

WANG, X., WANG, L., ZHANG, Z., CHEN, K., JIN, Y., YAN, Y. and LIU, J. 2022. Sparse data-extended fusion method for sea surface temperature prediction on the East China Sea. *Applied sciences* [online], 12(12); intelligent computing and remote sensing, article 5905. Available from: <https://doi.org/10.3390/app12125905>

Sparse data-extended fusion method for sea surface temperature prediction on the East China Sea.

WANG, X., WANG, L., ZHANG, Z., CHEN, K., JIN, Y., YAN, Y. and LIU, J.

2022

© 2022 by the authors. Licensee MDPI, Basel, Switzerland.

Article

Sparse Data-Extended Fusion Method for Sea Surface Temperature Prediction on the East China Sea

Xiaoliang Wang¹, Lei Wang^{1,*}, Zhiwei Zhang¹, Kuo Chen¹, Yingying Jin¹, Yijun Yan²  and Jingjing Liu³ 

¹ East Sea Information Center, State Oceanic Administration, Shanghai 200136, China; wangxiaoliang@ecs.mnr.gov.cn (X.W.); zzw@ecs.mnr.gov.cn (Z.Z.); chenkuo@ecs.mnr.gov.cn (K.C.); jinyingying@ecs.mnr.gov.cn (Y.J.)

² National Subsea Centre, Robert Gordon University, Aberdeen AB10 7AQ, UK; y.yan2@rgu.ac.uk

³ College of Information Science & Technology, Zhejiang Shuren University, Hangzhou 310015, China; liujingjing@zjsru.edu.cn

* Correspondence: wangleidett727@163.com

Abstract: The accurate temperature background field plays a vital role in the numerical prediction of sea surface temperature (SST). At present, the SST background field is mainly derived from multi-source data fusion, including satellite SST data and in situ data from marine stations, buoys, and voluntary observing ships. The characteristics of satellite SST data are wide coverage but low accuracy, whereas the in situ data have high accuracy but sparse distribution. For obtaining a more accurate temperature background field and realizing the fusion of measured data with satellite data as much as possible, we propose a sparse data-extended fusion method to predict SST in this paper. By using this method, the actual observed sites and buoys data in the East China Sea area are fused with Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Version 5.0 SST data. Furthermore, the temperature field in the study area were predicted by using Long Short-Term Memory (LSTM) and Gate Recurrent Unit (GRU) deep learning methods, respectively. Finally, we obtained the results by traditional prediction methods to verify them. The experimental results show that the method we proposed in this paper can obtain more accurate prediction results, and effectively compensate for the uncertainty caused by the parameterization of ocean dynamic process, the discrete method, and the error of initial conditions.

Keywords: deep learning; association relationship; heterogeneous clustering; extended fusion



Citation: Wang, X.; Wang, L.; Zhang, Z.; Chen, K.; Jin, Y.; Yan, Y.; Liu, J.

Sparse Data-Extended Fusion Method for Sea Surface Temperature Prediction on the East China Sea.

Appl. Sci. **2022**, *12*, 5905.

<https://doi.org/10.3390/app12125905>

Academic Editors: Feng Gao, Jin Zheng and Qizhi Xu

Received: 19 April 2022

Accepted: 6 June 2022

Published: 10 June 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sea surface temperature (SST) is the basic tool for understanding local and global climate change. The study of SST changes in marginal seas under the background of global warming has always been an important topic in regional oceanography and climatology [1].

Variability of SST will cause weather phenomenon changes, such as tropical monsoons and cyclones. The inseparable connection between the variability of Indian Ocean SST and the Indian summer monsoon has long been known [2]. SST variability has a great impact on the habits of marine organisms. For instance, predicting fishing areas is based on the potential law of the SST variability [3]. Moreover, marine ecosystems are affected by SST variability. For example, due to climate warming, coral reefs have a lower growth rate in the west, and the strong positive Indian Ocean dipole mode events (IOD) also caused catastrophic damage to the coral reefs of the Mentawi Islands [4]. Therefore, predicting SST in advance can not only protect related resources, but also prevent major losses caused by certain marine disasters.

SST is mainly obtained from satellite remote sensing and in situ observation including marine stations, buoys, and voluntary observing ships et al. Satellite remote sensing data has the advantages of all-weather and high spatio-temporal resolution, which has gradually become an indispensable means of marine research. However, the accuracy of satellite

remote sensing inversion is relatively low. Data from in situ observation including ocean stations, buoys [5], and volunteer ships [6] have high accuracy but sparse and uneven distribution. As a result, it is difficult to obtain the continuous, accurate and uniform SST and finally the prediction accuracy of SST is much lower in the wide range. Therefore, how to effectively integrate the in situ observational data with remote sensing data becomes a big challenge in marine environmental monitoring and prediction. Furthermore, the problems of sparse and uneven distribution of in-situ observation data and low accuracy of satellite data are also troublesome.

At present, the main methods for fusing observation data and remote sensing data to become standardized grid data products include the interpolation method [7], the variational method [8], the Kalman filter method [9], and the deep learning method [10]. Among the interpolation methods, optimal interpolation (OI) is the most classic one. Reynolds et al. [11,12] applied OI into the fusion analysis of satellite SST with field observation data, and finally produced average SSTs of monthly, weekly, and daily.

