

Integrating Fourier Transform and Residual Learning for Arctic Sea Ice Forecasting

Louis Lapp

Baltimore Polytechnic Institute
Baltimore, MD, USA
2louislapp@gmail.com

Sahara Ali

University of Maryland Baltimore County
Baltimore, MD, USA
sali9@umbc.edu

Jianwu Wang

University of Maryland Baltimore County
Baltimore, MD, USA
jianwu@umbc.edu

Abstract—Arctic sea ice plays integral roles in both polar and global environmental systems, notably ecosystems, communities, and economies. As sea ice continues to decline due to climate change, it has become imperative to accurately predict the future of sea ice extent (SIE). Using datasets of Arctic meteorological and SIE variables spanning 1979 to 2021, we propose architectures capable of processing multivariate time series and spatiotemporal data. Our proposed framework consists of ensembled stacked Fourier Transform signals (FFTstack) and Gradient Boosting models. In FFTstack, grid search iteratively detects the optimal combination of representative FFT signals, a process that improves upon current FFT implementations and deseasonalizers. An optimized Gradient Boosting Regressor is then trained on the residual of the FFTstack output. Through experiment, we found that the models trained on both multivariate and spatiotemporal time series data performed either similar to or better than models in existing research. In addition, we found that integration of FFTstack improves the performance of current multivariate time series deep learning models. We conclude that the high flexibility and performance of this methodology have promising applications in guiding future adaptation, resilience, and mitigation efforts in response to Arctic sea ice retreat.

Index Terms—Arctic sea ice extent, Fourier Transform, machine learning, time series forecasting

I. INTRODUCTION

Arctic sea ice is a critical component of polar atmospheric and oceanic systems. Yet, our linear regression analysis reveals marked sea ice decline in the past several decades, averaging 0.081 million km² of loss per year. As a matter of fact, the Arctic has undergone warming at globally unprecedented rates, a phenomenon known as Arctic amplification [13]. Once Arctic temperature reaches the bifurcation point, estimated around 2°C, summer sea ice may be irreversibly lost [8]. However, the timeline of this decline is uncertain due to the chronic underestimation of current predictions [12]. The effects on the Arctic environment are numerous, including increases in local sea level, air temperatures, and moisture levels [3], [7]. Causal sequences have also been linked between sea ice loss, anomalous atmospheric patterns, and extreme weather events in Europe, Asia, and North America [5], [16].

Arctic ecosystems are directly impacted by the geophysical effects of sea ice loss, leading to plankton growth, and polar bear vulnerability [10], [17]. Humans have also been impacted, putting at risk the livelihood of indigenous peoples, and increasing shipping route accessibility [9], [11]. Thus, Arctic

sea ice loss is a prominent issue for people and ecosystems on both the polar and global levels, thereby warranting further research.

A. Related Work

The majority of current SIE prediction models rely on statistical, dynamical, machine learning (ML), and more recently, deep learning (DL) techniques. For reference, we use one month lead time Root Mean Square Error (RMSE) score as the evaluation metric for comparison. Using daily sea ice concentration data, Wang et al. implemented a vector autoregression technique to predict sea ice concentration and derived September mean SIE, achieving an RMSE of 0.45 million km² [14]. Using atmospheric and oceanic variables, Wang et al. implemented the Climate Forecast System (CFSv2), a fully-coupled physical model, to predict Arctic SIE, achieving an RMSE range of 0.2-0.6 million km² [15]. Using atmospheric and oceanic variables, Ali et al. implemented an ensemble of Long-Short Term Memory (LSTM) models to predict Arctic SIE, achieving an RMSE of 0.586 million km² [1]. Using a similar dataset, Ali et al. also conducted comparative analyses of ML and probabilistic models, finding that a Multiple Linear Regression model performed best with an RMSE of 0.433 million km² for predicting one month ahead SIE [2]. Kim et al. proposed a multitask Convolved Neural Network (CNN) LSTM technique using atmospheric, oceanic, and pixel-wise sea ice concentrations. Their DL model predicts derived Arctic SIE from concentration, achieving an RMSE of 0.303 million km² [6]. Overall, DL has demonstrated significant potential in SIE prediction. However, the enlisted models fail to directly take advantage of intra-yearly and inter-yearly trends, likely contributing to poor extremum accuracy and inadequate intra-extrema performance. In addition, only select models have SIE visualization capabilities, and none have implemented an algorithm to remove cyclic trends.

