667 - Incorporating Causality with Deep Learning in Predicting Short-Term and Seasonal Sea Ice

📅 Wednesday, January 31, 2024

🕒 3:00 PM - 4:30 PM

📍 *Hall E (100 Level, The Baltimore Convention Center)*

Abstract

Arctic sea ice (ASI) is playing a pivotal role in keeping global warming under control. However, the recently amplified decreasing sea ice trend has become a major concern. Since satellites started monitoring the ASI in 1979, every decade the Arctic has lost 13.1% of sea ice and the Arctic's September Sea Ice Extent (SIE) is now almost half compared to 1979. If this trend continues, it will be ice-free by 2050. Due to its wide-ranging effects, forecasting Arctic sea ice extent is of utmost importance. Accurate forecasts are essential to understanding the effects of global climate change, protecting the polar ecosystem, determining marine shipping routes, assisting indigenous communities, etc. In a world that is rapidly changing, foreseeing sea ice changes enables proactive environmental, economic, and social responses.

Current machine learning (ML) based sea ice forecasting models are not better than typical statistical models when the lead time is more than two months. This is because these models are built on correlation and do not consider the actual cause of changes. Causal models, on the other hand, take into account cause-and-effect relationships between the atmospheric variables. Causal discovery techniques can directly identify these causal connections from observational data. Two popular causal discovery algorithms are Granger Causality (GC) and PCMCI+ (which is an extended version of the PC (Peter-Clark) algorithm with Momentary Conditional Independence). The main aim of this research is to predict Arctic SIE for lead times of 1-6 months using daily and monthly Arctic sea ice data. The objectives of this work are to identify the causal connections between Arctic sea ice and ocean-atmospheric variables using GC and PCMCI+, use the identified causal features to build deep learning models to predict SIE, and then compare the performance of the causal deep learning models against traditional deep learning models (which are trained on both causal and non-causal features).

We used the daily SIE data from 01/01/1979 to 12/31/2018 and monthly SIE data from 01/1979 to 08/2021 for the Pan-Arctic region, that is, $25°$ N. This dataset is curated by combining ERA-5 reanalysis and National Snow and Ice Data Centre (NSIDC) datasets, and then spatially averaged to create time series data. In addition to SIE, the datasets contain 10 ocean-atmospheric variables: *surface pressure, wind velocity, specific humidity, air temperature, shortwave radiation, longwave radiation, rainfall, snowfall, sea surface temperature,* and *sea surface salinity*. After applying the Granger Causality algorithm to these datasets, it identified all features except *Sea Surface Temperature (SST)* as the causal features (for both daily and monthly data). PCMCI+ algorithm identified *Longwave Radiation, Snowfall, Sea Surface Temperature, Sea Surface Salinity, Surface Pressure*, and *Sea Ice Extent* for the daily data and *Longwave Radiation, Sea Surface Temperature*, and *Sea Ice Extent* for the monthly data. Figure 1 (attached) shows the causal graphs identified by PCMCI+ for daily and monthly data.

Since we're dealing with time series data of sea ice, Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) are two popular and effective models for time series analysis. We used a hybrid GRU-LSTM model for predicting SIE as combining these two models is proven to improve the performance. The input layer has 21 neurons because PCMCI+ found that SIE has a maximum time lag of 21 for the causal connections. The first hidden layer is the GRU layer with 64 neurons and 20% dropout, and the second hidden layer is the LSTM layer having 128 neurons and 20% dropout. The third hidden layer is a dense layer with 64 neurons and the output layer has 1 neuron which predicts SIE. Based on the model we choose for daily or monthly data, we feed the appropriate features into the input layer.

The daily and monthly GRU-LSTM models are trained on the data up to 2013 from which 10% data is used for model validation. The remaining data is to test the performance of the model. We trained the models using *mean_squared_error* loss function, and *Adam* optimizer with a batch size 64 for 100 epochs. To evaluate the performance of the models, we used Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Correlation (R^2). The results are shown in Figure 2 (attached). For the daily data, we have three GRU-LSTM models based on the features they are trained on: all features, GC causal features only, and PCMCI+ causal features only. GRU-LSTM model trained on causal features identified by GC performs better (lower errors and higher R2) than the other two models in most of the lead times. In similar ways, for the monthly data, we have four models that were trained on: all features, GC features, PCMCI+ features identified for daily data, and PCMCI+ features identified for monthly data. The combination of the later three causal models (which considered only causal features) outperforms the model trained on all features (both causal and non-causal).

