30. Chen, C.; Lei, B.; Ma, Y.L.; Duan, R. Investigating sound speed profile assimilation: An experiment in the Philippine Sea. *Ocean Eng.* **2016**, *124*, 135–140. [[CrossRef](#)]
31. Liu, Y.; Chen, Y.; Meng, Z.; Chen, W. Performance of single empirical orthogonal function regression method in global sound speed profile inversion and sound field prediction. *Appl. Ocean Res.* **2023**, *136*, 103598. [[CrossRef](#)]
32. Sun, D.; Yu, M.; Cai, K. Inversion of ocean sound speed profiles from travel time measurements using a ray-gradient-enhanced surrogate model. *Remote Sens. Lett.* **2022**, *13*, 888–897. [[CrossRef](#)]
33. Bianco, M.J.; Gerstoft, P.; Traer, J.; Ozanich, E.; Roch, M.A.; Gannot, S.; Deledalle, C.A. Machine learning in acoustics: Theory and applications. *J. Acoust. Soc. Am.* **2019**, *146*, 3590–3628. [[CrossRef](#)] [[PubMed](#)]
34. Yang, H.; Lee, K.; Choo, Y.; Kim, K. Underwater acoustic research trends with machine learning: Ocean parameter inversion applications. *J. Ocean Eng. Technol.* **2020**, *34*, 371–376. [[CrossRef](#)]
35. Jain, S.; Ali, M. Estimation of Sound Speed Profiles Using Artificial Neural Networks. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 467–470. [[CrossRef](#)]
36. Nowruzi, H.; Ghassemi, H. Using artificial neural network to predict velocity of sound in liquid water as a function of ambient temperature, electrical and magnetic fields. *J. Ocean Eng. Sci.* **2016**, *1*, 203–211. [[CrossRef](#)]
37. Huang, W.; Liu, M.; Li, D.; Yin, F.; Chen, H.; Zhou, J.; Xu, H. Collaborating Ray Tracing and AI Model for AUV-Assisted 3-D Underwater Sound-Speed Inversion. *IEEE J. Ocean. Eng.* **2021**, *46*, 1372–1390. [[CrossRef](#)]
38. Yu, X.; Xu, T.; Wang, J. Sound Velocity Profile Prediction Method Based on RBF Neural Network. In Proceedings of the China Satellite Navigation Conference (CSNC) 2020 Proceedings, Chengdu, China, 22–25 November 2020; Volume III, pp. 475–487.
39. Bianco, M.; Gerstoft, P. Dictionary learning of sound speed profiles. *J. Acoust. Soc. Am.* **2017**, *140*, 1749–1758. [[CrossRef](#)]
40. Hua, X.; Cheng, L.; Zhang, T.; Li, J. Interpretable deep dictionary learning for sound speed profiles with uncertainties. *J. Acoust. Soc. Am.* **2023**, *153*, 877–894. [[CrossRef](#)]
41. Li, H.; Qu, K.; Zhou, J. Reconstructing sound speed profile from remote sensing data: Nonlinear inversion based on self-organizing map. *IEEE Access* **2021**, *9*, 109754–109762. [[CrossRef](#)]
42. Ou, Z.; Qu, K.; Liu, C. Estimation of sound speed profiles using a random forest model with satellite surface observations. *Shock Vib.* **2022**, *2022*, 2653791. [[CrossRef](#)]
43. Ou, Z.; Qu, K.; Shi, M.; Wang, Y.; Zhou, J. Estimation of sound speed profiles based on remote sensing parameters using a scalable end-to-end tree boosting model. *Front. Mar. Sci.* **2022**, *9*, 1051820. [[CrossRef](#)]
44. Santos, M.S.; Soares, J.P.; Abreu, P.H.; Araujo, H.; Santos, J. Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [research frontier]. *IEEE Comput. IntelligeNce MagaziNe* **2018**, *13*, 59–76. [[CrossRef](#)]
45. Ying, X. An overview of overfitting and its solutions. In *Proceedings of the Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2019; Volume 1168, p. 022022.
46. Jin, G.; Liu, F.; Wu, K.; Chen, C. Deep Learning-based framework for expansion, recognition and classification of underwater acoustic signal. *J. Exp. Theor. Artif. Intell.* **2020**, *32*, 205–218. [[CrossRef](#)]
47. Moradi, R.; Berangi, R.; Minaei, B. A survey of regularization strategies for deep models. *Artif. Intell. Rev.* **2020**, *53*, 3947–3986. [[CrossRef](#)]
48. Rich, C. Multitask Learning. *Mach. Learn.* **1997**, *28*, 41–75. [[CrossRef](#)]

49. Yang, Y.; Hospedales, T.M. Deep Multi-task Representation Learning: A Tensor Factorisation Approach. In Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017; pp. 1–12. [[CrossRef](#)]
50. Dong, C.; Xu, G.; Han, G.; Bethel, B.J.; Xie, W.; Zhou, S. Recent developments in artificial intelligence in oceanography. *Ocean.-Land-Atmos. Res.* **2022**, *2022*. [[CrossRef](#)]
51. Miao, Y.; Zhang, X.; Li, Y.; Zhang, L.; Zhang, D. Monthly extended ocean predictions based on a convolutional neural network via the transfer learning method. *Front. Mar. Sci.* **2023**, *9*, 1073377. [[CrossRef](#)]
52. Rostami, M.; Kolouri, S.; Eaton, E.; Kim, K. Deep transfer learning for few-shot SAR image classification. *Remote Sens.* **2019**, *11*, 1374. [[CrossRef](#)]
53. Liang, T.; Zhao, Q.; Lv, Q.; Sun, H. A novel wind speed prediction strategy based on Bi-LSTM, MOOFADA and transfer learning for centralized control centers. *Energy* **2021**, *230*, 120904. [[CrossRef](#)]
54. Chen, J.; Sun, J.; Li, Y.; Hou, C. Object detection in remote sensing images based on deep transfer learning. *Multimed. Tools Appl.* **2022**, *81*, 12093–12109. [[CrossRef](#)]
55. Bejani, M.M.; Ghatee, M. A systematic review on overfitting control in shallow and deep neural networks. *Artif. Intell. Rev.* **2021**, *54*, 6391–6438. [[CrossRef](#)]
56. Yin, F.; Pan, L.; Chen, T.; Theodoridis, S.; Luo, Z.Q.T. Linear Multiple Low-Rank Kernel Based Stationary Gaussian Processes Regression for Time Series. *IEEE Trans. Signal Process.* **2020**, *68*, 5260–5275. [[CrossRef](#)]
57. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, Long Beach, CA, USA, 4–9 December 2017.
58. Zhang, C.; Liu, Z.; Lu, S.; Li, H.; Sun, C.; Wu, X. *User Manual of GDCSM Argo Gridded Data Set*, 3rd ed.; China Argo Real-time Data Center: Hangzhou, China, 2021; p. 19.
59. Liu, F.; Ji, T.; Zhang, Q. Sound Speed Profile Inversion Based on Mode Signal and Polynomial Fitting. *Acta Armamentarii* **2019**, *40*, 2283–2295.

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