In light of these shortcomings, this study applies climatic trends and ML to predict Arctic SIE one month in advance with comparable or improved results to current models. This paper has the following contributions: 1) We propose FFTstack, an iterative Fourier Transform-based system for modeling cyclic temporal trends; 2) We further propose streamlined architectures, which accept both multivariate time series and spatiotemporal data, for ensembling FFTstack with linear re-

TABLE I
OVERVIEW OF DATASETS

Dataset	Domain	Interval	Shape	Shape details	Description
2D time series [1]	01/1979 to 08/2021	Monthly	(512, 10)	(time, variable)	Multivariate time series of 9 Arctic meteorological variables and sea ice extent
3D time series [4]	01/1979 to 08/2021	Monthly	(512, 448, 304)	(time, longitude, latitude)	Spatiotemporal time series of SIE images

TABLE II
2D DATASET VARIABLES

Variable	Range	Unit
Air temperature	[200, 350]	K
Sea surface temperature	[200, 350]	K
Surface pressure	[400, 1100]	hPa
Specific humidity	[0, 0.1]	kg/kg
Shortwave down radiation	[0, 1500]	W/m ²
Longwave down radiation	[0, 700]	W/m ²
Rain rate	[0, 800]	mm/day
Snowfall rate	[0, 200]	mm/day
Wind velocity	[0, 40]	m/s
Sea ice extent	[4, 13]	million km ²

gression, ML, and/or existing research models; 3) We perform a detailed comparative analysis of our proposed models and those produced by existing research. Our implementation code can be accessed at the GitHub repository¹.

II. DATASETS AND PROBLEM STATEMENT

A. Datasets

To forecast SIE, we use two datasets of varying shapes and structures. The multivariate architecture uses the same multivariate time series as in [1] that spans 01/1979 to 08/2021. It includes daily entries (n=15584) of nine Arctic atmospheric and oceanic variables, and sea ice extent, as enlisted in Tables I and II. Here, the target variable is sea ice extent. The monthly (n=512) and yearly (n=43) multivariate time series were derived from the daily data through averaging. For brevity, ‘2D data’ refers to the monthly version of this multivariate time series, and ‘2D architecture’ refers to the architecture capable of processing 2D data.

The 3D architecture used the Monthly Sea Ice Index, a spatiotemporal time series, provided by the National Snow and Ice Data Center in [4]. For this application, SIE images were used, with a concentration threshold of 15% and spatial resolution of 25km x 25km. The images are of shape (448, 304). This dataset spans the same time period as the previous. For brevity, ‘3D data’ refers to this spatiotemporal time series, and ‘3D architecture’ refers to the architecture capable of processing 3D data.

B. Problem Statement

Given N months of historic meteorological and sea-ice data X , learn a function to forecast sea ice extent Y for the next

one month, as shown in Equation (1) where $t \in (1, 2, \dots, N)$.

$$Y_{t+1} = f(X_{t-n}, X_{t-n+1}, \dots, X_t) \quad (1)$$

III. METHODS

This study evaluates the potential of leveraging temporal trends in data to predict Arctic SIE numerically and visually. First, we propose FFTstack, a novel approach to using Fourier Transform for modeling cyclic trends. Then, we explore the potential for applying FFTstack to Arctic sea ice prediction via integration with streamlined architectures. For the 2D architecture, we evaluate the potential to train an ML model on the deviations from the FFTstack-detected trends. We generate various baseline 2D models for comparison. For the 3D architecture, we apply FFTstack in a pixel-wise manner. A general overview of these processes is shown in Figure 1.

