A machine learning model of Arctic sea ice motions

Jun Zhai and Cecilia M. Bitz

[1]Department of Atmospheric Science, University of Washington
[2]Department of Atmospheric Science, University of Washington

Key Points:

- The response of Arctic sea ice motions to surface winds can be modeled using a convolutional neural network with predictors of the previous-day sea ice velocity and concentration and present-day surface wind.
- The superior performance of convolutional neural network suggests the importance of the inter-pixel (across space) dynamical connections of the sea ice motion compared to pixel-based baseline models that only consider local interactions.
- The success of the convolutional neural network model of sea ice motion holds promise for combining machine learning with physics-based models to simulate sea ice.

Corresponding author: , jzhai@uw.edu

Abstract

Sea ice motions play an important role in the polar climate system by transporting pollutants, heat, water and salt as well as changing the ice cover. Numerous physics-based models have been constructed to represent the sea ice dynamical interaction with the atmosphere and ocean. In this study, we propose a new data-driven deep-learning approach which utilizes a convolutional neural network (CNN) to model how Arctic sea ice moves in response to surface winds given its initial ice velocities and concentration a day earlier. Results show that CNN computes the sea ice response with a correlation of 0.82 on average with respect to reality, which surpasses a set of pixel-based predictions, such as persistence (PS), linear regression (LR), random forest (RF), multiple layer perceptrons (MLP) and CICE, a leading physics-based model. The superior predictive skill of CNN suggests the important role played by the connective patterns of the predictors of the sea ice motion.

Plain Language Summary

Sea ice, the frozen seawater that floats on the ocean, grows in each hemisphere's winter and retreats in the summer but does not disappear in the current climate. The sea ice coverage is discontinuous and thin enough to move with the winds and currents. These movements alter the sea ice cover and transport pollutants, heat, water and salt. Previous studies have advanced mathematical models that represent the physics of how sea ice moves in response to surface winds. In this study, we propose a different type of model, called a convolutional neural network (CNN), which is constructed purely from observational data without explicitly accounting for the underlying physical insights. Unlike other conventional data-driven models that are trained on each geographic local point independently, CNN takes into account how a given location connects to its neighbors via patterns. The superior performance of CNN suggests that local ice motions depend on a large-scale pattern of surrounding winds and that data-based methods are a promising alternative to physics-based models.

1 Introduction

As an essential element in polar climate and dynamics, the movement of sea ice distributes pollutants, transports heat, water and salt and affects the polar energy budget by modifying the ice cover. The factors that determine sea ice motion include the sea ice inertia, atmospheric and oceanic stress, the Coriolis force, the sea surface tilt and the internal ice force that arises from floe interactions such as collisions, rafting and deformation. Based on an old rule-of-thumb that sea ice moves with a speed of 2% of the surface wind speed and 45° to the right of the surface wind direction, Thorndike and Colony (1982) examined the relation between the ice velocity and geostrophic winds and found that for the long-term (multi-month) average, about half of the variance of the ice motion is directly related to the surface geostrophic wind, while the other half is due to the mean ocean circulation; on shorter time scales from days to months, geostrophic winds alone can explain more than 70% of the variance of sea ice motion in the central Arctic region. Particularly, at daily time scale the unexplained variance is even less than 20%. On diurnal time scale or within 400 km of the coasts, however, the ice inertia or coastal stress gradients can be as important as the wind stress, which disqualifies the wind-only approximation.

Using daily sea ice velocities recovered from passive microwave satellite images, Kimura and Wakatsuchi (2000) showed that a linear relation based on the surface geostrophic winds with a speed reduction factor generally explains 70% to 90% of the sea ice velocity variance over the Arctic, except along some coastal regions. The spatial variation of the speed reduction factor of the sea ice motion relative to the surface winds depends on the internal ice stress gradient, which depends on ice thickness and concentration. Gen-

erally speaking, thicker and higher sea ice concentration result in a greater internal ice stress gradient, leading to a greater speed reduction. For example, sea ice moves at 2% of the surface wind speed over the seasonal ice zones but at less than 0.8% over the interior Arctic (Kimura & Wakatsuchi, 2000).

