Study on prediction of SST and SSS in Southern Ocean by multi-layers ConvLSTM model（多層ConvLSTMモデルによる南極海の海面水温および海面塩分の予測に関する研究）

Master's Thesis

STUDY ON PREDICTION OF SST AND SSS IN SOUTHERN OCEAN BY MULTI-LAYERS CONVLSTM MODEL

March 2021

Graduate School of Marine Science and Technology
Tokyo University of Marine Science and Technology
Master's Course of Marine Resources and Environment
WEI ZITONG

Table of Contents

List of Figures and Tables

Chapter 1. Introduction

Recently, the marine environmental factors prediction has drawn considerable attention in various ocean-related fields such as global warming, fisheries, and oceanic environmental protection. Two types of approaches are used to predict the marine environmental factors. The first one is the physics-based numerical model. The second one is to discover the marine environmental factors variation through data rather than a physical constraint.

Salinity has been regarded as an important variable of the ocean that influences the hydrological cycle, ocean circulation, and climate changes. Variations in sea surface salinity (SSS) are key markers of changes in the surface freshwater fluxes, which provide useful information for understanding certain aspects of the hydrological cycle, including evaporation, precipitation, river runoff, and melting ice (Helber et al.,2010; Li et al., 2016; Qin et al., 2015; Rao, 2003; Wang & Zhang, 2012). On the other hand, salinity, along with ocean temperature, is required to compute ocean density and plays a significant role in modulating the climate (Ballabrera-Poy et al., 2002; Bao et al., 2018; de Boyer Montégut et al., 2007; Zhu et al., 2014). Salinity has also been used to track water mass movement and vertical mixing processes between the surface and subsurface layers (Durack et al., 2012). Therefore, it is critical to measure and analyze salinity variations and understand their potential relationship with the global hydrological cycle and the thermohaline circulation.

Some previous studies used machine learning such as support vector machine (SVM) for marine environmental factors prediction; however, deep learning algorithms recently are more accurate than the SVM when enough training data are available Recurrent Neural Network (RNN) are gradually being used to solve time series problems in speech recognition and machine translation. The long and short-term memory (LSTM) architecture improved on the RNN to overcome the RNN's memory shortcomings in long-term sequences. Srivastava et al. (2015) proposed an encoder–decoder LSTM model reconstructed and predicted to support a sequence of video data. Zhang et al. (2017) attempted the LSTM models to solve the marine environmental factors prediction problems. The existing LSTM models, only consider the time series of some isolated points and their spatial linkage was not considered. Lecun et al. (1998) officially proposed a convolutional neural network (CNN) architecture. Using the CNN algorithm, Krizhevsky et al. (2012) proposed Alex Net in image recognition in 2012. In

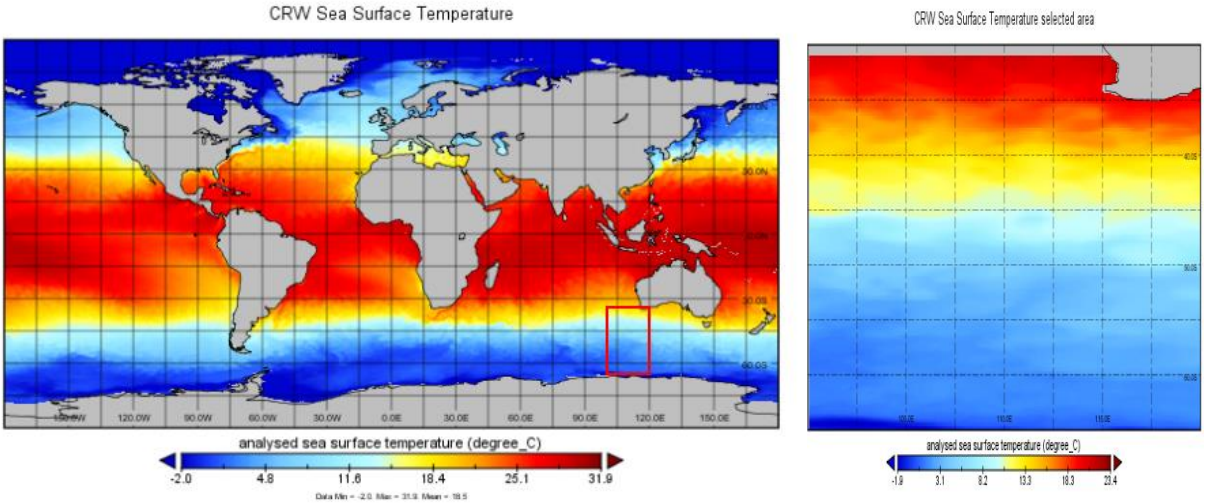
2015, Daverat et al. (2011) applied the CNN to identify the oceanfront. A CNN can extract the local spatial information from the image by sliding the convolution kernels. Therefore, it is suitable for extracting the features of spatial variation.

Based on the previous results, Shi et al. (2015) proposed a convolutional LSTM (ConvLSTM) for precipitation prediction. Yang et al. (2018) proposed a convolutional FC-LSTM (CFCC-LSTM) for marine environmental factors prediction. These studies only predict the temperature of sea surface, but other marine environmental factors are also important. Here, in this study proposes a model of multi-layers ConvLSTM (M-ConvLSTM) to predict Sea surface Temperature (SST) and salinity (SSS), comprising CNN, LSTM, and multi-layers stacking to consider the sea surface variations. The satellite data are used as label values to evaluate the prediction accuracy. Since the satellite time series data used in this study meet the prediction requirements of the time series model, multi-layers ConvLSTM model is applicable and is expected to possible to discover new variation of marine environmental factors through data rather than a physical constraint.

Chapter 2. Data

The NOAA Coral Reef Watch (CRW) global Sea Surface Temperature monthly average data set ranging from 2015 to 2019 (60 months) is collected from NOAA Ocean Watch Central Pacific Node (https://oceanwatch.pifsc.noaa.gov). In this study, the training of the ConvLSTM model, a typical rectangular area (20◦ ×35◦) of (E100◦–120◦, S30◦–65◦) in the Antarctic Ocean is selected here. Figure 1 shows the CRW temperature of sea surface (SST, in the depth of 0 m) in global and selected areas in January of 2019. And this area is the same as the region of the observations taken by T/R V Umitaka-Maru, as one of the Japanese Antarctic Research Expedition parties between 2011 and 2020. Figure 2 shows all the stations (bule dot) distribution for “Umitaka-Maru” in 2020.

The CTD data of T/R V Umitaka-Maru is verified and calibrated by AUTOSAL for the collected seawater, and the salinity is corresponding to WOCE accuracy. For the sea level salinity value of CTD, we used the salinity data obtained at sea level when recovering the CTD.



(a) (b)

Figure 1 CRW temperature of sea surface (0 m) in January 2019. (a) Global area. (b) Selected area.

