https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Data-driven surrogate modeling of high-resolution sea-ice thickness in the Arctic

Charlotte Durand[1], Tobias Sebastian Finn[1], Alban Farchi[1], Marc Bocquet[1], and Einar Òlason[2]
[1]CEREA, École des Ponts and EDF R&D, Île-de-France, France
[2]Nansen Environmental and Remote Sensing Center, 5007 Bergen, Norway
Correspondence: Charlotte Durand (charlotte.durand@enpc.fr)

Abstract. A novel generation of sea-ice models with Elasto-Brittle rheologies, such as neXtSIM, can represent sea-ice processes with an unprecedented accuracy at the mesoscale, for resolutions of around $10\,\mathrm{km}$. As these models are computationally expensive, we introduce supervised deep learning techniques for surrogate modeling of the sea-ice thickness from neXtSIM simulations. We adapt a convolutional UNet architecture to an Arctic-wide setup by taking the land-sea mask with partial convolutions into account. Trained to emulate the sea-ice thickness on a lead time of 12 hours, the neural network can be iteratively applied to predictions up to a year. The improvements of the surrogate model over a persistence forecast prevail from 12 hours to roughly a year, with improvements of up to $50\,\%$ in the forecast error. The predictability of the sea-ice thickness measured against a daily climatology additionally lays by around 8 months. By using atmospheric forcings as additional input, the surrogate model can represent advective and thermodynamical processes, which influence the sea-ice thickness and the growth and melting therein. While iterating, the surrogate model experiences diffusive processes, which result into a loss of fine-scale structures. However, this smoothing increases the coherence of large-scale features and hereby the stability of the model. Therefore, based on these results, we see a huge potential for surrogate modelling of state-of-art sea-ice models with neural networks.

1 Introduction

Caused by climate change, the Arctic sea ice is melting and becoming thinner (IPCC, 2022). Analysis from satellite observations (Laxon et al., 2013; Kwok, 2018; Kacimi and Kwok, 2022) as well as direct measurements (Renner et al., 2014) have shown a decrease of the sea-ice volume. This loss may be unprecedented over the past century (Schweiger et al., 2019). It can lead to changes in global ocean circulation (Kwok et al., 2013), weather patterns (Screen, 2013; Barnes and Screen, 2015; Jung et al., 2014) with potential impacts on global sea levels, coastal communities, and commercial activities. Sea-ice models are used to simulate and predict the changes in sea-ice cover and its effects on the Arctic and global climate. These models are based on a combination of observational data and theoretical understanding of the physical processes that govern sea-ice dynamics. They are essential conceptual and numerical tools to understand the past, current and future state of the Arctic sea-ice, and to identify the key processes that drive its changes.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Here, we present a novel way to make use of data coming from theoretical understanding of the physical processes: based on neural networks, we build a surrogate model for the sea-ice thickness as simulated by the Arctic-wide neXtSIM model (Rampal et al., 2016; Ólason et al., 2022).

Several sea-ice models of the Arctic, like CICE (Hunke et al., 2017) and SI3 (Sievers et al., 2022), are concurrently developed for operational purposes: short-term predictions for maritime road and weather forecast as well as long-term simulations for climate projections. The recent development of models based on brittle rheologies (Girard et al., 2011; Rampal et al., 2016; Dansereau et al., 2016), like neXtSIM, can represent the observed effects of small-scale processes onto the resolved mesoscale with $\sim 10\,\mathrm{km}$ horizontal resolution (Bouchat et al., 2022). Small-scale sea-ice dynamics also impact the global sea-ice mass balance (Boutin et al., 2022). Divergent features in the ice, like leads and polynyas, are associated with localized intense ocean heat loss that enhances sea-ice production in winter (Kwok, 2006; von Albedyll et al., 2022), accounting for about 30% of the total ice production in the Arctic Ocean. Consequently, models correctly representing the effects of such small-scale can have also an advantage in representing the thermodynamics of sea ice.

Geophysical models are computationally expensive, especially for operational forecasts. However, geophysical models can be partially or completely emulated using data-driven surrogate models. Such surrogate models can speed up the forecasting process, once their costly training phase is finished. Notably, the development of more powerful graphics processing units (GPU) in the past few years benefit the use of neural networks for surrogate modeling.

Over the past years, emulating or replacing geophysical models by neural networks has become a promising topic of research, with recent overviews by Bocquet (2023); Cheng et al. (2023). Emulating ERA5 data, the European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis product, recent examples of global-scale surrogate models adopt developments from computer vision by using graph neural networks (Keisler, 2022; Lam et al., 2022) and vision transformers (Bi et al., 2022; Nguyen et al., 2023).

By employing convolutional neural network architectures, Liu et al. (2020) and Andersson et al. (2021) have successfully shown that the sea-ice concentration can be predicted. Furthermore, Horvat and Roach (2022) and Finn et al. (2023) have recently presented neural network approaches to emulate wave-ice interactions and high-resolution sea-ice dynamics.

Encouraged by such examples, we introduce a neural network to emulate the sea-ice thickness from Arctic-wide neXtSIM simulations. Using a convolutional U-Net architecture, we train the network to predict the thickness for a lead-time of 12 hours based on initial thickness conditions and atmospheric forcings. This surrogate model can be then sequentially applied to obtain sea-ice thickness predictions for seasonal time-scales.

We concentrate the surrogate model on the sea-ice thickness, as it is an important quantity for the forecast of sea ice and, yet, difficult to predict, especially on short time-scales (Zampieri et al., 2018; Xiu et al., 2022). Nonetheless, the thickness contains useful information for seasonal forecast (Balan-Sarojini et al., 2021) with direct links to other important quantities, like the sea-ice concentration and sea-ice extent. Our results show that neural networks can predict the sea-ice thickness at the mesoscale for daily and monthly time-scales with predictability up to half a year. The neural network learns to represent the advection and thermodynamics needed for predictions at longer time-scales, whereas the effects of leads and subgrid-scale dynamics are diffused over time.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Our surrogate model is trained to minimize the $\mathcal{L}_2$ error. This type of error metric tends to smooth out features of the fields that lead to double penalty errors, like leads in sea ice. Explained differently, the surrogate model is trained to reduce errors. This diffusion process has been previously observed for deterministic neural networks (Ravuri et al., 2021), when optimized on $\mathcal{L}_2$ error, but also within many forecasting and data assimilation problems in geosciences (e.g., Amodei and Stein, 2009; Farchi et al., 2016; Vanderbecken et al., 2023). To quantify the diffusion, we propose in this paper an analysis based on the Power Spectral Density (PSD).

Section 2 introduces the dataset from which we learn the data-driven model and its structure. Section 3 presents the neural network framework, the choices we made about its architecture, and its optimization. Section 4 introduces the metrics for evaluating the results on the surrogate model. Section 5 delivers and discusses the results of the neural networks training, forecast skill abilities and advection capabilities of the surrogate model, as well as an analysis of the diffusion phenomenon introduced above. The discussions and conclusions are given in Section 6 and Section 7. Appendices provide technical information and further illustrations of the results.

2 Description of the dataset

Our goal is to train a neural network to surrogate model the sea-ice thickness (SIT) for a lead time of $12\,\mathrm{h}$. As training dataset, we extract the SIT from neXtSIM simulations and atmospheric forcings from the ERA5 reanalysis, which we introduce in the following.

2.1 neXtSIM model and the sea-ice thickness

NeXtSIM is a dynamic and thermodynamic sea-ice model (Rampal et al., 2016). It currently uses Brittle Bingham-Maxwell rheology (Boutin et al., 2022; Ólason et al., 2022) to emulate the mechanical behavior of sea ice. NeXtSIM can represent the observed fine-scale dynamics of sea ice, including its scaling and multifractal properties in space and in time (Rampal et al., 2019; Bouchat et al., 2022). The model is discretized on a Lagrangian triangular mesh. The model output is projected on a static quadratic grid, on which our surrogate model is based. The sea-ice model is coupled with the ocean part of NEMO, OPA, (version 3.6, Madec et al., 1998; Rousset et al., 2015). The model configuration is further detailed in Appendix A.

In this study, we extract only the sea-ice thickness, the variable predicted by the neural network. We rely on simulations from 2006 to 2018. As model area, the simulations use the regional CREG025 configuration (et al, 2021), a regional extraction of the global ORCA025 configuration developed by the Drakkar consortium (Bernard et al., 2006). This area encompasses the Arctic and parts of the North Atlantic down to $27°$ with a nominal horizontal resolution of $\frac{1}{4}^\circ$ ($\simeq 12\mathrm{km}$ in the Arctic basin). The outputs projected onto the static grid can be seen as two-dimensional images with 603×528 grid-cells. Without degrading its information content, we can crop the data to 512×512 grid-cells: lower latitudes are removed, as well as zones in Eastern Europe and America, where no sea ice appears. An example model snapshot is presented in Fig. 1.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

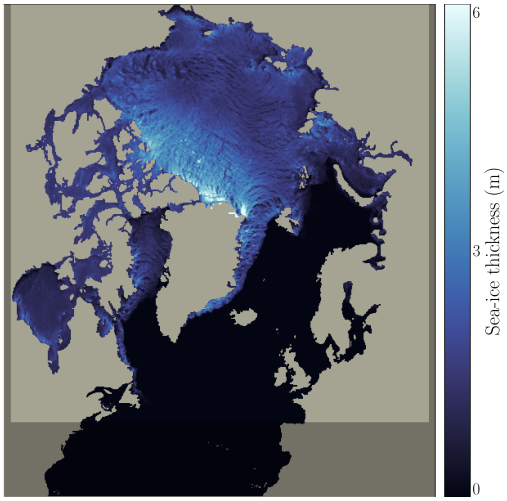

Figure 1. SIT simulated by neXtSIM at 15:00 UTC on 03/03/2009. The shaded area represent the cropped grid-cells that are further removed in order to keep a 512×512 grid cell SIT field, without degrading the information

2.2 Forcing fields

Several atmospheric forcings are added as input fields to the neural network, as the dynamics of the sea-ice thickness are especially driven by the atmosphere (Guemas et al., 2014): the subseasonal to interannual variability in the Arctic surface circulation is predominantly influenced by patterns in the atmospheric wind (Serreze et al., 1992). The atmospheric winds play a crucial role in shaping and driving the circulation patterns of the Arctic Ocean, which in turn affects the movement and distribution of sea ice. Additionally, fluctuations in the atmospheric surface temperature have a significant impact on the Arctic sea ice variability (Olonscheck et al., 2019). Changes in atmospheric temperature directly affect the growth, melt, and overall state of sea ice in the Arctic region. Warmer atmospheric temperatures accelerate sea ice melting, leading to reductions in ice extent and thickness, while colder temperatures can promote ice growth and expansion.

Based on these considerations, we supplemented to the sea-ice thickness the 2-meter temperature (T2M), and the atmospheric u- and v-velocities in $10\,\mathrm{meters}$ height (U10 and V10). Those forcings come from the ERA5 reanalysis dataset (Hersbach et al., 2020) of the ECMWF.

ERA5 forcings are interpolated on neXtSIM Lagrangian grid using a nearest neighbors scheme. Furthermore, to guide the temporal development of the sea ice, forcings at time $t + 6\mathrm{h}$ and $t + 12\mathrm{h}$ are added as predictors to the neural network, as commonly done in sea-ice forecasting (Grigoryev et al., 2022).

Because these forcings are also to guide the neXtSIM simulations, we assume a perfect knowledge of the forcing, albeit operational sea-ice forecasts use atmospheric forecasts as forcings. Note, T2M forcings from ERA5 are known to have an important bias during the freezing period (Yu et al., 2021; Wang et al., 2019; Køltzow et al., 2022; Nielsen-Englyst et al., 2021).

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Nonetheless, in order to stay as close to the configuration of the neXtSIM simulations, we maintain the ERA5 reanalysis as forcings.

3 Learning the dynamics of sea-ice thickness with neural networks

In this section, we will provide a description of the neural network structure, its input and output, the training process, and the various neural networks that were trained. During training, the neural network is trained in a supervised setting. The input to the network consists of the concatenated sea-ice thickness and atmospheric fields, whereas the predicted target is the increment in sea-ice thickness over the subsequent 12-hour period. One challenge in training the neural network is dealing with unavailable data points caused by land grid-cells. To address this challenge, a technique called partial convolution is employed.

3.1 Preparation of the dataset for supervised learning

Let us represent the sea-ice thickness at time t by $\mathbf{x}_t \in \mathrm{R}^{512\times512}$. The land-masked grid-cells are systematically assigned a value of zero thickness. For small signal levels, the noise induced by the imperfections of the neural network can overshadow the signal contained in the data. Consequently, to increase the signal in the target and decrease the auto-correlation, we chose a lead time of $12\,\mathrm{hours}$, rather than the lead time of $6\,\mathrm{hours}$ as possible by the dataset.

The neural network is trained to predict the increment in SIT instead of the absolute SIT. The increments of the SIT $\mathbf{y}_{t+\Delta t}$ for $\Delta t = 12\,\mathrm{hours}$ are given by the difference to a persistence forecast,

$$\mathbf{y}_{t+\Delta t} \triangleq \mathbf{x}_{t+\Delta t} - \mathbf{x}_t. \tag{1}$$

Based on the current SIT $\mathbf{x}_t$ and given forcings $\mathbf{F}$, our objective is to construct a neural network $f_{\boldsymbol{\theta}}(\mathbf{x}_t, \mathbf{F})$ with its parameters $\boldsymbol{\theta}$, which predicts the SIT increment $\mathbf{y}^{\mathrm{f}}_{t+\Delta t}$,

$$\mathbf{y}^{\mathrm{f}}_{t+\Delta t} = f_{\boldsymbol{\theta}}(\mathbf{x}_t, \mathbf{F}). \tag{2}$$

The neural network is hereby trained to approximate the real increment estimated from neXtSIM simulations (1) such that $\mathbf{y}_{t+\Delta t} \approx \mathbf{y}^{\mathrm{f}}_{t+\Delta t}$ approximately holds.

A table detailing the inputs and target for the neural network is shown in Tab. 1. In order to represent the temporal development of the sea-ice thickness with the neural network, we add to the inputs the SIT and the atmospheric forcings at $t - \Delta t$. When the neural networks are trained on those fields at time $t - \Delta t$ and t, there are called later 'with 2 inputs'. Otherwise, the neural networks are trained 'with 1 input' which correspond to the last three columns of the inputs described in the table.

Data from 2009 to 2016 is used for training, giving 11584 training samples. 2017 is used for the validation of the learned neural network and all the preliminary tests of the surrogate model. 2018 is used as year for testing: the results were evaluated once and only once on this year at the end of the study, after the hyperparameters were chosen for the neural network. For longer forecasts, to evaluate seasonal forecasts, another test dataset was selected between 2006 and 2008.

The input and target data are normalized by a global per-variable mean and standard deviation. These statistics are estimated over the entire training dataset and applied to all datasets.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Table 1. Inputs and targets for the neural networks. The table shows the predictors, including sea-ice thickness (SIT) at different time steps, and atmospheric variables: 2-meter temperature (T2M), 10-meter wind components (U10 and V10). The target is the increment of the SIT $12\,\mathrm{hours}$ later ($\Delta t = 12\,\mathrm{h}$). The evaluated neural networks use either the last 3 columns as input, learning with a single timestep for SIT (x_t), or all columns, learning with both $x_{t-\Delta t}$ and x_t. Note, as the SIT is the predicted quantity, there are no SIT values in the inputs for time steps larger than t.