One disadvantage of the OI method is that the background error covariance matrix is static, and the advantage of the variational method over OI is that the variational method [13] introduces a nonlinear observation operator, which avoids the limitation of the linear relationship between observed quantity and analytical quantity in the OI method [14]. The variational method is a global optimization whereas OI is a local optimization. In addition, the variational method can also introduce some dynamic constraints. Variational methods mainly include three-dimensional variation (3DVAR) and four-dimensional variation (4DVAR). 4DVAR is an improvement from 3DVAR. Due to the limitations of the existing optimization algorithm, its convergence is reduced due to the introduction of mode constraints, and because of the roughness of the initial guess field, it doesn't converge to a global minimum, but converges to a local extreme value, which seriously affects the expected effect of the 4DVAR method.

The Kalman filter (KF) was originally proposed by Kalman [15]. Compared with the variational method, the advantage of KF is that there is no need to write the accompanying pattern, and the disadvantage is that it needs to store a high-dimensional background error covariance matrix and requires an accurate estimation of the model error.

In recent years, the rapid development of artificial intelligence has brought new ideas for ocean data fusion. Machine learning is an important method to realize artificial intelligence. It is a method of using algorithms to analyze and learn from data, and then make decisions and predictions about phenomena in the real world. As a technology for realizing machine learning, deep learning has the advantage of proposing a method for the computer to automatically learn pattern features, and integrating feature learning into the model building process. It reduces the incompleteness caused by man-made design features and has more advantages compared with OI, 4DVAR, and KF, conceptually. At present, machine learning has a very wide range of applications [16–18], such as language recognition, image recognition, data mining, and expert systems, etc. Some machine learning algorithms are used in the assimilation of ocean remote sensing data, such as genetic algorithms [19] and neural networks [20]. These data fusion methods have shown certain advantages in different aspects. However, they only assimilate data of type and homology. In addition, the past statistical methods rely more on regression analysis for spatial modeling, and the model input parameters are relatively single.

To tackle the challenges and issues mentioned above, this paper proposed a new framework for SST prediction, where the gate recurrent unit (GRU) [21] is employed to classify data according to the characteristics of remote sensing data. Then the sparse measurement point observation data in the same category is used to fuse with the large-scale remote sensing data, resulting a temperature background field with promising accuracy of prediction.

The rest of this paper is organized as follows: Relevant definitions involved in this work are introduced in Section 2. The whole process of proposed framework for SST prediction, including: (1) the SST correlation calculation method based on deep learning,

(2) the heterogeneous clustering method, (3) the sparse data expansion and fusion method, and (4) the SST prediction method based on deep learning, is described in detail in Section 3. Section 4 presents the experimental results and discussions. Finally, conclusions and future prospects are summarized in Section 5.

2. Problem Description and Definition

For the observed SST data, the time series data is collected from in-site measurement points. For the satellite SST data, it provides a spatially complete SST dataset, and each grid point has a continuously long time series of SSTs.

First, we obtain the characteristics of the time series on each grid point of satellite SST data, and calculate the association relationship of each grid point. Secondly, the grid points are clustered by taking the association relationship as the features and using the location of each in-site measurement point as the cluster center. Thirdly, based on the clustering results, the large-scale remote sensing data are expanded and fused. Finally, the SSTs are predicted based on the fusion results.

The specific process is shown in Figure 1 below.

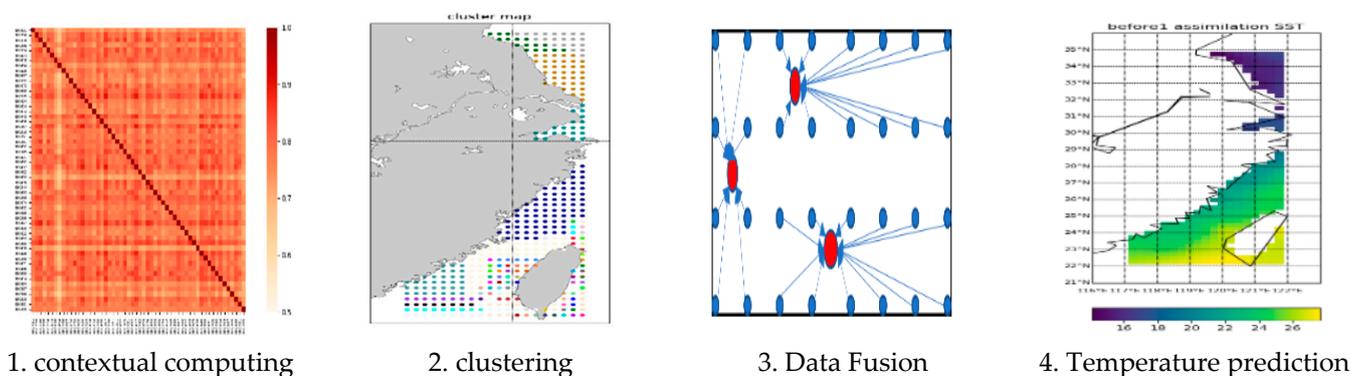


Figure 1. Data processing.

Large-scale remote sensing data and sparse observation data in this paper are defined as spatio-temporal series.