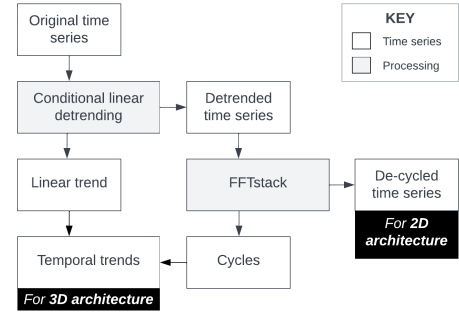


Fig. 1. Overview of Methodology

A. Data Preprocessing

To begin, both 2D and 3D data were split into 80:20 ratio, where approximately 80% was reserved for training (n=408), i.e. from 01/1979 to 12/2012, and 20% was set aside for testing (n=104), i.e. from 01/2013 to 08/2021.

Each monthly time series variable, x , underwent conditional detrending, collectively forming ‘detrended data’. First, a linear regression model was trained on the yearly time series, x_{year} . If the trend was statistically significant ($p < 0.05$), the linear regression model was trained on the monthly time series. The trend was subtracted from the monthly time series to obtain a detrended time series for that variable. If the trend was not statistically significant, the average of the monthly time series was subtracted from the monthly time series itself to center it around zero. The final trend, whether linear regression or average, is denoted by l . Thus, only variables with statistically significant inter-yearly trends were detrended,

¹github.com/big-data-lab-umbc/sea-ice-prediction/tree/main/fftstack-icmla

thereby preventing variables with trends occurring by chance from being detrended as well, as shown in Equation (2).

$$l_x = \begin{cases} \text{linreg}(x) & , p(\text{linreg}(x_{year})) < 0.05 \\ x - \bar{x} & , p(\text{linreg}(x_{year})) \geq 0.05 \end{cases} \quad (2)$$

B. FFTstack

Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) are computationally efficient implementations of mathematical transforms that conduct analysis from time to frequency and frequency to time domains, respectively.

1) *Composite Fourier Transform Model*: Individual FFT-based cycles, c , were generated with an FFT-based model for the monthly detrended time series of every variable, x . Two parameters were specified per variable: number of iterations, a , and threshold, b . The following process was iterated a times. First, FFT was conducted on the detrended time series. The resulting complex frequency domain values were converted to real numbers and normalized. By scaling FFT output to a set range between 0 and 1, normalization eased the detection of a threshold.

Next, a binary mask of booleans removed noise from the FFT output, where any frequencies with a normalized amplitude greater than or equal to the specified threshold were assigned 1, while all others were assigned 0. Then, the binary mask was applied to the FFT frequencies by element-wise multiplication. IFFT was subsequently conducted on the cleaned FFT frequencies, effectively returning cycles that modeled existing oscillation. If the number of iterations exceeded one ($a > 1$), the process would repeat with the residual of the cycles serving as the new input, thereby removing leftover oscillation from the previous iteration. The sequential removal of trends in these residuals was more robust against overfitting the FFT model than determining a more precise threshold. Thus, a combination of the two was used instead, as illustrated in Equation (3).

$$c_x = \sum_{i=1}^a \text{IFFT} \left(\text{FFT} \left(x - \sum_{j=1}^i c_j \right) * \left[\text{norm} \left(\left| \text{FFT} \left(x - \sum_{j=1}^i c_j \right) \right| \right) \geq b \right] \right) \quad (3)$$

FFTstack returned either the sum of all cycles generated by each iteration, called 'composite cycles', or the residual of the composite cycles, called 'de-cycled data'. The 2D architecture will use the de-cycled data, while the 3D architecture will use the composite cycles instead.

2) *Grid search*: Grid search was employed to optimize the FFT-based model. First, detrended was further split into an 80:20 ratio, with approximately 80% reserved for training ($n=324$), i.e. from 01/1979 to 12/2005, and 20% reserved for validation ($n=84$), i.e. from 01/2006 to 12/2012. Due to the generally strong periodicity of Arctic meteorological variables and conservative modeling technique of FFTstack, additional splitting was not considered necessary to ensure a well-fit

model, as substantiated by Section IV-2. Then, the algorithm determined the parameter combination with the lowest residual metric for the training set. The trend in the training data was extrapolated to the testing data and de-cycled accordingly. The search space encompassed a number of iterations and threshold over 45 combinations, as shown in Table III.