Similar conclusions were drawn for Antarctic sea ice from other studies. For example, in the Weddell Sea, nearly 70% to 95% of the variance of daily ice drift velocity can be linearly related to the wind velocity, except when the wind speed is below 3.5 m/s (Kottmeier et al., 1992). In addition, statistically significant relations between the sea ice velocity and local winds were detected from observations in most sectors over the Antarctic (Holland & Kwok, 2012).

Several studies have come up with equations-based models of varying complexity that utilize a wind-ice relation to model sea ice motion. A simple linear relational rule was described by Thorndike and Colony (1982) as follows:

$$\begin{bmatrix} U \\ V \end{bmatrix} = F \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} \bar{c_u} \\ \bar{c_v} \end{bmatrix}, \tag{1}$$

where (U,V) and (u,v) are the sea ice velocity and wind velocity, respectively, for the same day; F and θ are the speed reduction factor and turning angle, respectively; ($\bar{c_u}$, $\bar{c_v}$) is the daily mean surface ocean current velocity. Park and Stewart (2016) introduced a more complex nonlinear analytical model that calculates the ice velocity given that the surface winds are known and the ocean geostrophic currents are weak.

Physics-based models compute the sea ice velocity from the ice momentum equation (Hibler, 1979), which is primarily driven by surface stresses from winds and ocean currents. Specifically, the sea ice velocity tendency is modeled as

$$m\frac{\partial \mathbf{u}}{\partial t} = \nabla \cdot \vec{\sigma} + \vec{\tau_a} + \vec{\tau_o} - \hat{k} \times mf\mathbf{u} - mg\nabla H_o, \tag{2}$$

where $\vec{\sigma}$, $\vec{\tau_a}$ and $\vec{\tau_o}$ are the internal stress tensor, atmosphere stress, and ocean stress, respectively; m is the combined mass of ice and snow per unit area; H_o is the sea surface height; and $\mathbf{u}$ represents (u, v). The last two terms on the right hand side are the Coriolis force and sea surface tilt term. The parameterization for the atmosphere and ocean stresses contains the sea ice concentration as a multiplicative factor in order to accommodate the free drift scenario in low ice concentration regions.

Though having achieved a certain amount of success, physics-based models face challenges in practice, which include, but are not limited to, justification for assumptions about the sea ice rheology and the requirement for high-resolution observations to verify the large set of parametrizations. In this study, our goal is to model the sea ice response in an alternative approach that is data-driven and free of physical assumptions. Specifically, we aim to construct a convolutional neural network (CNN) model, in order to find how Arctic sea ice responds to the future surface winds in general given its current sea ice velocity and concentration. The trained CNN model has been shown to outperform the persistence and a set of pixel-based baseline predictions. The practice here aims to recover as much of the general relation between the surface winds and the Arctic sea ice motion as possible from the limited measurement.

The structure of the paper is as follows: Section 2 describes the concept and architecture of the CNN as well as the data used for training and testing; section 3 documents the path we took to select the input predictors for the CNN model; section 4 concludes with the training results and discussions.

2 Methodology and Data

CNN is a class of deep-learning neural network first introduced to classify the handwritten documents in the field of computer vision (LeCun et al., 1998) and later mostly

applied to image analysis. Inspired by the biological process whereby animals visualize images through the connectivity between neurons, CNN breaks-down an image composed of complex patterns into smaller and simpler pattern feature blocks by employing a mathematical operation called "convolution". The basic architecture of a CNN consists of the following elements:

- **an input layer** that is a tensor with dimensions of (number of image samples)×(image height)×(image width)×(number of input predictors/channels),
- **convolutional layers** that convolve the original images with kernels and activation functions into smaller-sized features maps,
- **pooling layers** that further reduce the dimensions of the feature maps via down sampling, and
- **a fully connected layer** that, via multiple layer perceptrons, fully connects the flattened (i.e., collapsing a multi-dimensional array into one dimension) output tensor from the last convolutional layer to the terminal output of the CNN.

