https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

Automatic detection of Arctic polynyas using hybrid supervised-unsupervised deep learning

Céline Heuzé[1] and Carmen Hau Man Wong[1]

[1]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden

Correspondence: Céline Heuzé (celine.heuze@gu.se)

Abstract. Polynyas are areas with no- or thin-ice within the ice pack. They play a crucial role for the Earth system, from deep water to cloud formation, causing large gas exchanges, and acting as hotspots for marine life. Yet their monitoring in the Arctic is challenging because polynya detection is non-trivial, owing to the Arctic's complex geometry. Recently, a labelled dataset was released in which daily winter Arctic sea ice concentration since 1978 was turned into a polynya mask. After oversampling to reduce the class imbalance from 0.1 to 2.5%, we use this labelled dataset to train a Unet autoencoder to detect polynya pixels in daily sea ice concentration images. We further filter out the false positives in the marginal ice zone using an unsupervised Gaussian Mixture Model classifier. False negatives are virtually absent from 2012 onwards, when noise in the labelled dataset is reduced by combining ice concentration and thickness masks. False positives exhibit a significant trend with time and anticorrelation with the Arctic sea ice area. Coupled with our expert assessment of individual images, we argue that most "false positives" are in fact correct, detecting patterns of reduced ice within the changing, more unpredictable ice cover that the rigid traditional methods with fixed thresholds cannot identify. We also successfully apply our trained model to detect polynyas in daily and monthly climate model output at low computing costs. As Arctic sea ice continues to decrease, pushing traditional methods to their limits, we expect such machine-learning methods to become the norm.

1 Introduction

Polynyas, small openings in the pack ice, play a crucial role for the climate and ecosystem. Winter polynyas in particular, by exposing the warm ocean to the frigid atmosphere, act as "ice factories" (Smith Jr and Barber, 2007), with just a few polynyas being responsible for most of the sea ice production in the Arctic (Tamura and Ohshima, 2011; Preußer et al., 2016). The intense air-sea interaction also induces a strong heat- and moisture- loss to the atmosphere (Morales Maqueda et al., 2004; Zhou et al., 2023) that results in cloud formation locally (Monroe et al., 2021) and even impacts the large-scale atmospheric circulation (Gordon et al., 2007). In the ocean, the brine rejected during ice re-formation in the polynya causes deep water formation (Martin and Cavalieri, 1989; Ohshima et al., 2016). The combined effect of ocean mixing and air-sea exchanges further leads to intense gas exchanges and nutrient upwelling (Else et al., 2013; Marchese et al., 2017), making polynyas hotspots for marine life (Moore et al., 2023; Golledge et al., 2025).

Automatically detecting polynyas around Antarctica is comparatively easy, due to the zonal distribution of ocean - sea ice - land (e.g. Mohrmann et al., 2021). The Arctic is way more tedious. Just in the Nordic Seas, to the west sea ice can extend

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

to the southern tip of Greenland while to the east not begin until north of Svalbard. The land geometry is also more complex, with polynyas often forming in the lee of the region's many islands (Smith Jr and Barber, 2007). Finally, Arctic polynyas are much smaller than their Antarctic counterparts (Morales Maqueda et al., 2004). The result is that few real pan-Arctic polynya studies exist; instead, researchers tend to focus on a series of regions with recurrent polynyas (e.g. Tamura and Ohshima, 2011; Preußer et al., 2016). In these regions, they then use a fixed threshold in sea ice properties to distinguish polynyas from the pack ice. This threshold is chosen somewhat arbitrarily, ranging from 30% (e.g. Tamura and Ohshima, 2011) to more than 60% (e.g. Monroe et al., 2021) for sea ice concentration and 10 (e.g. Martin et al., 2004) to 30 cm (e.g. Smedsrud et al., 2006) for sea ice thickness. As throughout the Arctic sea ice concentration and thickness are dramatically decreasing with climate change (Stroeve and Notz, 2018; Jahn et al., 2024), it is unclear whether and for how long these thresholds will remain valid. There is an urgent need to try novel approaches, for full pan-Arctic automatic detection.

Enter machine learning. As recently reviewed by Bracco et al. (2025), examples abound of applications of machine learning methods to climate science, including in the polar oceans (Sonnewald et al., 2021). The most famous example for Arctic sea ice probably is IceNet (Andersson et al., 2021), a series of Unet networks that can forecast Arctic sea ice extent up to 6 months in advance. A Unet network consists of two convolutional neural networks (CNN), i.e. neural networks that identify patterns in blocks of an image. The first CNN, named the encoder, "zooms" on the image, while the second one, the decoder, "zooms back out" so that the result of Unet is a pixel-wise classification. Originally designed for medical applications (Ronneberger et al., 2015), this architecture has proven particularly suited to so-called anomaly detection, i.e. finding a rare event occurrence among a majority of normal data points, with applications ranging from detecting image forgery (Choudhary et al., 2024) or electricity theft (Aslam et al., 2020) to pest infestation in forests (Ye et al., 2022). Unet is a type of supervised classifier, which means that it needs labelled data to learn from. Wong et al. (2025) recently produced such a labelled dataset for winter Arctic polynyas since 1978, at a daily resolution. We here leverage this dataset and use it to train a Unet network to automatically detect Arctic polynyas. Following Liu et al. (2025)'s work on lead detection in the Arctic, we improve our results by combining Unet with a Gaussian Mixture Model (GMM), an unsupervised classification method widely used in climate science (e.g. Jones et al., 2023; Paçal et al., 2023; Poropat et al., 2024).

The objective of this manuscript is dual: First automatically detecting Arctic polynyas in satellite observations using hybrid supervised-unsupervised machine learning, then briefly investigating whether our new method can be applied to climate model data. We present the satellite and model data in section 2, in which we also describe the Unet (section 2.3) and GMM (section 2.4) models. We then evaluate and discuss their application to observations (section 3.1) and to climate model output (section 3.2). We summarise our findings and make final remarks in the conclusions.

2 Data and Model architectures

We here present the satellite- and model-based sea ice data used in this study, and the processing that was performed on them before they could be passed to our two machine-learning models, also described here.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

2.1 Observed and modelled sea ice data

The objective of this paper is the automatic detection of polynyas from individual Arctic sea ice concentration (SIC) maps. Our input data are the daily SIC from the National Snow and Ice Data Center (NSIDC). Briefly, the daily SIC is obtained from satellite microwave radiometers processed with the NASA Team algorithm (Cavalieri et al., 1999). We use the daily data from November 1978 to April 2023 at a 25 km resolution. We generated the daily sea ice area and sea ice extent from this timeseries, using the common threshold of 15% concentration for the sea ice extent and no threshold for the area.

For supervised learning, a labelled dataset is required. Here we use the daily polynya mask produced by Wong et al. (2025), based on the NSDIC SIC described above and hence on the same spatial grid as our input data. Wong et al. (2025) also produced a second mask, based on the Soil Moisture Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) derived sea ice thickness (SIT Paţilea et al., 2019). Briefly, Wong et al. (2025) detect the pack ice (SIC > 0% or SIT > 0 cm) using a flood-fill algorithm, seeded in the Atlantic and Pacific oceans. Within the pack ice north of 65°N, they then identify polynyas as any pixel with SIC $\leqslant$ 50% or SIT $\leqslant$ 20 cm. On the masks, the pixel value is 0 if that pixel is not a polynya, and 1 if it is.

We use their SIC mask for training, validation, and testing our model. We use their SIT mask to better analyse our false positives (model says there is a polynya, but the mask says there is none) and false negatives (model says there is no polynya, but the mask says there is) after 2012, i.e. when SIT data became reliable. The two masks are available at a daily resolution but only for the winter months, November - April included. We therefore also limit our analysis to these months. The dataset hence consists of 7300 daily images.

