Interpretable Spatiotemporal Forecasting of Arctic Sea Ice Concentration at Seasonal Lead Times

Matthew Beveridge
Cambridge, MA
mattjbev@gmail.com

Lucas Pereira
ITI/LARSyS, Técnico Lisboa
Portugal
lucas.pereira@iti.larsys.pt

Abstract

There are many benefits from the accurate forecasting of Arctic sea ice, however existing models struggle to reliably predict sea ice concentration at long lead times. Many numerical models exist but can be sensitive to initial conditions, and while recent deep learning-based methods improve overall robustness, they either do not utilize temporal trends or rely on architectures that are not performant at learning long-term sequential dependencies. We propose a method of forecasting sea ice concentration using neural circuit policies, a form of continuous time recurrent neural architecture, which improve the learning of long-term sequential dependencies compared to existing techniques and offer the added benefits of adaptability to irregular sequence intervals and high interpretability.

1 Introduction

Polar ice caps play an important role in maintaining a balanced global climate. Perhaps most importantly, the sea ice serves as insulation to regulate temperature, moisture, and solar radiation on the Earth's surface and in the atmosphere, [20, 33]. Sea ice cover in the Arctic has declined rapidly as a result of anthropogenic warming over the preceding century [26, 28], which has been noted as a leading cause of near-surface temperature amplification in the Arctic [19, 13, 27, 15, 25] and changes in mid-latitude meteorology well beyond the boundaries of the Arctic circle [7, 31]. The melting trend is predicted to continue, with the current decadal rate of decline at 13.16% in September and 2.67% in March [30]. Effectively forecasting Arctic ice cover will prove crucial for communities residing within the region who are most susceptible to food insecurity induced by algal blooms in the absence of ice that are toxic to their food supply [1]. Further, accurate Arctic sea ice prediction may correlate to improved navigation within the region and weather estimates elsewhere in the world [14], although the latter is still in dispute [3].

The most common methods for predicting Arctic sea ice concentration (SIC) are dynamical models which tend to couple ice, ocean, and atmospheric data into a deterministic prediction, however, they are highly susceptible to noise in the initial conditions [22] and rarely outperform statistical models at lead times beyond two months [1]. Deep generative models, a form of statistical model, have also had success in the tangential task of precipitation nowcasting due to their probabilistic nature and consequent ability to mimic ensemble methods [23]. The recent wide availability of remote sensing data has allowed the development of more robust, deep learning-based models to predict SIC as well as other sea ice characteristics. The seminal learned models for this task examine spatial relationships in ice concentration using an ensemble of convolutional networks that improve predictions of sea ice extent (SIE), derived from SIC, at lead times up to six months compared to baseline persistence and climatological models [1]. Naturally, subsequent works utilize the time domain and tackle the forecasting task using spatiotemporal sequence models in the form of convolutional long-short term memory networks (ConvLSTMs) [30, 2, 18]. Such ConvLSTM models for SIC are typically evaluated at subseasonal lead times and outperform baseline statistical and dynamical methods under

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2022.

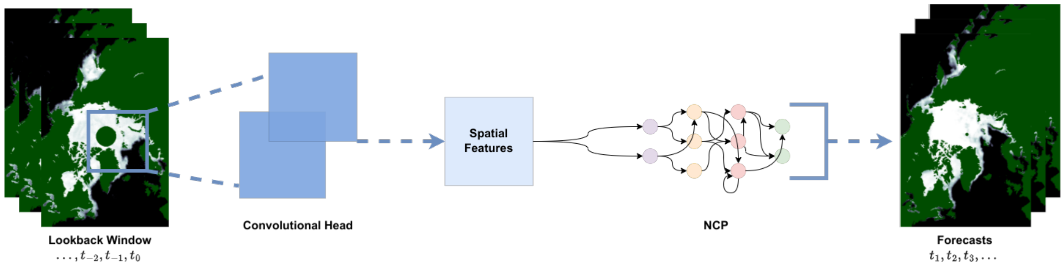

Figure 1: Sequential forecasting pipeline. Given a sequence of input SIC observations for a region, the model first uses a series of convolutional layers to embed spatial features before being passed through the recurrent NCP model. The output is a series of predictions at future time steps.