Our objective is to predict present-day sea ice velocity (ui_1, vi_1) as accurately as possible with a parsimonious set of predictors as features input to the CNN. We investigate the importance of a list of candidate input predictors:

- present-day surface wind velocity (ua_1, va_1),
- previous-day sea ice velocity (ui_0, vi_0), and
- previous-day sea ice concentration (c_0) and thickness (s_0).

All of the predictors except sea ice thickness are readily available from a combination of satellite-based observations and atmospheric reanalysis from 1990 to 2018. Unfortunately, satellite-based observations of thickness prior to 2019 have high uncertainty, are discontinuous in time, and are completely unavailable in summer. Because we deemed observations of thickness to be insufficient for our purpose, we first test the importance of predictors with data sets taken exclusively from a climate model by constructing a series of CNN models with various combinations of the predictors listed above. Fortunately, the process of this feature exploration, as will be discussed in Section 3, reveals a combination lacking s_0 as one of the optimal feature input combinations. Therefore, we reconstruct the CNN model based on a data sets available from satellite-based observations and atmospheric reanalysis.

For climate model output, we use one historical simulation of the Community Earth System Model version 2 (CESM2)(Danabasoglu et al., 2020) in the Coupled Model Intercomparison Project 6 (CMIP6). The resolution is a nominal 1° in all components, which, for sea ice variables, is a 320x384 grid, with a global average resolution of 64 km approximately, and a finer resolution of 3 to 11 km over the Arctic. The 850-hPa winds are remapped to the sea ice grid to make consistent pixel-by-pixel alignments.

For reanalysis and satellite-based observations, we use the 10-m surface winds from the Japanese 55-year Reanalysis (Kobayashi et al., 2015), which has a resolution of 47 km on average in the Arctic, Polar Pathfinder Version 4 Daily Sea Ice Motion Vectors (Tschudi et al., 2019) and passive microwave sea ice concentrations from the NASA team (Cavalieri et al., 1996). The ice data are regridded from their 25-km EASE-grid to the JRA55 grid for consistent pixel-wise alignments.

We use the data sets that span 1990 to 2018 and split it into a training set (1990 to 2014), a validation set (2015 to 2016) and a testing set (2017 to 2018). The architecture of the CNN is shown in Fig.1: the input is composed of five channels, which are the present-day surface zonal and meridional velocities, the previous-day ice zonal and meridional velocities, and the previous-day ice concentration over the Arctic field of size 40×640. In this example, the input combination does not include thickness. The input is then pro-