Inputs				Target
SIT($t-\Delta t$)	SIT(t)	-	-	Δ SIT($t+\Delta t$)
T2M($t-\Delta t$)	T2M(t)	T2M($t+\Delta t/2$)	T2M($t+\Delta t$)	
U10($t-\Delta t$)	U10(t)	U10($t+\Delta t/2$)	U10($t+\Delta t$)	
V10($t-\Delta t$)	V10(t)	V10($t+\Delta t/2$)	V10($t+\Delta t$)	

3.2 Neural network architecture

Convolutional neural networks (CNN) are largely used in computer vision and have shown to be scalable to high-dimensional datasets (Pinckaers et al., 2022). These networks are based on convolutional layers, designed to recognize translation-invariant patterns. In the case of sea-ice thickness, the neural networks need to detect, e.g., leads, as well as the marginal ice zone, irrespective of their actual locations.

The UNet (Ronneberger et al., 2015) is a convolutional neural network architecture with an encoder-decoder structure. In the encoding part, convolutional layers and max-pooling layers are stacked in order to extract spatially more and more compressed features. As convolutional layers are localized by definition, the spatial compression helps the network to extract more globalized features. The number of successive resolution reduction defines the depth of the UNet. At the lowest resolution, in the bottleneck, several convolutional layers are stacked with 256 features (channels). In the decoding part, the features are up-sampled through a nearest neighbor interpolation and convolutional layers. Skip connections couple the encoding and decoding part at the same resolution level to facilitate training and to retain fine-granular information in the network. This neural network architecture is designed to extract multiscale features, notably present in sea-ice dynamics (Rampal et al., 2019). The UNet used here is described in detail in Appendix C and schematically outlined in Fig. 2.

The last layer of the neural network is a linear function without any activation, as we cast learning of the SIT increment as a regression problem. For all other layers, the *mish* activation function (Misra, 2019). As opposed to the more-often used rectified linear unit (relu), mish is a continuously differentiable function and has been previously proven to be effective in computer vision tasks (Bochkovskiy et al., 2020; Zhang et al., 2019), demonstrating improvements in training CNNs, particularly in addressing issues such as gradient explosion and gradient dispersion.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

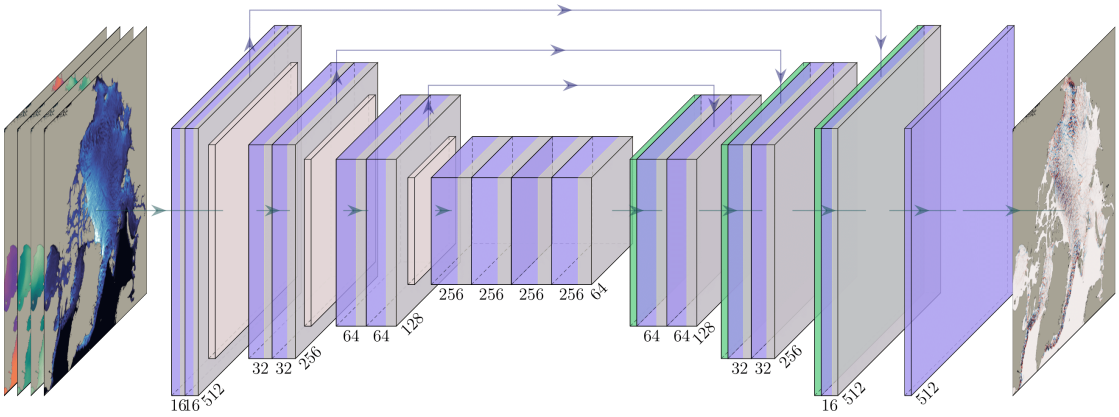

Figure 2. Architecture of the UNet-based neural networks. The UNet consists of three levels of depth with image sizes of 512, 256, and 128, in x- and y-direction. The input of the UNet includes either 10 or 14 channels, depending on whether only the current time step (x_t) or both the current and previous time steps (x_t and $x_{t-\Delta t}$) are used, alongside their associated atmospheric forcings. The input channels comprise sea-ice thickness, air velocities, and temperature. The number of channels for each convolution is indicated below it, with the first block having 16 channels. The upward arrows represent skip connections, allowing the neural network to retain information from earlier stages and incorporate it into subsequent stages, bypassing the bottleneck.

3.3 Partial convolution

As we can see in Fig. 1, the information on sea-ice thickness is only defined for ice and ocean grid-cells. Land grid-cells are masked. When performing two-dimensional convolutions on land cells, the presence of masked values has a detrimental effect on the local averages computed during the convolution operation. The convolution kernels then also includes the land cells with an assigned value of 0. One solution is to use partial convolutions (Liu et al., 2018) in every convolutional layer of the neural network. The key idea of partial convolutions is to separate the missing points from informative ones during convolutions, such that the results of convolutions only depend on sea and ice grid cells; land grid cells are simply omitted in the convolutional kernel. Let us see how it works in a simple example for a single convolution window.

Let us define $\mathbf{W} \in \mathbb{R}^{k_s \times k_s}$ and $b \in \mathbb{R}$, as the weights and bias of a convolution filter. k_s is the kernel-size of each convolution, always set to 3, except for the last layer of the neural network where it is set to 1. $\mathbf{X} \in \mathbb{R}^{k_s \times k_s}$ represents the pixel values (or feature activation values) being convoluted and $\mathbf{M} \in \mathbb{R}^{k_s \times k_s}$ is the corresponding binary mask which indicates the validity of each pixel/feature value: 0 for missing (land) pixels and 1 for valid (ocean and ice) pixels. The output of the proposed partial convolution $x' \in \mathbb{R}$, computed in a convolution window, is then

$$x' = \begin{cases} \mathbf{W}^{\mathsf{T}}(\mathbf{X} \odot \mathbf{M}) \frac{\mathrm{sum}(\mathbf{1})}{\mathrm{sum}(\mathbf{M})} + b & \text{if } \sum_{i,j} \mathbf{M}_{i,j} > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

where $\odot$ is an element-wise multiplication and $\mathbf{1}$ is a matrix of ones that has the same shape as $\mathbf{M}$. In comparison, a normal convolution would be defined as

$$x' = \mathbf{W}^{\mathsf{T}}\mathbf{X} + b, \tag{4}$$

independent of the validity of the grid cells.

From Eq. (3), we can see that the results of the partial convolution only depend on the valid input values (as $\mathbf{X} \odot \mathbf{M}$). The scaling factor $\mathrm{sum}(\mathbf{1})/\mathrm{sum}(\mathbf{M})$ adjusts the results as the number of valid input values for each convolution varies. It has been used previously in order to recover missing regions from observational datasets (Kadow et al., 2020). In this study, the goal is not to recover data, but to avoid artifacts near land caused by the underestimation in normal convolutions. The algorithm for partial convolution is further described in Appendix C1.

3.4 Global constraint on loss training

Emulating physical systems with neural networks can lead to a non-physical response (Beucler et al., 2021). In order to reduce a systematic bias of the surrogate model and to ensure that the neural network can correctly predict the global amount of sea ice, we add to the loss an additional penalty term. The non-penalized loss is defined by a pixel-wise mean-squared error (MSE), with $\mathbf{x}$ and $\mathbf{y}$ two vectors of dimension (N_x, N_y)

$$\mathcal{L}_{\mathrm{local}}(\mathbf{x},\mathbf{y}) = \mathrm{MSE}(\mathbf{x},\mathbf{y}) = \frac{1}{N_x \cdot N_y} \sum_{i,j}^{N_x,N_y} (x_{i,j} - y_{i,j})^2. \tag{5}$$

The penalization term is defined by the squared difference between the global mean of $\mathbf{x}$ and $\mathbf{y}$

$$\mathcal{L}_{\mathrm{global}}(\mathbf{x},\mathbf{y}) = (\bar{\mathbf{x}} - \bar{\mathbf{y}})^2 = \left(\frac{1}{N_x \cdot N_y} \sum_{i,j}^{N_x,N_y} x_{i,j} - y_{i,j} \right)^2. \tag{6}$$

This term weighted against the local loss with the help of a scalar λ

$$\mathcal{L}(\mathbf{x},\mathbf{y}) = \mathcal{L}_{\mathrm{local}}(\mathbf{x},\mathbf{y}) + \lambda \mathcal{L}_{\mathrm{global}}(\mathbf{x},\mathbf{y}). \tag{7}$$

λ is manually tuned to 100. The local loss is approximatively 4 orders of magnitude larger than the global loss. By setting $\lambda = 100$, the global loss is $1\,\%$ of the local loss. In the following parts, we show the results for $\lambda = 0$ and $\lambda = 100$. The former case will be called *unconstrained* and the latter *constrained*

3.5 Neural network training

The neural networks are trained on a single NVIDIA A100 GPU with a batch size of eight samples. As optimizer, AdamW (Loshchilov and Hutter, 2017) is used with a learning rate of $\gamma = 5 \times 10^{-5}$ and a weight decay, scheduled with a 3 steps piecewise constant decay, starting at $w = 1 \times 10^{-6}$. If the loss in the independent validation dataset plateaus for 20 epochs, the training is stopped early.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

We trained four different neural networks, as described in Table 2. By setting λ to either 0 or 100, we switch the additional loss function constrain on or off, checking its influence on the performance. Additionally, we test if additional temporal guidance by giving an additional time step as input helps the neural network to predict the increment in the sea-ice thickness.

Table 2. Comparison of the trained UNet-based neural networks in the study. Four neural network configurations are evaluated, varying in the number of inputs and the presence of a constraining term in the loss function. The inputs include either x_t alone or both $x_{t-\Delta t}$ and x_t, representing sea-ice thickness and atmospheric variables. The addition of a constraining term in the loss function regulates the neural network based on the global sea-ice thickness.

Neural network	Constraint (λ)	Inputs
NN 1 input - unconstrained	0	x_t (10 channels)
NN 1 input - constrained	100	x_t (10 channels)
NN 2 input - unconstrained	0	$x_{t-\Delta t}$, x_t (14 channels)
NN 2 input -constrained	100	$x_{t-\Delta t}$,x_t (14 channels)

4 Surrogate modeling and evaluation methods

4.1 Surrogate modeling

To emulate the physical model $\mathcal{M}^{\mathrm{p}}$, we built a surrogate model $\mathcal{M}^{\mathrm{s}}$ by applying the neural network $f_{\boldsymbol{\theta}}(\cdot,\cdot)$ that predicts the sea-ice thickness increment. Initializing the model with given initial conditions, $\mathbf{x}_{t_0}$, and given forcings $\mathbf{F}$, the surrogate model propagates the sea-ice thickness forward in-time, predicting the sea-ice thickness $\Delta t = 12\,\mathrm{hours}$ later,

$$\mathbf{x}^{\mathrm{f}}_{t_0+\Delta t} = \mathbf{x}_{t_0} + f_{\boldsymbol{\theta}}(\mathbf{x}_{t_0}, \mathbf{F}) = \mathcal{M}^{\mathrm{s}}(\mathbf{x}_{t_0}). \tag{8}$$

Note, for the ease of notation, we incorporate the forcings into the surrogate model. Using the forecasted state $\mathbf{x}^{\mathrm{f}}_{t_0+\Delta t}$ as next initial conditions, we can cycle the surrogate model and predict the sea-ice thickness for longer lead-times than Δt,

$$\mathbf{x}^{\mathrm{f}}_{t_0+N\Delta t} = \underbrace{\mathcal{M}^{\mathrm{s}} \circ \mathcal{M}^{\mathrm{s}} \circ \cdots \circ \mathcal{M}^{\mathrm{s}}}_{N\text{ times}}(\mathbf{x}_{t_0}) \tag{9}$$

$$= \mathbf{x}_{t_0} + f_{\boldsymbol{\theta}}(\mathbf{x}_{t_0} + f_{\boldsymbol{\theta}}(\ldots, \mathbf{F}), \mathbf{F}). \tag{10}$$

The forecast at longer lead times is consequently the initial conditions plus a recursive increment term.

Our baseline for the model comparison is constantly predicting the initial conditions without any increment, a so-called persistence forecast, i.e. the sea-ice thickness is unchanged over time. It is a commonly used baseline in sea-ice forecasting, as the auto-correlation of the sea-ice thickness in time is high up to a 1 month-lead time (Lemke et al., 1980; Blanchard-Wrigglesworth et al., 2011). We also compare the surrogate model to the daily climatology, computed on a day-of-year basis over the complete training dataset.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

4.2 Evaluation metrics for the surrogate

The goal of the surrogate model is to predict as accurately as possible sea-ice thickness over longer lead times than 12 hours, i.e. after several iterations of the surrogate model. We define the forecast skill of the surrogate at the k-th iteration by computing the root-mean-squared error (RMSE) between the predicted SIT and the actual SIT as simulated by neXtSIM,

$$\mathrm{RMSE}(k) = \frac{1}{N_s}\sum_{n=1}^{N_s}\sqrt{\frac{1}{N_x \cdot N_y}\sum_{i,j}^{N_x,N_y}\left(\mathbf{x}^{\mathrm{f}}_{t_n+k\Delta t} - \mathbf{x}_{t_n+k\Delta t}\right)^2}. \tag{11}$$

The RMSE between the prediction $\mathbf{x}^{\mathrm{f}}_{n+k\Delta t}$ and the simulation $\mathbf{x}_{n+k\Delta t}$ is computed over all pixels (i, j) of the field of size (N_x, N_y), for each sample n of the test set containing N_s trajectories, initialized at time t_n.

The global RMSE is calculated by averaging the RMSE values obtained when the whole sea-ice thickness fields are treated as single data point. It represents the discrepancy in averaged sea-ice thickness between the prediction and the simulation. By considering the global RMSE, we can assess the performance of the surrogate model in accurately reproducing the average sea-ice thickness compared to the reference model.

In order to quantify systematic errors of the surrogate model, we compute its mean error (bias). This metric tells about the ability of the neural network to correctly estimate the total amount of sea-ice in the full domain,

$$\mathrm{bias}(k) = \frac{1}{N_s}\sum_{n=1}^{N_s}\frac{1}{N_x \cdot N_y}\sum_{i,j}^{N_x,N_y}\left(\mathbf{x}^{\mathrm{f}}_{t_n+k\Delta t} - \mathbf{x}_{t_n+k\Delta t}\right). \tag{12}$$

The sea-ice extent (SIE) can be derived from the sea-ice thickness. We define a threshold $\sigma_{\mathrm{acc}} = 0.1\,\mathrm{m}$ for the SIT (see Appendix B for its definition) above which a grid point is considered as sea ice. By obtaining a classification mask between ice and no ice, we can easily define an accuracy metric based on the SIE. We define two terms: the first one $N_{>\sigma_{\mathrm{acc}}}$ indicates the number of pixels where $\mathbf{x}_{t_n+k\Delta t}$ and $\mathbf{x}^{\mathrm{f}}_{t_n+k\Delta t}$ disagree on the presence of sea ice, and the second one $N_{<\sigma_{\mathrm{acc}}}$ where the models disagree on the presence of open water. The accuracy is averaged over all N_s samples,

$$\mathrm{acc}_{\mathrm{SIE}}(k) = \frac{1}{N_s}\sum_{n=1}^{N_s}\left(1 - \frac{N_{>\sigma_{\mathrm{acc}}}(t_n+k\Delta t) + N_{<\sigma_{\mathrm{acc}}}(t_n+k\Delta t)}{N_{\mathrm{sea-ice\,pixels}}}\right). \tag{13}$$

4.3 Quantification of the diffusion effect

Diffusion can impact the accuracy and fidelity of the surrogate model's predictions. Excessive diffusion may lead to the loss of important details and reduce the model's ability to capture complex patterns. By quantifying diffusion, we can evaluate the model's performance and how the diffusion process evolves with increasing lead time.

