Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 12, No. 2, 2024

Sequence prediction of missing data of spatio-temporal series of marine environmental elements based on ST-ResNet model

Yuanfeng Zhong

College of Computer Science and Technology, Qingdao University, Qingdao 266071, China

Abstract: Spatial data of marine environmental elements are of great significance to the marine industry and marine safety, and they are also affected by external factors. This results in a new method for spatial and temporal sequence prediction of marine environmental elements based on the ST-ResNet model. Specifically, the temporal distance, period, and trend of marine regions are modelled by residual networks, and different residual sub-networks are used for each characteristic, in each sub-network, different weights are assigned to different branches and regions by simulating the changes of marine data, and the output is finally fitted to obtain the output. The ST-ResNet model was validated on the ERA5 dataset as an example, and compared with other prediction models, the ST-ResNet model achieved better results, which verified the effectiveness of the method in predicting the temporal and spatial sequences of marine environmental elements.

Keywords: Marine environmental elements; Deep learning; ST-ResNet model.

1. Introduction

Spatio-temporal series forecasting refers to the task of predicting serial data involving both temporal and spatial dimensions. This type of data can cover multiple moments and multiple spatial locations, and spatio-temporal series forecasting usually aims to predict values or trends at future moments and spatial locations [1]. Spatio-temporal series prediction of marine environmental elements is a complex and important task because spatio-temporal series prediction of marine environmental elements usually needs to take full account of the dynamics, complexity and diversity of the marine system.

The spatio-temporal sequences of marine environmental elements exhibit two distinctive features, namely spatio-temporal coupling and geographic correlation [2]. Spatio-temporal coupling refers to the fact that the characterisation data of marine environmental elements have both temporal and spatial attributes, and exhibit a gradually increasing spatio-temporal resolution with the continuous development of technology. This feature allows us to observe and understand the evolution of marine phenomena in time and space in greater detail. Geographic correlation, on the other hand, indicates the presence of significant proximity effects in oceanic processes, i.e., the existence of close interrelationships between neighbouring spatial locations. This correlation leads to complex patterns of heterogeneity and correlation in the marine environment at different spatial and temporal scales. Addressing how to simultaneously take into account the spatio-temporal coupling and geographic correlation properties of multidimensional ocean spatio-temporal processes is a key issue in achieving accurate ocean spatio-temporal predictions. This involves integrating data from different moments and locations to build more comprehensive and accurate spatio-temporal models. At the same time, in-depth studies of the interactions between different geographical locations in ocean processes are needed to better understand and predict the evolution of the marine environment. Such an integrated study is essential for improving the understanding of the spatial and temporal dynamics of the oceans, advancing the development of marine science and providing reliable decision support for future marine resource management and environmental protection.

In recent years, Adversarial Networks [3] and Variational Self-Encoders [4] are widely used in image recognition and image enhancement, etc. Typically, architectures based on Variational Self-Encoders and Adversarial Networks are coupled to multiple convolutional layers used to capture the semantic features of the input data and represent them in fewer dimensions. In contrast, the inherited stack self-encoders have achieved better results in the prediction of effective wave heights [5]. Long and short-term memory networks with encoder-decoder structure have achieved good results in performing sequence-to-sequence prediction [6]. In literature [7] the data representation of wind speed is obtained through an autoencoder and finally the prediction is done using long short-term memory network. Deep neural network is the process of mapping the extracted deep feature information into output values [8], many methods perform this mapping with simple convolutional or fully connected layers, which makes it difficult to resolve the deep features due to the large amount of spatio-temporal prediction output data. In this study, based on deep learning, we comprehensively analysed the characteristics of marine environmental element data, introduced a deep learning model based on ST-ResNet[9], and used convolutional network to capture the spatial information of marine environmental element data, so that it can accurately reflect the changes of marine dynamics to achieve the spatio-temporal sequence prediction of marine environmental elements. The deep learning-based spatio-temporal sequence models ConvLSTM[10], PredRNN[11], and MIM[12] models are selected as control experiments, and the spatial sequence prediction is performed by inputting the effective wave height, maximum wave period, and maximum wave direction data with a temporal resolution of 1 hour and a spatial resolution of 0.5° in the ERA5 marine environmental elements in the years 2019-2021.

