Forecasting Arctic Sea Ice Concentration using Long Short-term Memory Networks

THUNCHANOK PHUTTHAPHAIBOON

Computer Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand, phut.thunchanok@gmail.com, (+66)924231928

PEERAPON SIRIPONGWUTIKORN

Computer Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand, peerapon.sir@kmutt.ac.th

PRIYAKORN PUSAWIRO

Computer Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand, priyakorn.pus@kmutt.ac.th

Due to global warming, Arctic sea ice is now declining, and this loss is a self-accelerating process that speeds up sea ice melting and the severity of climate change. Accurate and timely sea ice information is critically important for better monitoring of global climate. Publicly available multi-source, multi-scale, and high-dimensional sea ice data from satellites is a game changer that allows researchers to better understand the Arctic through more sophisticated methods. This study proposes two Long short-term memory (LSTM) networks for sea ice concentration (SIC) forecasting in the arctic area over 1-, 3-, 6-, and 9-month forecast horizons. The first network forecasts the SIC of each grid in a single output with the grid coordinate must be supplied as an additional input, while the second network forecasts the SIC of all grids at once in a single output. The models with and without atmospheric and oceanic variables as external predictors were trained by using 43 years of data and tuned by using random search strategies. The model performance was evaluated and compared based on the root mean square errors and weighted absolute percentage errors to determine the impact of using climate variables in the prediction and arrive at the best-performing forecast model.

CCS CONCEPTS • Applied Computing~Physical Sciences and Engineering~Earth and Atmospheric Sciences • Computing Methodologies~Machine Learning~Machine Learning Approaches

Additional Keywords and Phrases: Sea Ice Concentration Forecast, Spatial Time Series Forecasting, Machine Learning, Long Short-term Memory

1 INTRODUCTION

Even though the Arctic is unreachable to most people, sea ice in the Arctic is a key part of life support systems on the Earth and humans depend on it for many reasons. For instance, Arctic sea ice acts as a huge white surface reflecting most of the Sun's radiation into space which helps keep the planet's climate in balance. The melting of ice will release significant amounts of carbon dioxide stored in the form of permafrost into the atmosphere. Sea ice also controls the movement of ocean currents to globally regulate the ocean temperature and is beneficial for indigenous communities

and native wildlife by protecting shorelines and providing ice for hunting. However, Arctic sea ice has been shrinking over 40 years [1], from nearly 6.90 million km^2 in 1979 to about 4.72 million km^2 in 2021. The decreased sea ice exacerbates global warming by reinforcing the ice to melt faster. Massive loss of sea ice is going to force people and animals to be in more endangered situations. A common metric to measure sea ice is the Sea Ice Concentration (SIC), which indicates the density of the sea ice in space and ranges 0-1 or 0%-100%, corresponding to ice-free to full-ice conditions. The SIC data is currently available from remote sensing satellites. Accurate forecasting of SIC allows for better monitoring of global climate, management of ocean and coastal resources, and security of indigenous communities and biodiversity in the Arctic.

Various models have been developed to forecast SIC in the Arctic region. These models can be classified into three categories: physical models, statistical models, and machine learning models. Examples of physical models include SEAS5 [2] and Sea-ice-ocean-atmosphere model (ArcIOAM) [3]. However, physics-based models incurred high computational costs and were susceptible to erratic forecasts due to physical biases. For instance, values of observed annual sea ice extent for 8 out of 12 years from 2009 to 2020 deviated from the interquartile range of predicted sea ice extent of physical models submitted to the Sea Ice Prediction Network (SIPN) [4]. Statistical models such as Vector Autoregression (VAR) [5] and Bayesian Logistic Regression [6] have also been applied to forecast the SIC value and the presence or absence of SIC value above 15%. Compared to physical and statistical models, forecast models based on neural networks and deep learning could address nonlinearity in sea ice time series [7] and had less computational costs once the models are fitted [8]. Previous research works showed that Long Short-Term Memory (LSTM) outperformed Multi-Layer Perceptron (MLP) [9] but LSTM had not presented a significant difference in predictability compared to Gated Recurrent Unit (GRU) when they were used to forecast daily SIC [10]. LSTM-based models were aplenty adopted in many sea ice studies. For instance, attention-based LSTM networks were used in [11] and [12] to make 1-month SIC forecasts which illustrated that the attention mechanism slightly reduced short-term forecast errors compared to traditional LSTM.