To analyze the smoothing of features across multiple iterations, we want to find a metric that can describe the evolution of these features across different scales. Mathematicians have proposed several metrics for quantifying multifracality, such as box-counting algorithms and computation of fractal dimensions (Xu et al., 1993). For two-dimensional geophysical fields, the power spectrum density (PSD) has the ability to detect spatial properties over the different space scales (Lovejoy et al., 2008).

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

This quantity allows for a quantitative assessment of the changes in the features and its multiscale distribution as a function of the forecast lead time. Let $\mathbf{x}$ be a snapshot of the sea-ice thickness at a given time, either from neXtSIM or from the surrogate model. We define the PSD of $\mathbf{x}$ by:

$$P(k_x, k_y) = \|\mathrm{dft}(\mathbf{x})(k_x, k_y)\|^2, \tag{14}$$

where $\mathrm{dft}(\mathbf{x})$ is the discrete Fourier transform of $\mathbf{x}$. The PSD P is indexed by the spatial wave numbers k_x and k_y. The energy as a function of the wave vector is in turn related to P via

$$E(k_x, k_y) = (P(k_x, k_y))^2. \tag{15}$$

The power-law behavior of a field's energy spectrum can be justified by the underlying self-similarity or fractal nature of the image. Fractals are patterns or objects that display similar structures and statistical properties at various scales. In the case of an image, this means that certain statistical characteristics, such as texture or pixel intensity variations, repeat themselves across different scales. This energy spectrum can be identified with a power-law,

$$E(\mathbf{k}) \sim \|\mathbf{k}\|^{-\beta}, \tag{16}$$

where β is called a *spectral exponent*. In practice, this exponent can be numerically estimated by a linear regression between $\ln E$ and $\ln \|\mathbf{k}\|$. The power-law nature of the energy spectrum reflects the scaling properties of the field, where the statistical variations remain consistent regardless of the scales being observed. The power-law exponent β determines the degree of self-similarity and how quickly the energy decreases as the frequency or spatial scale increases.

We define the spectral exponents' ratio Q_β after t iterations by

$$Q_\beta(t) = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\beta^i_{\mathrm{surr}}(t)}{\beta^i_{\mathrm{neXtSIM}}(t)}, \tag{17}$$

the average over the full testing set of the ratio between the spectral exponent of predicted fields from the surrogate model and the spectral exponent of the fields as simulated with neXtSIM at the same time. If the surrogate model exhibits processes which lead to over-diffusion, then the spectral exponent of the predicted SIT should be larger than that of the actual SIT, resulting into a ratio $Q_b > 1$. Hence, the ratio corroborates the emergence of diffusion in the forecast.

5 Numerical results

In this section, the trained neural networks are assessed with their predictive performance in the test dataset, how they can be used for surrogate modeling on daily-time scales, and how they perform for seasonal scales. Additionally, the diffusive behavior of the neural networks is quantified. Hereafter, we present the results for each model based on their best validation loss.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

5.1 Short-term forecasting

In this paragraph, we will assess the performance of the surrogate model on a short-term timescale, specifically up to a one-month lead time. The metrics mentioned in Sec 4.2 are computed using the 2018 test dataset. They are reported in Table 3. We believe that the global RMSE serves as a proxy for the consistency of the surrogate model, which we define as the averaged sea-ice thickness in the domain. Based on this idea, we anticipate that the globally-constrained neural network will demonstrate improved performance for forecast lead times exceeding 12 hours.

Table 3. Statistical indicators to assess the performance of the surrogate models. The table shows the results for two lead time scenarios: 12 hours and 15 days. Two types of surrogate models are evaluated: those with 1 input (representing sea-ice thickness at time x_t) and those with 2 inputs (with SIT and atmospheric forcings at time $(x_{t-\Delta}, x_t)$). The models are trained with and without the addition of constrains, represented by a regularization parameter (Constrains). The evaluation metrics include RMSE, global RMSE, and SIE accuracy (ACC). Climatology and persistence baselines are included for comparison. Bold numbers indicate the best performing model in a given column.

		12 hours lead time		15 days lead time		
Surrogate	Constraint (λ)	RMSE ↓	Global RMSE ↓	RMSE ↓	Global RMSE ↓	ACC ↑
Climatology	-	3.76×10^{-1}	1.44×10^{-1}	0.376	1.44×10^{-1}	0.953
Persistence	-	6.20×10^{-2}	1.46×10^{-2}	0.293	4.91×10^{-2}	0.949
1 input	0	4.12×10^{-2}	6.57×10^{-3}	0.197	$\mathbf{6.96 \times 10^{-5}}$	0.963
1 input	100	4.17×10^{-2}	$\mathbf{3.88 \times 10^{-4}}$	**0.195**	8.53×10^{-4}	**0.970**
2 input	0	$\mathbf{3.56 \times 10^{-2}}$	1.81×10^{-3}	0.216	1.41×10^{-3}	0.966
2 input	100	3.63×10^{-2}	1.74×10^{-3}	0.202	8.77×10^{-5}	0.966

The introduction of a global penalization term in the constrained neural network reduces the global RMSE by one order of magnitude within 12 hours compared to the absence of penalization. However, the impact of the global loss term (as defined in Eq. 6) on the RMSE relatively small compared to the influence of including additional time steps.

On average, for the constrained neural network with one timestep as input, we observe a 33 % improvement after 12 lead hours, and a 33 % improvement after 15 lead days over the persistence on the RMSE. For the constrained neural network with two timesteps in the input, we observe a 41 % improvement after 12 lead hours, and a 31 % improvement after 15 lead days over the persistence. It is worth noting that all surrogate models exhibit significant improvements over climatology in forecasting sea-ice dynamics for a 15-day period. On average, these improvements amount to an 46 % enhancement compared to relying solely on climatology-based predictions for RMSE. The impact of adding the penalization term slightly reduce the RMSE of the surrogate after 12 hours but improves the surrogate RMSE after 15 days, of respectively 1% and 6 % for either one or two inputs (see Fig. 3a for the comparative evolution of the RMSE for 1 input up to a lead time of 25 days). The major improvement of adding the penalization term come from the global RMSE evaluation: after 12 hours, the global RMSE has improved by a factor 9.4 for the one input surrogate.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

We note that the global RMSE after 15 days is better in average for the unconstrained surrogate. Yet, as seen in Fig. 3, the strong advantage of the constrained surrogate is the important reduction of the bias standard deviation, represented in transparency in panel (b). This improvement is further supported by evaluating the averaged SIT over the entire year, as illustrated in panel (c). During periods of significant sea-ice production and melting, the surrogate model with the global constraint exhibits a closer alignment with the neXtSIM output, indicating a higher level of accuracy. These findings suggest that integrating a global constraint term during the optimization process enhances the surrogate model's ability to capture and reproduce the complex dynamics associated with sea-ice formation and melting.

As the neural network is trained for a 12 hours lead time, the first iteration of the surrogate model corresponds to its targeted lead time. In this first iteration, we observe that stacking two timesteps in the inputs of the neural network improves the surrogate by 13 % in terms of RMSE. In preliminary tests, we observed no further gain in performance with more than two timesteps as input.

The forecast skill for up to a lead time of 25 days highlights the overall improvement of the constrained surrogate model compared to the persistence forecast, Fig. 4, as similarly observed in Tab 3. For the constrained surrogates, the two inputs surrogate gain 14 % after 12 hours over the one input surrogate, but the results are reversed after 15 days, with a 3 % improvement for the one input surrogate, as we can also observe in the Fig. 4. If those results leaded us to select the surrogate model with two inputs, we need to remind that both biases are close to 0 and thus perfectly acceptable. While it is expected that the neural network gives a better RMSE with more inputs in the neural network, we hypothesize that increasing the number of inputs in the neural network leads to a higher likelihood of errors being introduced in the input data. As we cycle the neural network, the predictions from previous iterations are used as inputs for subsequent iterations. If there are errors or inaccuracies in these predictions, they can propagate and accumulate over time, potentially leading to a degradation in the quality of the inputs. As the surrogate learns the dynamics with "perfect" conditions, we increase the error of the inputs by having two timesteps as the input of the neural network after several iterations. Even if the two timesteps neural network provides better results for the first iterations and for the global RMSE, it seems more relevant to focus on longer lead times for model selection. In the next paragraphs, we will only present results for a surrogate model with one timestep in the input.

The evaluation of sea-ice extent (SIE) accuracy supports and strengthens our previous findings, providing additional evidence for the reliability of the results. In particular, we observe that the constrained neural network with one timestep as input consistently outperforms other models in predicting SIE. The higher accuracy achieved by the constrained neural network with one timestep as input suggests that this configuration effectively captures the relevant information and patterns necessary for accurate sea-ice dynamics prediction. This indicates that the specific constraints imposed during training, along with the inclusion of a single timestep as input, contribute to a better understanding and modeling of the underlying dynamics of SIT. The consistent performance of this model across different evaluation metrics and scenarios further validates its reliability and robustness. This surrogate configuration is able to capture the essential features and patterns of SIT dynamics, enabling more accurate predictions compared to other configurations.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

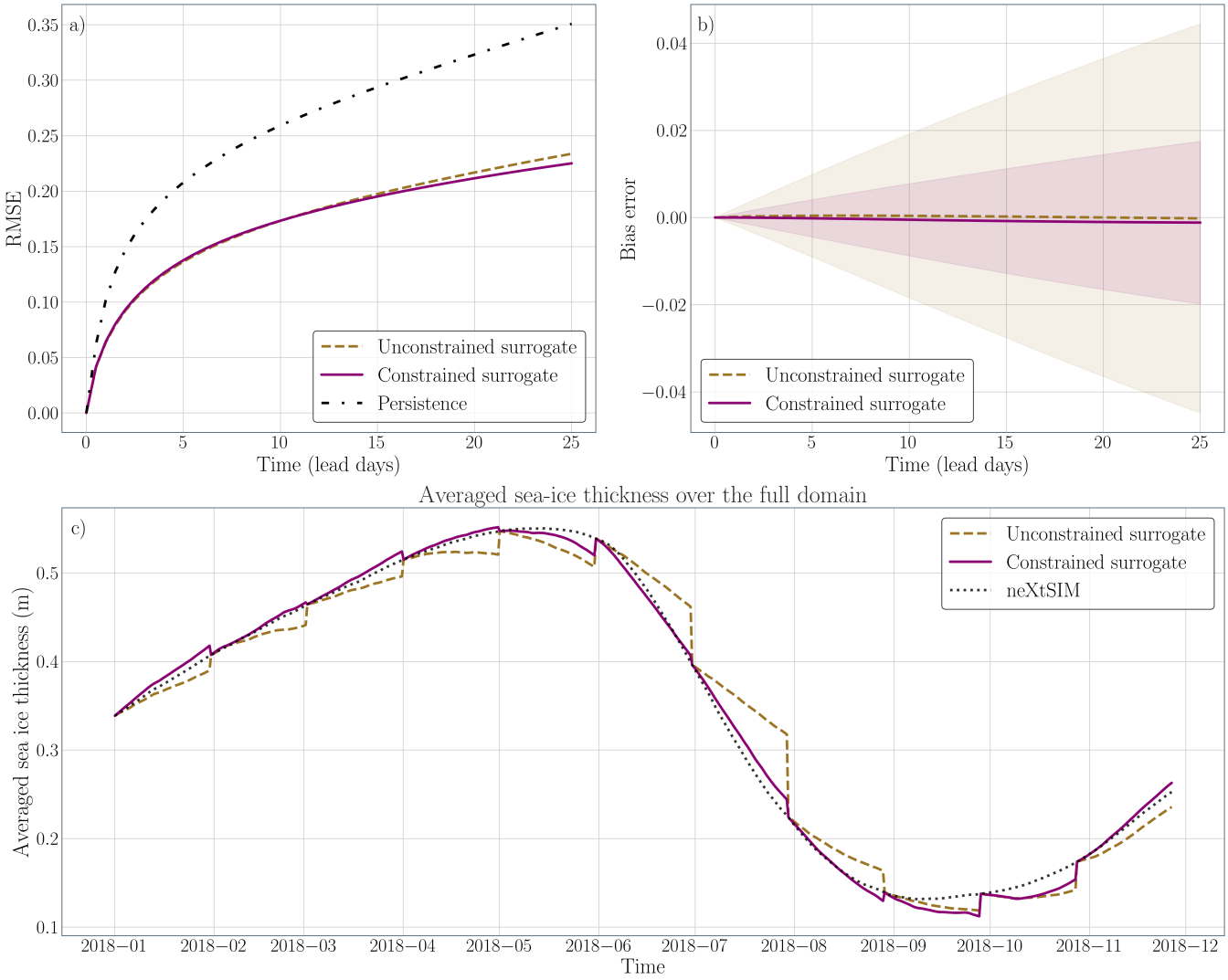

Figure 3. Analysis of the additional constraint during neural network training on the surrogate model over several iterations. In panel (a), the forecast skill of the surrogate model is depicted with solid lines representing the average. The unconstrained neural network is represented by the green line, while the constrained network is shown in purple. The black line represents persistence. Panel (b) displays the bias error associated with the surrogate, in transparency is represented the standard deviation, to outline the variance reduction of the constrained surrogate. In panel (c), the global conservation of sea-ice is plotted. The full-year trajectory is constructed by concatenating 60 forecast iterations. Every 30 days, the forecast is initialized using neXtSIM at the corresponding time and run for 60 iterations. The surrogate models are compared to the neXtSIM output over the same period.

