Multi-Task Spatiotemporal Deep Learning Based Arctic Sea Ice Prediction

REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering

Jamal Bourne Jr.[1], Michael Hu[2], Eliot Kim[3], Peter Kruse[4], Skylar Lama[5], Sahara Ali[6], Yiyi Huang[7], Jianwu Wang[6]

[1]Department of Mathematics and Computer Science, McDaniel College
[2]Department of Computer Science, Georgia Institute of Technology
[3]Nelson Institute of Environmental Studies Department of Statistics, University of Wisconsin-Madison
[4]Department of Accounting, Business, and Economics, Juniata College
[5]Department of Atmospheric and Oceanic Science, University of Maryland, College Park
[6]Department of Information Systems, University of Maryland, Baltimore County
[7]Science Systems and Applications, Inc.

Technical Report HPCF–2021–11, hpcf.umbc.edu > Publications

Abstract

Important natural resources in the Arctic rely heavily on sea ice, making it important to forecast Arctic sea ice changes. Arctic sea ice forecasting often involves two connected tasks: sea ice concentration at each pixel and overall sea ice extent. Instead of having two separate models for two forecasting tasks, in this report, we study how to use multi-task learning techniques and leverage the connections between ice concentration and ice extent to improve accuracy for both prediction tasks. Because of the spatiotemporal nature of the data, we designed two novel multi-task learning models based on CNNs and ConvLSTMs, respectively. We also developed a custom loss function which trains the models to ignore land pixels when making predictions. Our experiments show our models can have better accuracies than separate models that predict sea ice extent and concentration separately, and that our accuracies are better than or comparable with results in the state-of-the-art studies.

Key words. Arctic Sea Ice; Machine Learning; Deep Learning; Multi-Task Learning

1 Background

Arctic sea ice is essential to oceanic currents, atmospheric processes, and polar ecosystems. Historically, Arctic sea ice has exhibited annual fluctuations, reaching minimum concentrations in September and maximum concentrations in March. Despite this precendent of seasonal fluctuation, over the past few decades there has been a much larger decline in sea ice during months with minimum sea ice concentration (SIC). The September Arctic sea ice has disappeared by almost 50% in the past 40 years. Since 1981, observations show SIC declining at a rate of 13.1% per decade. In 1980 the September SIC was 7.67 million square kilometers, while the most recent record low to date was in 2012 at 3.57 million square kilometers [28]. These declines in sea ice intensified starting in September of 2002. With new summer lows followed by insufficient winter recovery, sea ice has only declined faster especially as global warming increases causes yearly temperature extremes in the atmosphere and ocean. On current trends, there are possibilities of an ice-free September for the Arctic Ocean as early as mid-21st century [4, 27].

Declining sea ice extent will have extreme consequences for various stakeholders. On the regional scale, estuaries and wildlife in the Arctic will suffer due to a lack of sea ice. Wildlife such as polar bears, arctic foxes, walruses, and many other species rely heavily on ice for their livelihoods. This direct effect on Arctic estuaries and wildlife creates a chain of events that eventually affects humans

as well, specifically when it comes to food supply [29]. Another impact of melting sea ice is the opening of more direct shipping routes between Asia, Europe, and North America. However, these time saving routes also pose many other dangers. Ocean that has now become open by the loss of sea ice is unexplored, meaning it could be dangerous for crews to sail across these routes. Along with these routes being dangerous for humans to cross, they could also be homes of ecosystems that have been protected by large amounts of sea ice. Uncovering large areas of the ocean that have been protected by ice for hundreds of years poses great danger to these once-protected ecosystems.

On a global scale, the decline in sea ice will also affect the atmosphere and climate as a whole. The Arctic is known as the world's refrigerator, helping to cool our atmosphere and prevent rapid warming [29]. Sea ice contributes to this phenomenon due to its high albedo. Albedo is the amount of radiation from a light source that is reflected by an object's surface. Objects with lighter colored surfaces and high latent heat values (requires more energy to heat) have higher albedos. Sea ice, being a high-albedo surface, naturally reflects solar radiation from the sun back into the atmosphere [19,32]. This aspect of sea ice is essential to the Arctic and our climate. However, as sea ice has declined, there is less ice to reflect solar radiation back into the atmosphere, causing warmer Arctic temperatures and altering global climatic patterns, specifically by creating heat waves.

While the Arctic may experience heatwaves, other parts of the world will in turn experience the polar vortex. The polar vortex contains cold air in the northern hemisphere by a jet stream with extremely strong winds [37]. As the polar vortex weakens, this cold air can move south and affect other areas of the world with extreme cold temperatures. These cold temperatures pose threats to various human activities, especially agriculture.

The consequences of melting Arctic sea ice are extremely detrimental not only to the Arctic but also the rest of the world. When such extreme changes occur in an important part of our world's climate, many different snowball effects occur. Because of these dangers, it is extremely pertinent to study the changes in SIC to learn more about what is causing them and what the future holds for the climate.

2 Related Works

2.1 Sea Ice Prediction Methods

Accurate predictions of Arctic sea ice extent and concentration have proven difficult, and the acceleration of climate change has further exacerbated this challenge [34]. Numerical, statistical, and machine learning methods have been used to make sea ice predictions, but there is still room for more accurate methods to be explored. Regression techniques have provided adequate sea ice predictions results for up to 7 month lead times, using sea ice data [20] or additional predictors [15]. Wang et al. used vector autoregression, involving a multivariate time series model, to predict daily summertime sea ice at an intraseasonal timescale of 20-60 days. Reasonable accuracies were achieved in this study using past sea ice as the sole predictor [36]. All of these methods have been able to capture general trends in Arctic sea ice, but a more efficient and accurate method for predicting is necessary for applicability.

Deep learning techniques have become useful in climate modeling, especially for incorporating complex data sources. According to recent studies, deep learning models are better suited than statistical or numerical methods to capture the complex interactions between environmental predictors which impact ice concentrations [31], [24], [3], [22]. However, deep learning techniques still have room for improvement when it comes to accurately modeling both sea ice concentrations and sea ice extent.

2.2 Multi-task Learning Models

Similar to this study, other studies have predicted sea ice using models similar to CNN and ConvLSTM. Chi et al. (2017) used two deep learning models, a multilayer perceptron and a long-short term memory model, to predict monthly 2015 Arctic sea ice concentrations [7]. Sea ice concentration for the preceding 12 months was the only input variable used for their models, but this study achieved an RMSE of 8.89% using LSTM for 2015 sea ice concentrations at 25km x 25km resolution. Kim et al. (2020) trained a convolutional neural network (CNN) as well as random forests to make one month-ahead monthly ice predictions. Monthly meteorological and past ice values were used as predictors. The CNN had the best performance, with an overall RMSE of 5.76% for predictions of Arctic sea ice from 2000-2017 [23]. Liu et al. (2021) compared the performance of CNN and ConvLSTM models in predicting Arctic sea ice concentrations at the daily scale for 2018. The spatial domain was divided into 20 sub-grids, and the two previous days were used to predict the next day's ice concentrations. The CNN had an average RMSE of 8.058%, and the ConvLSTM had an improved 6.942% RMSE for 2018 sea ice [25]. These studies highlight the promise of deep learning models for producing accurate ice predictions at a high spatiotemporal resolution.

This study builds upon the importance of using deep learning techniques in Arctic sea ice prediction by creating a more complex architecture while implementing multi-task models and a custom loss function to improve accuracy. With these unique approaches, this study's models are able to produce more accurate predictions with comparable or lower sea ice extent and concentration RMSE values than previous studies.

3 Data

This study uses sea ice, atmospheric, and meteorological data from 1979 through 2020 covering the Arctic Ocean and adjacent land areas.

