



Article

Prediction of Sea Surface Temperature Using U-Net Based Model

Jing Ren ^{1,†} , Changying Wang ^{1,†}, Ling Sun ^{2,3,4,*}, Baoxiang Huang ¹, Deyu Zhang ¹, Jiadong Mu ¹ and Jianqiang Wu ¹

¹ College of Computer Science and Technology, Qingdao University, Qingdao 266071, China; renjing@qdu.edu.cn (J.R.); wcing@qdu.edu.cn (C.W.); baoxianghuang@qdu.edu.cn (B.H.); zhangdeyu@qdu.edu.cn (D.Z.); mujiadong@qdu.edu.cn (J.M.); wujianqiang@qdu.edu.cn (J.W.)

² National Satellite Meteorological Center (National Centre for Space Weather), Beijing 100081, China

³ Innovation Center for FengYun Meteorological Satellite (FYSIC), Beijing 100081, China

⁴ Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, CMA, Beijing 100081, China

* Correspondence: sunling@cma.gov.cn

† These authors contributed equally to this work.

Abstract: Sea surface temperature (SST) is a key parameter in ocean hydrology. Currently, existing SST prediction methods fail to fully utilize the potential spatial correlation between variables. To address this challenge, we propose a spatiotemporal UNet (ST-UNet) model based on the UNet model. In particular, in the encoding phase of ST-UNet, we use parallel convolution with different kernel sizes to efficiently extract spatial features, and use ConvLSTM to capture temporal features based on the utilization of spatial features. Atrous Spatial Pyramid Pooling (ASPP) module is placed at the bottleneck of the network to further incorporate the multi-scale features, allowing the spatial features to be fully utilized. The final prediction is then generated in the decoding stage using parallel convolution with different kernel sizes similar to the encoding stage. We conducted a series of experiments on the Bohai Sea and Yellow Sea SST data set, as well as the South China Sea SST data set, using SST data from the past 35 days to predict SST data for 1, 3, and 7 days in the future. The model was trained using data spanning from 2010 to 2021, with data from 2022 being utilized to assess the model's predictive performance. The experimental results show that the model proposed in this research paper achieves excellent results at different prediction scales in both sea areas, and the model consistently outperforms other methods. Specifically, in the Bohai Sea and Yellow Sea sea areas, when the prediction scales are 1, 3, and 7 days, the MAE of ST-UNet outperforms the best results of the other three compared models by 17%, 12%, and 2%, and the MSE by 16%, 18%, and 9%, respectively. In the South China Sea, when the prediction ranges are 1, 3, and 7 days, the MAE of ST-UNet is 27%, 18%, and 3% higher than the best of the other three compared models, and the MSE is 46%, 39%, and 16% higher, respectively. Our results highlight the effectiveness of the ST-UNet model in capturing spatial correlations and accurately predicting SST. The proposed model is expected to improve marine hydrographic studies.

Keywords: sea surface temperature (SST); spatiotemporal prediction; deep learning; convolutional neural network; U-Net



Citation: Ren, J.; Wang, C.; Sun, L.; Huang, B.; Zhang, D.; Mu, J.; Wu, J. Prediction of Sea Surface Temperature Using U-Net Based Model. *Remote Sens.* **2024**, *16*, 1205. <https://doi.org/10.3390/rs16071205>

Academic Editors: Chung-Ru Ho and Jorge Vazquez

Received: 27 February 2024

Revised: 17 March 2024

Accepted: 26 March 2024

Published: 29 March 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

Sea surface temperature (SST) is an integrated result of solar radiation, ocean thermal and dynamical processes, and air–sea interactions, and is an important physical parameter for the study and understanding of the oceans [1–8]. Studies have shown that sea surface temperature anomalies in tropical seas can influence the summer Asian monsoon and determine precipitation patterns [9]. Sea surface temperature anomalies can also act as triggers for ENSO events [10]. In addition, SST has an impact on tropical cyclone activity [11]. Therefore, SST prediction is an important topic in ocean research. Existing

products provide past and present SST data, and SST projections can provide future projections based on these data, providing an efficient and accurate complementary tool for the short-term monitoring and forecasting of the marine environment. This is of great practical importance for the management and protection of marine ecosystems, for early warning of marine hazards and for supporting shipping and offshore operations. However, SST is affected by a variety of factors such as heat flux, radiation, and ocean currents, so the accurate prediction of SST remains a challenge.

SST prediction methods have witnessed a notable improvement in accuracy in recent years [12–25]. Current methods for predicting SST can be categorized into two groups [26], the first are numerical methods and the second are data-driven methods. The numerical methods are based on physical conditions and processes, using mathematical models such as partial differential equations to describe the variations of sea surface temperature [27]. The method is not only more complex, but also requires multiple factors affecting the ocean surface temperature as inputs. Data-driven methods include traditional statistical methods, machine learning, and deep learning methods, etc., and such models learn patterns of sea surface temperature variation directly from data. The prediction of sea surface temperature is usually solved as a time series problem. Examples include support vector regression (SVR) [28] and multi-layer perceptrons (MLP) [29]. With the development of deep learning, it has evolved from the original RNN to improved networks such as long short-term memory (LSTM) and gated recurrent unit (GRU). These networks have been applied to SST prediction tasks with good performance [30–37]. Among the more classical ones, Zhang et al. [38] proposed a fully-connected network model based on LSTM (FC_LSTM) for SST prediction. Xie et al. [39] proposed a GRU encoder-decoder model (GED) based on SST code and dynamic influence link for SST prediction. Usharani [40] improved the accuracy of LSTM-based SST prediction by introducing a new loss function in LSTM. Jia et al. [41] predicted and analyzed SST in the East China Sea using LSTM. However, the above models only utilize the temporal information of SST, thereby neglecting the spatial information. This easily leads to a loss of a significant amount of important information, which impacts the prediction accuracy. Therefore, there is a need for a model that can fully utilize the spatiotemporal information of SST.

Recently, a newly proposed convolutional network architecture, namely the U-net model, has demonstrated remarkable ability in analyzing spatial variability patterns [42]. UNet, as well as various extended versions of UNet, have also been applied to problems such as cloud or ocean vortex studies [43,44], but rarely to SST prediction studies. Multi-scale convolution allows the data to be analyzed by using different sized convolution kernels at the same time, enabling the model to understand the details of the information and the larger context, helping the model to look at the data from multiple perspectives, capturing more comprehensive information, and thus making more accurate predictions [42,45,46]. Xiao et al. [47] applied the Convolutional LSTM (ConvLSTM) model to the spatio-temporal prediction of the East China Sea SST, and confirmed the spatial information processing capability and time series data processing capability of ConvLSTM. The Atrous Spatial Pyramid Pooling (ASPP) module [48] is able to effectively understand different features in the image ranging from small to wide by simultaneously capturing the multi-scale information of the image using different scales of dilated convolution, which enhances the model's ability to capture multi-scale spatial features in the data. Based on the content above, we propose an improved version of the UNet architecture, namely the Spatiotemporal UNet (ST-UNet). In this model, during the encoding stage, a multi-scale convolutional fusion block, combining multi-scale convolutional feature block with ConvLSTM, is employed to comprehensively capture both temporal and spatial characteristics [49]. Additionally, an ASPP module is utilized at the network bottleneck to further leverage spatial information. Finally, in the decoding stage, a multi-scale convolutional feature block is used to generate prediction results. Through this model, comprehensive utilization of spatiotemporal information in sea surface temperature data can be achieved. We utilize the 13-year SST time series in the Bohai and Yellow Seas, as well as the South China Sea,

and compare them with several representative SST prediction models. The experimental results show that ST-UNet has the best performance.

2. Materials and Methods

In this section, we present the details of our ST-UNet model. We first introduce the various components utilized to construct the network, and then provide a comprehensive explanation of the entire architecture.