2. ST-ResNet model

The ST-ResNet model mainly consists of the marine environmental element data analysis network and the external factor analysis network, in which the marine environmental element data analysis network consists of the trend sub-network, the cycle sub-network. The proximity sub-network is composed of the proximity sub-network. The structure of the model is shown in Figure 1:

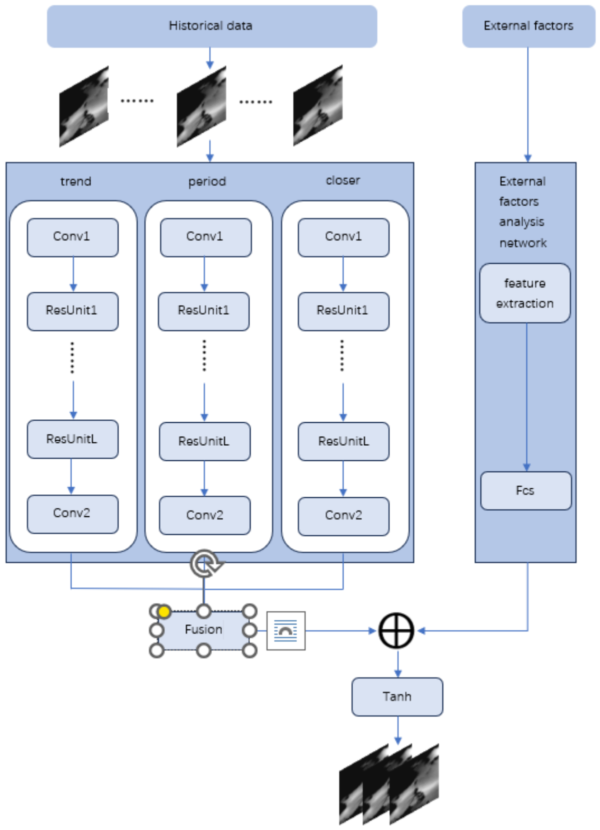


Fig. 1 ST-ResNet model

The transformation of each time interval of the marine environmental elements into a multi-channel image-like matrix. The time axis is divided into three segments, representing recent time, near history and distant history. The channel flow matrices in each time segment are then fed into the first three components to model the three temporal attributes mentioned above: proximity (closeness), period and trend, respectively. The first three components share the same network structure as the convolutional neural network. The second is a sequence of residual units, a structure that reflects the spatial dependence between near and far regions. In the external component, influences, such as wind speed affecting wave height, are extracted from an external dataset by feeding them into a two-layer fully connected neural network. The outputs of the first three components are fused into a parameter matrix-based x_Res, which assigns different weights to the results of different components in different regions. x_Res is further integrated with the output of the external component x_Ext. Finally, the aggregation is mapped to [-1,1] by a tanh activation function which yields faster convergence than the standard logistic function during backpropagation learning. the main advantages of the ST-ResNet model are:

(1) The ST-ResNet model uses a convolution-based residual network to model the spatial dependence between any two regions in the ocean in terms of distance and proximity, while ensuring that the prediction accuracy of the model is not affected by the deep structure of the neural network

(2) The temporal properties of marine environmental elements are grouped into four main categories, i.e., temporal proximity, period, trend, and external influences, and the ST-ResNet model uses three residual networks to model each of these properties.

(3) ST-ResNet dynamically aggregates the outputs of the above three networks, assigning different weights to different branches and regions.

(4) Better consideration of the influence of external factors.

