Supplementary material of "Calibration of short-term sea ice concentration forecasts using deep learning"

Cyril Palerme[1], Thomas Lavergne[1], Jozef Rusin[1], Arne Melsom[1], Julien Brajard[2], Are Frode Kvanum[1], Atle Macdonald Sørensen[1], Laurent Bertino[2], and Malte Müller[1]

[1]Norwegian Meteorological Institute, Oslo, Norway
[2]Nansen Environmental and Remote Sensing Center, Bergen, Norway

Correspondence: Cyril Palerme (cyril.palerme@met.no)
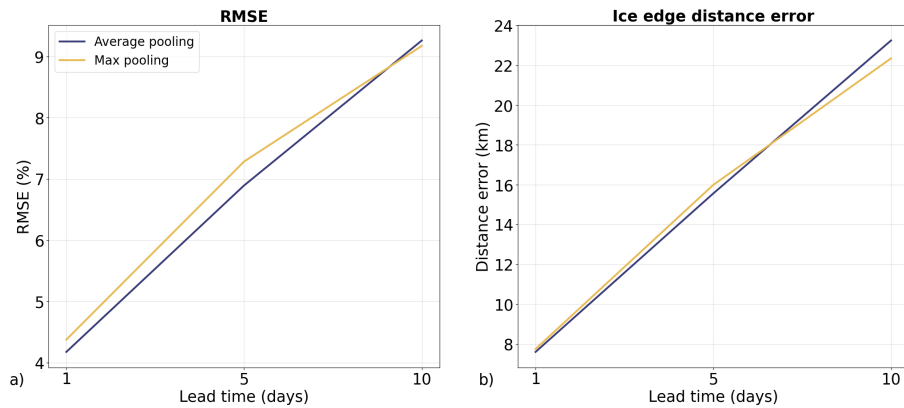
Figure S1. Comparison of the performances of deep learning models with the Attention Residual U-Net architecture using either average pooling or max pooling in 2021 (validation period) when the forecasts are evaluated using the RMSE (a) and the ice edge distance error (b).
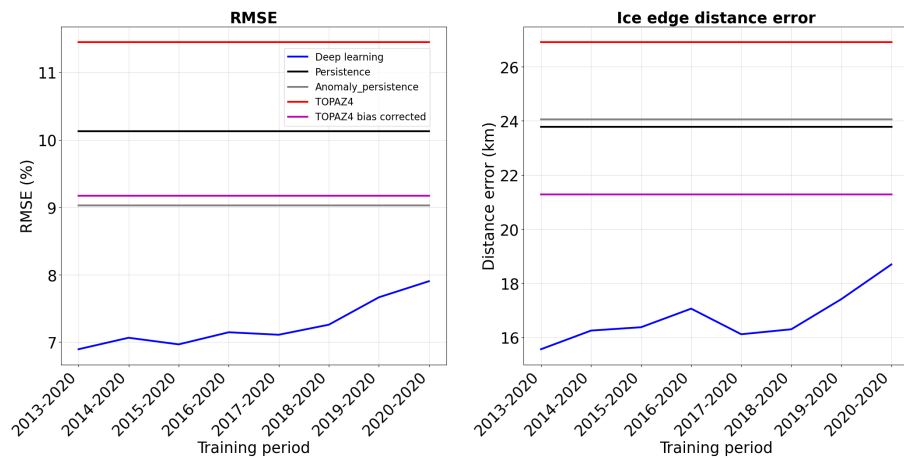
Figure S2. Performances of 5-day forecasts from deep learning models (with the Attention Residual U-Net architecture) trained over different periods (see x axis). This figure shows the performances of the deep learning models in 2021 (validation period). The model trained using the period 2013-2020 has the best performances for the RMSE and the position of the ice edge.
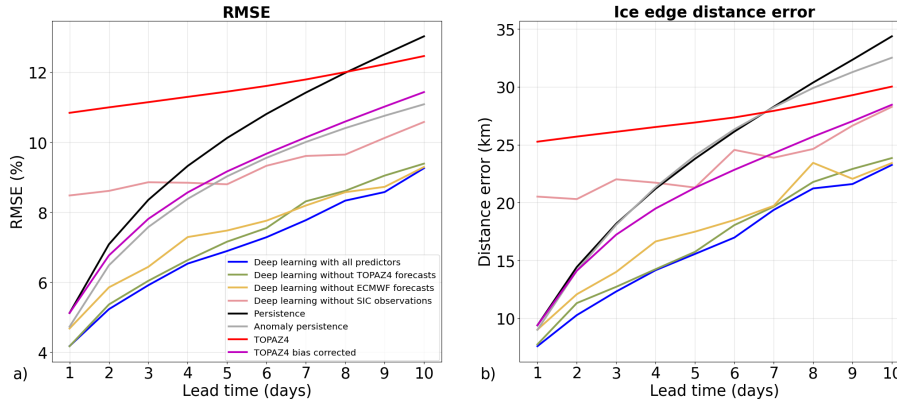
