

Classification of Sea-Ice Concentration from Ship-Board S-Band Radar Images Using Open-Source Machine Learning Tools

Elizabeth Westbrook¹, Peter Gaube^{2,4}, Emmett Culhane³, Frederick Bingham¹, Astrid Pacini², Carlyn Schmidgall^{2,4}, Julian Schanze⁵, and Kyla Drushka²

¹University of North Carolina - Wilmington North Carolina, United States of America

²Applied Physics Laboratory - University of Washington - Seattle Washington, United States of America

²Applied Physics Laboratory - University of Washington - Seattle Washington, United States of America

³Woods Hole Oceanographic Institution - Woods Hole Massachusetts, United States of America

²Applied Physics Laboratory - University of Washington - Seattle Washington, United States of America

⁴School of Oceanography - University of Washington - Seattle Washington, United States of America

²Applied Physics Laboratory - University of Washington - Seattle Washington, United States of America

⁵Earth and Space Research - Seattle Washington, United States of America

Correspondence: Peter Gaube (pgaube@uw.edu)

Abstract. To gain context on the ambient sea ice field during the 2022 NASA Salinity and Stratification at the Sea Ice Edge (SASSIE) expedition we developed a machine learning model to predict sea ice cover classification from screen captures of a ship-board S-band navigation radar. The SASSIE expedition measured ocean surface properties and air-sea exchange approximately 400 km north of Alaska in the Beaufort Sea for 20 days, during which time screen captures from the shipboard S-band radar were collected. Our goal was to analyze these images to determine when the ship was approaching sea ice, in the ice, or in open water. Here we report on the development of a machine learning method built on the PyTorch software packages to classify the amount of sea ice observed in individual radar images on a scale from C0-C3. C0 indicates open water and C3 is assigned to images taken when the ship was navigating through thick sea ice in the marginal ice zone. The method described here is directly applicable to any radar images of sea ice and allows for the classification and validation of sea ice presence or absence. Furthermore, this method uses a standard marine navigation radar that is not generally used to measure sea ice and thus opens the opportunity to categorize sea ice concentration using the type of navigation radar installed on most vessels.

1 Introduction

In the Arctic, the advance of the autumn sea ice has become slower and occurs later in the season, whereas the retreat of the summer ice has become faster and occurs earlier in the season, both as a direct result of climate change (Stroeve and Notz, 2018). As more of the ocean surface is exposed to the atmosphere during longer periods of open water, air-sea heat and momentum fluxes, the freshwater cycle, surface albedo feedback, primary production, and regional and global climate, as well as human and ecological health are changing (Lannuzel et al., 2020). The NASA Salinity and Stratification at the Sea Ice Edge (SASSIE) mission aims to study the role that surface salinity and stratification have on controlling the freezing of sea ice (Drushka et al., 2024). The SASSIE expedition took place in the marginal ice zone during fall 2022 in the western Beaufort

20 Sea around 72° – 73.5° N, 154° – 145° W (Fig. 1). This area was generally covered with sea ice year-round until the late 2000s, when the marginal ice zone moved northward during the summer melt season. Drastic changes in sea ice cover are occurring throughout the Arctic and direct observations of sea ice are needed to help understand and document the changes.

Direct observation of sea ice concentration is important because passive microwave satellite measurements struggle to quantify SIC during the melt season (e.g., Kern et al., 2020). Ship-based S-band radar has been shown to be useful in the identification of sea ice, although other wavelengths are typically better for sea-ice characterizations (Haykin et al., 1985). Lund et al. (2018) used a shipboard X-band marine navigation radar to detect and track sea ice. Ku-band radar was used to characterize the types of sea ice observed, including first-year ice, multilayer ice, icebergs, or the shadows cast by icebergs (Orlando et al., 1990). However, classifying individual sea-ice radar images is labor intensive. Previous studies have used satellite observations together with machine-learning methods to detect and classify ice using synthetic aperture radar (SAR) (de Gélis et al., 2021; Stokholm et al., 2022) along with multi-spectral imagery (Rogers et al., 2024). Building on these methods we develop a machine-learning method to characterize images captured by a shipboard S-band radar that would otherwise be used for navigational purposes only.

The classification of sea ice types from S-band radar images lends itself well to the use of machine learning models in part because the images are invariant to rotation. The ubiquity of machine-learning classification methods in Earth Science has resulted in open-source tools that can be implemented on diverse datasets without the need to develop complex software required to build such models. In this project we use the open-source PyTorch machine learning library <https://pytorch.org/>.

2 Methods

2.1 S-Band Radar Data Collection

The SASSIE expedition took place aboard the *R/V Woldstad* which had a bow-mounted Furuno 2137S-BB S-band marine navigation radar. The primary purpose of the S-band radar is to provide the ship’s captain with imagery of the surrounding ice field to assist in safely navigating through the ice. We recorded images of the radar by splitting the signal between the display and a screen capture device (Epiphan AV.io HD). This video feed was recorded using a command line video conversion program called “ffmpeg” available at ffmpeg.org that saved screenshots as jpeg files typically every 60 s while the ship was in and around the ice. During autonomous vehicle deployments on the ice, images were acquired every 10 s to map the evolution of the ice at higher resolution (Drushka et al., 2024). While the ship was in ice-free water for more than 1 day, S-band acquisition was paused.

2.1.1 S-Band Radar Data Processing

Details on how the screen capture device was set up and the steps used to convert the images (Fig. 2a) into geolocated bit maps (Fig. 2b) are described in detail in Drushka et al. (2024) and summarized below. The metadata was extracted from the images using MATLAB’s computer vision toolbox. Metadata include the spatial range of the image, which varied in size from 1.4

km to 22 km. The pixels outside of this range were removed from each image, and pixels within the range were georectified and stored as Geo-TIFF images. The pixels were classified according to their red, green and blue (RGB) values. Color values consistent with yellow to red color ranges were classified as “sea ice”, and those consistent with blue and green colors were classified as “no data” because that is the color used by the Furuno display to indicate range rings and radar scanning lines that were not retained for the analysis.