2.1. Convolutional Neural Networks

Convolutional neural networks (CNN) is a class of feedforward neural networks (FNNs) containing convolutional computation with deep structure, which is one of the representative algorithms of deep learning, and it has a powerful ability to capture spatial structure information in a hierarchical manner. In the ocean region, the data size of neighbouring regions will affect each other, i.e. there is a correlation between neighbouring regions, and the underlying pattern is captured by convolutional neural network to obtain the dependence of data changes in neighbouring regions. From there, dependencies over longer distances or even over the entire sampling area are captured through multiple convolutions.

2.2. Residual module

Using activation functions (ReLU activation functions) and regularisation techniques, deep convolutional networks can compromise training and require a very deep network to capture very large region-wide dependencies. When the input size is 32×32 and the kernel size of the convolution is fixed at 3×3, more than 15 consecutive convolutional layers are required if region-wide dependencies (i.e., each node in the higher layers depends on all nodes in the input) are to be modelled. To address this problem, residual learning is used in the model, which has been shown to be very effective for training ultra-deep neural networks with more than 1000 layers.

When a direct convolution scheme is adopted, the size of the data frame remains constant after each convolutional layer, such that the network can theoretically be extended indefinitely. The goal of this paper is to predict the whole ocean area, so only a deeper network is needed to capture the dependencies in the whole ocean area, and the larger the size of the network, the more layers of the network are needed. The LOSS of the training set generally decreases with the increase of the number of network layers, but when the number of network layers is greater than a certain value, if the depth of the network is increased, the LOSS of the training set will increase, which is the phenomenon of gradient disappearance (explosion) that often occurs in the convolutional network. Adding residual units to the convolutional network can effectively solve the problem of network accuracy decrease (error increase) caused by the network depth. The principle is that if a convolutional network increases the number of layers by constant mapping, then the training error of the network after the number of layers increases will not be larger than the error when it does not increase the number of constant mapping layers. That is to say, after adding residual units to the network, the error will not become larger, and it is very likely to be reduced. The residual structure is shown in Figure 2.:

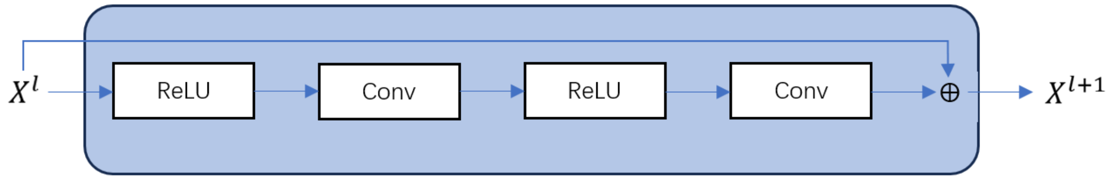


Fig. 2 Residual structure

In the ST-ResNet model, L residual cells are stacked on Conv1 as follows.

$$X_c^{(l+1)} = X_c^{(l)} + F(X_c^{(l)}; \theta_c^{(l)}) \qquad (1)$$

Where F is the residual function (i.e., "ReLU activation function + convolution") and θ^((l)) includes all learnable parameters in the 1st residual unit.

The cycle and trend components are constructed using the same operation. When the period is p, the period correlation sequence is [Xt-lp-p, Xt-(lp-1)-p, ---, Xt-p]. With the convolution operation and residual operation, the output of the period component is X(L+2)p. Meanwhile, the output of the trend component is X(L+2)q and the input is [Xt-lq-q, Xt-(lq-1)-q, ----, Xt-q]. where lq is the length of the trend correlation sequence and q is the trend span. Note that p and q are actually two different types of cycles. The p in the corresponding data is equal to an hourly interval, describing the daily periodicity, and the q is equal to a one-day interval, revealing the daily trend.

2.3. External component

The external factor analysis network consists of an input layer and two fully connected layers. The first fully connected layer receives the input data and performs the first step of feature fusion. The second layer is used to scale the output of the network to the size of the road network for subsequent fusion operations. Since elements of the marine environment can be affected by many complex external factors, such as effective wave height data being affected by wind speed. In Fig. 3, the data of effective wave height versus wind speed from 1 to 10 January 2021 is shown, from which it can be seen that the variation of effective wave height is affected by wind speed.