Definition of Spatio-temporal Series: Let $D = \{X^1, X^2, \dots, X^n\}$ be spatio-temporal series data, where $X^i = \{x_1^i, x_2^i, \dots, x_n^i\}$ is a piece of time-series data. Define the position information of each time data of i in X^i , and x_j^i ($1 < j < n$) is the data at the j th time point of the i th time series, and n means the number of grid points.

Definition of Fuzzy Correlation: This is to analyze the relationship between the two points from the perspective of time and space. Each time series data point has multiple samples, and each sample can be divided into different categories. The relationship between the samples can be defined by the proportional relationship between the categories.

Definition of Associated Metric: An associated metric m takes as input a tuple $(\times 1, \times 2)$, and produces an output r ; the larger r , the higher the association.

In previous studies, the clustering methods mainly focused on exploring the optimal clustering center on the satellite data. In this paper, we propose a fusion-based clustering method, wherein the cluster centers are determined by in-site measurement, and each grid point from satellite data will be aligned with certain cluster centers based on the association.

Definition of Heterogeneous Clustering: Given a time series correlation matrix R of remote sensing data, the required number of clusters k , which is also the number of observation points, and the clustering results are obtained by Cluster (k, R) method.

Definition of Expansion Fusion: After clustering, the sparse and uneven distributed observation data can be expanded to cover the global ocean, where the SST value from the satellite on each grid point will be replaced by its associated cluster center. Better clustering performance will lead to higher expansion ability.

3. Implementation of the Proposed Approach

The extended fusion method to predict SST for sparse data proposed in this paper is mainly divided into four stages. The first stage is data association calculation. This stage mainly uses the time series classification method of deep learning to classify different data points, and then performs association calculations on the classification results between different data points. The second stage is the cluster analysis of the data. The result of the association calculation is used as the distance between the data points, and then K-means clustering algorithm is used to cluster the data points, where the data center point is a heterogeneous data point. The third stage is to continue the assimilation calculation of the data through the clustering results. Each type of data is assimilated based on the clustering center and the association relationship. The fourth stage is the data prediction. By using the assimilated data for data training, the pros and cons of the assimilation methods are judged by comparing the data prediction effects before and after the assimilation. Detailed descriptions of the implementation are discussed as follows.

3.1. Calculation Method of Association Relationship Based on Deep Learning

This part calculates the data association relationship between different grids of remote sensing data. Neural networks have the ability to mine and discover the relationship between data. By using neural network algorithms, we can train large amounts of data and build relationships between data. Therefore, it has certain advantages in the calculation of the data association relationship between different grids of large-area remote sensing. In this paper, we use deep learning classification algorithms to classify spatiotemporal series data, determine the degree of association between data points, and use neural networks to construct data relationships.

Suppose the spatio-temporal series data is $D = \{X^1, X^2, \dots, X^n\}$, where $X^i = \{x_1^i, x_2^i, \dots, x_n^i\}$ is a piece of time series data, and define the category of each data sample of X^i as the i th type, and x_j^i is the i th data sample of the j th time series. By using the neural network algorithm for data training, $f(\cdot)$ can be calculated. Through data prediction of different data, the prediction category of each data sample can be obtained. The category sequence for the time series X^i data prediction can be defined as $C^i = \{c_1, c_2, \dots, c_n\}$, for different times. Sequence data points can be calculated by calculating the degree of penetration of other categories to this category. Therefore, the calculation method of data association is as follows:

$$r_i^j = \sum_{i=1}^n [c_i = c_j] / n \tag{1}$$

where n is the number of samples of data point X^i , and r_i^j represents the relationship between data point X^j and data point X^i .

The neural network used in this paper is the Gated Recurrent Unit (GRU) network. The GRU network mainly includes the reset gate as Equation (2) and the update gate as Equation (3). The update gate is used to control the extent to which the state information at the previous moment is brought into the current state. The larger the value of the update gate, the more state information from the previous moment is brought in. The reset gate is used to control the degree of ignoring the state information at the previous moment. A smaller reset gate will produce a higher degree of ignoring.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{2}$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{3}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t]) \tag{4}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{5}$$

$$y_t = \sigma(W_o \cdot h_t) \tag{6}$$

In the above formula, Z_t and r_t represent the update gate and reset gate respectively. In essence, it is the same as Long-Short Term Memory (LSTM). The output h_{t-1} of the previous moment $t-1$ is combined with the input x_t of the current moment to calculate various attenuation coefficients. The slight difference is that the linear transformation does not use bias. Because the memory state is also h_{t-1} , it is directly updated. The final output network result h_t is as shown in Equation (5). This result is also the memory state of the network. The output result at each step is y_t is as Formula (6).

We use a two-layer GRU network plus a fully connected layer network structure for network training in this paper; the network structure is show in Figure 2 below.