Sea ice concentration data was provided by the National Snow and Ice Data Center and obtained from the Nimbus-7 SSMR and DMSP SSM/I-SSMIS passive microwave data version 1 (http://nsidc.org/data/NSIDC-0051) [6]. This dataset is generated daily in the polar stereographic projection using a grid box of 25 km x 25 km dating from October 1978 to present time. The dataset of sea ice concentration produces an uncertainty of about +-5% in the Arctic winter when sea ice tends to reach its peak in concentration levels. During summer months, this uncertainty increases to about +-15% as there are more melt ponds present which can skew data collection [6]. This concentration data was considered to be the ground truth during modeling.

Due to the European Centre for Medium-Range Weather Forecasts (ECMWF)'s reliability and consistency in independent observations over other global reanalysis products, atmospheric and meteorological variables were obtained from ECMWF's ERA-5 global reanalysis product (https://cds.climate.copernicus.eu/cdsapp!/home) [5, 9]. ERA-5 was produced using 4D-Var data assimilation in CY41R2 of ECMWF's Integrated Forecast System (IFS). With these systems, 137 hybrid sigma/pressure (model) levels in the vertical were used with top level at 0.01 hPa [5]. In this study salinity was excluded as it was not observed in the year 2019. Information on these atmospheric variables along with sea ice concentration are listed in Table 3.1.

The inclusion of each atmospheric and meteorological variable was based on their physical impact on sea ice trends. Air temperature is the main driver of changes in sea ice, and record low sea ice extents during recent melting seasons have been associated with warmer atmospheric temperatures [30]. The inclusion of sea surface temperature and 2 meter air temperature in the dataset provide comprehensive information regarding oceanic and atmospheric heat relevant for

Feature	Source	Units	Range
Sea Ice Concentration	NSIDC	% per pixel	0-100
Surface Pressure	ERA5	Pa	40000-110000
10m Wind Speed	ERA5	m/s	0-40
Near-Surface Humidity	ERA5	kg/kg	0-0.1
2m Air Temperature	ERA5	K	200-350
Shortwave Radiation	ERA5	W/m^2	0-1500
Longwave Radiation	ERA5	W/m^2	0-300
Rain Rate	ERA5	mm/day	0-800
Snow Rate	ERA5	mm/day	0-200
Sea Surface Temperature	ERA5	K	200-350

Table 3.1: Input Features for CNN and ConvLSTM models. All features are monthly-averaged and one-month lagged.

sea ice. Studies have also shown that Arctic circulation and wind patterns have seasonally varying relationships with sea ice [10, 17]. For example, poleward winds specifically play a key role in transporting heat to the Arctic, which contributes to ice melt [2,21,35]. Precipitation trends are also connected to sea ice patterns. In recent years, earlier rainfalls during spring have triggered earlier snowmelt and, via feedback loops, earlier Arctic ice melt [11, 26]. The complexity of atmospheric, oceanic, and sea ice interactions is illustrated in [16, 18], which highlights the pathway by which regional differences in atmospheric pressure facilitate increased Arctic humidity, which in turn enables higher levels of longwave radiation to reach the sea surface, leading to earlier melting of sea ice. Thus, each predictor impacts Arctic sea ice through complex physical interactions in the ocean and atmosphere.

3.1 Data Exploration

To begin our background research, we created climatologies and anomalies to visualize and analyze the dataset. We calculated the average sea ice extent for each month from 1979 through 2018, shown in Fig. 3.1. With this data, we were then able to calculate anomalies in specific years to identify years where sea ice extent was substantially lower than the average of all of the years of data used. In performing anomaly calculations of sea ice, we were able to see that 2012 had record low sea ice extent values (Fig. 3.2). Although the anomalies only reach a 6% difference from the monthly averages, it must be noted that the monthly averages were calculated over a 40 year time period where sea ice extent has changed dramatically. Knowing this, we can better identify the extremes that such anomaly values indicate.

With this information, we were able to further analyze which variables could have the largest impact on sea ice extent. One variable that seems to correlate the most with low sea ice extent is T2m values (Fig. 3.3). Along with related research that shows the affects of temperature on SIE [30], Figure 3.3 shows much higher temperatures 2 meters above sea surface level occurring in summer months, which can be correlated to the decline in SIE in September. This correlation, however, is not as evident in other variables such as sea surface temperature (SST). Although SST does correlate with the changes in our climate and can affect sea ice extent, it does not have as strong of an impact as other factors such as T2m.

Time series analysis was conducted to provide baseline models for comparison with our deep learning results. By using Seasonal Autoregressive Integrated Moving Average (SARIMA), we were able to build a forecasting model based on time series seasonality of sea ice concentration. Our model predicted future average sea ice concentration, but the results were not promising. Not only

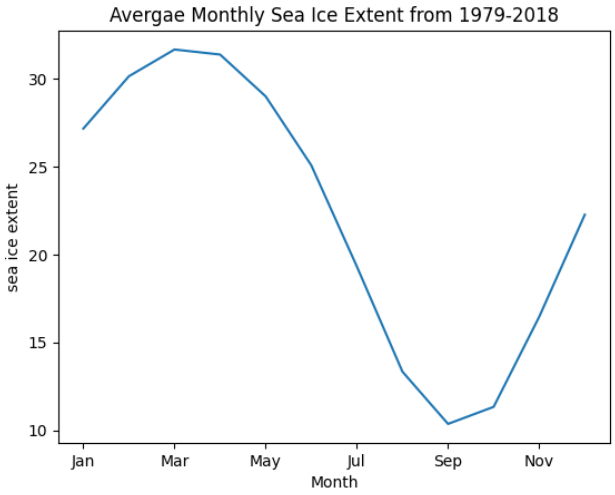

Figure 3.1: Monthly averaged sea ice extent values from 1979-2018 were averaged together grouping by month to create a 40 year climatology representing sea ice extent percentage values. These values were further used in calculating sea ice extent anomalies.

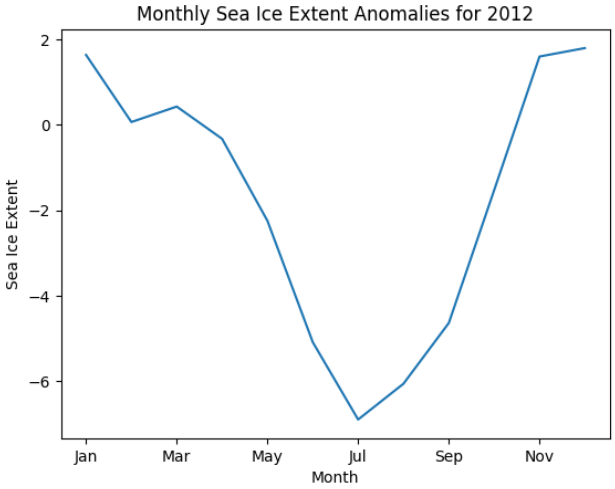

Figure 3.2: Plot of anomalies created by averaging monthly values of sea ice concentration from 1979-2018 and subtracting those values from the average monthly sea ice extent values for 2012 resulting in 12 total anomaly values measured by percentage of sea ice concentration.

did the model produce negative values, but it also strongly underestimated the Arctics freezing seasons.

Along with using SARIMA, we forecasted sea ice extent using a Vector Autoregression (VAR) model. This statistical model captures the fluctuations and changes in data over time. VAR proved to be fairly accurate in comparison to SARIMA. Results showed a basic understanding of the changes in sea ice extent averages through seasons but seemed to struggle when calculating maximums and minimums in the time series (Fig. 3.4).

