Machine-learning bias correction of passive-microwave sea-ice concentration using SAR improves Arctic analyses and forecasts

[1]Davi Mignac, Matthew J. Martin

[1]*Met Office, Exeter, UK*

Abstract

The accuracy of polar ocean prediction depends strongly on the initialisation of sea-ice concentration (SIC). Operational systems rely predominantly on passive-microwave retrievals, which are spatially smooth, relatively coarse, and subject to systematic biases—most notably across the marginal ice zone and during the summer melt, when surface melt ponds suppress the microwave signal. Machine-learning techniques now enable the derivation of high-resolution SIC retrievals from synthetic-aperture radar (SAR), which resolve these features in fine detail but provide only sparse and intermittent Arctic coverage. This study combines the complementary strengths of the two observation types. A Random Forest model is trained to bias-correct swath AMSR2-L2 SIC towards a machine-learning SAR product (ML-SAR), yielding a corrected observation dataset that preserves the spatiotemporal coverage of AMSR2-L2 while attaining SAR-like accuracy.

Relative to the SAR reference on a 1-year independent test set, the Random Forest model reduces the AMSR2-L2 SIC root-mean-square error from 0.16 to 0.05 and largely removes the bias (mean difference −0.10 to −0.004). The corrected product is assimilated in the Met Office global FOAM ocean–sea-ice forecasting system within a set of four data assimilation (DA) experiments: a control based on the operational AMSR2-L3; AMSR2-L2 alone; AMSR2-L2 combined with ML-SAR (AMSR2-ML-SAR); and the bias-corrected AMSR2-L2 combined with ML-SAR (AMSR2-BC-ML-SAR). Analyses and five-day forecasts spanning October 2024 to July 2025 are evaluated against SIC from regional ice charts, the MASIE ice-edge product, the SMOS-CS2-S3 and BGEP mooring sea ice thickness (SIT).

AMSR2-BC-ML-SAR is consistently the most skilful configuration. It increases the Critical Success Index across all concentration categories—most markedly during summer and for the highest-concentration class—and produces the lowest Integrated Ice Edge Error relative to MASIE, with the improvement maintained across all forecast lead times. AMSR2-BC-ML-SAR also produces a sharper SIC gradient at the ice edge, more faithfully reproducing the well-defined edges in the ice charts. The largest improvements occur during the melt season and in regions of sparse native SAR coverage. The principal trade-off is a moderate overestimation of sea-ice thickness, around 0.2 m on average against SMOS-CS2-S3 and BGEP that accumulates through spring, since no SIT observations are assimilated in those DA experiments.

Overall, the results demonstrate that a learned AMSR2-L2-to-SAR bias correction produces a sharper and less biased ice edge and concentration field in both analyses and forecasts, combining the spatial completeness of passive-microwave observations with the fidelity of SAR and offering a computationally inexpensive, observation-based pathway towards improved Arctic sea-ice prediction.