Surrogate Modelling for Sea Ice Concentration using Lightweight Neural Ensemble

Anonymous submission

Abstract

The modeling and forecasting of sea ice conditions in the Arctic region are important tasks for ship routing, offshore oil production, and environmental monitoring. We propose the adaptive surrogate modeling approach named *LANE-SI* (**L**ightweight **A**utomated **N**eural **E**nsembling for **S**ea **I**ce) that uses ensemble of relatively simple deep learning models with different loss functions for forecasting of spatial distribution for sea ice concentration in the specified water area.

Experimental studies confirm the quality of a long-term forecast based on a deep learning model fitted to the specific water area is comparable to resource-intensive physical modeling, and for some periods of the year, it is superior. We achieved a 20% improvement against the state-of-the-art physics-based forecast system SEAS5 for the Kara Sea.

Introduction

In recent years, Arctic modeling has become a focus of increased interest for researchers in various scientific fields. The sea ice melting process, in conjunction with global warming, determines the ecological state of the region, including biodiversity conservation. Ice concentration distribution plays a key role in laying tracks by icebreakers and advancing cargo and research ships.

Predictive modeling of ice conditions in the Arctic is considered in different time scales. Short-term forecasts involve the assimilation of operational information and reproducing accurate data for decision-making. This type of forecasting is limited to intervals from several hours to several months. Long-term forecasting is carried out one or several seasons in advance to enable advance planning, which is essential in the industrial sector (Harsem et al. 2015). In predicting the timing of opening and freezing of the water area, statistical modeling is carried out based on retrospective observational data (Wu et al. 2023).

Physical modeling became the classical approach to the simulation of natural environments because of advantages on the spatial and temporal scales and the ability to predict and analyze different scenarios (Chen et al. 2023). It utilizes complex numerical methods to solve the equations of motion, thermodynamics, and radiation (Stroeve et al. 2007). Despite the robustness and explainability of physical modeling, its application is often complicated by high computational cost, the need for boundary and initial conditions for many variables, and complex parameterization for a specific territory. Years of active development of earth remote sensing have contributed to the accumulation of spatiotemporal datasets for environmental systems, which can be used to train various data-driven models, including deep neural networks. While deep learning models can provide high-quality environmental forecasts (Lam et al. 2023), the computational cost for models' training to satisfy the task-specific requirements (water area, time and spatial resolution, forecast length, objective function) can be excessive. For this reason, we aimed to provide a more lightweight surrogate alternative to existing solutions for sea ice forecasting.

This paper presents a deep learning approach *LANE-SI* that uses an ensemble of convolutional neural networks (CNN) for long-term predictive modeling of sea ice concentration. Satellite data with post-correction - OSI SAF Global Sea Ice Concentration (SSMIS) product (Tonboe et al. 2016) was taken as training data. Results are presented for the Kara Sea and part of the Barents Sea region of the Arctic. The developed model is lightweight due to its simple architecture and is also undemanding in input data - only the pre-history of the target parameter is required. A comparison was made with the ECMWF's fifth-generation physics-based seasonal forecast system - SEAS5 (Johnson et al. 2019). The comparison results demonstrate that the surrogate model reproduces the absolute values of ice concentration and their spatial distribution with a quality comparable to the SEAS5 forecast. The surrogate model reproduces the position of the ice edge better than the SEAS5 forecast, both on average contour points position and the position of each point separately.

The source code and sample datasets are available on GitHub to reproduce experimental results. The link is provided in ANONYMIZED. The paper is structured as follows: Section 2 reviews the key related work, Section 3 describes the proposed method for the design of surrogate sea ice forecasting models named *LANE-SI*, Section 4 contains the experimental results, and Section 5 presents concluding remarks for the proposed methods.

Related Work

The classic method for modeling sea ice parameters is numerical modeling based on systems of differential equations. Over the years of technological development, a large number of both specific solutions that reproduce local ice

dynamics (Girard et al. 2011) and general methods of numerical hydrometeorological forecasting (Wang, Chen, and Kumar 2013; Zhang and Rothrock 2003) have been developed. State-of-the-art physical models like NEMO-SEAICE (Madec et al. 2017) make it possible to reproduce ice conditions both at the global (Hvatov et al. 2019) and regional (Pemberton et al. 2017) levels. Considering the universality of numerical models, it should be noted that their adaptation to a specific territory requires significant effort from a subject specialist and considerable computing resources, as well as setting boundary and initial conditions, which raises the need for high-quality initial data, which is not always available for the water area interest.