The purpose of doing this analysis prior to implementing deep learning techniques was to get an understanding of the data we were working with. By plotting changes in the data and running the data through time series analysis, we were able to understand seasonal changes in the data and gain knowledge that would be useful in further analyzing results produced in the deep learning models. Knowing seasonality changes along with the basic structure of the data proved to be essential in troubleshooting our later deep learning work.

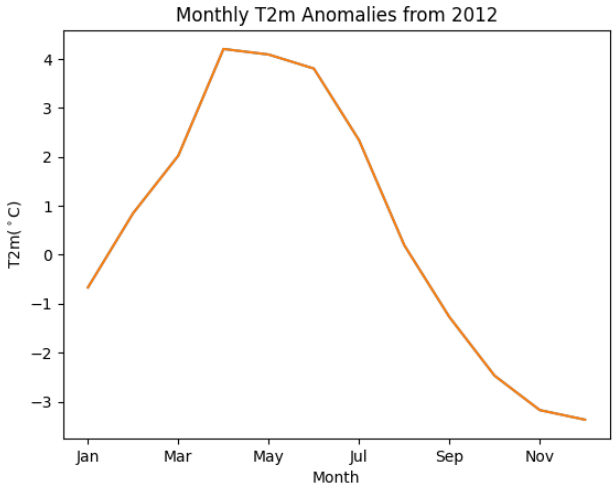

Figure 3.3: Plot of anomalies created by averaging monthly values of T2m from 1979-2018 and subtracting those values from the average monthly T2m values for 2012 resulting in 12 total anomaly values.

After completing the time series analysis we were able to implement deep learning techniques in our data, building neural networks in order to predict future sea ice extent.

3.2 Data Preprocessing

All variables were averaged from a daily resolution to the monthly scale. Prior to model-specific pre-processing, the dataset had 504 images, each with 448 by 304 grid cells and 10 channels, corresponding to the 10 input features details in Table 3.1.

3.3 Convolutional Neural Network Data Prepreocessing

CNN models were trained on the first 407 months of the data (January 1979 - November 2012) and validated on the last 96 months (January 2013 - November 2020), with a one-month lead time. Each image in the dataset was considered to be an individual training example and was used to predict per-pixel sea ice concentrations for the next month. For example, the image corresponding to January 1979 was used to predict ice concentrations for February 1979. Thus, the training dataset learned per-pixel sea ice concentrations for February 1979 - December 2012, and the validation dataset predicted per-pixel sea ice concentrations for February 2013 - December 2020.

3.4 Convolutional LSTM Data Preprocessing

In order to fully capture the spatio-temporal nature of our data using a Convolutional LSTM, heavy data preprocessing was necessary. The model was trained on the first 408 months of the data and validated on the last 95 months of the data. In Keras, ConvLSTM2D layers require 5 dimensional inputs of shape *(samples, timesteps, rows, columns, features)*. To reshape the data, a stateless rolling window was applied to the training and testing data, creating 396 samples of 12 months each. Sample one contained months 1-12, sample two contained months 2-13, and the last sample contained months 395-407. The final shape of the training input data was 396 samples with 12 months of 448×304 pixel images, each containing 11 feature measurements at each pixel.

Figure 3.4: VAR model forecast vs actual values for variables; ice extent, shortwave radiation, longwave radiation, rain, surface pressure, humidity, sea surface temperature, temperature, wind, and snow.

Similarly, the final shape of the test input data was 84 samples with 12 months of 448×304 pixel images each, all containing 11 feature measurements at each pixel.

The validation data consisted of 396 images in the training set and 84 images in the test set. Each image contained the average sea ice concentration for the corresponding month in each pixel. The first sample of input data, consisting of the first 12 months of images, was used to predict the sea ice concentrations in the 13th month in the output data; the second sample was used to predict the SIC in the 14th month.

Including a rolling window with 12-month timesteps allowed the ConvLSTM to learn yearly variations and relationships in SIC, resulting in more accurate predictions.

4 Methods

4.1 Masked Loss Function

Neural networks use loss functions to measure the error in their predictions after each epoch. Once the error has been measured, the model optimizes the loss function using a process called back-propagation.

To help the model learn on ocean and ice pixels while ignoring land pixels, a custom loss function was implemented in the networks' architectures. A land mask was applied to each output of the network before loss was evaluated. In the mask, land pixels were given a value of 0 and non-land pixels were given a value of 1. Each predicted output and the mask were multiplied elementwise, resulting in land pixels being ignored when calculating the loss.

After applying the mask, the root mean squared error of the model's predicted and actual values was calculated. By applying land masks, the model learned to ignore land pixels in its calculations, allowing it to more accurately optimize sea ice concentrations for non-land pixels.

4.2 Convolutional Neural Network

CNNs are a type of deep learning model particularly suited for working with images, speech, and audio signals. Thus, we chose to implement it to process our per-pixel data, which is in the form of image data. CNNs are able to process multidimensional data. In our case, the input is a three-dimensional array, i.e., $height \times width \times channel$ $(448, 304, 11)$. CNNs consist of three types of layers: convolutional layers, pooling layers, and fully connected layers [12]. We will explain each layer type below in detail.

Convolutional layers are where the majority of computation is done. In this layer a small matrix of fixed weights, aka a kernel, is passed over an image to create a feature map of the image using the following equation:

$$g(x, y) = \omega \cdot f(x, y) = \sum_{dx=-a}^{a} \sum_{dy=-b}^{b} \omega(dx, dy) f(x + dx, y + dy) \tag{4.1}$$

Where *g(x, y)* is the feature map, $w(dx, dy)$ is the kernel, and $f(x, y)$ is the original image. First the kernel is applied to an area of the image, where the dot product of the input pixels and the filter are fed into an output array. Afterwards, the filter shifts by a stride, repeating the process until the kernel has swept across the entire image and created a feature map [33]. This allows the model to recognize patterns such as edges or curves which recur through the image.

Pooling layers are used for reducing the amount of parameters in the input. Similar to convolutional layers, pooling layers sweep a small matrix across the image; however this matrix does not contain weights, it instead applies an aggregation function to the values within the receptive field. The two types of pooling layers are max pooling and average pooling. Max pooling simply chooses the pixel in the receptive field with the maximum value and sends it to the output array. Average pooling calculates the average value within the receptive field as it moves across the image to send to the output array.

Fully connected layers connect output layers to nodes of the previous layer. The input image is not directly connected to the output.

The purpose of this CNN is to predict spatial average sea ice concentration per month using variable measures across all longitudes and latitudes as input. With the input of different variables and historic sea ice data, the model will produce predictions for monthly averaged sea ice concentrations in the Arctic.

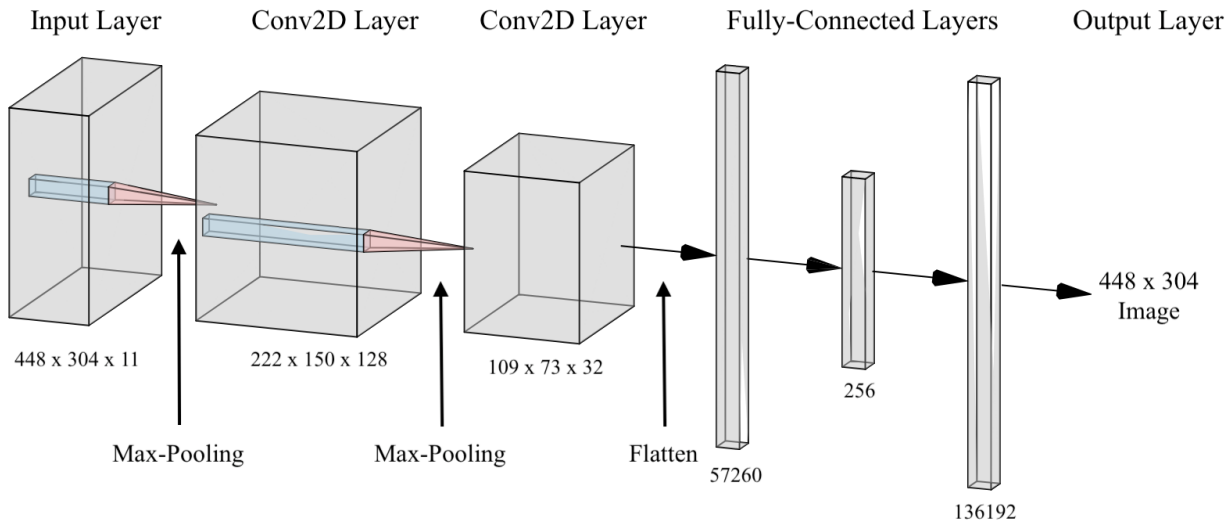

Figure 4.1: CNN model architecture.