two months, but degrade beyond that. This limited performance beyond a few months can be attributed to both the chaotic systems that drive Arctic SIC and the vanishing or exploding gradients that occur while training recurrent neural networks (RNNs) which limit the predictability of longer sequences. In this work, we propose a highly expressive spatiotemporal recurrent network architecture fit to the task of forecasting Arctic SIC. The resultant model structure improves interpretability insights into which regions of the Arctic are most influential in overall SIC, and can also be supplied with mixed memory architectures to mitigate the gradient issues of existing recurrent methods [16].

2 Learning to predict sea ice concentration

2.1 Neural architecture

Neural circuit policies (NCPs) have excelled as end-to-end (perception-to-actuation) controllers in the field of autonomous robotics [17]. They are a sparse but intentional wiring of liquid-time constant (LTC) recurrent network cells [11, 10], a form of continuous time recurrent model [21, 24], made up of four distinct layers of neurons whose names allude to their robotic origin: sensory, inter-neuron, command, and motor. When predicting Arctic SIC, sensory neurons extract contextual (spatial) relationships in the input features, inter and command neurons make a decision, and motor neurons execute the goal task of representing the ice concentration level. The relationships within this neuron structure are characterized by a semi-implicit ordinary differential equation (ODE) solver which yields a stable solution of the system, the result of which is folded into an RNN and trained via supervised learning [17].

LTC networks exhibit superior performance in modeling sequential and irregularly sampled data by combining a static network's depth dimension with an RNN's time dimension into a continuous vector field, allowing for parameter sharing and function approximation that is not possible with discrete neural networks [12]. Additionally, the inner neural ODEs have been shown to be highly expressible [6], which translates to the parent LTCs and NCPs and in turn enhances interpretability methods applied to these models.

The proposed architecture consists of a convolutional head, NCP backbone, and deconvolutional upsampling to produce output sea ice concentration maps as shown in Figure 1. Given a sequence of regularly or irregularly sampled observations composing the lookback window, the model makes predictions of future Arctic SIC. We henceforth refer to this model as ConvNCP.

2.2 Data sources

Since 1978, sea ice concentration readings have been collected over the poles by the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR), the Defense Meteorological Satellite Program (DMSP) -F8, -F11 and -F13 Special Sensor Microwave/Imagers (SSM/Is), and the DMSP-F17 Special Sensor Microwave Imager/Sounder (SSMIS) [5]. These data are collected and presented by the National Snow and Ice Data Center (NSIDC) in polar stereographic grids with resolution 25km $\times$ 25km at a daily frequency[1]. The NSIDC SIC dataset is the primary data source for training and evaluation of forecasting skill. Following [30], we discard data points prior to 1979 due to high

[1] `https://nsidc.org/data/nsidc-0051/versions/2`

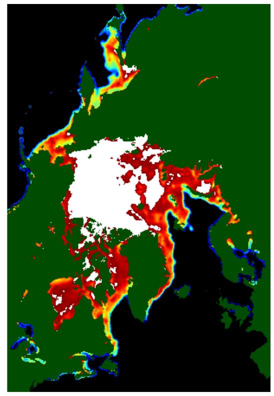

Figure 2: Artistic rendering of saliency map output from the learned NCP model over the Arctic. Heatmap areas represent the relative attention of the network used to identify which regions of the Arctic give the most insight into future SIC, green represents land mass, and white represents sea ice.

uncertainties and focus on the Arctic region within the area 31-90°N, 180°W-180°E. At the time of writing, the most recent data readings extend through May 31st, 2022.

Some prior works [1] also utilize the Coupled Model Intercomparison Project phase 6 (CMIP6) [8] to generate supplemental simulated data for the training of learned models. While this can significantly increase data volumes, it introduces the biases of the simulation during training that then need to be mitigated. We use NSIDC observations as our primary data source and reserve CMIP6 simulations as an additional option as needed. Non-sea ice data, such as ground permafrost levels [4], could also bolster model training.

