Environmental Modelling and Software 120 (2019) 104502

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data

Changjiang Xiao[a,b], Nengcheng Chen[a,*], Chuli Hu[e], Ke Wang[e], Zewei Xu[d], Yaping Cai[d], Lei Xu[a], Zeqiang Chen[a], Jianya Gong[a,c]

[a] State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, 430079, China
[b] College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China
[c] School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan, 430079, China
[d] CyberGIS Center for Advanced Digital and Spatial Studies, Department of Geography and Geographical Information Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, United States
[e] Faculty of Information Engineering, China University of Geosciences (Wuhan), Wuhan, 430074, China

ARTICLE INFO

Keywords:
Big data modelling
Spatiotemporal deep learning
ConvLSTM
Sea surface temperature (SST)
Field prediction
Time-series satellite data

ABSTRACT

重要的 整体的

Sea surface temperature (SST) is a vitally important parameter of the global ocean, which can profoundly affect the climate and marine ecosystems. To achieve an accurate and holistic prediction of the short and mid-term SST field, a spatiotemporal deep learning model is proposed which can capture the correlations of SST across both space and time. The model uses the convolutional long short-term memory (ConvLSTM) as the building block and is trained in an end-to-end manner. Experiments using 36-year satellite-derived SST data in a subarea of the East China Sea demonstrate that the proposed model outperforms the persistence model, the linear support vector regression (SVR) model, and two LSTM models with different settings, when judged using multiple statistics and from different perspectives. The results suggest that the proposed model is highly promising for short and mid-term daily SST field prediction accurately and conveniently.

使用ConvLSTM做的海洋温度的序列预测

1. Introduction and background

Sea surface temperature is one of the most important parameters of the world's ocean and plays a fundamental role in the exchange of energy, momentum, and moisture between the oceans and the atmosphere (Gentemann et al., 2000). Changes of SST can have profound effects on the global climate, marine ecosystem and even vegetation (Bouali et al., 2017; Cane et al., 1997; Castro et al., 2016; Chaidez et al., 2017; Friedel, 2012; Herbert et al., 2010; Rauscher et al., 2015; Yao et al., 2017). It can affect the precipitation distribution and further lead to droughts and floods (Rauscher et al., 2015; Salles et al., 2016). Daily SST, which shows surface thermal front and intensity, can be utilized to help detect marine ecosystems and assess the variability of such ecosystems, and to improve the understanding and qualification of the vertical structure of the water mass and the internal wave propagation. Besides, daily SST has been widely incorporated into hurricane models for anticipating tropical cyclone intensity (ESA). Therefore, it's vitally important and necessary to predict the future dynamics of SST, especially daily SST, in order to help identify potential extreme weather events in advance.

预测热带的气风强度

SST prediction is an active research topic. Methods for SST prediction can be summarized into three categories, including numerical methods, data-driven methods, and the combination of these two. The numeric methods are established based on physical, chemical and biological parameters as well as the complex interactions among them (Aparna et al., 2018; Liu and Fu, 2018; Stockdale et al., 2006). It uses mathematical models to describe the variations of SST. The numerical methods are usually quite sophisticated and predict SST at the ocean and even global scales with relatively low resolutions. Recent work has improved the SST and SST-related variables forecast potential of the physical model, namely, Nordic version of the Nucleus for European Modelling of the Ocean (NEMO-Nordic), in the Baltic Sea by assimilating high-resolution sea surface temperature data into the model (Liu and Fu, 2018). The data-driven methods attempt to solve the SST prediction problem from a data-centric perspective. They approximate the processes that generate the observed SST from patterns and relationships found in the observations. The learned patterns are then used to infer future SSTs. The data-driven methods require less ocean and

* Corresponding author. State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, 430079, China.
E-mail address: cnc@whu.edu.cn (N. Chen).

https://doi.org/10.1016/j.envsoft.2019.104502
Received 11 December 2018; Received in revised form 3 July 2019; Accepted 8 August 2019
Available online 13 August 2019
1364-8152/

resolution： 分辨率

atmospheric domain knowledge and are less sophisticated than the numerical methods. Besides, they are suitable for high-resolution SST predictions at relatively small scales. Methods in this category include traditional statistical methods such as the Markov model (Yan and Ants, 2000), the popular machine learning approaches, and artificial intelligent ones. Support vector machine (SVM) and neural network are widely used machine learning methods for SST prediction (Aparna et al., 2018; Lins et al., 2013; Patil and Deo, 2017; Tanggang et al., 1998; Wu et al., 2006; Zhang et al., 2017). Artificial intelligence methods for SST prediction include genetic algorithm (GA) and particle swarm optimization (PSO) (Lins et al., 2013; Neetu et al., 2011). To improve the prediction performance, the numeric methods are combined with the machine learning methods (Patil et al., 2016).

这个可以参考，写到师哥论文的背景介绍

In this paper, we focus on the high-resolution local SST field prediction in the East China Sea. As the numerical methods are too complex and involve the sophisticated interactions of a lot of variables, the much simpler data-driven methods are chosen. However, for traditional data-driven methods, most of them treat the observations in each location of an area independently without considering the influences of its surrounding locations, which may reduce the prediction performance. Besides, the prediction process is usually performed at each location individually, which is inconvenient (Zhang et al., 2017). Therefore, a prediction model that considers the spatiotemporal context and can predict the areal SST field holistically in an end-to-end manner is needed. To this end, the recently proposed recurrent neural network unit called convolutional long short-term memory (ConvLSTM) can be used due to its successful usage in similar applications such as precipitation prediction (Kim et al., 2017; Shi et al., 2015) and video frame prediction (Sautermeister, 2016). In these similar tasks, deep learning models based on the ConvLSTM unit which captures spatiotemporal correlations have achieved the state-of-the-art prediction performance. In this study, we adopt the ConvLSTM unit as the building block of our deep learning model to model the complex spatiotemporal dependencies of SST and further utilize the captured spatiotemporal dependencies for the SST field prediction task.

The goals of this paper are twofold: 1) developing a spatiotemporal deep learning model that can model and capture both the spatial and temporal dependencies of SST, and predict SST field accurately and holistically in an end-to-end manner; 2) investigating the applicability, effectiveness, and advantages of the developed model in predicting short and mid-term daily SST fields through experiments in a selected area in the East China Sea using 36-year time series satellite data.

The remainder of the paper is organized as follows. Section 2 describes the study area and satellite-derived time-series SST data used in this study. Section 3 presents the proposed spatiotemporal deep learning model for SST field prediction. The experimental results and thorough discussion are given in Section 4. Finally, Section 5 concludes the paper.

2. Study area and data

2.1. Study area

The study area locates in the East China Sea. The East China Sea is part of the famous Maritime Silk Road and one of the busiest seaways in the world. Understanding the patterns of SST and predicting future SST in this area is vitally important for safe seaway transportation, as well as the production and lives of people in China and its surrounding countries. As the East China Sea is located in the subtropical zone, the average annual water temperature is between 20 °C and 24 °C, with the annual temperature difference between 7 °C and 9 °C. The fluctuating characteristics of the SST in the East China Sea makes it suitable to be used for testing the prediction performance of the proposed spatiotemporal deep learning model. In this study, we choose a subarea of the East China Sea as our study area, covering the spatial extent of 27.5°N – 33°N, 123.5°E – 127.5°E, as is shown in Fig. 1. The reason why we choose this subarea is that it is the rectangular area that covers majority of the East China Sea without land pixels in it.