Despite its ability to predict sea-ice thickness over the full domain, the forecast skill is not homogeneous over the different regions of the Arctic. The delimitation of the region from the National Snow and Ice Data Center was interpolated on the

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

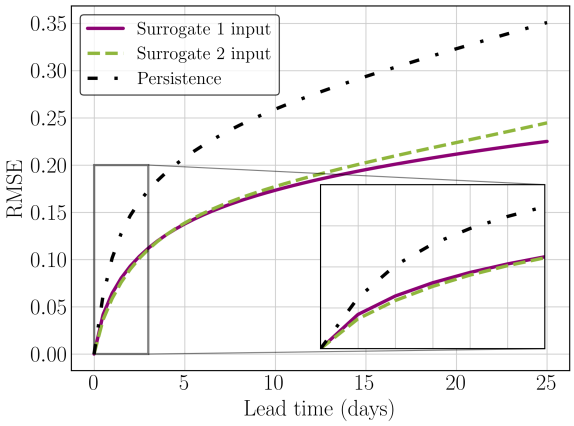

Figure 4. Comparison of Root Mean Square Error (RMSE) between the surrogate model and persistence approach for sea-ice thickness (SIT) prediction over a 25-day forecast horizon. The RMSE values are plotted, demonstrating the superior performance of the surrogate model in accurately predicting SIT compared to the persistence approach.

neXtSIM grid to then compute the forecast skill on the different regions, see Fig. 5 and numerical results for 25 lead days in Tab. 4. In Central Arctic, the surrogate forecast skill has an improvement of the RMSE of 31 % in average for a 25-day lead time over the persistence. The variability of the forecast skill is equal to 0.0590 and is 34 % lower than the variability of the persistence after the same lead-time. In Greenland Sea, the forecast skill of the surrogate is 35 % better than persistence for a 25-day lead time. The forecast skills of both the persistence and the surrogate are in average low because of the amount of sea-only pixels in this region during the full year. In every region, we systematically observe an improvement in both RMSE and its standard deviation of the surrogate over the persistence. This means that we improve over most samples the ability of the surrogate to predict the dynamics, across all regions. Notably, the Beaufort Sea exhibits a higher RMSE compared to the other regions. This discrepancy prompts further investigation into the surrogate model's limitations in accurately predicting sea ice dynamics near land in this region.

5.2 Advection

The surrogate model exhibits favorable advection properties, encompassing both large-scale and fine-scale dynamics. This successful advection can be attributed to the incorporation of atmospheric forcings in the model. The atmospheric forcings, which capture the influence of atmospheric conditions such as winds and temperatures, play a crucial role in driving the movement and behavior of sea ice.

The thickness field as well as the SIE are represented in Fig. 6 for neXtSIM (a) and the surrogate (b). Additional SIT fields are presented in the appendix for lead time of 5 days and lead time of 25 days in Fig. D1 and Fig. D2. The surrogate model seems to correctly advect the sea-ice sheet on the large scale. The inclusion of atmospheric forcings as inputs to the surrogate

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Table 4. Comparison of RMSE and its standard deviation (σ_{RMSE}) between the surrogate model and persistence for different regions. The table presents the mean RMSE and Std RMSE values for the Central Arctic, Greenland Sea, East Siberian Sea, Kara Sea, and Beaufort Sea as defined in Fig. 5.

	Surrogate		Persistence	
Regions	RMSE	σ_{RMSE}	RMSE	σ_{RMSE}
Central Arctic	0.267	0.0590	0.377	0.0892
Greenland Sea	0.139	0.0591	0.205	0.0786
East Siberian Sea	0.246	0.0727	0.366	0.0967
Kara Sea	0.230	0.1103	0.374	0.1822
Beaufort Sea	0.368	0.0921	0.553	0.1376

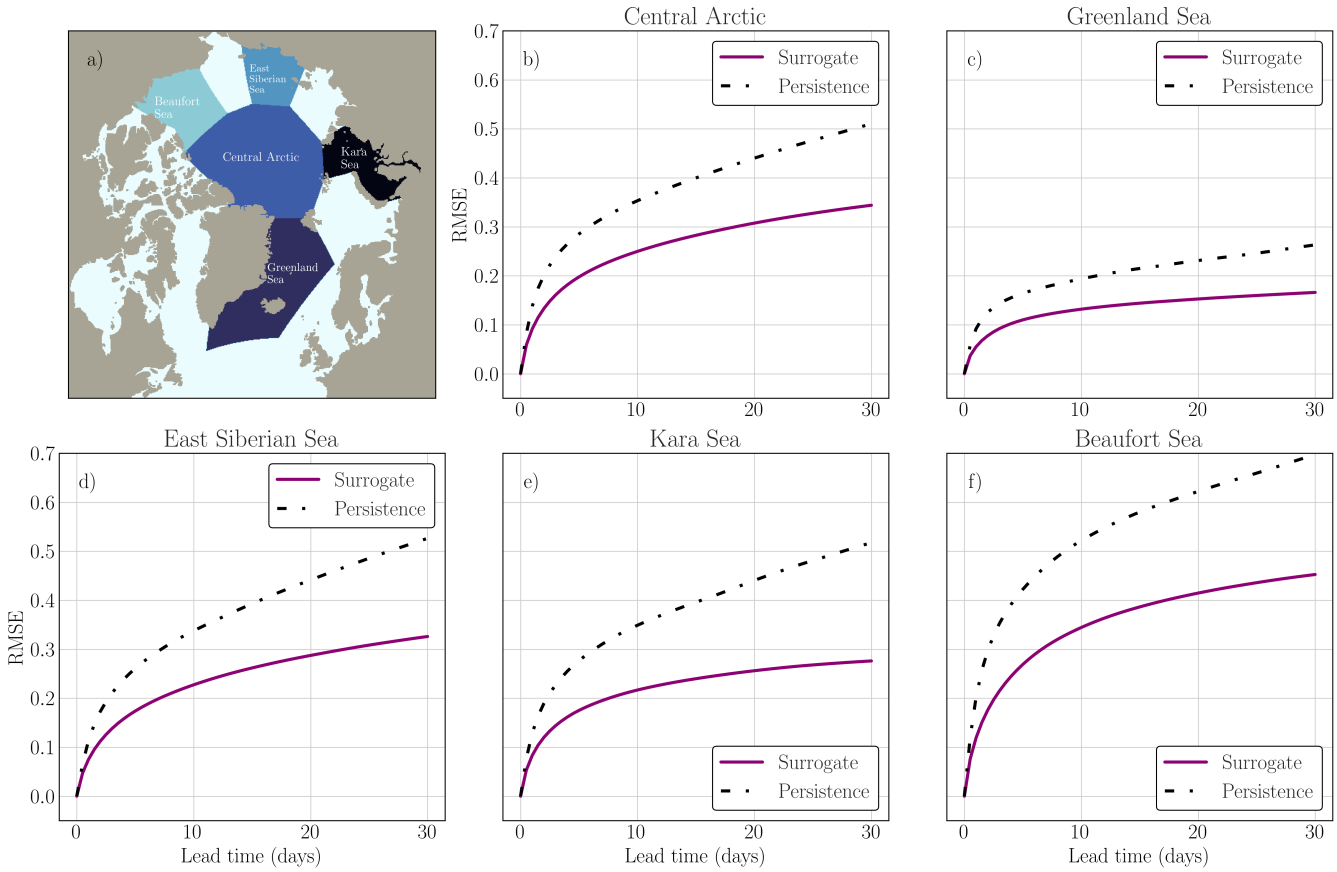

Figure 5. Regional forecast skills. Panel (a) illustrates the delineation of regions used for computing the regional forecast skill. Panels (b) to (f) show the averaged forecast skill over the full test year for specific regions: (b) Central Arctic, (c) Greenland Sea, (d) East Siberian Sea, (e) Kara Sea, and (f) Beaufort Sea. The purple line represents the surrogate model forecast skill, while the black line represents persistence.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

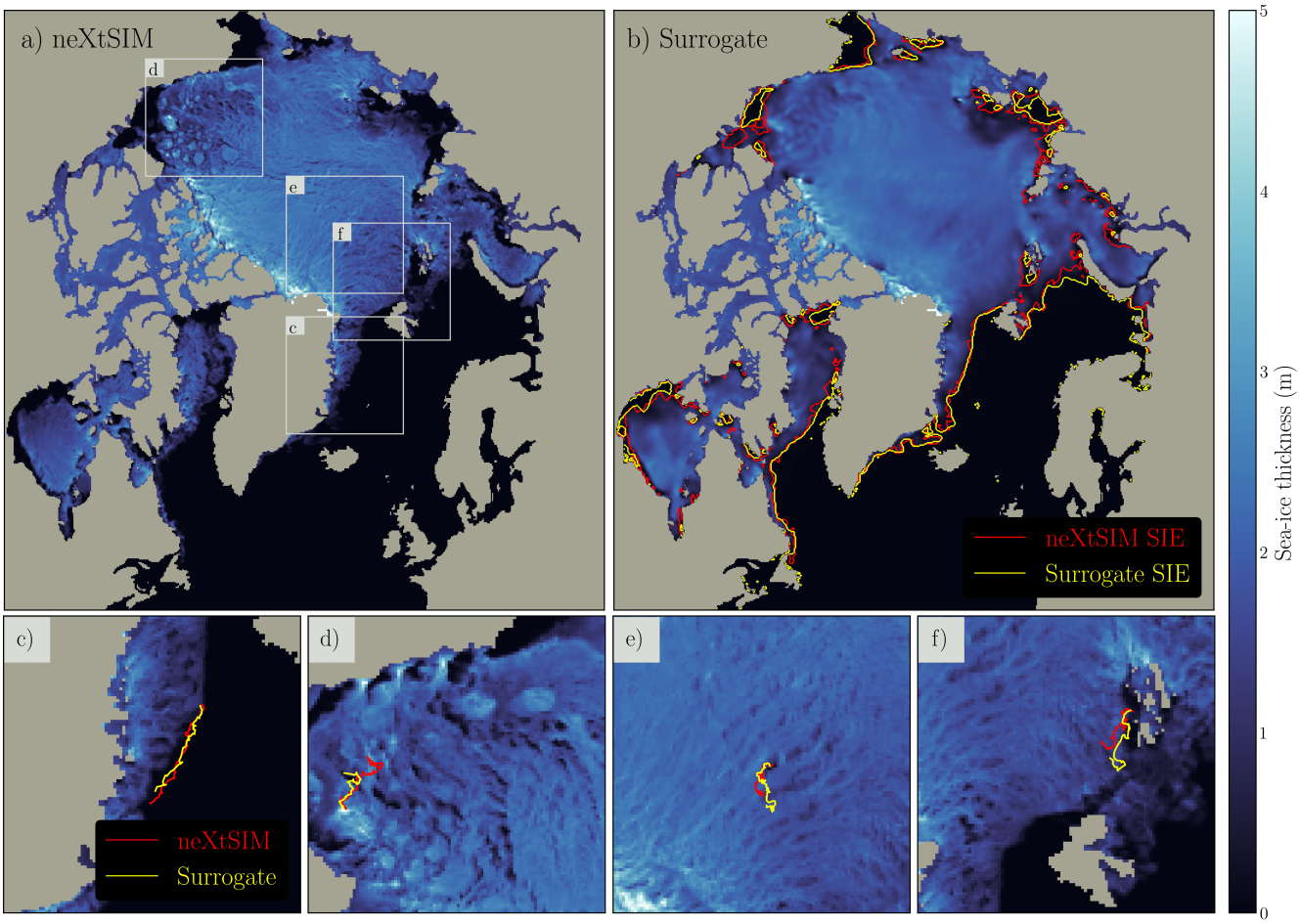

Figure 6. Evaluation of surrogate model advection performance. (a) neXtSIM sea-ice thickness (SIT) on May 30, 2018, with zoomed regions indicated for panels c), d), e), and f). (b) Surrogate model output on May 30, 2018, after a 30-day forecast initialized on May 15, with sea-ice extent (SIE) computed from neXtSIM (red) and the surrogate (yellow). Similar delimitation of sea-ice edges is observed in both curves. Panels c), d), e), and f) depict manual feature tracking in different Arctic regions: c) MIZ in Greenland Sea, d) Beaufort Sea, e) Central Arctic, and f) Barents Sea. The trajectories for 30 days are shown in red for neXtSIM and yellow for the surrogate model.

is crucial for capturing and learning the driving dynamics of sea-ice. By incorporating these forcings as inputs to the surrogate, we provide the model with the necessary information to better understand and predict the dynamic changes in sea-ice thickness. When training the surrogate model without incorporating any atmospheric forcings, the absence of advection becomes apparent. Without the driving influence of atmospheric conditions such as winds and temperatures, the surrogate model lacks the necessary information to simulate and reproduce the advection of sea ice, it tends to exhibit behavior similar to persistence.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

In order to verify this visual impression, we followed manually 4 remarkable features, (c) on the MIZ, (d) a feature in Beaufort Sea, (e) an important crack in central Arctic and (f) on the MIZ in the Barents Sea for 1 month prediction and compared the motion of those features between the surrogate model and the actual neXtSIM dynamics. The results depicting the sea-ice advection are illustrated in Fig. 6 in the lower panel. Notably, several features, particularly in the MIZ, demonstrate nearly identical displacements over this one-month period. These features appear to be accurately captured and reproduced by the surrogate model, reflecting its ability to simulate the advection of sea-ice. Slight deviations in trajectories are observed for features such as cracks within the sea-ice, but these differences do not indicate incoherent or erratic behavior.

5.3 Diffusion quantification

The observation of a smoothing effect on fine-scale features which increases with the forecast lead time aligns with our expectations. The optimization goal of training is to minimize the mean square error (MSE), which entails reducing discrepancies and errors by creating an average over the features. This smoothing effect can be seen in the upper panel of Fig. 7. While the surrogate model is able to predict the global and local advection patterns of sea-ice, it tends to average out the fine-scale features over successive iterations. The observed smoothing effect highlights the trade-off between capturing large-scale dynamics and preserving fine-scale features in the surrogate model. While the model may sacrifice some fine-scale details, it still retains the essential advection patterns and provides reliable predictions on a global scale.

This smoothness can be quantified by computing the power spectral density (PSD) (Hess et al., 2023; Neuhauser et al., 2022) and the Q_b ratio as defined in Sect. 4.3. The results are presented in 8. After 12 hours, the PSD of neXtSIM and the surrogate are close: the surrogate model exhibit similar multiscale properties than the physical model. We found the spectral exponent to be a good quantitative measurement of the diffusion process of the surrogate model. When the number of iteration of the surrogate model increase, the smoothness of the field is visually more important. This means that at high frequencies, we lose information, thus the PSD decrease on high frequency see Fig. 8 (b). The reduction in high-frequency PSD further supports the notion that the diffusion process within the surrogate model leads to a loss of detailed information and finer-scale features. When computing β, we see a fast increase of β when the number of cycle increase, 8 (c). In 10 lead-day time, we have an increase of the β exponent of 8% averaged over the full year. Interestingly, the spectral exponent does stabilize after 10 lead days and then slowly decrease. We hypothesize that the neural network has attained its resolution capacity for a correct advection of the sea-ice on the global scale by reducing the fine-scale dynamics that is inherently chaotic and stochastic.

