

In this study, the authors propose a data-driven model for short-term sea ice concentration (SIC) prediction based on a U-Net neural network architecture, termed SICUnet. The model extends a baseline U-Net by incorporating spatial-channel attention mechanisms. It is trained exclusively on satellite observations and uses daily SIC maps from seven consecutive days to predict SIC over the subsequent seven days. Longer forecast horizons are obtained through recursive application of the model, and results are presented up to a 35-day lead time (five recursive cycles). Performance evaluation over four test years indicates improved skill relative to other neural network architectures and to a reference numerical model.

The manuscript is generally clear and focuses primarily on the performance gains achieved through the proposed architectural modifications. However, the comparison with alternative approaches is somewhat limited, and the manuscript would benefit from a more critical discussion of the results, particularly with respect to non-conventional methodological choices (e.g., network architecture and loss function).

Below I outline several points that should be considered in revising the manuscript.

Dear Reviewer #1:

Thanks for your suggestions and comments. We have carefully revised the manuscript according to the comments. Your opinions are reasonable, greatly helping me improve my article. We respond to the review comments point by point. The response to the reviewer's comments is as follows:

More details should be provided on the resampling of TOPAZ5 data, as mismatches in spatial resolution between the model fields and the observations may introduce representation errors in the comparison.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

First, we agree with the reviewer's concern that inconsistencies in spatial resolution and projection grids may introduce representation errors. In this study, such errors arising from differences in projection methods and spatial resolution between numerical model outputs and satellite observations are indeed unavoidable. To mitigate these effects, we aimed to minimize the smoothing introduced during resampling. Accordingly, we followed the approach of *Palmer et al. (2024)*. In that study, nearest-neighbor interpolation was used to resample SIC

and related variables onto the TOPAZ4 projection grid. On this basis, we applied nearest-neighbor interpolation to resample the TOPAZ5 data using the Python tool `scipy.interpolate.griddata`.

Nevertheless, as highlighted by the reviewer, resampling TOPAZ5 data onto the NSIDC polar stereographic grid may introduce additional errors. These errors may affect the numerical model predictions used for comparison. In response to this issue, we additionally adopted the OSI-SAF SIC product as the reference benchmark. Predictions from both the numerical model and the deep learning model were generated on the OSI-SAF grid. This unified spatial framework improves spatial consistency and comparability across different predictive methods.

Therefore, to more rigorously and comprehensively assess the predictive performance of SICUNet, we designed a set of comparative experiments in Section 4.6 of the revised manuscript. In this experiment, SIC from OSI-450-a1 and OSI-430-a were used as inputs and outputs for training and prediction with SICUNet. The predictive performance of SICUNet was then compared with that of TOPAZ5 using OSI-SAF data as the observational reference. The experimental design and its results are described in detail in Section 4.6 of the revised manuscript.

The authors should discuss how the introduction of NIIEE in the loss function contributes to improved predictive skill.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

Following the reviewer's suggestion, we have added a discussion in the revised manuscript on the role of NIIEE in the loss function for enhancing model predictive performance. The specific content is as follows:

4.4 Contribution of NIIEE to SICUNet Predictive Performance

This study further analyzes the role of NIIEE in the loss function. It also evaluates its contribution to predictive performance through loss function ablation experiments. The experiments use the full loss function ($\text{MAE} + 0.01 \times \text{NIIEE}$) as the baseline. In the control experiment, only MAE is used to evaluate the impact of removing NIIEE on predictive performance.

The experiments show that incorporating NIIEE enhances the predictive capability of SICUNet at the sea ice margins during recursive forecasting. For short forecast horizons (1–21 days), the improvement in BACC is relatively limited. However, the positive effect of NIIEE on prediction becomes increasingly evident as recursive forecast steps increase. Over extended forecast horizons, BACC exhibits a modest increase. These results suggest that NIIEE primarily contributes to maintaining stability and spatial consistency in recursive forecasts. However, its impact on short-term predictive accuracy is relatively limited. This finding is consistent with previous work (Ren and Li, 2023). That study indicated that NIIEE provides limited improvement for 1–45 day forecasts, but its effect becomes more pronounced for 45–90 day predictions. Overall, these results indicate that incorporating the NIIEE loss function leads to more stable and spatially consistent recursive forecasts than using the MAE alone. The advantages of this approach become increasingly apparent as the recursive forecast steps lengthen.

Table 4: Overall prediction performance of SICUNet under different recursive prediction scenarios using only the MAE loss function (Unit: %)

Recursive Step	Output Lengh	MAE	BACC	NSE
	(days)			
0	7	1.22	97.28	98.11
1	14	1.51	96.62	97.22
2	21	1.73	96.08	96.42
3	28	1.88	95.71	95.81
4	35	2.10	95.21	94.96

NIIEE can quantify the differences between predicted and observed SIC in terms of spatial distribution. Unlike the IIEE that considers only SIE differences (Goessling et al., 2016), NIIEE accounts for the continuous SIC information instead of binarizing SIC. This allows for more comprehensive representation of sea ice spatial distribution. It also imposes explicit constraints on ice edges and the overall spatial patterns. Based on this design, SICUNet simultaneously

minimizes numerical errors and effectively constrains deviations in spatial structure during training. This enhances overall fitting consistency and stability in SIC prediction.

Additional details should be given regarding the alternative neural network architectures used for comparison, including whether their hyperparameters were optimized.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We apologize for the unclear statements in the manuscript and have added the relevant details in Section 4.5 of the revised manuscript:

In this study, ablation experiments were designed to evaluate the roles of residual connections and attention mechanisms in SICUNet. These experiments also aim to quantify their contributions to model performance. Specifically, residual connections and attention modules were independently removed from SICUNet. In the resulting variants, the residual module was removed for Residual U-Net, and the attention module was removed for Attention U-Net. The resulting changes in model performance were then analyzed.