3) *Residual metric*: The residual metric, RM , was designed to evaluate the size of the residual while preventing overfitting. Median Absolute Deviation (MAD) was selected as a robust measure of variability. The percent change between the MAD of the train and validation partition was calculated. If the percent change was greater than 0.1, the cycles were considered overfit and the metric was set to infinity. Otherwise, the metric was set equal to the average MAD of both the train and test time series, as shown in Equation (4). Because lower variability implied a smaller residual, a lower residual metric was desirable.

$$RM = \left[\frac{MAD(x_{train}) - MAD(x_{val})}{MAD(x_{train})} \leq 0.1 \right] * \left[\frac{MAD(x_{train}) + MAD(x_{val})}{2} \right] \quad (4)$$

C. Architectures

Two architectures were developed to accommodate the differing shapes and structures of SIE data. An overview of these architectures is available in Table IV.

1) *2D architecture*: The 2D architecture accepts multivariate time series data for univariate time series prediction of SIE. To accomplish this, an ML model was trained on the FFTstack-generated residuals, a process referred to as 'residual learning' (not to be confused with the DL term). We used Gradient Boosting Regressors (GBR), a type of ML algorithm that ensembles weak learners, due to their capacity for predicting complex patterns such as those found in the pattern-neutral residuals.

The proposed 2D GBR (de-cycled) model accepted 2D de-cycled data explained in Section III-B as input. Two baseline models were also generated for comparison, with the 2D GBR (original) model and 2D GBR (detrended) model accepting unaltered 2D data and 2D detrended data, respectively.

First, an expanding window splitter with two folds was applied to the inputted training data to ensure consistent performance. In Fold 1, the 2D data was split into a 2:3 ratio, where approximately 66% was reserved for training ($n=204$), i.e. from 01/1979 to 12/1995, and 33% was set aside for validation ($n=104$), i.e. from 01/1996 to 08/2004. In Fold 2, the 2D data was split into a 3:4 ratio, where approximately 75% was reserved for training ($n=308$), i.e. from 01/1979 to 08/2004, and 25% was set aside for validation ($n=100$), i.e. from 09/2004 to 12/2012.

Then, grid search optimized a recursive autoregressive GBR for lowest Root Mean Squared Error (RMSE), determining optimal hyperparameter and lag combinations. The search spaces are provided in Table III.

TABLE III
HYPERPARAMETER SEARCH SPACES

Model	Used by	Parameter	Values Format: (start, end, step=1) or [item_1, ... , item_n]
Composite FFT	FFTstack, 2D/3D architectures	n_iteration	(1, 5)
		threshold	(0.1, 0.9, 0.1)
Gradient Boosting	2D architecture only	n_estimators	(25, 500, 25)
		max_depth	(1, 5)
		learning_rate	[1, 0.5, 0.1, 0.05, 0.01]
		lags	(12, 60, 12)

TABLE IV
SUMMARY OF 2D AND 3D ARCHITECTURES

Model	Type	Architecture
2D GBR	Baseline	2D data $\rightarrow$ GBR $\rightarrow$ SIE numerical prediction
2D GBR (detrended)	Baseline	2D data $\rightarrow$ linear detrending + GBR $\rightarrow$ SIE numerical prediction
2D FFTstack	Proposed	2D data $\rightarrow$ linear detrending + FFTstack $\rightarrow$ Cyclic trends
2D GBR (de-cycled)	Proposed	2D data $\rightarrow$ linear detrending + FFTstack + GBR $\rightarrow$ SIE numerical prediction
3D FFTstack	Proposed	3D data $\rightarrow$ FFTstack $\rightarrow$ SIE image prediction

As denoted in Equation (5), the final prediction of the 2D GBR (de-cycled) model was obtained by adding the linear trend, L , composite cycles, C , and optimized GBR predictions based on n past months of de-cycled data, R .

$$Y_{t+1} = (L + C)_{t+1} + GBR(R_{t-n}, R_{t-n+1}, R_t) \quad (5)$$

2) *3D architecture*: The 3D architecture applies FFTstack to images to forecast the temporal trends contained in each pixel. First, the 3D data was flattened and transposed, such that the training and testing data had shapes of (136192, 408) and (136192, 104), respectively. Then, FFTstack was applied to each of the 136192 pixel time series, with cycles of temporal trends returned as the final output. Finally, all pixel time series predictions were rearranged into a 3D shape, such that the outputted training and testing forecasts occurred in shapes of (408, 448, 304) and (104, 448, 304), respectively.