Sea clutter is noise near the center of the radar image that results from surface waves reflecting the radar signal (e.g., prominent yellow-red spot in the center of Fig. 2a). Sea clutter appears in every radar image but its extent and intensity varies depending on factors including the sea state and the range setting of the radar. Visual inspection showed that sea clutter typically does not appear at a distance greater than 5% of the image’s range. In regions of heavy ice cover, sea clutter overlaps with sea ice. As a result, sea clutter could not be removed with a simple mask and was instead removed as follows. RGB values within 5% of the image range were summed as a function of distance from the center of each image (ΣRGB); sea clutter appears as a peak in ΣRGB . The edge of the sea clutter was defined as the distance from the center at which sea clutter could no longer be detected, which was identified as the distance at which ΣRGB dropped below 5% of its peak value. If no values of ΣRGB were below this threshold, the threshold was increased iteratively until ΣRGB dropped below the threshold. Pixels at smaller distances from center were flagged as sea clutter. The flagged images were stored as Level-4 (L4) Geo-TIFF images (e.g., Fig. 2b). In total, more than 24,000 L4 files were produced from the processing of the S-band images.

2.2 Classification of Sea-Ice Concentration Using L4 Geo-TIFF Images

The classification of sea ice along the ship track of the SASSIE expedition was first attempted as a simple ratio of black to white pixels. This did not work for the following reasons:

2.2.1 Noise in the L4 Images from Unidentified Sea Clutter

Due to turbulent water around the ship and noise from the S-band radar, it is common to see what looks to be small ice floes around the ship that are not actually ice. These range in intensity from a few splotches (Fig. 3a) to a full ring of white around the ship (Fig. 3b). This occurred mostly in cases where the ship was in open water, but also happened sometimes when the ship was in ice (Fig. 3c). These are likely sea clutter that was not identified by the algorithm described above. Since this sea clutter is unidentified, it is not possible to distinguish from genuine ice floes. However, the appearance of unidentified sea clutter and its typical proximity to the ship itself make it possible to recognize with a ML model.

2.2.2 Distortion of Ice Due to the Leading Edge Detection

When the ship approaches a patch of sea ice, the amount of ice in the floe is underestimated due to the radar’s detection of the leading edge of the ice, that is, the ice floes that are closest to the ship. The ship’s radar bounced off of the leading ice floes creating an acoustic shadow, so ice floes in the shadow were not detected. This is revealed in two radar captures taken 3 minutes apart as the ship enters a dense patch of ice (Fig. 4). This leading edge of the ice floe is shown in red and in Fig. 4b

blue arrows indicate ice floes missing or underestimated because they were in the shadow of larger ice floes. The result of this leading edge effect is an underestimation of ice concentration in close proximity to the ship when simple ratios of black and white pixels are used to analyze these images. In the machine learning model described below we are able to account for this distortion by including numerous examples in the training data.

2.2.3 Radar Setting Changes Impact How Sea-Ice is Observed

Since the primary purpose of the S-band radar was to allow for safe navigation of the *R/V Woldstad* through the ice field, radar settings were changed throughout the campaign depending on ice conditions, weather conditions, and captain preference. Ice fields on the radar look different depending on these settings, specifically the set range of radar detection. Two cases where the ship is in heavy ice with different ranges on the radar are shown in Fig. 5. When the radar is ‘zoomed-out’ to show a large area, ice near the ship appears much smaller and ice around the edges of the frame is not picked up on the radar (Fig. 5a) whereas when zoomed in, more ice is detected near the ship (Fig. 5b). The method we propose below is insensitive to this effect as the model is trained on images at all zoom levels.

2.3 Developing a Classification Scheme and Training Dataset

These issues with using S-band radar to classify ice led us to abandon our simple ratio estimation and develop a machine learning model to classify the images. To this end, a subset of 1,100 of the 24,000 L4 Geo-TIFF images were chosen in non-sequential order to remove near-duplicates to develop a training data set that were labeled using a simple 4-level classification with ice concentration ranging from a zero (C0) to three (C3). The training data set included roughly an equal number of images from each of the four classification levels. We chose to use ice levels versus directly estimating ice coverage because our interest is to identify whether the ship was in ice, around ice, or not in ice. This 4-level classification has a value of C0 when there is no ice at all in the frame of the radar image and the ship is in open water (Fig. 6a). Noteworthy variation in the L0 images was observed due to different amounts of noise present on the radar image due to deployed instruments (Wave Gliders, SWIFT buoys, etc.; see Drushka et al 2024). The C1 ice concentration includes images with just a few individual ice floes in the frame (Fig. 6). For radar captures where the ship was in low to medium concentrations of ice (~10% or more), or there is high concentration of ice in the frame but the ship is not passing through it, a C2 classification was used (Fig. 6c). Finally, a C3 classification was used to indicate that the ship is passing through sea ice with concentration in excess of about 30% (Fig. 6d). Higher sea ice concentrations were not encountered because the ship does not have ice breaking capabilities.

2.4 The Machine Learning Model

The machine learning (ML) model is well suited to use images that contain random anomalies like those from sea clutter in the S-Band radar images (Qu et al., 2022), and convolutional neural networks (CNNs) are particularly well-suited for tasks like this where visual patterns are perceptible to humans but difficult to formalize into specific rules. Since CNNs can learn these insights directly from labeled data (e.g., human-labeled ice levels), the model can be trained without requiring hand-crafted rules or

filters. This is ideal in situations where human experts can label examples but cannot precisely specify what distinguishes one level of ice from another in terms of pixel properties. CNNs bridge this gap by mapping visual patterns to labels directly, learning filters that capture various levels of detail—from edges and textures to larger structural patterns—enabling the network to distinguish between visual nuances like the difference between low-ice concentrations and sea-clutter, that resist more simplistic rules-based methods.

The machine learning (ML) model we used for image classification is an off-the-shelf implementation of the VGG-19 network in PyTorch (Simonyan and Zisserman, 2014). The VGG-19 is designed for the ImageNET benchmark task and we found it to be a good match to the problem we faced. To reduce the complexity of the model to suit our task we reduce the final fully connected layer from 100 to 4 dimensions, and initializing the model with pre-trained weights (Yosinski et al., 2014). Predictions are generated using the softmax activation function on the final fully connected layer (Bridle, 1990). VGG-19 is a complex and computationally expensive architecture, and so risks overfitting and can require the use of multi-core GPUs on large datasets (Huang et al., 2017). Here, we are able to run the model to convergence on a local CPU in less than an hour, as the single-channel image dataset is just 26MB on disk. We prevent overfitting by regularizing the model through the use of dropout and data augmentation. Dropout in a CNN is a regularization technique that randomly deactivates a fraction of neurons during training to prevent overfitting and improve the network’s ability to generalize to new data (Srivastava et al., 2014). Our best performing model (Fig. 1) used dropout probabilities of 0.6 on the two fully connected layers, such that 60% of the neurons in these layers are randomly set to zero for each training example. In addition to dropout, we regularized the model using a data augmentation protocol that took advantage of the rotational invariance of the images. The S-band radar observes a circular region around the vessel, and thus the assigned class should be invariant to any rotational changes to the input. Here, we used an augmentation protocol that rotated 75% of the images vertically and/or horizontally by up to 50%, such that just 12.5% of input images are fed to the model in their original orientation (as in Bridle, 1990). The model was trained for 15 epochs using a batch size of 64, a train/test split of 85/15, and 5-fold cross-validation. The model achieved 91% accuracy on the test data, with class-specific performance shown in Table 1 and confusion matrix in Figure 7. The model shows high performance on all 4 classes and does not misclassify any of the testing data by more than one ice level (ex. C0 are only misclassified as C1).