4.3 Convolutional Long Short Term Memory Network

ConvLSTM architecture combines the spatial recognition capabilities of convolutional neural networks and the temporal modeling capabilities of long short-term memory models (LSTM) to produce an output which takes spatial and temporal patterns into account. LSTMs use matrix multiplication on each gate in an LSTM cell; ConvLSTMs replace this matrix multiplication with convolutions, allowing the model to capture underlying spatial features in multi-dimensional data [13].

By combining convolutions and LSTM gates, a ConvLSTM is able to capture both spatial and temporal patterns in our data. Values for each of the ConvLSTM gates are calculated using the following equations:

$$c_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \tag{4.2}$$

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \tag{4.3}$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f), \tag{4.4}$$

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o) \tag{4.5}$$

$$h_t = o_t \circ \tanh(c_t) \tag{4.6}$$

where $*$ represents a convolution and $\circ$ represents a Hadamard transformation.

The structure of ConvLSTM networks is nearly identical to the structure of LSTM networks; however, ConvLSTM networks utilize 3D tensors for gates, inputs, and outputs. ConvLSTMs include *memory cells*, represented as c_t in Equation 4.2, which are modified by certain gates. The *input gate*, i_t in Equation 4.3, allows the memory cell to accumulate information. The *forget gate*, F_t in Equation 4.4, allows the memory cell to disregard the past cell status, c_{t-1}. The *output gate*,

O_t in Equation 4.5, controls whether the latest memory cell value will be fed to the node's *final state*, h_t in Equation 4.6 [14].

Similar to the CNN, the ConvLSTM will predict monthly averaged SIC over all longitudes and latitudes. The main difference in the ConvLSTM is the addition of temporality to the spatial inputs which produce monthly outputs that represent the prediction for Arctic sea ice concentration.

4.4 Multi-Task Models

Multi-task learning is a subset of machine learning where multiple tasks are learned by a shared model [8]. Our network was trained to produce monthly image predictions of SIC for each pixel while also predicting a single sea ice extent value.

4.4.1 Multi-Task ConvLSTM and CNN

Both the CNN and ConvLSTM also feature multi-task learning. A branched architecture is implemented to allow the model to learn on multiple tasks. The models contain a shared root, where data is input, and two subsequent branches which produce the SIC image and sea ice extent outputs. The input root for the ConvLSTM consists of one ConvLSTM2D layer with 8 filters of size 5×5 and ReLU activation, followed by two alternating max pooling and convolution layers. The only difference for the CNN is that the first layer is a convolutional layer; everything else that follows replicates the architecture of the ConvLSTM. The max pooling layers contain filters of size 4×4, while the convolutional layers contain 128 and 32 size 5×5 filters respectively. The convolutional layers use ReLU activation. The data is then flattened and propagated through a dense layer with 256 nodes and ReLU activation.

The image branch of the architecture receives the model's root output and propagates it through a dense layer of size $448 \times 304 = 136192$ with linear activation. The data is then reshaped into an image of size 448 rows × 304 columns. Each pixel in the image contains an SIC measurement, scaled as a percentage out of 100.

The extent branch also receives the model's root output; it propagates the root output vector through 4 dense layers of size 128, 32, 8, and 1 respectively, returning a single sea ice extent result for each input sample. The first 3 dense layers include a ReLU activation function, while the output layer utilizes linear activation for regression

Two equally weighted loss functions are used to optimize the models. The image branch is optimized using the custom masked loss function described in section 4.1, while the sea ice extent branch is optimized using MSE loss. The performance of the models are evaluated using the RMSE metric for both branches.

4.5 Post-Processed Results

Due to results producing outlier values when running both the ConvLSTM and CNN with and without the multi-task implementation, post-processing was utilized to create more accurate results. The main aspects of this study's post processing were: removing values below zero and above 100, setting a mask for land and open ocean values, and dealing with missing values in the North Pole Hole.

Values below 0 were set to 0 and values above 100 were set to 100. This step was necessary as the model results contained considerable noise in areas surrounding sea ice, while also increasing the amount of ice in heavier concentrations. After post-processing the data to ignore the noise, results showed strong cohesion with actual values and produced much lower RMSE values.

By adding a mask to the predicted SIC values over land, noise in areas without ice were further eliminated. The mask works by multiplying values over land pixels by 0 and values over sea pixels by 1 in order for results to only focus on sea ice values rather than outliers that were skewing the accuracy.

The last aspect in post-processing to note was filling the North Pole Hole. The North Pole Hole is a region in the Arctic where satellite imagery does not produce observations. Because of this hole in the data, either filling it with interpolated values or leaving it as zero produced inaccurate results as the model learned data that is not observed in the actual data. Instead, in post processing, the north pole hole was recognized as NaN values so that the model simply ignored the area.

5 Results

In comparing results from this study's two models, RMSE and NRMSE were used to analyze which models had better accuracy. RMSE and NRMSE were calculated using the following equations:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{N}} \tag{5.1}$$

$$NRMSE = \frac{RMSE}{\bar{y}} \tag{5.2}$$

Another factor to note in comparing the models is that they are all trained on the years 1979-2012 and tested from 2013-2020. More information on this can be found in Section 3.2 which discusses data preprocessing.

5.1 CNN

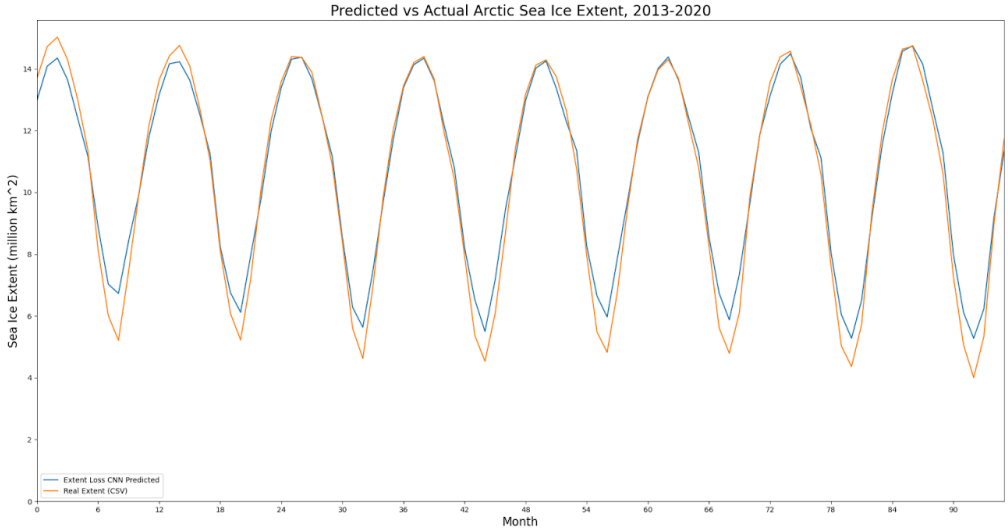

Figure 5.1: CNN derived predicted vs. Actual SIE values million km^2.