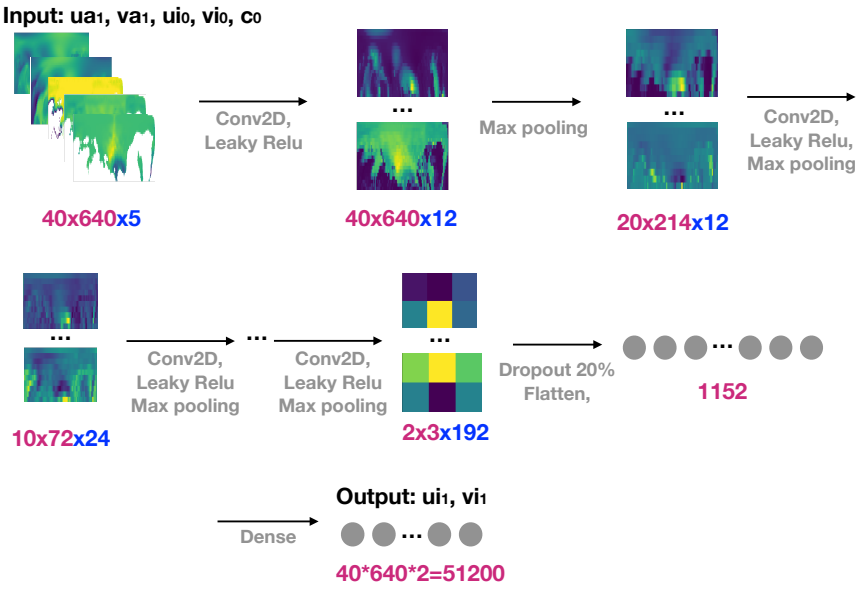

Figure 1. The CNN architecture. In each layer, the numbers in red are the dimensions of each channel, i.e. (image height)×(image width), and the numbers in blue specify the number of input predictors/channels.

cessed through five consecutive layers repeating a block unit: a 2D convolutional layer, a LeakyReLU (Leaky Rectified Linear Unit) layer and a 2D max-pooling layer. The output of this five-time repeated block unit is then passed to a 20% drop-out layer before getting flattened to a one-dimensional (1D) vector with a length of 1152. The final step is to regress the flattened 1D vector to another 1D layer of a size 51200, which is the concatenated 1D vector of two flattened 40×640 images that represent the present-day zonal and meridional ice velocities.

The kernel size for the 2D convolutions and max poolings are (2,3) or (2,2), where the latter is used for later layers since the image size is shrinking as the convolution progresses, and strides = (1,1) uniformly for all convolutions and max poolings. The convolutional layers all use a linear activation followed by a LeakyReLU added as a nonlinear activation with a negative slope coefficient, α=0.1. A root mean square error normalized by the standard deviation, i.e. the second term in Equation (4), is used as the loss function for optimization with the Adam optimizer. The total number of trainable parameters is 59,132,828.

Implemented in Python with the Tensorflow/Keras library (https://keras.io), the training takes 50 epochs with a batch size of 365 days on 9862 daily frames (from 1990 to 2016), the last two years of which (2015 to 2016) are used for validation. Then a fresh new set of daily frames (2017 and 2018) are used for final testing. To avoid obtaining accidental results due to fixed test sets, we have also performed a set of randomizations by shuffling the years for training, validation and testing. For example, one of the randomized set uses 1992 to 2016 for training, 2017 2018 for validation, and 1990 to 1991 for testing, etc. The results confirm that the prediction skill is not sensitive to the randomization and maintains stable metric scores. Therefore, here we present one set of training only by focusing on the results obtained using 1990 to 2016 for training and validation and 2017 to 2018 for final testing.

	M1	M2	M3	M4	M5	M6	M7
ua_1	✓	✓	✓	✓	✓		✓
va_1	✓	✓	✓	✓	✓		✓
ui_0	✓	✓				✓	✓
vi_0	✓	✓				✓	✓
c_0	✓		✓		✓		✓
s_0	✓		✓				
Corr	0.828	0.829	0.813	0.8001	0.813	0.787	0.828
PPP	0.436	0.438	0.413	0.388	0.413	0.381	0.437

Table 1. Seven CNN models, i.e. M1, M2, ..., M7 with different input combinations constructed for comparisons. The prediction skill evaluations of testing using Corr and PPP for the seven models are listed on the bottom rows.

To evaluate the CNN performance, we set-up four baseline models for comparison: persistence (PS), linear regression (LR), random forest (RF) and multiple layer perceptron (MLP). PS predicts the present-day sea ice velocity the same as that of the previous-day. LR, RF and MLP all regress the present-day sea ice velocity on the five input predictors used by CNN by using each time snapshot at each pixel as an independent sample for fitting. It is worth noticing that one advantage of CNN over these baseline models lies in the fact that instead of processing each pixel individually, CNN processes them in blocks that distill the neighboring connections into useful feature information.