3 Conclusion

Forecasting Arctic sea ice concentration has been a notoriously difficult task due to the inherently low predictibility of SIC [9] and fragile initial condition requirements. With the proposed ConvNCP architecture we aim to improve the following compared to existing statistical, dynamical, and learned models: (1) improved performance at long lead times, (2) increased robustness to noisy observations, and (3) novel interpretability capabilities for identifying the key drivers of fluctuations in Arctic SIC, as demostrated in Figure 2. For greater usability, a natural follow-up is to quantify uncertainty in ConvNCP predictions. Future extensions of this work include (1) investigating transformer architectures [29] which also solve the gradient issues of RNNs but may not be well suited for time series tasks due to permutation invariance [34, 32], and (2) the extension of the ConvNCP model to related forecasting tasks such as sea ice extent, sea ice thickness, and precipitation.

3.1 Pathway to climate impact

The developments of this work are most impactful to climate scientists as, to the best of our knowledge, no prior spatiotemporal climate forecasting methods offer the advanced explainability of our ConvNCP. Downstream methods can be developed to best utilize these outputs and better model climate elsewhere, e.g., mid-parallel meteorology. In combination with the SIC forecasts, this is a dual benefit to parties who wish to use the outputs of ConvNCP to enact change; since it is claimed that sea ice determines both weather and sea levels elsewhere in the world, improved ice predictions can be utilized by governments, NGOs, and nonprofits to legislate environmental policy while saliency maps can be used in part to identify scientific areas of interest. Lastly, with the growing abundance of environmental data available, machine learning will find its way into evermore applications in climate science. There has been a rapid rise in interdisciplinary research at the intersection of machine learning and climate, and future initiatives such as the EU's Copernicus initiative[2] will provide vast datasets to support these collaborations. ConvNCP for Arctic sea ice prediction can serve as a stepping stone to modern, useful learning-based methods for understanding remote sensing data.

[2] `https://www.copernicus.eu/en`

Acknowledgments and Disclosure of Funding

Lucas Pereira has received funding from the Portuguese Foundation for Science and Technology (FCT) under grants CEECIND/01179/2017 and UIDB/50009/2020. We thank Ryan Shubert for helpful discussions.

References

[1] Tom R Andersson et al. "Seasonal Arctic sea ice forecasting with probabilistic deep learning". In: *Nature communications* 12.1 (2021), pp. 1–12.

[2] Nazanin Asadi et al. "Probabilistic Gridded Seasonal Sea Ice Presence Forecasting using Sequence to Sequence Learning". In: *The Cryosphere Discussions* (2021), pp. 1–17.

[3] Elizabeth A Barnes and James A Screen. "The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?" In: *Wiley Interdisciplinary Reviews: Climate Change* 6.3 (2015), pp. 277–286.

[4] J. Brown et al. *Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2*. 2002.

[5] D. J. Cavalieri et al. "Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 2". In: *NASA National Snow and Ice Data Center Distributed Active Archive Center* (2022).

[6] Ricky TQ Chen et al. "Neural ordinary differential equations". In: *Advances in neural information processing systems* 31 (2018).

[7] Clara Deser et al. "The seasonal atmospheric response to projected Arctic sea ice loss in the late twenty-first century". In: *Journal of Climate* 23.2 (2010), pp. 333–351.

[8] Veronika Eyring et al. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: *Geoscientific Model Development* 9.5 (2016), pp. 1937–1958.

[9] Virginie Guemas et al. "A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales". In: *Quarterly Journal of the Royal Meteorological Society* 142.695 (2016), pp. 546–561.

[10] Ramin Hasani et al. "Closed-form continuous-depth models". In: *arXiv preprint arXiv:2106.13898* (2021).

[11] Ramin Hasani et al. "Liquid time-constant networks". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 9. 2021, pp. 7657–7666.

[12] Ramin Hasani et al. "Closed-form continuous-time neural networks". In: *Nature Machine Intelligence* (2022), pp. 1–12.