The base convolutional neural network attained an RMSE of 12.005% SIC for 2013 through 2020. After post-processing was applied to the predictions, the RMSE decreased to 7.106%. Implementing

SIE loss in conjunction with SIC loss for the CNN resulted in a small improvement in SIC prediction. While the test RMSE increased to 12.228%, the post-processed test RMSE fell to 6.993%, the lowest SIC RMSE among all models.

Sea ice extent values were calculated from the predicted SIC images of the base CNN and extent loss CNN. The base CNN resulted in an extent test RMSE of 0.868 million km^2, a significant improvement over the baseline VAR model. The extent loss CNN resulted in an extent test RMSE of 0.600 million km^2, a major improvement over the base CNN model. Overall, the extent loss CNN resulted in lower SIC and SIE errors compared to the base CNN, indicating the benefit of incorporating both SIC and SIE errors in the model loss function.

Figure 5.1 shows the predicted SIE values for the extent loss CNN compared to the real SIE values for 2013-2020. The model was able to predict the March SIE maxima with a high degree of accuracy. However, the model had a significant and consistent overestimate of the September SIE minima.

5.2 ConvLSTM

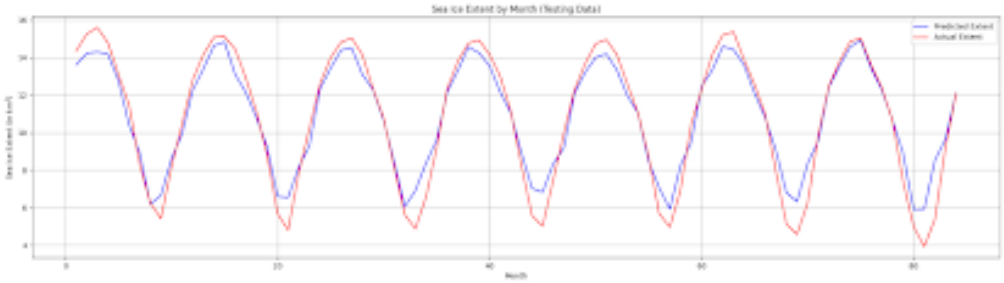

Figure 5.2: ConvLSTM derived predicted vs. Actual SIE values million km^2.

The convolutional LSTM produced an RMSE of 11.478% on image SIC testing data from 2014 to to 2020; its NRMSE was 1.116. After post-processing, the RMSE dropped to 8.162% per pixel; the associated NRMSE dropped to .8609. As seen in Figure 5.2, the convolutional LSTM's SIC prediction error values are lower than or comparable to all models from related works. Sea ice extent (SIE) for each month can also be derived from the ConvLSTM's predicted SIC maps. The ConvLSTM had a derived SIE RMSE of .9376 million km^2 and a corresponding NRMSE of .0866 for the testing data. Figure 5.2 shows the derived predicted vs. actual SIE values; the ConvLSTM captures fluctuations over time in SIE, but its predicted local minimum and maximum sea ice extent values are consistently inaccurate. This inaccuracy in the melting and freezing seasons can be seen in many other related models as well. Because of the large fluctuations in these maximum and minimum periods, it is more difficult for models to predict and is one of the main reasons for our study's implementation of multi-task learning.

5.3 Multi-Task CNN

The multi-task convolutional neural network resulted in improved SIE prediction and less accurate SIC prediction. After post-processing, the image SIC RMSE was 7.394%, the highest SIC error among all three CNN models. However, the SIE RMSE was 0.515 million km^2, the lowest SIE error among the CNN models. These results indicate that adding a separate extent branch helped the model target SIE learning, but at the potential detriment of SIC learning.

Figure 5.3 shows the predicted SIE values for the multi-task CNN compared to the real SIE values for 2013-2020. Compared to the extent loss CNN predictions shown in Figure TODO, the

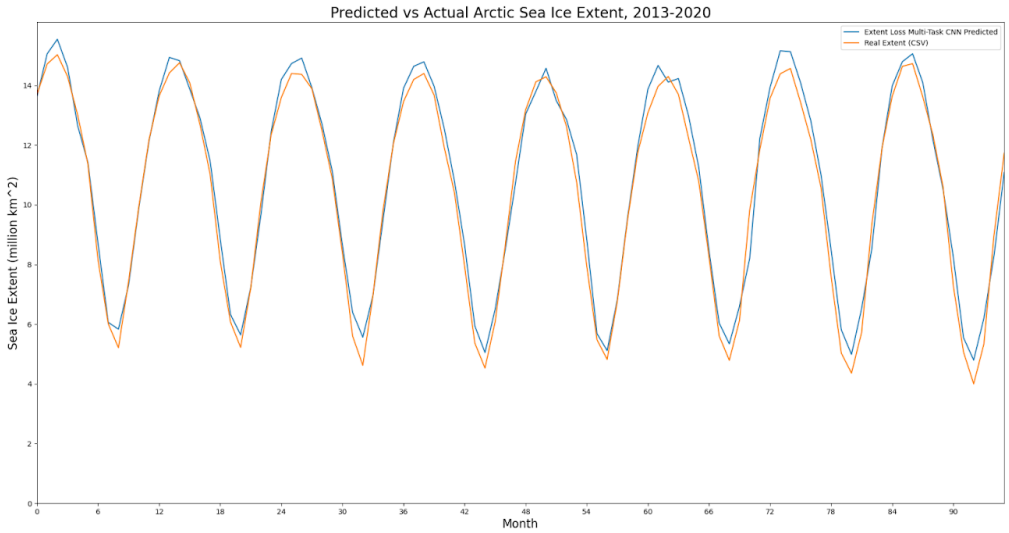

Figure 5.3: Multi-task CNN derived predicted vs. Actual SIE values in million km^2.

multi-task CNN featured improved prediction of the September SIE minima, but slightly worse performance of the March SIE maxima. This result is a key difference between the derived SIE predictions from the base and extent loss CNN and the SIE predictions from the multi-task CNN.

5.4 Multi-Task ConvLSTM

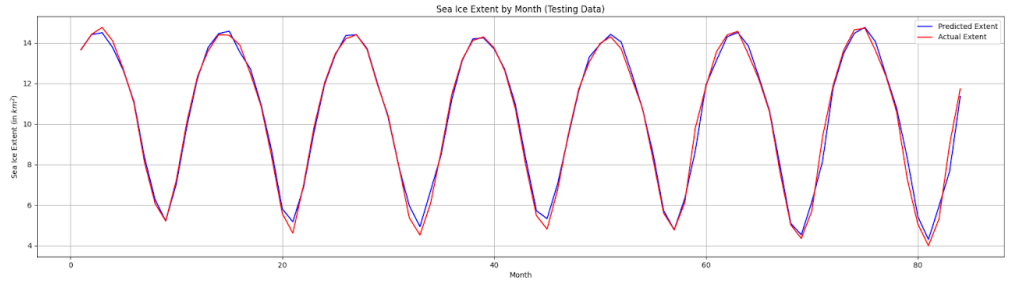

Figure 5.4: Multi-task ConvLSTM derived predicted vs. Actual SIE values million km^2.

For image SIC prediction, the Multi-Task ConvLSTM had an RMSE value of 10.785% per pixel and a subsequent NRMSE value of 1.0491 on the testing data. After post-processing, the image RMSE dropped to 7.192% per pixel with an associated NRMSE value of 0.759. For SIE prediction on the testing data, the Multi-Task ConvLSTM yielded an RMSE of 0.44071 million km^2 with a subsequent NRMSE of 0.062. As seen in Figure 5.4, the model accurately captures sea ice extent in all seasons, including the spring maxima and fall minima. Out of all networks we developed, the Multi-Task model had the lowest RMSE and NRMSE for SIC prediction before post-processing and second-lowest RMSE and NRMSE for SIE prediction.