As for skill metrics, the Pearson correlation (Corr),

$$corr_{x,y} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2}\sqrt{\sum_i^n (y_i - \bar{y})^2}}, \tag{3}$$

and Potential Prognostic Predictability (PPP),

$$PPP = 1 - \frac{\sqrt{\overline{(y_i - x_i)^2}}}{\sqrt{\overline{(x_i - \bar{x})^2}}}, \tag{4}$$

are used to quantify the model skill. For a given sample size n of a random variable x indexed with i, the skill of its prediction y can be quantified as (1) the covariance between the prediction and the truth scaled by their individual standard deviations (i.e., Corr) and (2) the the percentage of the true standard deviation explained by y (i.e., PPP). The "$^-$"s in Equation (3) and (4) denote the sample mean. Both Corr and PPP have a range from 0 to 1, with 1 being a perfect skill and 0 no skill. The difference between the metrics is that Corr quantifies how much x and y vary together regardless of their relative magnitudes to each other, while PPP takes the predicted magnitude into account for the accuracy.

3 Feature exploration

To explore the effectiveness of the candidate predictors in a nonlinear CNN model, we train seven CNN models with different input combinations using the historical simulations of CESM2 with the resulting prediction skill evaluations, as shown in Table 1. A comparison across the seven models draws the following conclusions:

- M3 and M5 are better than M4, indicating that in the absence of the sea ice velocity persistence (ui_0, vi_0) as predictor, including the sea ice concentration as predictor boosts the prediction accuracy.
- M3 and M5 are no different in terms of rounded metrics, indicating that there is no added benefit to the prediction from including thickness as predictor.
- Both M2 and M3 are better than M4, meaning that combining predictors of either ice conditions of concentration and thickness (c_0 and s_0) or velocity persistence (ui_0 and vi_0) with surface winds (ua_1 and va_1) enhances the prediction accuracy.
- M2 is slightly better than M3, indicating that the concentration and thickness might be correlated with the ice velocity field. This is further supported by the fact that adding ice conditions to M2 does not significantly improve the prediction as shown in the comparison between M1 and M2

For the remainder of our study, we select the best two combinations with the least number of predictor inputs, M2 and M7, for training using the observational and reanalysis data.

4 Results

By training M2 and M7 using JRA reanalysis data for surface winds and satellite-derived data for ice velocities and conditions, we find that M7 is slightly better (by 5%) than M2 in terms of the prediction metrics. Therefore we adopt M7 as our final CNN model and present its results here.

From the learning curve (Fig.2c,d), we see that the minimisation of the loss function is saturated at around 50 epochs with approximately 50% of the standard deviation unexplained, which corresponds to a correlation of 0.88 and a PPP of 0.5 approximately with respect to the truth. The training and validation curves closely link to each other to show improvement with the number of epochs, confirming that the model does not suffer from overfitting.

The prediction skill are evaluated in two fashions: local-wise and pattern-wise. Local-wise evaluation measures how well the model predicts the time series at a given pixel and presents a spatial distribution of the prediction metrics, as shown in Fig.2a,b. Pattern-wise evaluation, on the other hand, quantifies how well the prediction recovers the spatial pattern for each predicted field, quantified by computing Corr or PPP on each field flattened as a 1D vector and averaged over an interval such as a month (Fig3).

For local-wise evaluation, we evaluate the predicted sea ice velocity in terms of speed and angle. Figure 2a shows the spatial distribution of the averaged difference between the predicted and true speed. We see that most of the prediction errors of the CNN model manifest underestimating the speed by 5 cm/s or so (approximately 10% of the average speed). In general, most of the underestimations occur away from the central Arctic and are greatest (up to 25 cm/s) near the sea ice edge in the Greenland and Barents seas. The significant deviations in the angle prediction, on the other hand, are mostly along the Siberian coast of the Arctic with an angular deviation of up to 100°(i.e. $cos^{-1}(-0.2) \approx 100°$) from the truth (Fig.2b). On average, the angular error by CNN prediction over the interior of the central Arctic is approximately 32°(i.e. $cos^{-1}(0.85) \approx 32°$).

