

A Term Paper On

Application of Multi-Linear Perceptron Neural Network for Predicting Antarctic Sea-Ice Concentration using Meteorological Variables

By

Meghraj Goswami

2022A2B41869H

Under the supervision of

Prof. Jagadeesh Anmala

Submitted in Partial Fulfilment of the Requirements of
CE F417: Application of AI in Civil Engineering



**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN)**

HYDERABAD CAMPUS

(November 2025)

Acknowledgements

I would like to express my sincere gratitude to Professor Jagadeesh Annala, for his guidance, support, and encouragement throughout the semester. His expertise and insights have been instrumental in enhancing my understanding of the subject matter.

Abstract

This project uses feedforward Multi-Layer Perceptron (MLP) Neural Network models for forecasting Sea-Ice concentration (SIC) using meteorological variables derived from ERA5 reanalysis: sea-surface temperature (SST), 2m air temperature (T2M), 10m u and v wind components (U10, V10) and mean sea-level pressure (MSL). The MLP models are trained using various combinations of input variables and model hyperparameters. The performance of the models is evaluated using statistical metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). This project highlights the potential of MLP Neural Networks in accurately predicting SIC, which is crucial for understanding climate dynamics all over the world.

Contents

1	Introduction	4
1.1	Background	4
1.2	Study Area	5
2	Literature Review	6
3	Data and Methodology	7
3.1	Data Sources	7
3.2	Data Preprocessing	7
3.3	MLP Neural Network Architecture	7
3.4	Model Training and Evaluation	8
4	Results	9
4.1	Best One-Layer Model	9
4.2	Best Two-Layer Model	10
5	Discussion	11
5.1	Analysis of Results	11
6	Conclusion	12
7	Bibliography	13

1 Introduction

1.1 Background

Sea-ice plays an important role in the global climate system, influencing heat exchange, ocean circulation and weather patterns. Sea-Ice concentration in Antarctica varies seasonally, reaching its maximum in the Antarctic winter (July to September) and its minimum in the Antarctic summer (January to March). This variation is strongly influenced in non-linear ways by various meteorological variables, including temperature (SST), wind patterns (U10, V10), and atmospheric pressure (MSL). Accurate prediction of SIC is crucial for understanding climate dynamics and policy-making related to climate change.

Traditional methods for predicting SIC often rely on physical models that can be computationally expensive and require extensive calibration, and may not capture the complex non-linear relationships between SIC and meteorological variables. Moreover, the accuracy of these models can be limited by the resolution and quality of input data.

On the other hand, machine learning techniques, particularly neural networks, are known to capture complex patterns in data. Multi-Layer Perceptron (MLP) Neural Networks are a type of feedforward artificial neural network (ANN) that can model non-linear relationships between input and output variables. MLPs consist of multiple layers of interconnected neurons that can learn from large datasets through iterative training and optimization. Due to this, they are generally well-suited for cases where the input data is not at par with the quality or resolution required by physical models.

This project aims to explore the application of MLP Neural Networks for predicting Antarctic Sea-Ice concentration using atmospheric and oceanic variables as input variables derived from ERA5 reanalysis data. The performance of the MLP models are evaluated using various statistical metrics, and the results are compared to gauge the relative effectiveness of the models.

1.2 Study Area

The study area for this project is the Antarctic region, specifically the Southern Ocean surrounding the continent of Antarctica. Antarctica occupies an area of about 14 million square kilometers, which is mostly covered by 29 million cubic kilometers of ice - about 90% of the world's ice and 80% of its fresh water.

The Antarctic sea-ice cover is typically divided into five sectors: the Ross Sea, the Weddell Sea, the Indian Ocean sector, the West Pacific sector and the Bellingshausen-Amundsen Seas. Each sector exhibits distinct sea-ice characteristics due to variations in local atmospheric and oceanic conditions.

The weather in Antarctica is characterized by high-speed katabatic winds, which are created due to cold and dense air blowing downhill from the polar plateau. These winds influence formation and distribution of sea-ice in the surrounding Southern Ocean.

Antarctica also experiences extreme temperature patterns due to being plunged into darkness during the Antarctic winter and experiencing continuous sunlight during the Antarctic summer. Depending on the region, temperatures in Antarctica fluctuate between -80°C in winter to 10°C in summer. Precipitation is generally low, with most of it falling as snow.