Finally, to demonstrate the wider applicability of our method, we apply it to the global coupled model MIROC6 (Tatebe et al., 2019). We chose this model for two reasons: it provides both daily and monthly sea ice concentration (“siconc”); and the brief polynya analysis of Heuzé et al. (2023) revealed that it features large and small polynyas in the Arctic. The model’s nominal horizontal resolution for the sea ice grid is 1/3°, which in the Arctic is approximately 30 km. We used the last 30 years of SIC from its historical run, i.e. 1 January 1985 - 31 December 2014.

2.2 Input preparation and oversampling

The original NSIDC grid has 448 by 304 pixels. It extends far south into the open Pacific and Atlantic oceans, and includes many land pixels (orange on Fig. 1). For the unsupervised detection of the open ocean described at the end of this section, we keep the grid as it is. But for most of the work, we instead use a reduced grid with 192 x 192 pixels, chosen so that it includes the entirety of the Arctic Ocean (blue on Fig. 1). On both grids, we set all NaNs to 1, i.e. SIC = 100%.

We originally wanted to preserve the time dimension, assuming that the model would learn from the day-to-day polynya formation and shape evolution. Unfortunately, polynyas are rare events, so on each daily image on average polynya pixels covered less than 0.05% of the image, barely rising to 0.1% after image size restriction. After unsuccessful training, we reluctantly gave up on the time dimension and instead created an oversampled training dataset of 1000 scenes consisting of a random mix of:

– 100 scenes that combine 500 randomly selected daily images;

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

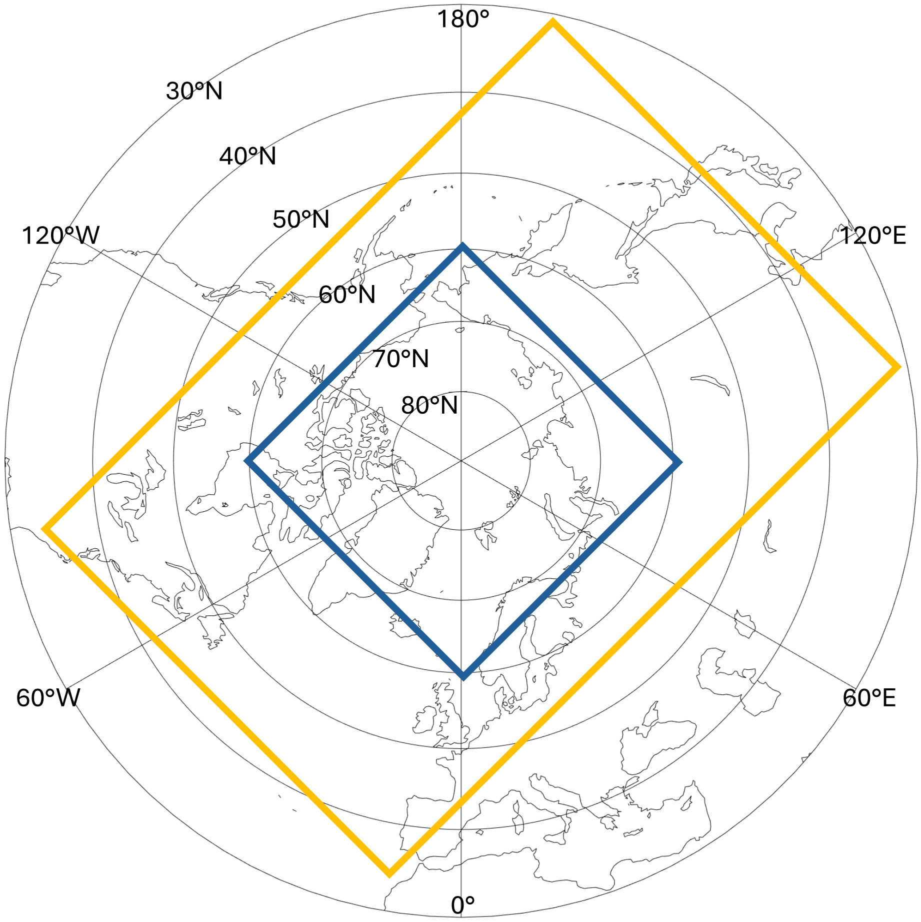

Figure 1. The two stereographic grids used in this study: The original 448 x 304 pixel grid from NSIDC used for the unsupervised network is shown in orange; the reduced 192 x 192 pixel grid used for the supervised learning, in blue. For increased readability, all subsequent figures will show only the rectangular/square regions with pixel data, rotated.

- 200 scenes that combine 200 randomly selected daily images;
- 400 scenes that combine 100 randomly selected daily images;
- 200 scenes that combine 50 randomly selected daily images;
- and 100 scenes that combine 10 randomly selected daily images.

95 The polynya mask was set to 1 on the oversampled scene if the grid cell was 1 on at least two of the original images. Visual inspection (not shown) revealed that two images was the minimum required to reduce noise. Where this new, oversampled polynya mask was set to 1, the sea ice concentration was set to the mean sea ice concentration of only the images where the original polynya mask was 1. The sea ice concentration elsewhere was set to the mean of all images elsewhere. Visual

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

inspection (not shown) revealed that the transition polynya/not polynya looks more natural on these oversampled scenes if prior to averaging we set all concentrations lower than 15% to 0.

The resulting oversampled dataset has an average of 2.5% polynya pixels per scene. This is still a large class imbalance, but it is on par with what most anomaly detection algorithms can manage (Krawczyk, 2016; Ghosh et al., 2024).

2.3 Supervised binary classification polynya / not polynya with a Unet autoencoder

Using Keras (Chollet and The Keras Team, 2015), we implemented a Unet autoencoder for the supervised classification of our highly imbalanced dataset. See schematic on Fig. 2:

- The encoder is made of two levels with two 2D-convolutional layers of filter size 16 (first level) then 32 (second level) and kernel size of 3 with a relu activation, followed by a 2D maxpooling with a pool size of 2 by 2;
- The bottleneck on the third level is made of two 2D-convolutional layers of filter size 64, also with a kernel size of 3 and a relu activation;
- The decoder is made of two levels of upsampling done by conv2Dtranspose of filter size 32 (second level) then 16 (top level), kernel size of 2, and 2 strides; a concatenation; and two 2D-convolutional layers of filter size 32 (second level) then 16 (top level) and kernel size of 3 with a relu activation. The output layer is a final 2D convolution of filter size 1, kernel size 1, and a sigmoid activation.

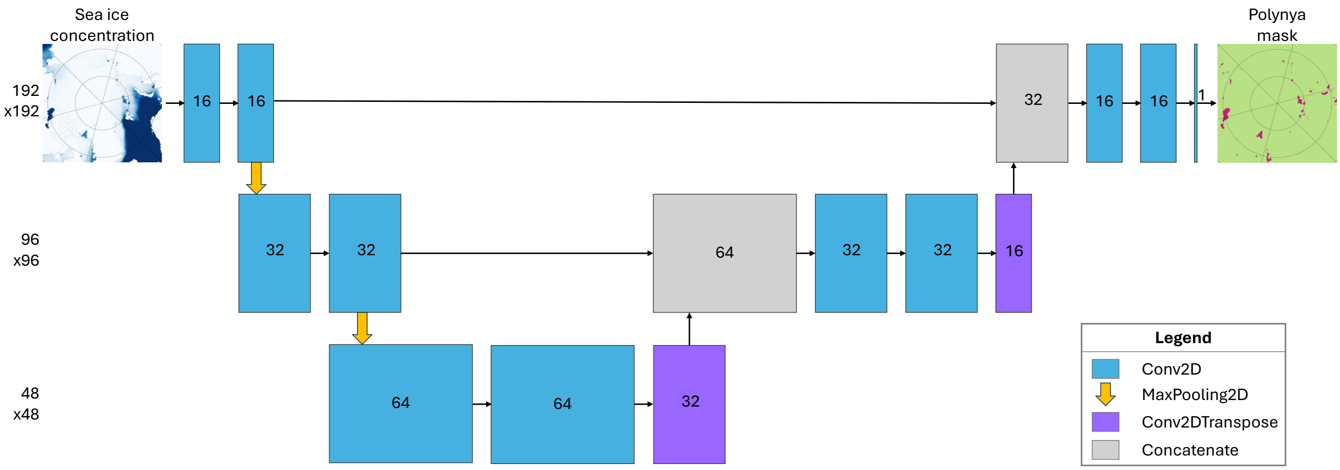

Figure 2. Schematic architecture of our Unet implementation. Numbers indicate the output dimensions. See text for the settings of each layer.