5.4 Long-term forecast

In this section, we will discuss the ability of the surrogate model to forecast the dynamics of the sea-ice thickness at a seasonal scale. Seasonal forecast of arctic sea-ice is complex (Sigmond et al., 2013), and even more so on a high resolution grid. While previous results were presented with at most 60 iterations of the surrogate model, we present here runs of the surrogate with 720 iterations which correspond to 360 days forecast of the sea-ice. Those forecast are initialized from January 2006 to January 2008, with several initialization periods evenly distributed during that period. We have 25 year-long runs to evaluate

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

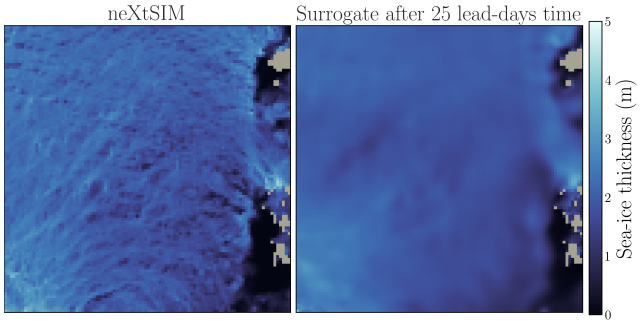

Figure 7. Comparison of neXtSIM (a) and surrogate (b) runs for a 30-day period in the Central Arctic. The fine-scale dynamics observed in panel (a) are smoothed in the surrogate model (b).

the surrogate model (see Fig. 9). In the appendix see Fig. D3, we propose some snapshots of the seasonal forecast model over the full year, accompanied by the SIE delimitation. The surrogate model is stable over the full year, and the RMSE reach the climatology after 5 months. In the bottom panel of the figure, the global averaged SIT for the surrogate model aligns with both the neXtSIM output and the climatology-based approach. The non-negligible bias of the climatology come from the fact that the daily climatology is computed over $2009 - 2016$ and is directly linked to sea-ice thinning. This consistency is partially attributed to the low bias exhibited by the constrained surrogate model, as demonstrated earlier. The constrained model's reduced bias helps maintain a physically realistic behavior of the sea-ice, even during long-term forecasts. The low error values observed during each iteration contribute to maintaining the physical integrity and conservation of the sea-ice in the surrogate model. This indicates that the surrogate model's predictions remain consistent with the overall dynamics of sea-ice, supporting its ability to capture the essential characteristics and behavior of the system. However, it is worth noting that when testing the surrogate model on a two-year forecast, difficulties were encountered in accurately predicting the dynamics beyond one year due to the constant, albeit slow, increase of the error of the surrogate.. In terms of sea-ice extent (SIE) prediction, the surrogate model demonstrates the capability to accurately forecast the edge of the sea-ice throughout the year, regardless of the initialization period, as shown in Figure 10. As anticipated, the surrogate model performs significantly better than persistence during periods of high variation, particularly during summer and autumn.

6 Discussion

Our proposed surrogate model based on a UNet neural network can emulate the large-scale sea-ice thickness as simulated by neXtSIM on daily and seasonal timescales. The main advantage of the emulator is the computational time needed for a forecast. Once the neural network is trained, computing one iteration of the surrogate, a $12\,\mathrm{hours}$ forecast, takes approximately $72\,\mathrm{ms}$ on a single NVIDIA A100 GPU. A forecast for one year takes under $1\,\mathrm{min}$. This opens the perspective to run a large ensemble of simulations for complex sea-ice models, which can facilitate data assimilation. Note that the training of the neural network

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

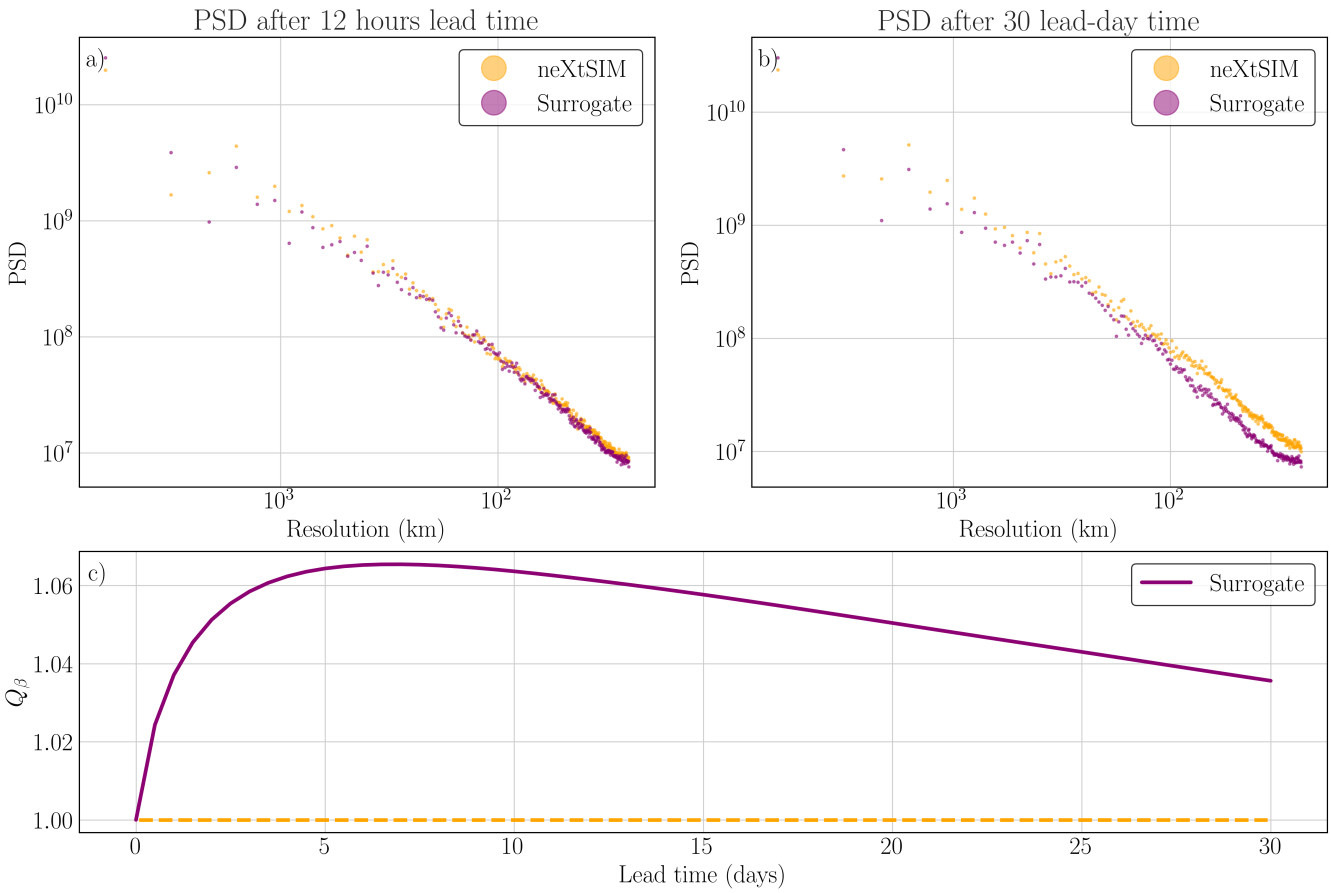

Figure 8. Evaluation of Power Spectrum Density and Diffusion Process in Surrogate Modeling. (a) Power spectrum density of one neXtSIM output (blue points) and one surrogate output after 12 hours (purple points), indicating a close match. (b) Power spectrum density after 30 lead-day time of the surrogate model compared to neXtSIM at the same time, showing lower values for high spatial frequencies in the surrogate model. (c) Q_β values quantifying the diffusion process for the surrogate, with the blue line representing neXtSIM for comparison.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

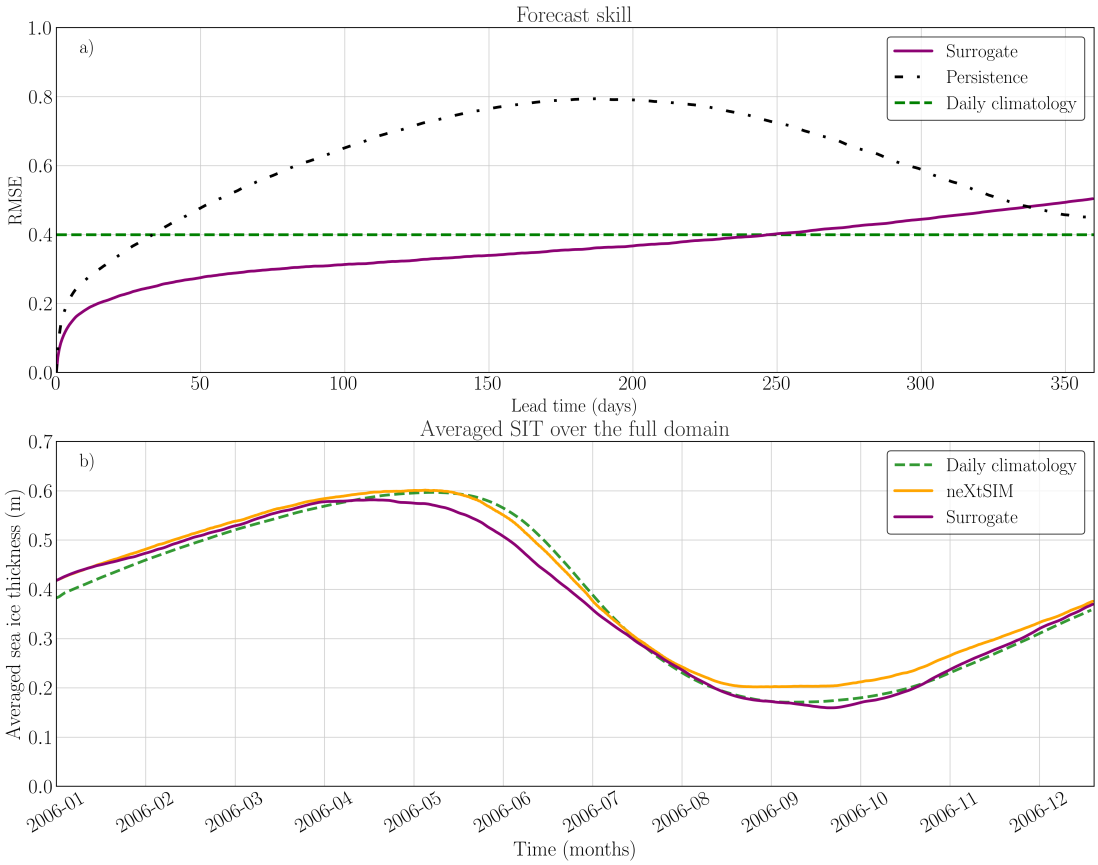

Figure 9. Performance evaluation of constrained surrogate model and comparison with persistence and climatology for seasonal forecast. (a) Forecast skill of the constrained surrogate model for year-long forecasts, based on 25 runs with different initialization times. The surrogate model is compared against persistence and daily climatology. (b) Global averaged sea-ice thickness throughout the year, starting in January 2006, comparing the surrogate model, the physical model neXtSIM, and the daily climatology. The seasonality of the SIT is well-preserved by the surrogate

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

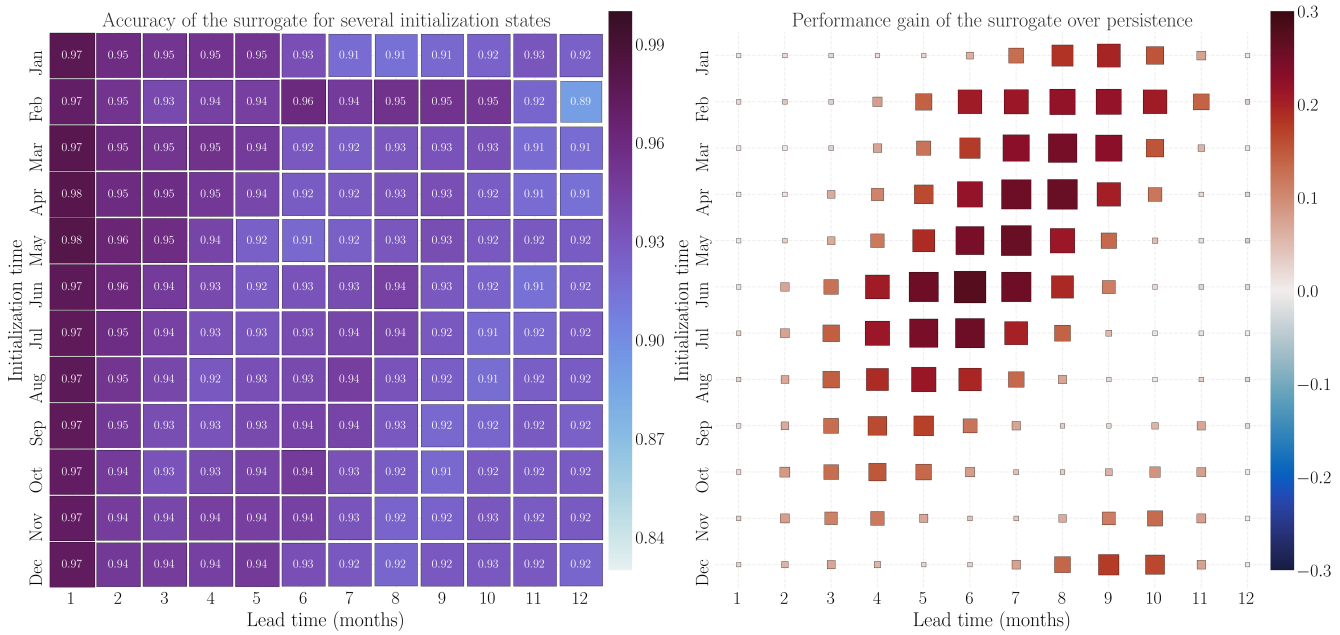

Figure 10. Estimation of sea-ice extent forecasting performance for different initialization times and forecast horizons. The left panel illustrates the sea-ice extent (SIE) accuracy on the $2006-2008$ test dataset with varying initialization times. The right panel displays the relative difference between the SIE accuracy of the surrogate model and persistence.

remains slightly costly, around $20\,\mathrm{h}$ on a single NVIDIA A100 GPU. Approximately, $10000\,\mathrm{hours}$ on NVIDIA A100 GPUs were necessary to conduct this study.

Using the PSD as a quantification of the diffusion shows that the diffusion process attains a threshold. We hypothesize that the surrogate model has reached its resolution capacity for correctly simulating the advection of sea-ice on a global scale. By reducing the inherently chaotic fine-scale dynamics, which exhibit an apparently stochastic behavior, the surrogate model achieves a balance between capturing the essential large-scale patterns and minimizing the impact of unpredictable fluctuations: chaotic and stochastic processes lead to a high sensitivity to initial conditions, making these processes difficult to model accurately. By prioritizing the larger-scale dynamics and averaging out fine-scale features, the surrogate model mitigates the influence of these chaotic and stochastic processes. This mitigation results into more stable and reliable predictions on a global scale. This hypothesis implies that the surrogate model focuses on capturing the dominant advection patterns that drive the overall behavior of sea ice, while sacrificing some of the finer details. Whereas this trade-off may result in a loss of information for fine-scale dynamics, it allows the model to provide valuable insights into global-scale advection patterns.

Regarding seasonal forecasts, the model is stable for lead times up to one year, even if the climatology has a smaller forecast error after 5 months. With this found ability to generate a year-long forecast, we can run a data assimilation scheme to estimate the initial conditions. Using the surrogate model allows us to simply compute an adjoint for the model by backpropagating through time. Thus, we can easily use a four-dimensional data assimilation scheme.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

This study only focuses on predicting the sea-ice thickness, an important variable especially in operating forecast. However, other variables, like the sea-ice velocity components, have important information for the prediction from the physical model. Using the interactions between different variables can provide valuable information to the surrogate model. Hence, learning to emulate these variables has the potential to improve the prediction of the sea-ice thickness. Nevertheless, multivariate modeling is a more complex objective for the neural network than univariate modeling.