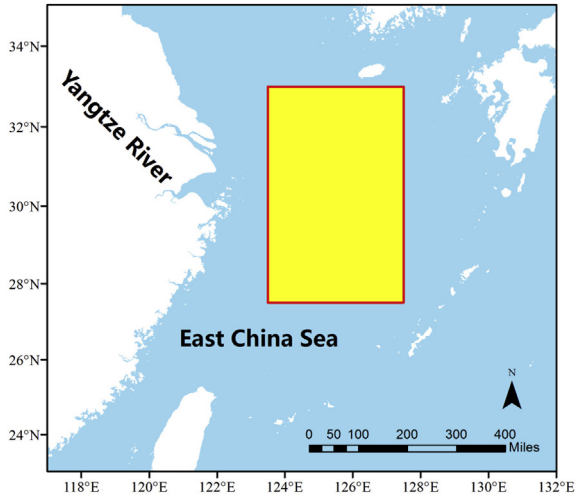

Fig. 1. The study area (in yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.2. Data

数据集可以借鉴

The data used is the daily Optimum Interpolation Sea Surface Temperature (daily OISST, version 2) produced by the National Oceanic and Atmospheric Administration (NOAA), with a spatial resolution of 1/4° latitude by 1/4° longitude. It's an analysis data using observations from multiple platforms, including satellites, ships, and buoys, on a regular global grid (Reynolds et al., 2007). At each grid, it contains a long time series of daily mean SST. The gaps are filled through interpolation, making it a spatially complete SST dataset. The satellite SST derived from the Advanced Very High Resolution Radiometer (AVHRR) is used to produce the adopted OISST which covers the global ocean from 89.975°S to 89.875°N, 0.125°E to 359.875°E, and covers temporal range from 1982/09/01 to present continuously. The subset of the OISST for the study area contains 13, 421 daily SST images from 1982/09/01 to 2018/05/30. Each image is with the size of 22 × 16 pixels. The original data are pre-processed and transformed into a 5-dimensional (5-D) tensor which is about 1.9 gigabytes (GB). Before feeding the tensor for training and testing the proposed spatiotemporal deep learning model and the LSTM models for comparison purpose, the tensor is normalized to the range [0,1]. For the linear SVR model, the tensor is normalized to the range [-1,1]. The normalization can make the data more centered, which helps to train the models more easily and quickly, and to further improve the performance of the models.

3. Formulation of the SST field prediction problem

In our research, the goal of SST field prediction is to forecast the next 10 days' SST field using the past few days' SST field (fixed prediction window). To achieve this goal, we adopt a rolling prediction scheme. In this scheme, we first use the SST field in the prediction window to forecast the next 1 day's SST field. Then, we move the time window 1 day ahead which incorporates the newly forecasted SST field as its latest element to predict the SST field one more day ahead. By repeating this process, 10-days-ahead SST field prediction can be achieved. Therefore, the key to achieving the goal is to predict the next 1 day's SST field.

滑动窗口

Suppose we observe the dynamic SST field over a spatial region which is represented by $M \times N$ grid, with M being the number of rows, and N being the number of columns. Inside each cell of the $M \times N$ grid,

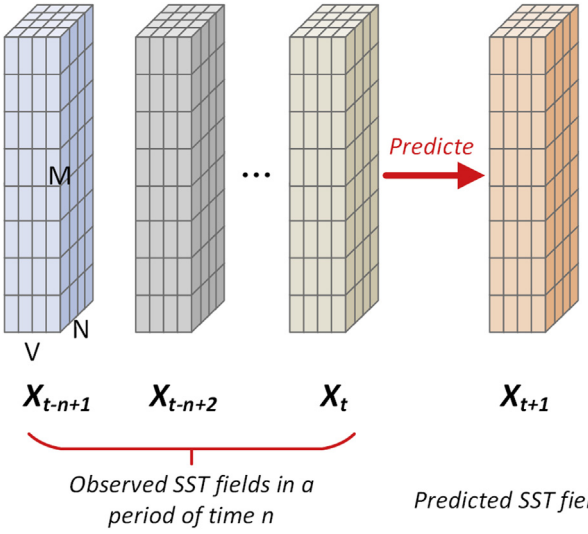

Fig. 2. Illustration of using the SST fields observed over a spatial region ($M \times N$ grids) in a period of time n to predict the SST field one-time-step ahead in the future over the same spatial region. $V = 1$ in the SST prediction task in this paper.

这个观测值以后能不能考虑不只是用温度，embedding的思想会不会好一些

there are V measurements (as we only use SST as the predicting variable, the value of V is 1), varying over time. Thus, the observation at any time slice t can be represented by a 3-D tensor $X_t \in R^{M\times N\times V}$, where R represents the domain of the observed SST. In a period of time n, we can obtain a spatiotemporal sequence of observations $X_{t-n+1}, X_{t-n+2}, ..., X_t$. The 1-day-ahead SST prediction problem is thus to forecast the most likely tensor in the future given the previous n-days observations, which can be formulated in Equation (1) and illustrated in Fig. 2:

$$X'_{t+1} = \underset{X_{t+1}}{\operatorname{argmax}}\, p(X_{t+1}|X_{t-n+1}, X_{t-n+2}, ..., X_t) \tag{1}$$

where $p(\cdot)$ is a conditional probability and X'_{t+1} is the predicted SST field at time $t+1$.

4. The spatiotemporal deep learning model for SST field modelling and prediction

4.1. Building block of the model

From Section 3 we can see that the SST field prediction involves the domains of both space and time. To deal with it, we adopt the ConvLSTM as the building block of our model. The ConvLSTM was proposed to overcome the drawbacks of fully-connected LSTM (FC-LSTM) in handling spatiotemporal data during which the spatial correlations are lost (Shi et al., 2015). Different from FC-LSTM, the ConvLSTM determines the future state of a cell in the grid by the inputs and past states of its local neighbours during which both the spatial and temporal correlations are captured and utilized. This is achieved by replacing the matrix multiplication operation used in FC-LSTM with convolution operation in the state-to-state and input-to-state transitions, as shown in Fig. 3.