Atmospheric pressure patterns in Antarctica also play a significant role in influencing sea-ice dynamics. The mainland is usually surrounded by a low-pressure belt, while the interior experiences high-pressure, creating conditions for the formation of the characteristic katabatic winds.

2 Literature Review

3 Data and Methodology

3.1 Data Sources

The data used in this project is taken from the ERA5 reanalysis dataset, which is maintained by the European Centre for Medium-Range Weather Forecasts (ECMWF). The dataset uses historical meteorological observations and weather prediction models to generate comprehensive data, from 1940 to the present. It has a spatial resolution of $0.25^\circ \times 0.25^\circ$ and a temporal resolution of 1 hour, although a monthly frequency has been chosen for this project. This makes it suitable for studying climate patterns and trends.

The specific variables used in this project are:

- Sea-Surface Temperature (SST)
- 2m Air Temperature (T2M)
- 10m U Wind Component (U10)
- 10m V Wind Component (V10)
- Mean Sea-Level Pressure (MSL)

The target variable for prediction is the Sea-Ice Concentration (SIC) in the Antarctic Southern ocean region (latitude range -55° to -75°).

The dataset spans a decade from January 2014 to December 2024, incorporating seasonal and inter-annual variability.

3.2 Data Preprocessing

In its raw form, the ERA5 dataset contains missing values, outliers and inconsistencies that need to be taken care of before training the MLP models. The following preprocessing steps are applied:

- **Handling Missing Values:** Missing values in the dataset are handled using imputing, interpolation and removal of records with excessive missing data. The mean of data points is used for imputation.
- **Train-Test Splitting:** The dataset is split into training (70%) and testing (30%) sets. The training set is used to train the MLP models, while the testing set is used to evaluate their performance.
- **Standardization:** The input variables are standardized, or in other words they are rescaled to have a mean of 0 and a standard deviation of 1. This ensures that all input variables contribute equally when training the model.

3.3 MLP Neural Network Architecture

A Multi-Layer Perceptron (MLP) Neural Network is a type of feedforward artificial neural network composed of an input layer, one or more hidden layers, and an output layer. Each layer consists of multiple neurons, and neurons in one layer are connected to those

in the next layer through weighted connections. Backpropagation is used to train the network by adjusting the weights based on the error difference between predicted and actual outputs. MLPs are widely used for supervised learning problems, such as regression and classification.

The MLP Neural Network architecture used in this project comprises the following components:

- **Input Layer:** The input layer consists of neurons corresponding to the number of meteorological variables used as input features (maximum of 5 in this case).
- **Hidden Layers:** The MLP architecture includes a number of hidden layers, each with varying number of neurons. The number of hidden layers and neurons per layer are hyperparameters that are defined by the user and can be tuned for optimal performance.
- **Output Layer:** The output layer consists of a single neuron that predicts the Sea-Ice Concentration (SIC) value.
- **Learning Rate:** The learning rate (α) controls the step size while updating the model weights during training. A suitable learning rate is chosen to ensure convergence.
- **Activation Functions:** Non-linear activation functions such as Rectified Linear Unit (ReLU), Sigmoidal or Hyperbolic Tangent (TanH) are used in the hidden layers to introduce non-linearity into the model.
- **Loss Function:** Mean Squared Error (MSE) is used as the loss function to measure the difference between predicted and actual SIC values during training.
- **Optimization Algorithm:** The Adaptive Moment Estimation (Adam) optimizer is used to update the weights of the network during training. Other optimizers like Stochastic Gradient Descent (SGD) can also be used.

3.4 Model Training and Evaluation

The MLP Neural Network models are trained using the preprocessed training dataset. The training process involves multiple iterations, or epochs, where the model iteratively perturbs and adjusts its weights to minimize the output of the loss function. The following steps are involved in model training and evaluation:

The trained models are evaluated using the testing dataset. The following statistical metrics are used to assess model performance:

- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- Coefficient of Determination (R^2)

The performance of different MLP models is compared based on the evaluation metrics using the following formula:

$$\frac{1}{RMSE} + \frac{1}{MAE} + 10R^2$$

4 Results

To fine-tune the optimal hyperparameters for the MLP Neural Network, an exhaustive grid search was conducted over a user-defined hyperparameter space. The hyperparameters considered in the grid search include:

- Input Variables: SST, T2M, U10, V10, MSL
- Architecture:
 - One-Layer and Two-Layer Networks
 - 2-6 Neurons in each hidden layer
- Activation Functions: ReLU, TanH, Sigmoidal
- Optimization Schemes: Adam, SGD
- Learning Rates: 0.001, 0.005, 0.01
- Maximum Epochs: 200
- Cross-Validation: 3-Fold

For the one-layer networks, a total of $4 * 3 * 2 * 3 = 72$ candidate models are tested, amounting to 216 training runs with 3-fold cross-validation. For the two-layer networks, a total of $16 * 3 * 2 * 3 = 288$ candidate models are tested, amounting to 864 training runs with 3-fold cross-validation.