Figure S3. Performances of the deep learning models with the Attention Residual U-Net architecture during 2021 (validation period). The deep learning models using all predictors are shown by the blue curves, the models which do not use predictors from TOPAZ4 sea ice forecasts (sea ice concentration forecasts and initial errors) are shown by the green curves, the models which do not use predictors from ECMWF weather forecasts (2-m temperature and wind) are shown by the yellow curves, and the models which do not use predictors from sea ice observations (AMSR2 sea ice concentration, AMSR2 sea ice concentration trend, and TOPAZ4 initial errors) are shown by the pink curves.
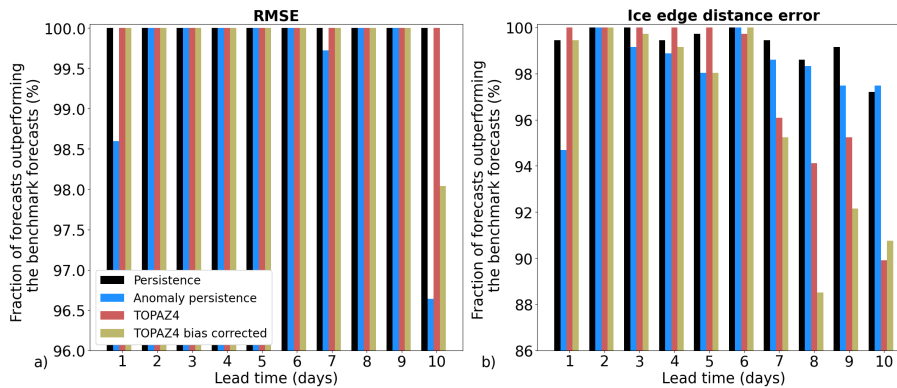
Figure S4. Fraction of days in 2021 (validation period) during which the forecasts from the models with the Attention Residual U-Net architecture outperform the different benchmark forecasts when the forecasts are evaluated with the RMSE (a) and with the ice edge distance error (b).
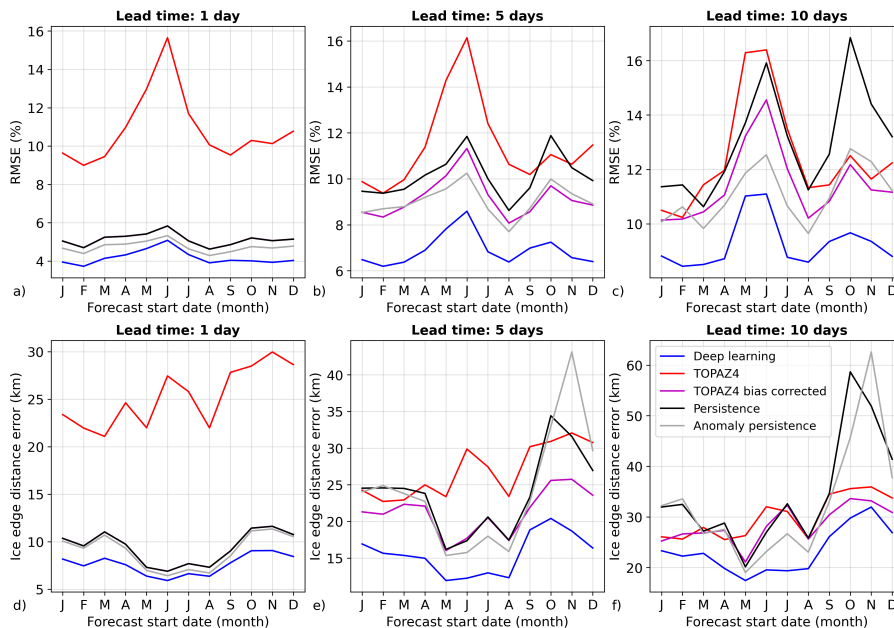
Figure S5. Seasonal variability in the performances of the deep learning models with the Attention Residual U-Net architecture in 2021 (validation period) for different lead times (1, 5, and 10 days) when the forecasts are evaluated using the RMSE (a, b, c) and using the ice edge distance error (d, e, f).