The working mechanism of the ConvLSTM including the gates (input gate, forget gate and output gate) and information flow can be expressed using Equation (2) – (7) with all the variables defined and explained in Table 1. In the equations, ' $\otimes$ ' denotes the convolution operation, and ' $*$ ' denotes the Hadamard product. σ is a sigmoid function used as the activation function applied to the weighted sum of the inputs of each gate. The weighted sum of the inputs of each gate is a 3-D tensor, and the σ is applied to each element of the 3-D tensor.

$$f_t = \sigma(W_{xf} \otimes X_t + W_{hf} \otimes H_{t-1} + W_{cf} * C_{t-1} + b_f) \tag{2}$$

$$i_t = \sigma(W_{xi} \otimes X_t + W_{hi} \otimes H_{t-1} + W_{ci} * C_{t-1} + b_i) \tag{3}$$

$$C'_t = \tanh(W_{xc} \otimes X_t + W_{hc} \otimes H_{t-1} + b_c) \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * C'_t \tag{5}$$

$$o_t = \sigma(W_{xo} \otimes X_t + W_{ho} \otimes H_{t-1} + W_{co} * C_t + b_o) \tag{6}$$

$$H_t = o_t * \tanh(C_t) \tag{7}$$

4.2. Structure of the model

To better capture the spatial and temporal correlations among SST fields at different time, we build a spatiotemporal deep learning model by stacking the ConvLSTM layers. As shown in Fig. 4, the model consists of four layers, including an input layer, two ConvLSTM layers, and a fully-connected layer as the output layer. Our experiments reveal that this neural network design can achieve the best prediction accuracy among a couple of designs. The input of the whole network is in the 5-D form expressed as (*samples, timesteps, width, height, features*). *samples* is the batch size for training. Empirically, the batch size is set to a value around 128. By trials and errors around 128, we set it to 150 which can achieve the best prediction accuracy and relatively high training efficiency. *timesteps* is the length of the time window of the input SST fields to predict the future SST field. In our method, we set it to 50 which is about 4 times of the prediction length (10 days) according to the periodical changes of temperature data (Zhang et al., 2017), meaning that we use the previous 50 days' spatiotemporal SST sequence to predict the 51st day's SST field. *width* and *height* are 22 and 16 respectively which are the spatial size of the input SST images. As we only use one observation variable, namely SST, as the predictor, *features* is set to 1. As there are no theoretical rules for determining the best number of convolutional kernels of a convolutional layer, the usually adopted method is to test the performance of the neural network model by varying the number of kernels in a given range. Using this scheme, we set the number of the convolutional kernels of both the two ConvLSTM layers to 12. Similarly, we vary the size of convolutional kernels in a given range, and finally find that the model can achieve best performance when the sizes of all convolutional kernels are set to 3×3. Zero paddings are used to ensure that the states have the same spatial dimensions (rows and columns) as the inputs. We use mean square error (MSE) as the loss function and Adam (Kingma and Ba, 2014) as the optimizer for training the model. The whole model is conveniently trained in an end-to-end manner.

这里也是只使用的温度

5. Experiments and discussion

5.1. Experiment setup

In this study, we compare the proposed spatiotemporal deep learning model with three different models, including LSTM, linear SVR and persistence model (only for the comparison of 1-day-ahead predictions). For the LSTM model, two different kinds of settings are compared. One is that we treat each pixel in an SST field as an individual sample and it contains only one feature which is called LSTM (1-feature) model. The other setting is that we treat all the pixels in an SST field as a sample and it contains 352 (22×16) features which is called LSTM (n-features) model. Both these two settings consider only the temporal correlations except the second one is an end-to-end trainable model. The linear SVR model adopts the same setting as the LSTM (1-feature) model. The persistence model simply uses the current SST field as the predictions of the next day.

这里的这种对比试验可以借鉴写到我后面那篇文章里面去，包括其他模型有哪些，是怎么设置的，甚至结果也可以借鉴

To build the dataset for training and testing the models, we first create the 5-D tensor as stated in Section 2.2. The tensor is arranged in the order of time along the *samples* dimension. Then, along the *samples* dimension of the 5-D tensor described in Section 2.2, we select the first 80% as the training samples, and the remaining 20% as held-out testing

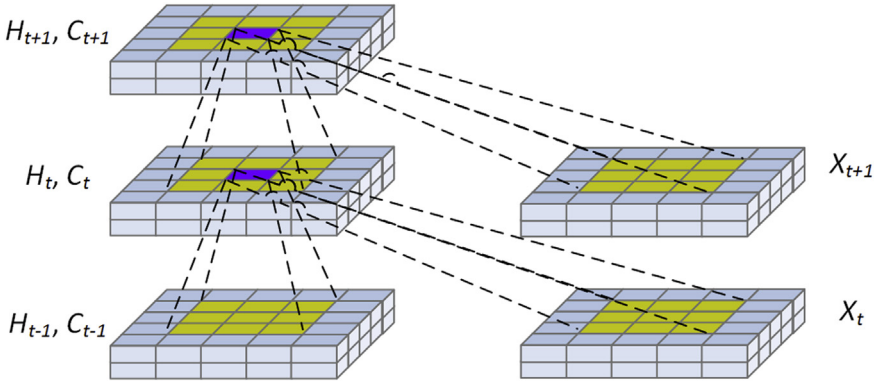

Fig. 3. Inner structure of ConvLSTM (Shi et al., 2015).

测试模型的指标也可以借鉴一下

samples. Further, 5% of the training samples are split out for validation purpose during training. Namely, the training and testing samples are not randomly chosen, but chosen in the order of time. The testing samples are samples in the future relative to the training samples. The actual number of samples for training, validation and testing are shown in Table 2. All of the samples are with the dimension of (51, 22, 16, 1). The proposed spatiotemporal deep learning model and two LSTM models are implemented with Keras using TensorFlow 1.5.0 (GPU version) as the backend and trained and tested on a computer with a single NVIDIA TITAN Xp GPU (Abadi et al., 2016; Chollet, 2015). The linear SVR is implemented using Scikit-learn 0.19.1 (Pedregosa et al., 2011). During training of the proposed spatiotemporal deep learning model, the loss function, namely MSE, decreases quickly and converges to a small value on both the training dataset and validation dataset, and the difference between the two converged values are small, indicating that the model is trained to a proper degree. Besides, the training results of the proposed model are relatively stable with respect to repeated training. Though they are different in the first few epochs for repeated training due to the random initialization of the model, they then converge and become stable. The difference among the converged values for repeated training are small as well.

Three indexes are selected to measure the performance of the different models for SST field prediction, including the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the Pearson correlation coefficient (r), which are defined as follows:

$$RMSE = \sqrt{\frac{\sum_1^n d_i^2}{n}} \tag{8}$$

$$MAPE = \frac{1}{n}\sum_1^n \left|\frac{y_i - y'_i}{y_i}\right| * 100 \tag{9}$$

$$r = \frac{\sum_1^n (y_i - \bar{y})(y'_i - \bar{y}')}{\sqrt{\sum_1^n (y_i - \bar{y})^2}\sqrt{\sum_1^n (y'_i - \bar{y}')^2}} \tag{10}$$

where d_i is the error vector calculated by the difference between the reference SST value y_i and the predicted SST value y'_i. $\bar{y}$ and $\bar{y}'$ are the mean value of the reference SST values and the predicted SST values, respectively. n is the total number of testing samples.

5.2. Experiment results and discussion

We conduct experiments for 1–10 days ahead predictions and compare the prediction performance of different models using different statistics and from different perspectives. Specifically, for the 1-day-ahead prediction, persistence model is incorporated for comparison.

Table 1
Definition and explanation of the variables used in Equation (2)–(7).