Compared to related research, the model had lower RMSE values than almost all comparable SIC prediction models. Additionally, the model had a lower NRMSE value than all related models.

Furthermore, as seen in table 6.2, the Multi-Task ConvLSTM had the second lowest sea ice extent prediction RMSE and NRMSE. It is apparent that the Multi-Task ConvLSTM provides accuracy comparable with or better than state-of-the-art models in both sea ice concentration and sea ice extent prediction domains.

6 Discussion

Research Group	Model	Data	Physical Variables	Temporal Resolution	Lead Time	RMSE	NRMSE
Liu [25]	ConvLSTM	25x25 km	✓	Daily	1 day	11.2%	NA
Liu [25]	CNN	25x25 km	✓	Daily	1 day	13.7%	NA
EGU Kim [23]	CNN	25x25 km	✓	Monthly	1 month	5.76%	16.15*%
RS Chi [7]	LSTM	25x25 km	X	Monthly	1 month	8.89%	NA
Big Data REU	ConvLSTM	25x25 km	✓	Monthly	1 month	7.482%	0.789%
Big Data REU	CNN	25x25 km	✓	Monthly	1 month	5.64%	NA
Big Data REU	Multi-Task ConvLSTM	25x25 km	✓	Monthly	1 month	7.197%	0.7586%
Big Data REU	Multi-Task CNN	25x25 km	✓	Monthly	1 month	7.167%	NA

Table 6.1: This table compares related research on deep learning techniques to this study's own results on predicting Arctic sea ice concentration [25], [23], and [7]. The "data" column correlates to the type of image resolution each study uses. "physical Variables" indicated by a check or "x" denotes if the study used other atmospheric and oceanic variables in their model's input. "Temporal Resolution" refers to the way the data is temporally averaged. "Lead Time" is the time period that each model predicts. The row highlighted in blue has the lowest RMSE value, indicating that it produces the best accuracy for predicting sea ice concentration. Finally RMSE and NRMSE are the error values used to compare each model's accuracy (NRMSE = RMSE/$\bar{y}$). These differences in the models are important to note as it affects accuracy results and can be the reason for better performance from certain models.

Finally, we assessed our own models along with models from related works to compare sea ice concentration and extent RMSE and NRMSE values (Tables 6.1 and 6.2). When comparing the ConvLSTM and CNN to their multi-task variants, it can be seen in Tables 6.1 and 6.2 that the introduction of multi-task learning produces slightly more accurate results. This comparison is both visible when comparing plots of multi-task models with models that exclude multi-task learning, and in comparing the numerical RMSE results of the models. Not only does the use of multi-task learning increase SIC prediction accuracy, but it also provides accurate sea ice extent predictions. Because the multi-task models are optimized on two loss functions measuring SIC and SIE prediction error, the model learns to accurately predict both factors over time.

Along with the implementation of multi-task variants, this study also uses a custom loss function that aids in the accuracy of predictions. By eliminating excess values produced over land and open ocean pixels, the models were able to focus on the important prediction values of SIC and SIE. By

Research Group	Model	Data	Physical Variables	RMSE	NRMSE
Ali [1]	Attention-based Ensemble LSTM	Spactially averaged daily and monthly inputs	✓	0.586 million km^2	5.67% (0.0567)
Big Data REU	LSTM	Spactially averaged daily inputs	✓	3.14 million km^2	3.05% (0.0305)
Big Data REU	CNN	25x25 km monthly averaged	✓	0.862 million km^2	NA
Big Data REU	Multi-Task ConvLSTM	25x25 km monthly averaged	✓	0.44071 million km^2	4.24% (0.0424)
Big Data REU	Multi-Task CNN	25x25 km monthly averaged	✓	0.536 million km^2	NA

Table 6.2: This table is a comparison of related work [1] results and different model accuracy results from this study for Arctic sea ice extent. The row highlighted in blue has the lowest RMSE value, indicating that it produces the best accuracy for predicting sea ice extent. We also must note that, to our best knowledge, these are the only machine learning studies that calculate SIE.

adding this aspect to our models, accuracy increased, causing our RMSE values to decrease. These results are not only visible when producing images due to the large amount of excess values over land and open ocean, but they are also numerically visible in RMSE results.

By creating novel approaches to spatiotemporally based deep learning models to predict Arctic sea ice concentration and extent, our results were comparable to and even better than some previous deep learning sea ice concentration models (Table 6.1). As for sea ice extent, to our best knowledge there is not a lot of previous research on deep learning techniques to predict sea ice concentration. Because of this lack of related work, we only compared our results to one other piece of literature along with our own models (Table 6.2).

In analyzing a comparison of all models and their subsequent RMSE values, we found that our models did not perform the way we had expected. When expanding the RMSE vales to be SIC monthly averaged RMSE from our testing years of 2013-2020, in Figure 6.1 it can be seen that RMSE values tend to peak in October and February. These peaks are expected due to the difficulty in predicting melting and freezing season minimums and maximums. However, we unexpectedly found that the CNN models had better RMSE values in comparison to the ConvLSTM and multi-task learning models. We also see similar results in Figure 6.2 which represents yearly averaged RMSE from 2013-2020. Once again we can see that the CNN models outperform the other models when comparing their RMSE values. These results show that our models still require further work to improve results for the multi-task learning models.

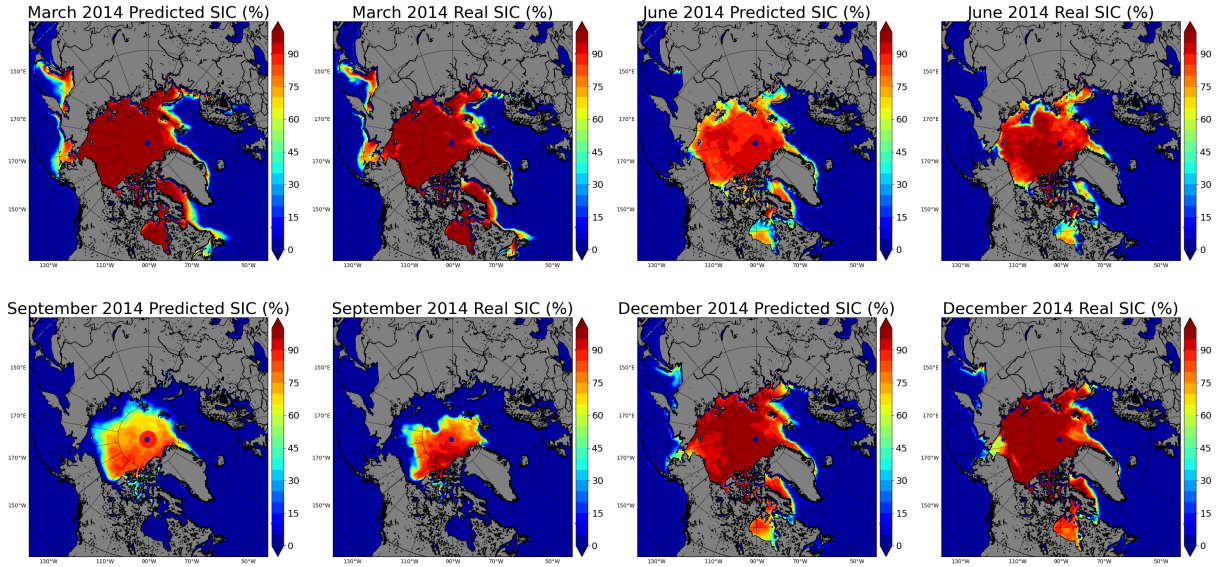

Figure 6.1: Multi-Task ConvLSTM SIC predictions and true values for March, June, September, and December 2014.