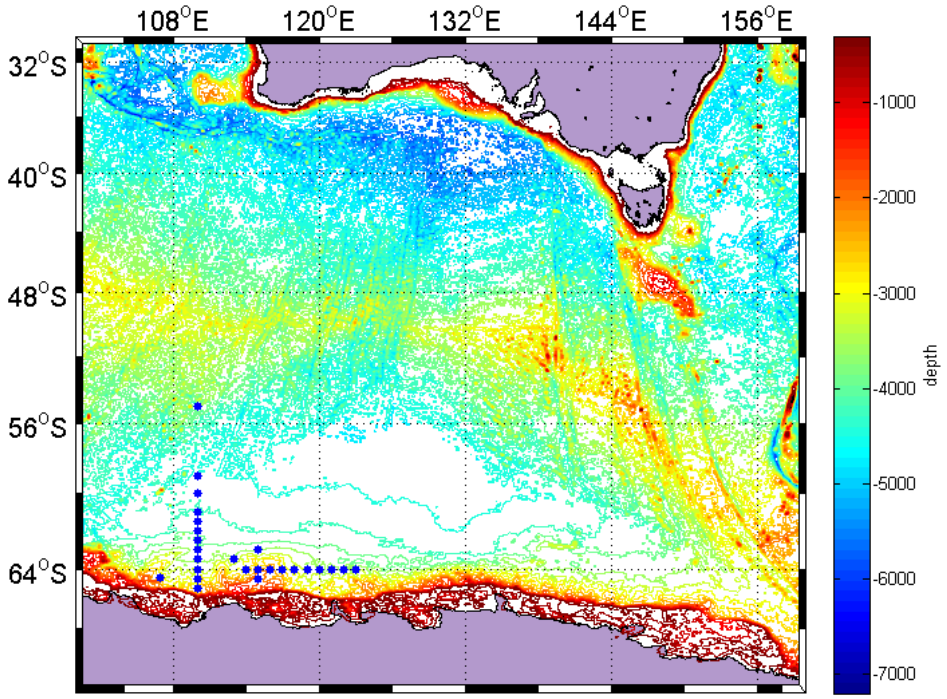


Figure 2 Stations distribution for TR/V Umitaka-Maru in 2020

The CRW operational near-real-time nighttime SST analysis product is a gap-filled, twice-weekly composite of global satellite nighttime AVHRR SST fields at 50 km resolution. In this study, Nighttime-only satellite SST observations are used to eliminate the issue of surface warming during the day and to avoid contamination from solar glare. Compared with daytime SST and day-night blended SST, nighttime SST provides more conservative and stable estimates in subsurface ocean condition. Nighttime SST also compare favorably with in situ SST at one meter depth (Montgomery and Strong, 1995), for example at times when corals bleaching occurs. The 50 km resolution data are derived by averaging multiple temperature observations (weighted by distance from pixel center, conditionally out to a maximum of 150 km), which are based on 4 km AVHRR Global Area Coverage (GAC) SST acquired daily (Skirving et al., 2006 b). In locations where cloudy conditions remain for extended periods of time, resulting in no new temperature observations, the most-recent high-quality SST value persists.

The BEC L3 products have been obtained from Operational products V2.0, which are generated from Level 2 v622 salinity (SMOS-BEC Team, 2017). The main features of SMOS is shown in Table 1. Two products are considered here: the monthly binned L3 map (weighted average) and the 9-day averaging objective analyzed L3 maps. The

monthly binned products are averaged at 1° × 1° spatial resolution, and the 9-day running objective analyzed L3 maps are provided daily at 0.25° × 0.25° spatial resolution. The time range of these products is January 2011 to recent.

The SMOS sea surface salinity 3-day data set ranging from 2010 to 2019 (120 months) is collected (https://oceanwatch.pifsc.noaa.gov/erddap/griddap/smos_3day). Training of the ConvLSTM model, a typical rectangular area (20◦ ×35◦) of (E100◦–120◦, S30◦–65◦) in the Antarctic Ocean is selected here. The figure 3 shows the SMOS salinity of sea surface (SSS, in the depth of 0 m) in global and selected areas in January of 2019.

Table 1.Main features of SMOS

Features	SMOS
Radiometer center frequency	
Antenna type	69 patch antennas (0.165 m)
Radiometer units	72 receivers
Scan geometry	Conically scanning: incidence angles varies from 0 to about 65°
Revisit time	3 days
Sampling time for 1 footprint	1.2 s
Calibration accuracy requirement	0.2–1 K
Sensors	A passive system (MIRAS, Y-shape array)
Swath	1,000 km
Spatial resolution	40 km
Stokes parameters	4 Stokes
Surface roughness correction	Use ancillary wind field (ECMWF)
The dielectric models	Klein and Swift (2003)
The ancillary SSS data	World Ocean Atlas 2009
Calibration scenes	A region in the Pacific Ocean

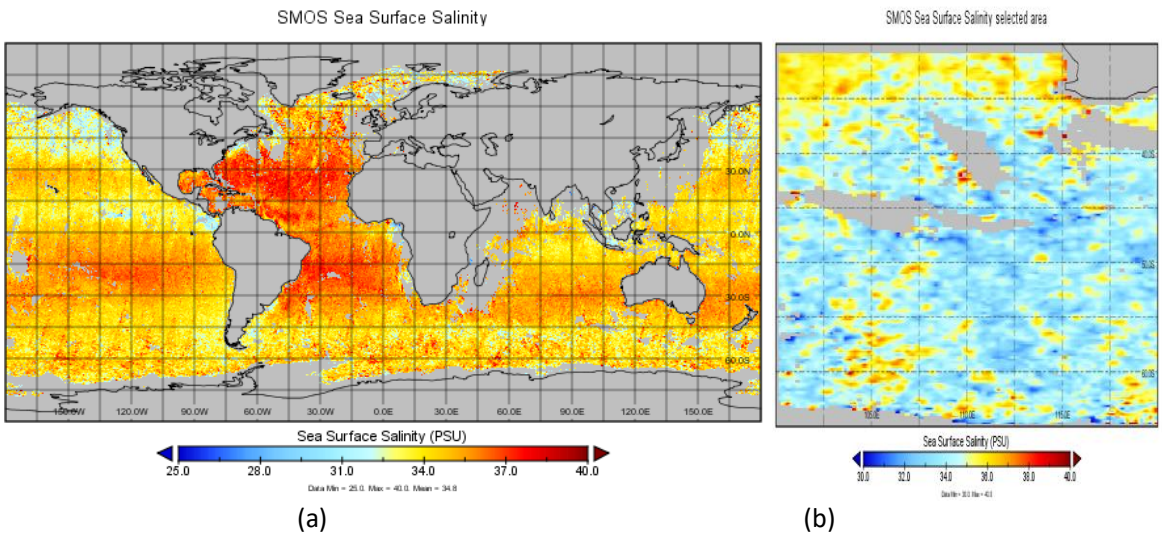


Figure 3 SMOS salinity of sea surface (0 m) in January 2019. (a) Global area. (b) Selected area.

We used normalized and standardized data for the neural network algorithm. The standardized approach in this letter is as follows:

$$x' = (x - \mu)/\sigma \qquad (1)$$

where the raw data are subtracted from the temperature mean (μ) and then divided by the variance (σ). The temperature values after standardization are normalized between – 1 and +1 to facilitate the network training.