To further validate the advantages of the model, this study compares SICUNet with U²Net. U²Net is a convolutional neural network derived from the U-Net architecture. Its core mechanism gradually deepens the network through a multi-layer nested U-block structure to extract multi-scale features. The network consists of five modules: input, encoder, decoder, connection and output (Qin et al., 2020), as shown in Fig. 9. Initially, the encoder captures spatiotemporal relationships from sequences of historical SIC data. It then produces downscaled spatiotemporal feature maps. Then, the connection module merges outputs from the encoder and decoder. Subsequently, the decoder progressively restores the spatial scale of the downscaled feature maps. Each decoder generates predictions at the original pixel level. Finally, the outputs from all decoders are concatenated along the channel dimension and convolved to generate the final prediction of Arctic SIC. In terms of network design, encoders 1–3 and decoders 1–3 follow a symmetric encoder–decoder structure. Convolutions are used to extract features. Max-pooling is used for downsampling. Bilinear interpolation is then employed to restore spatial resolution. Encoders 4 and 5 and decoder 4 incorporate dilated convolutions. This further expands the receptive field and allows the network to better capture

multi-scale spatiotemporal features. The input, output and intermediate channel numbers of the encoder and decoder are listed in Table 5.

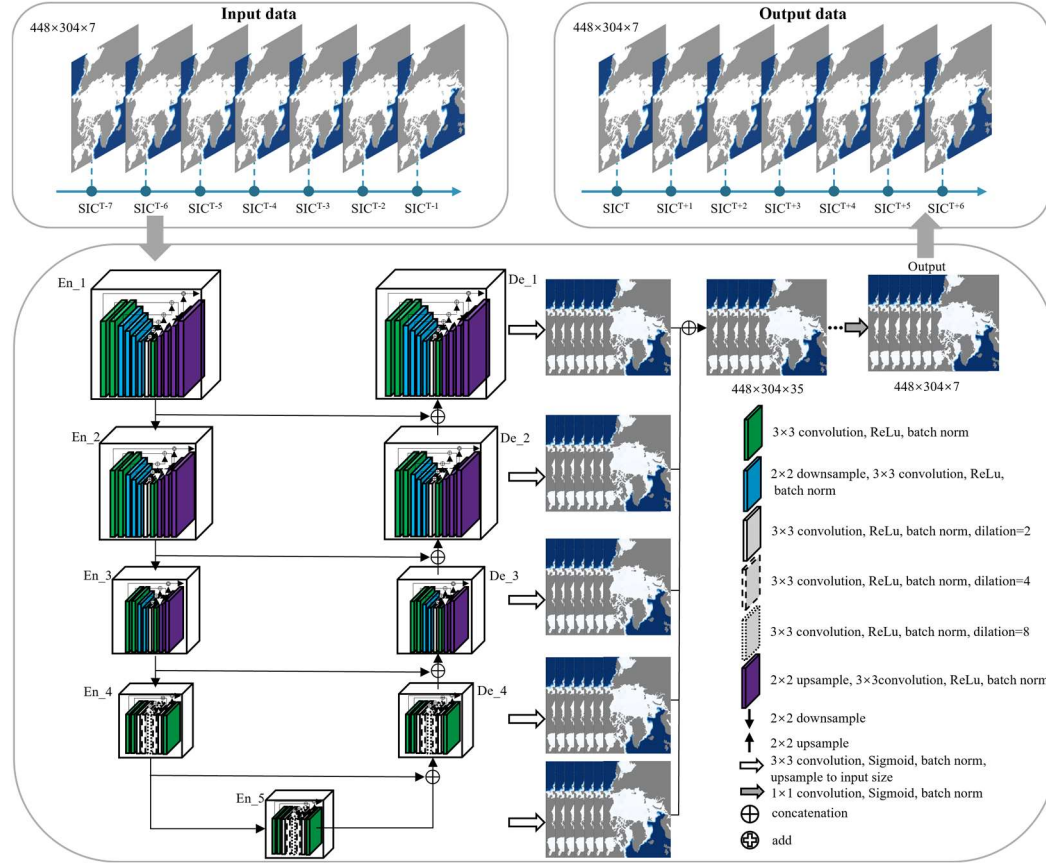


Figure 9 U²Net neural network structure and input–output data.

Table 5 Channel numbers of U²Net encoders and decoders

	encoder1	encoder2	encoder3	encoder4	encoder5	decoder4	decoder3	decoder2	decoder1
input channels	7	64	128	256	256	512	512	256	128
intermediate channels	32	32	64	128	128	128	64	32	16
output channels	64	128	256	256	256	256	128	64	64

To ensure comparability, the deep learning models used for comparison adopt the same hyperparameter selection strategy as SICUNet. The convolutional layers employ ReLU activation combined with He normal initialization to enhance training stability. The Adam optimization algorithm is used for model training, with the learning rate set at 0.0001. The output layer uses a Sigmoid activation, and the batch size is configured as 8.

TOPAZ5 forecasts are initialized from analyses that assimilate OSISAF SIC data. Therefore, when comparing SICUNet and TOPAZ5, both forecasts should be also evaluated against OSISAF observations; under these conditions, I would expect TOPAZ5 to perform better.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

Based on the reviewer’s recommendation, we included additional comparative experiments in the revised manuscript. In this experiment, SIC from OSI-450-a1 and OSI-430-a were used as inputs and outputs for training and prediction with SICUNet. The predictive performance of SICUNet was then compared with that of TOPAZ5 by using the OSI-SAF data as the reference truth. The details of the experimental design and results are presented in Section 4.6 of the revised manuscript, as follows:

4.6 Comparison with numerical models

It is noteworthy that current SIC products based on remote sensing observations contain uncertainties (Heinrichs et al., 2006; Ivanova et al., 2015; Ding et al., 2025). These may propagate through SICUNet and influence its predictions. However, the evaluation of SIC product uncertainty is a separate research topic and falls outside the focus of this work. In the presence of unavoidable observational uncertainties, it is especially important to maintain comparability among different prediction approaches using a common data reference. Considering that TOPAZ5 forecasts are initialized by assimilating OSI-SAF SIC data, this study adopts OSI-SAF SIC as both the input and validation reference for SICUNet. It is also used as the benchmark for evaluating TOPAZ5 forecasts. This ensures fairness and consistency across different prediction methods. The purpose of this configuration is to minimize the effects arising from differences in data sources. As a result, performance disparities can be mainly attributed to the forecasting methods themselves. This facilitates a clearer assessment of the predictive potential of data-driven models.