For pattern-wise evaluation, four baseline models (PS, LR, RF and MLP) are used for comparison. As shown in Fig.3, CNN outperforms all the baseline models that make pixel-based predictions, indicating the advantage gained from CNN's nonlinearity and the neighboring pixel connections. Specifically, the fact that CNN outperforms persistence (PS) confirms the influence of surface winds and ice conditions on the sea ice dynamic response.

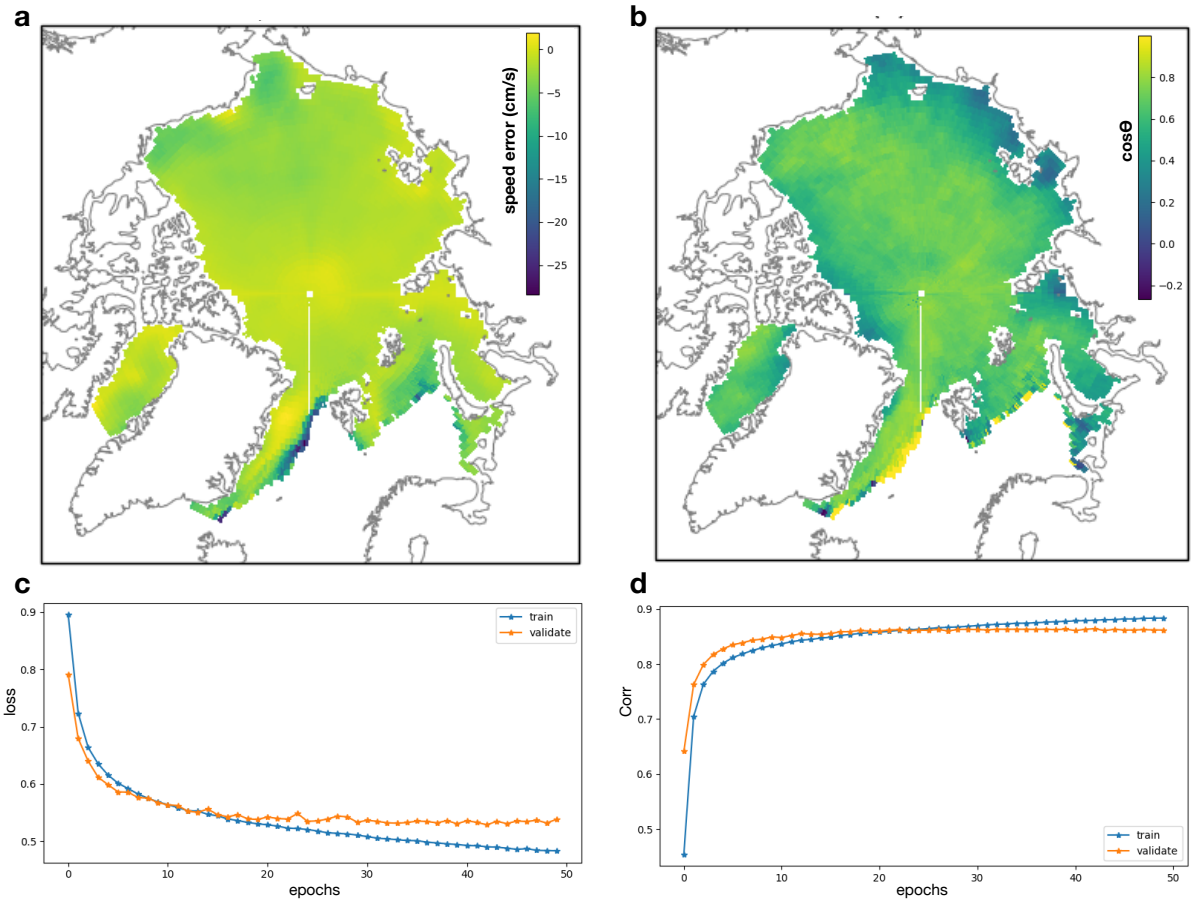

Figure 2. The top row shows the mean prediction errors in terms of **(a)** speed and **(b)** angle averaged over the test data set. Specifically, (a) shows the difference of the predicted speed and the true speed in cm/s with negative values associated with underestimating the speed; (b) shows the cosine of θ, the angle between the predicted velocity and the true velocity, with a $cos\theta$ closer to 1 associated with a smaller angular deviation from the truth and a negative $cos\theta$ associated with an angular deviation of greater than 90°. Both (a) and (b) are calculated for a given pixel at each time and temporally averaged over the two testing years. The bottom row shows **(c)** loss function and **(d)** Corr of prediction in the training and validation sets as indicated by the legends. The metrics are computed on each flattened 1-D sample frame.

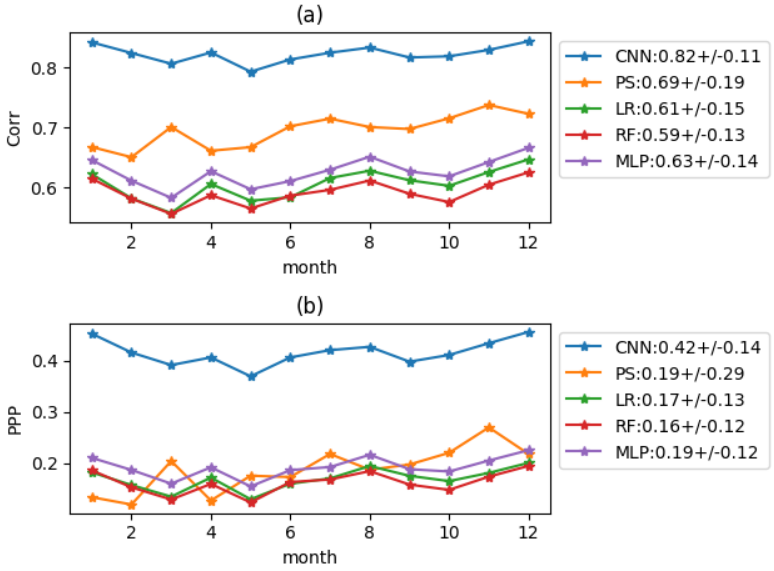

Figure 3. The averaged (a) Corr and (b) PPP for different models over each calendar month for the daily frames of the test set. The averages are computed as pattern-wise evaluation for each frame given a month. The mean and standard deviation (std) of Corr and PPP over the entire test set for each model are presented in the legends in the form of "mean+/-std".