Chapter 3. Multi-layers ConvLSTM Model

3.1 M-ConvLSTM Model

In artificial intelligence deep learning networks, traditional Recurrent Neural Network(RNN) and Long and Short-Term Memory(LSTM) are commonly used prediction models; however, for data with spatial-temporal variations, it is found that traditional prediction models lose spatial variation information. The Convolutional LSTM Network (ConvLSTM) model combines the advantages of Convolutional LSTM to extract graphic changes and LSTM to predict future data, which not only gives consideration to the temporal and spatial variation characteristics of parameters but also well predicts future data changes. In this paper, a neural network model was established with multi-layers ConvLSTM network as shown in Figure 4.

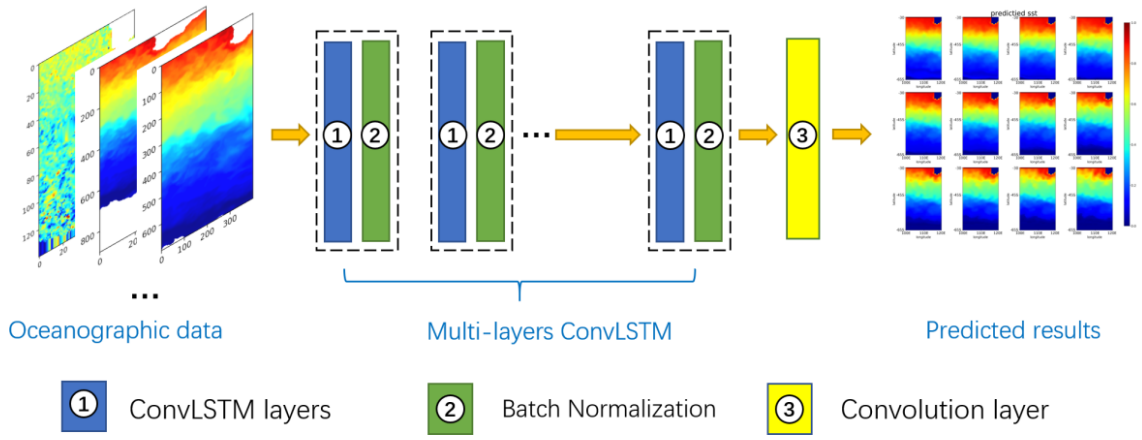


Figure 4 Schematic of the main process for ConvLSTM model

In this study, sea surface temperature data were used in the production of training samples. Data need to be normalized before input model is trained. We used the SST data from 2015 to 2019 as the training set and test set (60 batches of data in total), and the data from June to November 2019 as the verification set (6 batches of data). We input the superposition data of 6 months as the sample, and the SST of the next month as the sample label. Since the T/R V Umitaka-Maru observation were made at area as shown in Figure 2, the study area is mainly in the range of (30-65°S, 100-120°E), so we took this area as the study area.

3.2 ConvLSTM Layer

In this paper, the long-term time series data of marine environmental factors are used to predict the future situation. Since the objective laws of marine environmental factors in the past, present and future of selected sea areas remain basically unchanged, the interpretation of historical data can continue to the future. And the marine environment changes also meet the gradual change pattern. To sum up, the satellite time series data used in this paper meet the prediction requirements of the time series model.

The classical time series prediction models are RNN and LSTM. Long and short time memory network was proposed in 1997, which added two structures: cell state and gate structure on the basis of RNN. The cell state is used to record the time series characteristics, and the gate structure is used to control the information retention of the previous moment. The structure overcomes the difficulties of RNN training and gradient disappearance. Figure 5 show network structure.

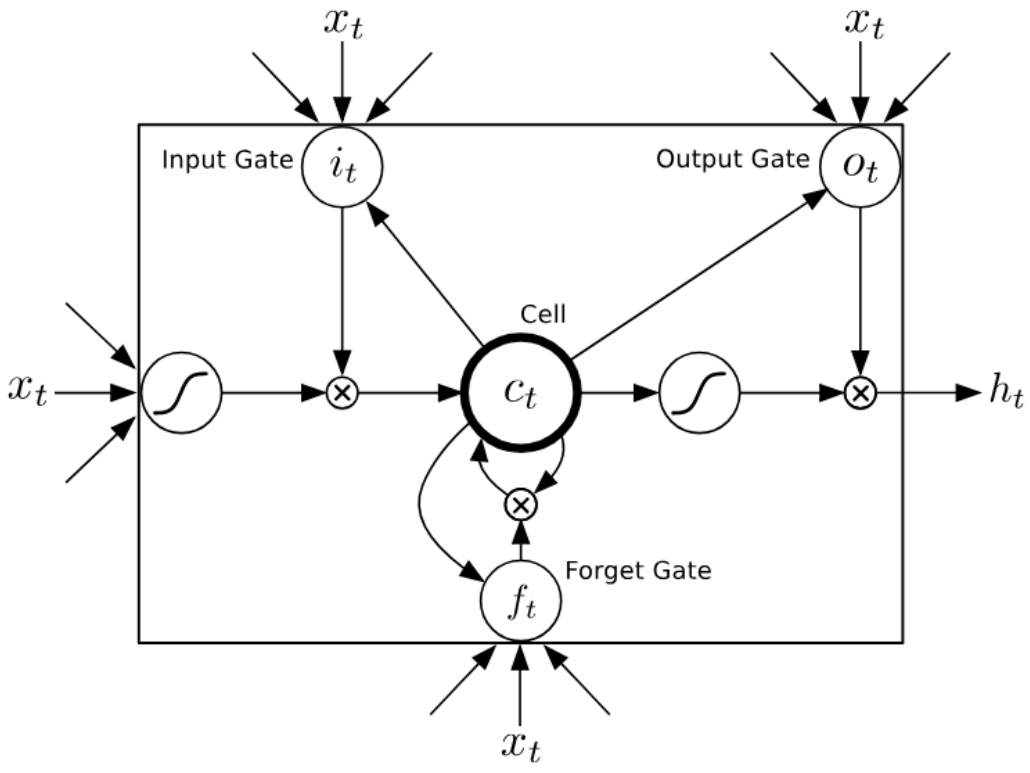


Figure 5 Schematic of the cell of LSTM

The main parameter formula of LSTM is as follows ("∘" representing Hadamard product)

$$i_t = \sigma(W_{wi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \quad (2)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \quad (3)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (4)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o) \quad (5)$$

$$h_t = o_t \circ \tanh(c_t) \quad (6)$$

Among them, *i*, *f* and *o* represent input gate, forget gate and output gate respectively; *x* represents network input; *W* represents network parameter, *b* represents offset; *h* represents hidden state. σ represents the sigmoid function.

Later, ConvLSTM added a convolution operation on the basis of the LSTM network, which can use convolution calculation to extract spatial features, and finally can extract and predict temporal and spatial features of time series.

ConvLSTM also uses a gate structure similar to LSTM to control the flow of information. The expressions of input gate, cell state, forget gate, and output gate are similar. Difference between LSTM and ConvLSTM is that the information input in each gate structure is replaced by a dot product with a convolution operation, and the cell state update still keeps the dot product unchanged. The weight part is put into the convolution kernel, and part is put into the loop kernel of the loop layer. The convolution kernel, W_i, W_f and W_o, is sliding the window in the spatial two-dimensional matrix, the convolution result obtained, according to the gate structure for input, update the cell state, forget, and output respectively.

$$i_t = \sigma(W_{wi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \quad (8)$$

$$c_t = f_t . c_{t-1} + i_t \circ tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \quad (9)$$

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o) \quad (10)$$

$$h_t = o_t . \tanh(c_t) \quad (11)$$

Among them, i_t is the input gate, c_t is the cell state, f_t is the forget gate, o_t is the output gate, h_t is the hidden layer output, and σ is the sigmoid function.