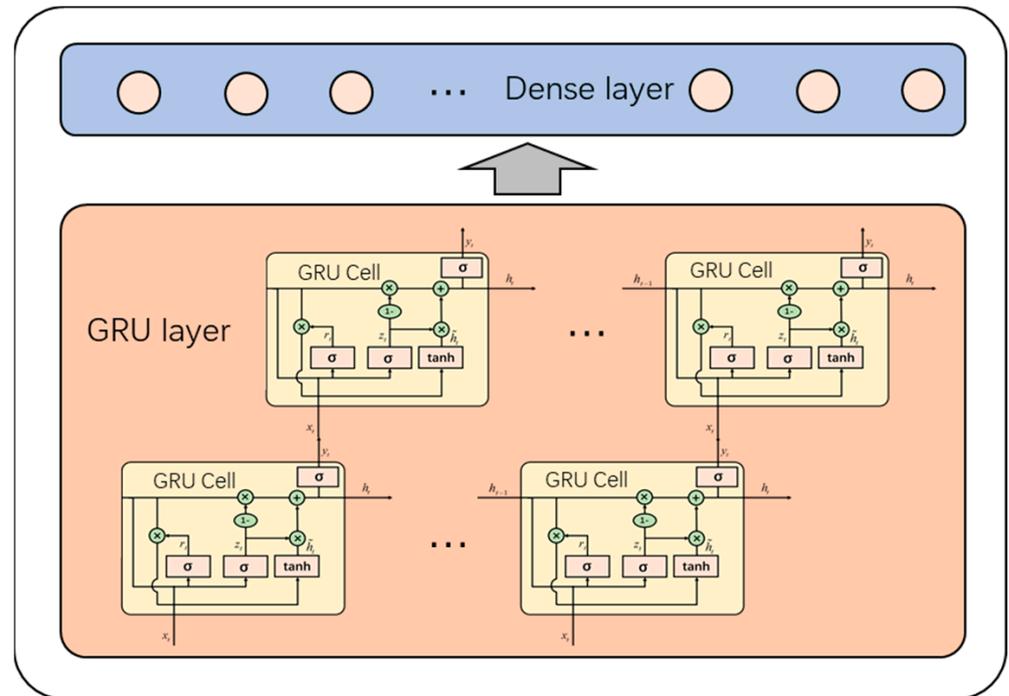


Figure 2. The network structure.

The output result of the first layer of GRUCell is used as the input of the second layer of GRUCell network, and the output of the last GRUCell is connected to the fully connected layer, and the classification probability of each point can be directly obtained. Among them, the two-layer GRU network can receive and process more data information, and the fully connected layer can map data hiding information to each data category.

3.2. Heterogeneous Clustering Method Based on Association

In order to explore the situation that observation data and remote sensing data have the same relationship in time and space, this section conducts cluster analysis on data that of the same type but from different sources. In this section, different sources of data were clustered based on the relational matrix between data which were mined based on neural network relationship and K-means method. The relational matrix can indicate the internal relationship between data and furthermore the distance between each grid point. Suppose that A comes from relatively accurate but sparse observational data, and B comes from relatively inaccurate but widely covered remote sensing data. According to the characteristics of the two types of data, use the data source of A as the clustering center of the data source of B, and then cluster the data source of B to prepare for the next step of data assimilation.

Assuming that the source data of A filters out some or all of the data $M = \{m_1, m_2, \dots, m_n\}$ as the original center point, then find the data point closest to these points in the B data source, as the center point $M' = \{m'_1, m'_2, \dots, m'_n\}$ of the B set, then cluster all the data in set B according

to the K-means clustering method. Among them the distance between the data is the data relationship calculated in the previous step, and cluster sets can be obtained.

3.3. Sparse Data Expansion Fusion Method

In order to obtain more accurate background field data so that the large-scale remote sensing data in each area can be fused and corrected with more accurate observation data, the fusion method in this section is to fuse the large-scale remote sensing data through accurate sparse data. The principle in the process of fusion is that the stronger the correlation, the stronger the degree of fusion of remote sensing data by measured data, and vice versa.

Because there are fewer precise data and more imprecise data, the fusion method is used to fuse precise data into remote sensing data. This is to achieve the purpose of calibrating fuzzy data and increasing its data volume.

The measured data is the temperature data measured at a specific location in the East China Sea, which is point data, whereas the remote sensing data is the temperature data inversion of the entire East China Sea, which is grid data. The measured data is located in some grid points of the remote sensing data. Using the measured temperature data in these grid points to assimilate the remote sensing data in the same area will make the assimilated data not only contain multi-source data, but also more accurate.

In the previous section, data clustering was completed, and n clustering sets were obtained. Set B with remote sensing data was fused and calibrated by set A with observation data.

Because set A and set B are both time series data, assuming that the central point data of a cluster set is X^i and the data to be fused is X^j , the fusion calculation formula is as follows:

$$X^j = r_i^j X^i + (1 - r_i^j) X^j. \tag{7}$$

Therefore, the assimilated data can be obtained, Figure 3 is a schematic diagram of the assimilation method of Equation (7).

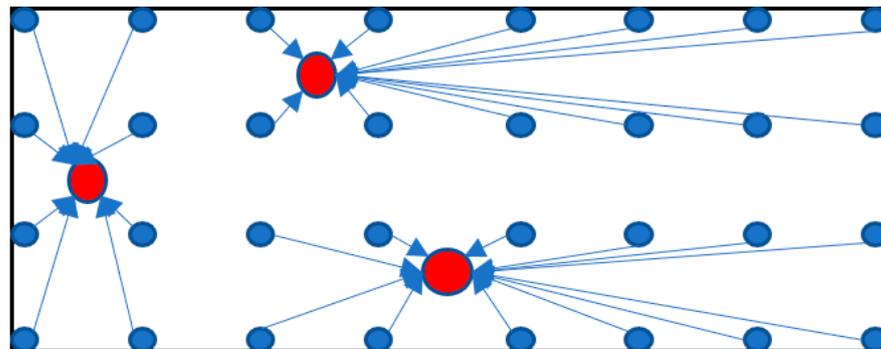


Figure 3. The red circle represents the measured data, the blue circle represents the remote sensing data, and the connected edges represent the fusion calculation of the remote sensing data with the measured data with the same cluster.