Variable	Definition and explanation	Dimension
X_t	The input 3-D tensor at time t.	3-D
H_{t-1}	The output 3-D tensor from the cell at time t-1.	3-D
H_t	The output 3-D tensor from the cell at time t.	3-D
C_{t-1}	The cell state at time t-1.	3-D
C'_t	The information that is stored in the new cell state C_t.	3-D
C_t	The cell state at time t.	3-D
f_t	Output of the forget gate. The value of each of its element is between 0 and 1. It controls the information that is forgotten in the old cell state C_{t-1}.	3-D
i_t	Output of the input gate. The value of each of its element is between 0 and 1. It controls how much of the information C'_t will be stored in the new cell state C_t.	3-D
o_t	Output of the output gate. The value of each of its element is between 0 and 1. It controls the information that is output (H_t) from the cell.	3-D
W_{xf}	The convolution kernel applied to the input 3-D tensor X_t in the forget gate.	3-D
W_{xi}	The convolution kernel applied to the input 3-D tensor X_t in the input gate.	3-D
W_{xc}	The convolution kernel applied to the input 3-D tensor X_t for creating the information C'_t that will stored in the new cell state C_t.	3-D
W_{xo}	The convolution kernel applied to the input 3-D tensor X_t in the output gate.	3-D
W_{hf}	The convolution kernels applied to the input 3-D tensor H_{t-1} in the forget gate.	3-D
W_{hi}	The convolution kernels applied to the input 3-D tensor H_{t-1} in the input gate.	3-D
W_{hc}	The convolution kernel applied to the input 3-D tensor H_{t-1} for creating the information C'_t that will be stored in the new cell state C_t.	3-D
W_{ho}	The convolution kernels applied to the input 3-D tensor H_{t-1} in the output gate.	3-D
W_{cf}	The weight that is applied to the old cell state C_{t-1} in the forget gate.	3-D
W_{ci}	The weight that is applied to the old cell state C_{t-1} in the input gate.	3-D
W_{co}	The weight that is applied to the new cell state C_t in the output gate.	3-D
b_f	The biases in the forget gate.	3-D
b_i	The bias in the input gate.	3-D
b_c	The bias for creating the information C'_t that will be stored in the new cell state C_t.	3-D
b_o	The bias in the output gate.	3-D

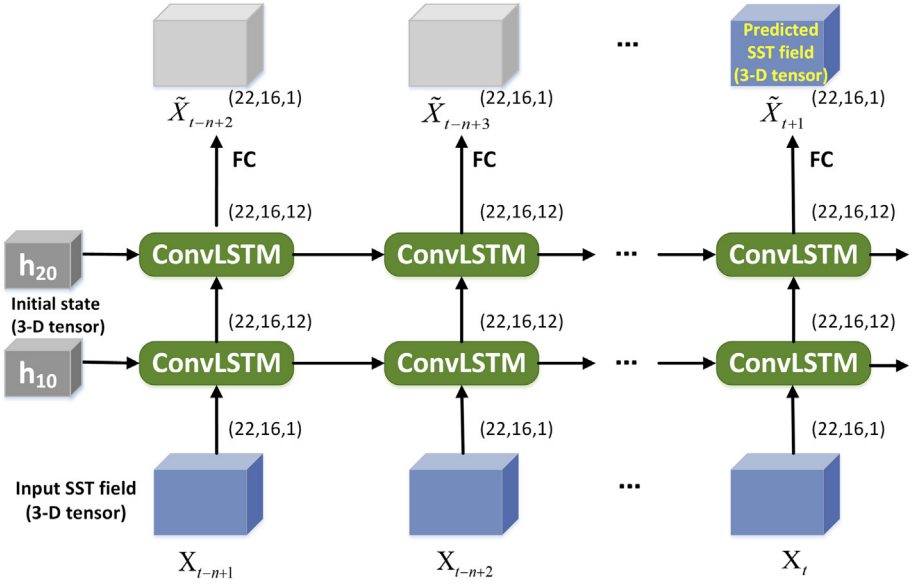

Fig. 4. Structure of the proposed spatiotemporal deep learning model for SST field prediction.

Table 2
Number of samples for training, validation and testing.

Total samples	Training samples	Validation samples	Testing samples
13370	10161	535	2674

Fig. 5 shows the spatial distribution of RMSEs over the study area for 1-day-ahead SST field prediction using different models. It can be seen that the proposed model performs better than the linear SVR, LSTM (1-feature) model, and the persistence model, and much better than the LSTM (n-features) model. It can also be seen that all the five models have their own largest prediction errors at nearly the same locations. This may be due to the abrupt changes of SST at these locations, as shown in Fig. 6. However, the prediction capability of the proposed model is less affected by this condition.

Fig. 7 shows the RMSE of predictions for the lead time of 1–10 days by different models. For each prediction lead, the RMSE is calculated over the whole study area on all the testing samples. The figure shows that the proposed model consistently outperforms all the other three models, which means that it can better capture the patterns hidden in historical SST fields and make more accurate predictions based on the captured patterns. The improvement of the RMSE of SST field predictions by the proposed model relative to the three models is given in Table 3. For linear SVR and LSTM (1-feature) model, the improvement increases along with the lead time of prediction while for the LSTM (n-features) model, the improvement decreases from the lead time of 1–7 days and then increases. One interesting finding from Fig. 6 is that the linear SVR model outperforms the LSTM (1-feature) model for the lead time of 1–4 days, but it performs worse than the LSTM (1-feature) model from the lead time of 5 days. A similar phenomenon exists between the linear SVR model and the LSTM (n-features) model. The reason may be that LSTM-based models are good at modelling long time dependencies of SST with their recurrent network structure and gating mechanisms. Fig. 8 and Table 4 show the comparison of MAPE of 1–10 days ahead prediction of SST field by different models, and the percentage improvement of MAPE by the proposed model relative to the other three models, respectively. They show similar patterns as those of the RMSE in Fig. 7 and Table 3.

In Fig. 9, the Pearson correlation coefficients (r) between the predicted SST field and the reference SST field for each prediction lead on all the testing samples are compared. The corresponding p values for the calculated r are all 0, showing the statistical significance of the calculated r. It shows that the r values of the proposed model are quite close to 1 and consistently bigger than those of the other three models, showing a stronger positive linear correlation between the predicted SST and the reference SST. Besides, the r values of the proposed model decrease more slowly than those of the other three models as the prediction lead increases from 1 to 10 days. The r values of the linear SVR model shows almost a straight downward trend as the prediction lead increases.