D. FFTstack Integration with Current Research

To evaluate its applicability to existing research, FFTstack was integrated into the frameworks of [1] and [2]. For both, FFTstack was applied to the input data. Then, the existing models were trained on the de-cycled data in accordance with the residual learning framework entailed in Section III-C1.

IV. RESULTS

We present results from two architectures encompassing five models which ensemble various linear regression, FFTstack, and ML algorithms. In addition, we evaluate the effect of integrating FFTstack with models from existing research.

1) *Performance metrics*: To evaluate the SIE prediction capability of both study and out-of-study models, performance metrics including RMSE, Normalized Root Mean Squared Error (NRMSE), calculated by dividing RMSE by the mean of the dataset, and R^2 were used. To evaluate long-term relevancy, we calculated the p-value of a linear regression model trained on averaged yearly predictions. A non-statistically significant

trend ($p \geq 0.05$) indicated the model maintained consistent performance over time.

The MAD percentage of initial was specifically used for the FFTstack model to evaluate the size of the residual relative to the input, as shown in Equation (6).

$$\% \text{ of initial} = \frac{MAD(x - [l_x + c_x])}{MAD(x)} \quad (6)$$

A smaller percentage suggested FFTstack removed more oscillation, implying a superior fit. Thus, FFTstack removed most oscillation from variables with a percentage less than 0.1, whereas only adequate removal occurred for those with a percentage less than 0.5. A percentage outlying above 0.75 suggested that FFTstack removed minor oscillation.

2) *2D FFTstack*: The variables enlisted in Table V exhibited statistically meaningful linear trends and were detrended. All variables had at least one parameter combination that attained a MAD percent change between training and validation less than 0.10, suggesting the optimal FFTstack models were not overfit. FFTstack demonstrated particularly strong performance for five variables, but poor performance for surface pressure, likely due to its low periodicity.

3) *Architectures*: As shown in Table VI, each optimized 2D model contained a lag value consistent with a year, suggesting the hyperparameter search space was sufficiently large. The baseline and test model were both slightly overfit, exhibiting higher test RMSE than train RMSE. The de-cycled model was well-fit, exhibiting a test RMSE less than train RMSE. The detrended and baseline model achieved successively higher R^2 . The de-cycled model achieved the lowest R^2 , likely due to the noisy nature of its input.

4) *Comparative Analysis*: The 2D GBR (de-cycled) model outperformed all the study models in all performance metrics, as shown in Table VII. This model exhibited NRMSE percent decreases of 53.31%, 37.47%, and 40.63% and R^2 percent increases of 2.80%, 1.12%, and 1.43% for the 2D GBR

TABLE V
DETRENDER AND FFTSTACK MODEL STATISTICS AND PARAMETER CONFIGURATIONS

Variable	Detrender	FFTstack			
	<i>P-value</i>	<i>Parameters</i>	<i>% change</i>	<i>Residual metric</i>	<i>% of Initial</i>
Air temperature	<0.001	{‘n_iteration’: 2, ‘threshold’: 0.10}	0.079	0.594	0.054891
Sea surface temperature	<0.001	{‘n_iteration’: 1, ‘threshold’: 0.05}	0.077	0.159	0.201613
Surface pressure	0.75	{‘n_iteration’: 1, ‘threshold’: 0.25}	0.049	1.574	0.786065
Specific humidity	<0.001	{‘n_iteration’: 2, ‘threshold’: 0.10}	0.078	0.069	0.083597
Shortwave down radiation	0.25	{‘n_iteration’: 2, ‘threshold’: 0.05}	-0.044	0.685	0.009984
Longwave down radiation	<0.001	{‘n_iteration’: 4, ‘threshold’: 0.55}	0.009	1.903	0.043795
Rain rate	0.03	{‘n_iteration’: 2, ‘threshold’: 0.10}	0.058	0.068	0.244662
Snowfall rate	0.007	{‘n_iteration’: 4, ‘threshold’: 0.85}	0.094	0.047	0.257701
Wind velocity	0.38	{‘n_iteration’: 4, ‘threshold’: 0.75}	-0.030	0.083	0.269971
Sea ice extent	<0.001	{‘n_iteration’: 2, ‘threshold’: 0.25}	0.081	284681.194	0.099539