3.4. Prediction and Inspection Method Based on Radiation Fusion

After assimilating, more data can be obtained. By using neural network for data training, it can better fit the real changes of the data.

Suppose the original data is D , and the data after assimilation is D' . Divide the original data and the assimilated data into a training set and a test set according to the same time and space. The original data are divided into $trainX$, $trainY$, $testX$ and $testY$, the assimilated data are divide into $trainX'$, $trainY'$, $testX'$, and $testY'$.

By training the neural network model on the original dataset we have the following:

$$D \xrightarrow{trainX, trainY} f(D).$$

The data training network model after assimilation is

$$D' \xrightarrow{\text{train}X' \text{train}Y'} f'(D').$$

After the training is completed, the assimilated data is tested through two models, so the absolute error value can be obtained by using the original data model.

$$MAE = f(\text{test}X', \text{train}Y')$$

After assimilation, the data model can be obtained as follows:

$$MAE' = f'(\text{test}X', \text{train}Y').$$

By comparing MAE and MAE' , the effect of the assimilation model and the prediction model can be seen.

The root mean square error ($RMSE$) and mean absolute error (MAE) are mainly used as the criteria for evaluating the quality of the model. The root mean square error can measure the deviation between the observed value and the true value; the root mean square error represents the average value of the absolute error. The formula is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i^p - X_i^r)^2} \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i^p - X_i^r|. \quad (9)$$

4. Experimental Results and Analysis

The experimental data of this paper come from two parts, The first part is satellite SST data, which comes from National Oceanic and Atmospheric Administration (NOAA) high-resolution ocean surface temperature data, in which the resolution is 0.25° . SST data include daily data from 1 January 2016 to 31 December 2016. The range is from 22.125° N to 34.875° N, 117.125° E to 121.875° E, basically covering the East China Sea. The second part is the actual measured data of 55 location points in the East China Sea. The location points are within the range of the first part of the data, and the time is the same as that of the first part of the data.

Experiment related parameters: the number of hidden layer units in the first layer is 256, the number of hidden layer units in the second layer is 64, the Adam optimizer is used for learning, the learning rate is 0.0001, and the cross-entropy loss function is used. In the experiment, the window and the step size are 100 and 30, respectively, and the fully connected output layer is 373 (consisting of 373 points), and the probability of each point is output.

According to the correlation matrix, the data is classified into two types, and the two types of data are assimilated according to the corresponding remote sensing data.

The ultimate goal of all algorithms is to extend the algorithm to new data. If you train the model based on all the data, you will not be able to measure the effect of the model on the new data. In order to solve this problem, the data is usually divided into two parts: training set and test set. Divide the data into 70% training set and 30% test set. The training set refers to the dataset that has been assimilated by remote sensing data and site data, and the test set is the dataset that has not been assimilated by remote sensing data and site data.

It can be seen from the Figure 4 below that the MAE basically no longer changes when the number of iterations reaches 300, which means that the model has basically reached the fit.

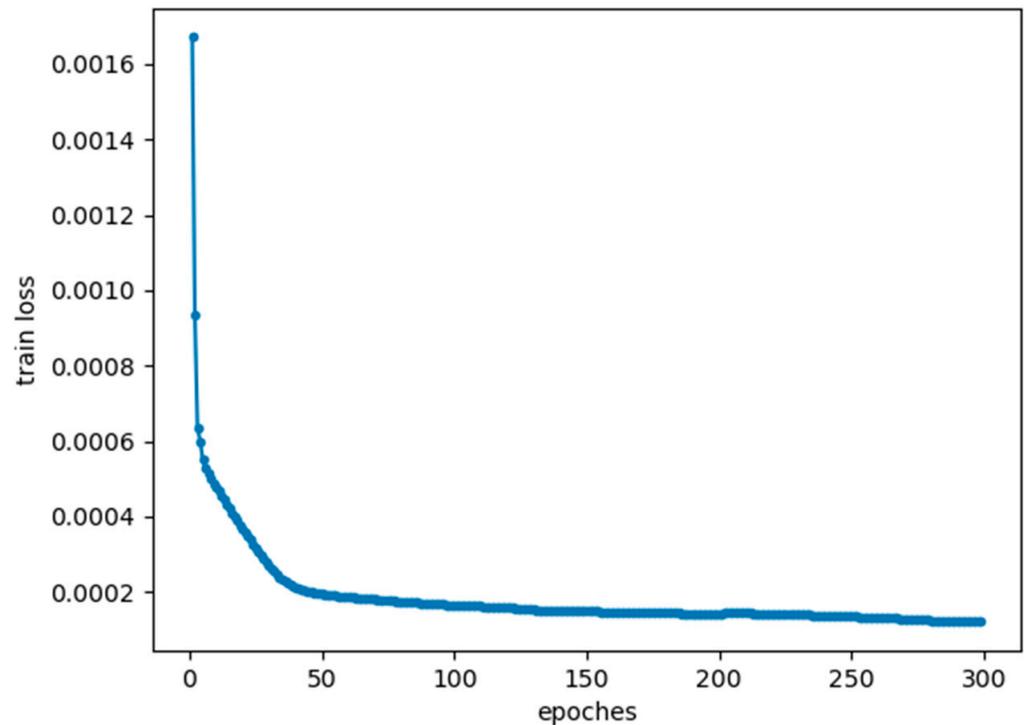


Figure 4. Relationship between iteration times and loss function.