The experiments show that the ConvLSTM network can better capture the spatio-temporal correlation, and it is always better than the traditional fully connected LSTM network.

3.3 Model structure

The figure 6 shows a flow chart of a ConvLSTM and Convolutional Neural Network (CNN) combined network model for forecasting Marine surface environmental factors. With the output of three ConvLSTM and two CNN layers, the model has a very good performance in the prediction of large-scale seawater temperature.

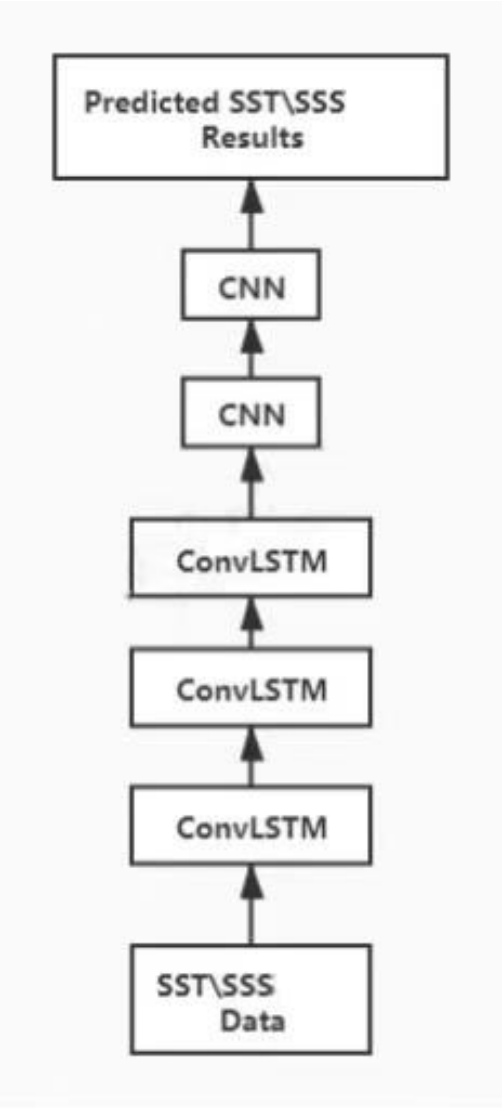


Figure 6 ConvLSTM model flow chart

3.4 Loss function

In this paper, *MSE* loss function is used to test the accuracy of prediction. In addition, in order to judge the performance of the neural network model, we added that mean absolute error (*MAE*) and Dice coefficient were used as parameters to evaluate the performance of the model, and the formula was as follows:

$$MAE = \frac{1}{n} \times \sum_{t=1}^{n} \left| X_t - Y_t \right| \tag{12}$$

$$MSE = \frac{1}{n} \times \sum_{t=1}^{n} \left(X_t - Y_t \right)^2 \tag{13}$$

Where: $X_t$is the predicted value; Y_t is the true value. The smaller MEA is, the more accurate model prediction is. The smaller the RMSE means the more stable the model prediction performance. The Dice coefficient, named after Lee Raymond Dice, is a set similarity measurement function. It is usually used to calculate the similarity between two samples (values in the range of [0, 1]). The definition of Dice coefficient, *Dice_coef*, is as follows:

$$Dice\_coef = \frac{2|x \cap y|}{|x| + |y|} = \frac{2TP}{FP + 2TP + FN} \tag{14}$$

Where x is the input sample and y is the output sample. And loss of Dice coefficient, *Dice_coef_loss*, is as follows:

$$Dice\_coef\_loss = 1 - Dice\_coef\_loss \tag{15}$$

Chapter 4. SST Predicted Results and Analysis

4.1 Predicted Results

Figure 7 and the Figure 8 are the comparison charts of SST forecast in the second half of 2019. Figure 7 is the real SST label data, and Figure 8 is the prediction result corresponding to the model. The Softmax function is used in the forecast result output, so the output result is a probability value (between [0,1]). We can find from the figure that the prediction results of the model are looks consistent with the actual verification.

The basic characteristics of the water temperature distribution are well reflected in the predicted values, but there is a tendency for the north-south temperature gradient to appear more strongly. Also, although the front area can be seen to undulate, the small vortex structure seems to be smoothed out considerably.