Fig. 10 compares the prediction performance of different models from the perspective of the spatial distribution of RMSEs of prediction over the study area. It shows that for each prediction lead, the RMSEs of the proposed model are lower than those of the other three models and

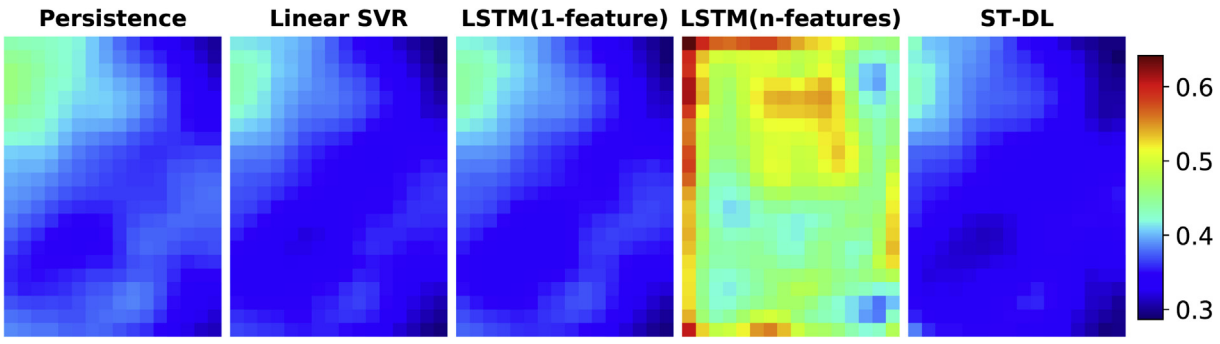

Fig. 5. Spatial distribution of RMSEs in the study area for 1-day ahead prediction on all the testing samples.

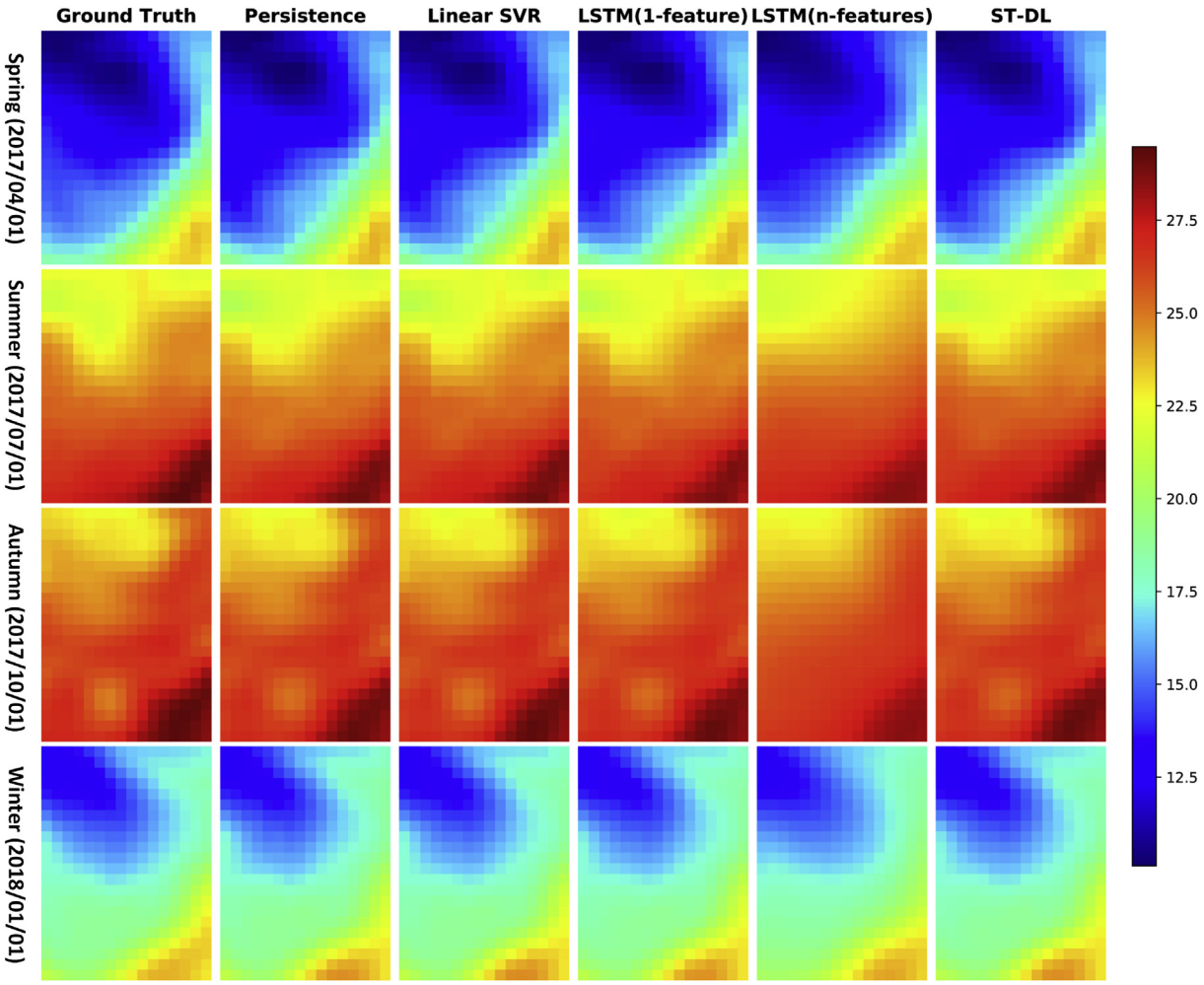

Fig. 6. 1-day-ahead prediction of SST field on a selected date in different seasons.

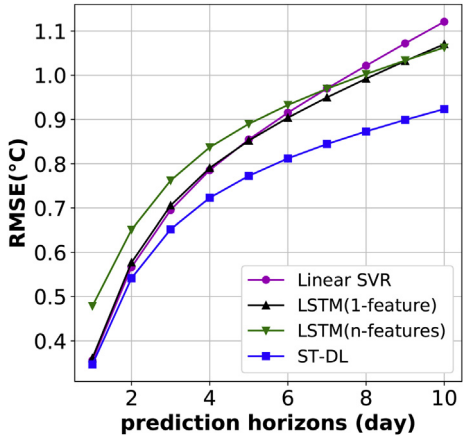

Fig. 7. Comparison of RMSE for 1–10 days ahead SST field prediction using different models.

are more evenly distributed over the study area. The RMSEs of the linear SVR model, LSTM (1-feature) model and LSTM (n-features) model increase quickly at locations where SST changes quickly (as shown in Fig. 6) as the prediction lead increases. In contrast, the proposed model has consistently and stably better performance at these locations.

To investigate the prediction performance from a holistic perspective, we compare the distribution of the prediction errors of 1–10 days ahead as a whole. It is achieved by conducting the kernel density estimation (KDE) of the prediction errors of different models using the Gaussian kernel (Terrell and Scott, 1992). The KDE curves of prediction errors of different models are shown in Fig. 11. The figure reveals that the proposed model has a much denser KDE curve with its mean almost equivalent to 0, showing that the errors of the proposed model spread more densely around 0 than the other three models do. Table 5 is also from a holistic perspective and shows the performance statistics calculated on all the 1–10 days ahead SST field prediction of each model, which demonstrates the outperforming prediction capability of the proposed model as well. Besides, the LSTM (1-feature) model predicts

Table 3
Percentage improvements of the RMSE by the proposed spatiotemporal deep learning model relative to the linear SVR model, the LSTM (1-feature) model, and the LSTM (n-features) model for each prediction lead.