2.1. Data

The experimental data utilized in this study were sourced from the National Oceanic and Atmospheric Administration (NOAA) optimum interpolation SST (OISST) dataset, accessible at <https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html> (accessed on 19 July 2023). The dataset includes daily, weekly, and monthly mean SST data from September 1981 to the present. The dataset offers extensive coverage of the entire global ocean region and is continually updated to ensure its accuracy and relevance. In the study of this paper, we selected the daily mean SST dataset of the Bohai and Yellow Seas ($33.07^{\circ}\text{N}\sim 41.0^{\circ}\text{N}$, $117.35^{\circ}\text{E}\sim 125.3^{\circ}\text{E}$) and the South China Sea ($12.125^{\circ}\text{N}\sim 19.875^{\circ}\text{N}$, $112.125^{\circ}\text{E}\sim 119.875^{\circ}\text{E}$) as the experimental dataset [50]. The area boxed out in Figure 1 is the ocean area where the SST data were collected in the experiment. The daily mean SST for this dataset was obtained by combining SST observations from different sources (including satellites, buoys, ship observations, etc.) using an optimal interpolation method and a daily mean calculation. The time span of this dataset is from January 2010 to December 2022, with 4748 days of data and the spatial resolution is $0.25^{\circ} \times 0.25^{\circ}$. Where the ground part is the default value, we set it to 0. In our experiments, we use a total of 4383 days of data from 2010 to 2021 to train the model. Subsequently, we evaluate the performance of the model by testing its predictions using data for a total of 365 days in 2022.

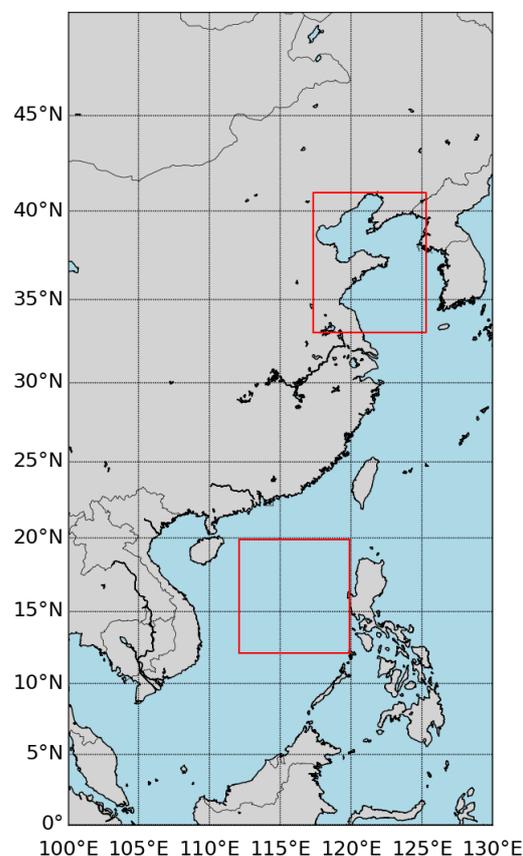


Figure 1. The target region for sea surface temperature (SST) data in the experiment.

2.2. Methods

2.2.1. Multi-Scale Convolutional Feature Block

Our block, depicted in Figure 2a, divides the data stream into parallel convolutions with different kernel sizes and subsequently reconnects the branches. The purpose of this design is to extract diverse features by applying multiple kernel sizes to the same data. In this research paper, we utilize three parallel branches with $1 \times 1 \times 1$, $3 \times 3 \times 3$, and $5 \times 5 \times 5$ convolution kernels. To reduce the computational cost, we approximate a $5 \times 5 \times 5$ convolution with two consecutive $3 \times 3 \times 3$ convolutions, Inspired by [51]. Next, the outputs of the various branches are concatenated and merged into a $1 \times 1 \times 1$ convolution. Consequently, the network learns over time to favor the branch with the most suitable kernel. Lastly, we apply the ReLU activation function to the output of this block. This block serves as the fundamental building block in the decoding phase of our network.

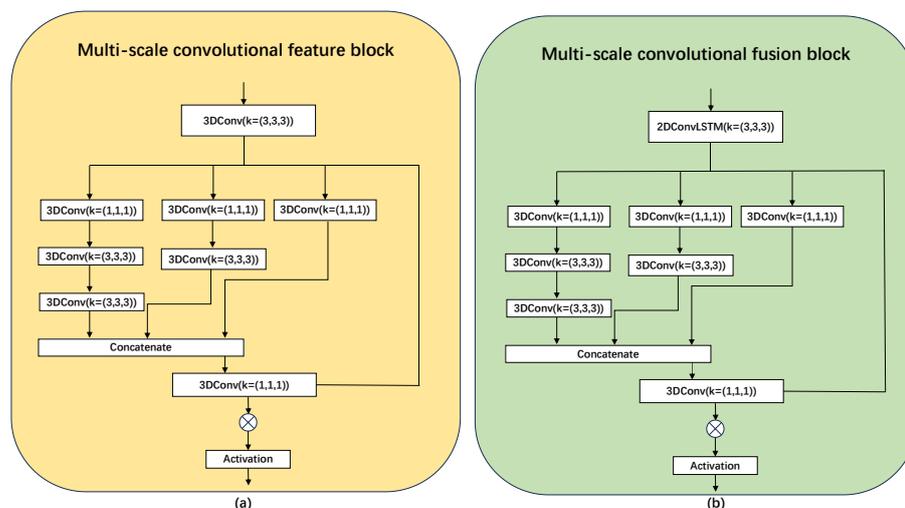


Figure 2. (a) Multi-scale convolutional feature block. Different kernel sizes are convolved in parallel to extract features at different scales. (b) Multi-scale convolutional fusion block. Replacing Conv3D with ConvLSTM to capture long-term dependencies and dynamics in time series based on the previous block. Where 3DConv stands for 3D convolution, k stands for convolution kernel size, Concatenate stands for splicing, and Activation stands for activation function. In addition, 2D ConvLSTM with $k = (3, 3, 3)$ represents a 2D ConvLSTM using a 3×3 spatial convolution kernel with depth 3.

2.2.2. Multi-Scale Convolutional Fusion Block

As shown in Figure 2b, our block improves on multi-scale convolutional feature block. We recognized that in order to address temporal dependencies in time series, it is not sufficient to use only convolution. Therefore, we decided to introduce ConvLSTM at the beginning of the branch instead of the previous convolutional layer. This choice is based on the memory capability of ConvLSTM, which can effectively model temporal relationships and better capture long-term dependencies and dynamics in time series. With this improvement, our model can better model dynamic changes in features. This block plays the role of a core building block in our network coding phase and its importance cannot be ignored. By introducing ConvLSTM, our block is able to more fully exploit the underlying patterns in the time series, improving the performance and modeling capabilities of the model.

2.2.3. Atrous Spatial Pyramid Pooling (ASPP)

Atrous Spatial Pyramid Pooling (ASPP) is a mechanism capable of capturing multi-scale information. It is realized by using multiple parallel convolutional branches. Each branch uses the same convolutional kernel size, but with different expansion rates (6, 12, and 18). The expansion rate is the number of intervals at which zero values are inserted into the convolution kernel during the convolution process. By performing convolution

operations at different expansion rates, information can be captured at different scales. In addition to the convolution branch with different expansion rates, it also includes a branch for extracting global context information. In this branch, the features of the entire image are downscaled through global pooling. These global features are then merged with the features from other branches using reshaping and upsampling operations. The resulting extracted features are concatenated and combined using a $1 \times 1 \times 1$ convolution. A visualization of this mechanism is shown in Figure 3.

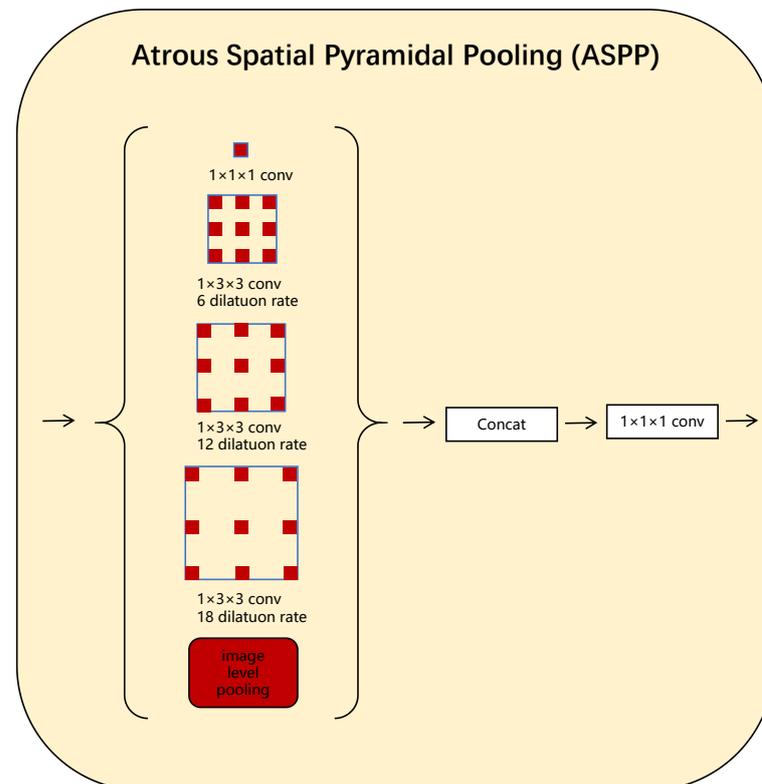


Figure 3. Atrous Spatial Pyramidal Pooling (ASPP) block. Different expansion rates enable the network to extract multi-scale information in order to fully extract spatial information.