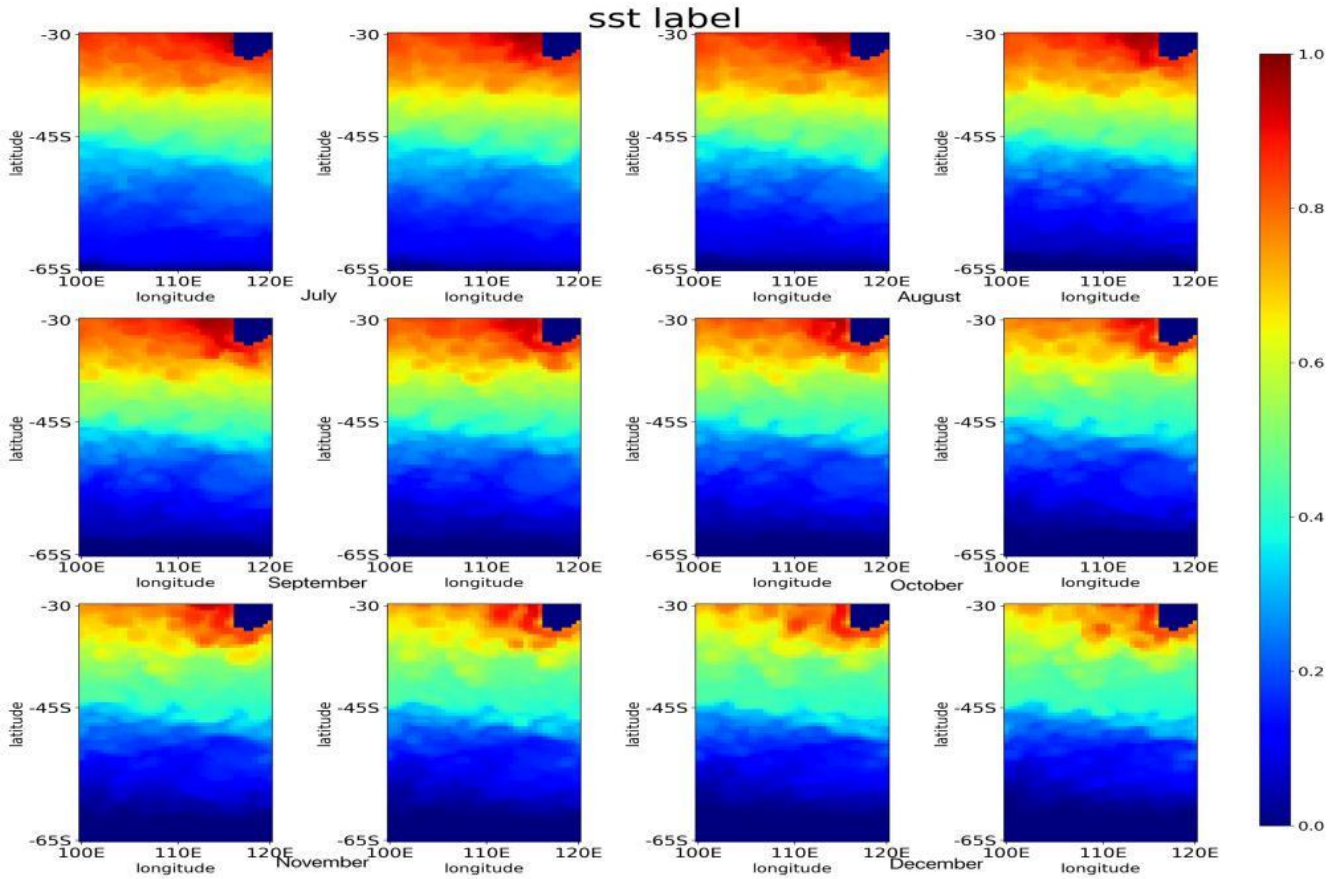


Figure 7 Real SST label data.

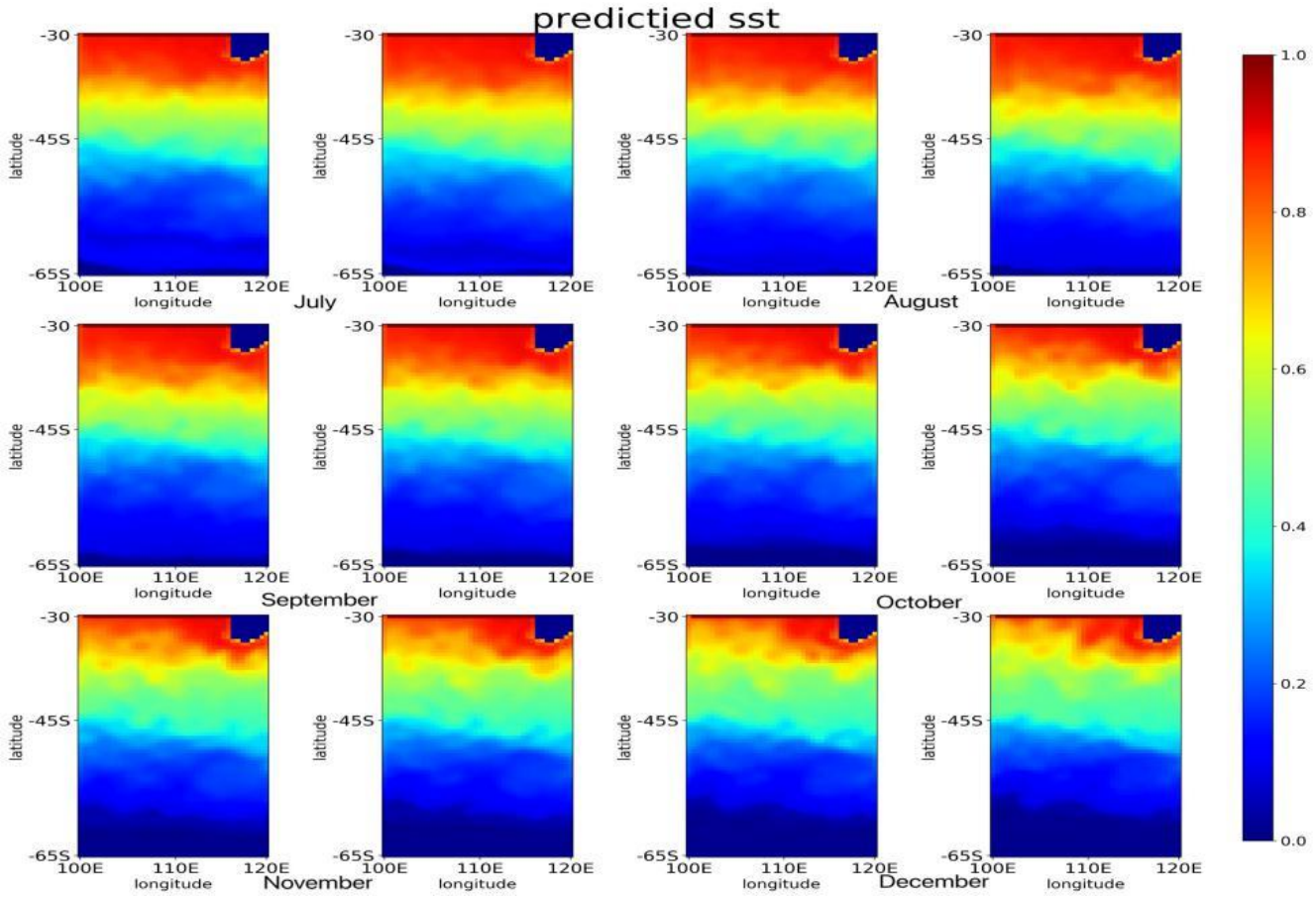


Figure 8 Prediction result of SST.

4.2 Evaluation of models

Sea surface temperature data were used in the production of training samples. Data need to be normalized before input model is trained. We used the SST data from 2015 to 2019 as the training set and test set (60 batches of data in total), and the data from June to November 2019 as the verification set (6 batches of data). We input the superposition data of 6 months as the sample, and the SST of the next month as the sample label. The study area is mainly in the range of (30-65°S, 100-120°E), so we take this area as the test area. The input sample is the tensor of (2,6,60,60,5), where 2 is Batch_Size,6 is the time mark of the input month, (60,60) is the spatial resolution of 0.5°×0.5°, and the height and width of the study area. Enter the SST data as the sample label. The experimental results obtained by the network model are shown in the table 2.

Table.2 Model parameters and error comparison

No.	Spatial resolution	Time resolution	MAE	MSE	Dice_coef(%)
1	0.5°×0.5°	15d	0.0366	0.0022	99.48

Month	July	August	September	October	November	December
Parameter	Month1	Month 2	Month 3	Month 4	Month 5	Month 6
MSE	0.001062	0.001527	0.002428	0.002984	0.001865	0.001921
MAE	0.02577	0.03078	0.04157	0.04491	0.03386	0.03139
R^2	0.9848	0.9802	0.9694	0.9582	0.9720	0.9695
Regression variance	0.9850	0.9908	0.9897	0.9859	0.9880	0.9844

Figure 9 shows the change of the MSE loss function corresponding to the training data and the value data during the SST model training process. MSE represents the accuracy of training in the model. The accuracy of this model on the training data for the first 5 rounds of training increased rapidly, and there was no major fluctuation in the next 40 rounds of training. After 50 rounds of training, the MSE kept stable. The accuracy of the training data increasing steadily with the increase of the training numbers, and the MSE of this model approaches 0.