Due to the class imbalance, we used a weighted binary cross entropy as loss function and monitored the precision and F1 score. We used an Adam optimiser. The oversampled set was randomly split between 70% for training, 15% for validation and 15% for testing. The best model was chosen based on its performance on the validation set.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

We varied the weight from 0 to 100 with an interval of 5, and obtained the lowest number of false positives with a weight of 15 - the number of false negatives was not significantly affected. Similarly, we varied the batch size from 2 to 40, and obtained the fewest false positives with a batch size of 5. For the F1 score, the classifier is so certain of its results that varying the threshold had no impact on the performance; we kept the default threshold of 0.5. For each setting, we trained an ensemble of 100 models.

The final performances, on the validation set, are an F1 score of 0.98 and precision of 0.97. We obtained 438 false negatives and 3 943 false positives, compared to 123 283 total polynya pixels in the oversampled validation set.

2.4 Filtering via unsupervised detection of the ocean and MIZ by Gaussian Mixture Models

With our Unet autoencoder, we obtain very few false negatives but a large number of false positives. Rather than modifying the architecture of the network, we filter its daily output using a second, unsupervised machine learning method: clustering using Gaussian Mixture Models. We use the Scikit-learn implementation (Pedregosa et al., 2011), with 3 components, a tolerance of 0.001, up to 200 iterations, and 3 initialisations using kmeans.

The setting that had the most dramatic impact of the performances was the number of components, i.e. how many different classes the model should detect. We varied the number from 3 to the 80 used by Liu et al. (2025). Three was chosen intuitively, assuming the network would split the data into open ocean, land, and pack ice classes. We expected the best performances for 4 or 5 classes, assuming the marginal ice zone (MIZ) would be a class of its own, but saw that for any class number larger than 3, polynyas and MIZ were bundled together in the same class. We therefore used 3 classes.

The input data was the original 448 x 304 pixel dataset, in order to maximise the amount of open ocean pixels the network could learn from. Prior to training, we set the SIC lower than a given threshold to 0, to merge the MIZ and open ocean into one class. We tested thresholds from 0 to 60%, starting from 0 and increasing by 5%, feeding the GMM 2 to 10 images at a time, and obtained the best agreement with Wong et al. (2025)'s masks with a threshold of 40% for 2 days. Two days is the minimum the GMM can take, and probably yielded the best results because the MIZ changes very rapidly. The higher the SIC threshold, the better the MIZ was filtered out; however, for any threshold larger than 40% polynyas started being detected as belonging to that open ocean / MIZ class, which is also why we stopped testing at 60%. For the modelled data, 40% yielded satisfactory results, but owing to model biases, 15% was the best compromise to remove the MIZ without removing polynyas as well. When working with monthly modelled data, the GMM was fed two months at a time; for daily modelled data, 2 days, as for the observations.

For each timestep, we ran an ensemble of 10 models. Since class numbers are randomly affected (see example on Fig. 3), we identify for each ensemble member the class number corresponding to the middle of the North Atlantic Ocean (median class number in the white band on Fig. 3, approx 48 - 59N and 47W - 10W). We then mark each pixel as "open ocean or MIZ" if strictly more than 5 of the 10 ensemble members put it in the open ocean class. We chose to detect the open ocean class rather than the pack ice because although the majority of the classification results looked like the example of Fig. 3, we noticed while randomly checking individual timesteps that there were occasions where half of the pack ice was detected as land. Only the

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

open ocean / MIZ was consistently its own class. Besides, the pack ice class is noisy, detecting “ice" off the southeast coast of England for example on Fig. 3.

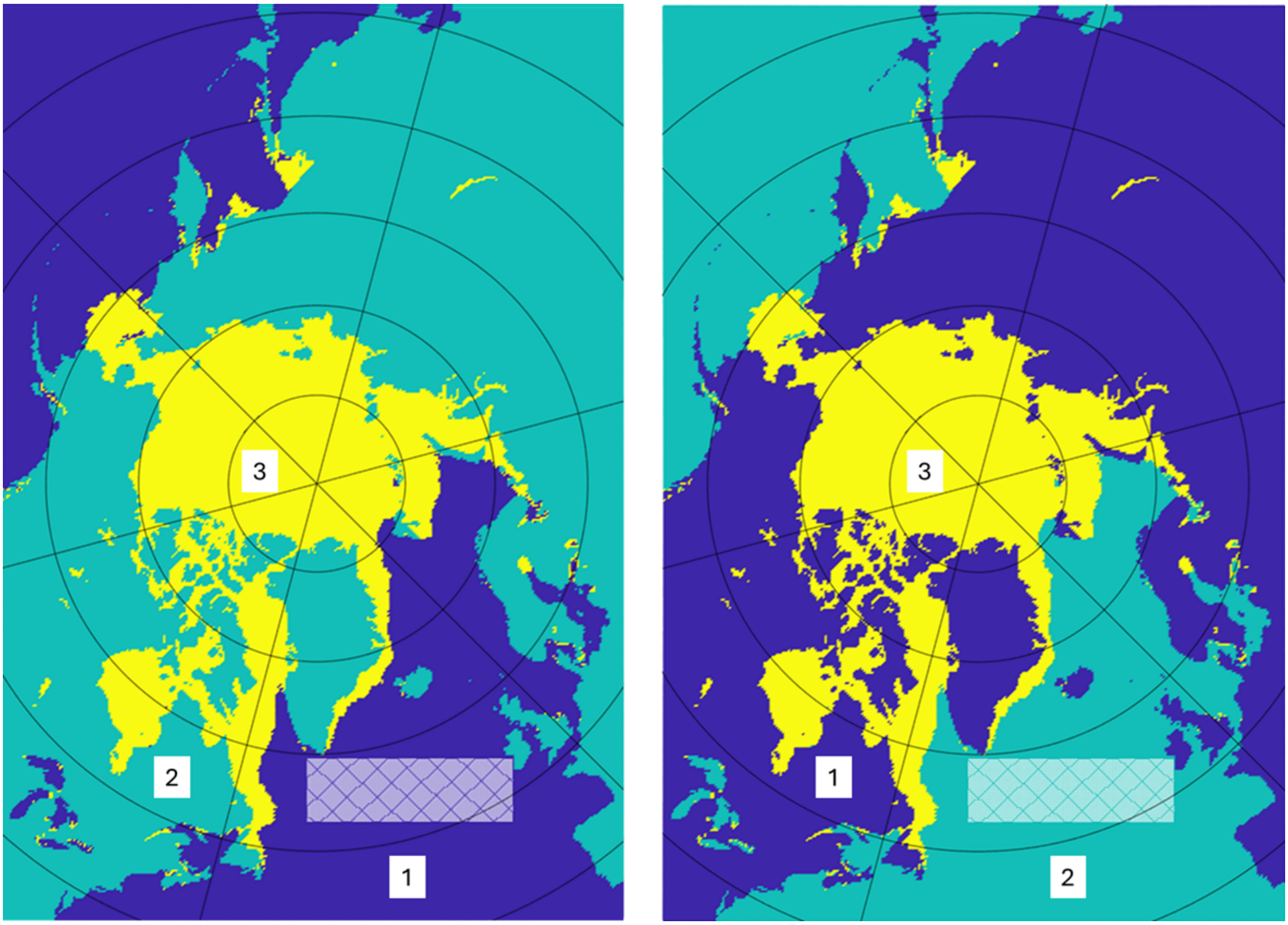

Figure 3. Two runs of the GMM unsupervised classification on the same random timestep (29 April 1984), illustrating that classes at the same location can randomly be affected different numbers: Here classes 1 (dark blue) and 2 (green) are swapped. The white band over the North Atlantic is that used to detect the open ocean / MIZ class number.