3 Validation

3.1 Photo and Logbook Validation

Throughout the SASSIE expedition, notes documenting the state of the ice were recorded in the ship’s log (N=31) and geolocated photos (N=10) were taken from the deck of the *R/V Woldstad* while the ship was in ice. To validate the predictions of our model we compiled a ‘ground-truth’ dataset from the images and logbook entries. Figures 8a-c provides examples of photos from the cruise that are classified as C1, C2, and C3, respectively. Observations from the photos and ship’s log were tabulated and compared with the moving hourly average from the S-band ice classification product, which is shown in Fig. 9 below revealing general agreement between the model classification (black curve) and the ship-board photos (blue points) and ship’s log (red points). In general the model estimates agree with both the photo and log book observations, however the point-wise

comparison lacks the granular detail required to compare the high-frequency S-band radar observations to validation data and the time stamps of the log book entries were often rounded to the nearest hour, resulting in some mismatch between the log book observations and our S-band based characterization.

3.2 Comparisons to Regional Ice Coverage Maps

150 To assess the validity of the model predictions of sea ice classification we compare our estimates to those of a derived daily gridded sea ice product generated from operational sea ice maps made by the National Weather Service Alaska Sea Ice Program (grASIP). The details of how the product is produced are given in Pacini et al. (2025). Following the shiptrack, general agreement is found between sea ice concentration estimated from the model developed here and that of the grASIP (Fig. 10). Regression of our classification to the daily grASIP sea ice concentrations yields an $r^2 = 0.78$.

155 3.3 Comparison to Temperature and Salinity

The S-band ice products are compared with temperature and salinity measurements from the shipboard thermosalinograph (Drushka (2024)) and the ‘salinity snake’ which recorded data at the sea surface throughout the field campaign (Schanze (2024)). Time series of hourly moving S-band product average together with temperature and salinity measurements were constructed and indicate that generally the warmest temperatures were observed when the ship was in open water and the
160 coolest when the ship was near sea ice, with generally good agreement between the ice classification and the temperature (Fig. 11).

Visualizing both the temperature and salinity as a function of sea-ice classification (Fig. 12) we can observe that generally C2 or C3 classifications are numerous in fresh and cool water, while C0 and C1 radar captures occurred in generally warmer and saltier water. The highest number of C3 classifications (our highest ice concentration) occurred in water cooler than 0
165 degrees C and fresher than 24, which suggests that our automated algorithm is indeed detecting sea ice and not just sea clutter.

4 Discussion and Conclusions

To classify sea-ice concentration into four categories during a 20-day expedition into the Beaufort Sea we employed an open-source machine learning model trained on a subset of manually classified S-band radar images. Our model was successful at predicting a subset of training data between 84% and 95% of the time when tested on a random 85/15 split of the training data
170 (Fig. 7). Using this model we classified over 24,000 images and compared these classifications to both photos taken from the ship as well as the Alaska Sea Ice Program gridded ice maps and shipboard measurements of temperature and salinity. This simple model yielded predictions that fit into the temperature and salinity expected for differing ice coverage. It is important to note that the middle two classifications (C1 and C2) were difficult to define in the training data and also yielded complex relationships with temperature and salinity suggesting that our methods are best suited for determining if there is ice or not,
175 versus the amount of ice present. Overall, this method provides an example of using existing machine learning methods to

create value-added data products from what generally is considered navigation radar and not traditionally used to characterize sea ice.

The distribution of ice classification from the S-band radar images in temperature and salinity space (Fig. 12) highlight the influence surface salinity has on ice concentration with the two highest categories of sea-ice (C3-C4) are almost exclusively
180 observed in water fresher than 25, regardless of temperature. Higher sea-ice classifications were observed in warmer waters (up to 2°C) only at salinity less than 24.5.

The methods developed here are applicable to any images or radar returns of sea ice. Existing radar and image observations of sea ice can use the methods developed here together with a curated training dataset to classify sea-ice concentrations. The application of methods developed here to existing ship-based radar images could provide important observational data of the
185 marginal ice zone where validation data can be hard to find (Pacini et al., 2025). Future efforts to extract sea-ice concentration from S-band radar images would be improved if continuous photographs are taken from the platform hosting the radar. Video or time-lapse images would help to build a more accurate training dataset and thus increase model precision.

5 Interactive computing environment

Code availability. All the codes used to train and run the ML model are available on the project repository [https://github.com/NASA-](https://github.com/NASA-SASSIE/seaice)
190 [SASSIE/seaice](https://github.com/NASA-SASSIE/seaice).

Data availability. The L4 Geo-TIFFS are available at NASA PO.DAAC, doi:10.5067/SASSIE-SBAND4. Shipboard temperature and salinity data are available at doi:10.5067/SASSIE-TSG2. Predictions from the ML model developed here are available also on the PO.DAAC, doi:10.5067/SASSIE-SBAND-ML.

Author contributions. EW, PG, EC, FB, CS, and KD contributed equally to the overall design of the project. EW and EC led the ML model
195 development. KD processed the raw images into L4 Geo-TIFFS and ship-board TSG data. JS processed the salinity snake data. AP created the gridded ASIP dataset, in collaboration with the National Weather Service. PG and EW led the writing of the manuscript.

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

Acknowledgements. The authors thank all the participants that made the SASSIE program possible including support from the NASA Physical Oceanography Program. We thank Nadya Vinogradova-Shiffer of NASA for her support throughout the SASSIE program. We

200 are grateful to the captains and crew of *R/V Woldstad* and to Support Vessels of Alaska Inc. We thank two anonymous reviewers for their comments, which strengthened the paper.