Since the TOPAZ5 forecast data begins in July 2021 (Hackett et al., 2025), this study selected data from three complete years (2022–2024) for comparison. This ensures the stability and representativeness of the results when evaluating the numerical model. The findings show that the MAE of TOPAZ5 exceeds that of SICUNet by more than 1.5%, while its BACC and NSE values are over 5% lower than those of SICUNet (Table 7). This highlights its advantages and stability in short-term SIC prediction.

Table 7: Prediction accuracy statistics of SICUNet and numerical simulation (Unit: %)

Model	MAE	BACC	NSE
-------	-----	------	-----

SICUNet	0.85	97.64	98.77
TOPAZ5	2.45	90.55	90.97

Section 4.6 requires clarification. Is the first week of a seasonal transition predicted by the first model and the second week by the second model, using predictors from a different seasonal regime? How would the authors handle a 35-day forecast spanning a seasonal transition?

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We apologize that the description in Section 4.6 (now Section 4.7 in the revised manuscript) of the manuscript was not sufficiently clear. We thank the reviewer for pointing out this critical issue. This study strictly follows the principle of seasonal consistency in both seasonal classification and model application.

Specifically, the period from April 1 to September 30 is defined as the melt season. Predictions during this period are performed using the melt-season model parameters. For the initial phase of the melting season (e.g., April 1–7), the model input consists of seven consecutive days of SIC from the end of March. The prediction process is conducted entirely using the melting-season model. Similarly, the period from October 1 to March 31 of the following year is defined as the freezing season. Predictions during this period are conducted using the freezing-season model parameters. During the onset of the freezing season (e.g., October 1–7), seven consecutive days of SIC data from the end of September are used as model input. The predictions are executed using the freezing-season model.

During the seasonal transition, it is indeed possible for a single 7-day prediction to span the boundary between the melting and freezing seasons. However, each prediction task is always performed based solely on the model parameters corresponding to a single season. Specifically, only the portion of the prediction results corresponding to the current season is retained. The cross-seasonal portion is discarded. Upon entering the next season, predictions are resumed using the model for that season.

The 35-day forecasts follow a recursive prediction scheme. Each forecasted step is used as input for the subsequent step. Accordingly, this study explicitly distinguishes between the direct

prediction phase and the recursive prediction phase in terms of temporal setup. The model first produces a direct 7-day prediction using seven consecutive days of observed SIC data, with no recursion involved at this stage. Recursive prediction begins on the 8th day. The model uses the predictions from the previous phase as input and progresses step by step to generate a longer forecast sequence. Therefore, in the recursive experiments for the melting and freezing seasons, April 8 and October 8 were set as the starting dates for recursive prediction. The predictions for April 1–7 and October 1–7 were obtained using 7-day direct forecasts. Within this setup, the 35-day recursive predictions generally do not cross the seasonal transition. If any sequences extend across the season boundary, only the predictions within the current season are kept. The cross-seasonal portions are truncated. This ensures that the seasonal context and model parameters remain consistent throughout the forecasting process.

Section 5 is entitled “Discussion”, but it primarily presents additional results obtained using extra input features. The manuscript would benefit from a more substantive discussion that engages with the main findings, limitations, and broader implications of the results, rather than introducing new result-focused analyses.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We thank the reviewer for the valuable suggestion. In accordance with your recommendation, the original “Discussion” section has been merged into the “Results and Analysis”. Additional discussion content has been included in the revised manuscript, as detailed below:

5 Discussion

This study relies exclusively on satellite observational data and develops an enhanced U-Net framework for short-term SIC forecasting. The framework incorporates residual connections and attention mechanisms to enable end-to-end short-term prediction of SIC. The results indicate that deep learning approaches can achieve strong short-term predictive capability without the support of numerical models or external dynamical field inputs. These findings suggest that rational network architecture design represents a key avenue for enhancing predictive performance. Moreover, the proposed framework offers useful guidance for the development of efficient and scalable sea ice forecasting systems.

It is important to emphasize that this study primarily focuses on exploring and validating the feasibility of predicting SIC using deep learning methods. However, the operational application of deep learning methods for SIC prediction remains limited (Palerme et al., 2025; ECMWF, 2025). Their stability, generalization capability, and long-term applicability require further evaluation. Therefore, gradually translating such models from the research stage to operational applications remains challenging. There is an urgent need for in-depth research on data quality control, model stability assessment, and model framework design.

From a methodological perspective, NIIEE is introduced into the loss function to enhance consistency in both the spatial structure and extent of sea ice. This constraint contributes to better representation of sea ice edges and overall spatial patterns in the predictions. However, NIIEE is essentially an empirical constraint. It does not explicitly incorporate thermodynamic or dynamic processes. In future studies, more physically consistent constraints could be explored. For example, simplified thermodynamic or dynamic conservation laws may be integrated into the loss function or network framework. This could improve both physical interpretability and long-term predictive stability.

Moreover, this study relies solely on satellite observations as model inputs and focusing on short-term prediction tasks. To extend applicability to longer-term or seasonal scales, future work could consider incorporating atmospheric and oceanic reanalysis data as additional inputs. This extension would allow a more systematic assessment of deep learning approaches under longer forecast horizons and more complex climate conditions.

Minor comments:

The authors should cite the works of Durand and Finn listed below. Although these approaches are model emulators and not purely observation-driven, they represent relevant surrogate modeling efforts for sea ice prediction and provide important context for the present study.

<https://doi.org/10.5194/tc-18-1791-2024>

<https://doi.org/10.1029/2024MS004395>

<https://doi.org/10.48550/arXiv.2508.14984>

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me

improve my article.