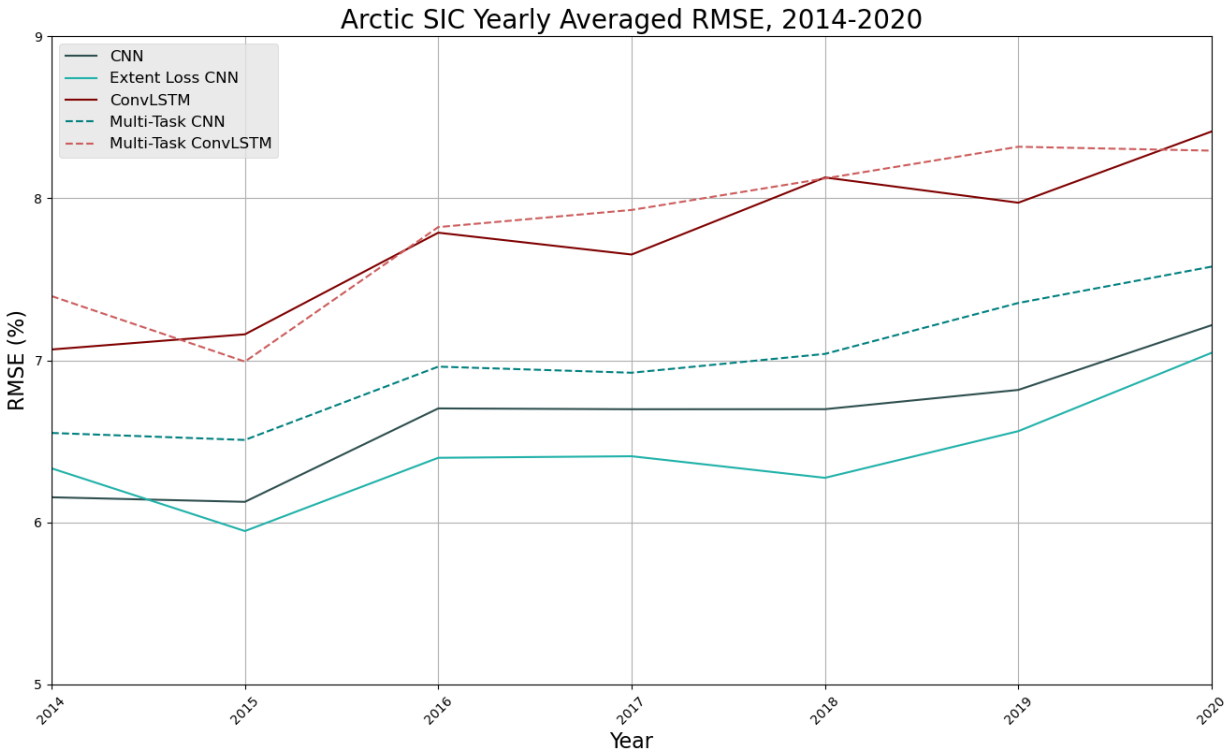

Figure 6.3: Yearly Averaged RMSE values for each model in million km^2

6.1 Future Work

Although our results were comparable to previous literature on sea ice prediction using deep learning, there is still improvement that could be made to our models to produce even more accurate results. One aspect that could be improved is the need for post processing. The use of a scaled

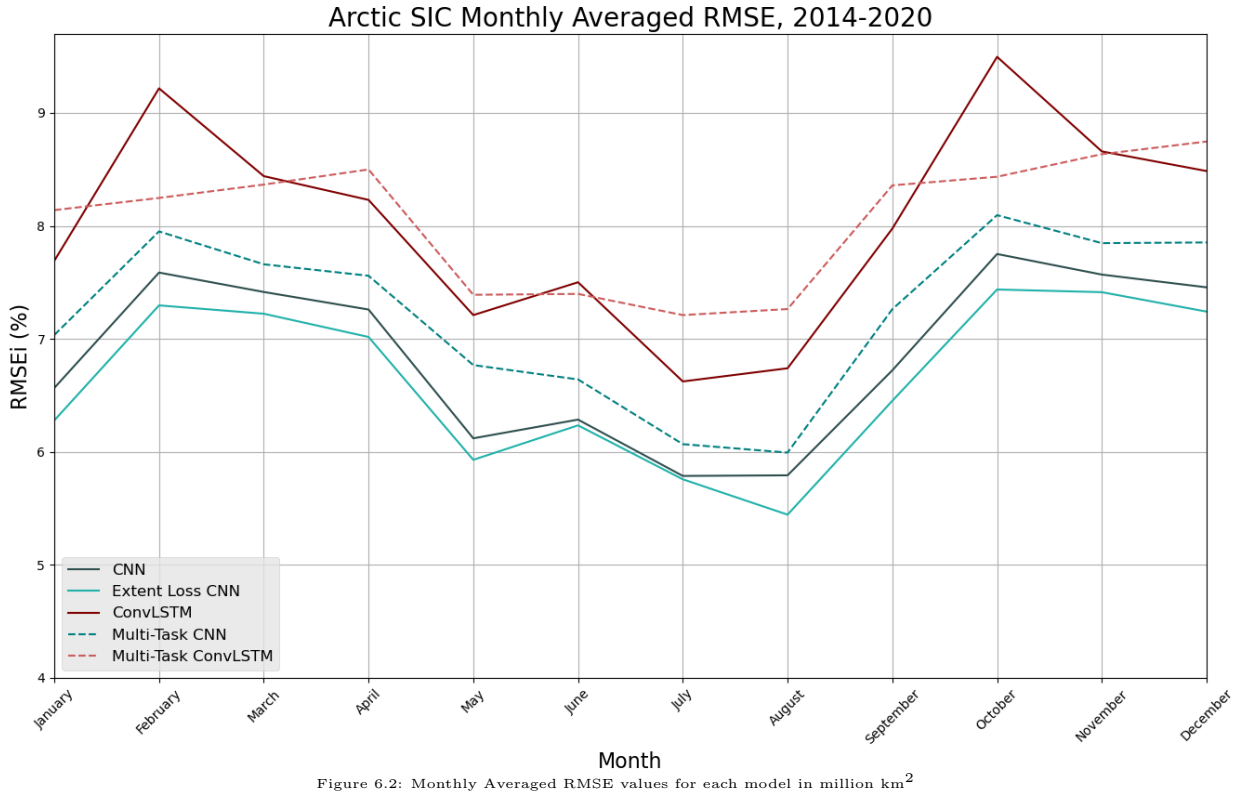

Figure 6.2: Monthly Averaged RMSE values for each model in million km^2

loss function could be a solution to this issue and eliminate the need for post processing. We also believe that, overall, our models could improve with more testing with varying lead times and further hyperparameter tuning on larger-scale studies. Our current models all run with a lead time of just 1-month, but we have seen slight improvement when varying the models' lead times and believe that with more testing, a better lead time could be used to achieve more accurate results. Along with varying the lead time, hyperparameter tuning alongside larger-scale testing could reduce issues with overfitting and aid in more accurate SIE and SIC predictions.

Acknowledgments

This work is supported by the grant "REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering" from the National Science Foundation (grant no. OAC–2050943). Co-author Ali and Wang additionally acknowledge support by the grant "CAREER: Big Data Climate Causality Analytics" from the National Science Foundation (grant no. OAC–1942714). The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS–0821258, CNS–1228778, and OAC–1726023) and the SCREMS program (grant no. DMS–0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources.