Any pixel that our Unet autoencoder said was a polynya but was found to be in this open ocean / MIZ class was set to 0 or “not a polynya". The effect of this filtering is quantified in the next section.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

3 Results and Discussion

3.1 Detection of Arctic winter polynyas in sea ice concentration observations

After training, we fit the best model to the original 7300 daily sea ice concentration images and compare the model's classification to the SIC polynya mask of Wong et al. (2025). The performances are obviously less impressive than on the validation, oversampled set (Table 1, first column): We obtain 1.4 million false positive pixels, which is 5 times the number of true positive pixels in the dataset (262 067). After applying the GMM-produced open ocean / MIZ filter (Table 1, second column), the number of false positives is more than halved and the already-low number of false negatives further decreases to just over 10 000.

Metric	Unet only	+ GMM	+ SIT mask
False negatives	12 776	10 048	6 271
False positives	1 475 630	601 633	630 011

Table 1. Number of false negatives (missed polynya pixels) and false positives (pixels incorrectly identified as a polynya) after the three main steps of the pipeline: Straight out of the supervised classification (Unet, section 2.3); after applying the Gaussian Mixture Models-produced mask (GMM, section 2.4); and after correcting the original sea ice concentration mask by merging it with the sea ice thickness one (SIT). For comparison, there are 107 578 polynya pixels in the dataset after combining the sea ice concentration and thickness masks, and just over 269 million pixels in total.

A visual inspection of our results made us suspect noise in the masks as the main reason for the number of false negatives. To reduce this noise, we combine the SIC and SIT masks of Wong et al. (2025) from 2012 onwards, i.e. when SIT became widely available in the Arctic. The effect is striking: The number of false negatives drops to near-zero after 2012 (Fig. 4a).

The number of false positives in contrast seems to have an increasing trend with time, even after applying the filter (Fig. 4b). In fact, we do identify a trend significant at 99% of +277 false positive pixels per year, which corresponds to +12 763 pixels over our 46-year timeseries. Visual inspection (e.g. Fig. 5) further suggests that the years with a reduced ice cover coincided with the years with most false positives. This is confirmed by a correlation analysis: with 99% significance, the daily number of false positives and the daily sea ice area (SIA) / sea ice extent (SIE) are strongly anticorrelated (R=-0.76 / -0.71), i.e. the less sea ice, the more false positives. Fitting a regression line through this result, this gives -50 / -57 false positive pixels per million km^2 of Arctic sea ice. As pictured on Fig. 5b, the correlation holds with yearly values too, reaching -0.78 / -0.80 between the yearly false positives and the yearly (winter only) minimum SIA / SIE.

The visual inspection also shows that many of these false positives seem correct. We show two examples on Fig. 5, and first discuss the areas circled in green. For both dates, although our method (magenta) correctly identifies the same polynyas as in the labelled masked (green, hereafter referred to as the traditional method), the polynyas detected by our method are larger. When visually comparing to the original sea ice concentration data, we agree more with our method than with the traditional one. The area circled in blue on Fig. 5b is a special case: it was also detected as a polynya by Wong et al. (2025) but has been

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

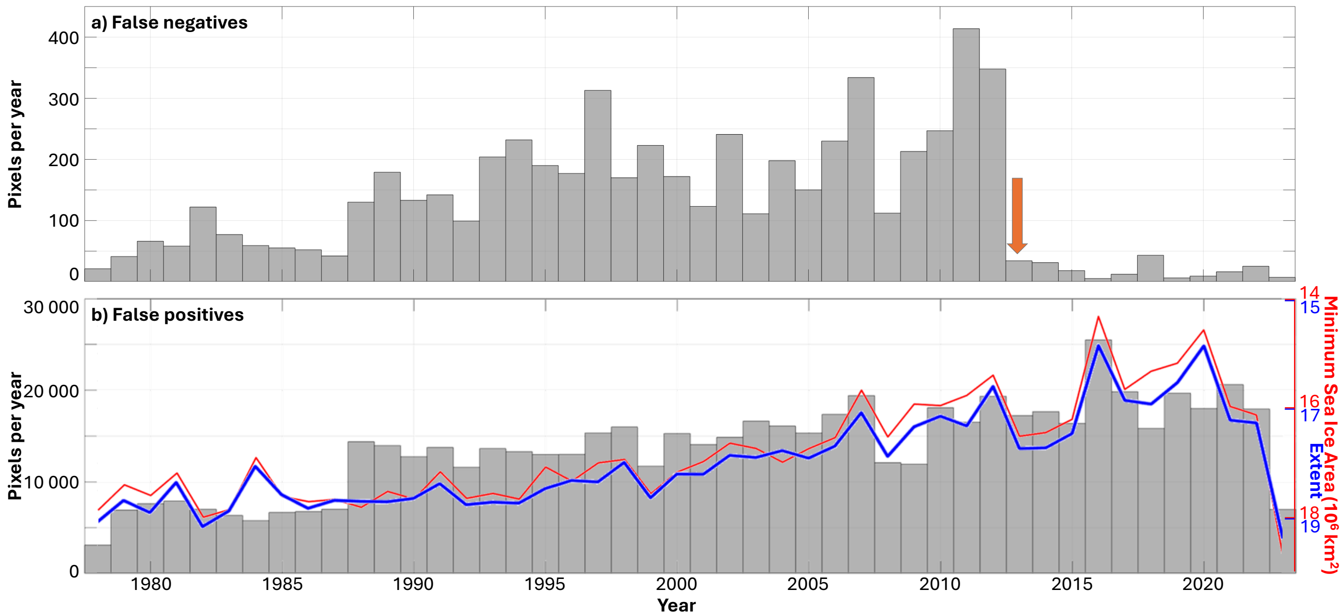

Figure 4. a) False negatives and b) false positives for each year, in number of pixels. Note the different y-axes. The orange arrow on a) highlights the beginning of reliable sea ice thickness observations, and therefore the effect of merging the SIC and SIT labelled datasets. The red and blue lines on b) are the minimum sea ice area and extent, respectively (winter months only), flipped vertically to highlight the correlation. Note that 2023 stops in April, hence the higher sea ice values.

removed during their post-processing since it is in a river. Hence technically, our two methods agree there as well. Moving to the areas where the traditional method found no polynya at all, circled in black on Fig. 5, our visual inspection reveals that these could qualify as a polynya, if one defines a polynya as an area where sea ice concentration strongly decreases compared to the surrounding ice. We do acknowledge that our method is not perfect: the MIZ mask absorbs some of these areas, if the surrounding ice itself has a comparatively low concentration, but leaves out as definitely false positive some areas with complex MIZ. Nonetheless, our results highlight the limitation of the traditional way of detecting polynyas using a fixed threshold, and probably explains why so many different sea ice concentration thresholds are used in the literature (30 to more than 60% in e.g. Tamura and Ohshima, 2011; Campbell et al., 2019; Mohrmann et al., 2021; Zhou et al., 2022; Bennett et al., 2024). Since convolutional neural networks detect shapes and gradients, they are better suited to this pattern recognition exercise. The increase of not-that-false positives with increasing climate change also suggests that as sea ice has entered a new regime (Stroeve and Notz, 2018; Sumata et al., 2023) and become more variable (Dörr et al., 2023), polynyas become harder to predict with traditional methods whereas the machine learning methods are not affected.