The main contributions of this paper are two-fold. First, novel LSTM networks for coordinate-based forecasting and the-entire-Arctic forecasting are proposed. Unlike the previous works where future SICs were forecasted by using only the history of SIC [9, 10], atmospheric and oceanic variables are incorporated as external predictors. Additionally, the models generate 1-, 3-, 6-, and 9-month ahead forecasts simultaneously. Second, the performances of the two LSTM networks with and without external predictors are investigated to assess the impact of external predictors on forecast accuracy and arrive at the best-performing forecast model. The remainder of the paper is organized as follows: Section 2 describes the dataset used in the modeling and explores their characteristics. Section 3 describes data preprocessing, model construction, and model training and tuning. Section 4 presents the results and discussion. Finally, the conclusion is presented in Section 5.

2 DATA SOURCES

SIC and related variables were obtained from ERA5 [13], a global reanalysis dataset published by European Centre for Medium-Range Weather Forecasts (ECMWF). The dataset contains hourly and monthly atmospheric, oceanic, and land surface climate variables from satellite observations. Forty-three years of data from the first launch of the satellite in 1979 to the year 2021 were taken for the model development. Original values of climate quantities retrieved from ERA5 are represented in the form of latitude-longitude grid projection with a spatial resolution of 30 km $\times$ 30 km, which amounts to 721 $\times$ 1440 grid points covering the entire Earth's surface. Figure 1 shows an example of the sea ice coverage

using cylindrical map projections on the Earth's surface in March 1979. Because the Arctic was our focus, only the Earth's surface above the 40th parallel north (red line) was selected, corresponding to 210 x 1440 grid points.

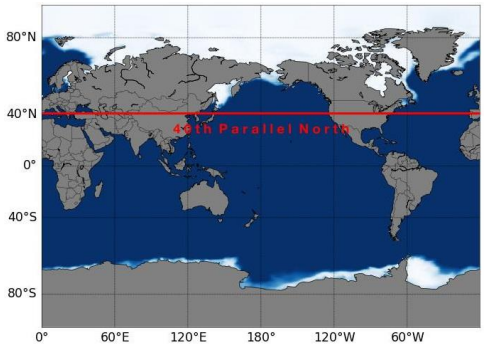


Figure 1: 40th parallel north or the line at 40 degrees north latitude which was approximately the lowest latitude where sea ice formed in the Northern Hemisphere

The variables used to develop the forecasting models are summarized in Table 1. The SIC is the model output while the other four variables are the external predictors, including sea surface temperature (SST), air temperature 2 meters above the surface (T2M), height above the surface with pressure 500hPa (Z500), and Albedo, or the proportion of light reflected by the surface (FAL). These external predictors were shown to be related to SIC based on scientific findings in previous studies [14, 15, 16]. Sample time series plots of the SIC and the predictors at a specific grid cell are depicted in Figure 2. The plots show that all the variables have strong seasonal components with a yearly seasonal period. The SIC shows intermittent periods of zeros during summer months, which poses a significant challenge in the modeling.

Table 1: Variables used in SIC forecasting models

Variable	Abbreviation	Value Range
Sea ice concentration	SIC	0.00 - 1.00 %
Sea surface temperature	SST	269.37 - 305.05 Kelvin
2m air temperature	T2M	221.23 - 310.79 Kelvin
500hPa geopotential height	Z500	47370.34 - 58288.11 m^2s^{-2}
Albedo	FAL	0.04 - 0.87 %

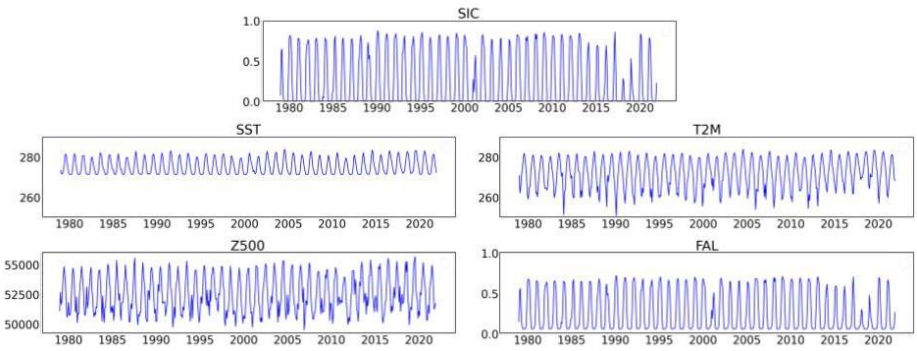


Figure 2: Sample time series plots of SIC and predictors at a geographic coordinate of 45 degrees north latitude 9.25 degrees west longitude