2.2.4. ST-UNet

We chose the UNet architecture as the basis of our model because it has been shown to be very effective in solving image-to-image mapping tasks. Initially developed for medical image segmentation purposes, the U-Net architecture shares similarities with a self-encoder structure. The encoding stage involves a sequence of downsampling operations aimed at extracting essential feature information. The decoder stage classifies each pixel to reconstruct the segmented output. In our proposed encoder for ST-UNet, the multi-scale convolutional fusion blocks alternate with pooling operations. This configuration enables the network to effectively capture spatial and temporal dependencies within a set of 2D images during the contraction phase. The first input dimension (temporal dimension) is then reduced from a time step (35 in our case) to a prediction step (1, 3, and 7 in our case, respectively). This reduction occurs before the data are forwarded to the expanded phase. The extracted and merged features are then employed in the decoder to reconstruct the image. In the decoder, we follow a similar structure to the encoder, using alternating multi-scale convolutional feature blocks and upsampling. Ultimately, with this architectural design, our ST-UNet model is able to fully utilize temporal and spatial information for the prediction task, while maintaining the model performance at the lowest possible computational cost.

To achieve multi-scale feature learning, we integrate the ASPP module with the convolutional blocks at the network's bottleneck. This strategic placement occurs at a point where the data possesses a highly abstract representation. By adopting this approach, we enable the network to effectively leverage and comprehend information of varying scales from this representation, avoiding the need for larger kernels and excessive computational resources. In addition, we apply dropout techniques at bottlenecks to motivate the network to learn sparser data representations, reduce training time, and avoid the possibility of overfitting. Thus, the input shape of our network is $T \times H \times W \times F$ and the output shape is $L \times H \times W \times F$, where T is the number of time steps (lags), L is the prediction steps, H and W are the height and width of the image, and F is the number of features or elements, which we consider as channels in the network. Using a prediction step of 1 as an example, we achieve the reduction of the first input dimension by applying a convolution with a kernel size of $\text{lags} \times 1 \times 1$ and employing effective padding. The comprehensive architecture of our model is depicted in Figure 4.

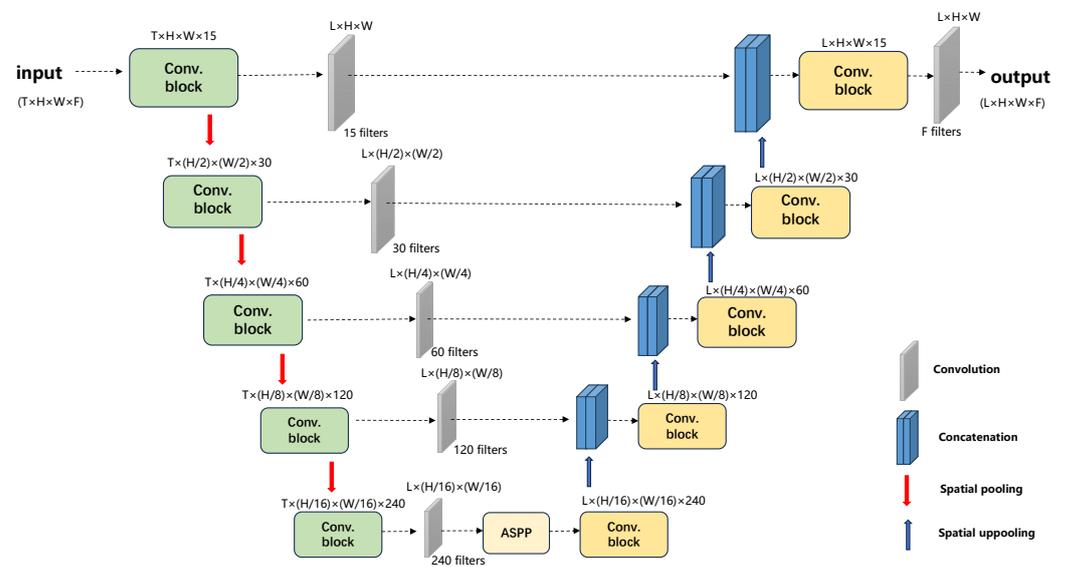


Figure 4. The complete ST-UNet architecture. For simplicity, the multi-scale convolutional feature block and the multi-scale convolutional fusion block are shown as different colored “Conv. block” according to the previously shown figure. The annotations on these blocks describe the output dimensions, where T denotes the time steps (lags), L is the prediction steps, H and W denote the height and width of each image, and F denotes the number of predicted features or elements.

2.3. Data Processing

To promote faster convergence during training and mitigate distribution discrepancies, we implement MaxMinScaler normalization to restrict the value range of the dataset to 0 to 1:

$$X_{norm} = \frac{X - Min}{Max - Min} \quad (1)$$

where X_{norm} is the normalized value, Min is the minimum value of the data in the dataset we used and Max is the maximum value.

2.4. Experiments Settings and Evaluation Indices

In our experimental configuration, we utilize Adam optimizer for training purposes. The batch size is set to 64, the training epoch is set to 100, and the learning rate is set at 0.001. To assess the performance of the model, we employ three metrics: mean absolute error (MAE), mean square error (MSE) and coefficient of determination (R^2). MAE provides an accurate representation of the actual prediction error, while MSE is highly sensitive to errors and effectively gauges prediction accuracy. The R^2 measures the performance of

our model in explaining the variance of the target variable in order to assess its prediction accuracy and reliability.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{pred}(i) - y_{true}(i)| \quad (2)$$

$$MSE = \sum_{i=1}^n (y_{pred}(i) - y_{true}(i))^2 \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{pred}(i) - y_{true}(i))^2}{\sum_{i=1}^n (y_{pred}(i) - \bar{y}_{true})^2} \quad (4)$$

where $y_{pred}(i)$ and $y_{true}(i)$ represent predicted and actual values, respectively.

3. Result

3.1. Accuracy Analysis

The ST-UNet was compared to the CFCC-LSTM (combined FC-LSTM and convolution neural network) model [52], GRU encoder–decoder (GED) [39] and memory graph convolutional network (MGCN) [53]. To fairly compare model performance, we align the sliding window length with the GED. Specifically, we configure the length of the sliding window to be 35 to predict the daily average SST for the subsequent 1, 3, and 7 days. To show the advantage of ST-UNet more clearly, we specifically indicate the best predictions in the table by highlighting them in bold. The experimental results on the Bohai Sea and Yellow Sea data sets are shown in Table 1. The comparison reveals that the ST-UNet, as proposed in this paper, exhibits varying degrees of superior performance over other models across different prediction scales, with the MGCN model following closely behind. Moreover, when the prediction scales are 1, 3, and 7 days, respectively, the MAE of the ST-UNet is higher than that of the MGCN model by 17%, 12%, and 2%, respectively (0.2085 vs. 0.2503, 0.3316 vs. 0.3779, and 0.5087 vs. 0.5210). To validate the reliability of ST-UNet, this research conducts a comparative analysis of the four models mentioned earlier, utilizing the South China Sea dataset. The experimental outcomes, which demonstrate the performance of each model, are documented in Table 2. The findings indicate that ST-UNet continues to exhibit superior performance compared to the other models across various prediction scales, which also proves the generality of our model. By comparing the data in Tables 1 and 2 together, we find that the prediction performance of all the methods seems to be better on the South China Sea dataset than on the Bohai Sea and Yellow Sea datasets. The preliminary analysis may be due to the fact that the South China Sea, with its lower latitude and geographic location near the equator, is able to absorb more heat throughout the year and has relatively less climatic variability compared to the Bohai Sea and the Yellow Sea. Therefore, the model can relatively simply predict SST in the South China Sea.