References

[1] Sahara Ali, Yiyi Huang, Xin Huang, and Jianwu Wang. Sea Ice Forecasting using Attention-based Ensemble LSTM. *arXiv e-prints*, page arXiv:2108.00853, July 2021.

[2] Ramdane Alkama, Ernest Koffi, S. Vavrus, Thomas Diehl, Jennifer Francis, J. Stroeve, Giovanni Forzieri, Timo Vihma, and Alessandro Cescatti. Wind amplifies the polar sea ice retreat. *Environmental Research Letters*, 15:1–16, 11 2020.

[3] T. Anderson, M. Peres-Ortiz, B. Paige, and J. Hosking et. al. Seasonal arctic sea ice forecasting with probabilistic deep learning. February, year =.

[4] Julien Boé, Alex Hall, and Xin Qu. September sea-ice cover in the arctic ocean projected to vanish by 2100. *Nature Geoscience*, 2(5):341–343, 2009.

[5] C3S. ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. Technical report, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2017.

[6] D.J. Cavalieri, C.L. Parkinson, P. Gloersen, and H.J. Zwally. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1. Technical report, NASA DAAC at the National Snow and Ice Data Center, 1996.

[7] J. Chi and H. Kim. Prediction of arctic sea ice concentration using a fully data driven deep neural network. *Remote Sensing*, December, year =.

[8] Michael Crawshaw. Multi-task learning with deep neural networks: A survey. *CoRR*, abs/2009.09796, 2020.

[9] Dick P Dee, S M Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, d P Bauer, et al. The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Q. J. Roy. Meteor. Soc.*, 137:553–597, 2011.

[10] Qinghua Ding, Axel Schweiger, Michelle L'Heureux, David S Battisti, Stephen Po-Chedley, Nathaniel C Johnson, Eduardo Blanchard-Wrigglesworth, Kirstin Harnos, Qin Zhang, Ryan Eastman, et al. Influence of high-latitude atmospheric circulation changes on summertime arctic sea ice. *Nature Climate Change*, 7(4):289–295, 2017.

[11] Tingfeng Dou, Cunde Xiao, Jiping Liu, Wei Han, Zhiheng Du, Andrew Mahoney, Joshua Jones, and Hajo Eicken. A key factor initiating surface ablation of arctic sea ice: Earlier and increasing liquid precipitation. *The Cryosphere*, 13:1233–1246, 04 2019.

[12] IBM Cloud Education. What are convolutional neural networks? https://www.ibm.com/cloud/learn/convolutional-neural-networks, October 2020.

[13] Jong-Min Yeom et al. Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional lstm networks for south korea. *Environmental Research Letters*, 15(9), 2020.

[14] Xingjian Shi et al. Convolutional lstm network: A machine learning approach for precipitation nowcasting. 2015.

[15] S. Horvath, J. Stroeve, B. Rajagopalan, and W. Kleiber. A bayesian logistic regression for probabilistic forecasts of the minimum september arctic sea ice cover. *Earth and Space Science*, 7:e2020EA001176, 2020.

[16] Sean Horvath, Julienne Stroeve, Balaji Rajagopalan, and Alexandra Jahn. Arctic sea ice melt onset favored by an atmospheric pressure pattern reminiscent of the north american-eurasian arctic pattern. *Climate Dynamics*, 04 2021.

[17] Yiyi Huang, Qinghua Ding, Xiquan Dong, Baike Xi, and Ian Baxter. Summertime low clouds mediate the impact of the large-scale circulation on arctic sea ice. *Communications Earth & Environment*, 2(1):1–10, 2021.

[18] Yiyi Huang, Xiquan Dong, Baike Xi, and Yi Deng. A survey of the atmospheric physical processes key to the onset of arctic sea ice melt in spring. *Climate Dynamics*, 52(7):4907–4922, 2019.

[19] Yiyi Huang, Xiquan Dong, Baike Xi, Erica K Dolinar, and Ryan E Stanfield. The footprints of 16 year trends of arctic springtime cloud and radiation properties on september sea ice retreat. *Journal of Geophysical Research: Atmospheres*, 122(4):2179–2193, 2017.

[20] M. Ionita, K. Grosfeld, P. Scholz, R. Treffeisen, and G. Lohmann. September arctic sea ice minimum prediction - a skillful new statistical approach. *European Geosciences Union*, March 2019.

[21] Liisi Jakobson, Timo Vihma, and Erko Jakobson. Relationships between sea ice concentration and wind speed over the arctic ocean during 1979–2015. *Journal of Climate*, 32, 08 2019.

[22] J. Kim, K. Kim, J. Cho, Y. Kang, H. Yoon, and Y. Lee. Satellite-based prediction of arctic sea ice concentration using a deep neural network with multi-model ensemble. *Remote Sensing*, 11, December 2018.

[23] Y. Kim, H. Kim, D. Han, S. Lee, and J. Im. Prediction of monthly arctic sea ice concentrations using satellite and reanalysis data based on convolutional neural networks. *European Geosciences Union*, March 2020.

[24] M. Li, R. Zhang, and K. Liu. Machine learning incorporated with casual analysis for short-term prediction of sea ice. *Frontiers in Marine Science*, May, year =.

[25] Q. Liu, R. Zhang, Y. Wang, H. Yan, and M. Hong. Daily prediction of the arctic sea ice concentration using reanalysis data based on a convolutional lstm network. *Marine Science and Engineering*, March, year =.

[26] Alexa Marcovecchio, Ali Behrangi, Xiquan Dong, Baike Xi, and Yiyi Huang. Precipitation influence on and response to early and late arctic sea ice melt onset during melt season. *International Journal of Climatology*, n/a(n/a).

[27] François Massonnet, Thierry Fichefet, Hugues Goosse, Cecilia M Bitz, Gwenaëlle Philippon-Berthier, Marika M Holland, and P-Y Barriat. Constraining projections of summer arctic sea ice. *The Cryosphere*, 6(6):1383–1394, 2012.

[28] NASA. Arctic sea ice minimum. https://climate.nasa.gov/vital-signs/arctic-sea-ice, July 2021.

[29] NSIDC. National snow and ice data center: Wildlife. https://nsidc.org/cryosphere/seaice/environment/mammals.html, April 2020.

[30] Dirk Olonscheck, Thorsten Mauritsen, and Dirk Notz. Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nature Geoscience*, 12, 05 2019.

[31] Z. Petrou and Y. Tian. Prediction of sea ice motion with convolutional long short-term memory networks. *IEEE Transactions on Geoscience and Remote Sensing*, April, year =.

[32] Aku Riihelä, Terhikki Manninen, and Vesa Laine. Observed changes in the albedo of the arctic sea-ice zone for the period 1982–2009. *Nature Climate Change*, 3(10):895–898, 2013.

[33] Piortr Skalsdi. Gentle dive into math behind convolutional neural networks. https://tinyurl.com/5fb635ac, April 2019.

[34] J. Stroeve, T. Markus, L. Boisvert, J. Miller, and A. Barrett. Geophysical research letters: Changes in arctic melt season and implications for sea ice loss. *Advanced Earth and Space Science*, 41, February 2014.

[35] Jia Wang, Jinlun Zhang, Eiji Watanabe, Moto Ikeda, Kohei Mizobata, John E Walsh, Xuezhi Bai, and Bingyi Wu. Is the dipole anomaly a major driver to record lows in arctic summer sea ice extent? *Geophysical Research Letters*, 36(5), 2009.

[36] L. Wang, X. Yuan, M. Ting, and C. Li. Predicting summer arctic sea ice concentration intraseasonal variability using a vector autoregressive model. *American Meteorological Society*, February, year =.

[37] D. Waugh, A. Sobel, and L. Polvani. What is the polar vortex and how does it influence weather? *American Meteorology Society*, pages 37–44, January 2017.