Following the suggestion from the reviewer, we have added the above three references in the Introduction. This provides a more comprehensive overview of the current state of Arctic sea ice prediction research. The corresponding changes have been highlighted in red in the manuscript.

Line 96: the manuscript states that the residual connection adds the input directly to the output of two convolutional layers plus a CBAM block; however, Figure 2 appears to show an additional convolutional layer along the residual path. This discrepancy should be clarified.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We apologize for any confusion caused by unclear statements in the manuscript.

The fundamental idea of residual connections is to perform element-wise addition of the input feature tensor of a residual block with its output features after several nonlinear transformations. This approach helps alleviate the vanishing gradient problem and improves training stability. In this study, the main branch consists of two convolutional layers and a CBAM attention module. The output features of the main branch are then added to the input features of the residual branch.

In the original manuscript, we failed to clearly explain a key technical detail. The convolution operations in the main branch can alter the number of channels in the feature map. As a result, the input and output cannot be directly added element-wise due to dimensional mismatch. This is a common design in deep residual networks (He et al., 2016). To address this issue, we introduced an additional 1×1 convolution layer in the residual branch. This layer serves only to match channel dimensions and does not incorporate any extra nonlinear activation. The main purpose of this layer is to make the addition operation mathematically valid. It also preserves the core principle of identity mapping in residual learning. This additional convolution layer for dimension alignment explains the apparent discrepancy between Figure 2 (actual structure) and the simplified depiction in line 96 of the original manuscript (idealized structure).

In response to the reviewer's comments, we have added the following details in the revised manuscript:

The ResNetCBAM consists of two convolutional layers in the main branch. The residual connection adds the input features to the output of the main branch. Because the convolution operations in the main branch may change the number of channels, a 1×1 convolution is introduced in the residual connection. This convolution aligns the input and output dimensions. This design enhances feature propagation and stabilizes gradient flow. The CBAM module is embedded after the second convolutional layer of the main branch to enhance attention to spatial and channel features (Fig. 2(a)).

Line 114: additional information would be helpful regarding how the hyperparameters were selected.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

The criteria for selecting hyperparameters were based on the training accuracy of the model. Currently, there is no clear method for hyperparameter selection; suitable hyperparameters can only be identified through continuous trial. However, certain tuning patterns can be summarized during this process. If the model exhibits low training accuracy, adjustments to the network structure and training regimen are typically applied to enhance performance. This may include increasing the number of convolutional channels, appropriately deepening the encoder–decoder hierarchy, or adjusting batch size or learning rate. Accurate predictions are achievable only after the convolutional neural network reaches an adequate level of training accuracy. In terms of network design, the encoder and decoder structures of U-Net fully account for the spatial features of the input data. The number of downsampling layers is set based on how many times the input grid size can be evenly divided. This ensures spatial scale consistency and effective feature alignment during repeated downsampling and upsampling. As a result, the skip connections can operate stably. In this study, the encoder–decoder hierarchy is configured such that the input features can be evenly divided.

ReLU activation functions are uniformly applied in the convolutional layers of the encoder and decoder. Their non-saturating property in the positive domain helps alleviate the vanishing gradient problem and improves training efficiency. The convolutional kernels are initialized using the He normal method. This initialization is statistically compatible with the ReLU

activation and helps maintain stable signal variance propagation in deep networks. Furthermore, Batch Normalization is applied after each convolutional layer to stabilize training and speed up convergence. It also improves the robustness of the model against variations in the distribution of different batches of samples. The model is trained using the Adam optimizer (Kingma and Ba, 2014). This gradient-based adaptive optimization method effectively accelerates convergence and enhances training stability.

In the output layer of SICUNet, a Sigmoid activation function is applied. This constrains the output within the $[0,1]$ range. It ensures that the predictions align with the physical values of SIC. The batch size is set to 8 to maintain training stability. This choice also takes into account training efficiency and the limitations of available hardware resources. Experiments showed that increasing the batch size further significantly raises memory usage. This may exceed the hardware capacity, whereas an excessively small batch size substantially reduces training efficiency. Considering training stability, computational resource constraints and efficiency requirements, this study selected a batch size of 8. This value represents a compromise between these factors.

We have revised and supplemented Section 3.2. All modifications are explicitly marked in the revised version. The revised content is presented as follows:

The performance of SICUNet depends not only on input and output data but also on hyperparameter settings. The model parameters are first randomly initialized. Subsequently, they are adjusted according to training accuracy to achieve an appropriate configuration. When training accuracy is low, several optimization strategies can be employed. These include adjusting the number of convolutional channels and appropriately deepening the encoder–decoder hierarchy. In addition, the batch size and learning rate can be adjusted to further improve the performance of SICUNet. The final selection of hyperparameters is determined through a comprehensive evaluation of the training and validation sets. In this study, the numbers of convolutional channels in the encoder–decoder are sequentially set to 64, 128, 256, 512, 1024, 512, 256, 128, 64, and 7. The convolutional layers in both the encoder and decoder adopt ReLU activation. He normal initialization is used to support stable training. SICUNet parameters are updated using the Adam optimization algorithm (Kingma and Ba, 2014). The learning rate is set to 0.0001. The batch size is set to 8 to balance training stability and efficiency.

It also considers hardware resource constraints. An early stopping strategy is applied during training. SICUNet is trained using a supervised learning approach, with the training workflow shown in Fig. 3.

The caption of Figure 2 should be improved by clarifying that panel (b) refers exclusively to the CBAM block.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We apologize for this oversight. We have improved the description of Figure 2:

Figure 2: (a) Structure of ResNetCBAM in SICUNet; (b) Framework of the Convolutional Block Attention Module.

Line 167: the term “7-day recursive forecasting step” is unclear. As described, SIC observations from 2–8 September are used to predict 9–15 September in a single forward pass, without recursion. The authors should clarify the terminology.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We sincerely apologize for the unclear wording. The term “7-day recursive prediction step” used here does not refer to recursive forecasting, but rather to a direct prediction of SIC for the subsequent seven days. To avoid confusion, we have revised this expression to “7-day direct prediction” in the manuscript.