Prediction lead (day)	Linear SVR (%)	LSTM (1-feature) (%)	LSTM (n-features) (%)
1	2.49	3.94	**27.48**
2	4.41	6.16	16.81
3	6.22	7.61	14.45
4	7.96	8.53	13.63
5	9.62	9.36	13.21
6	11.28	10.18	12.93
7	12.92	11.07	12.87
8	14.53	11.99	12.91
9	16.11	12.87	12.96
10	**17.62**	**13.71**	13.10

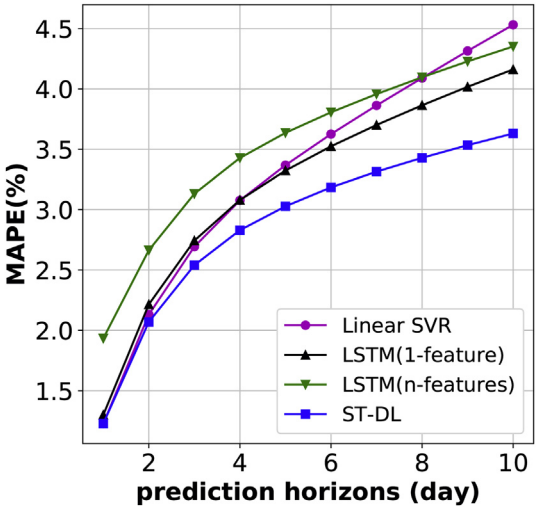

Fig. 8. Comparison of MAPE for 1–10 days ahead SST field prediction using different models.

Table 4
Percentage improvements of the MAPE by the proposed spatiotemporal deep learning model relative to the linear SVR model, the LSTM (1-feature) model, and the LSTM (n-features) model for each prediction lead.

Prediction lead (day)	Linear SVR (%)	LSTM (1-feature) (%)	LSTM (n-features) (%)
1	−0.03	5.52	**36.49**
2	2.87	6.56	22.32
3	5.69	7.42	18.90
4	8.03	8.12	17.45
5	10.10	8.89	16.74
6	12.16	9.63	16.33
7	14.21	10.46	16.25
8	16.17	11.26	16.30
9	18.11	12.05	16.43
10	**19.89**	**12.76**	16.58

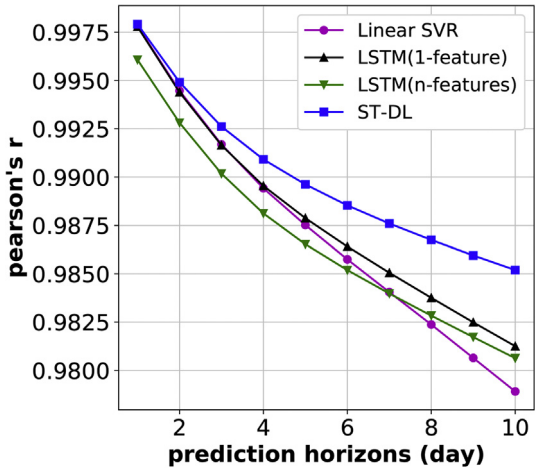

Fig. 9. Comparison of r for 1–10 days ahead SST field predictions using different models.

better than the linear SVR model considering long prediction leads holistically which is in line with what has been previously observed from Fig. 7.

From the above results using different statistics and from different perspectives, we can see that the proposed model outperforms the liner SVR model, the LSTM (1-feature) model, and the LSTM (n-features) model for the SST field prediction task. The reason is that the proposed model has both spatial modelling capability by using convolutional operation and temporal modelling capability by using recurrent network structure and gating mechanisms. It determines the future state of a cell in the grid by the inputs and past states of its local neighbours during which both the spatial and temporal correlations of SST are captured and utilized, which however is not inherently implemented in other models.

As we aim to predict the short and mid-term daily SST field, the experiments only investigate the prediction performance of different models for 1–10 days ahead. However, the model can also be utilized for longer-term SST field prediction, which may be explored in the future. Based on current results, longer trends can be inferred as well: (1) The LSTM (1-feature) model and LSTM (n-features) model will outperform the linear SVR model due to that the LSTM-based models are good at modelling long time dependencies with their recurrent network structure and gating mechanisms. (2) The proposed spatiotemporal deep learning model will consistently outperform the linear SVR model, the LSTM (1-feature) model, and the LSTM (n-features) model due to its strong abilities of modelling spatiotemporal dependencies of SST.

6. Conclusions

SST is a key physical parameter of the world ocean and plays an important role in the air-sea interactions. Changes of SST can have profound effects on the marine ecosystem, climate change, and may even lead to extreme weather events such as tropical storms, floods, and droughts. To achieve short and mid-term SST field prediction accurately and conveniently, a spatiotemporal deep learning model is proposed. The model can capture both the spatial and temporal correlations of time-series SST fields and make predictions in an end-to-end manner. The 36-year daily SST time series data derived from the AVHRR satellite sensors in a subarea of the East China Sea are used to train and test the model. Thorough comparisons with the persistence model (for the comparison of 1-day-ahead prediction only), the linear SVR model, the LSTM (1-feature) model, and the LSTM (n-features) model using different statistics including RMSE, MAPE, Pearson correlation coefficient and KDE from both the lead and holistic perspective demonstrates that the proposed spatiotemporal deep learning model consistently outperforms the other models for 1–10 days ahead prediction of SST fields. Besides, the proposed model can directly predict the SST fields while the linear SVR model and the LSTM (1-feature) model have to make predictions pixel by pixel which is inconvenient. The LSTM (n-features) model can directly predict the SST fields but it cannot capture the spatial correlations of SST which makes it perform poorly in the SST field prediction task. The results indicate that the proposed spatiotemporal deep learning model is promising in modelling the complex patterns of SST from historical observations, based on which accurate predictions of short and mid-term daily SST fields can be made.

By accurately predicting the daily resolution SST field using the proposed spatiotemporal deep learning model, the short and mid-term dynamics of fast changing SST can be caught. It can then be used to detect and assess the variability of the marine ecosystems based on the surface thermal front and intensity shown on the SST field maps. Besides, the predicted daily SST field can be incorporated into the hurricane models for anticipating tropical cyclone intensity (ESA) and utilized to track fast events, such as storms (Kuwano-Yoshida and Minobe, 2017; RSS; Zhou et al., 2015).