However, when R^2 is taken into account, the situation presents different characteristics. We observe that the Bohai and Yellow Seas also show higher values of R^2 despite their higher MSE and MAE, indicating that the model is able to capture the SST variability within this region. This phenomenon may be indicative of the fact that despite the larger prediction errors in absolute measures, the model is still able to effectively simulate the fluctuating trends in SST relative to the overall variability in this regions. In contrast, the South China Sea has relatively low R^2 values despite having a low MSE and MAE. This may reflect the fact that although the model predicts smaller errors in this region, the model has limited ability to capture these variabilities relative to the overall SST variability in the South China Sea. This may be related to the relatively more stable climatic conditions and smaller SST variability in the South China Sea, making even small prediction errors likely to appear more pronounced in the calculation of R^2 values.

Table 1. Prediction results of the Bohai Sea and Yellow Sea datasets.

Model	Metrics	Predictions Days		
		1	3	7
CFCC-LSTM	MSE (°C)	0.1334	0.3127	0.5843
	MAE (°C)	0.2347	0.4264	0.5971
	R ² (%)	99.73	99.44	98.85
GED	MSE (°C)	0.1305	0.3113	0.5811
	MAE (°C)	0.2459	0.4012	0.6026
	R ² (%)	99.74	99.45	98.86
MGCN	MSE (°C)	0.0985	0.2517	0.5146
	MAE (°C)	0.2503	0.3779	0.5210
	R ² (%)	99.78	99.49	98.95
ST-UNet	MSE (°C)	0.0823	0.2063	0.4674
	MAE (°C)	0.2085	0.3316	0.5087
	R ² (%)	99.83	99.59	99.05

Table 2. Prediction results of the South China Sea datasets.

Model	Metrics	Predictions Days		
		1	3	7
CFCC-LSTM	MSE (°C)	0.0714	0.1427	0.2495
	MAE (°C)	0.1778	0.2893	0.3856
	R ² (%)	96.97	93.16	86.01
GED	MSE (°C)	0.0681	0.1534	0.2619
	MAE (°C)	0.1718	0.2761	0.3646
	R ² (%)	97.00	93.18	85.73
MGCN	MSE (°C)	0.0533	0.1273	0.2230
	MAE (°C)	0.1626	0.2481	0.3383
	R ² (%)	97.32	94.19	88.68
ST-UNet	MSE (°C)	0.0286	0.0782	0.1868
	MAE (°C)	0.1192	0.2028	0.3269
	R ² (%)	98.62	96.15	90.50

In order to visualize the prediction performance of the model, this study now focuses on the SST prediction results for the onset and aftermath of specific extreme weather events and compares these results with the true values, which are selected from a test set with a forecast scale of 1 day. In particular, the Bohai and Yellow Seas for 15–22 September 2022 and the South China Sea for 30 October to 6 November 2022 have been selected for in-depth analysis. These dates were chosen because typhoons will pass through during this period. By comparing the predicted and actual values of the model, we can evaluate the performance of the model under complex weather conditions. The comparison results are shown in Figures 5 and 6. Since the input history length is 35 days, we visualize the first 35 days of the initial prediction date, so the first 35 days of each prediction date are shown in the figure. As shown in Figures 7 and 8, the red dots in the plots mark the path of action of the typhoon on that day. The slight difference is that the typhoons in Figure 7 are categorized into tropical storms by intensity, while those in Figure 8 are categorized into strong tropical storms and typhoons.

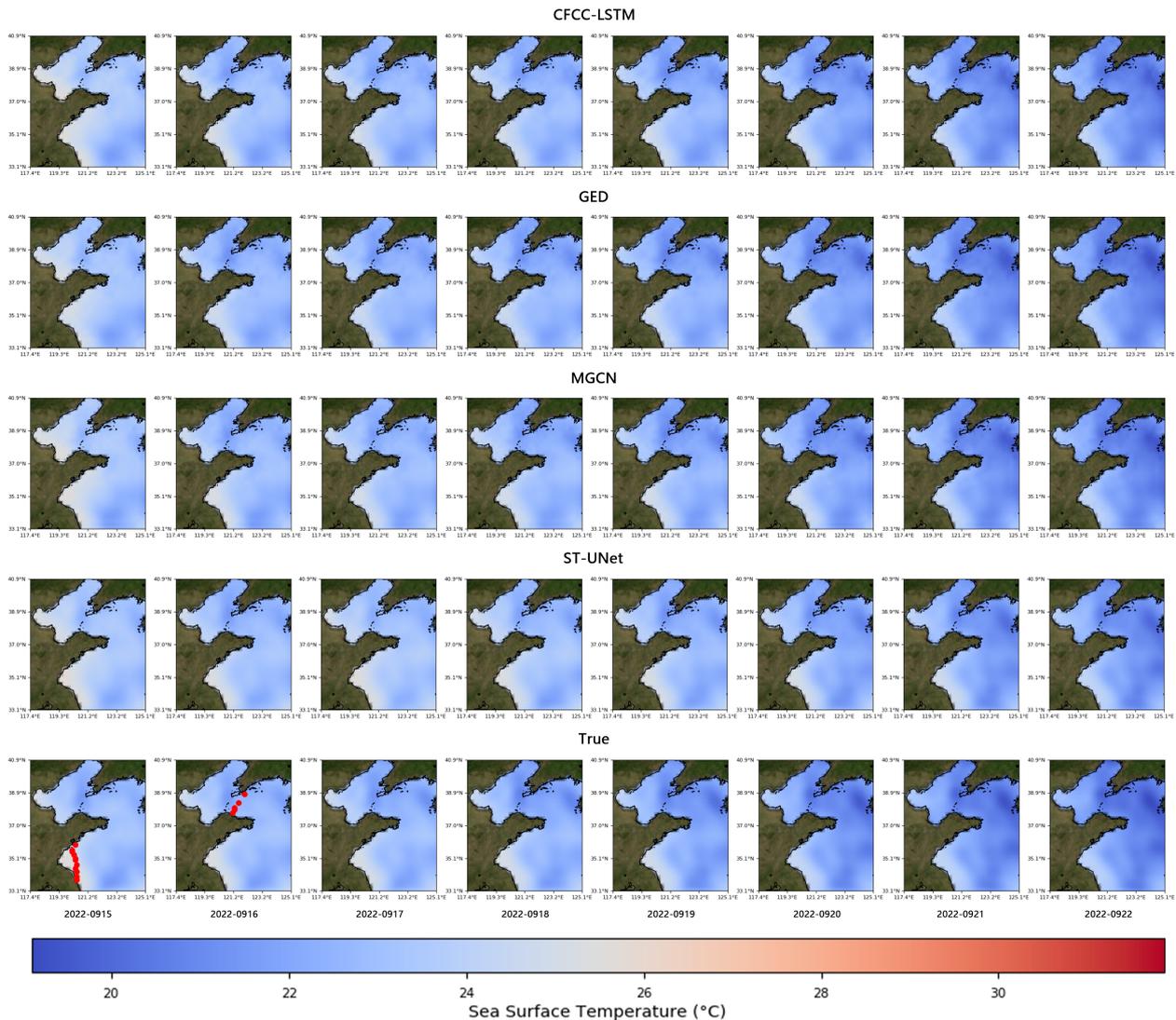


Figure 5. Visualization of SST in the Bohai and Yellow Seas predicted by four forecasting methods, with the red dots indicating the typhoon's path of movement on the day of the typhoon.

It is obvious from the data comparison between Tables 1 and 2 that our ST-UNet model's prediction performance in the South China Sea is significantly better than that in the Bohai Sea and the Yellow Sea at a prediction scale of 1 day. This is further confirmed by the comparative analysis in Figures 5 and 6. In particular, a careful analysis of these two figures reveals that the ST-UNet model exhibits overall higher prediction accuracy compared to other prediction methods, despite the differences in prediction ability in different sea areas.