Figure 5 shows the satellite SST data of the front area before assimilation, and Figure 6 shows the assimilated satellite SST data used by Formula (7).

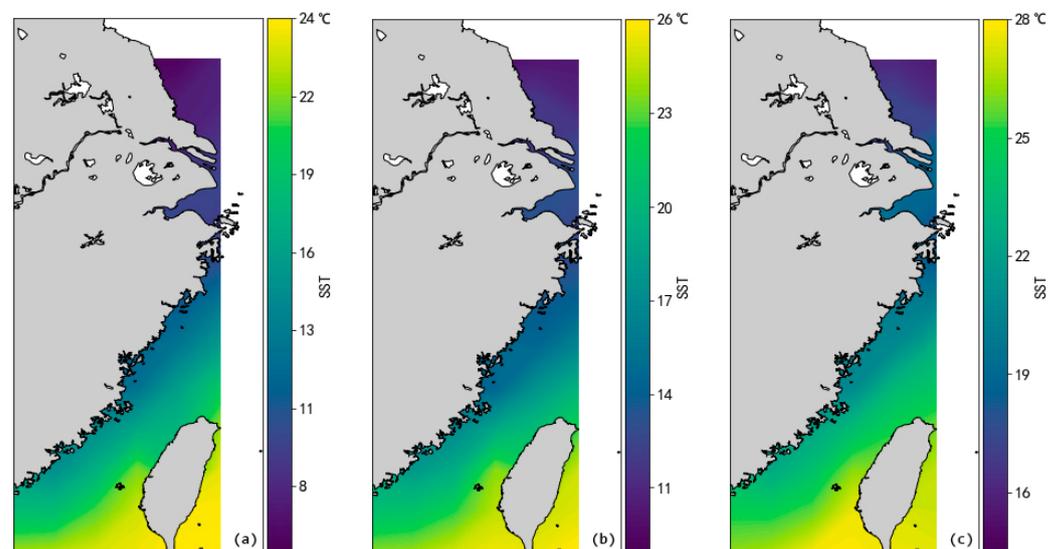


Figure 5. Average temperature in March, April, and May before assimilation. (a) Average temperature in March. (b) Average temperature in April. (c) Average temperature in May.

It can be seen that the data before and after the assimilation is basically consistent as a whole, so the stability of the data is well maintained, but there are still slight differences in some areas.

In order to verify the effectiveness of our method and the accuracy of the prediction results, this paper uses a comparative experiment method to verify the prediction results before and after assimilation. We use baseline methods for comparison: Gradient Boosting Decision Tree (GBDT) [22], Random Forest (RF) [23], and the Kalman Filter (KF) method. The methods are as follows:

GBDT: Apply the gradient descent tree to the regression problem to make predictions. Random Forest: Apply random forest to regression problems and make predictions. The experimental results are shown in the Table 1 below.

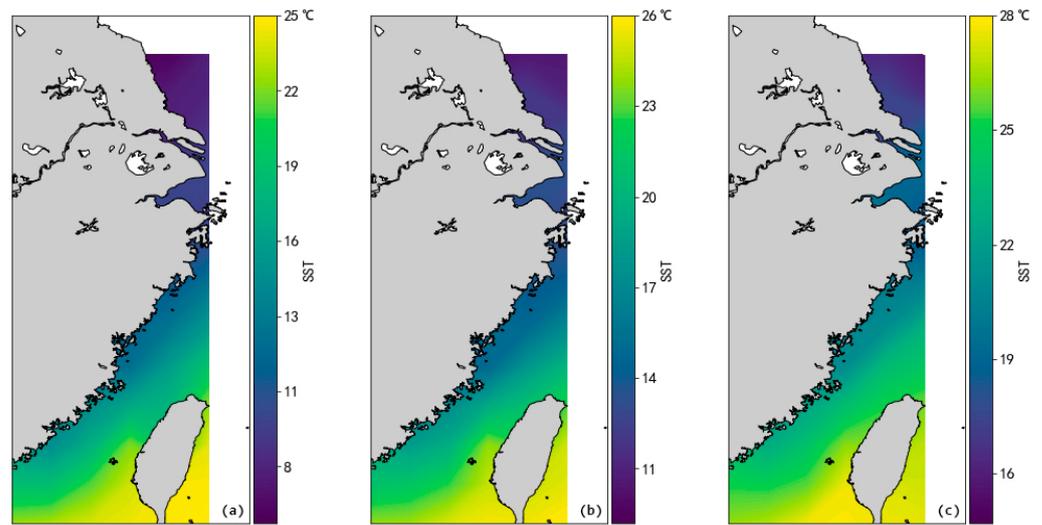


Figure 6. Average temperature in March, April, and May after assimilation. (a) Average temperature in March. (b) Average temperature in April. (c) Average temperature in May.

Table 1. Comparison of experimental results of different methods.