3 METHODS

3.1 Data preprocessing and preparation

To reduce computational costs during the model training, the resolution of the grid points was scaled down from 30-by-30 km^2 to 450-by-450 km^2 so that the number of grid points was reduced to $14 \times 96 = 1{,}344$. Since each scaled-down grid point consisted of 225 of 30-by-30 km^2 grid points, we obtained the values of each input variable in the scaled-down grid points by averaging 225 values from the original resolution. The SIC data also revealed that there existed many points with no sea ice, most of which were points in land and open seawater. Those grid points were excluded from the modeling, resulting in 810 usable grid points. As LSTM required normalized inputs, SIC and FAL variables were scaled by using min-max scaling between 0 and 1 as their range have exact bounds, whereas SST, T2M, and Z500 variables were scaled by using variance scaling (standardization) as their range have no exact bound.

3.2 Proposed LSTM Networks

Long Short-term Memory (LSTM) is a type of Recurrent Neural Network (RNN) designed to learn lone-term sequential data. It was invented by Hochreiter and Schmidhuber [17] to solve the problem that classical RNNs could not accurately make contextual predictions from long-term memory due to the vanishing gradient problem. In this work, two LSTM networks were developed with different approaches to concurrently forecast SIC at 1-, 3-, 6-, and 9-month forecast horizons. The first network is used to forecast the SIC of each individual grid point at the time while the second network is used to forecast the SIC of grid points in the entire Arctic all at once. The two model network architectures are described below.

3.2.1 Coordinate-based LSTM Network

The first LSTM network forecasts SIC for each grid at a time where the grid coordinate must be supplied as an additional input. As depicted in Figure 3(a), the model takes two sets of inputs. The first input fed to Input_1 layer consists of SIC, and optionally predictor variables, including SST, T2M, Z500, and FAL. The second input fed to Input_2 layer is the coordinate (latitude and longitude) of the grid point matching the first input. This second input is concatenated with the output from LSTM_3 layer before being fed to the dense layer to produce the SIC forecasts (1-, 3-, 6-, and 9-month forecast horizons). Tuples in each block denote the input and output dimensions of the layers. The first value in the tuple

is the number of time steps to train the LSTM unit and the second is the number of variables to feed to the LSTM unit. The quantity n_var is the number of features including the target variable fed to the model, and n_node is the number of LSTM units in the hidden layers to be tuned. Each input sample for the stacked LSTM layer (LSTM_1+LSTM_2+LSTM_3) is made up of overlapping windows of 12 lags of the five variables. As mentioned earlier, the input data period is 516 months, from January 1979 to December 2021. So, the number of input samples per grid point is 516-12-9 = 495, and the total number of input samples from all the grid points is 810 grid points x 495 samples = 400,950.

3.2.2 All-grid LSTM Network

The second LSTM network simultaneously predicts SIC values of the entire Arctic Ocean or of all 810 grid points for a specific forecast horizon. Figure 3(b) illustrates the model architecture. The model takes SIC, and optionally the predictor variables SST, T2M, Z500, and FAL data of all 810 grid points as the input to the stacked LSTM layer (LSTM_1+LSTM_2+LSTM_3). The output from the stacked LSTM layer is connected to the dense layer to produce the SIC forecasts. Each input sample for the stacked LSTM layer is made up of overlapping windows of 12 lags of the five variables of all 810 grid points. The total number of samples is 516-12-9 = 495.