There has been a rapid development of data-driven approaches in recent years. Model complexity varies from linear relationship searches (Kapsch et al. 2014) and regression models (Ahn et al. 2014) to deep learning networks (Chi and Kim 2017). Regression models are limited in their applicability due to the enormously increasing complexity when moving to grid calculations. Using neural networks based on simple convolutional (Kim et al. 2020) and more complex U-Net (Ali and Wang 2022) architectures provide high-resolution output images and allow a more detailed forecast. However, most high-quality existing solutions require a large amount of additional input data and computing power (Andersson et al. 2021; Ali and Wang 2022) for training and network prediction, significantly limiting their applicability for specific water areas.

It is known that combining a set of simple models into ensembles makes it possible to achieve higher quality forecasting, including in the subject area of hydrometeorology (Ali et al. 2021; Kim et al. 2018). We assume that ensembling of simple neural networks can lead us to better quality at lower computational costs (Zhou, Wu, and Tang 2002). Thus, the question arises of finding a balance between the complexity of the model and the quality of the forecast.

LANE-SI

To achieve a better quality of high-resolution long-term sea ice forecasting in a computationally cheap way, we developed the *LANE-SI* approach that allows designing the predictive model for sea ice in specific water areas. We used the convolutional neural network for a non-standard task - time-spatial forecasting. In contrast to LSTM architectures, which allow the extraction of features only from time series (Choi, De Silva, and Yamaguchi 2019), convolution emphasizes the spatial distribution of the parameter and reduces the risk of stagnation over time. The surrogate model was implemented to carry out long-term forecasting with a forecast horizon of one year. We aimed to make the model as flexible as possible to adapt to specific water areas to reduce the computational cost against neural surrogate models that cover the entire Arctic region (Andersson et al. 2021).

In LANE-SI, the dataset for training a surrogate data-driven predictive model for ice concentration is prepared similarly to the lagged transformation applied to time series. Previous k steps are used to predict n steps ahead using the pre-history of sea ice. Each time step is characterized by an image describing the concentration in each cell.

The surrogate model is represented by an ensemble that aggregates two CNNs trained with different loss functions and an inertial forecast - concentration values averaged for each day of the year over the previous five years.

CNN for sea ice forecasting

The basic structure of the deep neural networks used in *LANE-SI* include five convolution layers with ReLU activation functions (Agarap 2018). Postprocessing with cutting to the range was used due to the inconsistency of using the sigmoid activation function to ensure that the output image values range from 0 to 1 (Wang, Scott, and Clausi 2017). The structural scheme of CNN is presented in Figure 1. The specific size of layers is configured in an automated way.
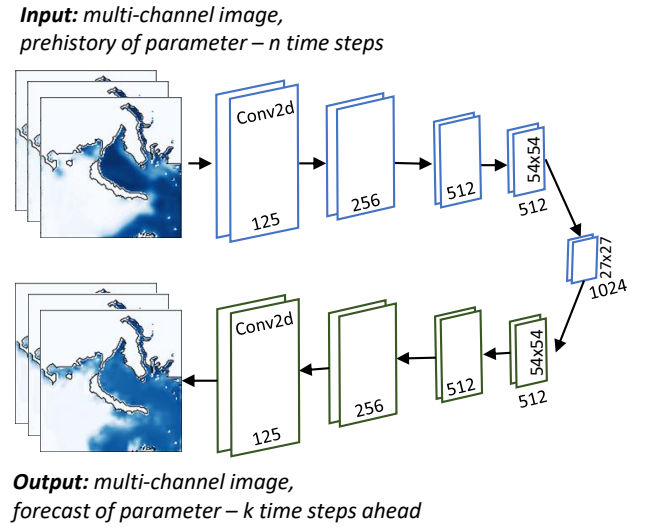

Figure 1: Architecture of deep model for sea ice concentration forecasting. The specific values of layer shape are described for the experimental domain (125x125) used in Section .