Figure 6: I recommend plotting cumulative distribution functions (CDFs) of absolute errors instead, as the probabilities discussed in lines 175–180 are not readily apparent in the current figure.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

As pointed out by the reviewer, the conclusions described in lines 175–180 of the original manuscript are not clearly discernible from Figure 6. In response to the reviewer’s suggestion, the figure is redrawn as a cumulative distribution function (CDF) of absolute error to provide a

clearer view of the error distribution.

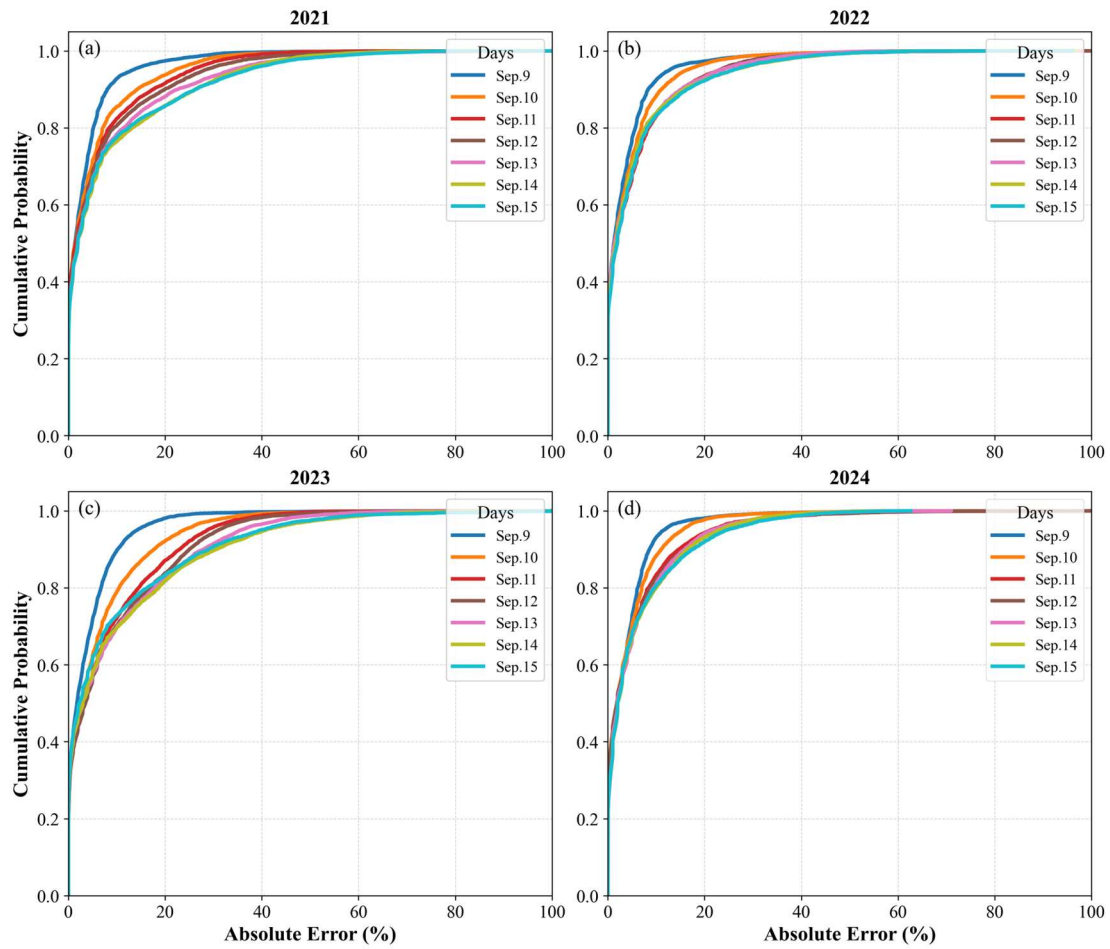


Figure 6: Cumulative distribution function of absolute errors between predicted and observed SIC.

The readability of Figures 4, 5, 6, 7, and 9 needs to be improved.

Response: Thanks for your suggestions and comments. Your feedback is greatly helping me improve my article.

We sincerely thank the reviewer for the constructive comments. Based on this suggestion, we carried out systematic revisions of the relevant figures in the manuscript. These revisions enhance the readability and clarity of their presentation.

It is worth noting that all figures presented in this manuscript were prepared at a resolution of 300 dpi. However, during the conversion of the Word document to PDF, figures containing multiple subpanels may exhibit a certain degree of visual distortion. This issue is difficult to completely avoid. Should the reviewer request them, the original high-resolution figures can be made available at any time. Nonetheless, following the reviewer's suggestions, we have

performed targeted optimizations for each figure, as detailed below.

Figure 4: Text labels were enlarged and emphasized. Line curves were thickened to improve visibility. The caption was supplemented to enhance figure readability and explanatory clarity.

Figure 5 contains numerous subpanels. This affected the overall clarity during PDF conversion. We further enlarged the text annotations and increased the overall resolution to improve visual quality.

Figure 6: Some information in the original figure was difficult to read directly. Therefore, the corresponding probabilistic characteristics were described numerically in lines 175–180 of the original manuscript. In response to the reviewer’s recommendation, Figure 6 was redrawn as a cumulative distribution function (CDF) of absolute errors to more clearly illustrate the error distribution.

Figure 7 is primarily used to evaluate sea ice edge differences between predictions and observations under extreme sea ice conditions during the testing period. Owing to the large number of subpanels in the original figure, overall readability was somewhat reduced. Accordingly, we performed additional statistical analysis of the BACC metric and added it to Figure 7 in a more concise format. The figure caption was also revised to emphasize the key points.

Figure 9: The original version displayed statistics across melting–freezing season transitions using multiple line plots. This resulted in numerous subfigures and some redundancy. Based on this consideration, the presentation of Figure 9 was adjusted. The original multiple subplots were consolidated into two subplots to more clearly convey the main statistical features. The caption was also revised accordingly.