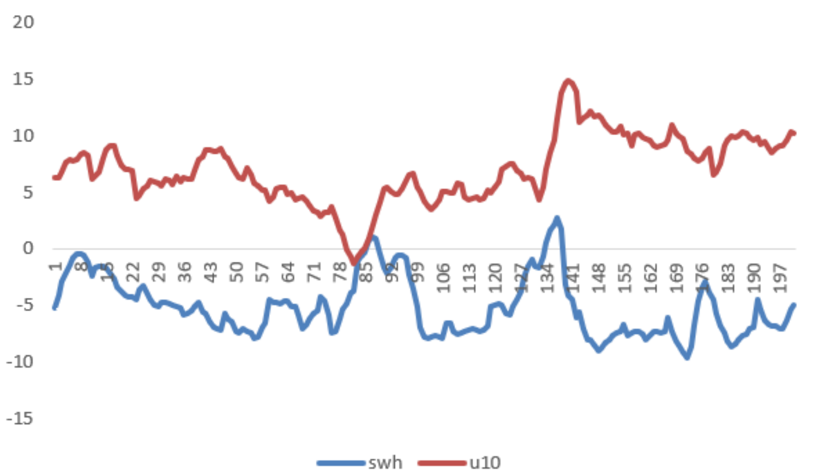


Fig. 3 Plot of effective wave height vs. wind speed

1.4 Parameter matrix based fusion

The formula for fusing subgrids (i.e., proximity, period, trend) is as follows.

$$x_{Res} = W_c * x_c^L + W_p * x_p^L + W_q * x_q^L \qquad (2)$$

where x_c^L is the input of the temporal proximity attribute on the L-layer of the network, and W_c is the weight matrix assigned to it; x_p^L is the input of the periodicity on the L-layer of the network, and W_p is the weight matrix assigned to it; and x_q^L is the input of the trending attribute on the L-layer of the network, and W_q is the weight matrix assigned to it.W_c, W_p, and W_q are all learnable parameters used to adjust for the variation effects of period, trend, and temporal proximity.

The outputs of the three subnets are summed with the outputs of the external components and mapped between [-1,1] using the tanh activation function. Define the predicted value x^_t for the tth time interval:

$$\hat{x}_t = \tanh(X_{Res} + X_{Ext}) \qquad (3)$$

Where tanh is the activation function that guarantees an output value between -1 and 1.

The mean square error between the optimised prediction matrix and the actual matrix is used as the loss function by training the network:

$$\mathcal{L}(\theta) = \|X_t - \hat{x}_t\|_2^2 \qquad (4)$$

where θ is a learnable parameter.

3. Empirical analysis

3.1. Experimental data

The measured data of marine environmental elements used in the study are based on the global wave parameter data based on buoys, altimeters, scatterometers, etc. provided by the European Centre for Medium-Range Weather Forecasts (ECMWF).The sea surface wind speed parameter of the ERA5 marine environmental reanalysis data has a latitude and longitude resolution of 0.25°, and the wave parameter data has a time resolution of 1 hour. The resolution of the sea surface wind speed parameter in the ERA5 ocean environment reanalysis data is 0.25° in latitude and longitude, the resolution of the wave parameter data is 0.5° in latitude and longitude, and the resolution of the time is 1 hour, and the data range is adopted from China's near-sea area (0°N~45°N, 105°E~135°E).

3.2. Experimental evaluation criteria

Root mean square error (RMSE,root mean square error) and mean absolute error (MAE,mean absolute error) were chosen to quantitatively evaluate the performance of the prediction model. Calculation formula:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \quad (5)$$

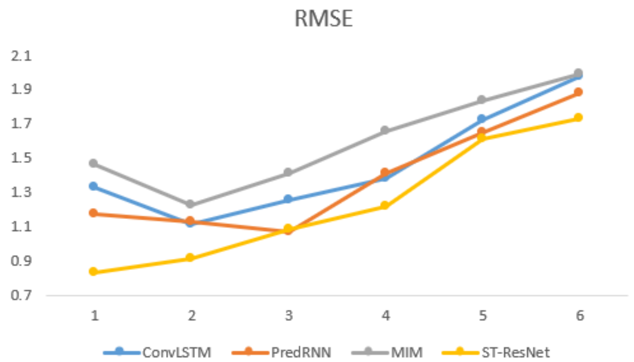
$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \quad (6)$$

Where m is the size of the test dataset; y_i denotes the ith true value of the marine environmental element; and (y_i)^ denotes the ith predicted value of the model.