In the modeling of spatiotemporal data, it is important not only to reproduce the absolute values but also to measure the proximity of the spatial distribution of values to the distribution in the target image. Therefore, CNNs were trained on two loss functions:

- **Mean Absolute Error (MAE, L1Loss)** - represents the closeness of the absolute values of each pixel of each prediction element with the target image;
- **Structural Similarity Index (SSIM)** - represents the similarity local patterns of pixel intensities (Wang et al. 2004) and shows similarity in the spatial distribution of the parameter at predicted and target images.

The CNN accepts multi-channel images with several years of ice concentration pre-history.At the output, the CNN provides images of the same spatial resolution as the input ones, but the number of channels is 52. It corresponds to a one-year forecast with a time sampling of seven days.

Neural ensembling of forecasts

The multi-model ensemble was used to improve the quality of the surrogate model in *LANE-SI*. The first CNN reflects the spatial distribution of the parameter (SSIM), and the second reproduces the parameter's absolute values (MAE). Also, a naive forecast in the form of repeating the average values for five years for each day of the year was added to reproduce long-term dynamics not represented in pre-history.

Three types of ensembling methods were tested by quality comparison with the same metrics (MAE, SSIM):

- Simple weighted ensemble, based on linear regression: each pixel in a predicted multi-channel image is considered an independent variable. Predicted values for each pixel of three separated models are summed with coefficients that provide less error on the delayed part of train data;
- CNN trained with MAE loss function uniting the elements of the ensemble in the form of multi-channel images into one prediction;
- CNN trained with SSIM loss function unites the ensemble's elements in the form of multi-channel images into one prediction.

The general scheme for an approach that combines single models into using neural ensembling is presented in Figure 2.

Experiments and Results

Dataset

We use OSI SAF Global Sea Ice Concentration (SSMIS) product as training data for sea ice concentration forecasting in specific water areas. The time resolution of the product is one day; however, for the problem of long-term forecasting, such discretization is redundant, and therefore the time resolution was sparsed to weekly. Data were normalized to a regular grid with a spatial resolution of 14 km using bilinear interpolation. An area with the Kara Sea and part of the Barents Sea was chosen as a test domain (as presented in Figure 3).

We conduct the experimental evaluation of *LANE-SI* for the Kara Sea using SSMIS for 1996 to 2022 years. Primary forecasting models were trained on the data time range from 1996 to 2009. The ensembling model was trained on predictions of single models from 2010 to 2015. They were additionally re-trained on the entire training set from 1996 to 2015 to obtain better forecasts using single models. Validation was made as an out-of-sample one-year-ahead forecast for 2006-2022 compared with naive forecast and actual data.

The training time for *LANI-SI* in the described setting on a local machine with an NVIDIA GeForce RTX 4080 video card and an AMD Ryzen 9 5950X processor (16 x 3.4 GHz) was approximately 450 minutes.

LANE-SI against separate data-driven models

The quality of the surrogate model (design with *LANE-SI*) for the Kara Sea was analyzed compared to the naive forecast for both the final ensemble model and individual ensemble components. Table 1 presents numerical estimates of metrics averaged over the years.

As the table shows, the surrogate ensemble model achieves a higher forecast quality than single models according to the SSIM metric for all years and almost all years according to the MAE metric.

Figure 4 presents an example of a spatial comparison of different data-driven forecasts (*LANE-SI*-designed ensemble and separate neural models) with actual data for a single forecast step.

This example shows the shortcomings of single models. The CNN with the L1 loss function reproduces the ice field with large granularity despite its good fit into the absolute values on average over the image. CNN with SSIM loss function and naive forecast tends to over-smooth the image and underestimate the absolute values of the parameter. Thus, the ensemble model combines single models while avoiding their inherent disadvantages.

LANE-SI against physics-based forecast

To objectively assess the quality of the surrogate model, it was compared with the state-of-the-art (SOTA) model based on differentiation equations - the predictive model SEAS5.

The maximum forecast horizon of the SEAS5 model is nine months (against 12 months for the design surrogate model). Therefore, we configured both forecasts to start on January 1 of each year. The error was calculated on average for each season (three months) out of three fully predicted.

Table 2 presents the comparison results for the entire SEAS5's available archives (2020-2022).

As the table shows, forecasting using a surrogate model designed by *LANE-SI* is comparable in quality to the physics-based forecast and, in some cases, can even surpass it due to the better adaptation to specific water areas.