The specific modifications and the revised figures are presented below.

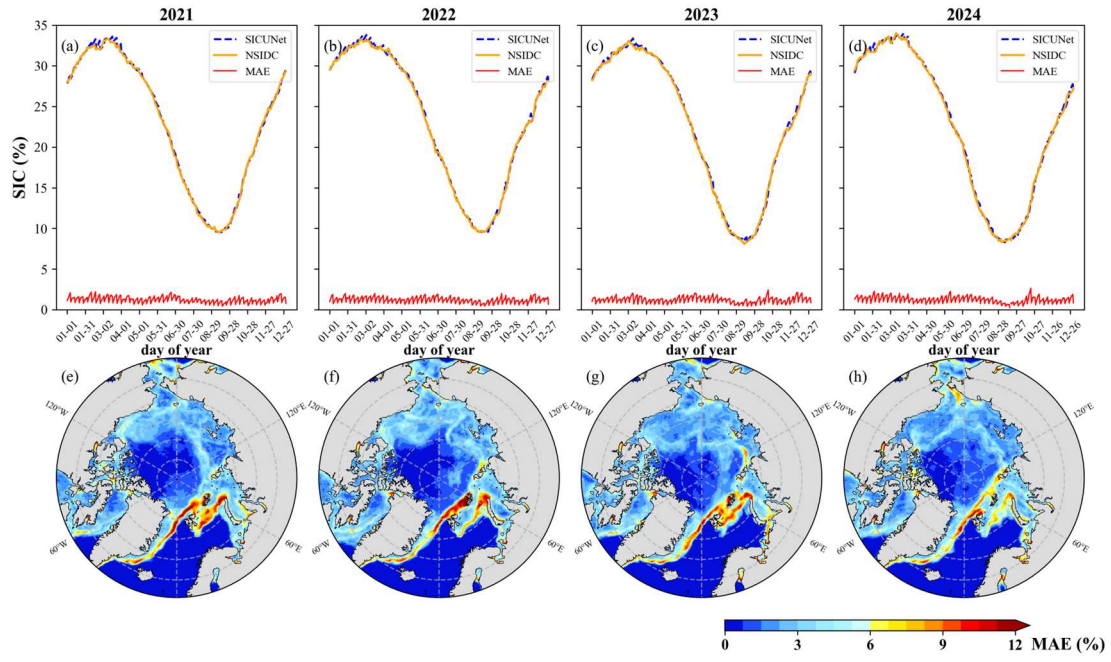


Figure 4: Prediction performance of SICUNet during the testing period. (a)–(d): SIC predictions, observations and MAE during the testing phase; (e)–(h): spatial distribution of annual mean MAE during the 2021–2024 test period.

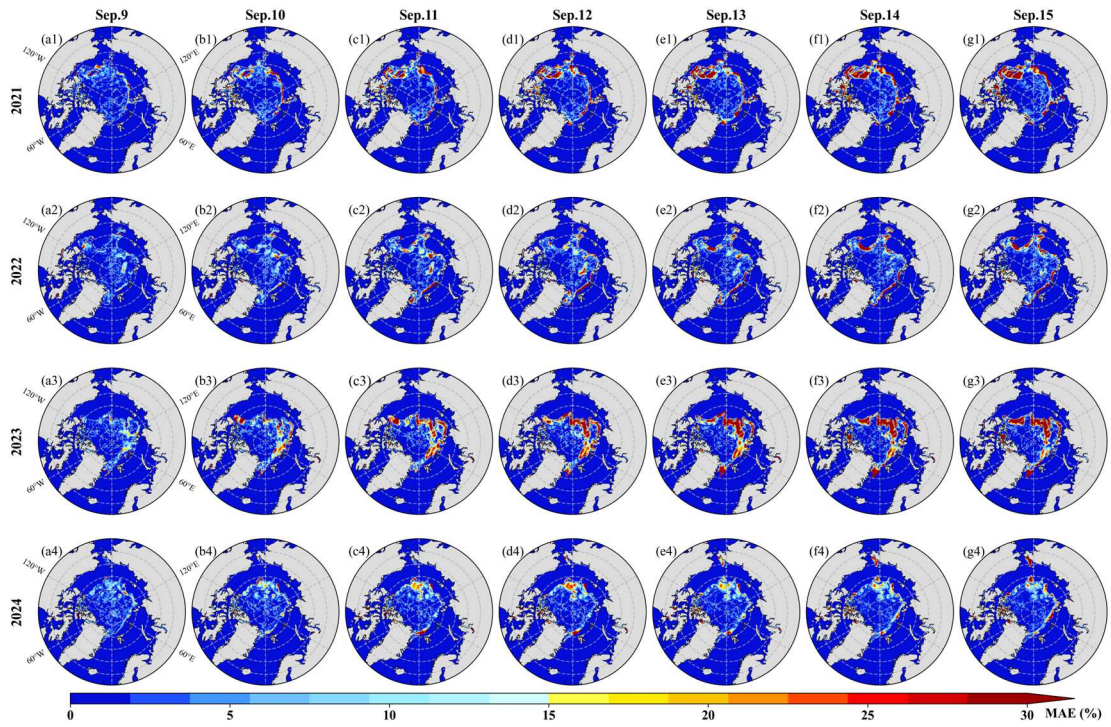


Figure 5: Spatial distribution of mean absolute error from September 9 to September 15 for each year during the testing period. (a1)–(g1), (a2)–(g2), (a3)–(g3) and (a4)–(g4) represent the spatial distribution of MAE from September 9 to September 15 in 2021, 2022, 2023 and 2024, respectively.