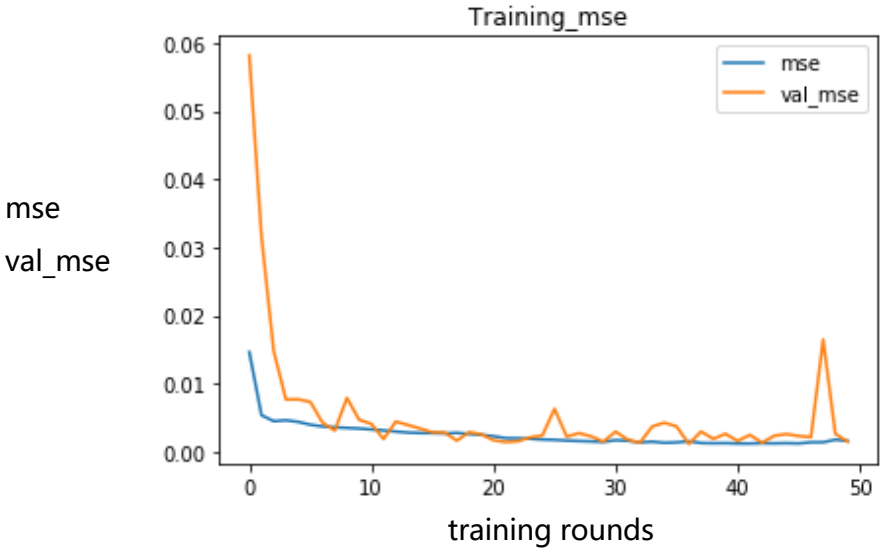


Figure 9 Accuracy of SST model training

Chapter 5. Application to SSS Prediction

Predicted results and evaluation of model：Figure 10 shows the SSS prediction results in 2020. The value of salinity anomaly is high near the northern boundary and low at the southern boundary, but is almost constant elsewhere. Figure 11 shows the change of the MSE loss function corresponding to the training data and the value data during the SSS model training process. MSE represents the accuracy of training in the model. It shows the initial stage of training. The accuracy of the training results was not improved. And the accuracy of the validation set fluctuates repeatedly. Eventually, the time gradient of salinity became extremely large, and as a result, the predicted value of salinity seemed to be almost constant.

As for the salinity obtained by satellite, as shown in Fig. 3 (b), the high salinity is seen as a patch, or is missing in a wide area. Area of such missing data changes from month to month, resulting in a large time gradient of salinity. Since it is necessary to filter error data and missing data from salinity data and utilize them, the method will be discussed later. It is difficult to predict because of the excessive noise of data in this area, which leads to the disappearance of gradient.

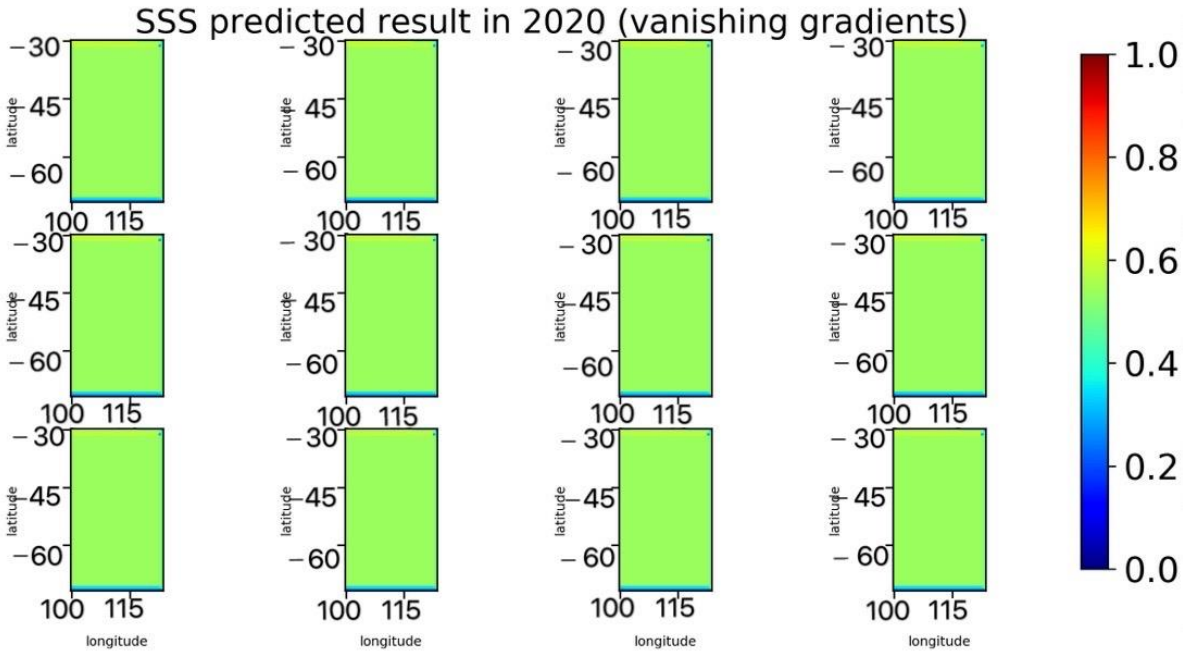


Figure 10 Prediction result of SSS

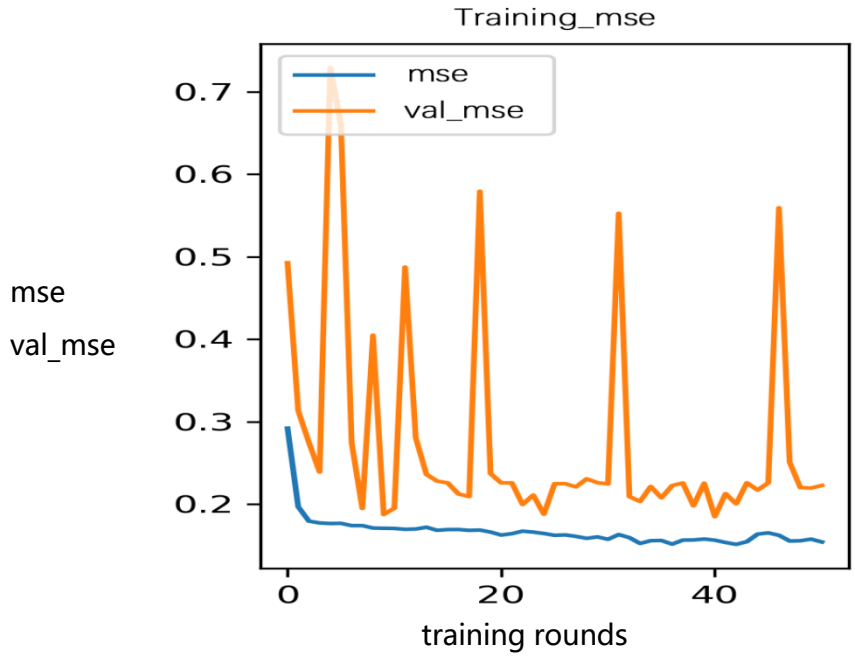


Figure 11 Accuracy of SST model training

Chapter 6. Discussion

In this study, we have applied the M-ConvLSTM model to satellite data of water temperature and salinity and tried to estimate the predicted values. As a result, very good prediction results were obtained for the water temperature data set, but didn't for the salinity. Salinity value did not converge, so correct results were not obtained. To clarify the difference of model application for SST and SSS in this area, water property obtained from the satellite data and in-situ CTD data were compared and analyzed.