References

- Bridle, J. S.: Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition, in: *Neurocomputing: Algorithms, architectures and applications*, pp. 227–236, Springer, 1990.
- 205 de Gélis, I., Colin, A., and Longépé, N.: Prediction of Categorized Sea Ice Concentration From Sentinel-1 SAR Images Based on a Fully Convolutional Network, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 5831–5841, <https://doi.org/10.1109/JSTARS.2021.3074068>, 2021.
- Drushka, K.: SASSIE Arctic Field Campaign Shipboard Thermos-salinograph Data Fall 2022, PO.DAAC, <https://doi.org/10.5067/SASSIE-TSG2>, 2024.
- 210 Drushka, K., Westbrook, E., Bingham, F., Gaube, P., Dickinson, S., Fournier, S., Menezes, V., Misra, S., Perez, J., Rainville, E. J., et al.: Salinity and Stratification at the Sea Ice Edge (SASSIE): An oceanographic field campaign in the Beaufort Sea, *Earth System Science Data Discussions*, 2024, 1–56, 2024.
- Haykin, S., Currie, B. W., Lewis, E. O., and Nickerson, K. A.: Surface-based radar imaging of sea ice, *Proceedings of the IEEE*, 73, 233–251, 1985.
- 215 Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q.: Densely connected convolutional networks, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.
- Kern, S., Lavergne, T., Notz, D., Pedersen, L. T., and Tonboe, R.: Satellite passive microwave sea-ice concentration data set inter-comparison for Arctic summer conditions, *The Cryosphere*, 14, 2469–2493, 2020.
- Lannuzel, D., Tedesco, L., Van Leeuwe, M., Campbell, K., Flores, H., Delille, B., Miller, L., Stefels, J., Assmy, P., Bowman, J., et al.: The 220 future of Arctic sea-ice biogeochemistry and ice-associated ecosystems, *Nature Climate Change*, 10, 983–992, 2020.
- Lund, B., Graber, H. C., Persson, P., Smith, M., Doble, M., Thomson, J., and Wadhams, P.: Arctic sea ice drift measured by shipboard marine radar, *Journal of Geophysical Research: Oceans*, 123, 4298–4321, 2018.
- Orlando, J. R., Mann, R., and Haykin, S.: Classification of sea-ice images using a dual-polarized radar, *IEEE Journal of Oceanic Engineering*, 15, 228–237, 1990.
- 225 Pacini, A., Steele, M., and Schreck, M.-B.: National Weather Service Alaska Sea Ice Program: Gridded ice concentration maps for the Alaskan Arctic, *EGUsphere*, 2024, 1–28, 2025.
- Qu, Q., Zhang, Y., Wang, X., Wang, X., and Wang, X.: A False Alarm Controllable Detection Method Based on CNN for Sea Surface Small Targets, *IEEE Sensors Journal*, 22, 15 077–15 087, <https://doi.org/10.1109/JSEN.2022.3182261>, 2022.
- Rogers, M. S., Fox, M., Fleming, A., van Zeeland, L., Wilkinson, J., and Hosking, J. S.: Sea ice detection using concurrent multispectral and 230 synthetic aperture radar imagery, *Remote Sensing of Environment*, 305, 114 073, <https://doi.org/https://doi.org/10.1016/j.rse.2024.114073>, 2024.
- Schanze, J.: SASSIE Arctic Field Campaign Shipboard Salinity Snake Data Fall 2022, PO.DAAC, <https://doi.org/10.5067/SASSIE-SNK2>, 2024.
- Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*, 2014.
- 235 Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting, *The journal of machine learning research*, 15, 1929–1958, 2014.

- Stokholm, A., Wulf, T., Kucik, A., Saldo, R., Buus-Hinkler, J., and Hvidegaard, S. M.: AI4SeaIce: Toward Solving Ambiguous SAR Textures in Convolutional Neural Networks for Automatic Sea Ice Concentration Charting, *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–13, <https://doi.org/10.1109/TGRS.2022.3149323>, 2022.
- 240 Stroeve, J. and Notz, D.: Changing state of Arctic sea ice across all seasons, *Environmental Research Letters*, 13, 103 001, 2018.
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H.: How transferable are features in deep neural networks?, *Advances in neural information processing systems*, 27, 2014.

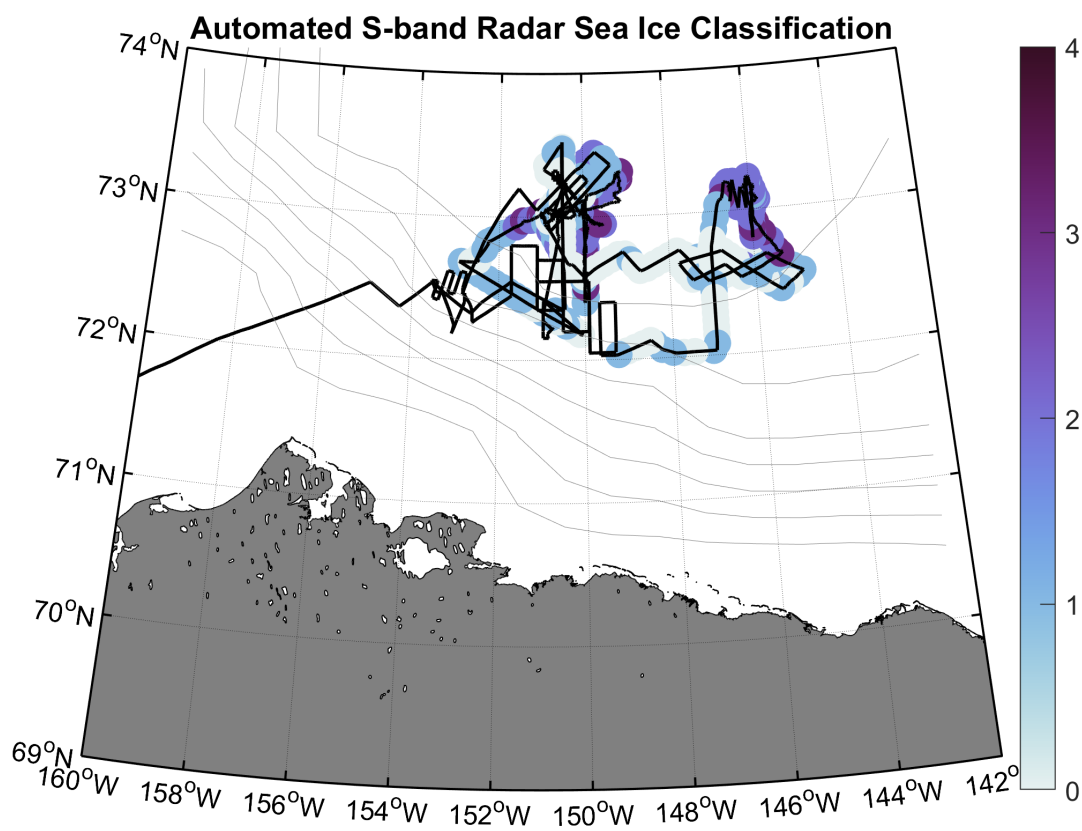


Figure 1. Automated S-band radar ice classification overlaid on the track of the *R/V Woldstad* during the September 2022 Salinity and Stratification at the Sea Ice Edge (SASSIE) expedition. S-band radar images were taken every five minutes but only every 50th record is shown for clarity. Bathymetry contours start at -500m and are spaced at 500m intervals.