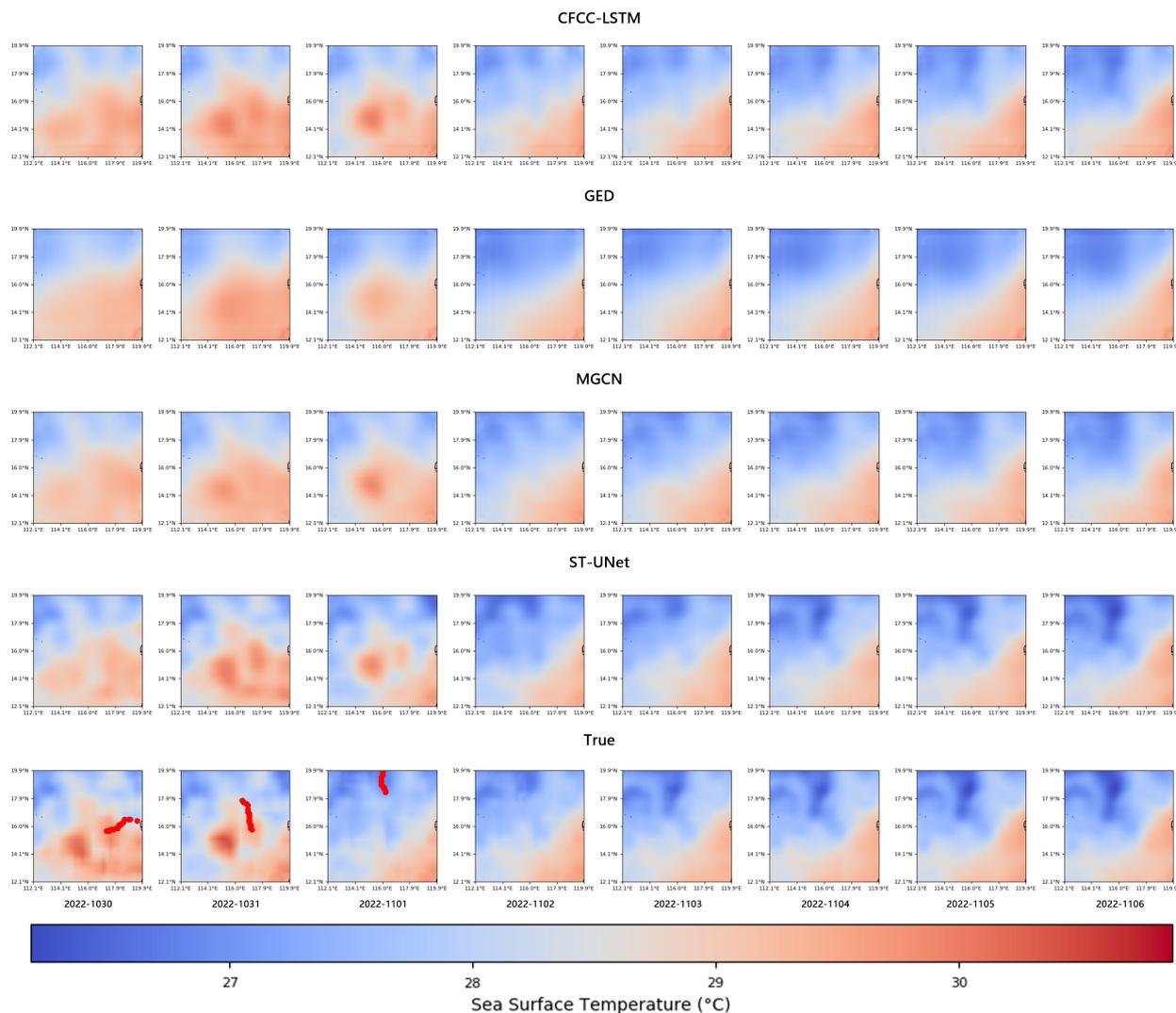


Figure 6. Visualization of SST in the South China Sea predicted by four forecasting methods, with the red dots indicating the typhoon's path of movement on the day of the typhoon.

However, as can be seen in Figure 6, a notable challenge we face is how to accurately predict SST changes during the transit of tropical cyclones. SST changes caused by tropical cyclones are extremely complex, and deep learning models still have limitations in capturing this complex ocean phenomenon. Despite the challenges in predicting extreme weather events, our ST-UNet model still outperforms existing models in SST prediction accuracy for other time periods. This demonstrates that we have made significant progress in improving the accuracy of deep learning for predicting sea surface temperature. In the future, we plan to continue to optimize our model and explore new ways to improve the model's ability to predict extreme weather events, especially the effects of tropical cyclone transits.

In conclusion, despite the challenges in predicting SST variations during tropical cyclones, our study still provides valuable insights into the application of deep learning in the field of sea surface temperature prediction and points to the direction of future research.

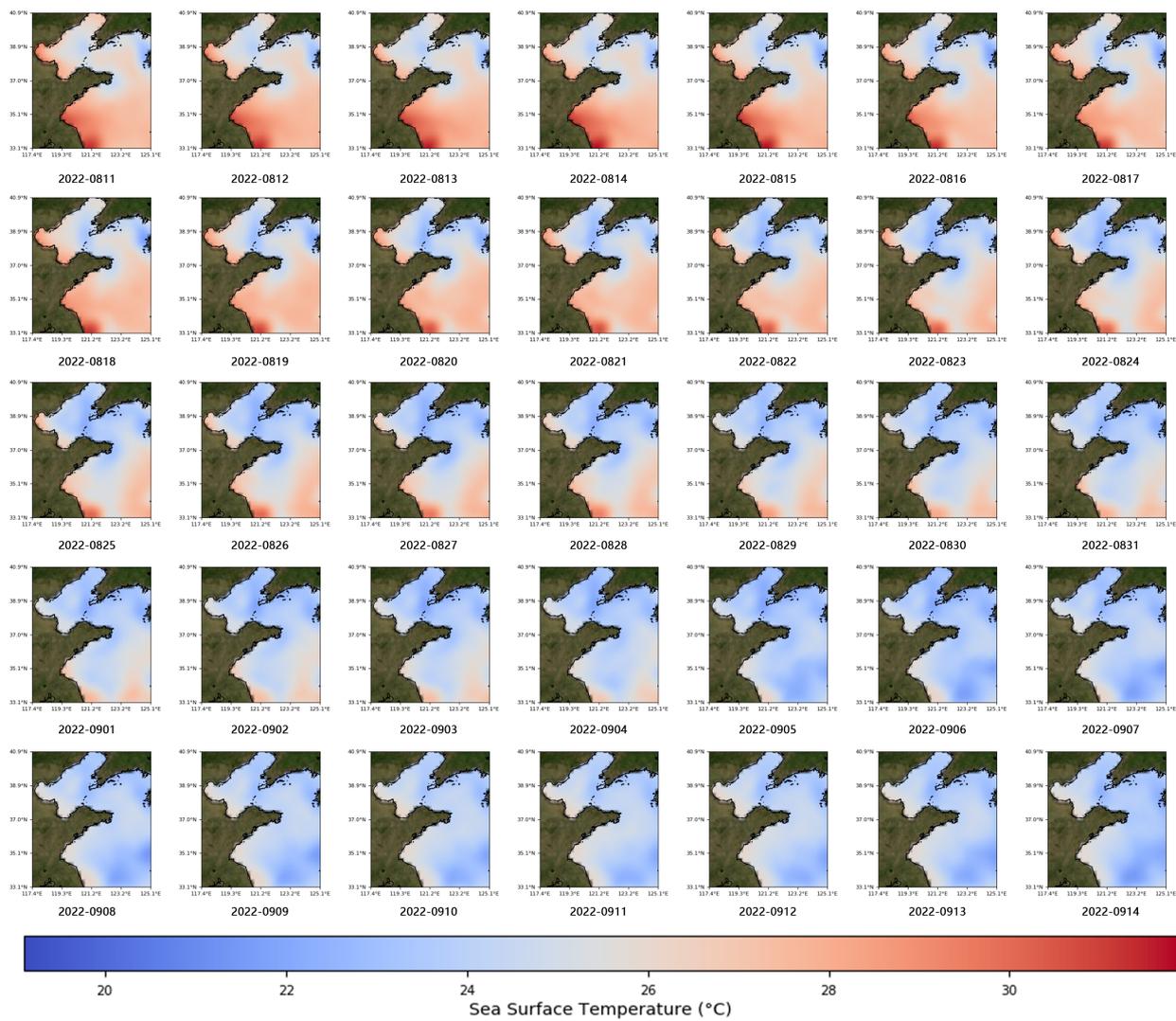


Figure 7. Visualization of SST in the Bohai and Yellow Seas from 11 August to 14 September 2022.