Admittedly, Arctic polynyas are more often detected using sea ice thickness since they can be covered with a thin layer of high concentration ice (e.g. Martin et al., 2004; Smedsrud et al., 2006; Tamura and Ohshima, 2011; Adams et al., 2013; Preußer et al., 2016; Ren et al., 2022). Since sea ice thickness data are only available since 2010s and that we have to combine images

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

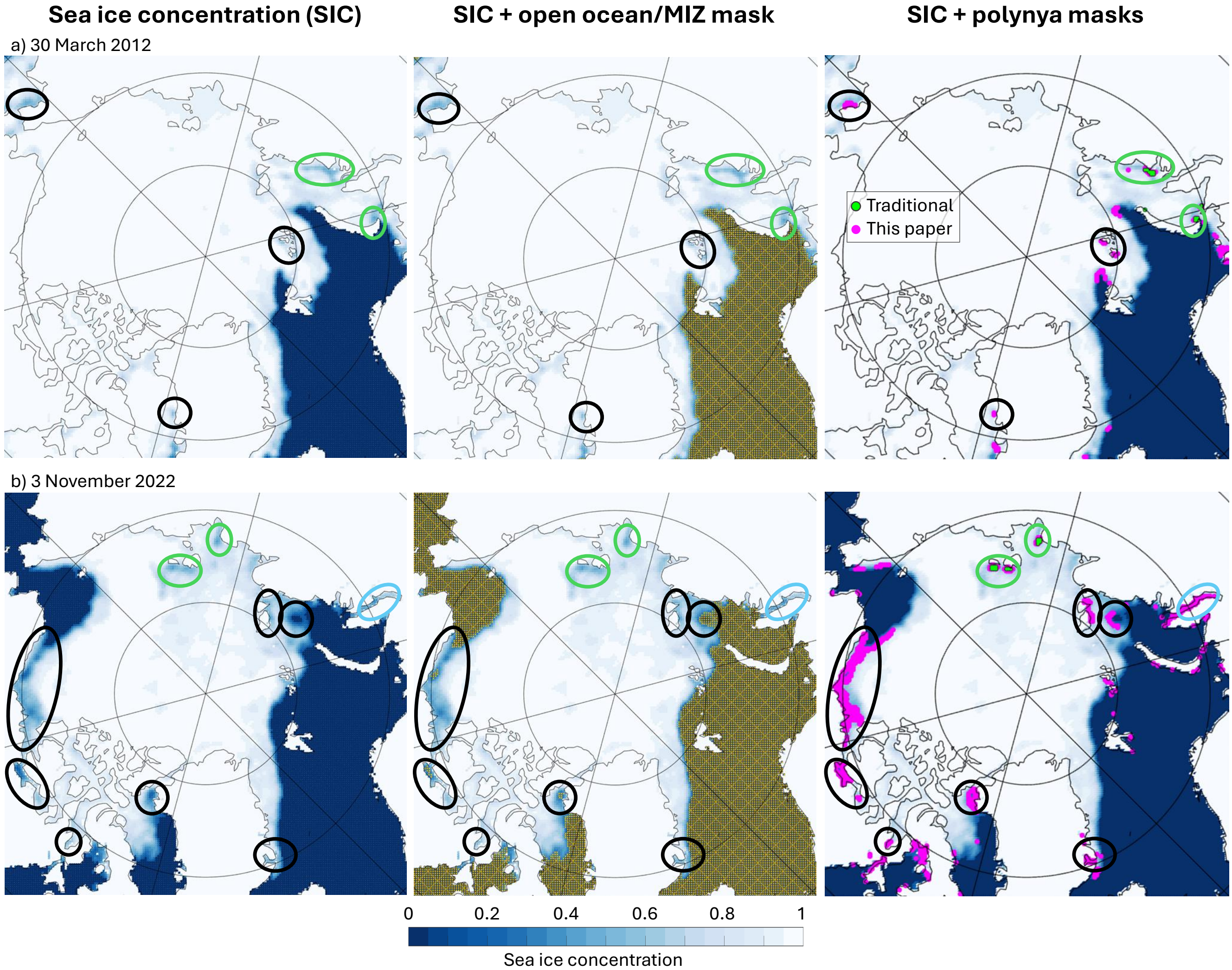

Figure 5. Two example results, on a) 30 March 2012 (top) and b) 3 November 2022 (bottom), of the original daily sea ice concentration data (left), with superimposed the resulting GMM-produced open ocean / MIZ mask (middle, orange asterisks) or the polynya masks (right, green and magenta). Green dots are the polynyas detected with a traditional threshold method, i.e. the labelled dataset we used for training, and magenta dots are the polynyas detected with our method. Green circles highlight areas where the two methods agree. Black circles highlight areas with false positives that we argue are correct, while the blue circle on b) is a special case, see text.

to create the oversampled training dataset, we could not train our model on SIT without grossly overfitting it. It should however be possible to perform transfer learning with minimum retraining, by min-max normalising the SIT so that it too has values
195 between 0 and 1, and most crucially finding an adequate threshold for Unet and the GMM. Similarly, we limited our study to winter polynyas because the labelled data were only available for the months November to April, but we see no reason why

our method would not function to detect summer polynyas from SIC as well, potentially after minor retraining or threshold adjustment for the GMM.

3.2 Application: Detection of winter Arctic polynyas in the global coupled model MIROC6

The ultimate objective of this exercise was for us to develop a method for automatically and rapidly detecting Arctic polynyas in modelled sea ice concentration output. We therefore test our scripts on one global coupled model which we know has winter Arctic polynyas: MIROC6. We do not have any truthing data for it; all assessment is based on visual inspection. We inspected every single monthly result, and a random quarter of the daily ones. As exemplified on Figs. 6 and 7, our model performs well, and equally well on the two temporal resolutions. As with the observations, the occasional false positive pixel survives the open ocean / MIZ filtering, but polynyas big and small are successfully detected. The fact that it works so well even on monthly data is surprising, given that both the brief and extensive studies of Heuzé et al. (2023) and Mohrmann et al. (2021), respectively, concluded that polynya detection on monthly modelled data is not recommended.

If anything, our visual inspection reveals that our results (magenta dots) are somewhat conservative; we would have considered flagging as polynya the regions on the Siberian shelf with reduced SIC of around 70%. Besides, to directly apply our trained model, we had to interpolate the modelled data onto the NSIDC polar stereographic grid and clearly introduced some interpolation artefacts (see the top left corner of all the panels, Figs. 6 and 7). One alternative could be to use only the unsupervised GMM to detect the open ocean / MIZ on the model's native grid, instead of the tedious traditional flood-fill, and then detect polynyas using a standard threshold method. This could be particularly advantageous if working with many models, since they have different sea ice concentration biases (Notz and Community, 2020). The one inconvenient is the computing time: it takes less than a minute to apply the already-trained Unet to 30 years of daily data, but more than 15h on our supercomputer to obtain the classes from the GMM. The fastest approach of them all would then be to do the opposite and use only our trained Unet, and minimise the false positives by analysing pre-determined sub-regions in the pack ice, away from the MIZ, as done on observations by e.g. Tamura and Ohshima (2011) or Preußer et al. (2016).

We acknowledge that MIROC6 has one of the highest resolutions of all models that participated in the Climate Model Intercomparison Project phase 6 (CMIP6, Eyring et al., 2016), and has been identified in the all-model assessments of monthly and daily Arctic sea ice concentration of Athanase et al. (2025) and Heuzé and Jahn (2024), respectively, as among the most accurate CMIP6 models. Our results should be taken as a proof of concept: had our method not worked on MIROC6, it would not have worked on the other CMIP6 models, but one will probably need to exercise caution if working with less accurate models. Daily sea ice thickness was not widely available in CMIP6; this is another reason why we developed our method on concentration instead. As with the observations, if daily thickness becomes routinely available for CMIP7 or if one wants to work with CMIP6 monthly thickness, we argue that our method can be used with minimum adjustment, provided one finds the suitable corresponding thresholds.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