4. Results and analyses

4.1. Predicted results for elements of the marine environment

In order to better demonstrate the model performance, the effective wave heights are fed into the model as input data. During the training phase, a loss function is used to evaluate the model performance and continuously optimise the internal parameters to achieve the best results. In order to increase the accuracy of model prediction for marine environmental element data, the root mean square error (RMSE) and the mean absolute error (MAE) are used as evaluation metrics for all models. The evaluation metrics of the comparison experiments at six consecutive moments are shown in Fig. 4, and in order to compare the evaluation metrics of each model more intuitively, the line graphs of MAE and RMSE at six consecutive prediction steps are shown in the figure. Overall, the prediction performance of the four models is better in the early stage than in the late stage due to the accumulation of continuous prediction errors. Comparing the four models, ST-ResNet based on ST has the lowest RMSE and MAE overall, and is significantly lower than ConvLSTM and MIM models, and also achieves a large advantage over the more effective PredRNN model. Especially when h=1, the MAE and RMSE of ST-ResNet model are 0.7634 and 0.8354 respectively, which are higher than 1.0792 and 1.1729 of PredRNN model, and achieve better results. Comparing the ConvLSTM with the PredRNN model, the PredRNN model achieves a greater advantage in feature extraction and obtains better prediction performance. When h=3, the MAE and RMSE of PredRNN model are significantly higher than ConvLSTM. For the MIM model when h=2, the prediction results are comparable to those of the ConvLSTM model, but when h gradually increases, the MAE and RMSE are much lower than those of the ST-ResNet model, which are 1.1083 and 1.2214, when h=4. It can be illustrated that the ST-ResNet model is able to efficiently extract more temporal and spatial information, and obtains better prediction performance.



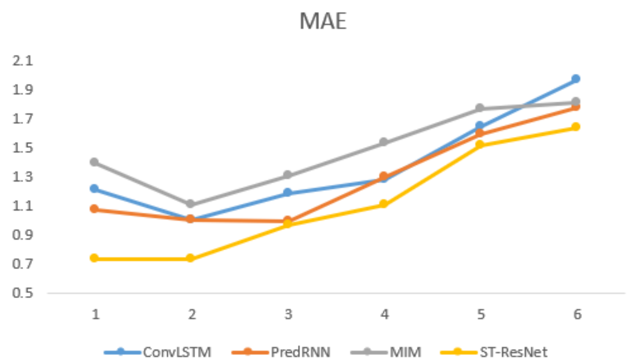


Fig. 4 Comparison of successive 6-step prediction algorithms

Table 1. Comparison of results of successive 6-step prediction algorithms

Model		1	2	3	4	5	6
ConvLSTM	MAE	1.213	1.0041	1.1885	1.2831	1.6472	1.9691
	RMSE	1.3326	1.1125	1.2541	1.3832	1.7254	1.9793
PredRNN	MAE	1.0792	1.0037	0.9957	1.3035	1.5946	1.7822
	RMSE	1.1729	1.1319	1.0678	1.4135	1.6507	1.8824
MIM	MAE	1.3993	1.1141	1.3119	1.5366	1.7726	1.8173
	RMSE	1.4622	1.2294	1.4125	1.661	1.8384	1.9926
ST-ResNet	MAR	0.7634	0.7356	0.9677	1.1083	1.5191	1.6382
	RMSE	0.8354	0.9103	1.0817	1.2214	1.6147	1.7312

4.2. Input Step Comparison

In order to verify the effects of different input steps on the model, the marine environmental element data of the past 6-48 time points were compared to predict the marine environmental element data of the future 12 time points, as shown in Table 2. The quantitative comparison results of RMSE and MAE at each time step with different time steps are given, and the prediction results are presented in the form of mean plus standard deviation.