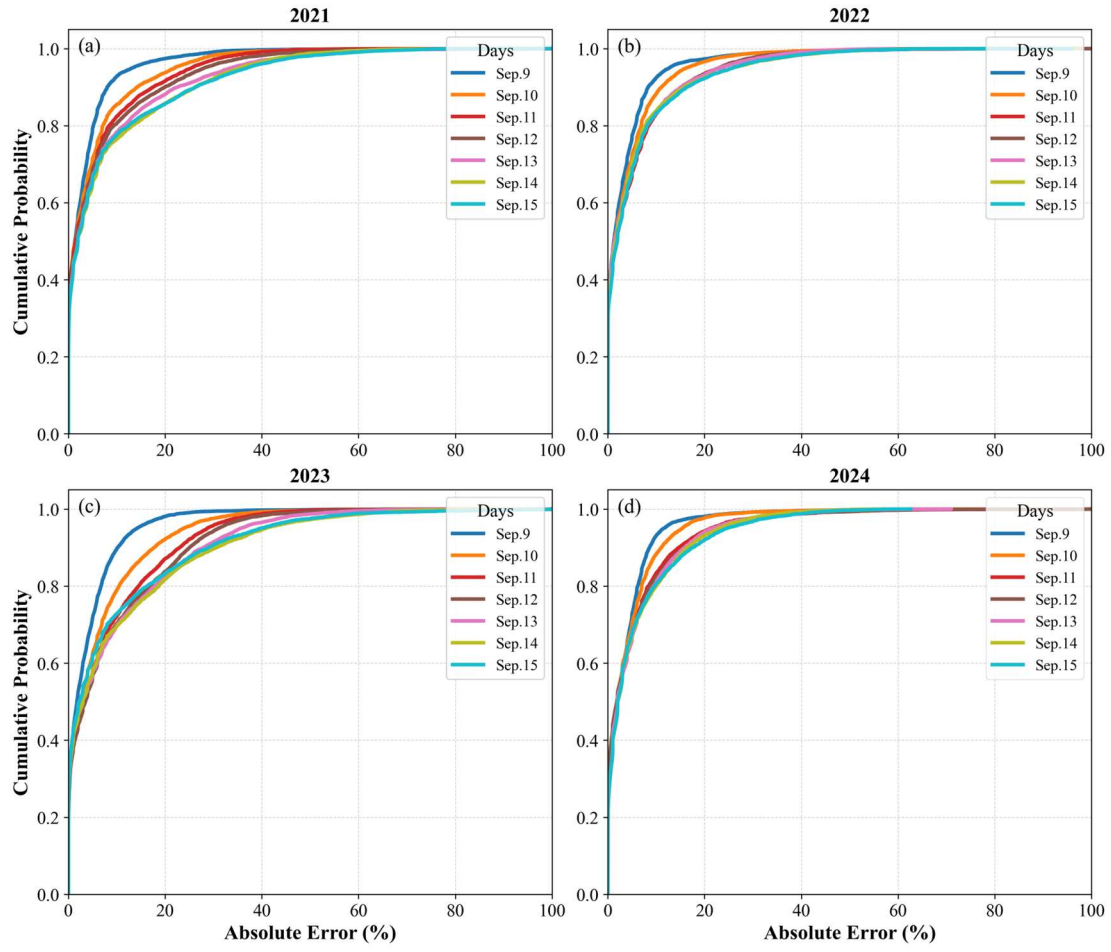


Figure 6: Cumulative distribution function of absolute errors between predicted and observed SIC.

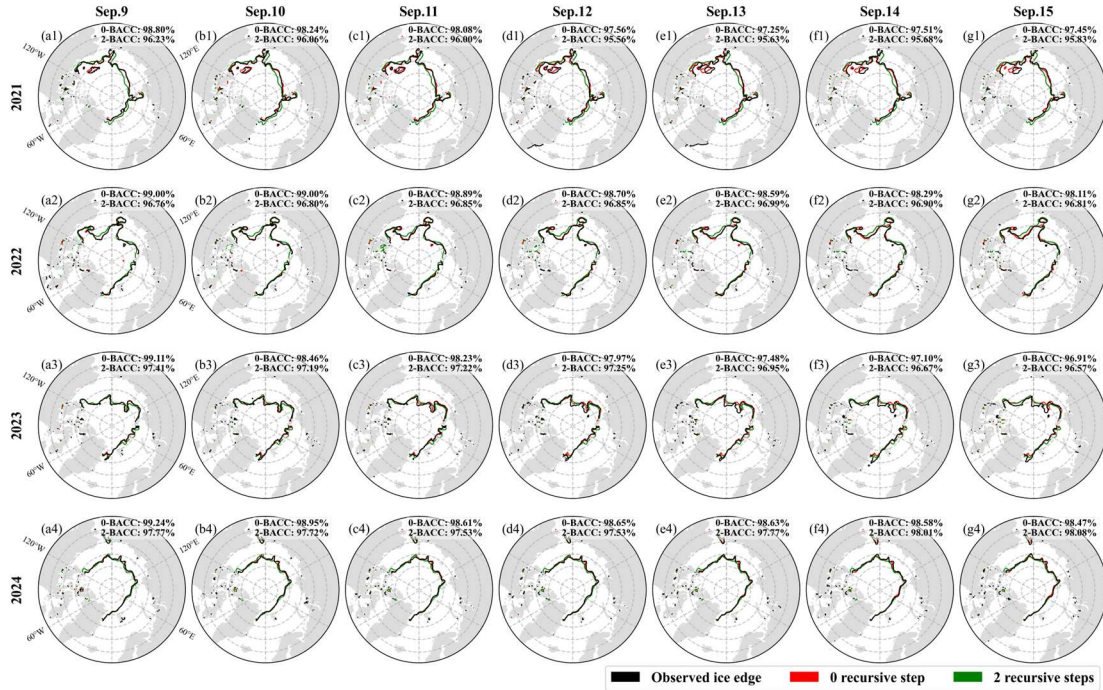


Figure 7: Sea ice edge from observations and SICUNet predictions at 0-step and 2-step recursive forecasts.