In this study, we use ERA5 reanalysis forcings, the same forcing product that has forced the simulations with neXtSIM. Preliminary results of running the surrogate model with a CMIP6 product show that the surrogate is still stable on the long term, Fig. D4. The forcing derived from the CMIP6 is a free-running simulation and, consequently, results into a different instantiation of the sea-ice thickness than with the ERA5 reanalysis product. Yet, the model is still stable for others types of forcing. Therefore, the surrogate model has learned to represent the large-scale sea-ice dynamics needed to simulate the sea-ice thickness on daily and seasonal timescales.

7 Conclusions

A neural network can emulate the sea-ice thickness at a resolution of $10\,\mathrm{km}$ as similarly simulated by neXtSIM. Trained for prediction of a $12\,\mathrm{h}$ lead time, the neural network can be iteratively applied for surrogate modelling to forecast the thickness for up to one year. The advantage of the surrogate model over a persistence forecast prevails from the daily timescales, with improvements of around $36\,\%$, to seasonal time scales with more than $50\,\%$ improvement.

We introduce a regularization method for the training of the neural network, constraining the deviations of the global averaged sea-ice thickness from the targeted simulations. This regularization reduces the bias of the neural network and increases the global consistency. The increased consistency then results into a decreased forecast error on daily to weekly timescales.

By adding atmospheric forcings, the surrogate model can represent advective and thermodynamical processes that influence the sea-ice thickness on a large, Arctic-wide, scale. Hence, the seasonal predictions with the surrogate have a predictabilty of up to 8 months, measured by comparison to the daily climatology.

When the surrogate model is iterated, it exhibits diffusive processes, which lead to a smoothing of the prediction. These processes are gaining importance during the first iterations, as shown by a power spectral density analysis. Whereas the smoothing induces a loss of fine-scale features, it allows the model to stay coherent for the large-scale dynamics that impact the sea-ice thickness. Thanks to this coherency, the surrogate model correctly manages to estimate the global amount of sea ice over the full Arctic. Consequently, the surrogate model offers a stable low-resolution adjoint for the sea-ice thickness in neXtSIM, for example useable in a variational data assimilation framework.

The surrogate model can make year-long forecasts within a minute. Therefore, the surrogate model presents itself as an opportunity to estimate a large ensemble of simulations. Such a large ensemble can allow us the assimilation of previously unused observations into the sea-ice thickness.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Appendix A: neXtSIM configuration

In this Appendix, we will describe more precisely the configuration of neXtSIM used. The model relies for its rheology on the combination of the Bingham-Maxwell constitutive model (Bingham, 1922) and the Maxwell-Elasto-Brittle (MEB) rheology (Dansereau et al., 2016). The model equations are solved on an adaptive Lagrangian triangular mesh (Rampal et al., 2016) using a finite element method with a re-meshing protocol. This method helps preserve the gradients in the sea-ice fields which can come from the formations of leads and ridge. The main parameters used for the model are presented in table A1. neXtSIM in this case is coupled with a sea model (NEMO).

Table A1. neXtSIM main parameters, see (Boutin et al., 2022) for more details about the model coupling.

Parameter	Symbol	Value
Ice-atmosphere drag coefficient	C_a	1.6×10^{-3}
Ice-ocean drag coefficient	C_w	6.7×10^{-3}
Scaling parameter for the ridging threshold	P	$3\,\mathrm{kPa/m^{3/2}}$
Main model time step	Δt_m	$450\,\mathrm{s}$
Time step for sea-ice dynamics solver	Δt	$6\,\mathrm{s}$
Maximum thickness of newly formed ice	h_{max}	$18\,\mathrm{cm}$
Sea-ice albedo	a_{ice}	0.57
Snow albedo	a_{snow}	0.8
Critical thickness parameter for ice grounding	k_1	5

Appendix B: Definition of the accuracy

In order to build an accuracy metric to evaluate the ability of the surrogate to predict the sea-ice thickness, it is necessary to define a threshold value to differentiate non sea-ice grid points from sea-ice grid points. Sea-ice experts commonly define the Marginal Ice Zone (MIZ) with sea-ice concentration between 0.15 and 0.8 (Strong, 2012; Comiso, 2006; Rolph et al., 2020). On one month of neXtSIM output (124 snapshots), we compute the grid-points included in this MIZ definition, and then we compute the cumulative distribution of SIT on those grid-points, see Fig. B1. We select a value of $\sigma_{\mathrm{acc}} = 0.1$ for the threshold on SIT to define sea-ice extent. If the grid-point has a SIT above σ_{acc}, it is considered a sea-ice pixel, otherwise it is considered either open sea or land.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

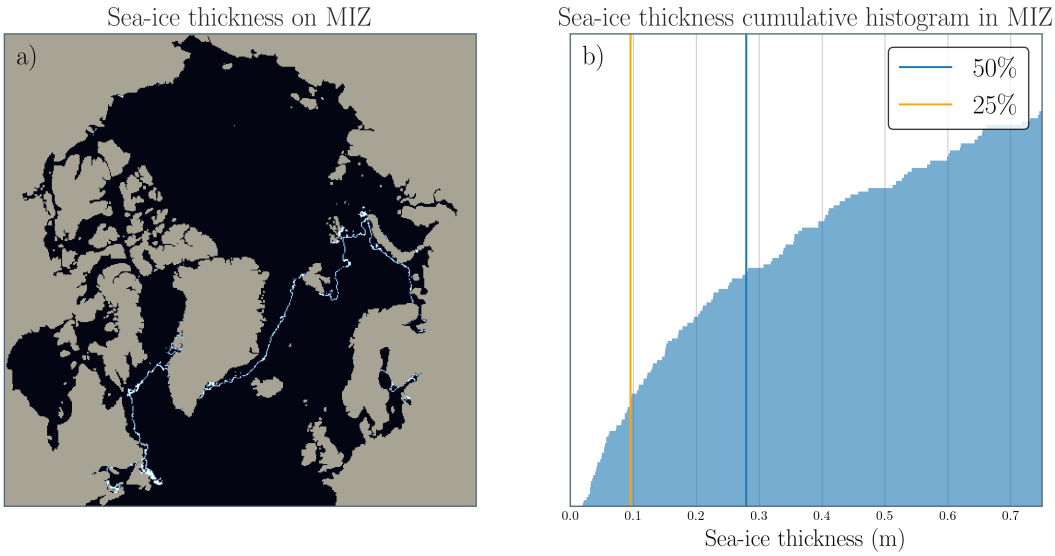

Figure B1. Choice of the SIT thickness threshold. On the left panel is shown the SIT thickness on the MIZ commonly defined with SIC. On the right panel is shown the cumulative distribution of this MIZ SIT. The blue vertical line outlines the mean of this distribution and the red line the 25% percentile which coincide with our chosen threshold for the SIT to define the SIE.

Appendix C: Neural networks architecture

We detail in this section the structure of the neural networks used in the paper. The models are implemented using Tensorflow and Keras. The next sections describe the implementation of partial convolution, and detailed structure of the UNet and ResNet neural networks.

C1 Partial Convolution algorithm

Instead of (Liu et al., 2020; Kadow et al., 2020) implementation of the partial convolution where the masks are convoluted alongside the images, we want to keep the mask constant only to represent the land around the sea-ice. Let's define M the mask for which 0 means a land pixel and 1 a valid pixel representing either ice or sea.

C2 UNet

The detailed structure of the neural network is described in table C1.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Algorithm C1 Partial Convolution pseudocode

Require: $X(n_b, n_x, n_y, n_c), M(n_x, n_y), k_s$
Compute image kernel
Compute zero-padding around (n_x, n_y)
$\tilde{M} = M +$ padding
Compute $W = \#k_s$
Compute $V = M \doteq 1_{windowsize}$
Compute $r_m = \frac{W \times M}{V + \epsilon}$
$\tilde{X} = X +$ padding
Apply Conv2D($\tilde{X} + \tilde{M}$)
$\tilde{X} = X \times r_m$
$\tilde{X} = \tilde{X} + b$
$\tilde{X} = \sigma(\tilde{X})$
return $\tilde{X}$

C3 Neural Network training

The losses as described in Section 3 are shown in Fig.C1 for the neural network with 1 input. The validation losses are plotted in transparency for each associated training losses. We do not observe overfitting which validate the size of the UNet with regard to the size of the dataset.

Appendix D: Surrogate modeling

In this section, we present more visual results for the surrogate model prediction and results with different type of forcings.

D1 Short-term forecast

In Fig. D1 and D2 we display several snapshots obtained from the constrained surrogate model alongside its corresponding neXtSIM state.

D2 Seasonal forecast

In Fig. D3 we can observe several snapshots obtained from the stable constrained surrogate model alongside its corresponding neXtSIM state. For every timestep shown, the surrogate model correctly manages to estimate the global state of the system. Yet, the smooth advection appears at the leading process. It is nonetheless a good approximation of the sea-ice structure during the full year.

As explained in Sect. 6, we evaluate our surrogate model approach with another type of forcings. The forcings are taken from the ECMWF-IFS-HR (25 km atmosphere and 25 km ocean) climate model (Roberts et al., 2017). The surface temperature and

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Stage	Layer	# params	n_x	n_y	$n_{channels}$
Input	PConv	3776	512	512	32
Down 1	PConv	9248	512	512	32
	PConv	9248	512	512	32
	PConv	9248	512	512	32
	BatchNormalization	128	512	512	32
	MaxPooling2D	0	256	256	32
Down 2	PConv	18496	256	256	64
	PConv	36928	256	256	64
	PConv	36928	256	256	64
	BatchNormalization	256	256	256	64
	MaxPooling2D	0	128	128	64
Bottleneck	PConv	14771	128	128	256
	PConv	59008	128	128	256
	PConv	59008	128	128	256
	PConv	59008	128	128	256
	BatchNormalization	1024	128	128	256
Up 2	UpSampling2D	0	256	256	256
	PConv	14752	256	256	64
	Concatenate	0	256	256	96
	PConv	55360	256	256	64
	PConv	36928	256	256	64
	BatchNormalization	256	256	256	64
Up 1	UpSampling2D	0	512	512	64
	PConv	18464	512	512	32
	Concatenate	0	512	512	64
	PConv	18464	512	512	32
	PConv	9248	512	512	32
Output	PConv	9248	512	512	32
	BatchNormalization	128	512	512	32
	PConv	33	512	512	1

Table C1. UNet architecture

velocities are taken, projected on neXtSIM grid and normalized to be able to be fit in the input of the surrogate model. The results are presented in Fig. D4. We see that there is after 50 days a constant bias differences between the surrogate with the different forcings, with the same global behavior. The surrogate model with CMIP6 forcings correctly handle the decrease of the SIE during September, but it has difficulties to match neXtSIM during the next refreezing period. We hypothesize this is

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

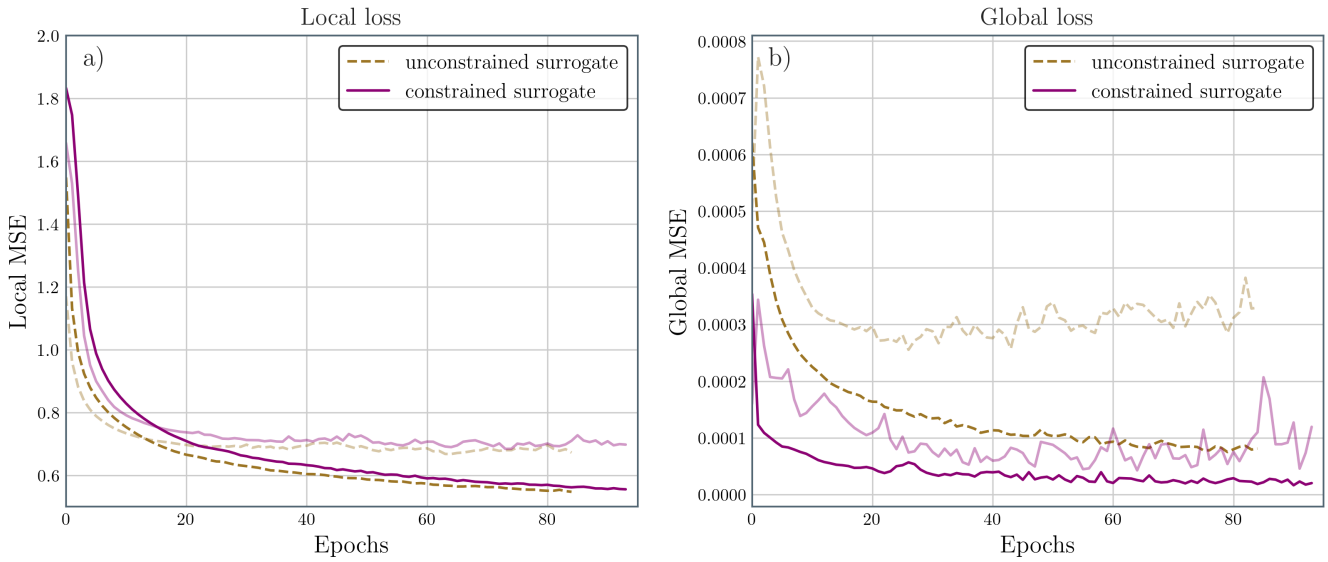

Figure C1. Training and validation (in transparency) losses for the neural networks with 1 timestep in the input. Yellow line represent the unconstrained neural network and purple the constrained neural network. We observe that adding the global term in the loss during training allows an important decrease of the global RMSE.

caused by the important bias difference between the forcings, neXtSIM being simulated with ERA5 forcings. In any case, our surrogate model remains stable when changing the forcings.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