3.2. Analysis of Module

In this section, we compare four different sea surface temperature prediction models: ST-UNet, UNet, UNet-ASPP, and UNet-Convblock. Among them, UNet refers to ordinary 3D UNet. UNet-ASPP refers to a model that is based on UNet and places an ASPP block at the bottleneck. UNet-Convblock refers to a model that is based on UNet and replaces normal convolution with a multi-scale convolutional feature block and multi-scale convolutional fusion block. With these comparison experiments, we aim to evaluate the impact of each module on the prediction performance. As in the experiments above, the past 35 days are used to predict the daily mean SST for 1, 3 and 7 days in the future, and the bolded items in the table indicate the best prediction. The results of the experiments on the Bohai and Yellow Sea datasets are shown in Table 3, and the results on the South China Sea dataset are shown in Table 4. Through the two tables, it can be seen that the ST-UNet proposed in this paper has the best effect, followed by UNet-Convblock, then UNet-ASPP, and finally UNet. This indicates that the multi-scale convolutional feature block and multi-scale convolutional fusion block, as well as the ASPP module are effective, and all of them make the prediction accuracy gained to a different degree on the basis of UNet. Among them, the multi-scale convolutional feature block and multi-scale convolutional fusion block are more effective than the ASPP module.

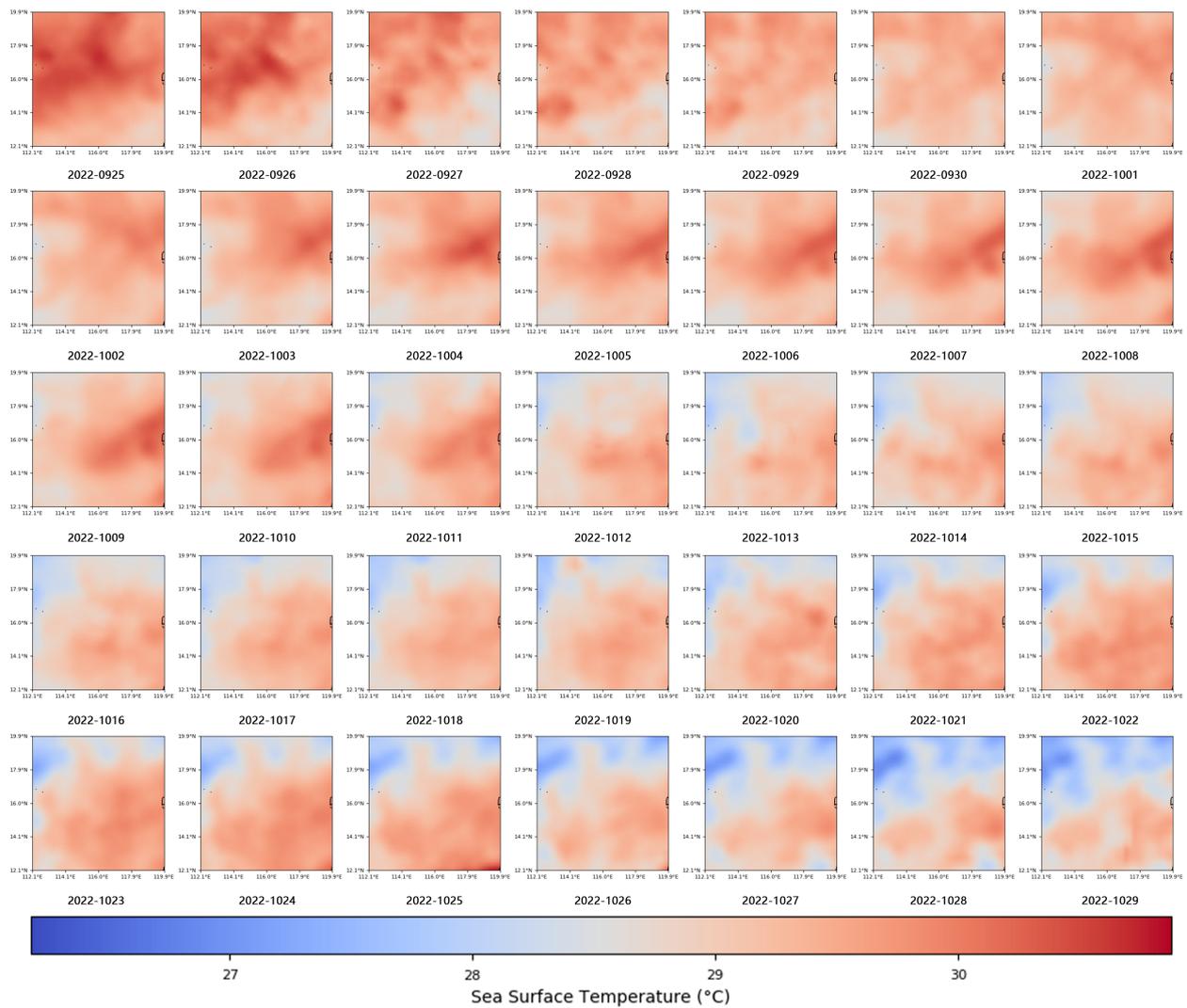


Figure 8. Visualization of SST in the South China Sea from 25 September to 29 October 2022.

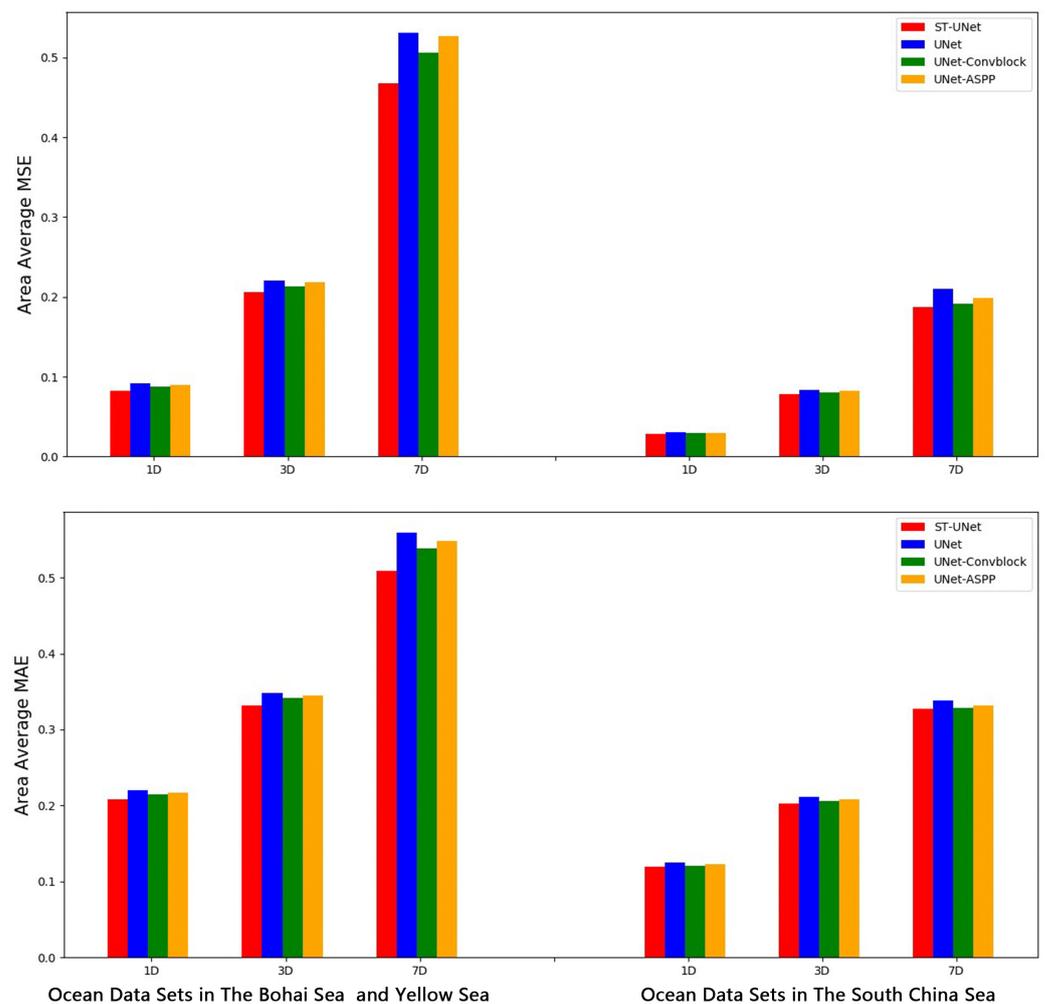
Table 3. Prediction results of the Bohai Sea and Yellow Sea datasets.