The experimental results show that the prediction error keeps decreasing between 6 and 24 time steps, and the prediction error is very close between 18 and 36 time steps. Therefore, considering the accuracy and computational efficiency, the time step of 24 is the best choice.

Table 2. Average evaluation results for different input step sizes

Step/h	MMAE	MRMSE
6	0.8812	0.9935
12	0.8749	0.9613
18	0.8712	0.9598
24	**0.8658**	**0.9492**
48	0.8721	0.9507

5. Conclusion

The ST-ResNet model is introduced for the prediction of spatially sequential column data in marine environmental elements by combining deep learning techniques for the measured data of marine environmental elements. The method in this paper improves the efficiency and performance of the model by constructing proximity, period, and trend sub-networks, dynamically aggregating the outputs of the three networks, assigning different weights to different branches and regions, and further combining with external factors. By comparing the experimental results, the ST-ResNet model shows a more superior prediction effect compared with ConvLSTM, PredRNN, and MIM deep learning models.The ST-ResNet model becomes an effective method to deal with the missing data of the marine environmental elements, which provides a useful reference for the research of the marine environment and the research of data processing.

References

[1] Dimoulkas I, Mazidi P, Herre L; EEM 2017 Prediction Competition: Using Autoregressive Models for Wind Power Generation Prediction. The 14th International Conference on European Energy Markets, 2017, (1-6).

[2] Wang Jieruo, Yu Jing, Sun Huiying, et al. Spatial and temporal distribution characteristics of coordinated development between China's marine economy and resource environment [J/OL]. Notice on Marine Lakes and Marshes, 1-9 [2024-03-27].

[3] Goodfellow I, Pouget Abadie J, Mirza M, etc. Generate adversarial networks [J]. ACM Communication, 2020, 63 (11): 139-144.

[4] Kingma D P, Welling M. Automatic Encoding Variational Bayes [C]//International Conference on Learning Representation. Isaka, NYarXivorg, 2013.

[5] Lu Xiaomin, Liu Fan, Cai Lihua, et al Integrated stack autoencoder and XGBoost for deep learning wave effective wave height prediction model J Computer and Modernization, 2021 (04): 32.

[6] Srivastava N, Mansimov E, Salakhudinov R. Unsupervised Learning Using LSTM Video Representation: International Conference on Machine Learning [Z]. two thousand and fifteen

[7] Zhu Q, Chen, Zhu, etc. Temporal and Spatial Wind Speed Prediction: A Deep Learning Method [J]. Energy, 2018, 11 (4): 705.

[8] Bengio Y, Courville A, Vincent P. Representative Learning: A New Perspective and Review [J]. IEEE Journal of Pattern Analysis and Machine Intelligence, 2013, 35 (8): 1798-1828.

[9] Zhang, Zheng Yong, Qi D. Deep spatiotemporal residual network for predicting urban crowd flow. Proceedings of the 31st AAAI Artificial Intelligence Conference. 2017, (1655-1661).

[10] Shi, Chen, Wang. Convolutional LSTM network: a machine learning method for precipitation forecasting. Proceedings of the 28th International Conference on Neural Information Processing Systems, Volume 1. 2015, 11 (802-810).

[11] Wang, Long, Wang, et al. Recurrent neural network predictive learning based on spatiotemporal LSTM [J]. Progress in Neural Information Processing Systems, 2017, 30.

[12] Wang, Zhang, Zhu, et al. Memory in Memory: A Predictive Neural Network for Learning High Order Non stationarity from Spatiotemporal Dynamics [C]//IEEE/CFF Conference on Computer Vision and Pattern Recognition Proceedings. 2019: 9154-9162.