Though future SST fields are predicted based on the discovered spatiotemporal patterns of SST from historical SST time-series data, and good results have been achieved in the subarea of the East China Sea, some physical limitations still exist. As we know, the SST is closely coupled with the atmosphere-ocean exchange of heat and momentum (Emery et al., 2001). The future SST is affected not only by the past SST, but also by other hydrological and meteorological conditions. Therefore, the prediction performance is limited by just using the SST itself as

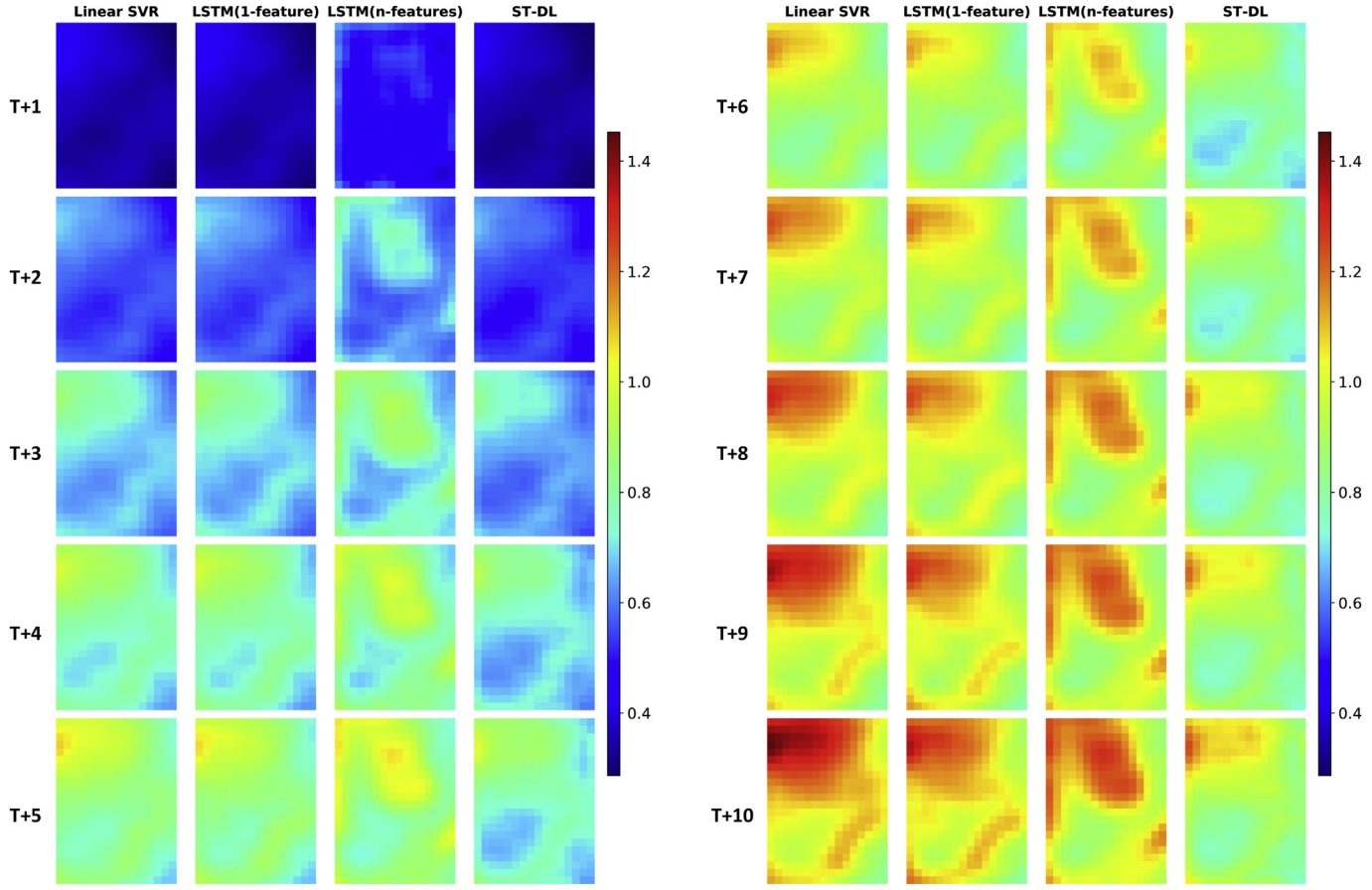

Fig. 10. Spatial distributions of RMSEs over the study area calculated on all testing samples for each prediction lead.

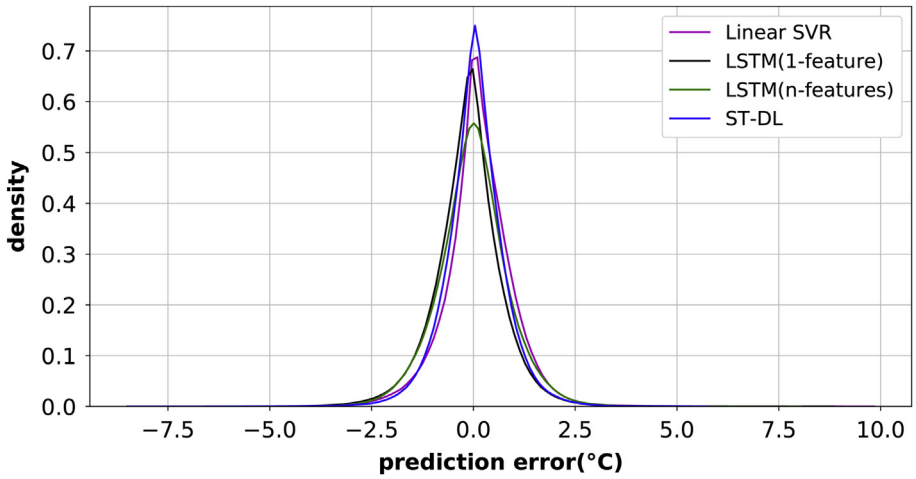

Fig. 11. Gaussian kernel density estimation of 1–10 days ahead SST field prediction errors holistically.

Table 5
Performance statistics calculated on all the predictions (1–10 days ahead as a whole) by each model.

Model	RMSE (°C)	MAPE (%)	*r*	p-value
Linear SVR	0.866	3.292	0.987	0
LSTM (1-feature)	0.850	3.193	0.988	0
LSTM (n-features)	0.880	3.523	0.987	0
ST-DL	**0.759**	**2.879**	**0.990**	0

the input prediction variable. However, the spatiotemporal deep learning model itself supports any number of input variables. In the future, if sufficient continuous and long time-series of these hydrological and meteorological data can be obtained, they can be easily incorporated into the model, and better prediction performance may be expected.

Although the spatiotemporal deep learning model is proposed to address the SST field prediction problem, it can also be applied to the prediction of other meteorological, environmental and atmospheric parameter fields, such as wind, soil moisture, atmospheric pollutants and so on (Ma et al., 2019; Wang et al., 2019; Zang et al., 2018). In the

future, high-resolution prediction of large extents of SST fields (e.g., ocean level) can be investigated based on the proposed spatiotemporal deep learning model and with the help of cyberGIS and supercomputing (Wang and Goodchild, 2019; Wright and Wang, 2011).

Declarations of interest

None.

Acknowledgment

This research was supported by the National Key Research and Development Program (No. 2018YFB2100500), the National Natural Science Foundation of China (NSFC) Program (No. 41890822), and the Creative Research Groups of Natural Science Foundation of Hubei Province of China (No.2016CFA003).

The authors would like to thank the following data and tool providers: NOAA/OAR/ESRL PSD for providing OISST-V2-AVHRR data (https://www.esrl.noaa.gov/psd/), Google for providing the open source machine learning framework TensorFlow (https://www.tensorflow.org/), Keras (https://keras.io/) for making it easier to use TensorFlow to build the spatiotemporal deep learning model, and Scikit-learn (http://scikit-learn.org/stable/) for providing the machine learning library based on which the experiments are conducted.

数据提供商

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation. USENIX Association, Savannah, GA, USA, pp. 265–283.