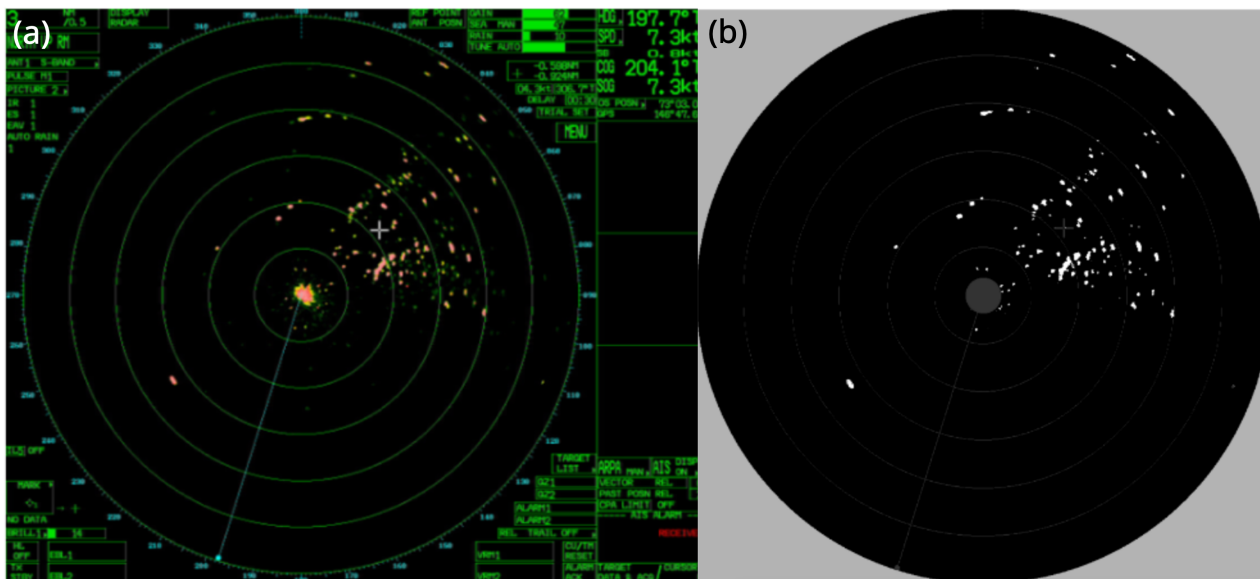


Figure 2. Example images recorded from the S-band radar display onboard the R/V Woldstad. The actual screen capture from the Furuno S-band radar display is shown in panel (a) and the accompanying L4 geolocated image used in this analysis in panel (b). The ship is at the center of the images and sea clutter is visible around the ship in panel (a). North is oriented up in both panels. The line clearly visible across the radar image is the current direction the radar is facing as it scans. Sea ice is visible as yellow and red pixels in panel (a) that are then converted into a geolocated binary image shown in panel (b).

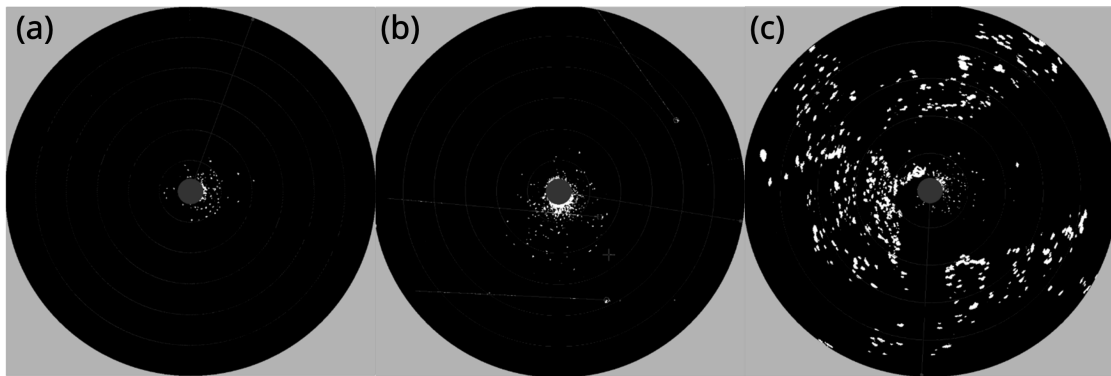


Figure 3. Example L4 Geo-TIFF images created during the S-Band radar data processing that show multiple examples of sea clutter. The range of the outermost rings are 20, 10, and 1 nautical mile, respectively, in panels a through c. The images are sourced from the training dataset, from which the text showing range, orientation, ship speed, etc. has been removed as part of the processing. (a) Example L4 Geo-TIFF image with minor unidentified sea clutter, (b) major unidentified sea clutter, and (c) unidentified sea clutter while in the ice. Examples in panels a and b had a sea ice classification of C0 and panel c was classified as C3.