Model	Metrics	Predictions Days		
		1	3	7
UNet	MSE (°C)	0.0918	0.2210	0.5308
	MAE (°C)	0.2205	0.3477	0.5589
	R ² (%)	99.78	99.53	98.96
UNet-ASPP	MSE (°C)	0.0901	0.2188	0.5273
	MAE (°C)	0.2172	0.3448	0.5484
	R ² (%)	99.80	99.54	98.97
UNet-Convblock	MSE (°C)	0.0876	0.2134	0.5056
	MAE (°C)	0.2149	0.3412	0.5379
	R ² (%)	99.82	99.57	98.99
ST-UNet	MSE (°C)	0.0823	0.2063	0.4674
	MAE (°C)	0.2085	0.3316	0.5087
	R ² (%)	99.83	99.59	99.05

Table 4. Prediction results of the South China Sea datasets.

Model	Metrics	Predictions Days		
		1	3	7
UNet	MSE ($^{\circ}\text{C}$)	0.0309	0.0835	0.2098
	MAE ($^{\circ}\text{C}$)	0.1250	0.2114	0.3382
	R^2 (%)	98.58	95.95	89.25
UNet-ASPP	MSE ($^{\circ}\text{C}$)	0.0292	0.0827	0.1989
	MAE ($^{\circ}\text{C}$)	0.1223	0.2080	0.3322
	R^2 (%)	98.60	96.10	90.00
UNet-Convblock	MSE ($^{\circ}\text{C}$)	0.0291	0.0801	0.1912
	MAE ($^{\circ}\text{C}$)	0.1210	0.2058	0.3289
	R^2 (%)	98.61	96.13	90.37
ST-UNet	MSE ($^{\circ}\text{C}$)	0.0286	0.0782	0.1868
	MAE ($^{\circ}\text{C}$)	0.1192	0.2028	0.3269
	R^2 (%)	98.62	96.15	90.50

To facilitate a more convenient and visually descriptive presentation of the experimental findings in this paper, we generated plots depicting the Mean Absolute Error (MAE) and Mean Squared Error (MSE) for the four models across different prediction intervals (1 day, 3 days, and 7 days) for datasets from two distinct regions. These plots, as illustrated in the Figure 9, enable a comparative analysis.

**Figure 9.** SST prediction results of the four models at different scales and sea areas.

4. Conclusions

In our study, the ST-UNet model is designed to address the dynamic spatial correlation problem that is often neglected by conventional sea surface temperature (SST) prediction models. By fusing convolutional kernels of different sizes, ST-UNet is able to generate more accurate feature mappings, an innovation that enables the model to synthesize different magnitudes of surrounding information and thus capture the spatial distribution properties of SST more effectively. Specifically, the large convolutional kernel is able to capture broader regional information, while the small convolutional kernel is able to capture detailed information, and this fusion strategy greatly enhances the model's ability to sense spatial variations in SST. In addition, we also introduce the ASPP module into the model to apply the convolutional operations of different receptive fields in parallel. The inclusion of the ASPP module further enhances the model's ability to process information at different scales, enabling ST-UNet to utilize the available spatial information in a more comprehensive way. This design not only improves the performance of the model in dealing with static features, but also strengthens the model's ability to capture temporal correlations in time-series data by incorporating the ConvLSTM module, thus realizing effective prediction of dynamic changes in SST.

The performance evaluations conducted on the Bohai and Yellow Seas, as well as the South China Sea SST datasets, consistently demonstrate the superior performance of ST-UNet over other models such as CFCC-LSTM, GED, and MGCN. That results hold true across various sea areas and prediction scales. Secondly, at the same prediction scale, the SST in the South China Sea is easier to predict than that in the Bohai Sea and Yellow Sea. In future studies, we will explore new methods to improve the model's ability to predict extreme weather events, especially the effects of tropical cyclone transits.

Author Contributions: Conceptualization, L.S. and J.R.; methodology, J.R. and C.W.; software, J.R. and B.H.; validation, C.W.; formal analysis, C.W., J.R. and D.Z.; investigation, J.R.; resources, L.S.; data curation, C.W. and J.M.; writing—original draft preparation, J.R.; writing—review and editing, L.S. and C.W.; visualization, J.R., C.W. and J.W.; supervision, L.S.; project administration, L.S.; funding acquisition, L.S. and C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China under Grant 2018YFB0504905, National Natural Science Foundation of China under Grant 42276203.

Data Availability Statement: For more information, please refer to the website: <https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html> (accessed on 19 July 2023).

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

References

1. Bao, X.; Wan, X.; Gao, G.; Wu, D. The characteristics of the seasonal variability of the sea surface temperature field in the Bohai Sea, the Huanghai Sea and the East China Sea from AVHRR data. *Acta Oceanol. Sin.* **2002**, *24*, 125–133.
2. Cao, L.; Tang, R.; Huang, W.; Wang, Y. Seasonal variability and dynamics of coastal sea surface temperature fronts in the East China Sea. *Ocean Dyn.* **2021**, *71*, 237–249. [[CrossRef](#)]
3. Trenberth, K.E.; Branstator, G.W.; Karoly, D.; Kumar, A.; Lau, N.C.; Ropelewski, C. Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res. Ocean.* **1998**, *103*, 14291–14324. [[CrossRef](#)]
4. Xie, S.P.; Deser, C.; Vecchi, G.A.; Ma, J.; Teng, H.; Wittenberg, A.T. Global warming pattern formation: Sea surface temperature and rainfall. *J. Clim.* **2010**, *23*, 966–986. [[CrossRef](#)]
5. Deser, C.; Alexander, M.A.; Xie, S.P.; Phillips, A.S. Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.* **2010**, *2*, 115–143. [[CrossRef](#)]
6. Kennedy, J.J.; Rayner, N.; Atkinson, C.; Killick, R. An ensemble data set of sea surface temperature change from 1850: The Met Office Hadley Centre HadSST. 4.0. 0.0 data set. *J. Geophys. Res. Atmos.* **2019**, *124*, 7719–7763. [[CrossRef](#)]
7. Donlon, C.; Robinson, I.; Casey, K.; Vazquez-Cuervo, J.; Armstrong, E.; Arino, O.; Gentemann, C.; May, D.; LeBorgne, P.; Piollé, J.; et al. The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 1197–1214. [[CrossRef](#)]