In addition to the metrics used for training (SSIM, MAE), a metric for comparing the similarity of the ice edge was used to assess the quality. The ice edge is the contour of the spatial position of ice with a concentration of more than 0.8. An expert can choose the threshold for switching to binary data (it depends on the scenario of the forecast usage). In this case, the ice edge's position is represented by a set of points with coordinates of a fixed length. The distance metric is calculated as a mean of the signed distance between each point of one contour and the nearest contour edge of the second contour (Brahmbhatt 2013). This metric was introduced because it is more indicative for assessing the spatial distribution of ice concentration values critical for navigation.

Figure 5 shows an example of the result of edge identification as 100 points with the distance metric calculation.

The distance values were displayed as boxplots for each prediction step to assess the range of agreement between the ice edge contour points. Figure 6 provides a visualization for the 2021 ice melt period.

The *LANE-SI*-designed surrogate model provides a prediction close to SEAS5 by the average value of edge distance criteria. Also, the edge distance criteria variance for each point is lower for the proposed surrogate model.
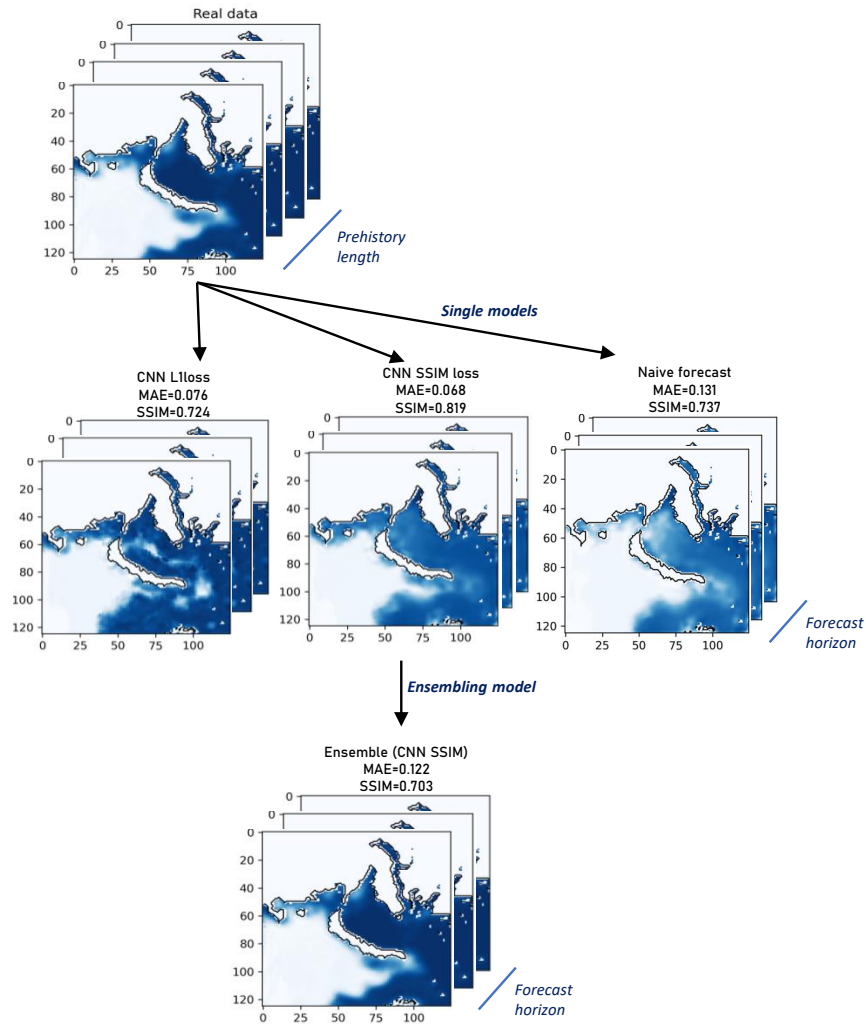
Figure 2: The structure of the proposed ensemble model (first layer - predictive model, second layer - neural ensembling model).
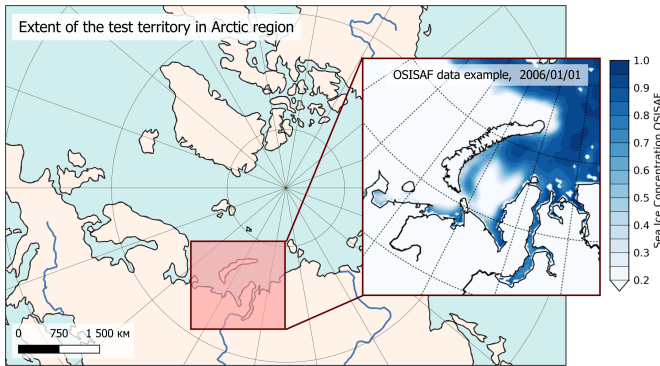

Figure 3: Spatial extent of testing domain with OSI SAF ice concentration data sample on 2006/01/01.