In addition, we also compute the Corr and PPP for the daily prediction of sea ice motion from a "stand-alone" sea ice simulation that is forced by prescribing the JRA55 reanalysis. The output is from the CICE5, the 5th version of the Los Alamos Sea Ice Model (Hunke et al., 2015), which is also the sea ice component of CESM2. However, we find a Corr of 0.27 and PPP of -0.47, very low compared to the predictions of the CNN model. This is very likely resulted from the difference in sampling between the CICE5 model and verification data sets, which makes the comparison not really legitimate: the CICE5 daily output is a snapshot at an instant in time, while the Polar Pathfinder ice motion is a daily composite composed of mosaic satellite images on a given day (Tschudi et al., 2019). In summary, CNN predicts the sea ice motions in response to the surface winds with the highest stable predictive skills compared to all the other pixel-based baseline models.

Furthermore, we also developed a CNN model to predict the sea ice motion simulated by the CICE5 model on its hourly timestep. In this case, (ua_1, va_1) were the present-hour winds from JRA55 (used to force CICE5) and (ui_0, vi_0) and c_0 were the previous-hour output simulated by CICE5. Because the CICE5 model's motion fields lack measurement error and the timestep was so small, we expected this CNN model trained on hourly CICE5 motion would yield an improved Corr and PPP over the results for predicting daily satellite-based motion. However, it turns out that the finer temporal resolution only leads to less than 6% of improvement in the prediction skill. This result suggests that an upper bound might exist for a data-driven model to learn the relationships that predict sea ice motion.