TABLE VI
2D GRADIENT BOOSTING REGRESSOR PARAMETER CONFIGURATIONS AND PERFORMANCE STATISTICS

Model	Train			Test	
	<i>Hyperparameters</i>	<i>Lags</i>	<i>RMSE</i>	<i>RMSE</i>	<i>R²</i>
GBR (original)	{‘learning_rate’: 0.1, ‘max_depth’: 2, ‘n_estimators’: 325}	36	565069.333	637720.768	0.965
GBR (detrended)	{‘learning_rate’: 0.05, ‘max_depth’: 2, ‘n_estimators’: 275}	12	409503.288	475498.916	0.981
GBR (de-cycled)	{‘learning_rate’: 0.05, ‘max_depth’: 3, ‘n_estimators’: 175}	12	358979.165	298025.108	0.610

TABLE VII
COMPARATIVE PERFORMANCE OF STUDY AND EXISTING RESEARCH MODELS

Model	Description	Data	Testing Performance			
			<i>RMSE (mil. km²)</i>	<i>NRMSE</i>	<i>R²</i>	<i>P-value</i>
Wang et al. [14]	Statistical model	Meteorological variables	0.45	-	-	-
Wang et al. [15]	Physical model	Meteorological variables	0.2–0.6	-	-	-
Ali et al. [1]	LSTM model	10 meteor. and SIE vars.	0.660	0.0628	0.963	1.0
[1] with FFTstack	LSTM + FFTstack	10 meteor. and SIE vars.	0.412	0.0400	0.986	0.72
Ali et al. [2]	Linear regression	11 meteor. and SIE vars.	0.453	0.0431	0.982	1.0
[2] with FFTstack	Lin. reg. + FFTstack	11 meteor. and SIE vars.	0.356	0.0346	0.990	0.72
Kim et al. [6]	CNN/LSTM model	10 meteor. and SIE vars. and SIE images	0.303	0.0292	-	-
3D FFTstack	See Table IV	SIE images	0.361	-	-	0.075
2D GBR (original)	See Table IV	9 meteor. vars. and SIE	0.638	0.0604	0.965	0.015
2D GBR (detrended)	See Table IV	9 meteor. vars. and SIE	0.475	0.0451	0.981	0.49
2D FFTstack	See Table IV	SIE variable	0.502	0.0475	0.978	0.14
2D GBR (de-cycled)	See Table IV	9 meteor. vars. and SIE	0.298	0.0282	0.992	0.29

(original), 2D GBR (detrended), and 2D FFTstack models, respectively. Thus, the 2D GBR (de-cycled) model was overall most adept at predicting SIE.

The averaged yearly performance of each model is charted in Figure 2B. The GBR (baseline) model exhibited a statistically significant long-term trend, suggesting its predictions are only meaningful in the short term. On the other hand, the 2D GBR (detrended), 2D FFTstack, and 2D GBR (de-cycled) models did not have statistically meaningful trends, suggesting longer-term prediction relevancy.

Figure 2C shows a month-of-year RMSE comparison of all 2D study models. The detrended model outperformed the de-cycled model in 50% of the months, including February, March, May, August, September, and December. Thus, the detrended model was more accurate in predicting extrema months, namely March and September. It should be noted that during these months, the detrended model exhibited only slightly better RMSE; the results were fairly comparable. In the remaining months, the de-cycled model tended to

outperform the detrended model, especially in October and November. Therefore, while the 2D GBR (detrended) model may have outperformed the 2D GBR (de-cycled) model in some months by small margins, the 2D GBR (de-cycled) model markedly outperformed the detrended model in the remaining months. Though the 3D architecture was outperformed by [6], it effectively removed most inter-yearly and seasonal cycles from the SIE images, as shown in Figure 3B.