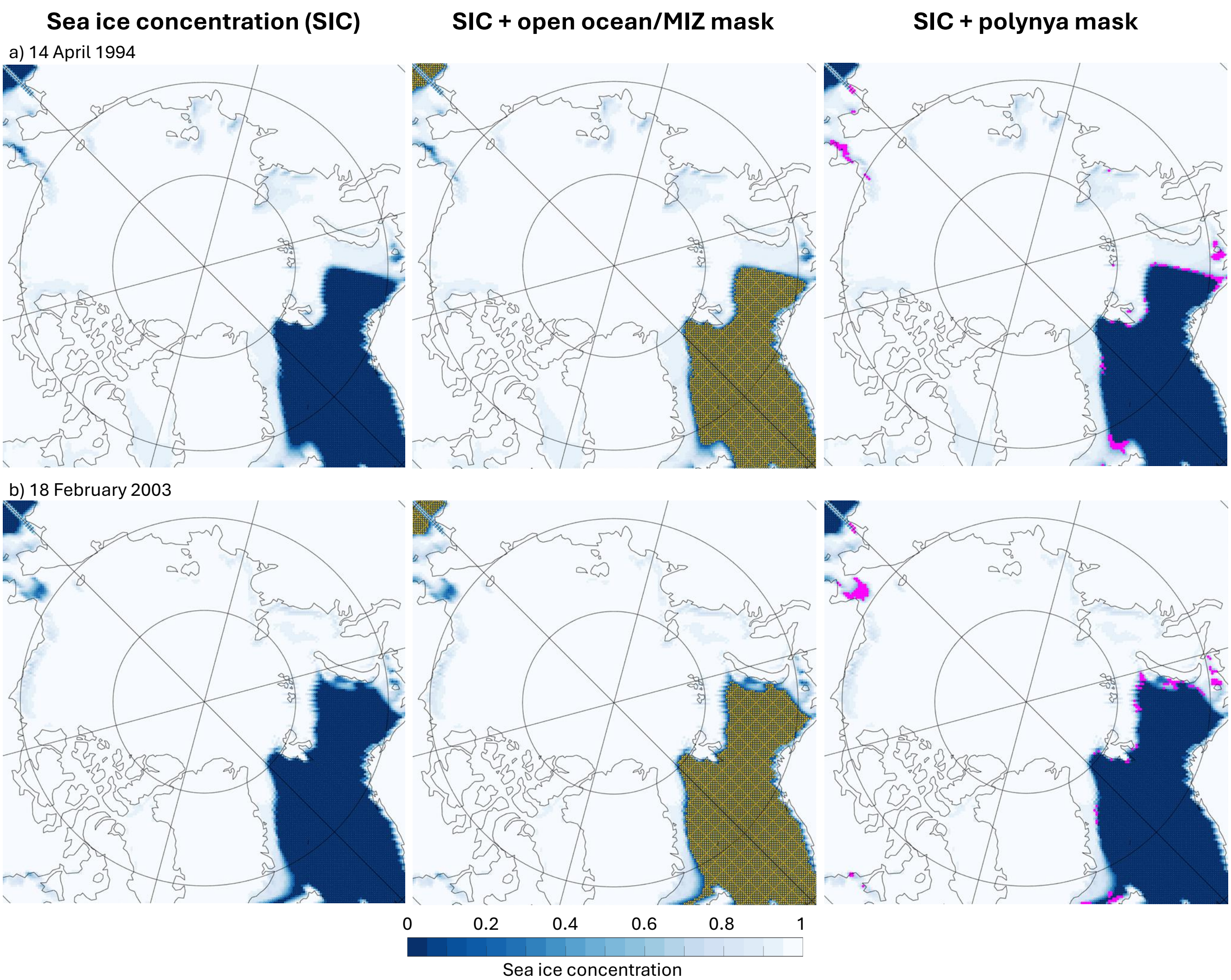

Figure 6. Two example results, on a) 14 April 1994 (top) and b) 18 February 2003 (bottom), of the model MIROC6's daily sea ice concentration data (left), with superimposed the resulting GMM-produced open ocean / MIZ mask (middle, orange asterisks) or the polynya mask (right, magenta).

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

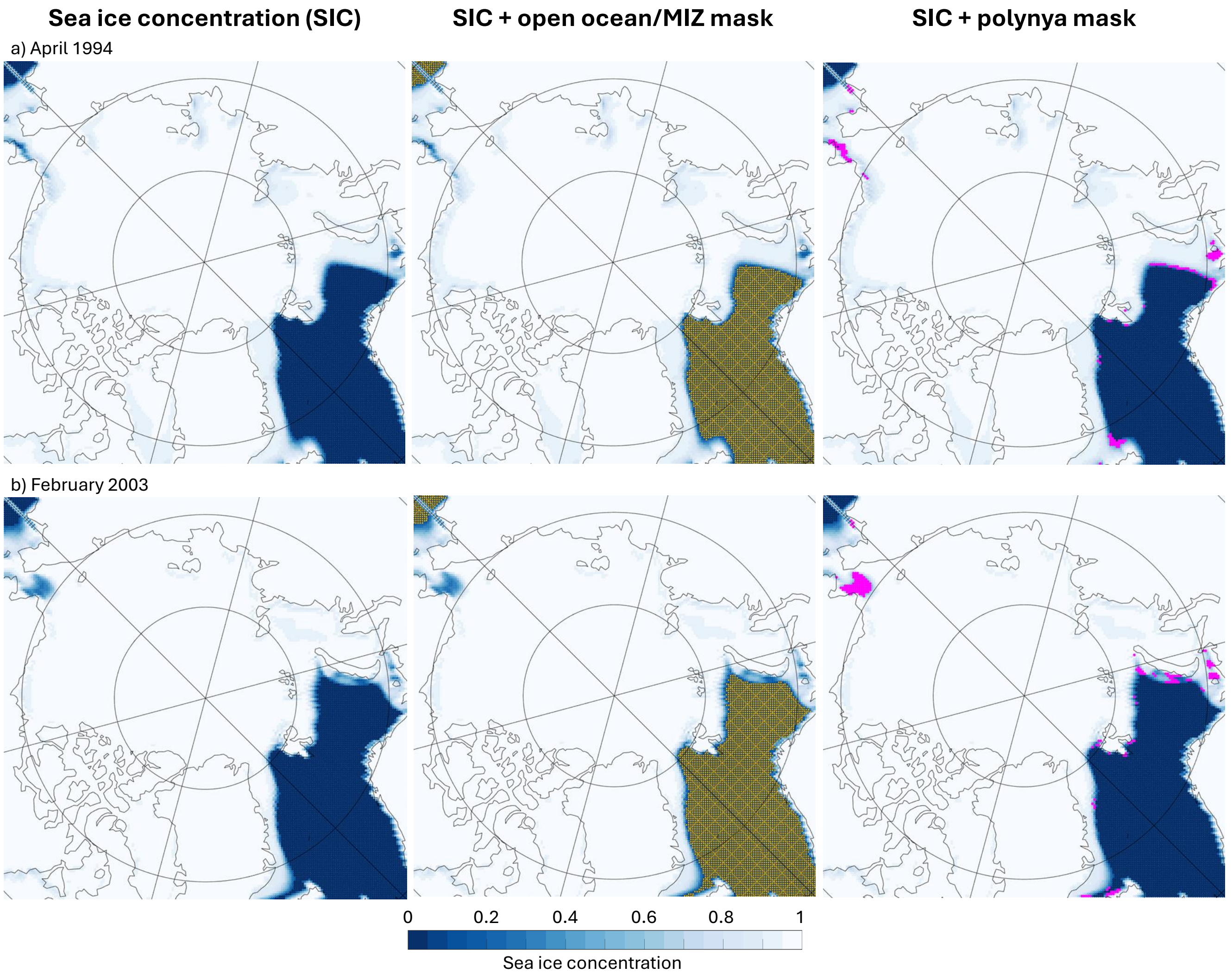

Figure 7. Same as Fig. 6 but using the model's monthly mean sea ice concentration.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

4 Conclusions

We successfully trained a hybrid supervised-unsupervised machine learning model to automatically detect polynyas in the Arctic from daily satellite images of sea ice concentration, despite polynyas occupying barely 0.1% of the image. The supervised network is a Unet autoencoder, trained on the labelled polynya masks of Wong et al. (2025). It returns many false positives in the marginal ice zone, so we filter its results using an unsupervised Gaussian Mixture Model classifier that detects the open ocean / MIZ on each image. By combining the labelled sea ice concentration and thickness masks over the period where they are both available (since 2012), we reduce their individual noise and show that we have virtually no false negative, i.e. all polynyas are detected. Our large number of false positives is to a small extent caused by the occasional complex unfiltered pattern in the MIZ, but is primarily due to the standard definition of a polynya. Our model detects areas with reduced sea ice cover surrounded by more compact ice, which we argue meet the criteria of a polynya, whereas traditional methods use fixed concentration thresholds for detection (e.g. Smedsrud et al., 2006; Tamura and Ohshima, 2011; Campbell et al., 2019). These not-that-false positives significantly increase with time (+277 per year) and are significantly anticorrelated to the Arctic sea ice area and extent (R = -0.76 and -0.71 for daily values), suggesting that as climate change continues impacting the sea ice cover (Stroeve and Notz, 2018), traditional definitions become less adapted, while the more flexible machine learning approaches can still successfully detect shapes and gradients.