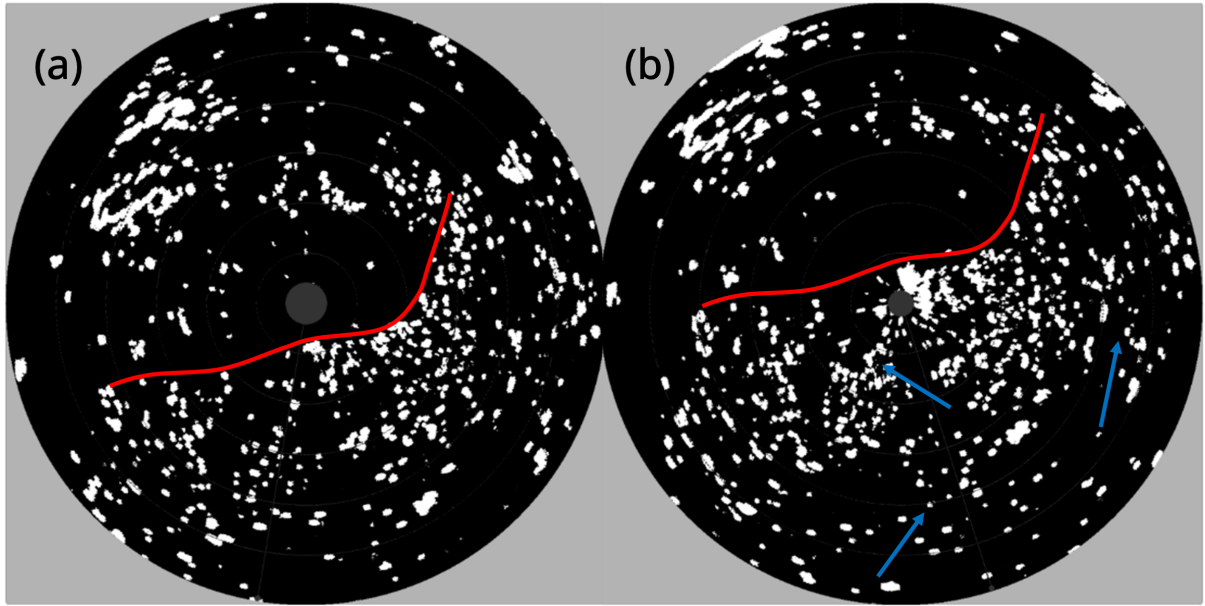


Figure 4. Example L4 Geo-TIFF image from September 9 2022 at (a) 17:51 and (b) 17:54 showing the effect of returns off of the leading edge of the ice pack (red curve) and returns off of ice not visible when first approaching the dense patch of ice (blue arrows in panel b). The range or geographic extent of these images is not reported in this figure as the images are sourced from the training dataset which did not include range, orientation or ship speed.

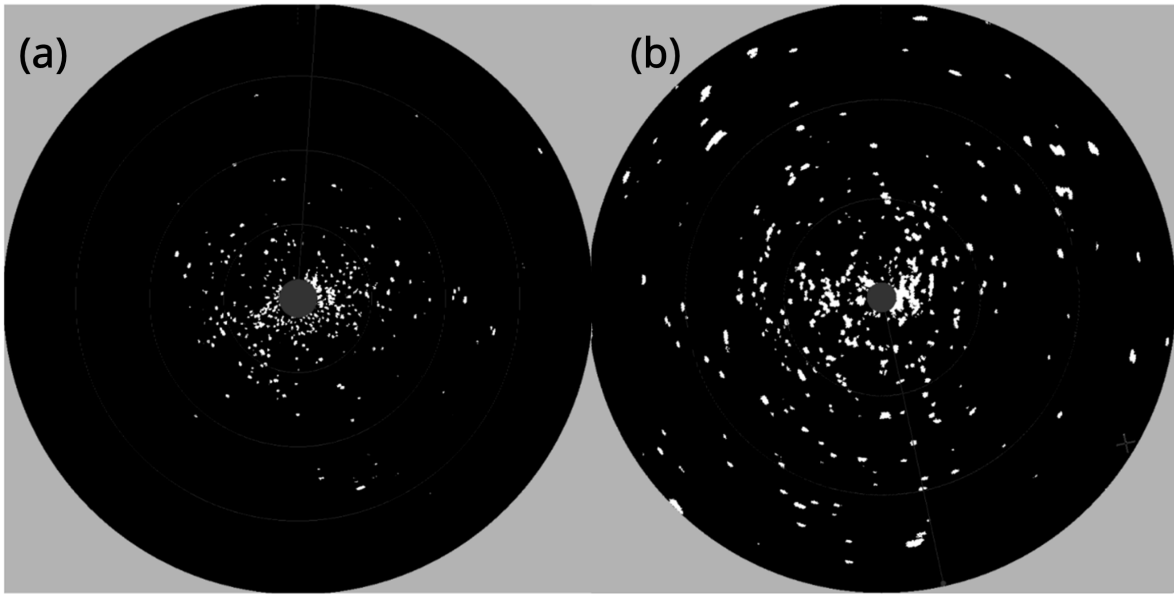


Figure 5. Example L4 Geo-TIFF images two minutes apart showing how changes in the radar setting by the ship’s captain, in this case changes in radar range with (a) being zoomed out (larger range) as compared to (b), affect the detection of sea ice. The range or geographic extent of these images is not reported in this figure as the images are sourced from the training dataset which did not include range, orientation or ship speed.