Firstly, the data of these two were matched, and then the least square linear regression method was used to analyze their correlation (Figure 12). The in-situ CTD data came from hydrological observations taken by T/R V Umitaka-Maru between 2011 and 2020.

The square of correlation coefficient, R^2, of SST between CRW satellite and CTD data was 0.81 (Figure 13). Accuracy of SST obtained by satellite showed good quality. On the other hand, square of correlation coefficient, R^2, between SSS from SMOS satellite and CTD data was 0.23. Variance of SSS from SMOS is larger in low salinity range of CTD data. Thus, one of the reasons why the SST data was well predicted by M-ConvLSTM is high accuracy for the satellite SST data itself.

Next, we have compared satellite data and CTD data from the TS relation. Figure 15 is SST and SSS diagram. The red dot shows the data from CTD and the blue dot shows CRW and SMOS data.

It can be seen that the TS of the CTD changes relatively monotonously, while the satellite data scattered widely especially in the low temperature range. This indicates that satellite salinity data is not accurate at low temperature range. Thus, the second reason is that the satellite data varies too much in the low temperature range. Another cause is that SSS obtained by satellite has more missing data as mentioned above than SST.

It is good not to use data that includes error data, but randomly appearance of missing data may give a lot of damage to the M-ConvLSTM calculation. Therefore, as a proposal in this study, we firstly conduct a water mass test and determine TS range for interested area by using CTD data, use satellite data within the TS range, and omit other data. Secondly, the salinity at the missing value is interpolated from the salinity estimated from the water temperature using the TS relationship. Finally, the salinity data is smoothed using a box filter.