Conclusion

This paper proposes an approach to predictive modeling of sea ice concentration named *LANE-SI*. It is based on an en-

semble of convolutional neural networks with different loss functions and naive forecast (that represent the average over the last several years).

Quality assessment was conducted using several objectives: (1) the coincidence of the absolute values of the parameter with actual data (MAE), (2) the coincidence of the spatial distribution of the parameter (SSIM), and (3) the distance between ice edges The results of experiments for Kara sea confirms that, according to the SSIM metric, an ensemble always gives better quality than single models; according to the MAE metric, in most cases, an ensemble always gives better quality. The comparison results for the average distance between actual and predicted ice edges indicate that, on average, over the contour and taking into account the scatter of the metric for each contour point, the surrogate model based on a convolutional network performs better in predicting the ice edge. Also, the developed model was compared with a physics-based forecast SEAS5. Experiments have shown that the *LANE-SI* approach allows forecasting with a quality close to the physics-based SOTA

Metric	Mean Absolute Error (MAE)				Structural Similarity Index (SSIM)			
Year	Surrogate	CNN L1Loss	CNN SSIM loss	Naive forecast	Surrogate	CNN L1Loss	CNN SSIM loss	Naive forecast
2016	0,064	**0,044**	0,048	0,086	**0,807**	0,726	0,796	0,723
2017	**0,056**	0,059	**0,056**	0,064	**0,810**	0,701	0,784	0,758
2018	**0,052**	0,061	**0,052**	0,06	**0,808**	0,690	0,776	0,768
2019	**0,045**	0,061	0,056	0,055	**0,823**	0,689	0,782	0,780
2020	0,085	**0,064**	0,079	0,104	0,509	**0,595**	0,522	0,469
2021	**0,07**	0,089	0,079	0,087	**0,526**	0,525	0,500	0,496
2022	**0,064**	0,074	0,082	0,085	0,532	**0,552**	0,492	0,494

Table 1: Comparison of surrogate ensemble model and single CNN models by MAE and SSIM metric mean for each year
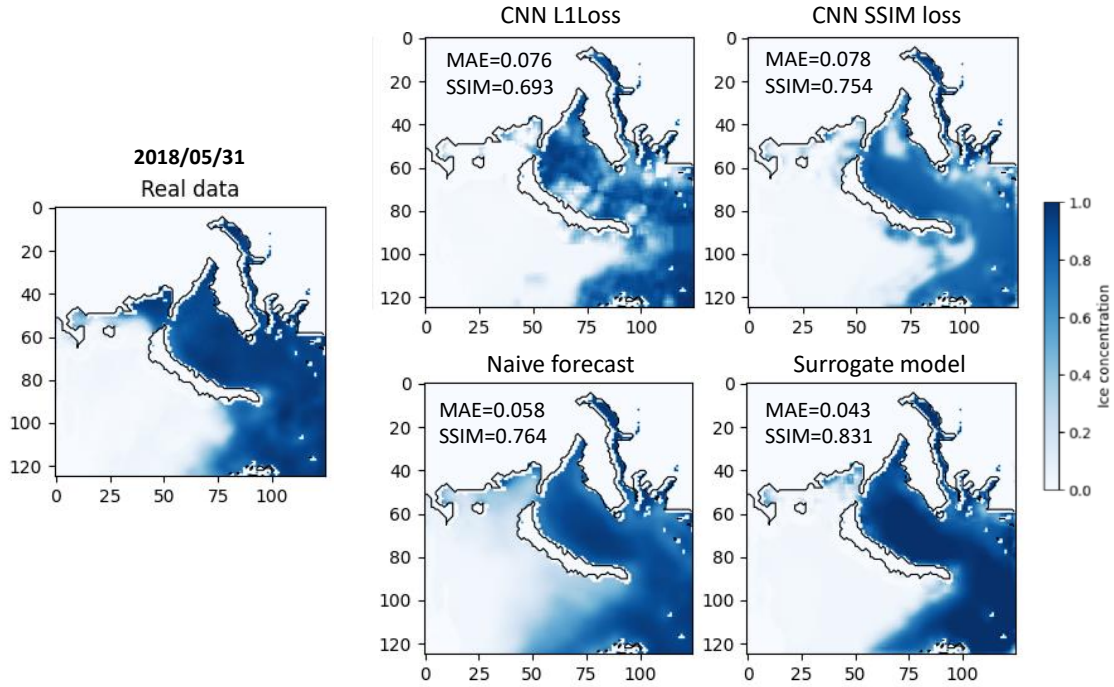