[13] WJ Ingram, CA Wilson, and JFB Mitchell. "Modeling climate change: An assessment of sea ice and surface albedo feedbacks". In: *Journal of Geophysical Research: Atmospheres* 94.D6 (1989), pp. 8609–8622.

[14] Thomas Jung et al. "Arctic influence on subseasonal midlatitude prediction". In: *Geophysical Research Letters* 41.10 (2014), pp. 3676–3680.

[15] Arun Kumar et al. "Contribution of sea ice loss to Arctic amplification". In: *Geophysical Research Letters* 37.21 (2010).

[16] Mathias Lechner and Ramin Hasani. "Learning long-term dependencies in irregularly-sampled time series". In: *arXiv preprint arXiv:2006.04418* (2020).

[17] Mathias Lechner et al. "Neural circuit policies enabling auditable autonomy". In: *Nature Machine Intelligence* 2.10 (2020), pp. 642–652.

[18] Yang Liu et al. "Extended-range arctic sea ice forecast with convolutional long short-Term memory networks". In: *Monthly Weather Review* 149.6 (2021), pp. 1673–1693.

[19] Syukuro Manabe and Ronald J Stouffer. "Sensitivity of a global climate model to an increase of CO2 concentration in the atmosphere". In: *Journal of Geophysical Research: Oceans* 85.C10 (1980), pp. 5529–5554.

[20] Gary A Maykut. "Large-scale heat exchange and ice production in the central Arctic". In: *Journal of Geophysical Research: Oceans* 87.C10 (1982), pp. 7971–7984.

[21] James Morrill et al. "Neural rough differential equations for long time series". In: *International Conference on Machine Learning*. PMLR. 2021, pp. 7829–7838.

[22] Rym Msadek et al. "Importance of initial conditions in seasonal predictions of Arctic sea ice extent". In: *Geophysical Research Letters* 41.14 (2014), pp. 5208–5215.

[23] Suman Ravuri et al. "Skilful precipitation nowcasting using deep generative models of radar". In: *Nature* 597.7878 (2021), pp. 672–677.

[24] Mona Schirmer et al. "Modeling irregular time series with continuous recurrent units". In: *International Conference on Machine Learning*. PMLR. 2022, pp. 19388–19405.

[25] James A Screen and Ian Simmonds. "The central role of diminishing sea ice in recent Arctic temperature amplification". In: *Nature* 464.7293 (2010), pp. 1334–1337.

[26] Mark C Serreze, Marika M Holland, and Julienne Stroeve. “Perspectives on the Arctic’s shrinking sea-ice cover”. In: *science* 315.5818 (2007), pp. 1533–1536.

[27] MC Serreze et al. “The emergence of surface-based Arctic amplification”. In: *The Cryosphere* 3.1 (2009), pp. 11–19.

[28] Julienne C Stroeve et al. “The Arctic’s rapidly shrinking sea ice cover: a research synthesis”. In: *Climatic change* 110.3 (2012), pp. 1005–1027.

[29] Ashish Vaswani et al. “Attention is all you need”. In: *Advances in neural information processing systems* 30 (2017).

[30] Jianfen Wei, Renlong Hang, and Jing-Jia Luo. “Prediction of Pan-Arctic Sea Ice Using Attention-Based LSTM Neural Networks”. In: *Frontiers in Marine Science* (), p. 918.

[31] Shuting Yang and Jens H Christensen. “Arctic sea ice reduction and European cold winters in CMIP5 climate change experiments”. In: *Geophysical Research Letters* 39.20 (2012).

[32] Ailing Zeng et al. “Are Transformers Effective for Time Series Forecasting?” In: *arXiv preprint arXiv:2205.13504* (2022).

[33] Jinlun Zhang et al. “Ensemble 1-year predictions of Arctic sea ice for the spring and summer of 2008”. In: *Geophysical Research Letters* 35.8 (2008).

[34] Haoyi Zhou et al. “Informer: Beyond efficient transformer for long sequence time-series forecasting”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 12. 2021, pp. 11106–11115.