Time	1 Day		3 Days		7 Days		14 Days	
Method	mae	rmse	mae	rmse	mae	rmse	mae	rmse
GBDT	0.852513307	1.23915234	1.300753798	1.720802339	1.611730345	2.147783524	2.186371979	2.778341512
RF	2.288058902	5.90082053	3.597995729	7.272703867	4.777261762	8.969278431	5.374765925	9.815889435
KF	0.389958846	0.618940732	0.646802982	1.080479235	0.900576924	1.46066052	1.260270091	1.842596171
Our method	0.373859393	0.542795005	0.57908497	0.866673458	0.766950071	1.196595851	1.248104149	1.819827343

It can be seen from the table that our method is better than the baseline method in an all-around way, and as the number of prediction days increases, the effect of our method is more obvious. In specific analysis, the effect of the RF method in the ocean sea temperature prediction method is relatively general, and the GBDT algorithm has a good effect when predicting one day, but as the number of prediction days increases, the effect becomes worse. In particular, the prediction results in the 14-day forecast, which proved the effectiveness of the assimilation method proposed in this paper.

In order to further verify the effectiveness of the idea of this paper, we added another set of experimental comparisons, the experimental data we selected are the same as the above experiments. Before data assimilation, the clustering method we choose comes from the complex network [24], the assimilation method comes from the method proposed in this paper, and the prediction methods are GRU and LSTM [25]. The experimental results of GRU and LSTM are shown in the Tables 2 and 3 below.

Table 2. Temperature prediction results under different prediction aging before assimilation.

GRU	Situation	1 Day	3 Days	7 Days	14 Days
All points	Before assimilation	0.6249	1.1209	1.8523	3.0078
	After assimilation	0.5490	0.9415	1.5366	2.7531

Table 3. Temperature prediction results under different prediction aging after assimilation.

LSTM	Situation	1 Day	3 Days	7 Days	14 Days
All points	Before assimilation	0.8310	1.3930	2.0955	3.1397
	After assimilation	0.7473	1.1665	1.7176	2.7794

We found that by calculating the average error of different methods, when the data are fused, regardless of the prediction time limit of 1, 3, 7, and 14 days, the results of temperature prediction are much better than those before fusion, which shows that the method proposed in this paper is more effective.

By calculating the average error of different methods, we find that based on the above analysis, considering that the satellite remote sensing data has a strong correlation locally, we cluster according to this correlation, then use the measured data to assimilate the remote sensing data in the cluster, and use the assimilation method proposed in this paper to assimilate the temperature data. Finally, the accuracy of temperature prediction results can be greatly improved.

5. Conclusions

Considering that the whole sea surface temperature has a certain correlation in local areas, this paper first clusters the sea surface temperature data retrieved by satellite remote sensing, then judges which clusters these stations are distributed among according to the location of the observation data, and finally uses the data-assimilation method proposed in this paper to assimilate the remote sensing data with the observation data to obtain the background data with high accuracy. Thus, high prediction accuracy is obtained.

- (1) Through the calculation of the correlation relationship between the sea surface temperature retrieved from the remote sensing data, the temperatures with the same law are correlated. It provides a more scientific and accurate way to assimilate the sea surface temperature with the same law for the measured data.
- (2) The method proposed in this paper has the characteristics of fast convergence speed and stable training process. It can be seen from Figure 4 that the training error of the model tends to converge after 300 iterations.
- (3) By comparing the calculation results of different methods, the method proposed in this paper shows the best results before and after data assimilation. Especially in the prediction of 3, 7, and 14 days, the method proposed in this paper reduces the error of rmae by about 0.3 compared with other methods.

Author Contributions: Data curation, Z.Z.; Formal analysis, K.C.; Investigation, Y.J.; Methodology, L.W.; Validation, J.L.; Writing—original draft, X.W.; Writing—review & editing, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [Tianjin enterprise postdoctoral innovation project merit funding project] grant number [TJQYBSH2018025], Marine Telemetry Technology Innovation Center of the Ministry of Natural Resources, Key Laboratory of Marine Environmental Survey Technology and Application, MNR, [Science and Technology Department of Zhejiang Province] grant number [LGG21F020008].

Institutional Review Board Statement: This study did not require ethical approval, so we choose to exclude this statement.

Informed Consent Statement: This study did not involve humans, so we also choose to exclude this statement.

Data Availability Statement: Data sharing not applicable.