Figure 4: Comparison of the spatial distribution of forecasted ice concentration by ensemble surrogate model and separate models with real data for 31.05.2018

model. While we used the Kara Sea for validation, *LANE-SI* can be used to design the forecasting model with a specified resolution and forecast horizon for any water area in the Arctic.

The future research development will be aimed at applying automated machine learning to automate the design of surrogate ensemble models for new conditions. Also, we will consider the applicability of *LANE-SI* to other environmental processes (e.g., weather forecasting).

References

Agarap, A. F. 2018. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.

Ahn, J.; Hong, S.; Cho, J.; Lee, Y.-W.; and Lee, H. 2014. Statistical modeling of sea ice concentration using satellite imagery and climate reanalysis data in the Barents and Kara Seas, 1979–2012. *Remote Sensing*, 6(6): 5520–5540.

Ali, S.; Huang, Y.; Huang, X.; and Wang, J. 2021. Sea Ice Forecasting using Attention-based Ensemble LSTM. *arXiv preprint arXiv:2108.00853*.

Ali, S.; and Wang, J. 2022. MT-IceNet-A Spatial and Multi-Temporal Deep Learning Model for Arctic Sea Ice Forecasting. In *2022 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)*, 1–10. IEEE.

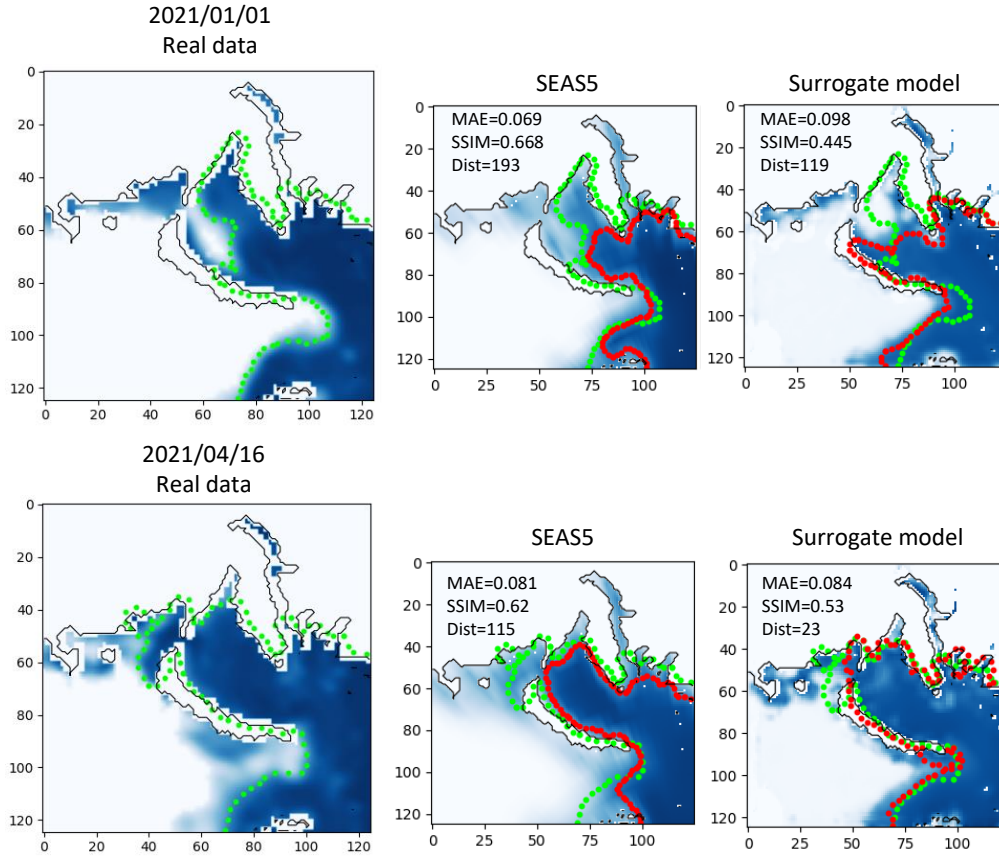Andersson, T. R.; Hosking, J. S.; Pérez-Ortiz, M.; Paige, B.;