Aparna, S.G., D'Souza, S., Arjun, N.B., 2018. Prediction of daily sea surface temperature using artificial neural networks. Int. J. Remote Sens. 39 (12), 4214–4231.

Bouali, M., Sato, O.T., Polito, P.S., 2017. Temporal trends in sea surface temperature gradients in the South Atlantic Ocean. Remote Sens. Environ. 194, 100–114.

Cane, M.A., Clement, A.C., Kaplan, A., Kushnir, Y., Pozdnyakov, D., Seager, R., Zebiak, S.E., Murtugudde, R., 1997. Twentieth-century sea surface temperature trends. Science 275 (5302), 957–960.

Castro, S.L., Wick, G.A., Steele, M., 2016. Validation of satellite sea surface temperature analyses in the Beaufort Sea using UpTempO buoys. Remote Sens. Environ. 187, 458–475.

Chaidez, V., Dreano, D., Agusti, S., Duarte, C.M., Hoteit, I., 2017. Decadal trends in Red Sea maximum surface temperature. Sci. Rep. 7 (1), 8144.

Chollet, F. Keras, 2015. https://github.com/keras-team/keras.

Emery, W.J., Castro, S., Wick, G.A., Schluessel, P., Donlon, C., 2001. Estimating sea surface temperature from infrared satellite and in situ temperature data. Bull. Am. Meteorol. Soc. 82 (12), 2773–2786.

ESA sea surface temperature. https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-slstr/applications/maritime-monitoring/sea-surface-temperature, Accessed date: 12 June 2019.

Friedel, M.J., 2012. Data-driven modeling of surface temperature anomaly and solar activity trends. Environ. Model. Softw 37, 217–232.

Gentemann, C., Wentz, F.J., Smith, D., Chelton, D., 2000. Satellite measurements of sea surface temperature through clouds. Science 288 (5467), 847–850.

Herbert, T.D., Peterson, L.C., Lawrence, K.T., Liu, Z., 2010. Tropical ocean temperatures over the past 3.5 million years. Science 328 (5985), 1530–1534.

Kim, S., Hong, S., Joh, M., Song, S.-K., 2017. DeepRain: ConvLSTM Network for Precipitation Prediction Using Multichannel Radar Data. arXiv preprint arXiv:1711.02316.

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.

Kuwano-Yoshida, A., Minobe, S., 2017. Storm-track response to SST fronts in the north-western pacific region in an AGCM. J. Clim. 30, 1081–1102.

Lins, I.D., Araujo, M., das Chagas Moura, M., Silva, M.A., Droguett, E.L., 2013. Prediction of sea surface temperature in the tropical Atlantic by support vector machines. Comput. Stat. Data Anal. 61, 187–198.

Liu, Y., Fu, W., 2018. Assimilating high-resolution sea surface temperature data improves the ocean forecast potential in the Baltic Sea. Ocean Sci. 14, 525–541.

Ma, H., Zeng, J., Chen, N., Zhang, X., Cosh, M.H., Wang, W., 2019. Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: a comprehensive assessment using global ground-based observations. Remote Sens. Environ. 231, 111215.

Neetu, Sharma, R., Basu, S., Sarkar, A., Pal, P.K., 2011. Data-adaptive prediction of sea-surface temperature in the Arabian sea. IEEE Geosci. Remote Sens. Lett. 8 (1), 9–13.

Patil, K., Deo, M.C., 2017. Prediction of daily sea surface temperature using efficient neural networks. Ocean Dyn. 67 (3–4), 357–368.

Patil, K., Deo, M.C., Ravichandran, M., 2016. Prediction of sea surface temperature by combining numerical and neural techniques. J. Atmos. Ocean. Technol. 33 (8), 1715–1726.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Rauscher, S.A., Jiang, X., Steiner, A., Williams, A.P., Cai, D.M., McDowell, N.G., 2015. Sea surface temperature warming patterns and future vegetation change. J. Clim. 28 (20), 7943–7961.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily high-resolution-blended analyses for sea surface temperature. J. Clim. 20 (22), 5473–5496.

RSS Storm watch. http://www.remss.com/storm-watch/, Accessed date: 12 June 2019.

Salles, R., Mattos, P., Iorgulescu, A.-M.D., Bezerra, E., Lima, L., Ogasawara, E., 2016. Evaluating temporal aggregation for predicting the sea surface temperature of the Atlantic Ocean. Ecol. Inf. 36, 94–105.

Sautermeister, B., 2016. Deep Learning Approaches to Predict Future Frames in Videos. Department of Informatics. Technische Universität München, pp. 129.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., Woo, W.-C., 2015. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 28 (NIPS 2015), pp. 802–810.

Stockdale, T.N., Balmaseda, M.A., Vidard, A., 2006. Tropical Atlantic SST prediction with coupled ocean–atmosphere GCMs. J. Clim. 19 (23), 6047–6061.

Tanggang, F.T., Hsieh, W.W., Tang, B., 1998. Forecasting regional sea surface temperatures in the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors. J. Geophys. Res. Ocean. 103 (C4), 7511–7522.

Terrell, G.R., Scott, D.W., 1992. Variable kernel density estimation. Ann. Stat. 20 (3), 1236–1265.

Wang, S., Goodchild, M.F. (Eds.), 2019. CyberGIS for Geospatial Discovery and Innovation, 1st. 118 Springer, Dordrecht, Netherlands.

Wang, W., Mao, F., Zou, B., Guo, J., Wu, L., Pan, Z., Zang, L., 2019. Two-stage model for estimating the spatiotemporal distribution of hourly PM1.0 concentrations over central and east China. Sci. Total Environ. 675, 658–666.

Wright, D.J., Wang, S., 2011. The emergence of spatial cyberinfrastructure. Proc. Natl. Acad. Sci. 108 (14), 5488–5491.

Wu, A., Hsieh, W.W., Tang, B., 2006. Neural network forecasts of the tropical Pacific sea surface temperatures. Neural Netw. 19 (2), 145–154.

Yan, X., Ants, L., 2000. Forecasts of tropical Pacific SST and sea level using a Markov model. Geophys. Res. Lett. 27 (17), 2701–2704.

Yao, S.-L., Luo, J.-J., Huang, G., Wang, P., 2017. Distinct global warming rates tied to multiple ocean surface temperature changes. Nat. Clim. Change 7, 486–491.

Zang, L., Mao, F., Guo, J., Gong, W., Wang, W., Pan, Z., 2018. Estimating hourly PM_1 concentrations from Himawari-8 aerosol optical depth in China. Environ. Pollut. 241, 654–663.

Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017. Prediction of sea surface temperature using long short-term memory. IEEE Geosci. Remote Sens. Lett. 14 (10), 1745–1749.

Zhou, G., Latif, M., Greatbatch, R.J., Park, W., 2015. Atmospheric response to the North Pacific enabled by daily sea surface temperature variability. Geophys. Res. Lett. 42 (18), 7732–7739.

这篇文章稍微总结一下： 真的实现了ConvLSTM完成了海洋平面的温度预测， 主要借鉴的地方就是人家的数据集，实验设置，还用评估标准等这些