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

References

1. Li, Y.; Wang, G.; Fan, W.; Kexiu, L.; Hui, W.; Tinz, B.; Storch, H.; Jianlong, F. The homogeneity study of the sea surface temperature data along the coast of the China Seas. *Acta Oceanol. Sin.* **2018**, *40*, 17–28. [\[CrossRef\]](#)
2. Vibhute, A.; Halder, S.; Singh, P.; Parekh, A.; Chowdary, J.S.; Gnanaseelan, C. Decadal variability of tropical Indian Ocean sea surface temperature and its impact on the Indian summer monsoon. *Theor. Appl. Climatol.* **2020**, *141*, 551–566. [\[CrossRef\]](#)
3. Abdul Azeez, P.; Raman, M.; Rohit, P.; Shenoy, L.; Jaiswar, A.K.; Mohammed Koya, K.; Damodaran, D. Predicting potential fishing grounds of ribbonfish (*Trichiurus lepturus*) in the north-eastern Arabian Sea, using remote sensing data. *Int. J. Remote Sens.* **2020**, *42*, 322–342. [\[CrossRef\]](#)
4. Kleypas, J. Climate change and tropical marine ecosystems: A review with an emphasis on coral reefs. *UNED Res. J.* **2019**, *11*, 24. [\[CrossRef\]](#)
5. Murakami, H.; Kawamura, H. Relations between Sea Surface Temperature and Air-Sea Heat Flux at Periods from 1 Day to 1 Year Observed at Ocean Buoy Stations around Japan. *J. Oceanogr.* **2001**, *57*, 565–580. [\[CrossRef\]](#)
6. Rossby, T.; Siedler, G.; Zenk, W. The Volunteer Observing Ship and Future Ocean Monitoring. *Bull. Am. Meteorol. Soc.* **1995**, *76*, 5–11. [\[CrossRef\]](#)
7. Crochiere, R.E.; Rabiner, L.R. *Multirate Digital Signal Processing*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1983.
8. Tong, C.C.; Jung, Y.; Xue, M.; Liu, C. Direct Assimilation of Radar Data with Ensemble Kalman Filter and Hybrid Ensemble-Variational Method in the National Weather Service Operational Data Assimilation System GSI for the Stand-Alone Regional FV3 Model at a Convection-Allowing Resolution. *Geophys. Res. Lett.* **2020**, *47*, e2020GL090179. [\[CrossRef\]](#)
9. Suma, K.; Kawahara, M. Estimation of boundary conditions for ground temperature control using Kalman filter and finite element method. *Int. J. Numer. Methods Fluids* **2015**, *31*, 261–274. [\[CrossRef\]](#)
10. Deng, L.; Yu, D. Deep Learning: Methods and Applications. *Found. Trends Signal Processing* **2014**, *7*, 197–387. [\[CrossRef\]](#)
11. Reynolds, R.W.; Smith, T.M. Improved Global Sea Surface Temperature Analyses Using Optimum Interpolation. *J. Clim.* **1994**, *7*, 929–948. [\[CrossRef\]](#)
12. Reynolds, R.W.; Smith, T.M.; Liu, C.; Chelton, D.B.; Casey, K.S.; Schlax, M.G. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *J. Clim.* **2007**, *20*, 5473–5496. [\[CrossRef\]](#)
13. Guan, Y.H.; Zhou, G.Q.; Lu, W.S.; Chen, J.P. Theory Development and Application of Data Assimilation Methods. *Meteorol. Disaster Reduct. Res.* **2007**, *30*, 938–950.
14. Yao, C.; Ru, W. *A Review of the Adjoint Method to Oceanic Numerical Simulation [EB/OL]*; China Science and Technology Paper Online: Beijing, China, 2006; pp. 771–775.
15. Kalman, R.E. A New Approach to Linear Filter and Prediction Problems. *J. Basic Eng.* **1960**, *82*, 35–45. [\[CrossRef\]](#)
16. Bonissone, P.P. *Machine Learning Applications*; Springer: Berlin/Heidelberg, Germany, 2015.
17. Yan, Y.; Ren, J.; Tschannerl, J.; Zhao, H.; Harrison, B.; Jack, F. Nondestructive Phenolic Compounds Measurement and Origin Discrimination of Peated Barley Malt Using Near-Infrared Hyperspectral Imagery and Machine Learning. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5010715. [\[CrossRef\]](#)
18. Yan, Y.; Ren, J.; Liu, Q.; Zhao, H.; Sun, H.; Zabalza, J. PCA-domain fused singular spectral analysis for fast and noise-robust spectral-spatial feature mining in hyperspectral classification. *IEEE Geosci. Remote Sens. Lett.* **2021**, *99*, 1. [\[CrossRef\]](#)
19. Leardi, R.; Boggia, R.; Terrile, M. Genetic algorithms as a strategy for feature selection. *J. Chemom.* **1992**, *6*, 267–281. [\[CrossRef\]](#)
20. Montavon, G.; Samek, W.; Müller, K.R. Methods for Interpreting and Understanding Deep Neural Networks. *Digit. Signal Processing* **2018**, *73*, 1–15. [\[CrossRef\]](#)
21. Disha, R.A.; Waheed, S. Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique. *Cybersecurity* **2022**, *5*, 1. [\[CrossRef\]](#)
22. Ying, Z.; Xu, Z.; Wang, W.; Meng, C. MT-GBM: A Multi-Task Gradient Boosting Machine with Shared Decision Trees. *arXiv* **2022**, arXiv:2201.06239.
23. Yuan, J.; Shi, M.; Wang, Z.; Liu, H.; Li, J. Random pairwise shapelets forest: An effective classifier for time series. *Knowl. Inf. Syst.* **2022**, *64*, 143–174. [\[CrossRef\]](#)
24. Wang, L.; Huang, Z.; Shi, S.; Chen, K.; Xu, L.; Zhang, G. Marine Multiple Time Series Relevance Discovery Based on Complex Network. In Proceedings of the International Conference on Neural Information Processing, Siem Reap, Cambodia, 13–16 December 2018; Springer: Cham, Switzerland, 2018.
25. Shi, S.; Wang, L.; Yu, X.; Xu, L. Application of long term and short term memory neural network in prediction of chlorophyll a concentration. *Haiyang Xuebao* **2020**, *42*, 134–142.