As proof-of-concept, we applied our method to detect polynyas in one CMIP6 global coupled model, MIROC6. In the absence of truthing data the validation was visual only but our expert eyes confirmed that it indeed works, successfully removing the MIZ and detecting polynyas large and small on both daily and monthly modelled sea ice concentration. This model has a high resolution, comparable to that of the observational product (Tatebe et al., 2019), and a somewhat realistic sea ice cover (Heuzé and Jahn, 2024); further studies will have to determine whether adjustments are necessary for coarser and/or more biased models. Similarly, we limited this study to sea ice concentration as observational sea ice thickness datasets (Paţilea et al., 2019) are not long enough yet to deal with such high-class-imbalance classification exercise, and to winter only since the labelled data covered only November to April each year (Wong et al., 2025). Nonetheless, this study paves the way for automatic detection at a somewhat low computing cost of Arctic polynyas, small features with such a large role for the Earth System (e.g. Smith Jr and Barber, 2007; Else et al., 2013; Moore et al., 2023).

Code and data availability. The scripts used in this manuscript and the best model we generated are available on Github at https://github.com/cheuze/Polynya_CNN2D_GMM; they will be licensed on Zenodo closer to publication. The observed sea ice concentration data from the National Snow and Ice Data Centre (dataset DOI 10.5067/MPYG15WAA4WX, DiGirolamo et al., 2022) are freely available at https://nsidc.org/data/nsidc-0051/versions/2, last accessed on 5 May 2025. The polynya masks produced by Wong et al. (2025) are in the process of becoming publicly available; in the meantime, they can be accessed at https://www.dropbox.com/scl/fi/fr0barqpe519e1nu6kl3q/daily_location.nc?rlkey=zh1ganta0noexwvvvch4sh3zr&dl=1 and https://www.dropbox.com/scl/fi/9a91cbsgrqj29o5ckb5zq/daily_location_SMOS.nc?rlkey=bdvlvay1k0rkql973u731ku0z&dl=1. The modelled sea ice concentration data from MIROC6 (dataset DOI 10.22033/ES-

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

GF/CMIP6.5603, Tatebe and Watanabe (2018), last accessed 5 May 2025) are freely available on any of the Earth System Grid Federation portals; we used the French one: https://esgf-node.ipsl.upmc.fr/search/cmip6-ipsl/, last accessed on 5 May 2025.

Author contributions. CH designed and conducted most of the study. CHMW obtained the data and produced the labelled, training dataset. CH wrote the original draft. Both authors revised the manuscript.

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

Acknowledgements. This work was funded by the Swedish National Space Agency grant no. 2022-00149 awarded to CH. The model training was enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies that support CMIP6 and ESGF.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

References

Adams, S., Willmes, S., Schröder, D., Heinemann, G., Bauer, M., and Krumpen, T.: Improvement and Sensitivity Analysis of Thermal Thin-Ice Thickness Retrievals, IEEE Transactions on Geoscience and Remote Sensing, 51, 3306–3318, https://doi.org/10.1109/TGRS.2012.2219539, 2013.

Andersson, T. R., Hosking, J. S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., Law, S., Jones, D. C., Wilkinson, J., Phillips, T., et al.: Seasonal Arctic sea ice forecasting with probabilistic deep learning, Nature communications, 12, 5124, 2021.

Aslam, Z., Javaid, N., Ahmad, A., Ahmed, A., and Gulfam, S. M.: A combined deep learning and ensemble learning methodology to avoid electricity theft in smart grids, Energies, 13, 5599, 2020.

Athanase, M., Köhler, R., Heuzé, C., Lévine, X., and Williams, R.: The Arctic Beaufort Gyre in CMIP6 Models: Present and Future, Journal of Geophysical Research: Oceans, 130, e2024JC021 873, https://doi.org/10.1029/2024JC021873, 2025.

Bennett, M. G., Renfrew, I. A., Stevens, D. P., and Moore, G. W. K.: The Northeast Water Polynya, Greenland: Climatology, Atmospheric Forcing and Ocean Response, Journal of Geophysical Research: Oceans, 129, e2023JC020 513, https://doi.org/https://doi.org/10.1029/2023JC020513, e2023JC020513 2023JC020513, 2024.

Bracco, A., Brajard, J., Dijkstra, H. A., Hassanzadeh, P., Lessig, C., and Monteleoni, C.: Machine learning for the physics of climate, Nature Reviews Physics, 7, 6–20, https://doi.org/10.1038/s42254-024-00776-3, 2025.

Campbell, E., Wilson, E., Moore, G., Riser, S., Brayton, C., Mazloff, M., and Talley, L.: Antarctic offshore polynyas linked to Southern Hemisphere climate anomalies, Nature, 570, 1–7, https://doi.org/10.1038/s41586-019-1294-0, 2019.

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, Journal of Geophysical Research, 104, 15,803–15,814, https://doi.org/10.1029/1999JC900081, 1999.

Chollet, F. and The Keras Team: Keras, [package] Github, https://github.com/fchollet/keras, 2015.

Choudhary, R. R., Paliwal, S., and Meena, G.: Image Forgery Detection System using VGG16 UNET Model, Procedia Computer Science, 235, 735–744, 2024.

DiGirolamo, N., Parkinson, C. L., Cavalieri, D. J., Gloersen, P., and Zwally, H. J.: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data. (NSIDC-0051, Version 2), https://doi.org/10.5067/MPYG15WAA4WX, 2022.

Dörr, J. S., Bonan, D. B., Årthun, M., Svendsen, L., and Wills, R. C.: Forced and internal components of observed Arctic sea-ice changes, The Cryosphere, 17, 4133–4153, https://doi.org/10.5194/tc-17-4133-2023, 2023.

Else, B. G. T., Papakyriakou, T. N., Asplin, M. G., Barber, D. G., Galley, R. J., Miller, L. A., and Mucci, A.: Annual cycle of air-sea CO2 exchange in an Arctic Polynya Region, Global Biogeochemical Cycles, 27, 388–398, https://doi.org/https://doi.org/10.1002/gbc.20016, 2013.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Ghosh, K., Bellinger, C., Corizzo, R., Branco, P., Krawczyk, B., and Japkowicz, N.: The class imbalance problem in deep learning, Machine Learning, 113, 4845–4901, https://doi.org/10.1007/s10994-022-06268-8, 2024.

Golledge, N. R., Keller, E. D., Gossart, A., Malyarenko, A., Bahamondes-Dominguez, A., Krapp, M., Jendersie, S., Lowry, D. P., Alevropoulos-Borrill, A., and Notz, D.: Antarctic coastal polynyas in the global climate system, Nature Reviews Earth & Environment, pp. 1–14, 2025.

Gordon, A. L., Visbeck, M., and Comiso, J. C.: A possible link between the Weddell Polynya and the Southern Annular Mode, Journal of Climate, 20, 2558–2571, 2007.

Heuzé, C. and Jahn, A.: The first ice-free day in the Arctic Ocean could occur before 2030, Nature Communications, 15, 1–10, https://doi.org/10.1038/s41467-024-54508-3, 2024.