The 2D GBR (de-cycled) model outperformed all other comparable models taken from literature. In comparison to this model, [1], [2], and [6] exhibited 50.26%, 31.55%, and 3.42% higher NRMSE, respectively. Comparatively small increases in R^2 were also present. In addition, when FFTstack was integrated into their frameworks, [1] and [2] underwent 36.31% and 19.72% decreases in NRMSE, respectively.

V. CONCLUSIONS

This study proposes FFTstack, a process that effectively detects and/or removes temporal trends from time series data.

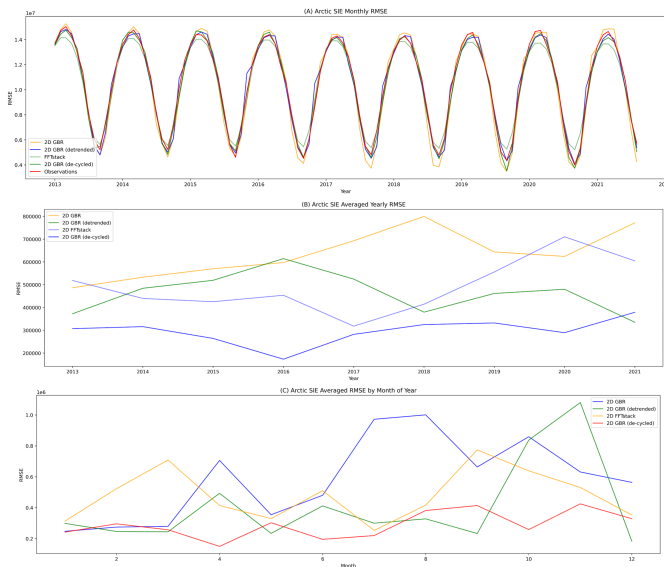


Fig. 2. Comparative RMSE analysis by (A) month, (B) year, and (C) month of year for 2D models

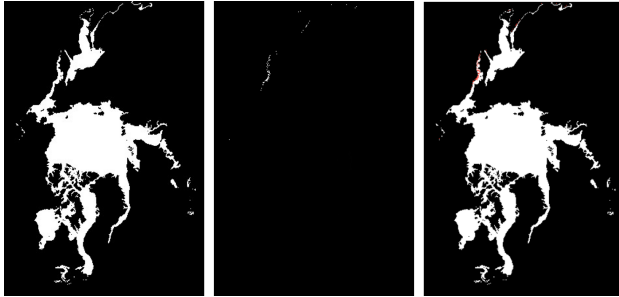


Fig. 3. From left to right, where white and red represent SIE and model error (02/2016 used as example): SIE input image, 3D FFTstack residual, 3D FFTstack composite cycles

By integrating FFTstack into a proposed 2D architecture, the 2D GBR (de-cycled) model outperformed all baseline models and existing state-of-the-art, ML, DL, and probabilistic models in RMSE, NRMSE, and R^2 metrics. Due to its superior performance, the proposed model can meaningfully contribute to predictive models of related research, such as predicting sea level rise, species decline, or anomalous weather patterns. In addition, the lack of a statistically meaningful long-term linear trend suggests that this model's predictions will remain applicable in the future, unlike the predictions of several existing models. In future research, we will investigate changes in model performance on the daily interval and/or at longer lead times. We will also implement a DL model to the 3D architecture, which currently only supports FFTstack, in order to complete the residual framework. FFTstack also improved the performance and long-term relevancy of existing models, demonstrating its widespread compatibility and utility in similar research. In conclusion, this research represents an important step in guiding mitigation, resilience, and adaptation efforts in response to Arctic sea ice decline. For example,

the 2D GBR (de-cycled) model can be used as mechanism of awareness, or a statistic guiding local and governmental action. Though more detailed and accurate data is needed, the 3D model can be used to map new shipping routes, understand changing ecosystems, and help indigenous peoples adjust to a changing environment. In short, understanding sea ice decline is the first step in understanding vastly more research, and it therefore stands as a first step to worthwhile global responses.