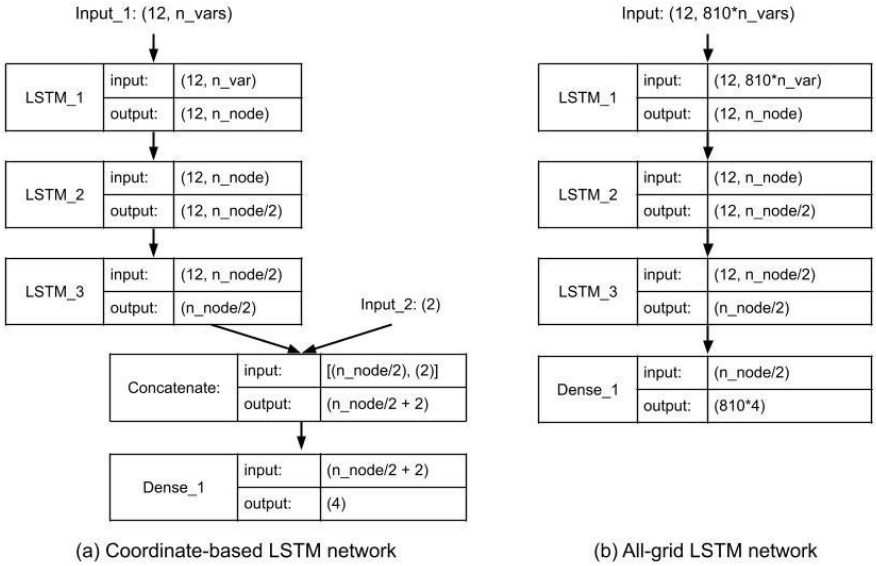


Figure 3: Two proposed LSTM networks for SIC forecasting

3.3 Hyperparameter tuning

The input samples were split into training, validation, and testing with the proportion of 80%, 10%, and 10% respectively. The hyperparameters were tuned by using the random search strategy [18] based on the root mean square errors (RMSE). The ranges of hyperparameters used in the tuning are shown in Table 2. A list of hyperparameter combinations was randomly generated, and each combination was evaluated by using growing-window walk-forward validation [19]. After all hyperparameter combinations in the list were evaluated, their RMSE scores were ranked to obtain the best-performing set of hyperparameters.

Table 2: Hyperparameter ranges used in the model tuning

Hyperparameter	Searching Range
Number of LSTM units in the stacked LSTM layer (n_node)	10 - 450 stepping by 5
Number of epochs to train model (n_epoch)	50 - 500 stepping by 50
Batch size	32, 64, 128, 256
Optimizer	'adam', 'Nadam'
Dropout	0.0 - 0.5 stepping by 0.1

4 EVALUATION RESULTS AND DISCUSSION

We considered scenarios with and without climate variables as external predictors, resulting in four models being evaluated. We refer to the coordinate-based LSTM networks with and without external predictors as Models M1a and M1b respectively. In other words, model M1a used only SIC as the input while model M1b uses SIC, SST, T2M, Z500, and FAL as the inputs. Similarly, the all-grid LSTM networks with and without external predictors are referred to as Models M2a and M2b respectively. All four models were tuned according to the method presented in section 3.3. Both RMSE and weighted absolute percentage error (WAPE) performance metrics are reported. A WAPE is a scale-independent forecast accuracy metric that is appropriate when the output contains zeros or values near zeros. It is calculated by taking the sum of the absolute errors and dividing it by the sum of observed values as

$$WAPE = \frac{\sum_{i,t}|\hat{y}_{i,t} - y_{i,t}|}{\sum_{i,t}|y_{i,t}|} \times 100$$

where $y_{i,t}$ denotes the observed value of grid point i at time t, and $\hat{y}_{i,t}$ denotes the predicted value of grid point i at time t.

Table 3 shows the forecast accuracies of the four models over the test data (February 2017 - December 2021). In terms of model complexity, the all-grid models required many more nodes in the LSTM layers as well as the number of epochs compared to the coordinate-based models because the formers were designed to learn from much larger input data. In addition, the models that included external predictors (M1b and M2b) required more nodes compared to the models which used only SIC as the input. A performance comparison of the two LSTM networks—coordinate-based and all-grid—presented an obvious difference. The all-grid LSTM networks gave more accurate SIC predictions with respect to RMSE and WAPE. However, the use of additional climate variables did not improve the forecast accuracies as anticipated, as model M1b outperformed M1a, but M3b did not outperform M3a. The most accurate forecasting was obtained from model M2a.

Table 3: Model tuning results and performance

Model	LSTM Network	Input Variables	Tuned Hyperparameters	RMSE	WAPE
M1a	Coordinate-based	SIC	n_node = 30, n_epoch = 100, batch size = 128, optimizer = 'adam', dropout = 0.0	0.097	10.145%
M1b	Coordinate-based	SIC, SST, T2M, Z500, and FAL	n_node = 85, n_epoch = 100, batch size = 128, optimizer = 'adam', dropout = 0.0	0.093	9.397%
M2a	All-grid	SIC	n_node = 300, n_epoch = 300, batch size = 64, optimizer = 'adam', dropout = 0.4	0.074	7.898%
M2b	All-grid	SIC, SST, T2M, Z500, and FAL	n_node = 380, n_epoch = 300, batch size = 64, optimizer = 'adam', dropout = 0.1	0.080	8.291%