Heuzé, C., Zanowski, H., Karam, S., and Muilwijk, M.: The deep Arctic Ocean and Fram Strait in CMIP6 models, Journal of Climate, 36, 2551–2584, https://doi.org/10.1175/JCLI-D-22-0194.1, 2023.

Jahn, A., Holland, M. M., and Kay, J. E.: Projections of an ice-free Arctic Ocean, Nature Reviews Earth & Environment, 5, 164–176, 2024.

Jones, D. C., Sonnewald, M., Zhou, S., Hausmann, U., Meijers, A. J., Rosso, I., Boehme, L., Meredith, M. P., and Naveira Garabato, A. C.: Unsupervised classification identifies coherent thermohaline structures in the Weddell Gyre region, Ocean Science, 19, 857–885, 2023.

Krawczyk, B.: Learning from imbalanced data: open challenges and future directions, Progress in artificial intelligence, 5, 221–232, https://doi.org/10.1007/s13748-016-0094-0, 2016.

Liu, W., Tsamados, M., Petty, A., Jin, T., Chen, W., and Stroeve, J.: Enhanced sea ice classification for ICESat-2 using combined unsupervised and supervised machine learning, Remote Sensing of Environment, 318, 114 607, https://doi.org/10.1016/j.rse.2025.114607, 2025.

Marchese, C., Albouy, C., Tremblay, J.-É., Dumont, D., d'Ortenzio, F., Vissault, S., and Bélanger, S.: Changes in phytoplankton bloom phenology over the North Water (NOW) polynya: a response to changing environmental conditions, Polar Biology, 40, 1721–1737, https://doi.org/10.1007/s00300-017-2095-2, 2017.

Martin, S. and Cavalieri, D. J.: Contributions of the Siberian shelf polynyas to the Arctic Ocean intermediate and deep water, Journal of Geophysical Research: Oceans, 94, 12 725–12 738, https://doi.org/https://doi.org/10.1029/JC094iC09p12725, 1989.

Martin, S., Drucker, R., Kwok, R., and Holt, B.: Estimation of the thin ice thickness and heat flux for the Chukchi Sea Alaskan coast polynya from Special Sensor Microwave/Imager data, 1990–2001, Journal of Geophysical Research: Oceans, 109, https://doi.org/https://doi.org/10.1029/2004JC002428, 2004.

Mohrmann, M., Heuzé, C., and Swart, S.: Southern Ocean polynyas in CMIP6 models, The Cryosphere, 15, 4281–4313, https://doi.org/10.5194/tc-15-4281-2021, 2021.

Monroe, E. E., Taylor, P. C., and Boisvert, L. N.: Arctic Cloud Response to a Perturbation in Sea Ice Concentration: The North Water Polynya, Journal of Geophysical Research: Atmospheres, 126, e2020JD034 409, https://doi.org/https://doi.org/10.1029/2020JD034409, e2020JD034409 2020JD034409, 2021.

Moore, G., Howell, S., and Brady, M.: Evolving relationship of Nares Strait ice arches on sea ice along the Strait and the North Water, the Arctic's most productive polynya, Scientific Reports, 13, https://doi.org/10.1038/s41598-023-36179-0, 2023.

Morales Maqueda, M. A., Willmott, A. J., and Biggs, N. R. T.: Polynya Dynamics: a Review of Observations and Modeling, Reviews of Geophysics, 42, https://doi.org/https://doi.org/10.1029/2002RG000116, 2004.

Notz, D. and Community, S.: Arctic sea ice in CMIP6, Geophysical Research Letters, 47, e2019GL086 749, https://doi.org/10.1029/2019GL086749, 2020.

Ohshima, K. I., Nihashi, S., and Iwamoto, K.: Global view of sea-ice production in polynyas and its linkage to dense/bottom water formation, Geoscience Letters, 3, 1–14, 2016.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

Paçal, A., Hassler, B., Weigel, K., Kurnaz, M. L., Wehner, M. F., and Eyring, V.: Detecting extreme temperature events using Gaussian mixture models, Journal of Geophysical Research: Atmospheres, 128, e2023JD038 906, 2023.

Paţilea, C., Heygster, G., Huntemann, M., and Spreen, G.: Combined SMAP–SMOS thin sea ice thickness retrieval, The Cryosphere, 13, 675–691, https://doi.org/10.5194/tc-13-675-2019, 2019.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and Vanderplas, J.: Scikit-learn: Machine learning in Python, the Journal of machine Learning research, 12, 2825–2830, 2011.

Poropat, L., Jones, D., Thomas, S. D., and Heuzé, C.: Unsupervised classification of the northwestern European seas based on satellite altimetry data, Ocean Science, 20, 201–215, 2024.

Preußer, A., Heinemann, G., Willmes, S., and Paul, S.: Circumpolar polynya regions and ice production in the Arctic: results from MODIS thermal infrared imagery from 2002/2003 to 2014/2015 with a regional focus on the Laptev Sea, The Cryosphere, 10, 3021–3042, https://doi.org/10.5194/tc-10-3021-2016, 2016.

Ren, H., Shokr, M., Li, X., Zhang, Z., Hui, F., and Cheng, X.: Estimation of Sea Ice Production in the North Water Polynya Based on Ice Arch Duration in Winter During 2006–2019, Journal of Geophysical Research: Oceans, 127, e2022JC018 764, https://doi.org/https://doi.org/10.1029/2022JC018764, e2022JC018764 2022JC018764, 2022.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, pp. 234–241, Springer, 2015.

Smedsrud, L. H., Budgell, W. P., Jenkins, A. D., and Ådlandsvik, B.: Fine-scale sea-ice modelling of the Storfjorden polynya, Svalbard, Annals of Glaciology, 44, 73–79, https://doi.org/10.3189/172756406781811295, 2006.

Smith Jr, W. O. and Barber, D.: Polynyas: Windows to the world, vol. 74, Elsevier, 2007.

Sonnewald, M., Lguensat, R., Jones, D. C., Dueben, P. D., Brajard, J., and Balaji, V.: Bridging observations, theory and numerical simulation of the ocean using machine learning, Environmental Research Letters, 16, 073 008, 2021.

Stroeve, J. and Notz, D.: Changing state of Arctic sea ice across all seasons, Environmental Research Letters, 13, 103 001, https://doi.org/10.1088/1748-9326/aade56, 2018.

Sumata, H., de Steur, L., Divine, D. V., Granskog, M. A., and Gerland, S.: Regime shift in Arctic Ocean sea ice thickness, Nature, 615, 443–449, https://doi.org/10.1038/s41586-022-05686-x, 2023.

Tamura, T. and Ohshima, K. I.: Mapping of sea ice production in the Arctic coastal polynyas, Journal of Geophysical Research: Oceans, 116, https://doi.org/https://doi.org/10.1029/2010JC006586, 2011.

Tatebe, H. and Watanabe, M.: MIROC MIROC6 model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.5603, 2018.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geoscientific Model Development, 12, https://doi.org/10.5194/gmd-12-2727-2019, 2019.

Wong, C., Heuzé, C., Ickes, L., and Zhou, L.: The spatio-temporal variability, trends, and drivers of winter Arctic polynyas, under review for Journal of Climate, https://doi.org/10.31223/X5VJ0P, 2025.

Ye, W., Lao, J., Liu, Y., Chang, C.-C., Zhang, Z., Li, H., and Zhou, H.: Pine pest detection using remote sensing satellite images combined with a multi-scale attention-UNet model, Ecological Informatics, 72, 101 906, 2022.

https://doi.org/10.5194/egusphere-2025-2747
Preprint. Discussion started: 23 June 2025

Zhou, L., Heuzé, C., and Mohrmann, M.: Early Winter Triggering of the Maud Rise Polynya, Geophysical Research Letters, 49, e2021GL096 246, https://doi.org/https://doi.org/10.1029/2021GL096246, e2021GL096246 2021GL096246, 2022.

Zhou, L., Heuzé, C., and Mohrmann, M.: Sea ice production in the 2016 and 2017 Maud Rise polynyas, Journal of Geophysical Research: Oceans, 128, e2022JC019 148, 2023.