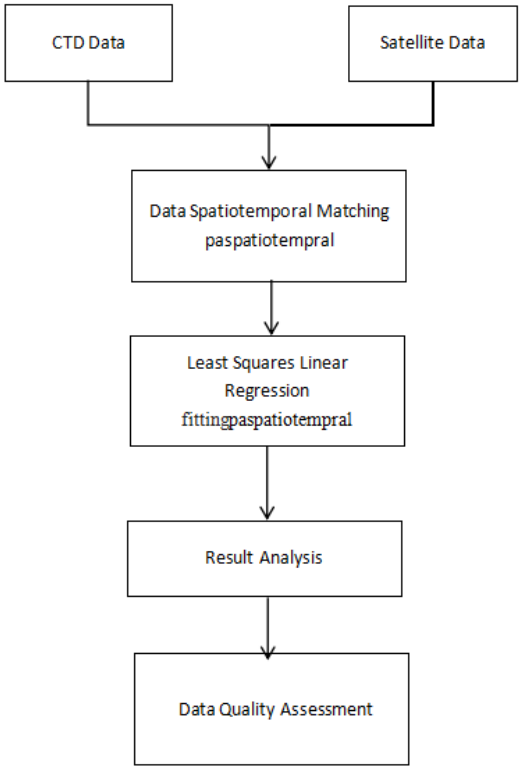


Figure 12 Satellite and CTD data comparison flow chart

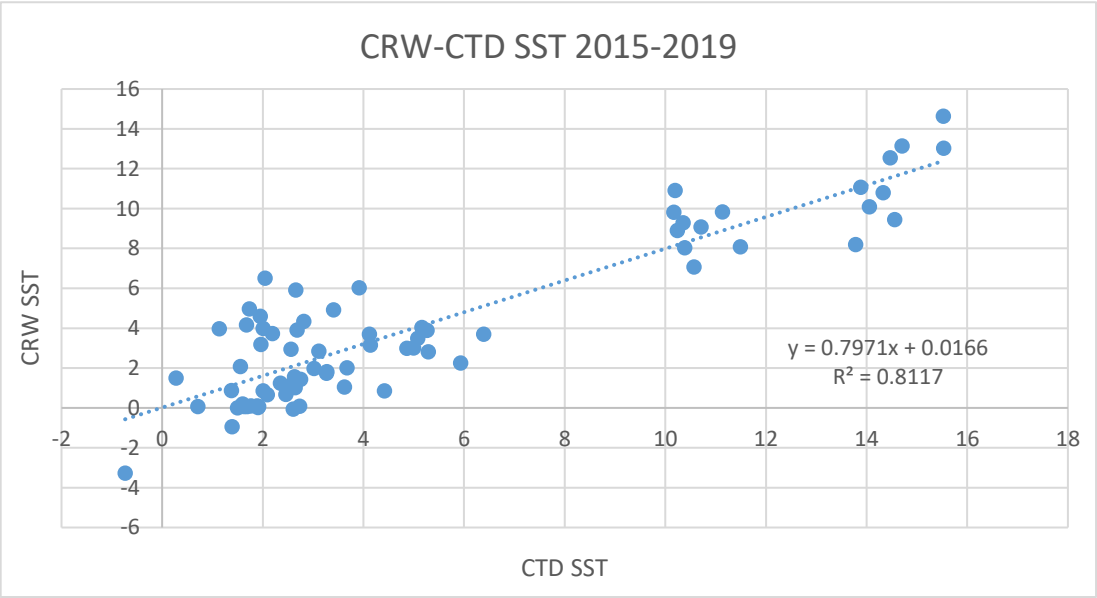


Figure 13 SST satellite and CTD comparison

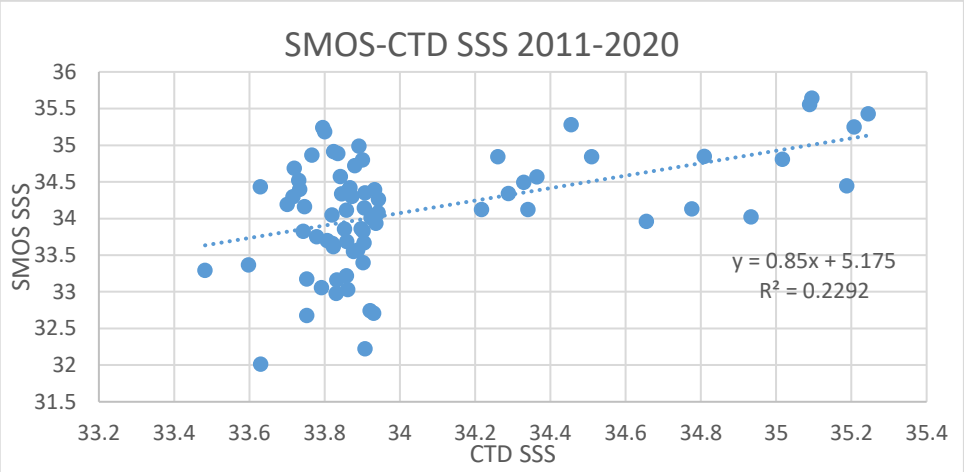


Figure 14 SSS satellite and CTD comparison

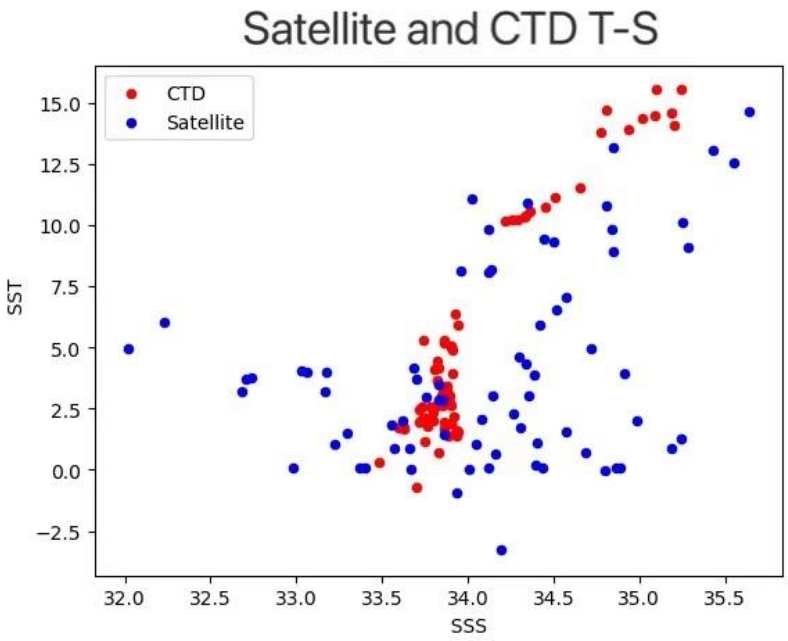


Figure 15 SST and SSS diagram

Chapter 7. Conclusion

In this study, the M-ConvLSTM model that combines CNN and LSTM with layer stacking is proposed to predict the marine surface environmental factors. It can learn the variation of temperature and salinity at the surface ocean. The experimental results show that the M-ConvLSTM offers good prediction accuracy. For several moments of prediction, the SST offer good accuracy as the time step increases. Possibly because the variations of SST are smaller than that of the SSS. For future work, the M-ConvLSTM should be able to improve accuracy in the marine surface environmental factors.

The results showed that the ConvLSTM offers good prediction accuracy for SST. It was considered that the method used here has some usage such as optimal interpolation of missing data. However, in order to further develop such a method and make it practical as like a weather forecast, it is necessary to combine it with the prediction linked to the laws of physics.

Satellite ocean temperature data has already been successfully commercialized. This kind of data changes regularly and with high accuracy. However, L-band is less sensitive to salinity. The result of remote sensing SSS data itself is not good, and the data can not meet the requirements of the prediction model. The artificial intelligence technology used in this study is a data-driven algorithm. It can be seen from the salinity data that the accuracy is low and the change is not regular, which leads to the failure to obtain the rule in machine learning. In the future, if the data accuracy is improved, it will be helpful to predict SSS.

Acknowledgements

Many people have made invaluable contributions, both directly and indirectly to my research. I would like to express my warmest gratitude to Prof. Kitade, my supervisor, for his instructive suggestions and valuable comments on the writing of this thesis. Without his invaluable help and generous encouragement, the present thesis would not have been accomplished. I also wish to thank Dr. Nemoto and Dr. Mizobata for their helpful comment.

Besides, I wish to thank my classmates in my laborary, who helped me search for reference.

Finally, I greatly appreciate my parents and friends' support and endless love. My heart swells with gratitude to all the people who helped me.

REFERENCES

Bengio, Y., P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” IEEE Trans. Neural Netw., vol. 5, no. 2, pp. 157–166, Mar. 1994.

Daverat, F., J. Martin, R. Fablet, and C. Pécheyran, “Colonisation tactics of three temperate catadromous species, eel Anguilla anguilla, mullet Liza ramada and flounder Plathychtys flesus, revealed by Bayesian multielemental otolith microchemistry approach,” Ecol. Freshwater Fish, vol. 20, no. 1, pp. 42–51, Mar. 2011.

Funk, C. C. and A. Hoell, “The leading mode of observed and CMIP5 ENSO-residual sea surface temperatures and associated changes in Indo-Pacific climate,” J. Climate, vol. 28, no. 11, 2015,Art. no. 150202132719008.

Graves, A. and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” Neural Netw., vol. 18, nos. 5–6, pp. 602–610, Jul. 2005.

Hochreiter, S. and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

Hu, Z.-Z., A. Kumar, B. Huang, W. Wang, J. Zhu, and C. Wen,“Prediction skill of monthly SST in the North Atlantic Ocean in NCEP Climate Forecast System version 2,” Climate Dyn., vol. 40, nos. 11–12,pp. 2745–2759, Jun. 2013.

Krizhevsky, A., I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proc. Int. Conf. Neural Inf. Process. Syst., 2012.

Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

Lins, I. D. et al., “Sea surface temperature prediction via support vector machines combined with particle swarm optimization,” in Proc. 10th Int. Probabilistic Saf. Assessment Manage. Conf., 2010, pp. 16–29.

Pascanu, R., T. Mikolov, and Y. Bengio. On the difficulty of training recurrent neural networks. In ICML,pages 1310–1318, 2013.

Ranzato, M., A. Szlam, J. Bruna, M. Mathieu, R. Collobert, and S. Chopra. Video (language) modeling: a baseline for generative models of natural videos. arXiv preprint arXiv:1412.6604, 2014.

Reyniers, M., Quantitative Precipitation Forecasts Based on Radar Observations: Principles, Algorithms and Operational Systems. Institut Royal M´et´eorologique de Belgique, 2008.

Sakaino, H., Spatio-temporal image pattern prediction method based on a physical model with

timevarying optical flow. IEEE Transactions on Geoscience and Remote Sensing, 51(5-2):3023–3036, 2013.

Shi, X. et al., “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in Proc. Neural Inf. Process. Syst. (NIPS) Conf., 2015.

Solanki, H. U., D. Bhatpuria, and P. Chauhan, “Integrative analysis of AltiKa-SSHa, MODIS-SST, and OCM-chlorophyll signatures for fisheries applications,” Mar. Geodesy, vol. 38, pp. 672–683, Mar. 2015.

Srivastava, N., E. Mansimov, and R. Salakhutdinov. Unsupervised learning of video representations using lstms. In ICML, 2015.

Srivastava, N., E. Mansimov, and R. Salakhutdinov, “Unsupervised learning of video representations using LSTMs,” in Proc. 32nd Int. Conf. Mach. Learn., 2015, pp. 843–852.

Stockdale, T. N., M. A. Balmaseda, and A. Vidard, “Tropical Atlantic SST prediction with coupled ocean-atmosphere GCMs,” J. Climate, vol. 19, no. 23, pp. 6047–6061, 2010.

Wang, Q., J. Lin, and Y. Yuan, “Salient band selection for hyperspectral image classification via manifold ranking,” IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 6, pp. 1279–1289, Jun. 2016.

Yang, Y., J. Dong, X. Sun, E. Lima, Q. Mu, and X. Wang, “A CFCC– LSTM model for sea surface temperature prediction,” IEEE Geosci.Remote Sens. Lett., vol. 15, no. 2, pp. 207–211, Feb. 2018.

Zhang, Q., H. Wang, J. Dong, G. Zhong, and X. Sun, “Prediction of sea surface temperature using long short–term memory,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 10, pp. 1745–1749, Oct. 2017.