The distributions of forecast errors at the 1-month forecast horizon are shown in Figure 4. The forecast errors of all the models appear heavy-tail symmetric but slightly right-skewed. The forecast error distributions of the other remaining forecast horizons are similar. While most of the forecast errors were clustered around zeroes, we found that large forecast errors occurred around the marginal ice zone (MIZ) over the Arctic subregions, including Beaufort Sea, Chukchi Sea, and East Siberian Sea. These areas were challenging for forecasting because they completely melted and stayed in an ice-free state in all previously reported summers. Moreover, models M1a and M1b underestimated SIC in September 2021, most likely due to changes in the SIC trends around MIZ after 2010. As models M1a and M1b aimed to forecast each grid individually without considering data from the other grids, the drastic change in a specific grid may confuse the learning. Figure 5(a) illustrates the average WAPE by months. All four models had a similar pattern of monthly WAPEs in which the errors stood very low at the beginning of the year and peaked in August. They all had trouble forecasting SIC from midsummer to fall but Models M2a and M2b can maintain WAPE of less than 10% in June and November. From Figure 5(b), model M2a appears to yield the overall-best forecast accuracies across all the forecast horizons.

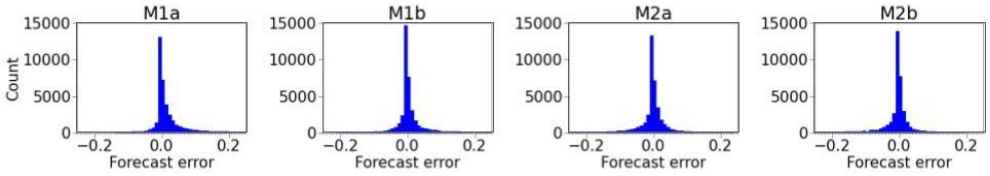
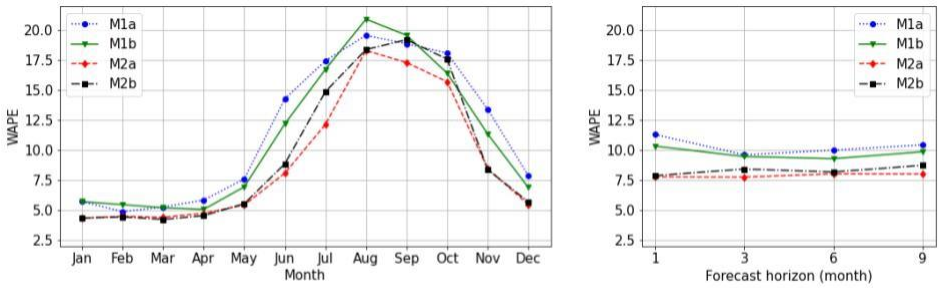


Figure 4: Distribution of forecast errors at 1-month forecast horizon



(a) Average WAPE by month at 1-month forecast horizon

(b) Average WAPE at different forecast horizons

Figure 5: Comparison of WAPEs from different forecast models

5 CONCLUSION

We proposed two LSTM networks for SIC forecasting in the Arctic area. The first network was designed to forecast the SIC of each grid point in a single output with the grid coordinate being supplied as an additional input, whereas the other was designed to concurrently forecast SICs of all grid points covering the Arctic in a single output. Four forecasting models based on coordinate-based and all-grid LSTM networks were trained using the past data of SIC with and without incorporating climate variables as external predictors over the data period of 464 months, and the model

hyperparameters were tuned by using the random search strategy. The evaluation results revealed that the model based on the all-grid LSTM network using only the historical SIC data yielded the best performance with an RMSE of 0.074 and a WAPE of 7.898%. Not only did the model outperform others, but also required lower computational costs. However, computational limitations restricted the model deployment in higher spatial resolution, which lessened the opportunity that the forecasts would be practically beneficial for stakeholders to anticipate future SIC trends. Further development of this research is to be more selective in choosing predictors, which could be done by performing feature importance analysis. Moreover, SIC time series in certain sets of grid points show similar patterns, which allows for clustering grid points and developing separate forecasting models for different clusters to improve overall forecast accuracies without incurring too much computational cost in model training.