(a1)-(g1), (a2)-(g2), (a3)-(g3) and (a4)-(g4) represent the sea ice edge from September 9 to September 15 in

2021, 2022, 2023 and 2024, respectively. The black curve represents the satellite-observed sea ice edge; the red curve denotes the 7-day direct prediction sea ice edge; the green curve indicates the 21-day recursive prediction sea ice edge. The 0-BACC and 2-BACC correspond to the binary accuracy for the 7-day direct prediction and 21-day recursive prediction, respectively.

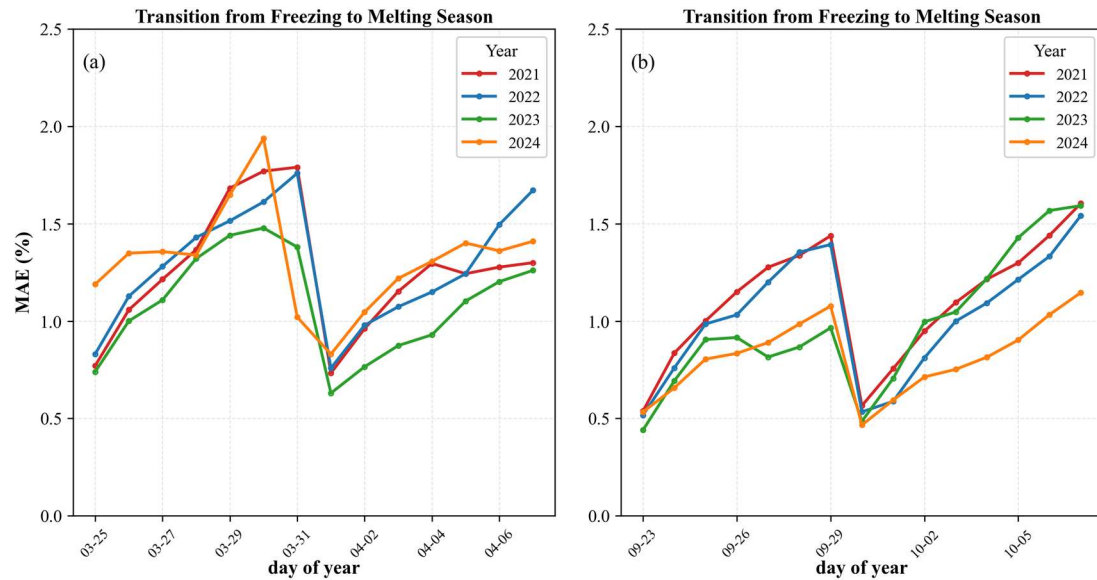


Figure 9: Mean absolute error during seasonal transitions. (a): mean absolute error during the transition from freezing to melting seasons from 2021 to 2024; (b): mean absolute error during the transition from melting to freezing seasons from 2021 to 2024.

We have carefully revised the expression of English and related contents of this paper. There are many changes in the article, but the structure and framework of this paper have not been changed. The relevant changes have been marked in red in the manuscript.

We appreciate for reviewer's warm work earnestly, and hope that the correction will meet with approval.

Once again, thank you very much for your comments and suggestions.

Reference

Ding, Y. R., Liu, X., Dai, X. F., Yang, Y., Yin, G. Y., Sun, H. P., and Guo, J. Y.: Spatiotemporal analysis of sea ice in the Weddell Sea of Antarctic based on GTWR, *Sci. Rep.*, 15, 5863. doi:

10.1038/s41598-025-90106-z, 2025.

ECMWF, “AIFS Machine Learning data,” European Centre for Medium-Range Weather Forecasts, 2025. [Online]. Available: <https://doi.org/10.21957/open-data>.

Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., and Jung, T.: Predictability of the Arctic sea ice edge, *Geophys. Res. Lett.*, 43, 1642-1650. doi: 10.1002/2015GL067232, 2016.

Hackett, B., Bertino, L., Ali, A., Burud, A., and Williams, T.: Copernicus Marine Environment Monitoring Service (CMEMS) ARCTIC_ANALYSISFORECAST_PHY_002_001 Product [data set], <https://data.marine.copernicus.eu/products/>, last access: 2025.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016, 770-778. doi: 10.1109/CVPR.2016.90, 2016.

Heinrichs, J. F., Cavalieri, D. J., and Markus, T.: Assessment of the AMSR-E Sea Ice-Concentration Product at the Ice Edge Using RADARSAT-1 and MODIS Imagery, *IEEE Trans. Geosci. Remote Sens.*, 44, 3070-3080. doi: 10.1109/TGRS.2006.880622, 2006.

Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, *Cryosphere*, 9, 1797-1817. doi: 10.5194/tc-9-1797-2015, 2015.

Kingma, D. P. and Ba, J. J. C.: Adam: A Method for Stochastic Optimization, doi: abs/1412.6980, 2014.

Palermé, C., Röhrs, J., Lavergne, T., Rusin, J., Kvanum, A. F., Macdonald Sørensen, A., Melsom, A., Brajard, J., Idžanović, M., Durán Moro, M., and Müller, M.: MET-AICE v1.0: an operational data-driven sea ice prediction system for the European Arctic, *Geosci. Model Dev.*, 18, 9751-9766. doi: 10.5194/gmd-18-9751-2025, 2025.

Palermé, C., Lavergne, T., Rusin, J., Melsom, A., Brajard, J., Kvanum, A. F., Macdonald Sørensen, A., Bertino, L., and Müller, M.: Improving short-term sea ice concentration forecasts using deep learning, *Cryosphere*, 18, 2161-2176. doi: 10.5194/tc-18-2161-2024, 2024.

Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., and Jagersand, M.: U2-Net: Going deeper with nested U-structure for salient object detection, *Pattern Recognit.*, 106, 107404. doi: 10.1016/j.patcog.2020.107404, 2020.

Ren, Y. and Li, X.: Predicting the Daily Sea Ice Concentration on a Subseasonal Scale of the Pan-Arctic During the Melting Season by a Deep Learning Model, *IEEE Trans. Geosci. Remote Sens.*, 61, 1-15. doi: 10.1109/TGRS.2023.3279089, 2023.