Figure 5: Visualization of the identified ice edge on real data (green points) in comparison with forecasts of SEAS5 and a surrogate model (red points)

Metric	Mean Absolute Error (MAE)		Structural Similarity Index (SSIM)	
Year and quater	SEAS5	Surrogate model	SEAS5	Surrogate model
2020Q1	**0,097**	0,103	**0,602**	0,516
2020Q2	0,113	**0,108**	**0,542**	0,474
2020Q3	0,082	**0,021**	**0,583**	0,571
2021Q1	**0,099**	0,103	**0,630**	0,521
2021Q2	**0,078**	0,094	**0,610**	0,490
2021Q3	0,063	**0,029**	**0,626**	0,560
2022Q1	0,103	**0,092**	**0,587**	0,540
2022Q2	0,095	**0,081**	**0,546**	0,508
2022Q3	0,074	**0,025**	**0,576**	0,563

Table 2: Comparison of surrogate ensemble model and SEAS5 forecast for first three quarters of each year

Elliott, A.; Russell, C.; Law, S.; Jones, D. C.; Wilkinson, J.; Phillips, T.; et al. 2021. Seasonal Arctic sea ice forecasting with probabilistic deep learning. *Nature communications*, 12(1): 5124.

Brahmbhatt, S. 2013. *Practical OpenCV*. Apress.

Chen, S.; Li, K.; Fu, H.; Wu, Y. C.; and Huang, Y. 2023. Sea Ice Extent Prediction with Machine Learning Methods and Subregional Analysis in the Arctic. *Atmosphere*, 14(6): 1023.

Chi, J.; and Kim, H.-c. 2017. Prediction of arctic sea ice concentration using a fully data driven deep neural network. *Remote Sensing*, 9(12): 1305.

Choi, M.; De Silva, L. W. A.; and Yamaguchi, H. 2019. Artificial neural network for the short-term prediction of arctic sea ice concentration. *Remote Sensing*, 11(9): 1071.

Girard, L.; Bouillon, S.; Weiss, J.; Amitrano, D.; Fichefet, T.; and Legat, V. 2011. A new modeling framework for sea-ice mechanics based on elasto-brittle rheology. *Annals of Glaciology*, 52(57): 123–132.

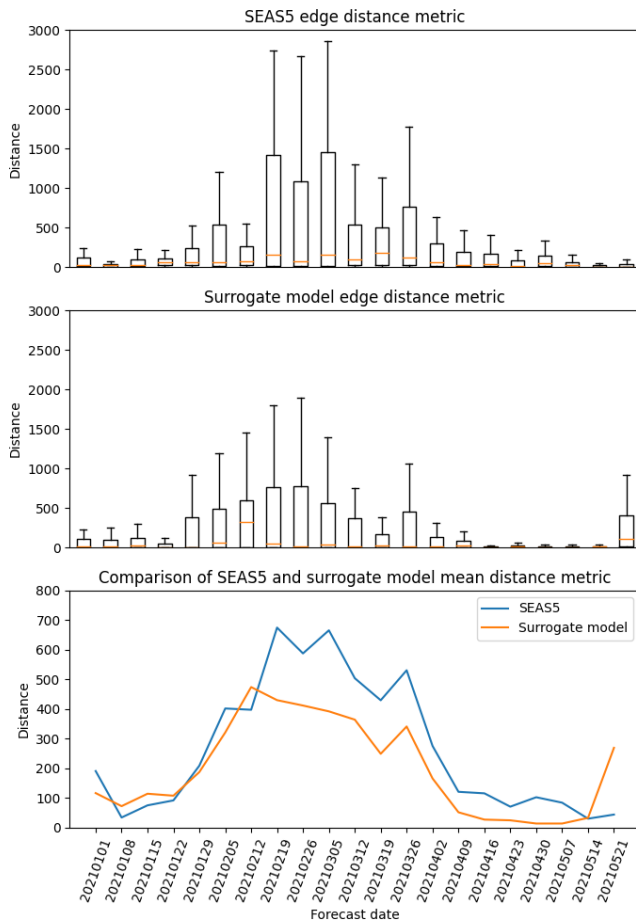Harsem, Ø.; Heen, K.; Rodrigues, J.; and Vassdal, T. 2015.