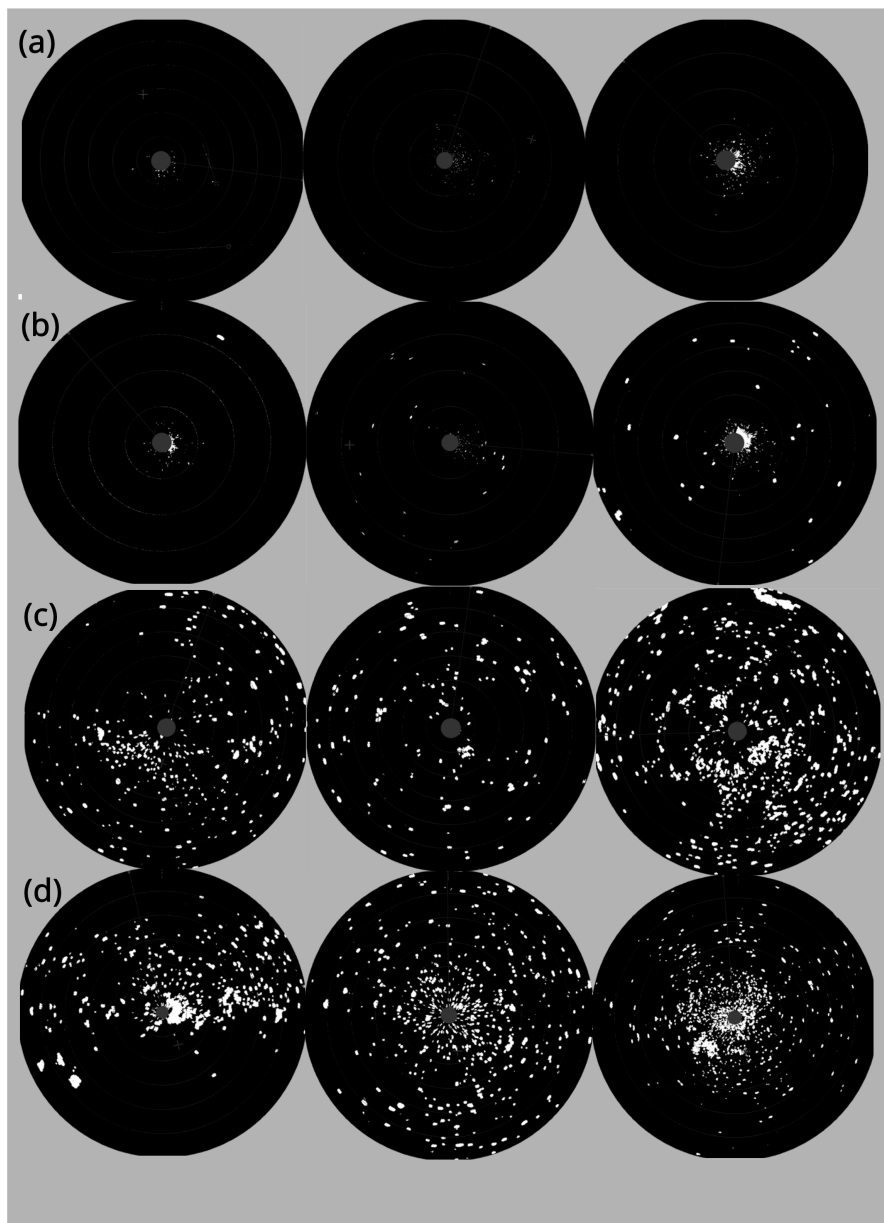


Figure 6. L4 Geo-TIFF images taken from the training dataset showing three examples each of (a) C0, (b) C1, (c) C2, and (d) C3 sea ice classifications.

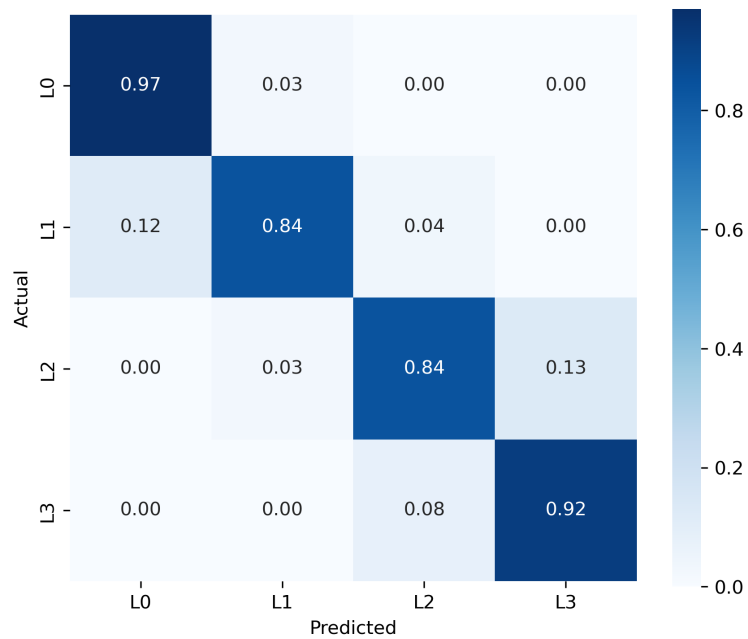


Figure 7. Confusion matrix of the predictive performance of the final CNN model used in the analysis. The model accuracy was estimated by testing the model on a subset of 15% of the images that were randomly held out of the training dataset. The class accuracy is shown along the diagonal. Notably, no examples in the testing data have been misclassified to an ice level that is more than 1 step away. For example, no C0 are classified into C2 or C3, no C2 are classified as C0, etc

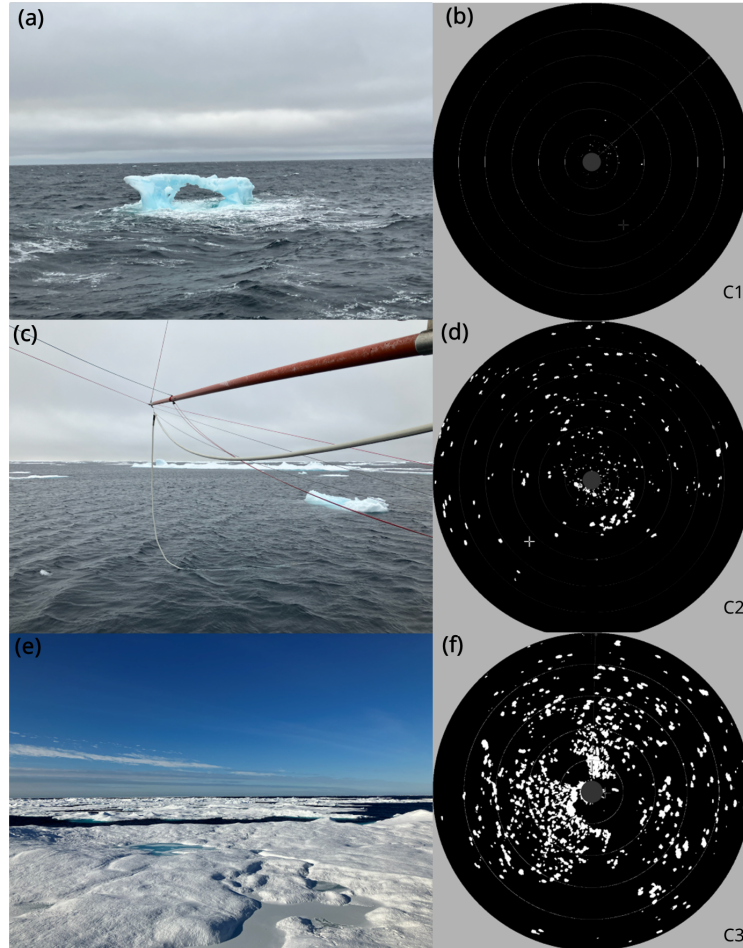


Figure 8. Example collocated photographs (left column) and associated L4 Geo-TIFF S-band radar images (right column) used for validating the model results showing (a,b) C1, (c,d), C2, and (e,f) C3 sea ice classifications.