5 Conclusions

In this study, we construct a deep-learning model, CNN, to predict how the Arctic sea ice responds to a given surface wind. Specifically, it takes the previous-day sea

ice velocity and concentration as well as the present-day surface wind as input features and predicts the present-day sea ice velocity. The result of prediction is evaluated against four pixel-fitted baseline models which are PS, LR, RF and MLP. It follows that CNN outperforms all four baseline models in terms of Corr and PPP as the accuracy metrics. The superior performance of CNN over the pixel-based models suggests the importance of connecting the neighboring local dynamics over the Arctic when modeling the sea ice rheology. The CNN model is also far superior at predicting the observed motion than the standard CICE5 model, which suggests there may be promise for combining a machine-learning method with a physics-based model.

Acknowledgments

All data used in this study are publicly available. For climate model output, we use one historical simulation (r1i1p1f1) of the Community Earth System Model version 2 (CESM2)(Danabasoglu et al., 2020) in the Coupled Model Intercomparison Project 6 (CMIP6). The 10-m surface winds from the Japanese 55-year Reanalysis are available through (Kobayashi et al., 2015). Polar pathfinder Sea ice velocity (Tschudi et al., 2019) and passive microwave sea ice concentrations (Cavalieri et al., 1996) are both from the National Snow and Ice Data Center (NSIDC).

References

Cavalieri, D. J., Parkinson, C., Gloersen, P., & Zwally, H. (1996). Sea ice concentrations from nimbus-7 smmr and dmsp ssm/i-ssmis passive microwave data, version 1. *Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center, doi*, *10*.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., ... others (2020). The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916.

Hibler, W. D. (1979, jul). A Dynamic Thermodynamic Sea Ice Model. *J. Phys. Oceangr.*, *9*(4), 815–846. Retrieved from `http://journals.ametsoc.org/doi/abs/10.1175/1520-0485{\%}281979{\%}29009{\%}3C0815{\%}3AADTSIM{\%}3E2.0.CO{\%}3B2` doi: 10.1175/1520-0485(1979)009⟨0815:ADTSIM⟩2.0.CO;2

Holland, P. R., & Kwok, R. (2012). Wind-driven trends in antarctic sea-ice drift. *Nature Geoscience*, *5*(12), 872.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., & Elliott, S. (2015). *CICE : the Los Alamos Sea Ice Model Documentation and Software User's Manual Version 5.1 LA-CC-06-012* (Tech. Rep.). Los Alamos National Laboratory. Retrieved from `http://oceans11.lanl.gov/trac/CICE`

Kimura, N., & Wakatsuchi, M. (2000). Relationship between sea-ice motion and geostrophic wind in the northern hemisphere. *Geophysical Research Letters*, *27*(22), 3735–3738.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... others (2015). The jra-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48.

Kottmeier, C., Olf, J., Frieden, W., & Roth, R. (1992). Wind forcing and ice motion in the weddell sea region. *Journal of Geophysical Research: Atmospheres*, *97*(D18), 20373–20383.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Park, H.-S., & Stewart, A. (2016). An analytical model for wind-driven arctic summer sea ice drift. *The Cryosphere*, *10*(1), 227–244.

Thorndike, A., & Colony, R. (1982). Sea ice motion in response to geostrophic

winds. *Journal of Geophysical Research: Oceans*, *87*(C8), 5845–5852.

Tschudi, M., Meier, W., Stewart, J., Fowler, C., & Maslanik, J. (2019). *Polar pathfinder daily 25 km ease-grid sea ice motion vectors, version 4 dataset 0116. nasa national snow and ice data center distributed active archive center.*