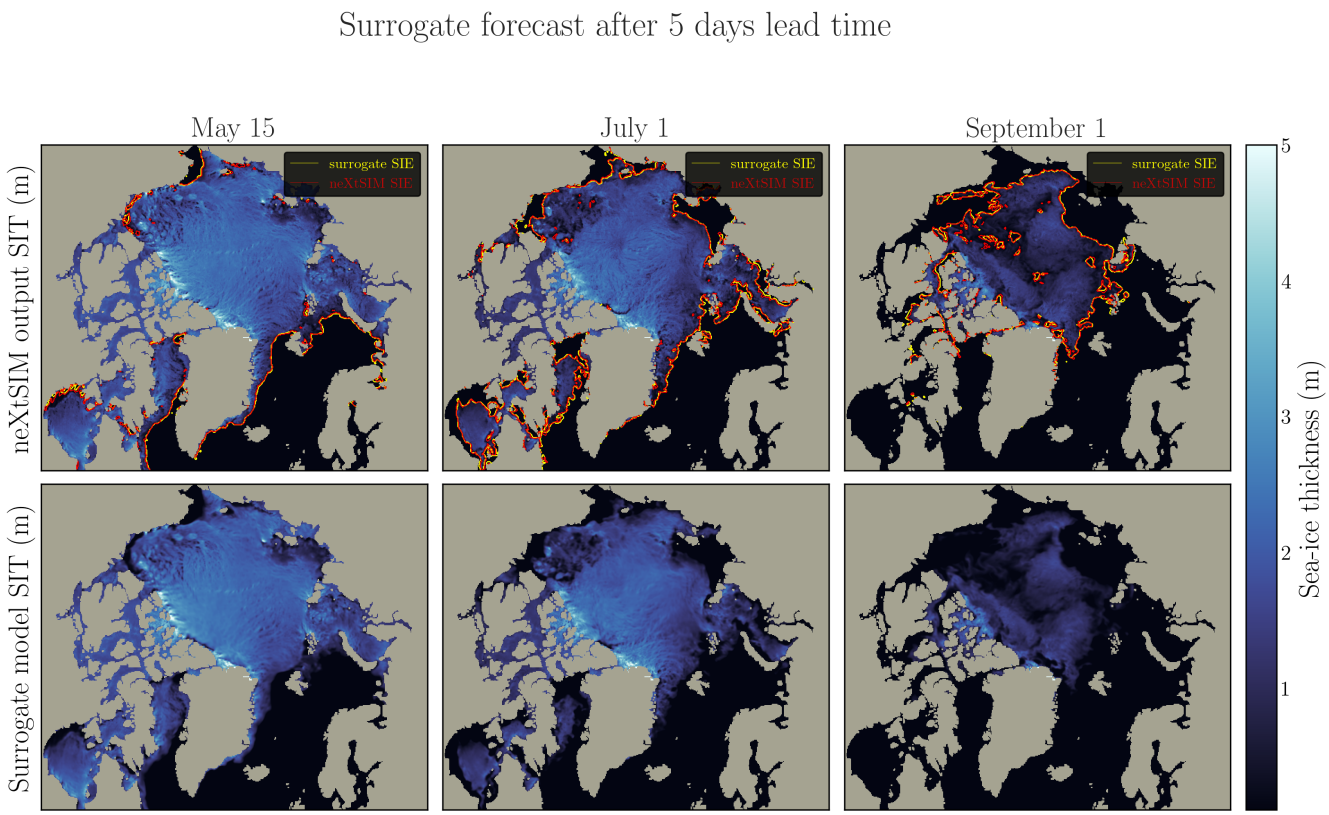

Figure D1. The figure presents the surrogate model using three different initialization states (May 15, July 1, and September 1). The surrogate model, with constrained inputs and 1-timestep configuration, is run for 5 lead days. The top panel illustrates the neXtSIM output, while the middle panel showcases the surrogate model output. Contour lines representing the sea-ice extent are displayed on the neXtSIM panel for both neXtSIM (red) and the surrogate (yellow), while the surrogate model examples omit these contours for clarity.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

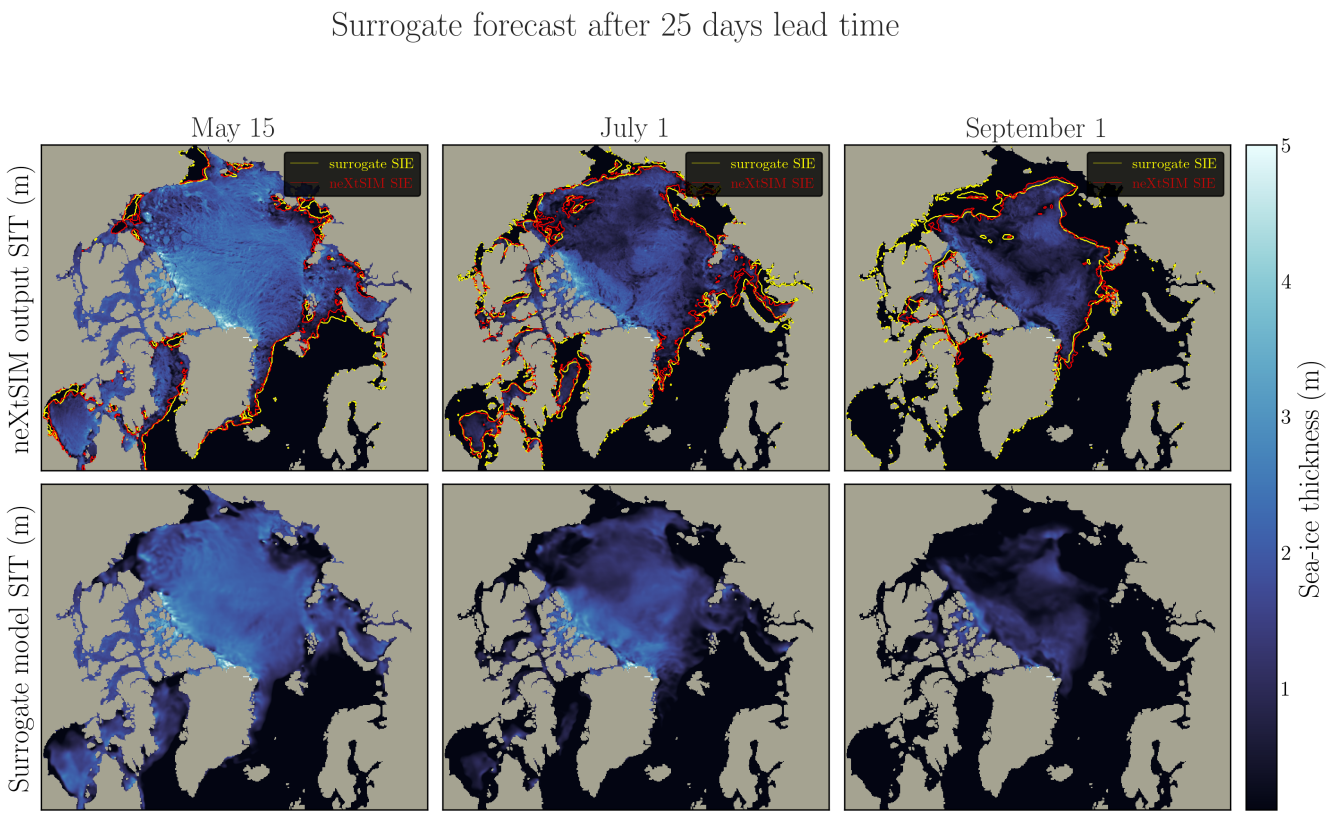

Figure D2. The figure presents the surrogate model using three different initialization states (May 15, July 1, and September 1). The surrogate model, with constrained inputs and 1-timestep configuration, is run for 25 lead days. The top panel illustrates the neXtSIM output, while the middle panel showcases the surrogate model output. Contour lines representing the sea-ice extent are displayed on the neXtSIM panel for both neXtSIM (red) and the surrogate (yellow), while the surrogate model examples omit these contours for clarity.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

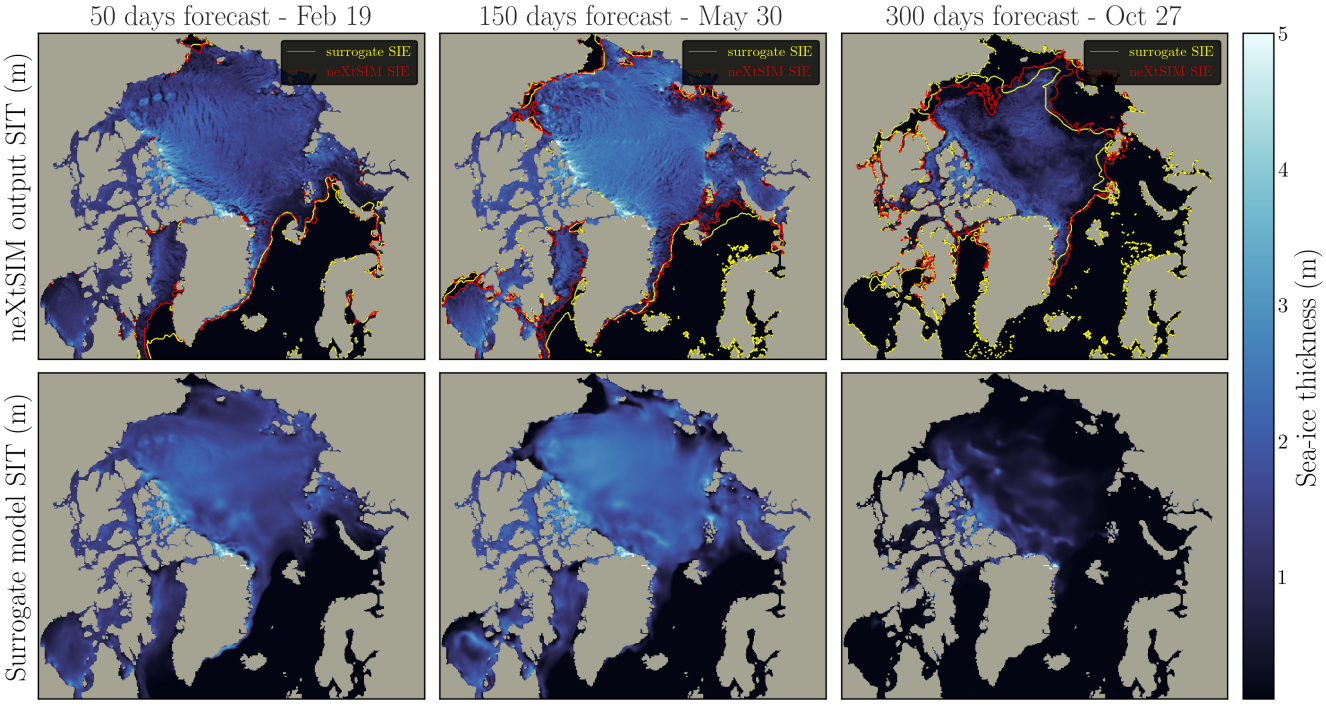

Figure D3. Snapshots for seasonal forecast of neXtSIM and the surrogate. The surrogate model is run starting from January 1st 2006, for 720 iterations. Results after 100, 300, and 600 iterations are presented in this figure. Results are plotted with the neXtSIM output above and the surrogate model in the middle panel. On the neXtSIM panel are plotted the contour of the sea-ice extent for both neXtSIM (red) and the surrogate (yellow). For better clarity, those contours are not represented on the surrogate model examples.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

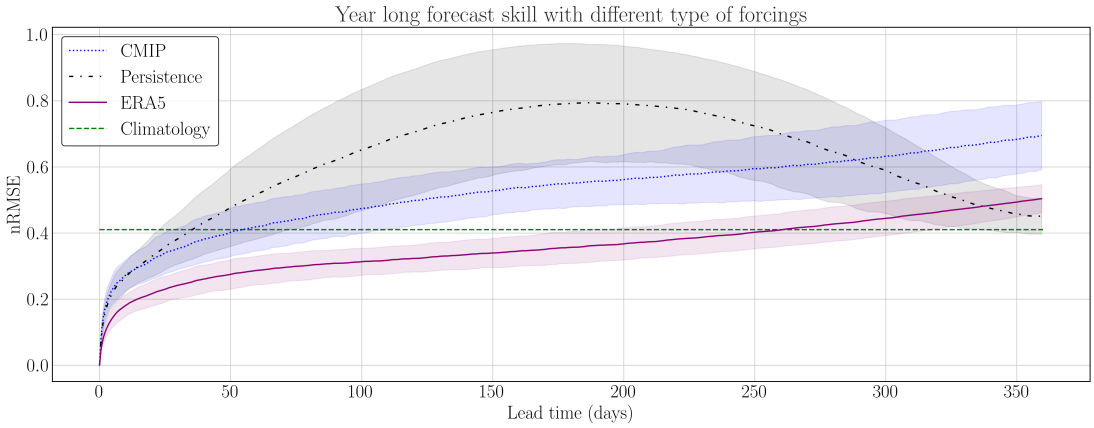

Figure D4. Seasonal forecast of the surrogate model while changing the type of forcings. The neural network remains exactly the same as the trained one. It is thus trained with ERA5 forcings. In the test part, the forcings are changed to CMIP6 forcings (blue line). This forecast skill is compared with neXtSIM and ERA5 forcings (red line), persistence (black line) and daily climatology (green line).

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Code and data availability. The authors will provide access to the data and weights of the neural networks upon request. The source code for the experiments and the neural networks is publicly available under https://github.com/cerea-daml/nextsim-surrogate. The outputs of neXtSIM model will be made available upon requests. Forcings data are publicly available in the Copernicus Data Store https://cds.climate.copernicus.eu

Video supplement. A video of the seasonal forecast for the year 2017 is available at https://doi.org/10.5446/62131

.

Author contributions. EO provided the data and the insigths about neXtSIM. CD, TSF, AF, and MB refined the scientific questions and prepared an analysis strategy. CD performed the experiments. CD, TSF, AF, and MB analyzed and discussed the results. CD wrote the manuscript with TSF, AF, MB, and EO reviewing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge the support of the project SASIP (grant n° 353)funded by Schmidt Futures – a philanthropic initiative that seeks to improve societal outcomes through the development of emerging science and technologies. This work was granted access to the HPC resources of IDRIS under the allocations 2021-AD011013069 and 2022-AD011013069R1 made by GENCI. The authors would like to thank Pierre Rampal, Laurent Bertino, Anton Korosov, Julien Brajard and their colleagues from NERSC for their insightful inputs. CEREA is a member of Institut Pierre-Simon Laplace (IPSL).

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

References

Amodei, M. and Stein, J.: Deterministic and fuzzy verification methods for a hierarchy of numerical models, Meteorological Applications, 16, 191–203, https://doi.org/10.1002/met.101, 2009.

Andersson, T. R., Hosking, J. S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., Law, S., Jones, D. C., Wilkinson, J., Phillips, T., Byrne, J., Tietsche, S., Sarojini, B. B., Blanchard-Wrigglesworth, E., Aksenov, Y., Downie, R., and Shuckburgh, E.: Seasonal Arctic sea ice forecasting with probabilistic deep learning, Nature Communications, 12, https://doi.org/10.1038/s41467-021-25257-4, 2021.

Balan-Sarojini, B., Tietsche, S., Mayer, M., Balmaseda, M., Zuo, H., de Rosnay, P., Stockdale, T., and Vitart, F.: Year-round impact of winter sea ice thickness observations on seasonal forecasts, The Cryosphere, 15, 325–344, https://doi.org/10.5194/tc-15-325-2021, 2021.

Barnes, E. A. and Screen, J. A.: The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?, WIREs Climate Change, 6, 277–286, https://doi.org/10.1002/wcc.337, 2015.

Bernard, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Sommer, J. L., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., and Cuevas, B. D.: Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution, Ocean Dynamics, 56, 543–567, https://doi.org/10.1007/s10236-006-0082-1, 2006.

Beucler, T., Pritchard, M., Rasp, S., Ott, J., Baldi, P., and Gentine, P.: Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems, Physical Review Letters, 126, https://doi.org/10.1103/physrevlett.126.098302, 2021.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q.: Pangu-Weather: A 3D High-Resolution Model for Fast and Accurate Global Weather Forecast, https://doi.org/10.48550/ARXIV.2211.02556, 2022.

Bingham, E.: Fluidity and plasticity, McGraw-Hill, 1922.

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., and DeWeaver, E.: Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations, Journal of Climate, 24, 231–250, https://doi.org/10.1175/2010jcli3775.1, 2011.

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M.: YOLOv4: Optimal Speed and Accuracy of Object Detection, https://doi.org/10.48550/ARXIV.2004.10934, 2020.

Bocquet, M.: Surrogate modelling for the climate sciences dynamics with machine learning and data assimilation, Front. Appl. Math. Stat., 9, https://doi.org/10.3389/fams.2023.1133226, 2023.