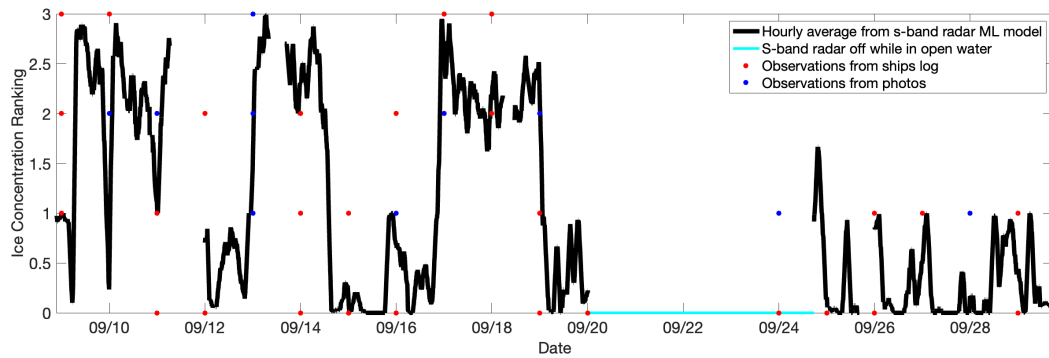


Figure 9. Hourly average ice classification (black line) plotted with ice classifications (0-3) assigned to times based on photos from the cruise and notes from the ship's log. Some of the time stamps on notes from the ship's log may be slightly off due to rounding of times when recording.

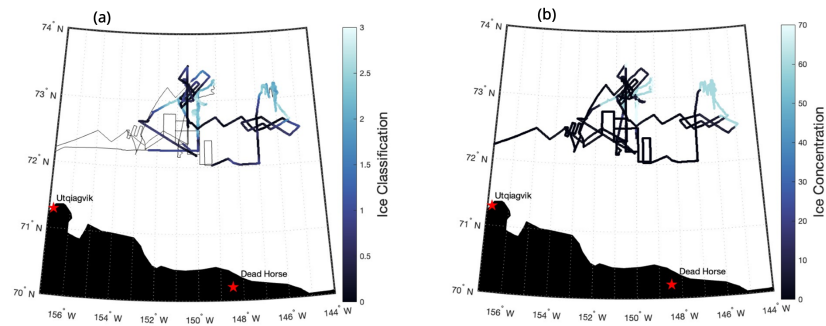


Figure 10. Maps of the track of the *R/V Woldstad* during the SASSIE campaign colored by hourly average ice classification from the S-band radar model estimates (a) and the ice concentration from grASIP (b). Grey ship track to the west in (a) are when the S-band radar was turned off.

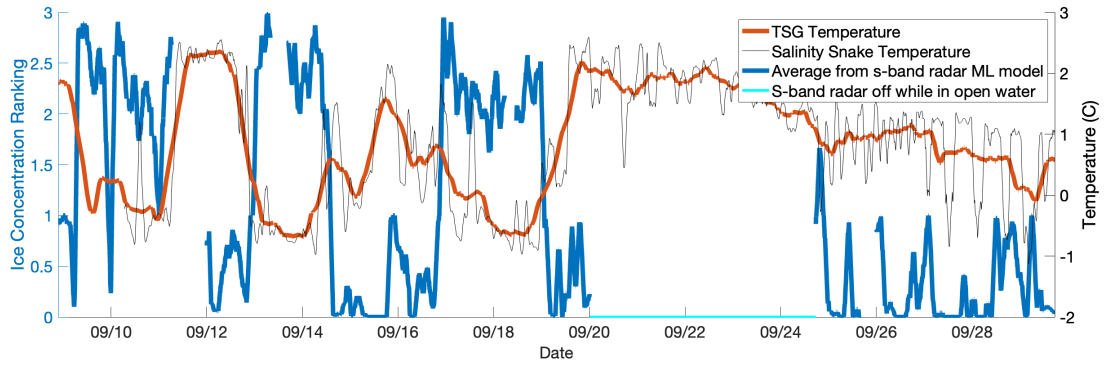


Figure 11. Temperature recorded by the shipboard TSG and Salinity Snake throughout the SASSIE field campaign.

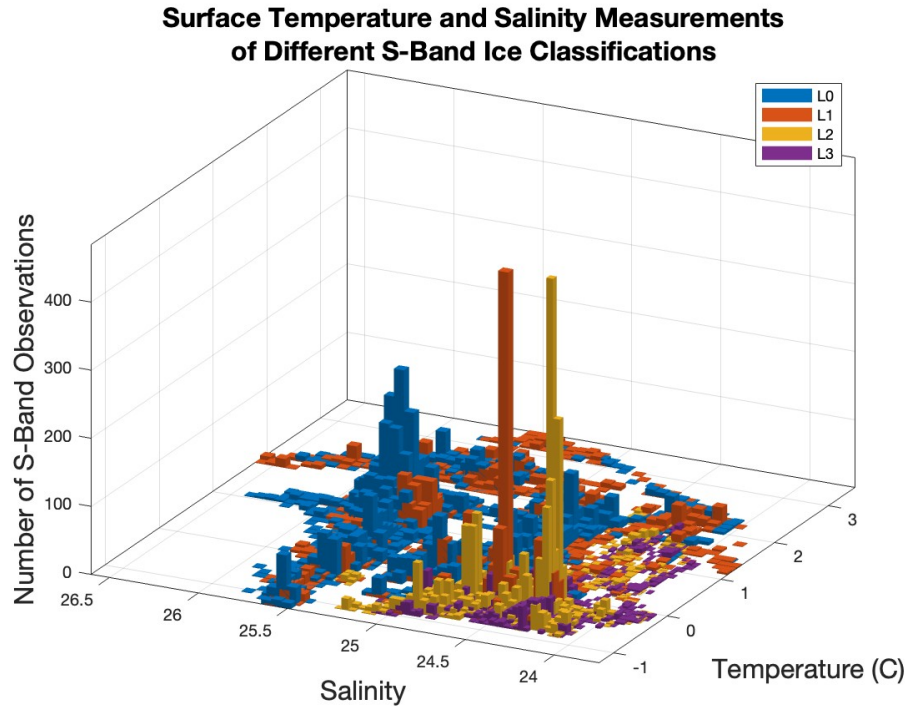


Figure 12. The number of measurements recorded from each ice classification in temperature and salinity bins as a 2D histogram. Bar color indicates sea ice classification with the z-axis indicating the number of observations in a given temperature and salinity bin. Bin size was chosen to ensure the representation of the full dynamic range of the data and to be able to differentiate between individual bins.

Table 1. Table showing the model accuracy at predicting a subset of the training data withheld for validation.

Classification	Accuracy (%)
C0	97
C1	92
C2	84
C3	84