REFERENCES

[1] The National Aeronautics and Space Administration. (n.d.). Arctic Sea Ice Minimum Extent | Vital Signs – Climate Change: Vital Signs of the Planet. NASA Climate Change. Retrieved October 8, 2022, from https://climate.nasa.gov/vital-signs/arctic-sea-ice/

[2] Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., & Monge-Sanz, B. M. (2019). SEAS5: The New ECMWF Seasonal Forecast System. Geoscientific Model Development, 12, 1087-1117. https://doi.org/10.5194/gmd-12-1087-2019

[3] Ren, S., Liang, X., Sun, Q., Yu, H., Tremblay, B., Lin, B., Mai, X., Zhao, F., Li, M., Liu, N., Chen, Z., & Zhang, Y. (2021). A Fully Coupled Arctic Sea-Ice–Ocean–Atmosphere Model (Arcioam V1.0) Based On C-Coupler2: Model Description and Preliminary Results. Geoscientific Model Development, 14, 1101-1124. https://doi.org/10.5194/gmd-14-1101-2021

[4] Wei, K., Liu, J., Bao, Q., He, B., Ma, J., Li, M., Song, M., & Zhu, Z. (2021). Subseasonal To Seasonal Arctic Sea-Ice Prediction: A Grand Challenge Of Climate Science. Atmospheric and Oceanic Science Letters, 14. https://doi.org/10.1016/j.aosl.2021.100052

[5] Wang, L., Yuan, X., Ting, M., & Li, C. (2016). Predicting Summer Arctic Sea Ice Concentration Intraseasonal Variability Using a Vector Autoregressive Model. Journal of Climate, 29(4), 1529-1543. https://doi.org/10.1175/JCLI-D-15-0313.1

[6] Horvath, S., Stroeve, J., Rajagopalan, B., & Kleiber, W. (2020). A Bayesian Logistic Regression for Probabilistic Forecasts of the Minimum September Arctic Sea Ice Cover. Earth and Space Science, 7(10). https://doi.org/10.1029/2020EA001176

[7] Li, M., Zhang, R., & Liu, K. (2021). Machine Learning Incorporated with Causal Analysis for Short-Term Prediction of Sea Ice. Frontiers in Marine Science, 8. https://doi.org/10.3389/fmars.2021.649378

[8] Hunke, E., Allard, R., Blain, P., Blockley, E., Feltham, D., Fichefet, T., Garric, G., Grumbine, R., Lemieux, J.-F., Rasmussen, T., Ribergaard, M., Roberts, A., Schweiger, A., Tietsche, S., Tremblay, B., Vancoppenolle, M., & Zhang, J. (2020). Should Sea-Ice Modeling Tools Designed for Climate Research Be Used for Short-Term Forecasting? Current Climate Change Reports, 6, 121-136. https://doi.org/10.1007/s40641-020-00162-y

[9] Chi, J., & Kim, H. (2017). Prediction of Arctic Sea Ice Concentration Using a Fully Data Driven Deep Neural Network. Remote Sensing, 9(12), 1305. https://doi.org/10.3390/rs9121305

[10] Choi, M., De Silva, L. W. A., & Yamaguchi, H. (2019). Artificial Neural Network for the Short-Term Prediction of Arctic Sea Ice Concentration. Remote Sensing, 11(9), 1071. https://doi.org/10.3390/rs11091071

[11] Wei, J., Hang, R., & Luo, J.-J. (2022). Prediction of Pan-Arctic Sea Ice Using Attention-Based LSTM Neural Networks. Frontiers in Marine Science, 9. https://doi.org/10.3389/fmars.2022.860403

[12] Ali, S., Huang, Y., Huang, X., & Wang, J. (2021). Sea Ice Forecasting using Attention-based Ensemble LSTM. https://doi.org/10.48550/arXiv.2108.00853

[13] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 Global Reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999-2049. https://doi.org/10.1002/qj.3803

[14] Timmermans1, M. L., & Labe, Z. (2020). Arctic Report Card 2020: Sea Surface Temperature. Arctic Report Card. https://doi.org/10.25923/v0fs-m920

[15] Liu, X., & Lu, C. (2021). Effects Of Sea Ice Change on The Arctic Climate: Insights from Experiments with A Polar Atmospheric Regional Climate Model. Journal of Water and Climate Change, 12(7), 2885-2893. https://doi.org/10.2166/wcc.2021.206

[16] Pistone, K., Eisenman, I., & Ramanathan, V. (2014). Observational Determination of Albedo Decrease Caused By Vanishing Arctic Sea Ice. Earth, Atmospheric, and Planetary Sciences, 111(9), 3322-3326. https://doi.org/10.1073/pnas.1318201111

[17] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[18] Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. Journal of Machine Learning Research, 13, 281-305.

[19] Schnaubelt, M. (2019). A Comparison of Machine Learning Model Validation Schemes for Non-Stationary Time Series Data. FAU Discussion Papers in Economics. https://doi.org/10.13140/RG.2.2.29545.24168