8. Oliver, E.C.; Benthuisen, J.A.; Darmaraki, S.; Donat, M.G.; Hobday, A.J.; Holbrook, N.J.; Schlegel, R.W.; Sen Gupta, A. Marine heatwaves. *Annu. Rev. Mar. Sci.* **2021**, *13*, 313–342. [[CrossRef](#)] [[PubMed](#)]
9. Fan, L.; Shin, S.I.; Liu, Z.; Liu, Q. Sensitivity of Asian Summer Monsoon precipitation to tropical sea surface temperature anomalies. *Clim. Dyn.* **2016**, *47*, 2501–2514. [[CrossRef](#)]
10. Ham, Y.G.; Kug, J.S.; Park, J.Y.; Jin, F.F. Sea surface temperature in the north tropical Atlantic as a trigger for El Niño/Southern Oscillation events. *Nat. Geosci.* **2013**, *6*, 112–116. [[CrossRef](#)]
11. Ralph, T.U.; Gough, W.A. The influence of sea-surface temperatures on Eastern North Pacific tropical cyclone activity. *Theor. Appl. Climatol.* **2009**, *95*, 257–264. [[CrossRef](#)]
12. Kug, J.S.; Kang, I.S.; Lee, J.Y.; Jhun, J.G. A statistical approach to Indian Ocean sea surface temperature prediction using a dynamical ENSO prediction. *Geophys. Res. Lett.* **2004**, *31*. [[CrossRef](#)]
13. Berliner, L.M.; Wikle, C.K.; Cressie, N. Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *J. Clim.* **2000**, *13*, 3953–3968. [[CrossRef](#)]
14. Kug, J.S.; Lee, J.Y.; Kang, I.S. Global sea surface temperature prediction using a multimodel ensemble. *Mon. Weather Rev.* **2007**, *135*, 3239–3247. [[CrossRef](#)]
15. Repelli, C.A.; Nobre, P. Statistical prediction of sea-surface temperature over the tropical Atlantic. *Int. J. Climatol. J. R. Meteorol.* **2004**, *24*, 45–55. [[CrossRef](#)]
16. Borchert, L.F.; Menary, M.B.; Swingedouw, D.; Sgubin, G.; Hermanson, L.; Mignot, J. Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6. *Geophys. Res. Lett.* **2021**, *48*, e2020GL091307. [[CrossRef](#)]
17. Colman, A.; Davey, M. Statistical prediction of global sea-surface temperature anomalies. *Int. J. Climatol. J. R. Meteorol. Soc.* **2003**, *23*, 1677–1697. [[CrossRef](#)]
18. Barnett, T.; Graham, N.; Pazan, S.; White, W.; Latif, M.; Flügel, M. ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean–atmosphere model. *J. Clim.* **1993**, *6*, 1545–1566. [[CrossRef](#)]
19. Davis, R.E. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.* **1976**, *6*, 249–266. [[CrossRef](#)]
20. Alexander, M.A.; Matrosova, L.; Penland, C.; Scott, J.D.; Chang, P. Forecasting Pacific SSTs: Linear inverse model predictions of the PDO. *J. Clim.* **2008**, *21*, 385–402. [[CrossRef](#)]
21. Gao, G.; Marin, M.; Feng, M.; Yin, B.; Yang, D.; Feng, X.; Ding, Y.; Song, D. Drivers of marine heatwaves in the East China Sea and the South Yellow Sea in three consecutive summers during 2016–2018. *J. Geophys. Res. Ocean.* **2020**, *125*, e2020JC016518. [[CrossRef](#)]
22. Costa, P.; Gómez, B.; Venâncio, A.; Pérez, E.; Pérez-Muñuzuri, V. Using the Regional Ocean Modelling System (ROMS) to improve the sea surface temperature predictions of the MERCATOR Ocean System. *Sci. Mar.* **2012**, *76*, 165–175. [[CrossRef](#)]
23. Xue, Y.; Leetmaa, A. Forecasts of tropical Pacific SST and sea level using a Markov model. *Geophys. Res. Lett.* **2000**, *27*, 2701–2704. [[CrossRef](#)]
24. Collins, D.; Reason, C.; Tangang, F. Predictability of Indian Ocean sea surface temperature using canonical correlation analysis. *Clim. Dyn.* **2004**, *22*, 481–497. [[CrossRef](#)]
25. Wolff, S.; O’Donncha, F.; Chen, B. Statistical and machine learning ensemble modelling to forecast sea surface temperature. *J. Mar. Syst.* **2020**, *208*, 103347. [[CrossRef](#)]
26. Patil, K.; Deo, M.; Ravichandran, M. Prediction of sea surface temperature by combining numerical and neural techniques. *J. Atmos. Ocean. Technol.* **2016**, *33*, 1715–1726. [[CrossRef](#)]
27. Peng, W.; Chen, Q.; Zhou, S.; Huang, P. CMIP6 model-based analog forecasting for the seasonal prediction of sea surface temperature in the offshore area of China. *Geosci. Lett.* **2021**, *8*, 8. [[CrossRef](#)]
28. Imani, M.; Chen, Y.C.; You, R.J.; Lan, W.H.; Kuo, C.Y.; Chang, J.C.; Rateb, A. Spatiotemporal prediction of satellite altimetry sea level anomalies in the tropical Pacific ocean. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1126–1130. [[CrossRef](#)]
29. Aparna, S.; D’souza, S.; Arjun, N. Prediction of daily sea surface temperature using artificial neural networks. *Int. J. Remote Sens.* **2018**, *39*, 4214–4231. [[CrossRef](#)]
30. Hou, S.; Li, W.; Liu, T.; Zhou, S.; Guan, J.; Qin, R.; Wang, Z. MIMO: A Unified Spatio-Temporal Model for Multi-Scale Sea Surface Temperature Prediction. *Remote Sens.* **2022**, *14*, 2371. [[CrossRef](#)]
31. Wei, L.; Guan, L.; Qu, L.; Guo, D. Prediction of sea surface temperature in the China seas based on long short-term memory neural networks. *Remote Sens.* **2020**, *12*, 2697. [[CrossRef](#)]
32. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
33. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
34. Jordan, M.I. Serial order: A parallel distributed processing approach. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1997; Volume 121, pp. 471–495.
35. Xu, S.; Dai, D.; Cui, X.; Yin, X.; Jiang, S.; Pan, H.; Wang, G. A deep learning approach to predict sea surface temperature based on multiple modes. *Ocean Model.* **2023**, *181*, 102158. [[CrossRef](#)]
36. Shao, Q.; Li, W.; Han, G.; Hou, G.; Liu, S.; Gong, Y.; Qu, P. A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea. *J. Geophys. Res. Ocean.* **2021**, *126*, e2021JC017515. [[CrossRef](#)]

37. Kim, M.; Yang, H.; Kim, J. Sea surface temperature and high water temperature occurrence prediction using a long short-term memory model. *Remote Sens.* **2020**, *12*, 3654. [[CrossRef](#)]
38. Zhang, Q.; Wang, H.; Dong, J.; Zhong, G.; Sun, X. Prediction of sea surface temperature using long short-term memory. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1745–1749. [[CrossRef](#)]
39. Xie, J.; Zhang, J.; Yu, J.; Xu, L. An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 740–744. [[CrossRef](#)]
40. Usharani, B. ILF-LSTM: Enhanced loss function in LSTM to predict the sea surface temperature. *Soft Comput.* **2023**, *27*, 13129–13141. [[CrossRef](#)]
41. Jia, X.; Ji, Q.; Han, L.; Liu, Y.; Han, G.; Lin, X. Prediction of sea surface temperature in the East China Sea based on LSTM neural network. *Remote Sens.* **2022**, *14*, 3300. [[CrossRef](#)]
42. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention, Proceedings of the MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015*; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [[CrossRef](#)]
43. Jiao, L.; Huo, L.; Hu, C.; Tang, P. Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation. *Remote Sens.* **2020**, *12*, 2001. [[CrossRef](#)]
44. Lguensat, R.; Sun, M.; Fablet, R.; Tandeo, P.; Mason, E.; Chen, G. EddyNet: A deep neural network for pixel-wise classification of oceanic eddies. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 1764–1767. [[CrossRef](#)]
45. Cui, Z.; Chen, W.; Chen, Y. Multi-scale convolutional neural networks for time series classification. *arXiv* **2016**, arXiv:1603.06995.
46. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500. [[CrossRef](#)]
47. Xiao, C.; Chen, N.; Hu, C.; Wang, K.; Xu, Z.; Cai, Y.; Xu, L.; Chen, Z.; Gong, J. A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. *Environ. Model. Softw.* **2019**, *120*, 104502. [[CrossRef](#)]
48. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
49. Fernández, J.G.; Abdellaoui, I.A.; Mehrkanon, S. Deep coastal sea elements forecasting using UNet-based models. *Knowl. Based Syst.* **2022**, *252*, 109445. [[CrossRef](#)]
50. Huang, B.; Liu, C.; Banzon, V.; Freeman, E.; Graham, G.; Hankins, B.; Smith, T.; Zhang, H.M. Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1. *J. Clim.* **2021**, *34*, 2923–2939. [[CrossRef](#)]
51. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
52. Yang, Y.; Dong, J.; Sun, X.; Lima, E.; Mu, Q.; Wang, X. A CFCC-LSTM model for sea surface temperature prediction. *IEEE Geosci. Remote Sens. Lett.* **2017**, *15*, 207–211. [[CrossRef](#)]
53. Zhang, X.; Li, Y.; Frery, A.C.; Ren, P. Sea surface temperature prediction with memory graph convolutional networks. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [[CrossRef](#)]

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.