ACKNOWLEDGMENT

Lapp acknowledges the Ingenuity Project with additional thanks to research coordinator Ms. Kowsar Ahmed and research director Dr. Nicole Rosen. Wang and Ali are supported by NSF grants OAC-1942714 and OAC-2118285.

REFERENCES

- [1] S. Ali, Y. Huang, X. Huang, and J. Wang. Sea ice forecasting using attention-based ensemble lstm. *arxiv e-prints*, page. *arXiv preprint arXiv:2108.00853*, 2021.
- [2] S. Ali, S. A. Mostafa, X. Li, S. Khanjani, J. Wang, J. Foulds, and V. Janeja. Benchmarking probabilistic machine learning models for arctic sea ice forecasting. In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, pages 4654–4657. IEEE, 2022.
- [3] M. Casas-Prat and X. L. Wang. Sea ice retreat contributes to projected increases in extreme arctic ocean surface waves. *Geophysical Research Letters*, 47(15):e2020GL088100, 2020.
- [4] K. K. W. N. M. M. S. Fetterer, F. and A. K. Windnagel. Sea ice index, version 3, 2017.
- [5] F. S. Hu, P. E. Higuera, J. E. Walsh, W. L. Chapman, P. A. Duffy, L. B. Brubaker, and M. L. Chipman. Tundra burning in alaska: linkages to climatic change and sea ice retreat. *Journal of Geophysical Research: Biogeosciences*, 115(G4), 2010.
- [6] E. Kim, P. Kruse, S. Lama, J. Bourne, M. Hu, S. Ali, Y. Huang, and J. Wang. Multi-task deep learning based spatiotemporal arctic sea ice forecasting. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 1847–1857. IEEE, 2021.
- [7] B. Light, T. C. Grenfell, and D. K. Perovich. Transmission and absorption of solar radiation by arctic sea ice during the melt season. *Journal of Geophysical Research: Oceans*, 113(C3), 2008.
- [8] I. Mahlstein and R. Knutti. September arctic sea ice predicted to disappear near 2 c global warming above present. *Journal of Geophysical Research: Atmospheres*, 117(D6), 2012.
- [9] T. Pearce, B. Smit, F. Duerden, J. D. Ford, A. Goose, and F. Kataoyak. Inuit vulnerability and adaptive capacity to climate change in ulukhaktok, northwest territories, canada. *Polar Record*, 46(2):157–177, 2010.
- [10] E. V. Regehr, C. M. Hunter, H. Caswell, S. C. Amstrup, and I. Stirling. Survival and breeding of polar bears in the southern beaufort sea in relation to sea ice. *Journal of animal ecology*, 79(1):117–127, 2010.
- [11] S. R. Stephenson, L. C. Smith, L. W. Brigham, and J. A. Agnew. Projected 21st-century changes to arctic marine access. *Climatic Change*, 118:885–899, 2013.
- [12] J. Stroeve, M. M. Holland, W. Meier, T. Scambos, and M. Serreze. Arctic sea ice decline: Faster than forecast. *Geophysical research letters*, 34(9), 2007.
- [13] T. Vihma. Effects of arctic sea ice decline on weather and climate: A review. *Surveys in Geophysics*, 35:1175–1214, 2014.
- [14] L. Wang, X. Yuan, M. Ting, and C. Li. Predicting summer arctic sea ice concentration intraseasonal variability using a vector autoregressive model. *Journal of Climate*, 29(4):1529–1543, 2016.
- [15] W. Wang, M. Chen, and A. Kumar. Seasonal prediction of arctic sea ice extent from a coupled dynamical forecast system. *Monthly Weather Review*, 141(4):1375–1394, 2013.
- [16] B. Wu, R. Zhang, R. D'Arrigo, and J. Su. On the relationship between winter sea ice and summer atmospheric circulation over eurasia. *Journal of Climate*, 26(15):5523–5536, 2013.
- [17] J. Zhang, Y. H. Spitz, M. Steele, C. Ashjian, R. Campbell, L. Berline, and P. Matrai. Modeling the impact of declining sea ice on the arctic marine planktonic ecosystem. *Journal of Geophysical Research: Oceans*, 115(C10), 2010.