Bouchat, A., Hutter, N., Chanut, J., Dupont, F., Dukhovskoy, D., Garric, G., Lee, Y. J., Lemieux, J.-F., Lique, C., Losch, M., Maslowski, W., Myers, P. G., Ólason, E., Rampal, P., Rasmussen, T., Talandier, C., Tremblay, B., and Wang, Q.: Sea Ice Rheology Experiment (SIREx): 1. Scaling and Statistical Properties of Sea-Ice Deformation Fields, Journal of Geophysical Research: Oceans, 127, e2021JC017 667, https://doi.org/10.1029/2021JC017667, 2022.

Boutin, G., Òlason Einar, and et al, R. P.: Arctic sea ice mass balance in a new coupled ice-ocean model using a brittle rheology framework, The Cryosphere, https://doi.org/https://doi.org/10.5194/tc-2022-142, 2022.

Cheng, S., Quilodran-Casas, C., Ouala, S., Farchi, A., Liu, C., Tandeo, P., Fablet, R., Lucor, D., Iooss, B., Brajard, J., Xiao, D., Janjic, T., Ding, W., Guo, Y., Carrassi, A., Bocquet, M., and Arcucci, R.: Machine learning with data assimilation and uncertainty quantification for dynamical systems: a review, IEEE/CAA Journal of Automatica Sinica, –, –, https://doi.org/–, in press, 2023.

Comiso, J. C.: Abrupt decline in the Arctic winter sea ice cover, Geophysical Research Letters, 33, n/a–n/a, https://doi.org/10.1029/2006gl027341, 2006.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Dansereau, V., Weiss, J., Saramito, P., and Lattes, P.: A Maxwell Elasto-Brittle Rheology for Sea Ice Modelling, The Cryosphere, 10, 1339–1359, https://doi.org/10.5194/tc-10-1339-2016, 2016.

et al, T.: $CREG025.L75-NEMO_r3.6.0$, https://doi.org/10.5281/ZENODO.5802028, 2021.

Farchi, A., Bocquet, M., Roustan, Y., Mathieu, A., and Quérel, A.: Using the Wasserstein distance to compare fields of pollutants: application to the radionuclide atmospheric dispersion of the Fukushima-Daiichi accident, Tellus B, 68, 31 682, https://doi.org/10.3402/tellusb.v68.31682, 2016.

Finn, T. S., Durand, C., Farchi, A., Bocquet, M., Chen, Y., Carrassi, A., and Dansereau, V.: Deep Learning of Subgrid-Scale Parametrisations for Short-Term Forecasting of Sea-Ice Dynamics with a Maxwell-Elasto-Brittle Rheology, EGUsphere, pp. 1–39, https://doi.org/10.5194/egusphere-2022-1342, 2023.

Girard, L., Bouillon, S., Weiss, J., Amitrano, D., Fichefet, T., and Legat, V.: A new modeling framework for sea-ice mechanics based on elasto-brittle rheology, Annals of Glaciology, 52, 123–132, https://doi.org/10.3189/172756411795931499, 2011.

Grigoryev, T., Verezemskaya, P., Krinitskiy, M., Anikin, N., Gavrikov, A., Trofimov, I., Balabin, N., Shpilman, A., Eremchenko, A., Gulev, S., Burnaev, E., and Vanovskiy, V.: Data-Driven Short-Term Daily Operational Sea Ice Regional Forecasting, Remote Sensing, 14, 5837, https://doi.org/10.3390/rs14225837, 2022.

Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., Fučkar, N. S., Germe, A., Hawkins, E., Keeley, S., Koenigk, T., y Mélia, D. S., and Tietsche, S.: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales, Quarterly Journal of the Royal Meteorological Society, 142, 546–561, https://doi.org/10.1002/qj.2401, 2014.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hess, P., Lange, S., and Boers, N.: Deep Learning for bias-correcting comprehensive high-resolution Earth system models, https://doi.org/10.48550/ARXIV.2301.01253, 2023.

Horvat, C. and Roach, L. A.: WIFF1.0: A Hybrid Machine-Learning-Based Parameterization of Wave-Induced Sea Ice Floe Fracture, Geoscientific Model Development, 15, 803–814, https://doi.org/10.5194/gmd-15-803-2022, 2022.

Hunke, E., Lipscomb, W., Jones, P., Turner, A., Jeffery, N., and Elliott, S.: CICE, the Los Alamos sea ice model, Tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2017.

IPCC: Polar Regions, in: The Ocean and Cryosphere in a Changing Climate, pp. 203–320, Cambridge University Press, https://doi.org/10.1017/9781009157964.005, 2022.

Jung, T., Kasper, M. A., Semmler, T., and Serrar, S.: Arctic influence on subseasonal midlatitude prediction, Geophysical Research Letters, 41, 3676–3680, https://doi.org/10.1002/2014gl059961, 2014.

Kacimi, S. and Kwok, R.: Arctic Snow Depth, Ice Thickness, and Volume From ICESat-2 and CryoSat-2: 2018–2021, Geophysical Research Letters, 49, https://doi.org/10.1029/2021gl097448, 2022.

Kadow, C., Hall, D. M., and Ulbrich, U.: Artificial intelligence reconstructs missing climate information, Nature Geoscience, 13, 408–413, https://doi.org/10.1038/s41561-020-0582-5, 2020.

Keisler, R.: Forecasting Global Weather with Graph Neural Networks, https://doi.org/10.48550/ARXIV.2202.07575, 2022.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Køltzow, M., Schyberg, H., Støylen, E., and Yang, X.: Value of the Copernicus Arctic Regional Reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic, Polar Research, 41, https://doi.org/10.33265/polar.v41.8002, 2022.

Kwok, R.: Contrasts in sea ice deformation and production in the Arctic seasonal and perennial ice zones, Journal of Geophysical Research, 111, https://doi.org/10.1029/2005jc003246, 2006.

Kwok, R.: Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018), Environmental Research Letters, 13, 105 005, https://doi.org/10.1088/1748-9326/aae3ec, 2018.

Kwok, R., Spreen, G., and Pang, S.: Arctic sea ice circulation and drift speed: Decadal trends and ocean currents, Journal of Geophysical Research: Oceans, 118, 2408–2425, https://doi.org/10.1002/jgrc.20191, 2013.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., Ravuri, S., Ewalds, T., Alet, F., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Stott, J., Vinyals, O., Mohamed, S., and Battaglia, P.: GraphCast: Learning skillful medium-range global weather forecasting, https://doi.org/10.48550/ARXIV.2212.12794, 2022.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophysical Research Letters, 40, 732–737, https://doi.org/10.1002/grl.50193, 2013.

Lemke, P., Trinkl, E. W., and Hasselmann, K.: Stochastic Dynamic Analysis of Polar Sea Ice Variability, Journal of Physical Oceanography, 10, 2100–2120, https://doi.org/10.1175/1520-0485(1980)010<2100:sdaops>2.0.co;2, 1980.

Liu, G., Reda, F. A., Shih, K. J., Wang, T., Tao, A., and Catanzaro, B.: Image Inpainting for Irregular Holes Using Partial Convolutions, CoRR, abs/1804.07723, http://arxiv.org/abs/1804.07723, 2018.

Liu, Y., Bogaardt, L., Attema, J. J., and Hazeleger, W.: Extended Range Arctic Sea Ice Forecast with Convolutional Long-Short Term Memory Networks, Monthly Weather Review, 2020.

Loshchilov, I. and Hutter, F.: Fixing Weight Decay Regularization in Adam, CoRR, abs/1711.05101, http://arxiv.org/abs/1711.05101, 2017.

Lovejoy, S., Tarquis, A. M., Gaonac'h, H., and Schertzer, D.: Single- and Multiscale Remote Sensing Techniques, Multifractals, and MODIS-Derived Vegetation and Soil Moisture, Vadose Zone Journal, 7, 533–546, https://doi.org/10.2136/vzj2007.0173, 2008.

Madec, G., Delecluse, P., Imbard, M., and Levy, C.: OPA 8 Ocean General Circulation Model - Reference Manual, Tech. rep., LODYC/IPSL Note 11, 1998.

Misra, D.: Mish: A Self Regularized Non-Monotonic Neural Activation Function, CoRR, abs/1908.08681, http://arxiv.org/abs/1908.08681, 2019.

Neuhauser, M., Verrier, S., and Mangiarotti, S.: Multifractal analysis for spatial characterization of high resolution Sentinel-2/MAJA products in Southwestern France, Remote Sensing of Environment, 270, 112 859, https://doi.org/10.1016/j.rse.2021.112859, 2022.

Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K., and Grover, A.: ClimaX: A foundation model for weather and climate, https://doi.org/10.48550/ARXIV.2301.10343, 2023.

Nielsen-Englyst, P., Høyer, J. L., Madsen, K. S., Tonboe, R. T., Dybkjær, G., and Skarpalezos, S.: Deriving Arctic 2 m air temperatures over snow and ice from satellite surface temperature measurements, The Cryosphere, 15, 3035–3057, https://doi.org/10.5194/tc-15-3035-2021, 2021.

Ólason, E., Boutin, G., Korosov, A., Rampal, P., Williams, T., Kimmritz, M., Dansereau, V., and Samaké, A.: A New Brittle Rheology and Numerical Framework for Large-Scale Sea-Ice Models, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021ms002685, 2022.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

Olonscheck, D., Mauritsen, T., and Notz, D.: Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations, Nature Geoscience, 12, 430–434, https://doi.org/10.1038/s41561-019-0363-1, 2019.

Pinckaers, H., van Ginneken, B., and Litjens, G.: Streaming Convolutional Neural Networks for End-to-End Learning with Multi-Megapixel Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44, 1581–1590, https://doi.org/10.1109/TPAMI.2020.3019563, 2022.

Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: A New Lagrangian Sea Ice Model, The Cryosphere, 10, 1055–1073, https://doi.org/10.5194/tc-10-1055-2016, 2016.

Rampal, P., Dansereau, V., Olason, E., Bouillon, S., Williams, T., Korosov, A., and Samaké, A.: On the multi-fractal scaling properties of sea ice deformation, The Cryosphere, 13, 2457–2474, https://doi.org/10.5194/tc-13-2457-2019, 2019.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., and Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672–677, https://doi.org/10.1038/s41586-021-03854-z, 2021.

Renner, A. H. H., Gerland, S., Haas, C., Spreen, G., Beckers, J. F., Hansen, E., Nicolaus, M., and Goodwin, H.: Evidence of Arctic sea ice thinning from direct observations, Geophysical Research Letters, 41, 5029–5036, https://doi.org/10.1002/2014gl060369, 2014.

Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., and Keeley, S.: ECMWF ECMWF-IFS-HR model output prepared for CMIP6 HighResMIP control-1950, https://doi.org/10.22033/ESGF/CMIP6.4945, 2017.

Rolph, R. J., Feltham, D. L., and Schröder, D.: Changes of the Arctic marginal ice zone during the satellite era, The Cryosphere, 14, 1971–1984, https://doi.org/10.5194/tc-14-1971-2020, 2020.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, CoRR, abs/1505.04597, http://arxiv.org/abs/1505.04597, 2015.

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., and Vivier, F.: The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities, Geoscientific Model Development, 8, 2991–3005, https://doi.org/10.5194/gmd-8-2991-2015, 2015.

Schweiger, A. J., Wood, K. R., and Zhang, J.: Arctic Sea Ice Volume Variability over 1901–2010: A Model-Based Reconstruction, Journal of Climate, 32, 4731–4752, https://doi.org/10.1175/jcli-d-19-0008.1, 2019.

Screen, J.: Influence of Arctic sea ice on European summer precipitation, Environmental Research Letters, 8, 4015–, https://doi.org/10.1088/1748-9326/8/4/044015, 2013.

Serreze, M. C., Maslanik, J. A., Barry, R. G., and Demaria, T. L.: Winter atmospheric circulation in the Arctic Basin and possible relationships to the great salinity anomaly in the northern North Atlantic, Geophysical Research Letters, 19, 293–296, https://doi.org/10.1029/91gl02946, 1992.

Sievers, I., Gierisch, A. M. U., Rasmussen, T. A. S., Hordoir, R., and Stenseng, L.: Arctic sea ice and snow from different ice models: A CICE–SI3 intercomparison study, The Cryosphere, https://doi.org/10.5194/tc-2022-84, 2022.

Sigmond, M., Fyfe, J. C., Flato, G. M., Kharin, V. V., and Merryfield, W. J.: Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system, Geophysical Research Letters, 40, 529–534, https://doi.org/10.1002/grl.50129, 2013.

Strong, C.: Atmospheric influence on Arctic marginal ice zone position and width in the Atlantic sector, February–April 1979–2010, Climate Dynamics, 39, 3091–3102, https://doi.org/10.1007/s00382-012-1356-6, 2012.

Vanderbecken, P. J., Dumont Le Brazidec, J., Farchi, A., Bocquet, M., Roustan, Y., Potier, E., and Broquet, G.: Accounting for meteorological biases in simulated plumes using smarter metrics, Atmos. Meas. Tech., 16, 1745–1766, https://doi.org/10.5194/amt-16-1745-2023, 2023.

https://doi.org/10.5194/egusphere-2023-1384
Preprint. Discussion started: 23 August 2023
© Author(s) 2023. CC BY 4.0 License.

von Albedyll, L., Hendricks, S., Grodofzig, R., Krumpen, T., Arndt, S., Belter, H. J., Birnbaum, G., Cheng, B., Hoppmann, M., Hutchings, J., Itkin, P., Lei, R., Nicolaus, M., Ricker, R., Rohde, J., Suhrhoff, M., Timofeeva, A., Watkins, D., Webster, M., and Haas, C.: Thermodynamic and dynamic contributions to seasonal Arctic sea ice thickness distributions from airborne observations, Elementa: Science of the Anthropocene, 10, https://doi.org/10.1525/elementa.2021.00074, 2022.

Wang, C., Graham, R. M., Wang, K., Gerland, S., and Granskog, M. A.: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution, The Cryosphere, 13, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019, 2019.

Xiu, Y., Luo, H., Yang, Q., Tietsche, S., Day, J., and Chen, D.: The Challenge of Arctic Sea Ice Thickness Prediction by ECMWF on Subseasonal Time Scales, Geophysical Research Letters, 49, https://doi.org/10.1029/2021gl097476, 2022.

Xu, T., Moore, I. D., and Gallant, J. C.: Fractals, fractal dimensions and landscapes — a review, Geomorphology, 8, 245–262, https://doi.org/10.1016/0169-555x(93)90022-t, 1993.

Yu, Y., Xiao, W., Zhang, Z., Cheng, X., Hui, F., and Zhao, J.: Evaluation of 2-m Air Temperature and Surface Temperature from ERA5 and ERA-I Using Buoy Observations in the Arctic during 2010–2020, Remote Sensing, 13, 2813, https://doi.org/10.3390/rs13142813, 2021.

Zampieri, L., Goessling, H. F., and Jung, T.: Bright Prospects for Arctic Sea Ice Prediction on Subseasonal Time Scales, Geophysical Research Letters, 45, 9731–9738, https://doi.org/10.1029/2018gl079394, 2018.

Zhang, Z.-H., Yang, Z., Sun, Y., Wu, Y.-F., and Xing, Y.-D.: Lenet-5 Convolution Neural Network with Mish Activation Function and Fixed Memory Step Gradient Descent Method, in: 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, IEEE, https://doi.org/10.1109/iccwamtip47768.2019.9067661, 2019.