4.1 Best One-Layer Model

Out of the defined hyperparameter space, the best-performing one-layer MLP model based on the evaluation metrics had the following hyperparameters:

- Input Variables: SST, T2M, U10, V10, MSL
- Architecture: One hidden layer with 4 Neurons
- Activation Function: TanH
- Optimization Scheme: SGD
- Learning Rate: 0.005
- Maximum Epochs: 200

The performance of this model on the testing dataset is as follows:

- RMSE: 0.1919
- MAE: 0.1099
- R^2 : 0.8705

4.2 Best Two-Layer Model

Out of the defined hyperparameter space, the best-performing two-layer MLP model based on the evaluation metrics had the following hyperparameters:

- Input Variables: SST, T2M, U10, V10, MSL
- Architecture: Two hidden layers with 5 Neurons each
- Activation Function: TanH
- Optimization Scheme: SGD
- Learning Rate: 0.001
- Maximum Epochs: 200

The performance of this model on the testing dataset is as follows:

- RMSE: 0.1923
- MAE: 0.1050
- R^2 : 0.8690

5 Discussion

5.1 Analysis of Results

The results obtained from the grid searches indicate that both the best one-layer and two-layer MLP models constrained in the defined hyperparameter space performed comparably well in predicting Antarctic SIC using meteorological variables. The RMSE and MAE values for both models are very similar, with the one-layer model slightly outperforming the two-layer model in terms of RMSE, while the two-layer model had a slightly better MAE. The R^2 values for both models are also very close, indicating that both models explain a significant portion of the variance in the SIC data.

The choice of activation function (TanH) and optimization scheme (SGD) appears to be effective for this specific problem, as both models utilized these configurations. The learning rates chosen for the best models (0.005 for one-layer and 0.001 for two-layer) also seem to have contributed positively to the training process.

The marginal differences in performance metrics suggest that increasing the complexity of the model by adding an additional hidden layer did not lead to a significant improvement. Therefore, guided by the principle of parsimony, the simpler one-layer network is the preferred model, and is termed a parsimonious model. This is likely because the relationship between the input variables and SIC is not complex enough to warrant a higher dimensional network, or the additional layer led to overfitting on the training data.

Overall, both models demonstrate the capability of MLP Neural Networks in capturing the non-linear relationships between input variables and the target. However, with near identical performance, the parsimonious model (the one-layer network) is preferred for the balance between simplicity and predictive accuracy for this specific application.

6 Conclusion

This project explored the application of MLP Neural Networks for predicting Antarctic SIC using five key meteorological variables derived from ERA5 reanalysis data. An exhaustive grid search was conducted over a defined hyperparameter space to identify the best-performing one-layer and two-layer MLP models.

The best performing models used:

- Hyperbolic Tangent Activation Function
- Stochastic Gradient Descent Optimization Scheme

and achieved:

- RMSE: 0.192
- MAE: 0.107
- R^2 : 0.87

The results indicate that even simple networks can effectively model large-scale climate phenomena like sea-ice variability when trained on physically meaningful input variables. More complex architectures did not yield significant improvements while increasing computational costs. This highlights the importance of model parsimony in machine learning applications for climate science.

Future work could explore additional input variables, longer training periods, and alternative neural network architectures like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) to further enhance predictive capabilities. In conclusion, this project demonstrates the potential of MLP Neural Networks as a valuable tool for climate modeling and prediction tasks.

7 Bibliography

- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023). *ERA5 hourly data on single levels from 1940 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.adbb2d47>
- Jena B, Kumar A, Ravichandran M, Kern S (2018). *Mechanism of sea-ice expansion in the Indian Ocean sector of Antarctica: Insights from satellite observation and model reanalysis*. PLoS ONE 13(10): e0203222. <https://doi.org/10.1371/journal.pone.0203222>
- 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*. Journal of Machine Learning Research, 12, 2825-2830. <http://jmlr.org/papers/v12/pedregosa11a.html>