Prediction of short-term or seasonal sea ice extent is important for various reasons including tracking global warming patterns. This work aims to develop a generalized deep-learning model that incorporates causality for both short-term and long-term SIE prediction. Our proposed model can utilize both correlated and causal features for predicting sea ice up to 6-month lead times. The results showed that out of 42 cases (7 models and 6 lead times), training the deep learning model using only causally related features improves the predictive capability. However, the predictive capability of these deep learning models can further be improved. Instead of training on all time lags, we can train the deep learning model by considering only the identified time lags by the PCMCI+ algorithm. Identifying and training on the exact time lag may particularly help in predicting summer
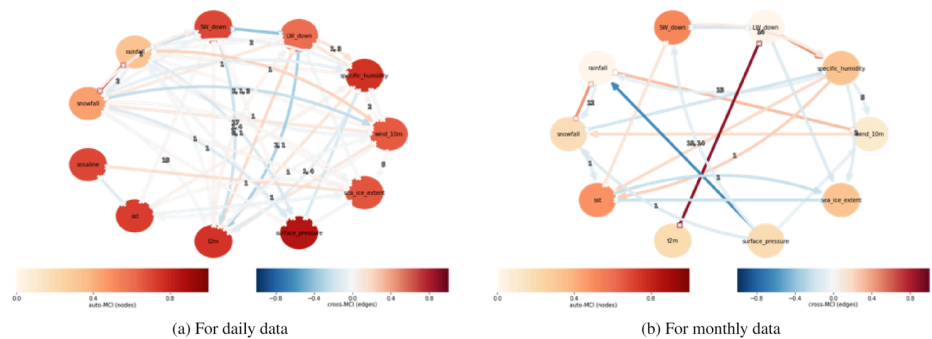
sea ice.

(a) For daily data   (b) For monthly data

FIG. 1: Causal graphs identified by PCMCI+ algorithm

.

(a) RMSE for the daily data   (b) MAE for the daily data   (c) R^2 for the daily data

(d) RMSE for the monthly data   (e) MAE for the monthly data   (f) R^2 for the monthly data
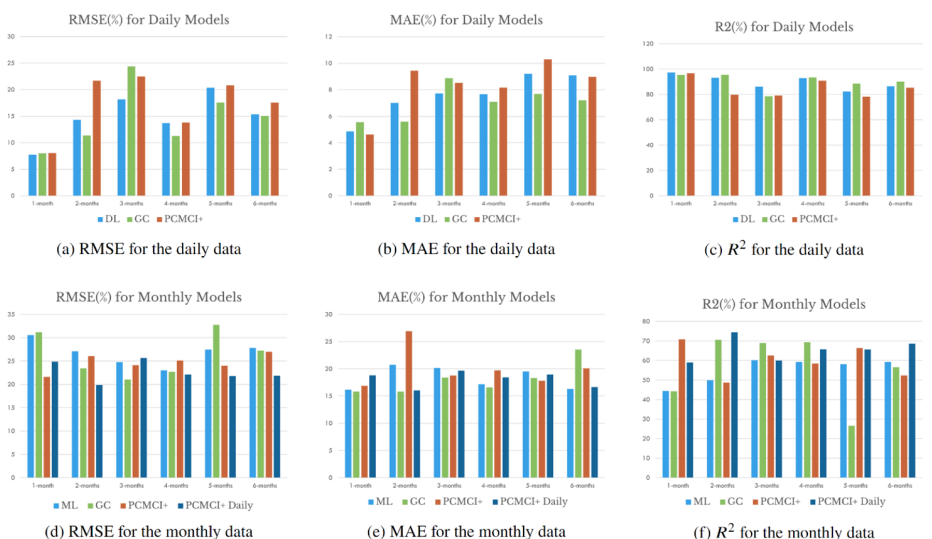
FIG. 2: Performance comparison of the models

Co-Authors

Emam Hossain (Presenter)
University of Maryland Baltimore County
Baltimore, MD
USA

Sahara Ali
University of Maryland Baltimore County
Baltimore, MD
USA

Yiyi Huang
University of Maryland Baltimore County
Baltimore, MD
USA

Nicole-Jeanne Schlegel
GFDL
Princeton, NJ

USA

W

Jianwu Wang

University of Maryland Baltimore County
Baltimore, MD
USA

S

Aneesh C. Subramanian

Univ. of Colorado
Boulder, CO
USA

G

Md Osman Gani

University of Maryland Baltimore County
Baltimore, MD
USA