Figure 6: Boxplots of distance criteria for each point of ice edge - comparison of SEAS5 and surrogate model prediction

Oil exploration and sea ice projections in the Arctic. *Polar Record*, 51(1): 91–106.

Hvatov, A.; Nikitin, N. O.; Kalyuzhnaya, A. V.; and Kosukhin, S. S. 2019. Adaptation of NEMO-LIM3 model for multigrid high resolution Arctic simulation. *Ocean Modelling*, 141: 101427.

Johnson, S. J.; Stockdale, T. N.; Ferranti, L.; Balmaseda, M. A.; Molteni, F.; Magnusson, L.; Tietsche, S.; Decremer, D.; Weisheimer, A.; Balsamo, G.; et al. 2019. SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3): 1087–1117.

Kapsch, M.-L.; Graversen, R. G.; Economou, T.; and Tjernström, M. 2014. The importance of spring atmospheric conditions for predictions of the Arctic summer sea ice extent. *Geophysical Research Letters*, 41(14): 5288–5296.

Kim, J.; Kim, K.; Cho, J.; Kang, Y. Q.; Yoon, H.-J.; and Lee, Y.-W. 2018. Satellite-based prediction of Arctic sea ice concentration using a deep neural network with multi-model ensemble. *Remote Sensing*, 11(1): 19.

Kim, Y. J.; Kim, H.-C.; Han, D.; Lee, S.; and Im, J. 2020. Prediction of monthly Arctic sea ice concentrations using satellite and reanalysis data based on convolutional neural networks. *The Cryosphere*, 14(3): 1083–1104.

Lam, R.; Sanchez-Gonzalez, A.; Willson, M.; Wirnsberger, P.; Fortunato, M.; Alet, F.; Ravuri, S.; Ewalds, T.; Eaton-Rosen, Z.; Hu, W.; et al. 2023. Learning skillful medium-range global weather forecasting. *Science*, eadi2336.

Madec, G.; Bourdallé-Badie, R.; Bouttier, P.-A.; Bricaud, C.; Bruciaferri, D.; Calvert, D.; Chanut, J.; Clementi, E.; Coward, A.; Delrosso, D.; et al. 2017. NEMO ocean engine.

Pemberton, P.; Löptien, U.; Hordoir, R.; Höglund, A.; Schimanke, S.; Axell, L.; and Haapala, J. 2017. Sea-ice evaluation of NEMO-Nordic 1.0: a NEMO–LIM3. 6-based ocean–sea-ice model setup for the North Sea and Baltic Sea. *Geoscientific Model Development*, 10(8): 3105–3123.

Stroeve, J.; Holland, M. M.; Meier, W.; Scambos, T.; and Serreze, M. 2007. Arctic sea ice decline: Faster than forecast. *Geophysical research letters*, 34(9).

Tonboe, R.; Lavelle, J.; Pfeiffer, R.-H.; and Howe, E. 2016. Product user manual for osi saf global sea ice concentration. *Danish Meteorological Institute: Copenhagen, Denmark.*

Wang, L.; Scott, K. A.; and Clausi, D. A. 2017. Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network. *Remote Sensing*, 9(5): 408.

Wang, W.; Chen, M.; and Kumar, A. 2013. Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system. *Monthly Weather Review*, 141(4): 1375–1394.

Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612.

Wu, D.; Tian, W.; Lang, X.; Mao, W.; and Zhang, J. 2023. Statistical Modeling of Arctic Sea Ice Concentrations for Northern Sea Route Shipping. *Applied Sciences*, 13(7): 4374.

Zhang, J.; and Rothrock, D. A. 2003. Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Monthly Weather Review*, 131(5): 845–861.

Zhou, Z.-H.; Wu, J.; and Tang, W. 2002. Ensembling neural networks: many could be better than all. *Artificial intelligence*, 137(1-2): 239–263.